# THE PHYLOGENETIC HISTORY OF HYPOTHALAMIC NEUROMODULATORS

EDITED BY: Jackson Cioni Bittencourt, Giovanne B. Diniz, David Lovejoy and Herbert Herzog

PUBLISHED IN: Frontiers in Neuroscience and Frontiers in Endocrinology

frontiers Research Topics

# frontiers



ISSN 1664-8714
ISBN 978-2-88971-222-9
DOI 10.3389/978-2-88971-222-9

## About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

## Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

## Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

## What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

# THE PHYLOGENETIC HISTORY OF HYPOTHALAMIC NEUROMODULATORS

Topic Editors:
**Jackson Cioni Bittencourt,** University of São Paulo, Brazil
**Giovanne B. Diniz,** California National Primate Research Center (CNPRC), United States
**David Lovejoy,** University of Toronto, Canada
**Herbert Herzog,** Garvan Institute

**Citation:** Bittencourt, J. C., Diniz, G. B., Lovejoy, D., Herzog, H., eds. (2021). The Phylogenetic History of Hypothalamic Neuromodulators. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88971-222-9

# Table of Contents

frontiers in Neuroscience

EDITORIAL
published: 23 June 2021
doi: 10.3389/fnins.2021.712448

# Editorial: The Phylogenetic History of Hypothalamic Neuromodulators

*Jackson C. Bittencourt*[1]*, *Giovanne B. Diniz*[2], *David A. Lovejoy*[3] and *Herbert Herzog*[4]

[1] Department of Anatomy, Institute of Biomedical Sciences, University of São Paulo, São Paulo, Brazil, [2] California National Primate Research Center, University of California, Davis, Davis, CA, United States, [3] Department of Cell and Systems Biology, University of Toronto, Toronto, ON, Canada, [4] Neuroscience Division, Garvan Institute of Medical Research, Darlinghurst, NSW, Australia

**Keywords: neurosecretion, peptides, neuromodulators, evolution, diencephalon, invertebrates, vertebrates, homologs**

## Editorial on the Research Topic

## The Phylogenetic History of Hypothalamic Neuromodulators

We are delighted to introduce this new Research Topic of "Frontiers in Neuroscience" that focuses on the phylogenetic history and evolution of a number of critical hypothalamic hormones associated with the neuroendocrine regulation of vertebrate energy homeostasis.

The first studies on the concept of neurosecretion were published almost a century ago, beginning with the work of Ernst Scharrer, who showed that the preoptic region of fishes possessed neuroendocrine properties associated with pituitary function. Together with his wife, Bertha, they reported that a neurosecretory system located between the *corpus cardiacum* and *corpus allatum* in insects may be homologous to the hypothalamo-pituitary system of vertebrates. However, it was not until the 1930s that the concept of neurosecretion became integrated into the understanding of how the hypothalamus could play a similar role in vertebrates. Later, Bargmann and Scharrer postulated that vasopressin and oxytocin were produced in the hypothalamus and released via the neural lobe of the pituitary gland. By the 1940s, Geoffrey Harris and John Green established that neuroendocrine factors could flow from the hypothalamus to the pituitary gland. These studies established the modern basis of hypothalamic and pituitary function.

Since then, these seminal studies have spawned advances in comparative neurobiology, endocrinology proteomics, and genomics. However, almost a century after the early studies, we still have much to learn. Our goal in this article collection was not to provide a comprehensive review of the current state of scientific understanding of the multitude of neurohormones associated with the hypothalamus and their role with organismal metabolism and homeostasis, but rather to provide a selective overview of some key studies across a variety of hypothalamic neurohormones and their functions. We anticipate that these publications will foster new research and stimulate new directions in the field of hypothalamus-associated physiology. Given this, we offer several novel articles in various aspects of hypothalamic function.

With respect to the phylogenetic origins of hypothalamic peptides, four studies have focused on the early phylogeny of these peptides. Building upon the initial work on the role of oxytocin and vasopressin release from the neurohypophysis, Odekunle and Elphick have established the early phylogenetic history of oxytocin- and vasopressin-like peptides in invertebrates and their structural relationship to those orthologs in vertebrates. Similarly, but focusing on the adenohypophysis, Cardoso et al. have established that PACAP, a peptide associated with the Secretin family of peptides which possesses structural similarity with peptides found in early evolving metazoans, likely obtained its final gene and peptide form in vertebrates. Another peptide family described in this volume is that of the corticotropin-releasing factor/hormone (CRF/CRH) system. Required for the central regulation of the hypothalalmus-pituitary-adrenal (HPA) axis, and possessing a

OPEN ACCESS

***Edited and reviewed by:***
*Hubert Vaudry,*
*Université de Rouen, France*

***Correspondence:**
*Jackson C. Bittencourt*
*jcbitten@icb.usp.br*

***Specialty section:***
*This article was submitted to Neuroendocrine Science, a section of the journal Frontiers in Neuroscience*

***Received:*** *20 May 2021*
***Accepted:*** *26 May 2021*
***Published:*** *23 June 2021*

***Citation:***
*Bittencourt JC, Diniz GB, Lovejoy DA and Herzog H (2021) Editorial: The Phylogenetic History of Hypothalamic Neuromodulators. Front. Neurosci. 15:712448. doi: 10.3389/fnins.2021.712448*

number of paralogues including CRF2/teleocortin; urocortin/urotensin-I/sauvagine; urocortin 2; and urocortin 3, the CRF system is one of the most ancient hypothalamic systems and serves as a model system to understand how invertebrate molecules evolved into vertebrate neuromodulators. Cardoso et al. from their studies in lampreys, introduce a new and detailed understanding of how CRF gene phylogeny, following the R2 hypothesis, led to the expansion of CRF-related peptides in vertebrates. Finally, Michalec et al. have suggested a novel phylogenetic relationship of secretin, CRF, and insulin-like peptides based on their relationship to the teneurin C-terminal associated peptides (TCAP), which have been suggested to among the earliest peptide precursors to CRF in the metazoans.

Although these studies provide some insight into the first wave of peptides discovered from the earliest studies indicated above, additional new studies provide novel interpretations of vital hypothalamic neuropeptides. Melanin-concentrating hormone (MCH), initially discovered in teleosts, and essential for modulating arousal and energy balance in vertebrates, in addition to numerous other functions, is the subject of a paper by Diniz and Bittencourt. This paper establishes that there are two paralogues of this gene and peptide in teleosts and provides a detailed understanding of these gene duplications and paralogues lead to new functions. Prolactin, likewise, has numerous physiological functions in vertebrates. Dobolyi et al. have provided an evolutionary rationale for the functions of prolactin concerning osmoregulation, growth, and reproduction in vertebrates. The orexins (hypocretins) are among the newer peptides discovered as hypothalamic factors regulating organismal physiology, energy metabolism and, importantly, sleep/wake cycles. Soya and Sakurai provide a novel review of this system with respect to their functions in vertebrates. Finally, due to the essential integratory function of the hypothalamus, most of its critical neurohormonal systems are subject to complex regulation during development. Schredelseker et al. have shown that the brain-specific homeobox transcription factor (Bsx) is required for the transcription of a number of critical energy regulating peptides, including neuropeptide Y, CRF, vasoactive intestinal peptide (VIP), and others, indicating that although these peptides belong to different gene systems, they may be regulated by a common ancestral transcription factor.

We hope the readers will utilize these publications to extend their own research in the structure, function, and phylogeny of these and other hypothalamic modulators.

## AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neuroscience

REVIEW
published: 26 November 2019
doi: 10.3389/fnins.2019.01280

# The Melanin-Concentrating Hormone (MCH) System: A Tale of Two Peptides

*Giovanne B. Diniz*[1,2] *and Jackson C. Bittencourt*[1,3]*

[1] *Departamento de Anatomia, Instituto de Ciências Biomedicas, Universidade de São Paulo, São Paulo, Brazil,* [2] *Department of Neurosurgery, Yale School of Medicine, New Haven, CT, United States,* [3] *Nucleo de Neurociencias e Comportamento, Instituto de Psicologia, Universidade de São Paulo, São Paulo, Brazil*

The melanin-concentrating hormone (MCH) system is a robust integrator of exogenous and endogenous information, modulating arousal and energy balance in mammals. Its predominant function in teleosts, however, is to concentrate melanin in the scales, contributing to the adaptive color change observed in several teleost species. These contrasting functions resulted from a gene duplication that occurred after the teleost divergence, which resulted in the generation of two MCH-coding genes in this clade, which acquired distinctive sequences, distribution, and functions, examined in detail here. We also describe the distribution of MCH immunoreactivity and gene expression in a large number of species, in an attempt to identify its core elements. While initially originated as a periventricular peptide, with an intimate relationship with the third ventricle, multiple events of lateral migration occurred during evolution, making the ventrolateral and dorsolateral hypothalamus the predominant sites of MCH in teleosts and mammals, respectively. Substantial differences between species can be identified, likely reflecting differences in habitat and behavior. This observation aligns well with the idea that MCH is a major integrator of internal and external information, ensuring an appropriate response to ensure the organism's homeostasis. New studies on the MCH system in species that have not yet been investigated will help us understand more precisely how these habitat changes are connected to the hypothalamic neurochemical circuits, paving the way to new intervention strategies that may be used with pharmacological purposes.

**Keywords: NEI, MCHR1, MCHR2, lateral hypothalamus, neuropeptides, vertebrates**

OPEN ACCESS

***Edited by:***
*Dan Larhammar,*
*Uppsala University, Sweden*

***Reviewed by:***
*Hervé Tostivint,*
*Muséum National d'Histoire Naturelle, France*
*Gina Leinninger,*
*Michigan State University, United States*

***Correspondence:**
*Jackson C. Bittencourt*
*jcbitten@icb.usp.br*

***Specialty section:***
*This article was submitted to Neuroendocrine Science, a section of the journal Frontiers in Neuroscience*

***Received:*** *07 August 2019*
***Accepted:*** *11 November 2019*
***Published:*** *26 November 2019*

***Citation:***
*Diniz GB and Bittencourt JC (2019) The Melanin-Concentrating Hormone (MCH) System: A Tale of Two Peptides. Front. Neurosci. 13:1280. doi: 10.3389/fnins.2019.01280*

**Abbreviations:** 3V, third ventricle; A13, dopaminergic group A13; aa, amino acid(s); ACTH, adrenocorticotropic hormone; AH, adenohypophysis; AHA, anterior hypothalamic area; AROM, antisense RNA overlapping MCH; CNS, central nervous system; CRF, corticotropin-releasing factor; CSF, cerebrospinal fluid; DHA, dorsal hypothalamic area; DHN, dorsomedial hypothalamic nucleus (non-mammalian); f, fornix; GH, growth hormone; GPR24, G protein-coupled receptor 24; ic/cp, internal capsule/cerebral peduncle; IHy, incerto-hypothalamic area; IR, immunoreactivity; LH, lateral hypothalamus; LHA, lateral hypothalamic area; LVR, lateral ventricular recess; MCH, melanin-concentrating hormone; MCHR1, melanin-concentrating hormone receptor 1; MCHR2, melanin-concentrating hormone receptor 2; ME, median eminence; mfb, medial forebrain bundle; MGOP, MCH gene-overprinted polypeptide; MGRP, MCH gene-related peptide; MYA, million years ago; NEI, neuropeptide glutamic acid-isoleucine; NGE, neuropeptide glycine-glutamic acid; NH, neurohypophysis; NLT, lateral tuberal nucleus (*nucleus lateralis tuberis*); *Npr-24,* neuropeptide receptor family 24 gene; ORX, orexin; PCR, proximal neurosecretory contact region; Pe, periventricular hypothalamic area; PHA, posterior hypothalamic area; *Pmch/pmch/PMCH*, melanin-concentrating hormone precursor (gene); Pmch/PMCH, melanin-concentrating hormone precursor (protein); *PMCHL1*, *PMCH*-Linked 1 gene; *PMCHL2*, *PMCH*-Linked 2 gene; PVO, paraventricular organ; SL, somatolactin; SLC-1, somatostatin-like receptor 1; SSTR, somatostatin receptor(s); Th, thalamus; VTN, ventral tuberal nucleus; ZI, zona incerta; α-MSH, α-melanocyte-stimulating hormone.

# INTRODUCTION

## The Melanin-Concentrating Hormone System

The melanin-concentrating hormone (MCH) system is a robust integrator of exogenous and endogenous information, modulating arousal, promoting motivated behaviors, and controlling energy balance (Diniz and Bittencourt, 2017; Chee et al., 2019), contributing to the appropriate sleep architecture (Ferreira et al., 2017a; Gao, 2018), and tethering energy status and reproductive physiology (Naufahu et al., 2013). While certain aspects of this system have been explored in length, such as its neuroanatomical aspects (Bittencourt and Diniz, 2018), several others are still open to investigation, including its role in parental behavior (Benedetto et al., 2014; Alachkar et al., 2016; Alhassen et al., 2019), or the mechanisms through which MCH is used to convey information within the central nervous system (CNS) (Noble et al., 2018). Although MCH is strongly linked to the roles above, its discovery, and hence its name is linked to an additional function performed in Teleosts: the control of skin color through the modulation of chromatophore activity (Kawauchi et al., 1983). MCH is synthesized by neurons in the Teleost hypothalamus and released in the bloodstream through the neurohypophysis (NH), conferring MCH the status of neurohormone. Upon reaching the melanophores in the scales, MCH promotes pallor, necessary for adaptive color change. Although only isolated in 1983, its existence was predicted almost 50 years prior, by Hogben and Slome (1931).

The identification of MCH in the chum salmon pituitary was the gateway for a plethora of discoveries regarding this system. Just six years after its description in Salmoniforms, Nahon et al. (1989) identified the mammalian *Pmch* gene, as well as other predicted peptides that originate from the *Pmch*-encoded precursor, PMCH: neuropeptide E-I (NEI) and neuropeptide G-E (NGE), following the nomenclature scheme of Tatemoto and Mutt (1981) (**Figure 1**). In that same year, Vaughan et al. (1989) isolated and sequenced the mammalian MCH peptide (**Figure 1**). In parallel, different peptides were identified originating from the teleost *Pmch* gene, such as neuropeptide E-V (Minth et al., 1989). Three years later, Bittencourt et al. (1992) published the first complete mapping of *Pmch* expression and MCH and NEI immunoreactivity (IR) in the rat brain.

While a wide range of vertebrates has been investigated in terms of *Pmch* expression and MCH synthesis, the precise origin of *Pmch* in vertebrates is unknown. As we will see in detail in the next section, lampreys are the earliest-diverging animals with evidence of MCH existence. Nahon et al. (1989) proposed that PMCH and the prepropeptide-A of the sea slug *Aplysia californica* are distantly related, given the 24% of sequence identity shared between the two prepropeptides and the matching sets of cleavage sites. This hypothesis, however, failed to gain traction in the literature, making the origin of *Pmch* in vertebrates an important open question that will require experimental investigation to be answered. In the lack of new data, the most parsimonious hypothesis is that the founder gene of *Pmch* originated in phylostratum 11, as determined by the phylostratigraphy method of Domazet-Lošo and Tautz (2010) and Domazet-Lošo et al. (2007).

The next breakthrough in the field came at the turn of the century, when reverse pharmacology studies identified GPR24/SLC-1 as the selective receptor for MCH, now known as MCHR1 (Bächner et al., 1999; Chambers et al., 1999; Lembo et al., 1999; Saito et al., 1999; Shimomura et al., 1999). Homology searches in genomic databases then revealed a second MCH receptor in 2001, now known as MCHR2 (An et al., 2001; Hill et al., 2001; Mori et al., 2001; Rodriguez et al., 2001; Sailer et al., 2001; Wang et al., 2001). Initially identified as the somatostatin-like coupled receptor 1, MCHR1 shares over 40% identity with somatostatin receptors (SSTRs) in the transmembrane domains (Kolakowski et al., 1996; Lakaye et al., 1998). Despite these similarities, somatostatin does not bind to MCHR1, with MCH acting as its only specific ligand. The level of similarity between MCHR1 and SSTRs is comparable to the identity between the paralogs MCHR1 and MCHR2, which share 44% identity in the transmembrane domain (An et al., 2001). As is the case with *Pmch*, the precise origin of MCHR-coding genes in the vertebrate lineage is unknown. Yun et al. (2015) suggested that putative ortholog genes of *Mchr* exist in the ascidian *Ciona intestinalis* and in the lancelet *Branchiostoma floridae*. Furthermore, in the same work, *B. floridae* and vertebrate *Mchr* paralogs were nested with the protostome neuropeptide receptor family 24 (*Npr-24*) of *Caenorhabditis elegans*, suggesting a possible distant relationship between MCH receptors and *Npr-24*. This falls in line with the idea that many neuropeptidergic families have roots in invertebrate species (Elphick et al., 2018).

In addition to the canonical MCH system, there are non-canonical transcripts that originate from the *Pmch*/*PMCH* genes. Toumaniantz et al. (1996) discovered an alternative-splicing product originating from those genes, the MCH-gene-overprinted-polypeptide (MGOP). In the antisense strand of the *Pmch*/*PMCH* genes, Borsu et al. (2000) identified the antisense-RNA-overlapping-MCH (AROM), a complex gene that originated coding and non-coding transcripts that may modulate gene expression (Moldovan et al., 2012). Finally, exclusively in the hominid lineage, two chimeric genes originated from *PMCH, PMCH-Linked* 1, and 2 (*PMCHL1*/*PMCHL2*), with putative transcription modulation activity (Courseaux and Nahon, 2001).

As just a glance reveals, MCH is part of a complex system, involved in numerous functions and with multiple canonical and non-canonical elements. This complexity stems from a rich evolutionary history, as significant genomic events influenced the *Pmch* gene and its ancillary elements. The MCH system, therefore, provides us with a window to look at those evolutionary events and how they shaped the vertebrate hypothalamus and its circuits. In this review, the canonical MCH peptidergic system will be reviewed, including the orthologs and paralogs of the *Pmch* gene, and the distribution of the *Pmch*-coded peptides within the nervous system. Due to the abundance of data, we will not include the MCH receptors and the non-canonical elements, except when they help us understand

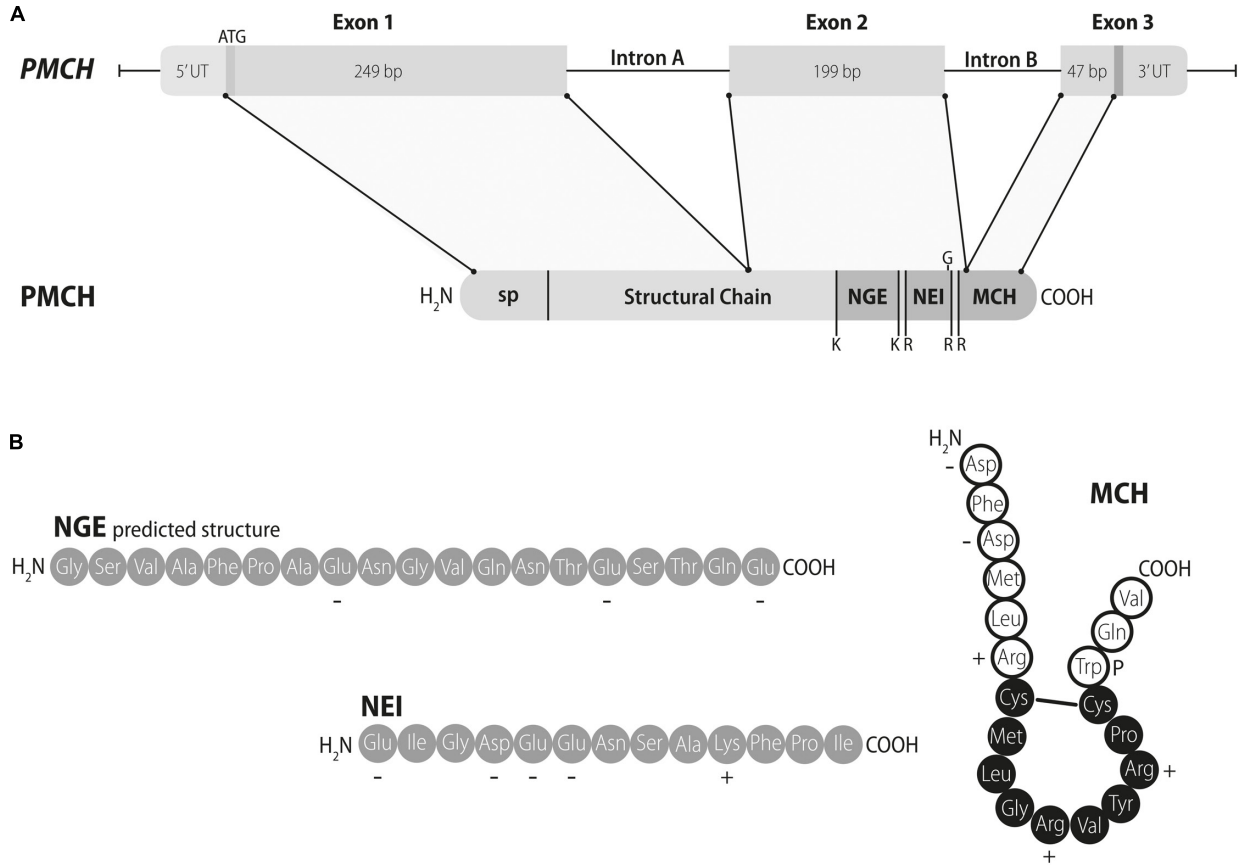


**FIGURE 1 |** The mammalian *PMCH* gene, PMCH protein, and the peptides of the MCH family. **(A)** The mammalian PMCH gene is composed of three exons and two introns. Both introns and exons have variable lengths, with the sizes depicted in this diagram corresponding to human *PMCH*. The transcribed product of *PMCH* is formed by five components: a signal peptide (sp), encoded in Exon 1; a structural chain, formed by both Exons 1 and 2; two MCH gene-related peptides (MGRPs), neuropeptide G-E (NGE) and neuropeptide E-I (NEI), both encoded in Exon 2; and a mature melanin-concentrating hormone (MCH), with 3 N-terminal residues found in Exon 2, a residue formed by Exon 2 and Exon 3 combining to generate its codon, and 15 C-terminal residues found in Exon 3. Both MCH and NEI can be generated through dibasic proteolytic sites, indicated in the diagram by RR and KR, and NGE may be cleaved at a single basic site K. Neuropeptide E-I contains an amidation site in its C-terminal portion, indicated by G. **(B)** Structure of the peptides generated from *PMCH*. Both NGE and NEI are linear peptides characterized by a predominance of negatively charged residues. The precise structure of NGE is unknown, given that this peptide has never been isolated and characterized. Mature MCH, on the other hand, is a cyclic peptide, due to a cysteine bridge formed between residues 7 and 16. The ring structure formed between the two cysteine residues, indicated by black circles, is essential for the binding activity of MCH to its receptors. Residue 17 plays an important role in potentiating MCH binding (indicated by the letter P). Adapted with permission from Bittencourt and Diniz (2018).

the peptidergic family. A brief description of the phylogenetic relationship between major clades is included in each section to help readers contextualize the information.

## THE MCH SYSTEM IN EARLY CHORDATES

### Chordata > Petromyzontiformes

Lampreys (order Petromyzontiformes) and hagfishes (order Myxiniformes) were the first vertebrates to diverge (**Figure 2**), with molecular estimates ranging from 797 to 489 MYA (Kumar and Hedges, 1998; Blair and Hedges, 2005; Kuraku and Kuratani, 2006; Kumar A. et al., 2017; Miyashita et al., 2019), while fossil data situates this divergence predating the Devonian, at least 360 MYA (Gess et al., 2006). Together, these two groups form superclass Cyclostomata, clade characterized by the lack of a jaw and internal branchial arches. Due to the very early divergence of Cyclostomata, species belonging to this clade are often considered models for the common vertebrate ancestor. Four species of lamprey have been used to study MCH: the European river lamprey (*Lampetra fluviatilis*), sea lamprey (*Petromyzon marinus*), brook lamprey (*Lampetra planeri*), and pouched lamprey (*Geotria australis*). Morphological data has been obtained in all species using a salmon MCH-directed antibody (Al-Yousuf and Mizuno, 1991; Bird et al., 2001). Little is known about MCH-related genes in Petromyzontids, despite the availability of the whole genome of *P. marinus* (Smith et al., 2013). Two genes are annotated as *Mchr1* (ENSPMAG00000005548.1) and *Mchr2* (ENSPMAG00000001735.1), but there is no gene annotated as *Pmch*. It should be noted, however, that scaffold GL478617 contains a sequence that codes for a peptide that has

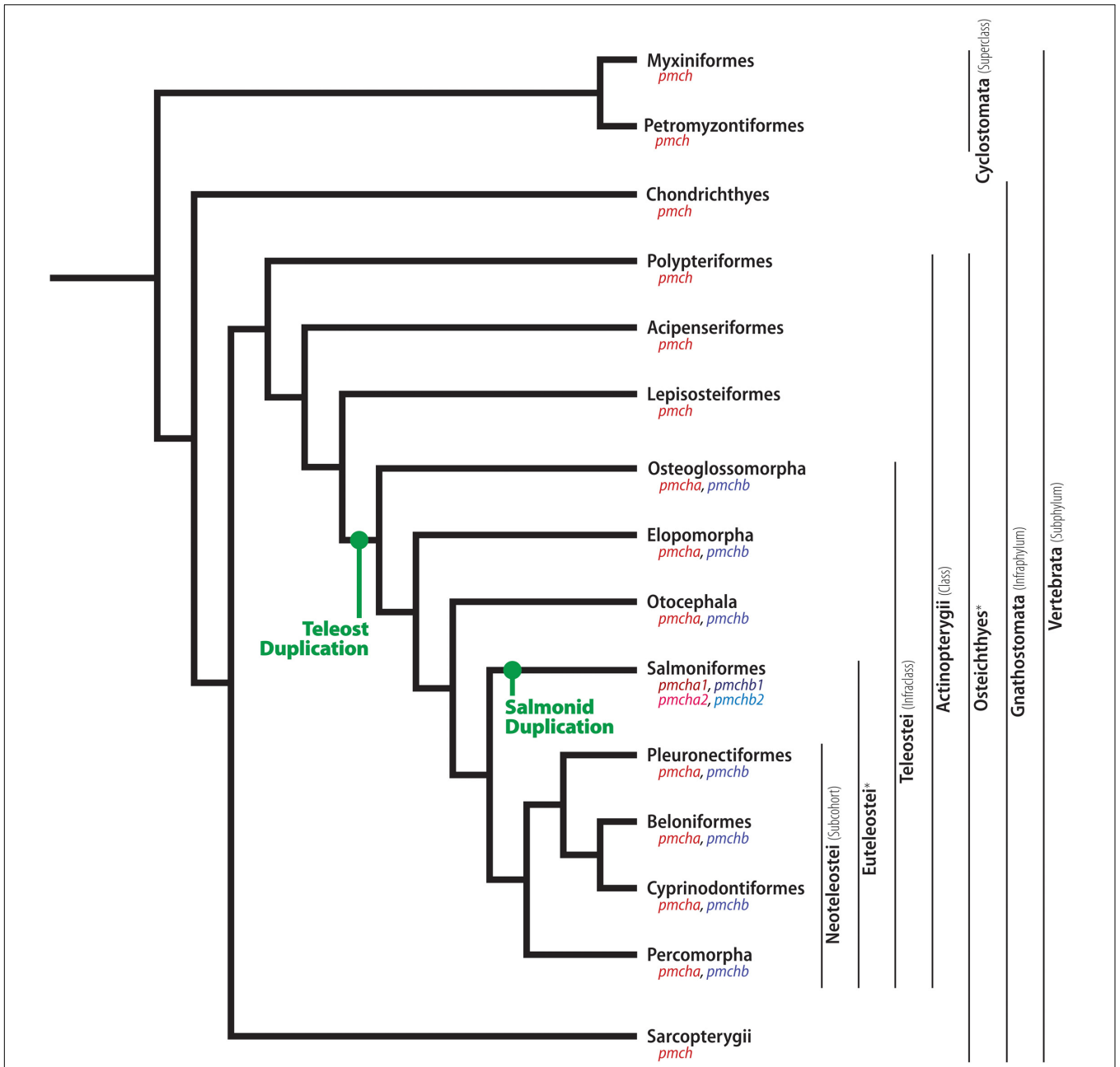


**FIGURE 2 |** Phylogenetic relationship between early-diverging Chordates and Actinopterygians. Illustrative cladogram showing the main phylogenetic relationship between clades with available information about MCH. Several clades have been omitted for clarity, and the cladogram is not meant to represent the absolute order of relationship between these groups, but rather the most probable organization based on current molecular data and information about the MCH system. The relative position of two *pmch* duplication events are indicated in green, and the resulting genes are indicated under each clade name. The ubiquitous vertebrate *pmch* and its homologs are indicated in red, while the teleost-specific *pmchb* is indicated in blue. The duplicated copies originated in the salmonid lineage are indicated in different shades of red and blue. To facilitate the reading of this review, relevant clades are indicated alongside their rank on the **right** side of the cladogram, encompassing the clades to the **left** of each black line. * Unranked clades.

13 identical residues when compared to Gnathostome *Pmch*, including a fully conserved cysteine ring, but experimental confirmation is necessary.

In Petromyzontids, MCH-immunoreactive cells are predominantly restricted to a single hypothalamic *locus*, the dorsomedial hypothalamic nucleus (DHN) of the posterior hypothalamus, occupying parts of the ependyma and subependyma, with only a few neurons found scattered toward the lateral hypothalamus (LH) (Al-Yousuf and Mizuno, 1991; Bird et al., 2001). The position of MCH cells in the hypothalamus

is illustrated in **Figure 3**. The DHN is found bordering the third ventricle (3V) and is part of the paraventricular organ (PVO), an ubiquitous non-mammalian structure that contains a myriad of neuroactive substances (Nozaki et al., 1983; Brodin et al., 1990; Tobet et al., 1995), plays a role in hypothalamic integration (Vigh-Teichmann and Vigh, 1989; Meurling and Rodríguez, 1990), has no blood-brain barrier, and is highly vascularized (Röhlich and Vigh, 1967). Exclusively in sexually maturing *L. petromizon*, a weakly-labeled group of neurons is found in the anterior basal telencephalon. Neurons immunoreactive to MCH in the petromyzontid DHN are frequently bipolar, with one axon projecting into the 3V *lumen*, and the other axon extending laterally toward the LH.

In addition to local lateral projections, three major innervation pathways are observed: anterior, toward the olfactory lobes; dorsal, toward the habenular nucleus; and posterior, toward the spinal cord (Bird et al., 2001). In *L. fluviatilis* and *L. planeri*, an additional innervation pathway is observed: ventral, toward the hypophysis. In lamprey, the hypophysis is continuous with the hypothalamic floor, as there is no portal system or *infundibulum*. In *L. fluviatilis*, immunoreactive fibers are observed in the proximal neurosecretory contact region (PCR) – the Petromyzontid homolog of the tetrapod median eminence (ME) – and in the proximal NH. Some immunolabeled axon terminals are also found close to the basement membrane that separates the NH and the *pars distalis* of the adenohypophysis (AH) (Al-Yousuf and Mizuno, 1991). It should be mentioned, however, that Baker and Rance (1983) did not find evidence of MCH activity in a bioassay using *L. fluviatilis* hypophyses.

Based on these data, Bird et al. (2001) make a series of morphofunctional correlates that, as it will be described later in this work, are extremely relevant for the understanding of MCH in mammals. These authors suggest that: (1) the 3V-contacting axon of MCH neurons, and their position near the ependyma and subependyma, could allow those neurons to sense biomarkers in the cerebrospinal fluid (CSF), or release MCH directly in the *lumen* to act on distant sites; (2) The ample distribution of laterally-projecting axons allows MCH neurons to exert widespread modulation within the CNS; (3) The moderate presence of MCH-ir terminals in the NH could be the substrate through which MCH modulates the release of other neuropeptides in a paracrine fashion, or influences the secretory action of the AH; (4) The NH MCH fibers observed in *L. petromizon* could originate from the telencephalic group of cells observed only in sexually maturing individuals of that species, and those neurons could play a role in physiological adaptation toward reproduction.

### Chordata > Gnathostomata > Chondrichthyes

Cartilaginous fish, members of the Chondrichthyes class, are the earliest diverging class of living jawed vertebrates (Gnathostomata), splitting from bony vertebrates (Osteichthyes) between 475 and 450 MYA (Venkatesh et al., 2014; Kumar S. et al., 2017; **Figure 2**). Extant animals are divided into two subclades, Holocephali (chimeras) and Elasmobranchii (sharks and rays), which split about 421 MYA (Renz et al., 2013).

Chondrichthyans have a single *pmch* gene composed of three exons, encoding a Pmch precursor that is 172 aa-long in the Holocephalan Australian ghost shark (*Callorhinchus milii* – *RefSeq: XP_007893646.1*) or 165 aa-long in the Elasmobranch scalloped hammerhead (*Sphyrna lewini* – *BAM63324.1*). At the N-terminus of this precursors sits a signal peptide that varies in length according to the species, and at the C-terminus is a mature 19 aa-long MCH that can be released by proteolytic activity through a dibasic Arg-Arg site (Mizusawa et al., 2012). In *C. milii*, it is possible that a second 14 aa-long peptide (Neuropeptide G-T) is cleaved from another dibasic pair upstream from the MCH-originating pair, but in *S. lewini* this peptide is 33 aa in length (Neuropeptide T-V). Since there is significant variation in the other peptides that may be produced from the *pmch* gene in addition to MCH, these peptides will all be grouped under an umbrella term: MCH gene-related peptides (MGRPs), a term borrowed from Gröneveld et al. (1993). The Chondrichthyan *pmch* gene is a perfect blueprint for the mammalian *PMCH*. Both human and *S. lewini* precursors have the same length (165 aa), a signal peptide in its N-terminus (which varies in length depending on the species), and a 19 aa-long mature MCH in the C-terminus that can be cleaved from an Arg-Arg pair. There is a single substitution between mammalian MCH and Chondrichthyes MCH ($Val^{19}$ in most mammals, $Asn^{19}$ in *S. lewini*, $Ile^{19}$ in *C. milii*), but the ring structure between $Cys^{7}$-$Cys^{16}$ is wholly preserved. The conservation of the ring sequence is likely linked to its importance for MCH binding to its receptors (Macdonald et al., 2000).

Regarding the distribution of immunoreactivity in Chondrichthyes, labeled cells are found in the dorsal wall of the posterior hypothalamus, and fibers were found exclusively inside the hypothalamus (Mizusawa et al., 2012; **Figure 3**). This distribution of MCH-synthesizing neurons is similar to what has been described for Cyclostomes, further reinforcing the dorsomedial posterior hypothalamus as the original *locus* of MCH synthesis in the chordate brain. No fibers are found in the hypophysis, and it is unclear if the hypothalamic-restricted distribution of fibers is a feature of the species or a methodological artifact.

## THE MCH SYSTEM IN THE RAY-FINNED FISH LINEAGE

### Gnathostomata > Osteichthyes > Actinopterygii

The superclass Osteichthyes contains all vertebrates with bony skeletons, splitting from Chondrichthyans at around 450 MYA (Venkatesh et al., 2014). Osteichtyans split early into two major clades, Actinopterygii (the ray-finned fishes) and Sarcopterygii (lobe-finned fishes), at around 435 MYA (**Figure 2**; Kumar S. et al., 2017). Actinopterygians then diverged into more than 25,000 known extant species, making it the most specious

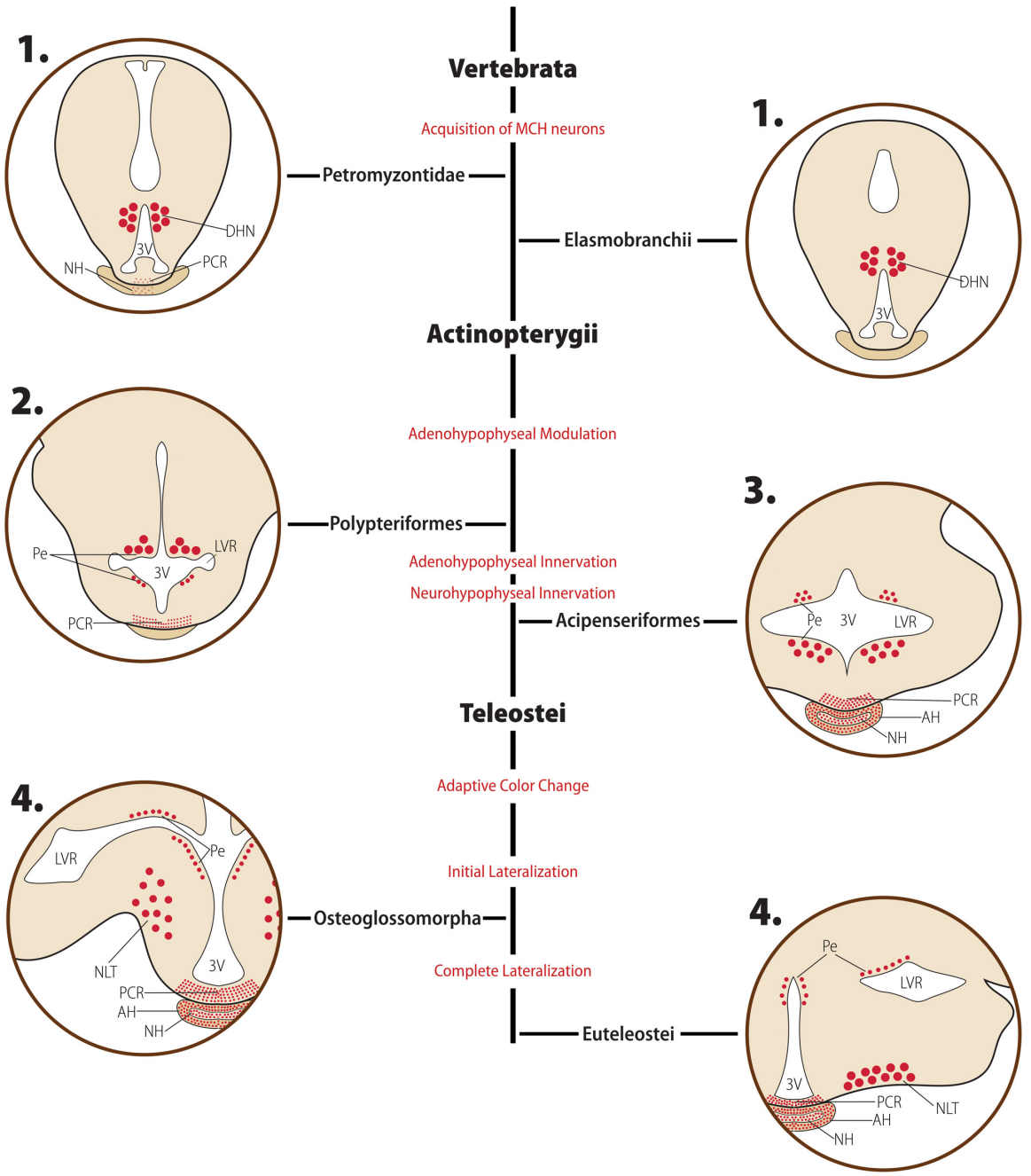


**FIGURE 3 |** Diagrammatic representation of the main morphological features of the MCH distribution in early-diverged Chordates and Actinopterygians. Major clades are indicated in bold font, while clade-wide events are represented in red. The diagrams for each clade are not meant to represent a single species or a single hypothalamic level, but rather a visual summary of what has been described for animals belonging to that clade. Red circles represent MCH neurons: small circles represent neurons with small somas, while large circles represent neurons with enlarged somas, typically associated with a neurosecretory profile. Fiber plexuses are represented by red dots. **1.** In Petromyzontids and Elasmobranchs, the majority of MCH neurons is found in the ependymal and subependymal layers (Pe) of the third ventricle (3V), within the dorsomedial hypothalamic nucleus (DHN). While there is no indication of hypophyseal innervation in Elasmobranchs, there are mixed reports about the presence of MCH in the proximal neurosecretory contact region (PCR) and neurohypophysis (NH) of Petromyzontids. **2.** After the Actinopterygii divergence, two major changes occurred in the MCH system. First, the periventricular neurons were divided in two groups by the emerging lateral recess, separating them into dorsal and ventral groups. Second, a clear innervation of the PCR develops, indicating the acquisition of a modulatory role over the adenohypophysis by MCH neurons at this stage. **3.** Shortly after the development of PCR innervation, MCH fibers started innervating the adenohypophysis (AH) and NH directly, creating the anatomical substrate necessary for a neurosecretory role. **4.** Around the time of the Teleost split, the final large-scale changes in the MCH system occurred. A lateral migration process started, with large, neurosecretory MCH neurons moving toward the *nucleus lateralis tuberis* (NLT) and densely projecting to the hypophysis from this ventrolateral position. This change occurred contemporaneously to the acquisition of a role in adaptive color change by MCH. Based on data from Baker and Bird (2002); Duarte et al. (2001) and Amano et al. (2003).

vertebrate clade, partially thanks to the Teleost radiation that occurred at the Cretaceous-Paleogene transition (Friedman, 2010; Sibert and Norris, 2015). Actinopterygians split into five clades: Polypteriformes, Acipenseriformes, Lepisosteiformes, Amiiformes, and Teleostei (**Figure 2**). There is uncertainty in the literature regarding the exact relationship between the clades of Actinopterygii (Kumar S. et al., 2017), but for this work it is only relevant that Polypteriformes, Acipenseriformes, Lepisosteiformes, and Amiiformes are all considered ancient with respect to Teleostei. Due to the importance of the latter to the understanding of MCH, those animals will be described in a separate section.

Regarding the makeup of *pmch* genes, there are no significant changes in non-Teleost Actinopterygii. The genes with available data are comprised of three introns, giving origin to 167–170 aa-long precursors (*Lepisosteus oculatus* – XP_015207385.1, *Erpetoichthys calabaricus* – XP_028672358.1). Mature MCH is 19 aa in length, with some divergence observed in the last residue ($Val^{19}$ in *L. oculatus*, $Ile^{19}$ in *E. calabaricus*), with a 13 aa-long MGRP possibly cleaved at an upstream Arg-Arg pair. Contrasting to the relatively unchanged genomic structure, the Actinopterygii divergence was a period of change in terms of the neuroanatomy of the MCH system. In the brain of the Polypteriform *E. calabaricus*, MCH IR is remarkably similar to what has been described for Petromyzontids. All immunoreactivity resides in the periependymal area, over the dorsal surface of the lateral ventricular recesses (LVR) and in the lateral wall of the 3V (**Figure 3**). These neurons are found within the PVO, close to the blood capillaries and contacting the ventricular cavity. Abundant fibers are found in the PCR, but not in the hypophysis (Baker and Bird, 2002). The distribution of immunoreactivity in the Acipenseriform starry sturgeon (*Acipenser stellatus*) is similar to that of *E. calabaricus*, with one striking difference: instead of stopping at the PCR, fibers continued toward the NH (Baker and Bird, 2002; **Figure 3**).

In the brain of the Lepisosteriform longnose gar (*Lepisosteus osseus)*, a significant change occurred. Instead of being concentrated in a single area, MCH neurons are found in two separate groups: the dorsomedial ventricular group, similar to what has been described previously, and a new group of neurons within the *nucleus lateralis tuberis* (NLT). Neurons in the ventricular group follow the same pattern as described for MCH neurons so far: predominantly bipolar, in association with the PVO, with one axon contacting the ventricular cavity and another branching in the lateral hypothalamus and other areas of the CNS. Fibers from the NLT, on the other hand, are found coursing through the basal hypothalamus toward the pituitary stalk, forming a plexus around blood capillaries of the PCR and NH (Baker and Bird, 2002).

The morphological aspects of the MCH system in the three aforementioned species are in accordance to some models of Actinopterygii divergence. Polypteriforms were likely the first clade to split, at around 407 MYA (Kumar S. et al., 2017), and the distribution of MCH-immunoreactivity in these animals closely resemble that of Petromyzontids and Elasmobranchii.

The next clade to split was likely Acipenseriformes, since *A. stellatus* has a very similar distribution of cellular bodies but differs from Polypteriforms by having a dense direct innervation of the NH. Since the lamprey *L. planari* also has a direct innervation of the NH, it is possible that this feature appeared independently in Petromyzontids and the common ancestor between Acipenseriforms and Lepisosteriforms. Alternatively, the innervation of the NH by MCH fibers may have first appeared in Petromyzontids as a plastic feature, becoming then fixed by the time of Acipenseriform divergence. An evolutive advantage of a direct NH innervation (interpreted as MCH release directly in the bloodstream) may have paved the way for the split between periventricular/dorsomedial and tuberal lateral groups to be positively selected, with NLT neurons becoming a magnocellular group that preferentially innervates the NH, which was then itself followed by the acquisition of adaptive color change.

## Actinopterygii > Teleostei

The infraclass Teleostei split from other Actinopterygii at around 320 MYA (Kumar S. et al., 2017). Teleosts can be divided into four major clades: Osteoglossomorpha, Elopomorpha, Otocephala, and Euteleostei (**Figure 2**). Most teleost species are part of Euteleostei, with extant members of the other three clades including bonytongues (Osteoglossomorpha), eels (Elopomorpha), and catfishes (Otocephala). Although there is some controversy regarding the exact phylogenetic relationship between Teleost clades, molecular and morphological data suggest that Osteoglossomorpha may have been the first clade to split, at 285 MYA, followed by Elopomorpha (265 MYA), and finally Otocephala and Euteleostei (230 MYA) (Kumar S. et al., 2017). Since several developments occurred in the MCH system shortly after the Teleostei split, and Euteleosts have been extensively used to probe the MCH system, they will be examined in the following section, with this section focusing on non-Euteleost Teleosts.

### Structural and Evolutionary Aspects

Considerable changes in terms of genomic makeup marked the divergence of Teleosts. This period was characterized by accelerated genome changes, including higher rates of gene-linkage disruption and chromosomal rearrangements (Ravi and Venkatesh, 2008), a large number of retroposition events (Fu et al., 2010), and a teleost-specific whole-genome duplication event (Jaillon et al., 2004; Meyer and Van de Peer, 2005). A second copy of *pmch* was generated during this period, likely through retroposition, given that the newly generated copy is intronless, a hallmark of retrocopies (Fu et al., 2010; Grzybowska, 2012; Casola and Betrán, 2017), and does not follow the expected chromosomal location for a copy originated from a whole-genome duplication (Kasahara et al., 2007). This duplication, when compounded with the historical order of MCH discoveries, creates substantial clutter in the nomenclature of *pmch* orthologs and paralogs. Therefore, we will use in this review a nomenclature that adheres to the prescribed gene nomenclature guidelines for the various species examined, at the expense of not using some of the original nomenclature used in the literature. A table of normalized terms is provided (**Table 1**). Henceforth, the gene most commonly identified as "pmch2" will be designated as *pmcha*, and "pmch1" will

**TABLE 1 |** Standardized nomenclature of *pmch* genes across vertebrate species.

| Gene abbreviation | Protein abbreviation | Diff. criteria |
|---|---|---|
| **Petromyzontid** | | |
| *Pmch* | Pmch | – |
| **Elasmobranchii** | | |
| *pmch* | Pmch | – |
| **Actinopterygii (non-teleost)** | | |
| *Pmch* | Pmch | – |
| **Teleost (non-salmonids)** | | |
| *Pmcha* | Pmcha | Three exons |
| *pmchb* | Pmchb | One exon |
| **Salmonids** | | |
| *pmcha1* | Pmcha1 | Three exons, MCH ($Ser^{19}$) |
| *pmcha2* | Pmcha2 | Three exons |
| *pmchb1* | Pmchb1 | One exon, NEV ($Gly^{3}$) |
| *pmchb2* | Pmchb2 | One exon |
| **Dipnoi** | | |
| *Pmch* | Pmch | – |
| **Lissamphibian** | | |
| *Pmch* | Pmch | – |
| **Sauropsid (non-aves)** | | |
| *Pmch* | Pmch | – |
| **Aves** | | |
| *PMCH* | PMCH | – |
| **Mammalian (non-primate)** | | |
| *Pmch* | PMCH | – |
| **Primate** | | |
| *PMCH* | PMCH | – |

*The gene nomenclature guidelines used to compose this table are: Sprague et al. (2006); Karimi et al. (2017), Kusumi et al. (2011); Burt et al. (2009), Blake et al. (2016) and Braschi et al. (2018).*

be designated as *pmchb*. This nomenclature revision has the added benefit of conciliating the nomenclature for the mature MCH produced by teleost paralogs ($MCH_A$ and $MCH_B$) with the standardized nomenclature for MCH receptors ($MCH_1$ and $MCH_2$) as defined by the Nomenclature Committee of the Union of Basic and Clinical Pharmacology (NC-IUPHAR) (Alexander et al., 2017).

The *pmcha* gene has a similar structure to both Elasmobranch and Mammalian *pmch/Pmch*, being comprised of three exons, having a 3' splice site of intron 2 in the same position, and displaying similar synteny. The Pmcha prepropeptide ranges from 147 to 157 aa in length (*Electrophorus electricus* – XP_026877406.1, *Paramormyrops kingsleyae* – XP_023672675.1), and through a dibasic cleavage site originates a mature $MCH_A$ that is 19 aa-long and 84.2% identical and 89.5% similar to mammalian MCH. Three substitutions are observed in the mature $MCH_A$, $Ile^2$ replaces $Phe^2$, $Val^9$ replaces $Leu^9$, and $Ala^{19}$ replaces $Asn^{19}$/$Ile^{19}$ in Elasmobranchii or $Val^{19}$ in Mammals. An additional aa substitution may have occurred in the common carp, *Cyprinus carpio* ($Ile^4$ replacing $Met^4$ – XP_018948604.1, XP_018948612.1). These changes are all conservative, and only the $Val^9$ substitution has occurred in the bioactive zone of $MCH_A$. The *pmchb* gene, on the other hand, is intronless and codes a Pmchb precursor that is 124 aa-long in all described species (*C. carpio* – XP_018967332.1). There is a remarkably low similarity between *pmcha* and *pmchb*. At the C-terminus of Pmchb is a mature $MCH_B$ that is 17 aa-long (two residues shorter in the N-terminus), has two substitutions on the N-terminal stretch before the ring structure ($Thr^2$ replaces $Met^4$ and $Met^3$ replaces $Leu^5$), the same substitution as $MCH_A$ inside the ring, and one non-conservative substitution in the C-terminal sequence outside the ring ($Glu^{16}$ replaces $Gln^{18}$).

Two aspects of the formation of $MCH_A$ and $MCH_B$ in the Teleost ancestor are worth noting. The first is the remarkable capability of neuromodulators to change shortly after being duplicated. Given the uniform distribution of *pmchb* in early Teleosts, it is clear that this newly generated copy underwent its substitutions before there were any significant splits in the Teleost lineage. The second remarkable aspect is how genomic duplications affect phylogenetic constraints. As the only source of MCH in non-Teleosts, the *pmch/Pmch* gene remained remarkably conserved from Elasmobranchs to Mammals. In the Teleost lineage, however, *pmcha* became a very dynamic neuropeptide, while *pmchb* became the most conserved paralog, even though its sequence differs significantly from non-Teleost *pmch/Pmch,* probably due to the acquisition of an adaptive color change role for MCH, which then acted to impose a phylogenetic constraint over *pmchb*.

It should be mentioned that attempts to clone the transcripts of *pmch* genes in some Otocephala species have not always resulted in the identification of two paralogs. In the goldfish *Carassius auratus*, two paralogs encoding highly similar *pmch* were found (Cerdá-Reverter et al., 2006), and a single paralog was found in *Schizothorax prenati* (Wang et al., 2016) and *C. carpio* (Xu et al., 2019). This apparent contradiction between the genomic databases and the attempts to clone *pmch* transcripts can be explained by a very low expression of *pmcha,* coupled to a high dissimilarity between *pmcha* and *pmchb*. Since most probes have been designed based on the well-known salmon sequence (of *pmchb*), it is easy to imagine that most probes would fail to identify *pmcha* transcripts. Furthermore, as we will see in the next section, the use of different antibodies to map MCH in goldfish is supportive of the existence of two structurally dissimilar MCH peptides.

## Anatomy

In the Osteoglossomorphs goldeye (*Hiodon alosoides*) and freshwater butterflyfish (*Pantodon buchholzi*), the bulk of MCH neurons is found in the basal hypothalamus, but instead of forming a neuronal sheet in the NLT area, neurons are found in the mid-hypothalamic region (**Figure 3**). Despite this difference, these neurons project to the NH, similar to the NLT group of other Actinopterygians. A second group of small neurons is found in the periventricular area of the LVR, clustered around the PVO. As the other periventricular groups described here, these neurons project to the ependyma and the ventricular cavity (Baker and Bird, 2002). In the Elopomorph European eel (*Anguilla anguilla*), 80% of MCH neurons are large and located in the NLT at the ventrolateral hypothalamus, below the LVR, with dense projections to the NH. This is the first significant shift of

cells to the ventral hypothalamus from the periventricular area, a transitional stage that will be repeated in several other species. A small group of neurons still resides at the dorsal surface of the LVR, but there is no apparent contact with the PVO (Baker and Bird, 2002). In Otocephala, the dominant group of MCH neurons is located in the NLT, with dense projections to the hypophysis, in addition to projections to the thalamus, pretectal region, preoptic area, and telencephalon. A second, small group of neurons is observed close to the LVR, near the junction between the LVR and 3V. The axons of those neurons course toward the vicinity of the PVO, but no direct contact with the ventricular cavity is observed (Bird et al., 1989; Baker and Bird, 2002). These descriptions were made using a salmon MCH-directed antibody; therefore, those descriptions are likely more relevant to $MCH_B$ than to $MCH_A$ in those species. Fortunately, some works provide some insight into the differences between $MCH_A$ and $MCH_B$ in terms of distribution.

In zebrafish (*Danio rerio)*, *pmchb* mRNA expression is found in the lateral and posterior NLT, in a group dorsal to the LVR, and the caudal zone of the periventricular hypothalamus. The distribution of *pmcha*, on the other hand, is more restricted, with *pmcha*-expressing neurons found exclusively in the anterior NLT. There is no overlap in the expression of the two *pmch* paralogs. Immunoreactivity to $MCH_A$ and $MCH_B$ was determined to follow a similar pattern, as revealed by the use of salmon MCH- and mammalian MCH-directed antibodies. An extensive network of $MCH_A$-ir fibers was found, including immunoreactivity in the dorsal nucleus of the ventral telencephalic area, the thalamus, the habenula, the periventricular nucleus of the posterior tuberculum, the posterior tuberal nucleus, and the *torus lateralis*. In the hypothalamus, fibers were found in the lateral and periventricular zones, the ventral hypothalamus close to the ME, and the hypophysis. In the mesencephalon, the periventricular gray zone of the optic *tectum* and the *torus semicircularis*, and in the rhombencephalon fibers were found in the *griseum centrale* and *locus coeruleus* (Berman et al., 2009). This widespread distribution of fibers, covering regions from the anterior telencephalon to the rhombencephalon will be found in mammals and other species. A similar dichotomy between *pmcha*/$MCH_A$ and *pmchb*/$MCH_B$ is observed in the goldfish (Huesa et al., 2005; Cerdá-Reverter et al., 2006; Matsuda et al., 2006; Tanaka et al., 2009).

These observations of MCH immunoreactivity in Teleostei leading to Euteleostei show a consolidation of the pattern that emerged in ancient Actinopterygii and, in particular, at the time of Lepisosteriform divergence. The NLT group of MCH neurons became the dominant group, and a strong innervation of the ME and the hypophysis developed. On the other hand, the periventricular group started drifting away of the periependymal area and the PVO, and lost contact with the ventricular cavity, but remained in the posteromedial part of the hypothalamus. The duplication of the *pmch* gene also impacted the anatomy of the system. While *pmchb* was retained in all previously described groups of MCH neurons, projections from $MCH_B$-synthesizing neurons became concentrated in the hypophysis. On the other hand, *pmcha* became restricted to a small group of neurons in the NLT, but those neurons have widespread fibers in the CNS.

## Teleostei > Euteleostei > Protacanthopterygii

The Euteleosts have been extensively investigated, due to the initial discovery of MCH happening in the chum salmon pituitary. The Euteleosts can be split into two clades: Protacanthopterygii and Neoteleostei, which includes Acanthopterygii (**Figure 2**). Order Salmoniformes, which includes trout and salmons, is part of the Protacanthopterygii, and will be the focus of this section. The members of the *Oncorhynchus* genus played a historical role in the discovery of MCH since the peptide was first isolated from the hypophysis of the chum salmon (*Oncorhynchus keta*) by Kawauchi et al. (1983).

### Structural and Evolutionary Aspects

Another major genetic shift occurred after the divergence of the Salmoniformes. In these animals, an additional duplication of the *pmch* genes occurred, conforming to the 4R theory, which states that a fourth whole-genome duplication event occurred in the Salmoniform lineage, between 25 and 100 MYA (Allendorf and Thorgaard, 1984), and this is reflected in gene databases (**Figure 2**). The nomenclature employed by the automated computational analyses, however, is often confusing and should be interpreted with care. In this review, the two copies of *pmcha* will be called *pmcha1* and *pmcha2*, and the two copies of *pmchb* will be called *pmchb1* and *pmchb2*. Detailed information is available for three species of Salmonids, the rainbow trout (*Oncorhynchus mykiss*), the coho salmon (*Oncorhynchus kisutch*), and the Chinook salmon (*Oncorhynchus tshawytscha*), with minimal variation between species. Both *pmcha1* and *pmcha2* are composed of three exons that encode Pmcha precursors that are 144 aa-long for *pmcha2* (*O. mykiss* – XP_021432849.1 {Chromosome 21}; *O. tshawytsha* – XP_024228562.1 {Chromosome 15}) and 146 or 147 aa-long for *pmcha1* (*O. mykiss* – XP_021419487.1 {Chromosome 15}; *O. tshawytsha* – XP_024232986.1 {Chromosome 17}). At the C-terminus is a mature $MCH_A$ that is 21 residues-long due to the insertion of two residues at the N-terminal ($Glu^1$ and $Ala^2$). There are also two conservative substitutions in the N-terminal stretch before the cysteine ring ($Leu^4$ replaces $Ile^2$ and $Glu^5$ replaces $Asp^3$) when compared to the carp $MCH_A$. A single substitution differentiates $MCH_{A1}$ from $MCH_{A2}$ ($Ser^{19}$ replaces $Trp^{19}$). The *pmchb1* and *pmchb2* genes, on the other hand, are very similar: both are intronless genes that code for 132 aa-long Pmchb (*O. mykiss* – XP_021446538.1 {Chromosome 29}, XP_021458034.1 {Chromosome 5}; *O. tshawytsha* – XP_024252615.1 {Chromosome 33}, XP_024237102.1 {Chromosome 20}). At the C-terminus of those preprohormones is a 17 aa mature $MCH_B$, with sequence identical to Otocephala $MCH_B$.

Attempts to clone MCH transcripts in salmonids, however, did not reproduce what is observed in the automatic annotation of gene databases. Most attempts to identify *pmch* genes in salmonids resulted in the identification of only *pmchb1* and *pmchb2* transcripts (Masao et al., 1988; Minth et al., 1989; Takayama et al., 1989; Nahon et al., 1991; Baker et al., 1995). Genes identified by library cloning agree well to the independent

information available at the online repositories. It is unclear, at this point, if all the salmonid *pmch* genes are expressed. One report in *O. tshawitscha* found similar patterns of expression for both *pmchb1* and *pmchb2* (Masao et al., 1988), while in *O. kisutch* and *O. mykiss* only *pmchb2* was found to be expressed, while *pmchb1* appears to be a silent gene (Nahon et al., 1991; Baker et al., 1995; Suzuki et al., 1995; Suzuki et al., 1997). Further studies are necessary to validate the automatic annotation reports of *pmcha1* and *pmcha2* in salmonids, and to detect if those genes are expressed.

### Anatomy

The distribution of $MCH_B$-ir cells has been reported for *O. mykiss* and *O. keta*. In these animals, the main group of $MCH_B$-ir neurons is found in the NLT, encircling the pituitary stalk (**Figure 3**). A dense network of fibers is directed to the hypophysis, while a few projections are found in the telencephalon, preoptic area, thalamus, and pretectal region. In the hypophysis, fibers predominantly innervated the NH, but could also be found in the AH. Smaller groups of neurons were found behind the pituitary stalk, extending between the basal hypothalamus and the LVR, and medially over the dorsal surface of the LVR, in close contact with the PVO (**Figure 3**; Naito et al., 1985; Bird et al., 1989; Baker et al., 1995; Suzuki et al., 1995; Baker and Bird, 2002). *In situ* hybridization supports the presence of cell bodies in the NLT and dorsal to the LVR (Baker et al., 1995). This distribution is highly compatible with that described for $MCH_B$ in zebrafish and other species of Otocephala.

## Teleostei > Euteleostei > Neoteleostei

The Neoteleostei is a large and complex clade of Euteleosts. Regarding the MCH system, all works have concentrated in a single taxon, Acanthomorpha, and within it, clade Percomorpha of superorder Acanthopterygii. The animals evaluated regarding the MCH system can be divided into two major clades, Carangimorpharia and Percomorpharia. Inside Carangimorpharia, order Pleuronectiformes was the first to diverge, followed then by the closely related orders Beloniformes and Cyprinodontiformes (**Figure 2**). Inside taxon Percomorpha, all species studies were part of order Perciformes.

### Evolutionary and Structural Aspects

The idea that the second duplication of MCH occurred specifically in the Salmonid lineage is reinforced by the observation of only two copies of *pmch* in Neoteleosts. Neoteleost *pmcha* encodes Pmcha precursors that range from 146 to 150 aa in length. The mature $MCH_A$ contained in these precursors also has a variable length, ranging from 21 aa (winter flounder, *Platichthys americanus*) to 25 aa (barfin flounder, *Verasper moseri*). There is no indication that an MGRP can be originated from Pmcha (Tuziak and Volkoff, 2012; Kang and Kim, 2013; Mizusawa et al., 2015). On the other hand, *pmchb* is an intronless gene, encoding a Pmchb protein that is shorter than Pmcha, ranging from 129 aa (*starry flounder, Platichthys stellatus*) to 136 aa (Nile tilapia, *Oreochromis niloticus*). In all available described and predicted cases, $MCH_B$ is 17 aa long and mostly conserved: tilapia $MCH_B$ is identical to Otocephala $MCH_B$, while other Neoteleosts have a single conservative substitution ($Asn^2$ replaces $Thr^2$), and the winter flounder has an additional substitution inside the ring sequence ($Gly^7$ replaces $Val^7$). In most cases, a large 22 or 23 aa-long MGRP precedes mature $MCH_B$, potentially cleaved in a single basic *locus* to originate smaller peptides (Gröneveld et al., 1993; Takahashi et al., 2004; Pérez-Sirkin et al., 2012; Tuziak and Volkoff, 2012; Kang and Kim, 2013; Hosomi et al., 2015). It should be noted, however, that Takahashi et al. (2004) found no evidence of MGRP synthesis using mass spectroscopy in *V. moseri* samples.

### Anatomy

The distribution of MCH has been examined in Pleuronectiformes (Baker and Bird, 2002; Amano et al., 2003; Amiya et al., 2008a,b), Cyprinodontiformes (Batten and Baker, 1988; Batten et al., 1999) and Perciformes (Mancera and Fernández-Llebrez, 1995; Batten et al., 1999; Duarte et al., 2001; Baker and Bird, 2002; Pandolfi et al., 2003; Cánepa et al., 2008). The overall distribution of MCH neurons remains mostly the same, with the main group of cells found in the NLT and a second, smaller group found close to the LVR (**Figure 3**). The principal group of fibers courses toward the hypophysis, with fibers reaching not only the NH but also the AH junction in the rostral and proximal *pars distalis*. Within the hypophysis, MCH-immunoreactive fibers contact multiple cellular types. A large number of fibers contain stained vesicles in the posterior neurohypophysis, making contact with pituicytes or the basement membrane of capillaries. Discontinuities within the neuro-intermediate basement membrane allow MCH-ir fibers to contact *pars intermedia* endocrine cells, including α-melanocyte-stimulating hormone (α-MSH) and somatolactin (SL) cells. In some instances, MCH fibers are observed in the adrenocorticotropic hormone (ACTH) cell zone and contacting growth hormone (GH) cells. Other cell types do not appear to be contacted by MCH-immunoreactive fibers (Batten and Baker, 1988; Batten et al., 1999). Smaller numbers of fibers are found projecting to other areas, such as the preoptic hypothalamus, thalamus, pretectal region, and telencephalon. In the sailfin molly (*Poecilia latipinna*), fibers are also found in the ventral telencephalon and olfactory bulb (Batten and Baker, 1988). Neurons above the LVR appear to project preferentially to non-hypophyseal targets. At least in Pleuronectiforms, neurons in the NLT are contacted by both gonadotropin-releasing hormone (GnRH)-ir and orexin-ir fibers, but reciprocal connections are only made to orexin neurons (Amiya et al., 2008a,b).

One interesting aspect of the MCH anatomy in Neoteleostei is the description of time-sensitive neurons during development. These neurons appear not to be present in Pleuronectiformes (Amano et al., 2003), but they are found in at least two Perciform species, the gilt-head bream (*Sparus aurata*) and the Cichlid *Cichlastoma dimerus*. In *S. auratus*, cells were found in the periventricular area of the medial hypothalamus from days 4 through 23 after hatching, disappearing after this time frame (Mancera and Fernández-Llebrez, 1995). In *C. dimerus*, the transient neurons were found in the *nucleus periventricularis*

posterior, starting at day 6 and disappearing by day 42 after hatching (Pandolfi et al., 2003).

## THE MCH SYSTEM IN THE LOBE-FINNED FISH LINEAGE

### Osteichthyes > Sarcopterygii > Dipnoi

As mentioned before, Actinopterygii and Sarcopterygii diverged at approximately 435 MYA (Kumar S. et al., 2017). The Sarcopterygii clade includes the tetrapod lineage, in addition to the clade containing their closest extant relatives, the lungfish, grouped in Subclass Dipnoi (**Figure 4**). Limited information is available about MCH in lungfish, as a single work has examined the distribution of MCH-ir elements in the West African lungfish (*Protopterus annectens*) (Vallarino et al., 1998). In this animal, diencephalic MCH neurons are found in two groups: a main group, located in the periventricular tuberal hypothalamus, and a second group described in the peripheral layers of the ventral hypothalamus (**Figure 5**). According to Croizier et al. (2013), this peripheral group of neurons corresponds to a migrated sheet of cells in contact with the dorsal periventricular hypothalamus. Another two groups have been found in lungfish: in the *subpallium* and the *pars intermedia* of the hypophysis. It is unclear, at this moment, if those extra-diencephalic groups represent actual *loci* of MCH neurons, with further studies necessary to ascertain their specificity. Regarding the fiber distribution, immunoreactive projections are found throughout the telencephalon, including the anterior olfactory nucleus, medial *subpallium*, and medial *pallium*. The preoptic, suprachiasmatic and caudal hypothalamus contain large numbers of fibers, while the thalamus receives a moderate-to-low number of fibers. The mesencephalon and rhombencephalon contain average numbers of fibers, except for the mesencephalic *tectum*, which received a large input, agreeing to what has been described for Actinopterygii. No fibers were observed in the ME or the hypophysis (Vallarino et al., 1998).

### Sarcopterygii > Tetrapoda > Lissamphibia

Tetrapods diverged from lungfish at approximately 413 MYA, and subclass Lissamphibia was the first group to diverge, including all extant amphibians, at approximately 350 MYA (**Figure 4**; Kumar S. et al., 2017). Amphibians, also classified as anamniotes, are semiaquatic, laying their eggs in the water. Subclass Lissamphibia is composed of three major groups: Gymnophiona (caecilians), Anura (frogs), and Caudata (salamanders). In terms of genetic composition, there are no major changes in Lissamphibians when compared to non-Teleost groups. A single *pmch* gene is found, encoding a Pmch precursor that ranges from 167 to 180 aa in length (*X. tropicalis* – XP_002936876.1, *N. parkeri* – XP_018415182.1). At the C-terminus of this Pmch is a mature MCH identical to mammalian MCH, which can be cleaved from a dibasic Arg-Arg pair. MGRPs can potentially be cleaved from a dibasic pair upstream of mature MCH, but they vary in length and sequence.

The distribution of MCH neurons in Lissamphibians follows a particular pattern that is not observed in other clades (Andersen et al., 1986; Francis and Baker, 1995; Lázár et al., 2002; Croizier et al., 2013). The major cluster of MCH neurons occurs in the dorsal periventricular nucleus, arranged as a subependymal sheet of cells. Although these neurons are found close to the PVO, they do not invade its limits. A second group of cells has been described in the ventral tuberal nucleus, at least in the common frog (*Rana temporaria*) (**Figure 5**). The lateral hypothalamus (LH) appears not to contain MCH neurons, at least in Anurans. Andersen et al. (1986) describe MCH-ir neurons in the ventral thalamic area of the marsh frog (*Rana ridibunda*), and Lázár et al. (2002) found MCH-IR in the posterior *tuberculum*, which Croizier et al. (2013) insightfully pointed out as harboring dopaminergic neurons. It is possible that these two groups represent the basis of what later became the zona incerta (ZI) and incerto-hypothalamic area (IHy) in mammals (Bittencourt et al., 1992; Sita et al., 2007). There appears to be substantial plasticity in the MCH system of Anurans, with an enlargement of certain subsets of neurons upon the transition from tadpole to adult and the appearance of neurons in the preoptic hypothalamus linked to the reproductive period (Francis and Baker, 1995). Areas that receive MCH input in Lissamphibia include the olfactory lobe, the habenular nucleus, the optic *tectum*, the ME, and the spinal cord. In addition to those areas, Andersen et al. (1986) found a dense plexus of fibers in the NH.

### Tetrapoda > Amniota > Sauropsida

Amniota is a group of vertebrates who have developed adaptations to lay their eggs in a terrestrial environment. The amnion membrane that gives name to the clade is a structure in the egg that forms a cavity filled with fluids around the embryo, providing the necessary hydration during development. This adaptation occurred at approximately 350 MYA and was a key development in the transition from an aquatic to a fully terrestrial life for vertebrates (Kumar S. et al., 2017). At approximately 320 MYA, the Amniota lineage split into Diapsida, who then later originated the Mammalia clade, and Sauropsida, which contains all extant reptiles and birds. Within the Sauropsida clade, order Squamata was the first to split, at approximately 280 MYA, and contains the extant lizards and snakes. The next split within Sauropsida was between Testudines and Archosauria, at approximately 250 MYA. Testudines contains the extant turtles, while Archosauria contains orders Crocodilia and Aves, which split at around 240 MYA (**Figure 4**; Kumar S. et al., 2017).

The genetic makeup of the MCH system remains strongly conserved in Sauropsida. Among Squamata, most animals from this clade have a single *pmch* formed by three exons that codes for a PMCH that is 164 aa-long in Serpentes (*N. scutatus* – XP_026550810.1, *P. bivittatus* – XP_007420018.1), and between 166 and 174 aa in other clades (*G. japonicus* – XP_015278938.1, *P. muralis* – XP_028601891.1). At the C-terminus of PMCH is a 19 aa-long mature MCH, which is identical to mammals in the common wall lizard (*Podarcis muralis*) and Schlegel's Japanese gecko (*Gekko japonicus*), or have a single substitution in residue 19 when compared to both humans and Elasmobranchs ($Ala^{19}$ replaces $Ile^{19}$ or $Val^{19}$). An additional substitution is observed in family Elapidae of serpents, where $Leu^{4}$ replaces $Met^{4}$.

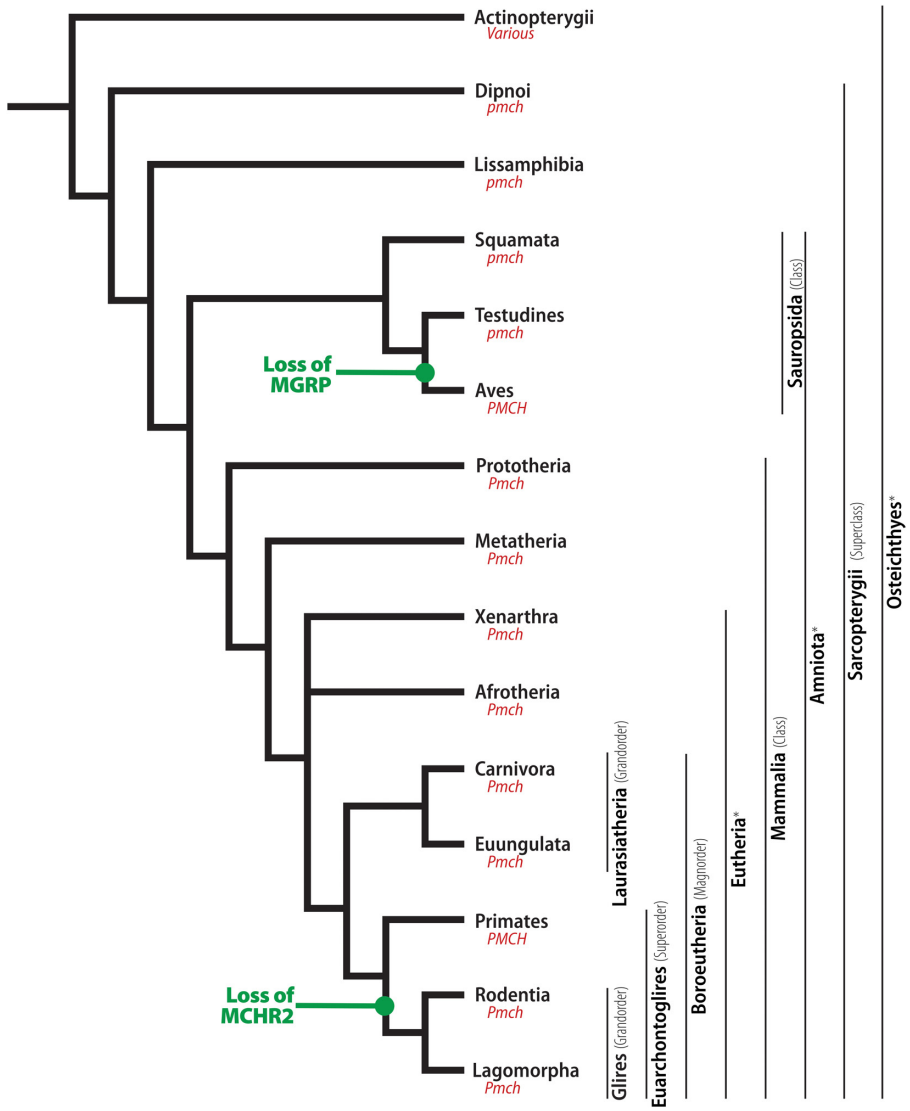


**FIGURE 4 |** Phylogenetic relationship between Sarcopterygians. Illustrative cladogram showing the main phylogenetic relationship between clades with available information about MCH. Several clades have been omitted for clarity, and the cladogram is not meant to represent the absolute order of relationship between these groups, but rather the most probable organization based on current molecular data and information about the MCH system. Two changes in the MCH system are indicated in the tree: the loss of an MCH gene-related peptide (MGRP) in birds and the loss of the MCH receptor subtype 2 (MCHR2) in Glires. Under each clade name is the standardized nomenclature for the pmch gene that respective clade. *Unranked clades.

In Testudines and Crocodilia, PMCH is similar to Squamata PMCH, 167 aa in length and with an MCH sequence that is identical to mammalian MCH (*P. sinensis* – XP_006138097.1, *C. porosus* – XP_019400295.1, *A. sinensis* – XP_006019570.1). In most species within Aves, the mature MCH structure is similar to the other Sauropsids, with occasional changes occurring in individual lineages, especially in positions 4 ($Ile^4$, $Thr^4$ or $Lys^4$ replacing $Met^4$) and 19 ($Ile^{19}$ or $Ala^{19}$ replacing $Val^{19}$). It is likely, therefore, that positions 4 and 19 represent "hotspots" that were frequently interchanges during evolution. In non-Avian Sauropsids, an MGRP that is 13 aa in length can be produced from a dibasic site, but in Aves, one of the Arg residues of this cleavage *locus* was replaced by a Glu residue. It is still possible that an MGRP may be produced from a single basic residue by different prohormone convertases, but this seems unlikely (Cardot et al., 1999).

Immunoreactivity to MCH has been described in a few Squamata and Testudines species: the common wall lizard (*P. muralis*), the grass and viperine snakes (*Natrix natrix* and *Natrix maura*), and the water turtle (*Chrysemis scripta elegans*), by Cardot et al. (1994). In Aves, species examined include chicken (*Gallus gallus domesticus*), guinea hens (*Numida meleagris*), quails (*Coturnix coturnix japonica*), gosling (*Anser domesticus*), ducks (*Cairina moschata*), and coots (*Fulica atra*) (Cardot et al., 1999). Cells immunoreactive to MCH were found in two major groups: at the dorsomedial periventricular nucleus, ventrolateral

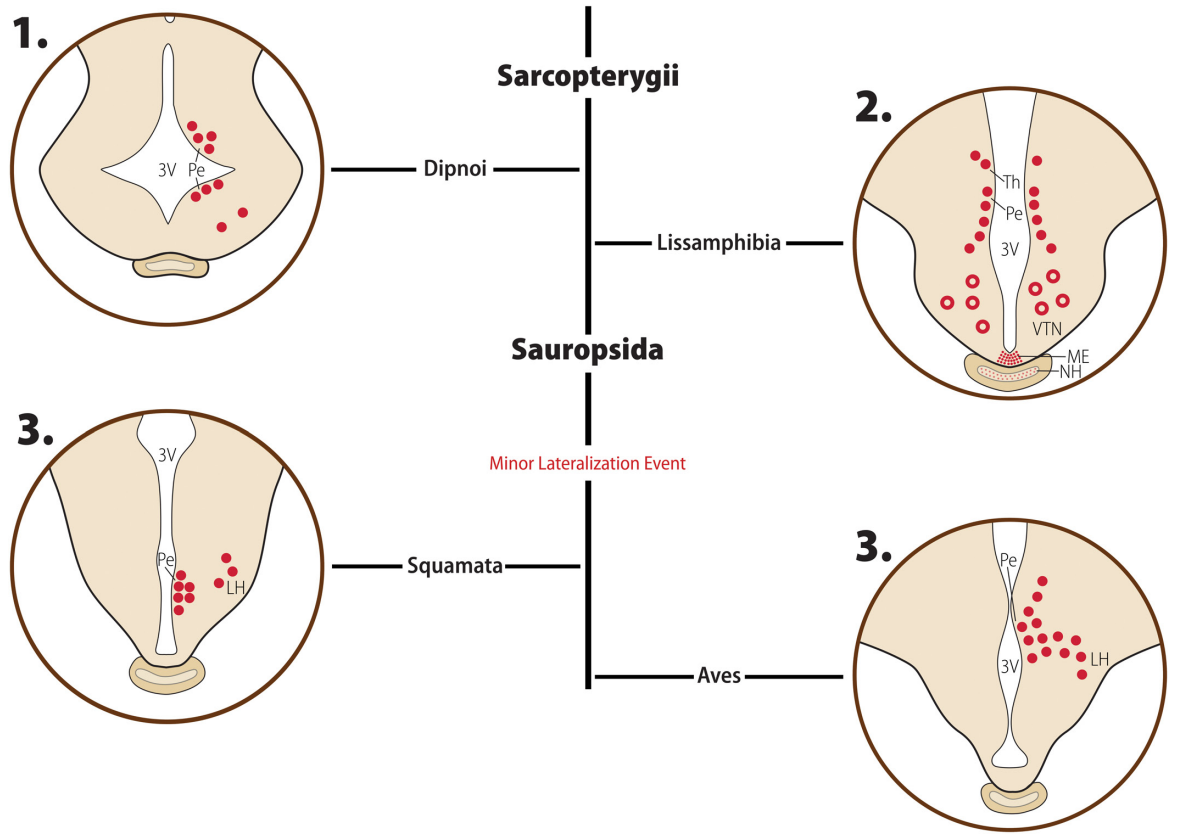


**FIGURE 5 |** Diagrammatic representation of the main morphological features of the MCH distribution in non-Mammalian Sarcopterygians. Major clades are indicated in bold font, while clade-wide events are represented in red. The diagrams for each clade are not meant to represent a single species or a single hypothalamic level, but rather a visual summary of what has been described for animals belonging to that clade. Full red circles represent MCH neurons, open red circles represent inconsistent information in the literature about MCH-ir neurons in the area. Fiber plexuses are represented by red dots, and inconsistent data about innervation is represented by red stars. **1.** Lungfish (Dipnoi) have a distribution of cells that closely resembles that of Elasmobranchs and early Actinopterygians, with neurons concentrated around the ventricle (Pe), and only scattered cells found in lateral areas. **2.** In Lissamphibians, neurons are still concentrated in the dorsomedial hypothalamus, but these cells are not as close to the ventricle as in other species, and some of the extend dorsally toward the ventral thalamus (Th). Fibers are observed in the Lissamphibian median eminence (ME). There is no consensus about the presence of MCH neurons in the ventral tuberal nucleus (VTN), represented by red open circles, or about the innervation of the neurohypophysis (NH), represented by red stars. **3.** In Sauropsids, a lateral migration event occurred, with some neurons found forming an arc shaped in the lateral hypothalamus (LH). These neurons do not surpass the number of periventricular neurons, however, making this a minor lateral migration event. Based on data from Vallarino et al. (1998); Cardot et al. (1994), Cardot et al. (1999); Francis and Baker (1995), and Lázár et al. (2002).

to the PVO, and in the LH. In some cases, a few cells were observed within the PVO. In the LH, cells were described to form an arc shape (**Figure 5**).

Regarding the distribution of fibers, non-avian Sauropsids are very similar to what is observed in mammals. Fibers are found in olfactory areas, such as the olfactory bulb, the olfactory *tuberculum*, and the piriform cortex; the septum, the diagonal band of Broca, the *paleostriatum*, the amygdala, parts of the cortex, the preoptic hypothalamus, the lateral zone of the hypothalamus, the pretectal area, the optic lobes and in several areas of the brainstem and spinal cord (Cardot et al., 1994). Immunoreactivity was also observed in the same areas when an anti-NEI antiserum was used, confirming the synthesis of an MGRP in non-avian Sauropsids. In Aves, the same basic plan was observed, but a few key differences are noted: projections to the olfactory system are more restricted in birds, the hippocampus receives less MCH fibers, the thalamus receives less dense projections, and although dense, projections to the brainstem are more constrained to specific areas, as opposed to the more diffuse projections observed in other Sauropsids. No staining was observed when an anti-NEI antibody was used (Cardot et al., 1999).

Summarizing these observations, non-avian Sauropsids developed a very similar pattern of projections to mammals, an example of convergent evolution facilitated by a shared common plan first observed in Lissamphibia. In the Aves lineage, however, projections were trimmed in some areas, while developed in others, likely to better work for the different needs of birds in their environment. Another important event was the loss of an MGRP in the Aves lineage, which combined to the great variability in the sequence of those peptides in other *phyla*, raises questions about the extent of functions performed by those peptides. We cannot discard the possibility, however,

that the loss of an MGRP in the base of Aves has facilitated the involution of that system in areas where its interaction with MCH was important.

## Tetrapoda > Amniota > Mammalia

Mammals originated from the sister clade of Sauropsida, Synapsida. Mammals are characterized by the acquisition of several morphological traits, including mammary glands, three bones in the inner ear, and hair. Another important development was the acquisition of a placenta, after 160 MYA, which separates the Eutherians from Prototherians (e.g., Platypus and Echidna) and Metatherians (e.g., modern marsupials) (**Figure 4**; Kumar S. et al., 2017). Among Eutherians, geographically distinct clades developed in between 105 and 100 MYA, including Xenarthra (e.g., anteaters and armadillos), Afrotheria (e.g., moles, shrews, tenrecs, manatees, and elephants), and Boroeutheria (e.g., rodents, primates, carnivores, ungulates), which likely split almost simultaneously (**Figure 4**; Nishihara et al., 2009). To us, the exact relationship between these groups is of little importance, as almost no information is available for Xenarthra and Afrotheria regarding the MCH system.

Regarding the *Pmch* gene, mammals have a very homogenous composition. The *Pmch* gene is formed by three exons and two introns, which contribute similarly to PMCH formation as in other vertebrates. Mature MCH is 19 aa-long, processed from a dibasic (Arg-Arg) pair, ending with a $Val^{19}$ in most mammals, but a few species have an $Ile^{19}$ (similar to Elasmobranchs), such as the platypus (*Ornithorhynchus anatinus* – XP_001508069.2), the pangolin (*Manis javanica* – XP_017536842.1), and a few primate species, such as the northern white-cheeked gibbon (*Nomaseus leucogenys* – XP_030677139.1) and the white-tufted-ear marmoset (*Callithrix jacchus* – XP_002752955.1). The only exception to that rule are some species of bats, where $Ile^5$ replaced $Leu^5$ (*E. fuscus* – XP_008142972.1, *M. brandtii* – XP_005880777.1). All substitution observed are conservative and happened outside the loop between Cys residues. Two MGRPs appear to be encoded in mammalian *Pmch*. Mature NEI is 13 aa-long and precedes MCH in the PMCH sequence. The sequence of NEI has been mostly conserved in mammals, with some variation observed in the first three residues, in particular in position 2. These changes are conservative in all cases, except for the rhinoceros (*Ceratotherium simun simun* – XP_014635845.1), where $Glu^1$ was replaced by a $Gly^1$. While NEI is produced from a dibasic Arg-Arg pair in Prototheria and Metatheria, Eutheria has a single substitution on the first Arg of the pair, allowing NEI to be cleaved from a Lys-Arg pair instead. Finally, NGE may be cleaved from a single basic residue, but evidence of its actual synthesis is lacking (for a review, see Bittencourt and Diniz, 2018).

The distribution of MCH-ir perikarya among mammals is familiar, but distinctive. Here, we observe a significant shift of MCH cells, which now are more numerously found in the LH, rather than the periventricular zone. This lateral migration is similar to what happened during the Actinopterygii differentiation and represents a second, independent shift in the position of MCH cells. Similar shifts occurred with less intensity in other clades, such as Dipnoi and Sauropsida, but the number of cells in lateral areas never surpassed the density of cells in the periventricular area in those groups. While the later migration of MCH cells in Actinopterygii occurred in parallel to an increase of MCH innervation of the hypophysis, the lateral migration observed in Mammals occurred concomitant to an expansion of MCH innervation throughout the CNS. These movements are strongly linked to hodological characteristics of the lateral areas: while the NLT has an intimate relationship with the *infundibulum*, the lateral hypothalamic area (LHA) of mammals acts as the bed nucleus of the medial forebrain bundle (*mfb*), a massive fiber bundle that connects the basal telencephalon to the hindbrain through ascending and descending fibers (Nieuwenhuys et al., 1982). The presence of MCH neurons in the LHA serves a double purpose: it allows these neurons to receive a massive amount of information, at the same time granting access to distant areas of the CNS.

The Mammalian LHA is a large structure, with a very intricate pattern of parcelation based on its connections and neurochemistry (Swanson et al., 2005; Hahn, 2010). Melanin-concentrating hormone-ir neurons are found in the LHA of all studied Mammals, but its relative position within this structure varies. This is likely related to the particular area of the *mfb* that needs to be accessed by MCH neurons, since the *mfb* has a well-defined topographical distribution within the LHA. The LHA group of cells also varies substantially in its rostrocaudal and mediolateral extent among Mammalian clades. Usually, a second diencephalic group is present, in the form of a dorsomedial cluster of cells, which varies among species in terms of its proximity to the ventricle and rostrocaudal position and extent. A ventral thalamic group is often observed, corresponding to the *zona incerta* (ZI). Other groups besides these three will be highlighted for the appropriate species in the next sections.

## Boreoeutheria > Laurasiatheria

Approximately 10 million years after the divergence of Boroeutheria, this clade split again into two groups: Laurasiatheria and Euarchontoglires. Laurasiatheria contains a diverse group of animals, including ungulates, bats, whales, and carnivores. There are only three species of Laurasiatherians with a described distribution of MCH neurons: the domestic cat (*Felis catus*), a Carnivore, and the Euungulates domestic pig (*Sus scrofa domesticus*) and sheep (*Ovis aries*) (**Figure 4**).

In Euungulates, MCH neurons have an extensive distribution in the rostrocaudal axis (Tillet et al., 1996; Chaillou et al., 2003; Chometton et al., 2014). This large extent results from the combination of the lateral group starting in the anterior hypothalamus, and the dorsomedial group prolonging into the anterior level of the ventral tegmental nucleus. The main group of MCH neurons in the LHA is found in its ventral part, between the ventral margin of the internal capsule and cerebral peduncle (ic/cp), the fornix (f), and the optic tract (**Figure 6**). Fewer MCH neurons are observed dorsal to the fornix, in the lateral part of the LHA. In the medial zone, a few immunoreactive cells are

found in the internuclear space between the dorsomedial and ventromedial hypothalamic nuclei, while the periventricular zone is mostly devoid of labeling. The dorsomedial periventricular group of MCH cells is first observed within the posterior hypothalamic area (PHA). The most significant difference between pigs and sheep concerns the subthalamic area, as no cells were reported in the sheep ZI, while both the rostromedial and ventral ZI were reported to contain immunoreactive cells in pigs (Tillet et al., 1996; Chaillou et al., 2003; Chometton et al., 2014).

The distribution of fibers in Euungulates is extensive, and the primary projection pathways resemble those observed in Lissamphibians and Sauropsids. The anterior pathway conducts fibers to ventral telencephalic and subtelencephalic structures, including the cortical fields ventral to the rhinal sulcus, the dorsal *subiculum*, the *taenia tecta*, the amygdala, and the medial *septum*. Only a light innervation is observed in the hippocampus proper, and the dentate gyrus is mostly devoid of fibers. A dorsal pathway allows a sparse innervation of the midline thalamic nucleus, the habenular nuclei, and the *subthalamus*. A dense descending pathway takes fibers to most of the midbrain and hindbrain, including the optic *tectum*, the *substantia nigra*, the reticular formation and the periventricular gray matter, up to the dorsal horn of the spinal cord. Local dense hypothalamic projections are also observed, mostly restricted to the lateral zone in sheep, but more widespread in pigs. Some fibers are observed in the external layer of the ME (Tillet et al., 1996; Chaillou et al., 2003; Chometton et al., 2014).

The domestic cat, a Carnivore, shows several distinctive characteristics in its distribution of MCH-ir perikarya when compared to Euungulates. The large rostrocaudal extent is not observed, with neurons restricted to the tuberal and mammillary levels of the diencephalon. At tuberal levels, the largest group of neurons is found in the LHA, and, within it, in the perifornical nucleus. These neurons are mostly found in the medial LHA, directly dorsal and ventral to the fornix, making the distribution of MCH neurons in the cat substantially more dorsal than what is observed in Euungulates. Some neurons are also observed in the ventral ZI. In the medial hypothalamus, neurons are found in the dorsal hypothalamic area (DHA), dorsal to the dorsomedial hypothalamic nucleus, and these neurons appear to be contiguous with neurons in the PHA (**Figure 6**). Immunoreactive neurons are not found at the level of the mammillary bodies (Torterolo et al., 2006; Badami et al., 2010).

While a pattern is easily observed for Euungulates, it is harder to discern an overarching pattern for Laurasiatherians. The two main group of cells observed in Lissamphibians are likely represented here, the LHA cells corresponding to the ventrolateral group, while the DHA/PHA group corresponds to the dorsal infundibular nucleus surrounding the 3V. In cats, however, the LHA cells migrated dorsally, occupying a dorsomedial position within the LHA. It is hard to determine, however, how representative of the Carnivores as a whole the distribution of cats is, as it is the only species of carnivore mapped to this point. The difference between Carnivora and Euungulata could correspond to the difference between feeding habits between these two clades, but more information is necessary to confirm this.

## Boreoeutheria > Euarchontoglires > Rodentia

At approximately 90 MYA, the Euarchontoglires clade split into two major groups: Glires, containing rabbits, hares, and Rodents; and Primates, which includes lemurs, monkeys, and humans. Glires then split into two major orders: Lagomorpha and Rodentia, at approximately 82 MYA (Kumar S. et al., 2017). The Glires clade is marked by the loss of one of the MCH receptor paralogs, making MCHR1 the only receptor found in those animals (**Figure 4**; Tan et al., 2002). No information about the MCH system is available for Lagomorphs, while Rodents have been amply investigated due to their popularity as animal models. Rodentia is a complex order, with several molecular techniques employed to try to define the relationships between members of this order (**Figure 7**). Although order Rodentia is composed of several groups, only Muroidea has been investigated in terms of MCH distribution, leaving animals like beavers, squirrels, guinea pigs, and jerboas still to be studied. The two largest groups of Muroids are the Murids and the Cricetids, with these two groups splitting at approximately 33 MYA. The Muroids include the traditional laboratory models, the brown rat (*Rattus norvegicus*) and the house mouse (*Mus musculus*), which split at approximately 21 MYA (Kumar S. et al., 2017). Both rats and mice have been used to describe the distribution of MCH-ir elements (Bittencourt et al., 1992; Elias et al., 2001; Diniz et al., 2019). The Cricetids include a diverse group of species, including voles (Arvicolinae), hamsters (Cricetinae), and deer mice (Neotominae) (**Figure 7**). The MCH system has been described in two hamsters, the golden hamster (*Mesocricetus auratus*) and the Siberian hamster (*Phodopus sungorus*) (Khorooshi and Klingenspor, 2005; Vidal et al., 2005), in addition to the Neotomine Mexican volcano mouse (*Neotomodon alstoni*) (Diniz et al., 2019).

As mentioned before, the presence of MCH in the LHA is shared among all Mammals, and Rodents are no exception. In all cases, the largest cluster of MCH neurons is found in the dorsolateral tuberal LHA, lateral to the medial part of the internal capsule and cerebral peduncle (**Figure 6**). In *R. norvegicus*, *N. alstoni*, and *P. sungorus*, a large number of MCH neurons is observed within the limits of the perifornical nucleus, while *M. musculus* and *M. auratus* have only scattered cells in this area. The mediolateral extent of MCH cells within the LHA is also variable, with a continuous band of neurons from the pericapsular part to the medial hypothalamus observed in *R. norvegicus*, while in *M. musculus* MCH cells are mostly restricted to the area lateral to the fornix. Dorsal to the LHA is a group of cells in the ventral ZI, observed in all Rodent species. The lateral cells are observed until the tuberomammillary level and disappear before the mammillary nuclei are fully formed. A comparison on the three-dimensional distribution of MCH cells in rats, mice, and volcano mice is available in **Figure 8**.

The medial zone of the hypothalamus, on the other hand, shows more variability between species. At the dorsalmost part of

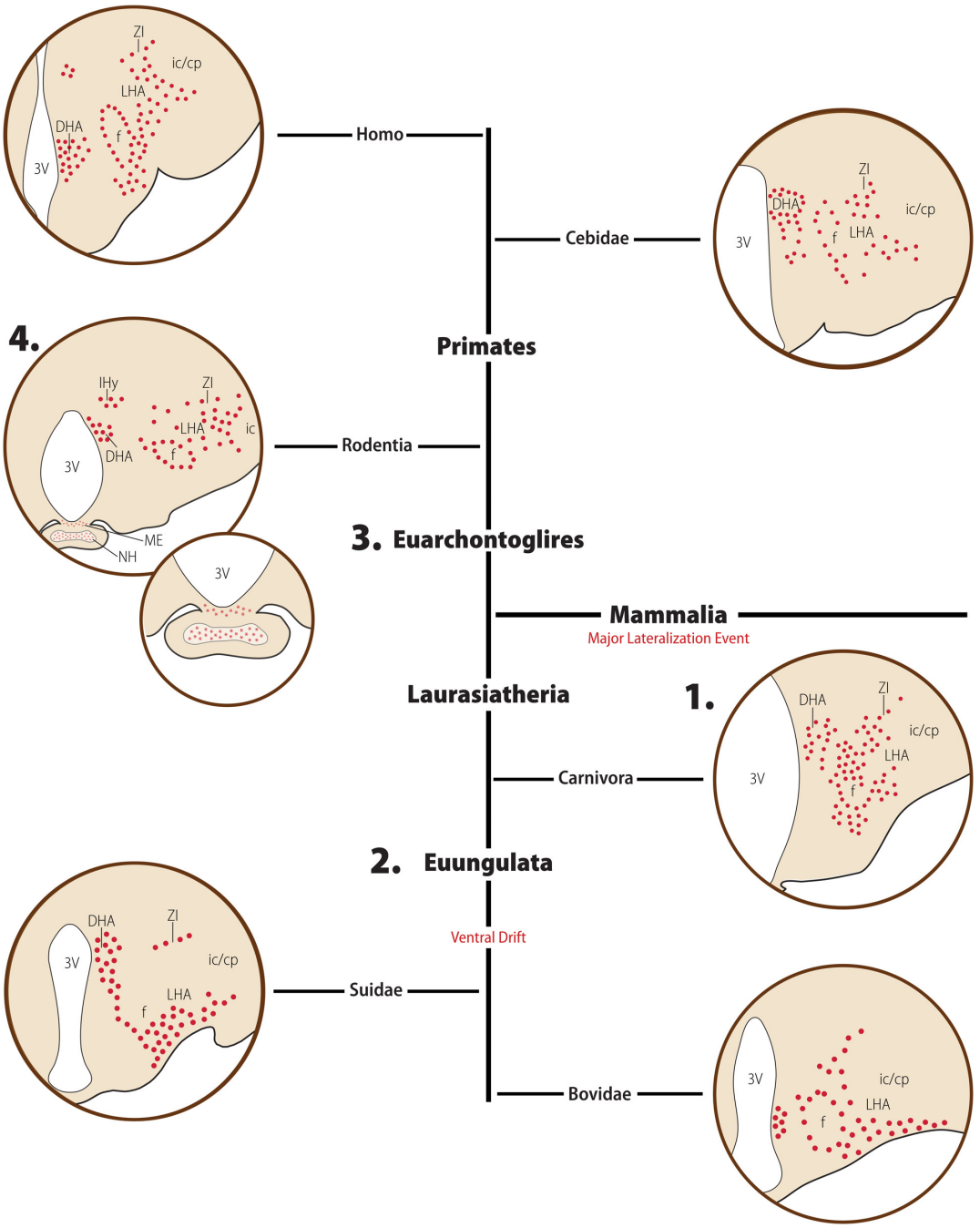


**FIGURE 6 |** Diagrammatic representation of the main morphological features of the MCH distribution in Mammals. Major clades are indicated in bold font, while clade-wide events are represented in red. The diagrams for each clade are not meant to represent a single species or a single hypothalamic level, but rather a visual summary of what has been described for animals belonging to that clade. Red circles indicate MCH neurons, red dots indicate neurosecretory fibers, and red stars indicate species-specific presence of fibers. The rat was chosen to represent rodents, to avoid the complexity that representing intra-Muroid variations would bring. In the Laurasiatheria branch, numerous MCH neurons are found in the ventrolateral hypothalamus, representing a ventral shift in this group when compared to Euarchontoglires. **1.** In Carnivora, there is a substantial group of labeled neurons in the medial hypothalamus, including the dorsomedial hypothalamic area (DHA), while the lateral part of the lateral hypothalamic area (LHA) is less densely populated. A few neurons are observed in the neighboring zona incerta (ZI). **2.** A very dense distribution of neurons in the ventral hypothalamus is observed in Euungulates. **3.** In Euarchontoglires, the dorsolateral LHA becomes the main locus of MCH neurons, with fewer neurons found in the medial hypothalamus and ZI. In all Mammals, there is a strong association of MCH neurons and the fornix and a DHA group of neurons, possible representing the ancestral group of periependymal neurons seen in Petromyzontids. **4.** In rodents, an additional group of neurons is observed in the incerto-hypothalamic area (IHy), and MCH-ir fibers are observed in the median eminence (ME) and neurohypophysis (NH) of the rat, but not in the closely related mouse. Based on data from Torterolo et al. (2006); Tillet et al. (1996), Chometton et al. (2014); Bittencourt et al. (1992), Bittencourt et al. (1998) and Krolewski et al. (2010).

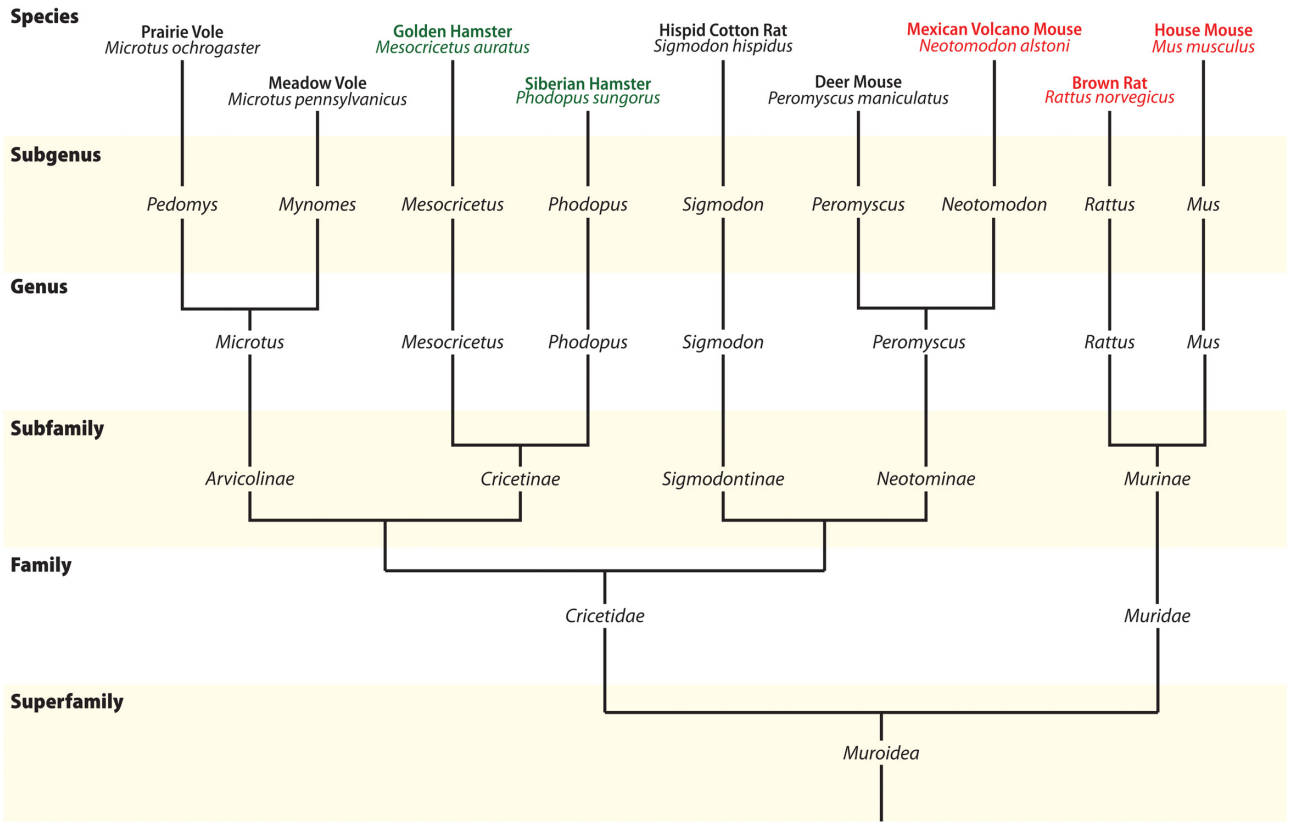


**FIGURE 7 |** Phylogenetic relationship between Muroids. Illustrative tree showing the phylogenetic relationship between members of Superfamily Muroidea, with several groups omitted for clarity. Commonly used animal models are indicated at the **top** of the tree, and species with available morphological data for MCH are indicated in green and red. Molecular data to build the tree comes from Fabre et al. (2012) and Platt et al. (2015). Reproduced with permission from Diniz et al. (2019).

the medial zone of the hypothalamus rests a small group of MCH neurons intermingled with dopaminergic neurons the A13 group, corresponding to the IHy (Sita et al., 2003, 2007). The existence of a neurochemically defined IHy has been demonstrated in *R. norvegicus* (Sita et al., 2003, 2007) and *M. musculus* (Diniz et al., 2019), and has been suggested in *N. alstoni* and *M. auratus* (Vidal et al., 2005; Diniz et al., 2019). Ventral to the IHy is one of the two dorsomedial groups of MCH cells found in Rodents, the anterior hypothalamic area group (AHA). In *R. norvegicus*, these cells are scarce and separated from the IHy, while numerous neurons are found in the AHA of *M. musculus* and *N. alstoni*, and this group is virtually continuous with the IHy. The second group of dorsomedial MCH cells is found in the PHA, in close association with the 3V. This group is observed in *R. norvegicus*, *N. alstoni*, and *P. sungorus*, but not in *M. musculus* or *M. auratus*. In addition to those groups, rats and mice have been shown to have an additional time-sensitive group of MCH neurons that is detectable only in the medial preoptic area, preoptic periventricular nucleus, and anterior paraventricular hypothalamic area of lactating animals (Knollema et al., 1992; Rondini et al., 2010; Alvisi et al., 2016; Ferreira et al., 2017b; Costa et al., 2019; Diniz et al., 2019). Several extra-diencephalic groups of MCH neurons have been identified exclusively in rats. In the basal forebrain, MCH neurons are found in the olfactory tubercle, and in the brain stem, immunolabeling is found in the laterodorsal tegmental nucleus, only in females, and in the paramedian pontine reticular formation (Bittencourt et al., 1992; Bittencourt and Diniz, 2018). Brainstem MCH neurons have also been observed in the cat (Costa et al., 2018). Regarding the distribution of fibers, almost all areas of the rat CNS receive MCH-ir fibers to some extent, except for some motor nuclei of the hindbrain (Bittencourt et al., 1992).

As the available data reveals, a substantial expansion of the MCH system occurred in Rodents. The LHA is the main group of neurons, as in other Mammals, but this group is the most extensive in Rodents, including the whole pericapsular LHA, and often the medial LHA and the perifornical nucleus. This distribution is more lateral than what is observed in Carnivores, and more dorsal than what is observed in Euungulates, and varies substantially among Rodents. As mentioned before, such variability may be linked to preferential access to some parts of the *mfb*, leading to richer innervations of some areas when compared to others. These differences may be linked to the vastly different habits and behaviors displayed by Mammals. The dorsomedial group of cells is also substantially variable within the rostrocaudal and mediolateral axes. In rats, dorsomedial MCH neurons are very close to the third ventricle, and neuroendocrine areas are densely innervated (Bittencourt et al., 1992;

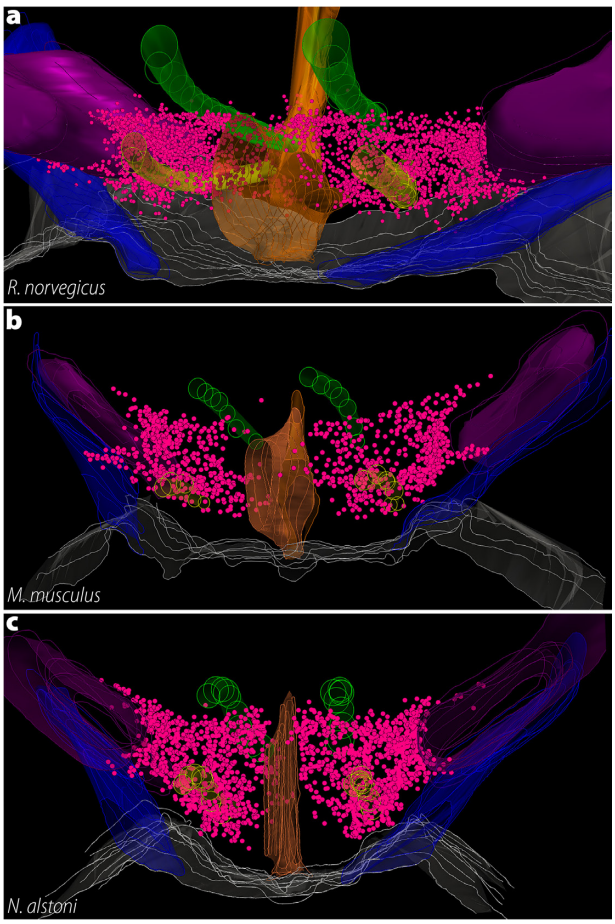


**FIGURE 8 |** The three-dimensional distribution of MCH neurons in Muroids. Hypothalamic reconstructions from *R. norvegicus* **(a)**, *M. musculus* **(b)**, and *N. alstoni* **(c)**. Each neuron is indicated by a magenta dot. Notable differences are observed when the three species are compared. The distribution of MCH neurons occupies virtually the whole medio-lateral extent of the tuberal hypothalamus, with neurons ranging from the ventricle border to the internal margin of the internal capsule. Only a small number of neurons is found ventral to the fornix. In contrast, the mouse periventricular nucleus and the medial part of the lateral hypothalamic area (LHA) are mostly devoid of neurons, with the anterior hypothalamic area and the dorsolateral LHA representing the largest groups. In *N. alstoni*, an intermediate profile is observed, with moderate numbers of neurons observed in the medial LHA and the periventricular area. The Mexican volcano mouse, however, distinguishes itself for a substantial presence of MCH neurons ventral to the fornix, in the perifornical area. Animated versions of these reconstructions can be found in Diniz et al. (2019). Structures: 3V, third ventricle (orange); f, fornix (yellow); ic, internal capsule (purple); mt, mammillothalamic tract (green); opt, optic tract (blue). Reproduced with permission from Diniz et al. (2019).

Diniz et al., 2019), while in mice these neurons drifted away from the periventricular zone and neuroendocrine areas appear to receive fewer fibers (Diniz et al., 2019).

Although the contact between MCH and the ventricular lumen first observed in Petromyzontids appears to have been lost in the early Sarcopterygii lineage, at least in Rodents, this contact appears to have been reacquired, resulting in a volume transmission mode of communication for MCH neurons (Noble et al., 2018). Other aspects of the distribution of MCH neurons in Rodents may have counterparts in distant species, including the time-sensitive appearance of MCH neurons linked to reproductive stage (Petromyzontids, Anurans), neurons in the basal forebrain (Dipnoi), neurons in the brainstem (Carnivora), and neurons in a transitional structure between the hypothalamus and the subthalamus, in proximity to dopaminergic neurons (Anurans). The multiple independent acquisitions of the same morphological aspects indicate that, despite the frequent occurrence of losses, it is likely that underlying properties of the peptidergic system facilitate the reacquisition of lost characteristics, despite substantial divergence in terms of habitat and behavior among species.

## Boreoeutheria > Euarchontoglires > Primates

Only two species of Primates have been examined in terms of MCH morphology: the tufted capuchin monkey, of the *Sapajus* genus, and humans (**Figure 4**). *Sapajus* spp. are new-world monkeys, member of Family Cebidae, which diverged at approximately 43 MYA. Member of genus Homo are believed to have split from their closest relatives, Pan, at approximately 6.7 MYA (Kumar S. et al., 2017). It should be noted that, in the original description of MCH in a new world monkey, the species has been identified as *Cebus apella* (Bittencourt et al., 1998). The taxonomic identification of the monkeys used in that experiment has been revised, as those animals are now more closely identified with members of genus *Sapajus*, and there is some controversy in the precise species definition (for a brief discussion on this subject, see Battagello et al., 2017). In *Sapajus* spp., MCH neurons are found exclusively in the diencephalon. These neurons are found from the caudal levels of the paraventricular nucleus up to the level of the medial mammillary nucleus, a distribution in the rostrocaudal axis that is slightly longer than Rodents, but shorter than Laurasiatherians. In the anterior tuberal hypothalamus, cells were observed dorsal to the fornix and in the dorsal part of the periventricular nucleus, in addition to the lateral ZI, but there is no indication that cells are found in the IHy area. At more caudal levels, the main group of neurons is observed occupying the LHA, but the medial zone is devoid of neurons (**Figure 6**). At mammillary levels, neurons are found dorsal to the medial mammillary nucleus and ventral to the mesencephalic aqueduct. Fibers were found in the medial mammillary nucleus, the external layer of the ME, and the lateral *globus pallidus*, but a complete mapping of fibers has not been published (Bittencourt et al., 1998).

Several authors investigated the distribution of MCH neurons in humans of male and female individuals, both through *in situ* hybridization for *PMCH* (Elias et al., 1998; Krolewski et al., 2010) or immunohistochemistry for MCH (Thannickal et al., 2007; Aziz et al., 2008). Due to how detailed is the description provided by Krolewski et al. (2010), it will be used as a base to describe the distribution of MCH in *H. sapiens*. Neurons expressing *PMCH* mRNA are first detected in the LHA, at the intermediate levels of the PVH. While Krolewski et al. (2010) identified these anterior neurons as belonging to the LHA, Elias et al. (1998) identified

them as part of the rostromedial ZI, possibly corresponding to the Muroid IHy. As the PVH nears its caudal end, the distribution of *PMCH* mRNA-expressing neurons expands to the lateral aspects of the LHA, and two new groups are observed: the dorsomedial hypothalamic nucleus and the DHA (Krolewski et al., 2010). Together, these groups give the impression that MCH neurons surround the whole tuberomammillary extent of the fornix (**Figure 6**; Elias et al., 1998). A dense cluster of MCH neurons in the DHA forms a ring-shaped structure around a central core of non-labeled cells (Krolewski et al., 2010). This observation is particularly interesting, as Diniz et al. (2019) recently proposed that, in *M. musculus*, MCH neurons form a ring-shaped structure in the LHA surrounding ORX neurons, a neurochemical division called the LHA shell. It is possible that a similar feature appears in humans, with MCH neurons surrounding another population of neurons. Finally, at the level of the mammillary bodies, the DHA continues into the PHA, which assumes a position close to the 3V.

In addition to the distribution of neurons, the number of MCH neurons has also been investigated, particularly in studies evaluating hypothalamic correlates in neurodegenerative diseases. Thannickal et al. (2007) report a loss of MCH neurons throughout the anteroposterior axis that is correlated with Parkinson's Disease stage, ranging from 12% of loss in Stage I to 74% of loss in Stage V. Aziz et al. (2008) performed a similar experiment in Huntington's Disease patients, but found no loss of MCH neurons in patients as compared to the controls. Regarding immunoreactive fibers, only Elias et al. (1998) reported fibers immunoreactivity, citing areas such as the cingulate and insular cortex, amygdala, hippocampus, anterior thalamic nucleus, preoptic hypothalamus, and mammillary bodies are recipients of fibers. These innervation fields are compatible with what has been described in other mapping studies for mammals.

Considering the data available, the distribution of MCH neurons in Primates appears to be mostly conserved. Both *S. apella* and *H. sapiens* share the same rostrocaudal extent, and the areas containing the most substantial numbers of neurons are the same. Some crucial questions are still open, however. It is unclear, at this point, if Primates have an IHy similar to Murids, where MCH neurons are found intermingled with tyrosine hydroxylase-synthesizing neurons. The second question is the presence of MCH neurons in the ZI, as Bittencourt et al. (1998) reported cells in the lateral ZI of *Sapajus* spp., but the same has not been reported in humans (Elias et al., 1998; Krolewski et al., 2010). Furthermore, there is no complete mapping of MCH fibers in the human brain and no physical evidence that NEI is synthesized in Primates. The overall pattern observed in Primates appears to be unique to this clade, as the rostrocaudal extent of MCH neurons is longer in Primates than it is in Rodents but shorter than what has been described for Laurasiatherians. It is a common feature between Euarchontoglires, however, the predominance of neurons dorsal to the fornix, and a lateromedial presence of neurons that spans the lateral and medial zones, with only the rat presenting a dense cluster of cells in the periventricular nucleus. There is a high number of labeled neurons around the fornix in several Euarchontoglires species (with the exception of mice and golden hamsters), but it is possible that the neurons observed in humans are actually part of other nuclei in the hypothalamus, compressed against the fornix due to the reduced lateromedial extent of the hypothalamus in the Primate lineage.

## CONCLUSION

The melanin-concentrating hormone peptidergic system is a versatile neurochemical system, demonstrated to play a varied range of roles within the Rodent CNS. Such functional diversity is only possible thanks to a complex morphological substrate, which has been investigated in a substantial number of species. This morphological substrate, however, is not a universal feature, varying among species and clades, likely reflecting differences in habitat and behavior. This interplay between MCH probably developed from the initial function of MCH: broadcaster of external stimuli. By contacting the ventricular lumen with fine projections while projecting to the lateral hypothalamus and thalamus, MCH neurons were able to coordinate metabolic signals present in the CSF with effector centers of the diencephalon. As vertebrates evolved to display more complex behaviors and a richer relationship with their environment, MCH neurons that could integrate not only endogenous signals, but also external inputs were positively selected, and this is reflected in a drift of MCH cells away from the ventricle. This aligns well with the idea that MCH is a major integrator of internal and external information, ensuring an appropriate response to ensure the organism homeostasis. It would not be possible, however, if all organisms shared identical MCH organizations, given the different challenges imposed by different aquatic/terrestrial environments, food chain positions, and reproductive strategies. New studies of the MCH system in species that have not been investigated yet will help us understand more precisely how these habitat changes are connected to the hypothalamic neurochemical circuits, paving the way to new intervention strategies that may be used with pharmacological purposes.

## AUTHOR CONTRIBUTIONS

GD and JB contributed with the writing of this article and approved it for publication.

## FUNDING

Financial support for this study was provided in the form of grants from the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, São Paulo Research Foundation; Grant Nos. 2010/52068-0 and 2016/02224-1 to JB and 2016/02748-0 to GD). We are thankful for the support provided by the Brazilian Coordenação de Aperfeiçoamento de Pessoal de Nível Superior CAPES, Office for the Advancement of Higher Education, Grant #848/15. JB was an Investigator with the Brazilian Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, National Council for Scientific and Technological Development, Grant # 426378/2016-4).

## ACKNOWLEDGMENTS

We would like to thank the colleagues who read and provided critical commentary on this work: Prof. Paul E. Sawchenko (Salk Institute for Biological Studies), Prof. Anthony N. van den Pol (Yale School of Medicine), and Dr. Luciane V. Sita (University of São Paulo). We would also like to thank Dr. Victor Luria (Harvard Medical School) for suggesting the phylostratigraphy positioning of Pmch.

## REFERENCES

Alachkar, A., Alhassen, L., Wang, Z., Wang, L., Onouye, K., Sanathara, N., et al. (2016). Inactivation of the melanin concentrating hormone system impairs maternal behavior. *Eur. Neuropsychopharmacol.* 26, 1826–1835. doi: 10.1016/j.euroneuro.2016.08.014

Alexander, S. P., Christopoulos, A., Davenport, A. P., Kelly, E., Marrion, N. V., Peters, J. A., et al. (2017). The concise guide to PHARMACOLOGY 2017/18: G protein−coupled receptors. *Br. J. Pharmacol.* 174, S17–S129. doi: 10.1111/bph.13878

Alhassen, L., Phan, A., Alhassen, W., Nguyen, P., Lo, A., Shaharuddin, H., et al. (2019). The role of olfaction in MCH-regulated spontaneous maternal responses. *Brain Res.* 1719, 71–76. doi: 10.1016/j.brainres.2019.05.021

Allendorf, F., and Thorgaard, G. (1984). "Tetraploidy and the evolution of salmonid fishes," in *in Evolutionary Genetics of Fishes*, ed. J. Turner, (New York, NY.: Plenum Press), 1–46.

Alvisi, R. D., Diniz, G. B., Da-Silva, J. M., Bittencourt, J. C., and Felicio, L. F. (2016). Suckling-induced Fos activation and melanin-concentrating hormone immunoreactivity during late lactation. *Life Sci.* 148, 241–246. doi: 10.1016/j.lfs.2016.02.038

Al-Yousuf, S. A., and Mizuno, N. (1991). Electron microscopic identification of axons containing melanin-concentrating hormone in the lamprey. *Lampetra fluviatilis L. Neurosci. Lett.* 128, 249–252. doi: 10.1016/0304-3940(91)90272-u

Amano, M., Takahashi, A., Oka, Y., Yamanome, T., Kawauchi, H., and Yamamori, K. (2003). Immunocytochemical localization and ontogenic development of melanin-concentrating hormone in the brain of a pleuronectiform fish, the barfin flounder. *Cell Tissue Res.* 311, 71–77. doi: 10.1007/s00441-002-0660-6

Amiya, N., Amano, M., Iigo, M., Yamanome, T., Takahashi, A., and Yamamori, K. (2008a). Interaction of orexin/hypocretin-like immunoreactive neurons with melanin-concentrating hormone and α-melanocyte-stimulating hormone neurons in brain of a pleuronectiform fish, barfin flounder. *Fish. Sci.* 74, 1040–1046. doi: 10.1111/j.1444-2906.2008.01622.x

Amiya, N., Amano, M., Yamanome, T., Yamamori, K., and Takahashi, A. (2008b). Effects of background color on GnRH and MCH levels in the barfin flounder brain. *Gen. Comp. Endocrinol.* 155, 88–93. doi: 10.1016/j.ygcen.2007.03.007

An, S., Cutler, G., Zhao, J. J., Huang, S.-G., Tian, H., Li, W., et al. (2001). Identification and characterization of a melanin-concentrating hormone receptor. *Proc. Natl. Acad. Sci. U.S.A.* 98, 7576–7581.

Andersen, A. C., Pelletier, G., Eberle, A. N., Leroux, P., Jegou, S., and Vaudry, H. (1986). Localization of melanin-concentrating hormone-like immunoreactivity in the brain and pituitary of the frog Rana ridibunda. *Peptides* 7, 941–951. doi: 10.1016/0196-9781(86)90119-1

Aziz, A., Fronczek, R., Maat−Schieman, M., Unmehopa, U., Roelandse, F., Overeem, S., et al. (2008). Hypocretin and melanin−concentrating hormone in patients with Huntington disease. *Brain Pathol.* 18, 474–483. doi: 10.1111/j.1750-3639.2008.00135.x

Bächner, D., Kreienkamp, H.-J., Weise, C., Buck, F., and Richter, D. (1999). Identification of melanin concentrating hormone (MCH) as the natural ligand for the orphan somatostatin−like receptor 1 (SLC−1). *FEBS Lett.* 457, 522–524. doi: 10.1016/s0014-5793(99)01092-3

Badami, V. M., Rice, C. D., Lois, J. H., Madrecha, J., and Yates, B. J. (2010). Distribution of hypothalamic neurons with orexin (hypocretin) or melanin concentrating hormone (MCH) immunoreactivity and multisynaptic connections with diaphragm motoneurons. *Brain Res.* 1323, 119–126. doi: 10.1016/j.brainres.2010.02.002

Baker, B., Levy, A., Hall, L., and Lightman, S. (1995). Cloning and expression of melanin-concentrating hormone genes in the rainbow trout brain. *Neuroendocrinology* 61, 67–76. doi: 10.1159/000126814

Baker, B. I., and Bird, D. J. (2002). Neuronal organization of the melanin−concentrating hormone system in primitive actinopterygians: Evolutionary changes leading to teleosts. *J. Comp. Neurol.* 442, 99–114. doi: 10.1002/cne.10074

Baker, B. I., and Rance, T. A. (1983). Further observations on the distribution and properties of teleost melanin concentrating hormone. *Gen. Comp. Endocrinol.* 50, 423–431. doi: 10.1016/0016-6480(83)90263-0

Battagello, D. S., Diniz, G. B., Candido, P. L., Da Silva, J. M., De Oliveira, A. R., Torres Da Silva, K. R., et al. (2017). Anatomical organization of urocortin 3-Synthesizing Neurons and immunoreactive terminals in the central nervous system of non-human primates [*Sapajus* spp.]. *Front. Neuroanat.* 11:57. doi: 10.3389/fnana.2017.00057

Batten, T. F. C., and Baker, B. I. (1988). Melanin-concentrating hormone (MCH) immunoreactive hypophysial neurosecretory system in the teleost Poecilia latipinna: light and electron microscopic study. *Gen. Comp. Endocrinol.* 70, 193–205. doi: 10.1016/0016-6480(88)90140-2

Batten, T. F. C., Moons, L., and Vandesande, F. (1999). Innervation and control of the adenohypophysis by hypothalamic peptidergic neurons in teleost fishes: EM immunohistochemical evidence. *Microsc. Res. Tech.* 44, 19–35. doi: 10.1002/(sici)1097-0029(19990101)44:1<19::aid-jemt4>3.3.co;2-c

Benedetto, L., Pereira, M., Ferreira, A., and Torterolo, P. (2014). Melanin-concentrating hormone in the medial preoptic area reduces active components of maternal behavior in rats. *Peptides* 58, 20–25. doi: 10.1016/j.peptides.2014.05.012

Berman, J. R., Skariah, G., Maro, G. S., Mignot, E., and Mourrain, P. (2009). Characterization of two melanin−concentrating hormone genes in zebrafish reveals evolutionary and physiological links with the mammalian MCH system. *J. Comp. Neurol.* 517, 695–710. doi: 10.1002/cne.22171

Bird, D. J., Baker, B. I., and Kawauchi, H. (1989). Immunocytochemical demonstration of melanin-concentrating hormone and proopiomelanocortin-like products in the brain of the trout and carp. *Gen. Comp. Endocrinol.* 74, 442–450. doi: 10.1016/s0016-6480(89)80042-5

Bird, D. J., Potter, I. C., Sower, S. A., and Baker, B. I. (2001). The distribution of melanin-concentrating hormone in the lamprey brain. *Gen. Comp. Endocrinol.* 121, 232–241. doi: 10.1006/gcen.2001.7609

Bittencourt, J. C., and Diniz, G. B. (2018). "Neuroanatomical Structure of the MCH System," in *Melanin-Concentrating Hormone and Sleep*, 1st Edn. eds S. R. Pandi-Perumal, P. Torterolo, and M. Jaime, (Berlin: Springer), 1–46. doi: 10.1007/978-3-319-75765-0_1

Bittencourt, J. C., Frigo, L., Rissman, R. A., Casatti, C. A., Nahon, J.-L., and Bauer, J. A. (1998). The distribution of melanin-concentrating hormone in the monkey brain (*Cebus apella*). *Brain Res.* 804, 140–143. doi: 10.1016/s0006-8993(98)00662-3

Bittencourt, J. C., Presse, F., Arias, C., Peto, C. A., Vaughan, J. M., Nahon, J.-L., et al. (1992). The melanin-concentrating hormone system of the rat brain: an immuno- and hybridization histochemical characterization. *J. Comp. Neurol.* 319, 218–245. doi: 10.1002/cne.903190204

Blair, J. E., and Hedges, S. B. (2005). Molecular phylogeny and divergence times of deuterostome animals. *Mol. Biol. Evol.* 22, 2275–2284. doi: 10.1093/molbev/msi225

Blake, J. A., Eppig, J. T., Kadin, J. A., Richardson, J. E., Smith, C. L., Bult, C. J., et al. (2016). Mouse genome database (MGD)-2017: community knowledge resource for the laboratory mouse. *Nucleic Acids Res.* 45, D723–D729. doi: 10.1093/nar/gkw1040

Borsu, L., Presse, F., and Nahon, J.-L. (2000). The AROM gene, spliced mRNAs encoding new DNA/RNA-binding proteins are transcribed from the opposite strand of the melanin-concentrating hormone gene in mammals. *J. Biol. Chem.* 275, 40576–40587. doi: 10.1074/jbc.m006524200

Braschi, B., Denny, P., Gray, K., Jones, T., Seal, R., Tweedie, S., et al. (2018). Genenames. org: the HGNC and VGNC resources in 2019. *Nucleic Acids Res.* 47, D786–D792. doi: 10.1093/nar/gky930

Brodin, L., Theordorsson, E., Christenson, J., Cullheim, S., Hökfelt, T., Brown, J. C., et al. (1990). Neurotensin−like peptides in the CNS of lampreys: chromatographic characterization and immunohistochemical localization with reference to aminergic markers. *Eur. J. Neurosci.* 2, 1095–1109. doi: 10.1111/j.1460-9568.1990.tb00021.x

Burt, D. W., Carrë, W., Fell, M., Law, A. S., Antin, P. B., Maglott, D. R., et al. (2009). The chicken gene nomenclature committee report. *BMC Genomics* 2(Suppl. 2):S5. doi: 10.1186/1471-2164-10-S2-S5

Cánepa, M., Pozzi, A., Astola, A., Maggese, M. C., and Vissio, P. (2008). Effect of salmon melanin-concentrating hormone and mammalian gonadotrophin-releasing hormone on somatolactin release in pituitary culture of Cichlasoma dimerus. *Cell Tissue Res.* 333:49. doi: 10.1007/s00441-008-0622-8

Cardot, J., Fellmann, D., and Bugnon, C. (1994). Melanin-concentrating hormone-producing neurons in reptiles. *Gen. Comp. Endocrinol.* 94, 23–32. doi: 10.1006/gcen.1994.1056

Cardot, J., Griffond, B., Risold, P. Y., Blähser, S., and Fellmann, D. (1999). Melanin−concentrating hormone−producing neurons in birds. *J. Comp. Neurol.* 411, 239–256. doi: 10.1002/(sici)1096-9861(19990823)411:2<239::aid-cne5>3.0.co;2-7

Casola, C., and Betrán, E. (2017). The genomic impact of gene retrocopies: what have we learned from comparative genomics, population genomics, and transcriptomic analyses? *Genome Biol. Evol.* 9, 1351–1373. doi: 10.1093/gbe/evx081

Cerdá-Reverter, J. M., Canosa, L. F., and Peter, R. E. (2006). Regulation of the hypothalamic melanin-concentrating hormone neurons by sex steroids in the goldfish: possible role in the modulation of luteinizing hormone secretion. *Neuroendocrinology* 84, 364–377. doi: 10.1159/000098334

Chaillou, E., Baumont, R., Fellmann, D., Tramu, G., and Tillet, Y. (2003). Sensitivity of Galanin−and Melanin−Concentrating Hormone−containing neurones to nutritional status: an immunohistochemical study in the ovariectomized ewe. *J. Neuroendocrinol.* 15, 459–467. doi: 10.1046/j.1365-2826.2003.00998.x

Chambers, J., Ames, R. S., Bergsma, D., Muir, A., Fitzgerald, L. R., Hervieu, G., et al. (1999). Melanin-concentrating hormone is the cognate ligand for the orphan G-protein-coupled receptor SLC-1. *Nature* 400, 261–265. doi: 10.1038/22313

Chee, M. J., Hebert, A. J., Briançon, N., Flaherty Iii, S. E., Pissios, P., and Maratos-Flier, E. (2019). Conditional deletion of melanin-concentrating hormone receptor 1 from GABAergic neurons increases locomotor activity. *Mol. Metab.* 29, 114–123. doi: 10.1016/j.molmet.2019.08.018

Chometton, S., Franchi, G., Houdayer, C., Mariot, A., Poncet, F., Fellmann, D., et al. (2014). Different distributions of preproMCH and hypocretin/orexin in the forebrain of the pig (*Sus scrofa* domesticus). *J. Chem. Neuroanat.* 61, 72–82. doi: 10.1016/j.jchemneu.2014.08.001

Costa, A., Castro-Zaballa, S., Lagos, P., Chase, M. H., and Torterolo, P. (2018). Distribution of MCH-containing fibers in the feline brainstem: Relevance for REM sleep regulation. *Peptides* 104, 50–61. doi: 10.1016/j.peptides.2018.04.009

Costa, H. C., Da-Silva, J. M., Diniz, G. B., Da-Silva Pereira, L., Gobbo, D. R., Da-Silva, R. J., et al. (2019). Characterization and origins of melanin-concentrating hormone immunoreactive fibers of the posterior lobe of the pituitary and median eminence during lactation in Long-Evans rat. *J. Neuroendocrinol.* 31:e12723.

Courseaux, A., and Nahon, J.-L. (2001). Birth of two chimeric genes in the hominidae lineage. *Science* 291, 1293–1297. doi: 10.1126/science.1057284

Croizier, S., Cardot, J., Brischoux, F., Fellmann, D., Griffond, B., and Risold, P. (2013). The vertebrate diencephalic MCH system: a versatile neuronal population in an evolving brain. *Front. Neuroendocrinol.* 34:65–87. doi: 10.1016/j.yfrne.2012.10.001

Diniz, G. B., Battagello, D. S., Cherubini, P. M., Reyes-Mendoza, J. D., Luna-Illades, C., Klein, M. O., et al. (2019). Melanin- concentrating hormone peptidergic system: comparative morphology between muroid species. *J. Comp. Neurol.* 527, 2973–3001. doi: 10.1002/cne.24723

Diniz, G. B., and Bittencourt, J. C. (2017). The melanin-concentrating hormone as an integrative peptide driving motivated behaviors. *Front. Syst. Neurosci.* 11:32. doi: 10.3389/fnsys.2017.00032

Domazet-Lošo, T., Brajkoviæ, J., and Tautz, D. (2007). A phylostratigraphy approach to uncover the genomic history of major adaptations in metazoan lineages. *Trends Genet.* 23, 533–539. doi: 10.1016/j.tig.2007.08.014

Domazet-Lošo, T., and Tautz, D. (2010). Phylostratigraphic tracking of cancer genes suggests a link to the emergence of multicellularity in metazoa. *BMC Biol.* 8:66. doi: 10.1186/1741-7007-8-66

Duarte, G., Segura-Noguera, M., Del Río, M. M., and Mancera, J. (2001). The hypothalamo-hypophyseal system of the white seabream Diplodus sargus: immunocytochemical identification of arginine-vasotocin, isotocin, melanin-concentrating hormone and corticotropin-releasing factor. *Histochem. J.* 33, 569–578.

Elias, C. F., Lee, C. E., Kelly, J. F., Ahima, R. S., Kuhar, M., Saper, C. B., et al. (2001). Characterization of CART neurons in the rat and human hypothalamus. *J. Comp. Neurol.* 432, 1–19. doi: 10.1002/cne.1085

Elias, C. F., Saper, C. B., Maratos−Flier, E., Tritos, N. A., Lee, C., Kelly, J., et al. (1998). Chemically defined projections linking the mediobasal hypothalamus and the lateral hypothalamic area. *J. Comp. Neurol.* 402, 442–459. doi: 10.1002/(sici)1096-9861(19981228)402:4<442::aid-cne2>3.3.co;2-i

Elphick, M. R., Mirabeau, O., and Larhammar, D. (2018). Evolution of neuropeptide signalling systems. *J. Exp. Biol.* 221:jeb151092. doi: 10.1242/jeb.151092

Fabre, P.-H., Hautier, L., Dimitrov, D., and Douzery, E. J. P. (2012). A glimpse on the pattern of rodent diversification: a phylogenetic approach. *BMC Evol. Biol.* 12:88. doi: 10.1186/1471-2148-12-88

Ferreira, J. G. P., Bittencourt, J. C., and Adamantidis, A. (2017a). Melanin-concentrating hormone and sleep. *Curr. Opin. Neurobiol.* 44, 152–158.

Ferreira, J. G. P., Duarte, J. C. G., Diniz, G. B., and Bittencourt, J. C. (2017b). Litter size determines the number of melanin-concentrating hormone neurons in the medial preoptic area of sprague dawley lactating dams. *Physiol. Behav.* 181, 75–79. doi: 10.1016/j.physbeh.2017.08.028

Francis, K., and Baker, B. I. (1995). Developmental changes in melanin-concentrating hormone in Rana temporaria. *Gen. Comp. Endocrinol.* 98, 157–165. doi: 10.1006/gcen.1995.1056

Friedman, M. (2010). Explosive morphological diversification of spiny-finned teleost fishes in the aftermath of the end-Cretaceous extinction. *Proc. R. Soc. B Biol. Sci.* 277, 1675–1683. doi: 10.1098/rspb.2009.2177

Fu, B., Chen, M., Zou, M., Long, M., and He, S. (2010). The rapid generation of chimerical genes expanding protein diversity in zebrafish. *BMC Genomics* 11:657. doi: 10.1186/1471-2164-11-657

Gao, X.-B. (2018). "The role of melanin-concentrating hormone in the regulation of the sleep/wake cycle: sleep promoter or arousal modulator?," in *Melanin-Concentrating Hormone and Sleep : Molecular, Functional and Clinical Aspects*, eds S. R. Pandi-Perumal, P. Torterolo, and J. M. Monti, (Cham: Springer International Publishing), 57–74. doi: 10.1007/978-3-319-75765-0_3

Gess, R. W., Coates, M. I., and Rubidge, B. S. (2006). A lamprey from the devonian period of South Africa. *Nature* 443, 981–984. doi: 10.1038/nature05150

Gröneveld, D., Hut, M. J., Balm, P. H. M., Martens, G. J. M., and Bonga, S. E. W. (1993). Cloning and sequence analysis of hypothalamus cDNA encoding tilapia melanin-concentrating hormone. *Fish Physiol. Biochem.* 11, 117–124. doi: 10.1007/BF00004557

Grzybowska, E. A. (2012). Human intronless genes: functional groups, associated diseases, evolution, and mRNA processing in absence of splicing. *Biochem. Biophys. Res. Commun.* 424, 1–6. doi: 10.1016/j.bbrc.2012.06.092

Hahn, J. D. (2010). Comparison of melanin-concentrating hormone and hypocretin/orexin peptide expression patterns in a current parceling scheme of the lateral hypothalamic zone. *Neurosci. Lett.* 468, 12–17. doi: 10.1016/j.neulet.2009.10.047

Hill, J., Duckworth, M., Murdock, P., Rennie, G., Sabido-David, C., Ames, R. S., et al. (2001). Molecular cloning and functional characterization of MCH2, a novel human MCH receptor. *J. Biol. Chem.* 276, 20125–20129. doi: 10.1074/jbc.m102068200

Hogben, L. T., and Slome, D. (1931). The pigmentary effector system. VI. The dual character of endocrine co-ordination in amphibian colour change. *Proc. R. Soc. Lon. Ser. B, Containing Pap. Biol. Character* 108, 10–53. doi: 10.1098/rspb.1931.0020

Hosomi, N., Masumoto, T., and Fukada, H. (2015). Yellowtail melanin-concentrating hormone 1: molecular cloning, tissue distribution, and response to fasting and feeding. *Aquacu. Sci.* 63, 169–178.

Huesa, G., Van Den Pol, A. N., and Finger, T. E. (2005). Differential distribution of hypocretin (orexin) and melanin−concentrating hormone in the goldfish brain. *J. Comp. Neurol.* 488, 476–491. doi: 10.1002/cne.20610

Jaillon, O., Aury, J.-M., Brunet, F., Petit, J.-L., Stange-Thomann, N., Mauceli, E., et al. (2004). Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype. *Nature* 431, 946–957. doi: 10.1038/nature03025

Kang, D.-Y., and Kim, H.-C. (2013). Functional characterization of two melanin-concentrating hormone genes in the color camouflage, hypermelanosis, and appetite of starry flounder. *Gen. Comp. Endocrinol.* 189, 74–83. doi: 10.1016/j.ygcen.2013.04.025

Karimi, K., Fortriede, J. D., Lotay, V. S., Burns, K. A., Wang, D. Z., Fisher, M. E., et al. (2017). Xenbase: a genomic, epigenomic and transcriptomic model organism database. *Nucleic Acids Res.* 46, D861–D868. doi: 10.1093/nar/gkx936

Kasahara, M., Naruse, K., Sasaki, S., Nakatani, Y., Qu, W., Ahsan, B., et al. (2007). The medaka draft genome and insights into vertebrate genome evolution. *Nature* 447, 714–719.

Kawauchi, H., Kawazoe, I., Tsubokawa, M., Kishida, M., and Baker, B. I. (1983). Characterization of melanin-concentrating hormone in chum salmon pituitaries. *Nature* 305, 321–323. doi: 10.1038/305321a0

Khorooshi, R. M. H., and Klingenspor, M. (2005). Neuronal distribution of melanin-concentrating hormone, cocaine-and amphetamine-regulated transcript and orexin B in the brain of the Djungarian hamster (Phodopus sungorus). *J. Chem. Neuroanat.* 29, 137–148. doi: 10.1016/j.jchemneu.2004.10.003

Knollema, S., Brown, E. R., Vale, W., and Sawchenko, P. E. (1992). Novel hypothalamic and preoptic sites of prepro-melanin-concentrating hormone messenger ribonucleic acid and peptide expression in lactating rats. *J. Neuroendocrinol.* 4, 709–717. doi: 10.1111/j.1365-2826.1992.tb00222.x

Kolakowski, L. F. Jr., Jung, B. P., Nguyen, T., Johnson, M. P., Lynch, K. R., Cheng, R., et al. (1996). Characterization of a human gene related to genes encoding somatostatin receptors. *FEBS Lett.* 398, 253–258. doi: 10.1016/s0014-5793(96)01160-x

Krolewski, D. M., Medina, A., Kerman, I. A., Bernard, R., Burke, S., Thompson, R. C., et al. (2010). Expression patterns of corticotropin−releasing factor, arginine vasopressin, histidine decarboxylase, melanin−concentrating hormone, and orexin genes in the human hypothalamus. *J. Comp. Neurol.* 518, 4591–4611. doi: 10.1002/cne.22480

Kumar, A., Bhandari, A., Sarde, S. J., and Goswami, C. (2017). Ancestry & molecular evolutionary analyses of heat shock protein 47 kDa (HSP47/SERPINH1). *Sci. Rep.* 7:10394.

Kumar, S., Stecher, G., Suleski, M., and Hedges, S. B. (2017). TimeTree: a resource for timelines, timetrees, and divergence times. *Mol. Biol. Evol.* 34, 1812–1819. doi: 10.1093/molbev/msx116

Kumar, S., and Hedges, S. B. (1998). A molecular timescale for vertebrate evolution. *Nature* 392, 917–920. doi: 10.1038/31927

Kuraku, S., and Kuratani, S. (2006). Time scale for cyclostome evolution inferred with a phylogenetic diagnosis of hagfish and lamprey cDNA sequences. *Zool. Sci.* 23, 1053–1065.

Kusumi, K., Kulathinal, R. J., Abzhanov, A., Boissinot, S., Crawford, N. G., Faircloth, B. C., et al. (2011). Developing a community-based genetic nomenclature for anole lizards. *BMC Genomics* 12:554. doi: 10.1186/1471-2164-12-554

Lakaye, B., Minet, A., Zorzi, W., and Grisar, T. (1998). Cloning of the rat brain cDNA encoding for the SLC-1 G protein-coupled receptor reveals the presence of an intron in the gene. *Biochim. Biophys. Acta Mol. Cell Res.* 1401, 216–220.

Lázár, G., Maderdrut, J. L., and Merchenthaler, I. (2002). Distribution of melanin-concentrating hormone-like immunoreactivity in the central nervous system of Rana esculenta. *Brain Res. Bull.* 57, 401–407.

Lembo, P. M. C., Grazzini, E., Cao, J., Hubatsch, D. A., Pelletier, M., Hoffert, C., et al. (1999). The receptor for the orexigenic peptide melanin-concentrating hormone is a G-protein-coupled receptor. *Nat. Cell Biol.* 1, 267–271.

Macdonald, D., Murgolo, N., Zhang, R., Durkin, J. P., Yao, X., Strader, C. D., et al. (2000). Molecular characterization of the melanin-concentrating hormone/receptor complex: identification of critical residues involved in binding and activation. *Mol. Pharmacol.* 58, 217–225.

Mancera, J., and Fernández-Llebrez, P. (1995). Development of melanin-concentrating hormone-immunoreactive elements in the brain of gilthead seabream (*Sparus auratus*). *Cell and Tissue Res.* 282, 523–526.

Masao, O., Chieki, W., Ikuo, O., Ichiro, K., and Hiroshi, K. (1988). Structures of two kinds of mRNA encoding the chum salmon melanin-concentrating hormone. *Gene* 71, 433–438.

Matsuda, K., Shimakura, S.-I., Maruyama, K., Miura, T., Uchiyama, M., Kawauchi, H., et al. (2006). Central administration of melanin-concentrating hormone (MCH) suppresses food intake, but not locomotor activity, in the goldfish. *Carassius auratus*. *Neurosci. Lett.* 399, 259–263.

Meurling, P., and Rodríguez, E. M. (1990). The paraventricular and posterior recess organs of elasmobranchs: a system of cerebrospinal fluid-contacting neurons containing immunoreactive serotonin and somatostatin. *Cell Tissue Res.* 259, 463–473.

Meyer, A., and Van de Peer, Y. (2005). From 2R to 3R: evidence for a fish−specific genome duplication (FSGD). *Bioessays* 27, 937–945.

Minth, C. D., Qiu, H., Akil, H., Watson, S. J., and Dixon, J. E. (1989). Two precursors of melanin-concentrating hormone: DNA sequence analysis and in situ immunochemical localization. *Proc. Natl. Acad. Sci. U.S.A.* 86, 4292–4296.

Miyashita, T., Coates, M. I., Farrar, R., Larson, P., Manning, P. L., Wogelius, R. A., et al. (2019). Hagfish from the Cretaceous Tethys Sea and a reconciliation of the morphological–molecular conflict in early vertebrate phylogeny. *Proc. Natl. Acad. Sci. U.S.A.* 116, 2146–2151.

Mizusawa, K., Amiya, N., Yamaguchi, Y., Takabe, S., Amano, M., Breves, J. P., et al. (2012). Identification of mRNAs coding for mammalian-type melanin-concentrating hormone and its receptors in the scalloped hammerhead shark Sphyrna lewini. *Gen. Comp. Endocrinol.* 179, 78–87. doi: 10.1016/j.ygcen.2012.07.023

Mizusawa, K., Kawashima, Y., Sunuma, T., Hamamoto, A., Kobayashi, Y., Kodera, Y., et al. (2015). Involvement of melanin-concentrating hormone 2 in background color adaptation of barfin flounder Verasper moseri. *Gen. Comp. Endocrinol.* 214, 140–148. doi: 10.1016/j.ygcen.2014.07.008

Moldovan, G.-L., Dejsuphong, D., Petalcorin, M. I. R., Hofmann, K., Takeda, S., Boulton, S. J., et al. (2012). Inhibition of homologous recombination by the PCNA-interacting protein PARI. *Mol. Cell* 45, 75–86. doi: 10.1016/j.molcel.2011.11.010

Mori, M., Harada, M., Terao, Y., Sugo, T., Watanabe, T., Shimomura, Y., et al. (2001). Cloning of a novel G protein-coupled receptor, SLT, a subtype of the melanin-concentrating hormone receptor. *Bioch. Biophys. Res. Commun.* 283, 1013–1018.

Nahon, J.-L., Presse, F., Bittencourt, J. C., Sawchenko, P. E., and Vale, W. (1989). The rat melanin-concentrating hormone messenger ribonucleic acid encodes multiple putative neuropeptides coexpressed in the dorsolateral hypothalamus. *Endocrinology* 125, 2056–2065.

Nahon, J. L., Presset, F., Schoepfer, R., and Vale, W. (1991). Identification of a single melanin−concentrating hormone messenger ribonucleic acid in coho salmon: structural relatedness with 7SL ribonucleic acid. *J. Neuroendocrinol.* 3, 173–183. doi: 10.1111/j.1365-2826.1991.tb00260.x

Naito, N., Nakai, Y., Kawauchi, H., and Hayashi, Y. (1985). Immunocytochemical identification of melanin-concentrating hormone in the brain and pituitary gland of the teleost fishes *Oncorhynchus* keta and Salmo gairdneri. *Cell Tissue Res.* 242, 41–48.

Naufahu, J., Cunliffe, A. D., and Murray, J. F. (2013). The roles of melanin-concentrating hormone in energy balance and reproductive function: are they connected? *Reproduction* 146, R141–R150. doi: 10.1530/REP-12-0385

Nieuwenhuys, R., Geeraedts, L. M. G., and Veening, J. G. (1982). The medial forebrain bundle of the rat. I. General introduction. *J. Comp. Neurol.* 206, 49–81.

Nishihara, H., Maruyama, S., and Okada, N. (2009). Retroposon analysis and recent geological data suggest near-simultaneous divergence of the three superorders of mammals. *Proc. Natl. Acad. Sci. U.S.A.* 106, 5235–5240. doi: 10.1073/pnas.0809297106

Noble, E. E., Hahn, J. D., Konanur, V. R., Hsu, T. M., Page, S. J., Cortella, A. M., et al. (2018). Control of feeding behavior by cerebral ventricular volume transmission of melanin-concentrating hormone. *Cell Metab.* 28, 55.e7–68.e7. doi: 10.1016/j.cmet.2018.05.001

Nozaki, M., Tsukahara, T., and Kobayashi, H. (1983). An immunocytochemical study on the distribution of neuropeptides in the brain of certain species of fish. *Biomed. Res. Tokyo* 4, 135–143.

Pandolfi, M., Cánepa, M., Ravaglia, M., Maggese, M., Paz, D., and Vissio, P. (2003). Melanin-concentrating hormone system in the brain and skin of the cichlid fish *Cichlasoma dimerus*: anatomical localization, ontogeny and distribution in comparison to α-melanocyte-stimulating hormone-expressing cells. *Cell Tissue Res.* 311, 61–69.

Pérez-Sirkin, D., Cánepa, M., Fossati, M., Fernandino, J., Delgadin, T., Canosa, L., et al. (2012). Melanin concentrating hormone (MCH) is involved in the regulation of growth hormone in *Cichlasoma dimerus* (*Cichlidae. Teleostei*). *Gen. Comp. Endocrinol.* 176, 102–111. doi: 10.1016/j.ygcen.2012.01.002

Platt, R. N., Amman, B. R., Keith, M. S., Thompson, C. W., and Bradley, R. D. (2015). What is peromyscus? Evidence from nuclear and mitochondrial DNA sequences suggests the need for a new classification. *J. Mammal.* 96, 708–719.

Ravi, V., and Venkatesh, B. (2008). Rapidly evolving fish genomes and teleost diversity. *Curr. Opin. Gen. Dev.* 18, 544–550. doi: 10.1016/j.gde.2008.11.001

Renz, A. J., Meyer, A., and Kuraku, S. (2013). Revealing less derived nature of cartilaginous fish genomes with their evolutionary time scale inferred with nuclear genes. *PLoS One* 8:e66400. doi: 10.1371/journal.pone.0066400

Rodriguez, M., Beauverger, P., Naime, I., Rique, H., Ouvry, C., Souchaud, S., et al. (2001). Cloning and molecular characterization of the novel human melanin-concentrating hormone receptor MCH2. *Mol. Pharmacol.* 60, 632–639.

Röhlich, P., and Vigh, B. (1967). Electron microscopy of the paraventricular organ in the sparrow (*Passer domesticus*). *Z. Zellforsch. Mikrosk. Anat.* 80, 229–245.

Rondini, T. A., Donato, J., Rodrigues, B. C., Bittencourt, J. C., and Elias, C. F. (2010). Chemical identity and connections of medial preoptic area neurons expressing melanin-concentrating hormone during lactation. *J. Chem. Neuroana.* 39, 51–62. doi: 10.1016/j.jchemneu.2009.10.005

Sailer, A. W., Sano, H., Zeng, Z., Mcdonald, T. P., Pan, J., Pong, S.-S., et al. (2001). Identification and characterization of a second melanin-concentrating hormone receptor, MCH-2R. *Proc. Nat. Acad. Sci.* 98, 7564–7569.

Saito, Y., Nothacker, H.-P., Wang, Z., Lin, S. H. S., Leslie, F., and Civelli, O. (1999). Molecular characterization of the melanin-concentrating-hormone receptor. *Nature* 400, 265–269.

Shimomura, Y., Mori, M., Sugo, T., Ishibashi, Y., Abe, M., Kurokawa, T., et al. (1999). Isolation and identification of melanin-concentrating hormone as the endogenous ligand of the SLC-1 receptor. *Biochem. Biophys. Res. Commun.* 261, 622–626.

Sibert, E. C., and Norris, R. D. (2015). New age of fishes initiated by the cretaceous-paleogene mass extinction. *Proc. Natl. Acad. Sci. U.S.A.* 112, 8537–8542. doi: 10.1073/pnas.1504985112

Sita, L. V., Elias, C. F., and Bittencourt, J. C. (2003). Dopamine and melanin-concentrating hormone neurons are distinct populations in the rat rostromedial zona incerta. *Brain Res.* 970, 232–237.

Sita, L. V., Elias, C. F., and Bittencourt, J. C. (2007). Connectivity pattern suggests that incerto-hypothalamic area belongs to the medial hypothalamic system. *Neuroscience* 148, 949–969.

Smith, J. J., Kuraku, S., Holt, C., Sauka-Spengler, T., Jiang, N., Campbell, M. S., et al. (2013). Sequencing of the sea lamprey (Petromyzon marinus) genome provides insights into vertebrate evolution. *Nat. Genet.* 45, 415–421. doi: 10.1038/ng.2568

Sprague, J., Bayraktaroglu, L., Clements, D., Conlin, T., Fashena, D., Frazer, K., et al. (2006). The zebrafish information network: the zebrafish model organism database. *Nucleic Acids Res.* 34, D581–D585.

Suzuki, M., Bennett, P., Levy, A., and Baker, B. I. (1997). Expression of MCH and POMC genes in rainbow trout (*Oncorhynchus mykiss*) during ontogeny and in response to early physiological challenges. *Gen. Comp. Endocrinol.* 107, 341–350.

Suzuki, M., Narnaware, Y., Baker, B. I., and Levy, A. (1995). Influence of environmental colour and diurnal phase on MCH gene expression in the trout. *J. Neuroendocrinol.* 7, 319–328.

Swanson, L. W., Sanchez-Watts, G., and Watts, A. G. (2005). Comparison of melanin-concentrating hormone and hypocretin/orexin mRNA expression patterns in a new parceling scheme of the lateral hypothalamic zone. *Neurosci. Lett.* 387, 80–84.

Takahashi, A., Tsuchiya, K., Yamanome, T., Amano, M., Yasuda, A., Yamamori, K., et al. (2004). Possible involvement of melanin-concentrating hormone in food intake in a teleost fish, barfin flounder. *Peptides* 25, 1613–1622.

Takayama, Y., Wada, C., Kawauchi, H., and Ono, M. (1989). Structures of two genes coding for melanin-concentrating hormone of chum salmon. *Gene* 80, 65–73.

Tan, C. P., Sano, H., Iwaasa, H., Pan, J., Sailer, A. W., Hreniuk, D. L., et al. (2002). Melanin-concentrating hormone receptor subtypes 1 and 2: species-specific gene expression. *Genomics* 79, 785–792.

Tanaka, M., Azuma, M., Nejigaki, Y., Saito, Y., Mizusawa, K., Uchiyama, M., et al. (2009). Melanin-concentrating hormone reduces somatolactin release from cultured goldfish pituitary cells. *J. Endocrinol.* 203, 389–398. doi: 10.1677/JOE-09-0330

Tatemoto, K., and Mutt, V. (1981). Isolation and characterization of the intestinal peptide porcine PHI (PHI-27), a new member of the glucagon–secretin family. *Proc. Nat. Acad. Sci. U.S.A.* 78, 6603–6607.

Thannickal, T. C., Lai, Y.-Y., and Siegel, J. M. (2007). Hypocretin (orexin) cell loss in Parkinson's disease. *Brain* 130, 1586–1595.

Tillet, Y., Batailler, M., and Fellmann, D. (1996). Distribution of melanin-concentrating hormone (MCH)-like immunoreactivity in neurons of the diencephalon of sheep. *J. Chem. Neuroana.* 12, 135–145.

Tobet, S., Nozaki, M., Youson, J., and Sower, S. (1995). Distribution of lamprey gonadotropin-releasing hormone-III (GnRH-III) in brains of larval lampreys (*Petromyzon marinus*). *Cell Tissue Res.* 279, 261–270.

Torterolo, P., Sampogna, S., Morales, F. R., and Chase, M. H. (2006). MCH-containing neurons in the hypothalamus of the cat: searching for a role in the control of sleep and wakefulness. *Brain Res.* 1119, 101–114.

Toumaniantz, G., Bittencourt, J. C., and Nahon, J.-L. (1996). The rat melanin-concentrating hormone gene encodes an additional putative protein in a different reading frame. *Endocrinology* 137, 4518–4521.

Tuziak, S. M., and Volkoff, H. (2012). A preliminary investigation of the role of melanin-concentrating hormone (MCH) and its receptors in appetite regulation of winter flounder (Pseudopleuronectes americanus). *Mol. Cell. Endocrinol.* 348, 281–296. doi: 10.1016/j.mce.2011.09.015

Vallarino, M., Trabucchi, M., Chartrel, N., Jäggin, V., Eberle, A. N., and Vaudry, H. (1998). Melanin−concentrating hormone system in the brain of the lungfish *Protopterus annectens*. *J. Comp. Neurol.* 390, 41–51.

Vaughan, J. M., Fischer, W. H., Hoeger, C., Rivier, J., and Vale, W. (1989). Characterization of melanin-concentrating hormone from rat hypothalamus. *Endocrinology* 125, 1660–1665.

Venkatesh, B., Lee, A. P., Ravi, V., Maurya, A. K., Lian, M. M., Swann, J. B., et al. (2014). Elephant shark genome provides unique insights into gnathostome evolution. *Nature* 505, 174–179. doi: 10.1038/nature12826

Vidal, L., Blanchard, J. H., and Morin, L. P. (2005). Hypothalamic and zona incerta neurons expressing hypocretin, but not melanin concentrating hormone, project to the hamster intergeniculate leaflet. *Neuroscience* 134, 1081–1090.

Vigh-Teichmann, I., and Vigh, B. (1989). The cerebrospinal fluid-contacting neuron: a peculiar cell type of the central nervous system. Immunocytochemical aspects. *Arch. Histol. Cytol.* 52, 195–207.

Wang, S., Behan, J., O'neill, K., Weig, B., Fried, S., Laz, T., et al. (2001). Identification and pharmacological characterization of a novel human melanin-concentrating hormone receptor, mch-r2. *J. Biol. Chem.* 276, 34664–34670.

Wang, T., Yuan, D., Zhou, C., Lin, F., Wei, R., Chen, H., et al. (2016). Molecular characterization of melanin-concentrating hormone (MCH) in *Schizothorax prenanti*: cloning, tissue distribution and role in food intake regulation. *Fish Physiol. Biochem.* 42, 883–893. doi: 10.1007/s10695-015-0182-2

Xu, J., Hou, F., Wang, D., Li, J., and Yang, G. (2019). Characterization and expression of melanin-concentrating hormone (MCH) in common carp (*Cyprinus carpio*) during fasting and reproductive cycle. *Fish Physiol. Biochem.* 45, 805–817. doi: 10.1007/s10695-018-0586-x

Yun, S., Furlong, M., Sim, M., Cho, M., Park, S., Cho, E. B., et al. (2015). Prevertebrate local gene duplication facilitated expansion of the neuropeptide GPCR superfamily. *Mo. Biol. Evol.* 32, 2803–2817. doi: 10.1093/molbev/msv179

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Endocrinology

REVIEW
published: 17 April 2020
doi: 10.3389/fendo.2020.00225

# Comparative and Evolutionary Physiology of Vasopressin/Oxytocin-Type Neuropeptide Signaling in Invertebrates

*Esther A. Odekunle and Maurice R. Elphick**

*School of Biological & Chemical Sciences, Queen Mary University of London, London, United Kingdom*

The identification of structurally related hypothalamic hormones that regulate blood pressure and diuresis (vasopressin, VP; CYFQNCPRG-$NH_2$) or lactation and uterine contraction (oxytocin, OT; CYIQNCPLG-$NH_2$) was a major advance in neuroendocrinology, recognized in the award of the Nobel Prize for Chemistry in 1955. Furthermore, the discovery of central actions of VP and OT as regulators of reproductive and social behavior in humans and other mammals has broadened interest in these neuropeptides beyond physiology into psychology. VP/OT-type neuropeptides and their G-protein coupled receptors originated in a common ancestor of the Bilateria (Urbilateria), with invertebrates typically having a single VP/OT-type neuropeptide and cognate receptor. Gene/genome duplications followed by gene loss gave rise to variety in the number of VP/OT-type neuropeptides and receptors in different vertebrate lineages. Recent advances in comparative transcriptomics/genomics have enabled discovery of VP/OT-type neuropeptides in an ever-growing diversity of invertebrate taxa, providing new opportunities to gain insights into the evolution of VP/OT-type neuropeptide function in the Bilateria. Here we review the comparative physiology of VP/OT-type neuropeptides in invertebrates, with roles in regulation of reproduction, feeding, and water/salt homeostasis emerging as common themes. For example, we highlight recent reports of roles in regulation of oocyte maturation in the sea-squirt *Ciona intestinalis*, extraoral feeding behavior in the starfish *Asterias rubens* and energy status and dessication resistance in ants. Thus, VP/OT-type neuropeptides are pleiotropic regulators of physiological processes, with evolutionarily conserved roles that can be traced back to Urbilateria. To gain a deeper understanding of the evolution of VP/OT-type neuropeptide function it may be necessary to not only determine the actions of the peptides but also to characterize the transcriptomic/proteomic/metabolomic profiles of cells expressing VP/OT-type precursors and/or VP/OT-type receptors within the framework of anatomically and functionally identified neuronal networks. Furthermore,

OPEN ACCESS

**Edited by:**
*David Lovejoy,*
*University of Toronto, Canada*

**Reviewed by:**
*Tatsushi Onaka,*
*Jichi Medical University, Japan*
*Qingchun Tong,*
*University of Texas Health Science Center at Houston, United States*

***Correspondence:**
*Maurice R. Elphick*
*m.r.elphick@qmul.ac.uk*

**Specialty section:**
*This article was submitted to Neuroendocrine Science, a section of the journal Frontiers in Endocrinology*

**Received:** *18 December 2019*
**Accepted:** *30 March 2020*
**Published:** *17 April 2020*

**Citation:**
*Odekunle EA and Elphick MR (2020) Comparative and Evolutionary Physiology of Vasopressin/Oxytocin-Type Neuropeptide Signaling in Invertebrates. Front. Endocrinol. 11:225. doi: 10.3389/fendo.2020.00225*

investigation of VP/OT-type neuropeptide function in a wider range of invertebrate species is now needed if we are to determine how and when this ancient signaling system was recruited to regulate diverse physiological and behavioral processes in different branches of animal phylogeny and in contrasting environmental contexts.

**Keywords: vasopressin, oxytocin, neuropeptide, receptors, diuresis, reproduction, social behavior, feeding**

# INTRODUCTION

## The Discovery and Functional Characterization of Vasopressin/Oxytocin-Type Neuropeptide Signaling in Mammals

The importance of the structural characterization of the pituitary neurohormones vasopressin (VP) and oxytocin (OT) was recognized in the award of the 1955 Nobel Prize for chemistry to Vincent du Vigneaud. This was the culmination of a programme of research dating back to 1895, when Oliver and Schafer reported that a substance in extracts of the pituitary gland elevates blood pressure when injected intravenously into dogs (1). It was established later that this vasopressor originates from the posterior pituitary or neurohypophysis (2). Dale then reported that a neurohypophysial substance triggers uterine contraction (3, 4) and other bioactivities of pituitary extracts were discovered, including stimulation of lactation (5, 6) and antidiuresis (7). Purification of these bioactive components of pituitary gland extracts revealed that the vasopressor and antidiuretic activity could be attributed to one substance (VP) and the uterotonic and lactation-promoting activity could be attributed to another substance (OT) (8). In the 1950s the amino acid sequences and secondary structures of VP (9–11) and OT (12, 13) were determined and both peptides were chemically synthesized (14, 15). This revealed that VP (CYFQNCPRG-$NH_2$) and OT (CYIQNCPLG-$NH_2$) are structurally very similar, with only two amino acids differing, indicative of a common evolutionary origin. Furthermore, both peptides have a disulphide bridge between the cysteine residues at positions one and six, which is a conserved feature of all VP/OT-type peptides that have been identified subsequently (**Table 1**).

OT and VP are derived from precursor proteins that contain the cysteine-rich proteins neurophysin-I and neurophysin-II, respectively. Neurophysins bind OT or VP and are required for normal targeting of these neuropeptides to the regulated secretory pathway (16). Three G-protein coupled receptors (GPCRs) mediate the effects of VP (V1aR, V1bR, V2R) and a single GPCR mediates the effects of OT (OTR) (17). Consistent with the actions of OT as a regulator of lactation and uterine tone, OTR is expressed in the mammary glands and uterus, respectively. Consistent with the actions of VP as a vasopressor and antidiuretic, V1aR is expressed in vascular smooth muscle and V2R is expressed in the kidney. However, all four receptors are also expressed in other tissues/organs and perhaps most notably OTR, V1aR, and V1bR are widely expressed in the brain (18). Investigation of the functional significance of VP/OT-type receptor expression in the brain has revealed that VP/OT-type signaling regulates reproductive and social behavior in humans and other mammals (19), discoveries that have broadened interest in VP/OT-type neuropeptides beyond physiology into psychology.

Whilst VP and OT are now perhaps best known for their roles in regulation of reproductive and social behavior, other brain-associated actions in mammals have also been discovered; for example, intracerebroventricular injection of OT inhibits food and fluid intake in rats (20). Thus, as with other neuropeptides, both VP and OT are pleiotropic in their actions and our understanding of their physiological roles in mammals requires an integrative physiological and behavioral perspective, as discussed recently by Leng and Russell (21). Furthermore, to understand not only what VP and OT do in mammals but also why they do what they do, an evolutionary and comparative perspective is needed. Accordingly, there is a rich history of research on VP/OT-type signaling in non-mammalian vertebrates. This has been reviewed extensively elsewhere but a brief overview is presented below to serve as a prelude to the main theme of this review, which focuses on research investigating the physiological roles of VP/OT-type neuropeptides in invertebrates.

## Evolution and Comparative Physiology of VP/OT-Type Neuropeptide Signaling in Non-mammalian Vertebrates

Analysis of genome sequence data from non-mammalian vertebrates has enabled reconstruction of the evolutionary history of VP/OT-type signaling in the vertebrate lineage. The most primitive extant vertebrates are the jawless fish (lampreys, hagfish; Agnatha) and analysis of the genome sequence of the lamprey *Lethenteron japonicum* revealed that it has a single gene encoding a VP/OT-type neuropeptide ("vasotocin;" CYIQNCPRG-$NH_2$) (22) (**Table 1**). This contrasts with jawed vertebrates (gnathostomes) that typically have two genes encoding VP/OT-type neuropeptides—one that is an ortholog of the mammalian VP gene and another that is an ortholog of the mammalian OT gene. Thus, it has been inferred that the VP-type and OT-type genes originated by tandem duplication of a single VP/OT-type gene in a common ancestor of the gnathostomes (22). Furthermore, this was preceded in a common ancestor of the vertebrates by a gene duplication that gave rise to two genes encoding VP/OT-type receptors. Then two rounds of whole-genome duplication during early vertebrate evolution gave rise to eight genes encoding VP/OT-type receptors, with subsequent lineage-specific gene loss and additional gene/genome duplication events resulting in the

**TABLE 1 |** Amino acid sequences of vasopressin/oxytocin-type neuropeptides in species that belong to a variety of bilaterian phyla/sub-phyla.

| Phylum/ Sub-phylum | Species (selected peptide names) | Sequence |
|---|---|---|
| Vertebrata | *Homo sapiens* (vasopressin) | CYFQNCPRG-$NH_2$ |
| | *Homo sapiens* (oxytocin) | CYIQNCPLG-$NH_2$ |
| | *Lethenteron japonicum* (vasotocin) | CYIQNCPRG-$NH_2$ |
| Urochordata | *Styela plicata* | CYISDCPNSRFWST-$NH_2$ |
| | *Ciona intestinalis* | CFFRDCSNMDWYR |
| Cephalochordata | *Branchiostoma floridae* | CYIINCPRG-$NH_2$ |
| Hemichordata | *Saccoglossus kowalevskii* | CFISDCARG-$NH_2$ |
| Echinodermata | *Strongylocentrotus purpuratus* | CFISNCPKG-$NH_2$ |
| | *Apostichopus japonicus* | CFITNCPLG-$NH_2$ |
| | *Asterias rubens* | CLVQDCPEG-$NH_2$ |
| | *Ophionotus victoriae* | CLVSDCPEG-$NH_2$ |
| Xenacoelomorpha | *Xenoturbella bocki* | CLVQGCPIG-$NH_2$ |
| | *Xenoturbella profunda* | CLVQGCPIG-$NH_2$ |
| | *Ascoparia* sp | CVIVACPRG-$NH_2$ |
| Arthropoda | *Locusta migratoria* | CLITNCPRG-$NH_2$ |
| | *Tribolium casteneum* | CLITNCPRG-$NH_2$ |
| | *Atta cephalotes* | CLITNCPRG-$NH_2$ |
| | *Camponotus floridanus* | CLIVNCPRG-$NH_2$ |
| | *Harpegnathos saltator* | CLITNCPRG-$NH_2$ |
| | *Lasius niger* | CLITNCPRG-$NH_2$ |
| | *Lasius neglectus* | CLITNCPRG-$NH_2$ |
| | *Cancer borealis* | CFITNCPPG-$NH_2$ |
| | *Portunus pelagicus* | CFITNCPPG-$NH_2$ |
| | *Strigamia maritima* | CYITNCPPG-$NH_2$ |
| | *Sarcoptes scabiei* | CFITNCPPA-$NH_2$ |
| Nematoda | *Caenorhabditis elegans* | CFLNSCPYRRY-$NH_2$ |
| Tardigrada | *Hypsibius dujardini* | CFVTNCPPG-$NH_2$ |
| | *Ramazzottius variernatus* | CFVTNCPPG-$NH_2$ |
| Mollusca | *Conus striatus* | CIIRNCPRG-$NH_2$ |
| | *Conus geographus* | CFIRNCPKG-$NH_2$ |
| | *Lymnaea stagnalis* | CFIRNCPKG-$NH_2$ |
| | *Octopus vulgaris* (cephalotocin) | CYFRNCPIG-$NH_2$ |
| | *Octopus vulgaris* (octopressin) | CFWTSCPIG-$NH_2$ |
| | *Sepia officinalis* (sepiatocin) | CFWTTCPIG-$NH_2$ |
| | *Sepia officinalis* (pro-sepiatocin) | CFFRNCPPG-$NH_2$ |
| | *Mizuhopecten yessoensis* | CFIRNCPPG-$NH_2$ |
| Annelida | *Erpobdella octoculata* | CFIRNCPKG-$NH_2$ |
| | *Eisenia fetida* | CFVRNCPTG-$NH_2$ |
| | *Platynereis dumerilii* | CFVRNCPPG-$NH_2$ |

*Insights into the physiological roles of many of these neuropeptides have been obtained and are discussed in this review.*

variable numbers of VP/OT-type precursor genes and VP/OT-type receptor genes that are found in extant vertebrates (23, 24).

Prior to the genome-sequencing era, a variety of VP/OT-type neuropeptides were identified in non-mammalian vertebrates. With the benefit of hindsight, the nomenclature that was chosen for VP/OT-type neuropeptides in non-mammalian vertebrates is potentially confusing. For example, the name "vasotocin" was given to peptides that are orthologs of VP and the names "mesotocin" and "isotocin" were given to peptides that are orthologs of OT (25). Nevertheless, the discovery of these peptides in non-mammalian vertebrates was very important because it enabled analysis of their physiological roles. For example, in teleost fish vasotocin has been found to have VP-like roles in osmoregulation and cardiovascular physiology as well as OT-like roles in regulation of reproduction (26). Furthermore, central administration of isotocin in goldfish inhibits food intake (27), consistent with the anorexigenic effect of OT in mammals (20). However, a detailed review of the physiological roles of VP/OT-type neuropeptides in non-mammalian vertebrates is beyond the scope of this article, and for this topic we refer readers to other reviews (25, 28–30).

## Discovery of VP/OT-Type Neuropeptide Signaling in Invertebrates

Immunocytochemical evidence that VP-like neuropeptides occur in invertebrates was first reported in the late 1970s. Thus, two cells immunoreactive with antibodies to VP and to neurophysin II were identified in the suboesophageal ganglion of the locust *Locusta migratoria* (31). Subsequently, a VP-like peptide (CLITNCPRG-$NH_2$) was purified from extracts of *L. migratoria* suboesophageal ganglia and, interestingly, both a monomeric peptide (F1) and an anti-parallel dimer of the F1 peptide (F2) were identified (32) (**Table 1**). In parallel with research on insects, the existence of VP-like substances in molluscan species was also reported (33, 34). Then in 1987, VP-like peptides named Lys-conopressin G (CFIRNCPKG-$NH_2$) and Arg-conopressin S (CIIRNCPRG-$NH_2$) were purified from the venom of the cone snails *Conus geographus* and *Conus striatus,* respectively, and sequenced (35) (**Table 1**). Thus, the discovery of VP-like peptides in both insect and molluscan species in 1987 provided the first definitive molecular evidence of the occurrence of VP/OT-type peptides in invertebrates, demonstrating the evolutionary antiquity of this neuropeptide family. Accordingly, the presence of VP-like immunoreactivity in insects, molluscs and a variety of other invertebrates was reported the following year (36).

A VP-like peptide identical in structure to Lys-conopressin G was purified from extracts of the pond snail *Lymnaea stagnalis* (**Table 1**) and, importantly, cloning and sequencing of the gene encoding the precursor of this peptide revealed evolutionary conservation of protein structure. Thus, as in vertebrate VP/OT-type precursors, the neuropeptide is located immediately after an N-terminal signal peptide and the C-terminal region of the precursor comprises a neurophysin domain (37). Furthermore, a G-protein coupled receptor that shares sequence similarity with vertebrate VP/OT-type receptors and that mediates the effects of Lys-conopressin in *L. stagnalis* was identified, revealing evolutionary conservation of an ancient neuropeptide-receptor signaling pathway (38).

The first genome sequence of an animal species was reported in 1998 with the sequencing of the genome of the nematode *Caenorhabditis elegans* (39) and subsequently genes encoding the VP/OT-type neuropeptide nematocin and two cognate receptors were identified in this species (40–42). Likewise, as the genomes

and/or transcriptomes of many other invertebrate species have been sequenced over the last two decades, genes/transcripts encoding VP/OT-type precursors and receptors have been identified in an ever-increasing variety of invertebrate taxa, as discussed in more detail below. It is in this context that here we go on to discuss in detail what is known about the occurrence, characteristics and physiological roles of VP/OT-type signaling in invertebrate taxa. In doing so, we build upon, complement, extend and update several related review articles that have been published previously (30, 43–47).

## COMPARATIVE PHYSIOLOGY OF VP/OT-TYPE NEUROPEPTIDE SIGNALING IN INVERTEBRATES

As a framework for investigation of the comparative physiology of VP/OT-type neuropeptide signaling, it is necessary to briefly introduce bilaterian phylogeny. The Bilateria comprise two major clades—the protostomes and deuterostomes. The protostomes are further sub-divided into the Ecdysozoa (e.g., arthropods, nematodes) and the Spiralia (e.g., molluscs, annelids), whilst the deuterostomes comprise vertebrates, invertebrate chordates (urochordates, cephalochordates), hemichordates and echinoderms (48) (**Figure 1**). In discussing the comparative physiology of VP/OT-type neuropeptide signaling below, we have arbitrarily elected to start with the ecdysozoans because it was in an insect species, the locust *L. migratoria*, that the first invertebrate insights into VP/OT-type neuropeptide structure and function were obtained (32, 49). We then progress to the spiralian protostomes before moving on to the invertebrate deuterostomes, which are of particular interest because of their close relationship with vertebrates. We discuss the phylum Xenacoelomorpha last because the phylogenetic position of this phylum is controversial, as discussed in more detail below.

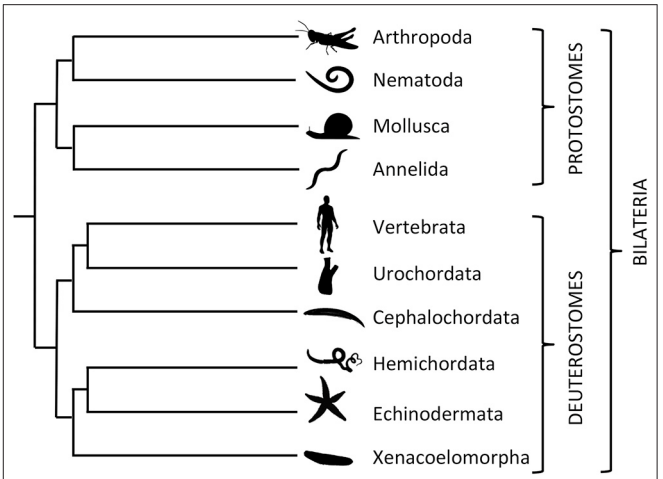


**FIGURE 1 |** Phylogenetic tree showing relationships of selected bilaterian phyla and sub-phyla discussed in this review. The Bilateria comprise two main clades: The protostomes, which include ecdysozoan phyla—Arthropoda (e.g., locust *Locusta migratoria* and fruit fly *Drosophila melanogaster*) and Nematoda (e.g., *Caenorhabditis elegans*)—and spiralian phyla—Mollusca (e.g., pond snail *Lymnaea stagnalis* and sea slug *Aplysia californica*) and Annelida (e.g., earthworm *Eisenia fetida* and ragworm *Platynereis dumerilii*). The deuterostomes, which include three chordate sub-phyla—Vertebrata (e.g., *Homo sapiens*), Urochordata (e.g., sea-squirt *Ciona intestinalis*) and Cephalochordata (e.g., lancelet *Branchiostoma floridae*)—and the xenambulacrarian phyla—Hemichordata (e.g., acorn worm *Saccoglossus kowalevskii*), Echinodermata (e.g., starfish *Asterias rubens*), and Xenacoelomorpha (e.g., *Xenoturbella bocki*). The branch lengths in the tree are arbitrary.

## Ecdysozoa

### Arthropoda

The purification and sequencing of a VP/OT-type neuropeptide from extracts of the suboesophageal ganglion of the locust *L. migratoria* provided definitive molecular evidence that this neuropeptide type exists in invertebrates (32). Subsequently, VP/OT-type neuropeptides have been identified in many other insects and the G-protein coupled receptors that mediate the effects of these peptides have been identified and pharmacologically characterized. Interestingly, however, VP/OT-type neuropeptide signaling has been lost in some insect lineages, including the model insect species *Drosophila melanogaster* (50).

Analysis of the distribution of VP-like immunoreactivity in the locust *L. migratoria* revealed a pair of immunostained neuronal cell bodies in the suboesophageal ganglion (31). Furthermore, use of radioimmunoassay methods revealed the presence of VP-like immunoreactivity in other regions of the locust body, including the brain, thoracic ganglia, abdominal ganglia, rectum, Malpighian tubules and hemolymph (51). Subsequently, a detailed analysis of the anatomy of the VP-like immunoreactive (VPLI) neurons of *L. migratoria*, combining use of immunohistochemistry with Lucifer Yellow or cobalt staining, revealed the processes of these neurons in the brain (optic lobe), suboesophageal ganglion, thoracic ganglia and abdominal ganglia (52). Furthermore, comparative analysis of 16 other grasshopper species revealed that the occurrence of a pair of VPLI neurons in the suboesophageal ganglion is a conserved feature of this family of insects, but with differences in the anatomy of their arborisations (53).

Insights into the potential physiological roles of the VP-type neuropeptide in locusts were first obtained with the discovery that injection of crude or purified extracts of suboesophageal ganglia causes a dose-dependent increase in the rate of dye excretion by Malpighian tubules *in vivo*. However, injection of mammalian VP had no effect on dye excretion in locusts and so it was unclear if the effect of ganglion extracts could be attributed to a VP-type peptide (54). With the isolation, sequencing and synthesis of the locust VP-type peptides F1 (monomeric peptide) and F2 (anti-parallel homodimer) (32, 49), it became possible to test the effects of these peptides on fluid secretion by Malpighian tubules *in vitro*. Interestingly, F1 was found to have no effect on urine production, whereas the anti-parallel dimer F2 appeared to have a stimulatory effect on diuresis (32). Subsequently, F1, F2 and a parallel dimer of F1 (PMd) were tested on *in vitro* preparations of Malpighian tubules from *L. migratoria* but none

of the peptides affected fluid secretion (55). Thus, the proposed physiological role of a VP-type peptide as a diuretic hormone in locusts remained unproven.

With respect to a potential role in regulation of diuresis, investigation of the electrophysiological properties of the pair of VPLI neurons that synthesize the VP-type neuropeptide in *L. migratoria* revealed that the activity of these neurons is not affected by the osmolality of perfusion salines. However, spiking activity of the neurons was found to be inversely related to light intensity, with the neurons receiving input from a pair of brain descending interneurons that form part of an extra-occular photoreceptor system (56). Further characterization of the VPLI neurons revealed that cholinergic inputs maintain the spiking activity of these neurons in the dark via both muscarinic and nicotinic acetylcholine receptors (57). Furthermore, stimulation of VPLI neuron activity was found to cause a reduction in cAMP levels in the locust brain. Biochemical analysis of isolated VPLI neurons revealed that they contain F1 and both the anti-parallel (F2) and parallel (PDm) dimers of F1, but only F1 was found to inhibit a forskolin-stimulated increase in cAMP in isolated locust neural membranes (58). In conclusion, over a period of almost two decades from the 1970s to the 1990s detailed insights into the biochemistry, anatomy, and physiology of VP-type peptides produced by the pair of VPLI neurons in the suboesophageal ganglion of the locust *L. migratoria* were obtained. More recently, a partial sequence of the gene encoding the VP-type peptide in *L. migratoria* has been reported (59). Further characterization of this gene would facilitate use of gene knockdown/knockout techniques to gain new insights into the physiological roles of VP-type signaling in locusts.

Sequencing of the *D. melanogaster* genome (60) revealed that VP/OT-type signaling has been lost in this species (50), which precluded its use as a model for functional studies on this neuropeptide signaling system. Likewise, genome sequencing revealed loss of VP/OT-type signaling in other insect species, including the mosquito *Anopheles gambiae* and the honey bee *Apis mellifera* (50). However, sequencing of the genome of the red flour beetle *Tribolium casteneum* revealed the existence of a gene encoding a VP/OT-type peptide precursor in an insect species (61–63). Furthermore, the *T. casteneum* VP/OT-type neuropeptide (**Table 1**) was named "inotocin" and its receptor was identified and characterized pharmacologically (50). Analysis of the expression of the genes encoding inotocin and its cognate receptor using qPCR revealed that in adult animals the expression level of both genes is highest in the head, but with receptor expression also detected in hind gut/Malpighian tubules. Interestingly, a developmental analysis revealed that expression of both genes is highest during the larval stage soon after hatching (50), which may indicate that VP/OT-type signaling is physiologically more important at this stage of development than in adults. Molecular and pharmacological characterization of the VP/OT-type signaling system in *T. castenuem* was also reported by Aikins et al. showing that the monomeric VP-type peptide (F1) is more potent as a ligand for the receptor than anti-parallel (F2) and parallel (D2) dimers of F1 (64). Consistent with previous findings from locusts, immunocytochemical analysis revealed expression of the VP-type peptide in a pair of neurons located in the suboesophageal ganglion of *T. castenuem,* with processes projecting anteriorly into the brain and posteriorly into the thoracic and abdominal ganglia of the ventral nerve cord. Furthermore, investigation of the physiological roles of VP/OT-type signaling in *T. castenuem* revealed that the monomeric F1 peptide stimulates diuresis in this species (64). However, neither the monomeric peptide (F1) nor dimers of F1 (F2, D2) stimulated fluid secretion when tested on *in vitro* preparations of Malpighian tubules from the another beetle species, *Tenebrio molitor*, consistent with previous findings from locusts (55). Interestingly, F1-stimulated fluid secretion was observed when Malpighian tubules were co-incubated with the central nervous system (including neurohormone secreting glands—the corpora cardiaca and corpora allata). Therefore, the authors concluded that the VP-type peptide stimulates diuresis in beetles not by acting directly on the Malpighian tubules but by stimulating the release of a diuretic substance from neural/neurohemal tissues (64), a mechanism of action that may also apply to other insects.

Most recently, advances in our knowledge of the physiological/behavioral roles of VP-type signaling in insects have been obtained through experimental studies on ants. Analysis of genome sequence data enabled identification of genes encoding a VP-type (inotocin) precursor in three ant species—*Atta cephalotes* (leaf-cutter ant), *Camponotus floridanus* (carpenter ant) and *Harpegnathos saltator* (Indian jumping ant) (65) (**Table 1**). Pharmacological characterization of a VP-type receptor in the black garden ant *Lasius niger* revealed that inotocin is a potent ligand for this receptor, with an $EC_{50}$ of 22 pM. Furthermore, it was discovered that an analog of *L. niger* inotocin (**Table 1**) containing a D-isomer of arginine at position 8 ([D-arg8]-inotocin) acts as a potent and selective antagonist of the human V1A receptor, providing a basis for translation of research on VP-type signaling in insects into potential clinical applications (66).

To gain insights into the physiological/behavioral roles of inotocin signaling in *L. niger*, inotocin receptor expression in the head of queen ants was measured. This revealed that expression levels were highest early in queen life when they experience crowded conditions in their mother nests before leaving to mate (67). Subsequently, a more detailed examination of the expression and physiological/behavioral roles of inotocin was reported, using the ant species *Lasius neglectus* as a model system (68) (**Table 1**). As reported previously in locusts (see above), inotocin expression was revealed in a pair of neurons in the suboesophageal ganglion. In addition, a detailed analysis of the expression levels of the inotocin precursor gene and the inotocin receptor gene in different developmental stages, castes and tissues/organs was reported. Furthermore, it was discovered that expression levels of both the inotocin precursor gene and the inotocin receptor gene are two-three fold higher in the summer than in the winter. Because feeding activity of ants is highest in the summer, the authors also examined gene expression with respect to feeding status and discovered that expression of the inotocin precursor gene is down regulated after a two-day starvation period. Knockdown of inotocin precursor gene expression using RNAi revealed changes in the expression of ∼100 other genes, many of which were found to be associated

with fat, protein and DNA metabolism. Furthermore, behavioral analysis of ants in which inotocin precursor gene expression had been knocked-down revealed increased locomotor activity and increased self-grooming in the brood chamber. Based on these findings, the authors concluded that inotocin signaling is important for regulation of energy status in association with locomotion and appetite in *Lasius neglectus* (68).

Complementary to Gruber and colleagues' detailed analysis of inotocin signaling in ant species of the genus *Lasius*, Koto et al. recently reported functional characterization of inotocin signaling in ant species of the genus *Camponotus* (69). The expression levels of the inotocin precursor gene and the inotocin receptor gene in different body parts were compared between mated queens, virgin queens, males, and workers with, for example, significantly higher levels of abdominal receptor gene expression detected in workers than in other castes. Furthermore, whole-body expression of both the precursor gene and the receptor gene were found to be upregulated as workers age and switch tasks from nursing to foraging. More specifically, inotocin precursor expression positively correlated with the overall level of activity, but not with time spent at food sources or in the nest. Consistent with previously reported findings from other insect species, immunohistochemical analysis revealed that inotocin is expressed in a pair of neurons in the suboesophageal ganglion. Furthermore, a cluster of inotocin-expressing neurons was also revealed in the protocerebrum of the brain. Analysis of inotocin receptor expression revealed high levels of transcript abundance in the fat body and use of mRNA *in situ* hybridization showed that these transcripts are specifically located in oenocytes, cells that are implicated in fatty acid and hydrocarbon metabolism. Experimental studies on *D. melanogaster* indicate that oenocytes produce cuticular hydrocarbons (CHCs) that are required for dessication resistance and pheromonal communication (70). Furthermore, the final step in CHC biosynthesis is catalyzed by the enzyme CYP4G1, which is expressed in *Camponotus* oenocytes. Therefore, inotocin signaling was investigated as a potential regulator of CYP4G1 and CHC biosynthesis in this species. Down-regulation of inotocin receptor expression was found to cause a reduction in *CYP4G1* expression and accordingly injection of inotocin receptor antagonists also caused a reduction in *CYP4G1* expression. Furthermore, injection of an inotocin receptor antagonist caused a reduction in hydrocarbon (alkane) synthesis and dessication resistance. Collectively, these findings led the authors to propose that when *Camponotus* worker ants start foraging, where there is an increased risk of dessication, inotocin signaling acts to stimulate synthesis of protective CHCs. In this context it is interesting that knockdown of inotocin expression in *L. neglectus* causes increased locomotor activity (68). Thus, inotocin signaling may act at a physiological level to stimulate synthesis of protective CHCs (*Camponotus*) and at a behavioral level to inhibit locomotor activity (*L. neglectus*) as integrated adaptive mechanisms to minimize dessication. These findings are of interest with respect to the anti-diuretic action of VP in mammals because they are suggestive of an evolutionarily ancient role of VP/OT-type signaling as a regulator of water homeostasis. What remains to be understood is the evolutionary/functional relationship between the water-preserving action of inotocin signaling in ants and the diuretic action of VP-type signaling in locusts and beetles (see above). To address this issue, experimental analysis of inotocin signaling in a wider range of insect species is now needed. Opportunities to do this have been afforded by sequencing of the transcriptomes/genomes of over 200 insect species and identification of genes encoding the inotocin precursor and/or inotocin receptor in these species. It should be noted, however, that VP/OT-type signaling has been lost in several insect lineages (71).

Currently, relatively little is known about VP/OT-type signaling in other arthropods. However, genes encoding VP/OT-type precursors and receptors have been identified in several non-insect arthropod taxa, including crustaceans, myriapods (millipedes and centipedes; e.g., *Strigamia maritima*) and chelicerates (scorpions, horseshoe crabs and mites; e.g., *Sarcoptes scabiei*, but with loss of this signaling system in spiders) (71) (**Table 1**). Furthermore, in several crustacean species the expression of these genes has been investigated in a functional context, as discussed below.

The stomatogastric nervous system, which controls feeding/digestion-associated processes in crustaceans, has been used as a model "simple" system in neurobiology to investigate how rhythmic motor-output is generated by a small neuronal network. Furthermore, neuropeptides have been identified that cause changes in motor output by altering synaptic connectivity within neural circuits (72). Expression of a VP/OT-type receptor has been detected in the stomatogastric ganglion of the crab *Cancer borealis* (73, 74) but the effects of the *C. borealis* VP/OT-type neuropeptide CFITNCPPG-$NH_2$ (**Table 1**) on the stomatogastric system have yet to be investigated.

A cDNA encoding the precursor of the VP/OT-type peptide CFITNCPPG-$NH_2$ in the blue swimming crab *Portunus pelagicus* (**Table 1**) has been sequenced and analysis of the expression of this precursor revealed transcripts in the eyestalk, brain, ventral nerve cord (VNC), intestine, gill and ovary. More specifically, transcripts were detected in oocytes and the VP/OT-type peptide in this species caused inhibition of steroid release from the ovary (75).

Expression of a VP/OT-type neuropeptide precursor in the shore crab *Carcinus maenas* has been examined in the context of ecdysis, which is regulated by release of ecdysteroid molting hormones from the endocrine Y-organs. Changes in Y-organ expression of the VP/OT-type neuropeptide precursor transcripts were detected over the molt cycle, but no changes in VP/OT-type receptor expression were observed (76).

Further studies are now needed to gain more detailed insights into the roles of VP/OT-type signaling in regulation of physiological processes in crabs and other crustaceans, with the potential for applicability in aquaculture of economically important species.

## Nematoda

Phylogenomic analysis of the sequences of VP/OT-type precursor proteins facilitated identification of a gene encoding the VP/OT-type peptide "nematocin" in the nematode *C. elegans* (40). Subsequently, mass spectrometric analysis of extracts of *C.*

*elegans* revealed that, by comparison with the majority of VP/OT-type neuropeptides in other taxa, nematocin has an unusual structure—CFLNSCPYRRY-$NH_2$ (**Table 1**). Thus, in common with other VP/OT-type peptides it has a disulphide bridge between the cysteines and a C-terminal amide group, but it is a C-terminally extended 11-residue peptide. There are two nematocin receptors in *C. elegans*, NTR-1 and NTR-2. Nematocin causes a dose-dependent increase in intracellular $Ca^{2+}$ and cAMP in cells transfected with NTR-1 but not in cells transfected with NTR-2. However, in cells transfected with both NTR-1 and NTR-2 nematocin causes a decrease in cAMP levels. Therefore, nematocin may activate alternative signaling pathways in *C. elegans* depending on whether NTR-1 is expressed alone or is co-expressed with NTR-2 (41, 42).

Analysis of the expression of nematocin in *C. elegans* revealed that in both sexes it is expressed in the AFD thermosensory neurons (which mediate thermotaxis), the DVA mechanosensory neuron (which regulates locomotion and posture), neurosecretory NSM cells, AVK interneurons and the pharyngeal neuron M5. Furthermore, in males nematocin is expressed in male-specific CP motor neurons that control turning behavior during mating. Analysis of the expression of the nematocin receptor genes revealed that in both sexes *ntr-1* and *ntr-2* are expressed in partially overlapping populations of head and tail neurons. For example, *ntr-1* is expressed in the left ASE (ASEL) gustatory neuron, the chemosensory neurons ASH and ADF and the PQR tail neuron. Furthermore, in males *ntr-1* and *ntr-2* are expressed in neurons and muscles that control and enable mating behavior. Thus, *ntr-1* is expressed in hook and tail sensory neurons that sense the vulva and hermaphrodite contact and in spicule protractor muscles that are involved in sperm transfer, whereas *ntr-2* is expressed in sensory-motor tail neurons that induce spicule penetration and muscle contraction for sperm transfer and in oblique muscles that promote prolonged vulval contact (41, 42). Informed by the patterns of expression of nematocin and its two receptors, the physiological roles of nematocin signaling in gustatory processes and male mating behavior were investigated.

With respect to gustation, *C. elegans* is normally attracted to low salt concentrations as an indicator of food availability. However, if wild-type worms experience low salt concentrations in the absence of food then they exhibit a change in behavior, showing reduced attraction to or even aversion to salt at low concentrations. Interestingly, this change in behavior was found to be impaired in worms with loss of function mutations in the genes encoding nematocin or NTR-1. Therefore, it was concluded that nematocin signaling is required for normal gustatory associative learning in *C. elegans* (41).

With respect to male mating behavior, if wild-type animals are placed in an arena with food and mating partners, they usually attempt to mate with the first hermaphrodite that they make contact with. First they make one or two turns around the body of the mating partner to locate the vulva and then sperm are usually transferred successfully within 5 min. In contrast, males with a mutated nematocin gene only attempted to mate after contact with several hermaphrodites and made more turns around the hermaphrodite before locating the vulva. In addition, the mutants were defective in turning behavior, maintenance of vulval contact and transfer of sperm. More specifically, deletion of nematocin expression in the DVA mechanosensory neuron was found to cause defects in the initial contact response and efficiency in locating the vulva, but turning behavior was not affected. Furthermore, generation of mutants lacking functional nematocin receptors revealed that NTR-1 and NTR-2 have overlapping roles in mediating the effects of nematocin as a regulator of mating behavior in *C. elegans*. Thus, nematocin signaling appears to increase the effectiveness of neural circuits in generating a sequence of behaviors that result in successful sperm transfer. Furthermore, it is proposed that nematocin signaling primes "neurons in a variety of local circuits to generate a neuroethological 'appetitive' function in mating" (42).

Male *C. elegans* also exhibit what is referred to as long-term mate search behavior, where animals will leave a bacterial food source if hermaphrodite mating partners are not also present. Interestingly, animals with mutations in genes encoding nematocin and its receptors exhibited deficiencies in this leaving behavior, providing further evidence of the importance of nematocin signaling for reproduction-associated behavior in *C. elegans* (42). The presence of larval progeny also increases the occurrence of adult *C. elegans* leaving a food source and this appears to be triggered by larval release of pheromones. Interestingly, worms with mutations in the *nematocin* or *ntr-1* genes exhibit reduced progeny-induced food-leaving behavior, indicating that nematocin is an important regulator of this behavior. This finding is interesting because it provides evidence of an evolutionarily ancient role of VP/OT-type neuropeptide signaling in parental-offspring social behavior (77).

Although research on VP/OT-type neuropeptide signaling in nematodes has primarily focused on *C. elegans* as a model system, it is of importance to also consider the occurrence of this signaling system in other nematodes and in particular parasitic species. Interestingly, phylogenetic analyses based on analysis of genome sequence data indicate that VP/OT-type signaling has been lost in *Brugia malayi*, which causes lymphatic filariasis in humans, *Ascaris suum*, an intestinal parasite in pigs, and *Trichinella spiralis*, which causes trichinosis in humans. It has been noted that these three species all lack a free-living larval stage outside the vector or host and therefore loss of VP/OT-type signaling may be reflective of this (45).

## Other Ecdysozoans

Analysis of genome sequence data has enabled identification of genes encoding VP/OT-type precursors in the tardigrades (water bears) *Hypsibius dujardini* and *Ramazzottius variernatus*. In both species the neuropeptide derived from the precursor protein has the predicted sequence CFVTNCPPG-$NH_2$ (**Table 1**). Furthermore, both species have two genes encoding VP/OT-type receptors (78). However, nothing is known about the physiological roles of VP/OT-type signaling in tardigrades. These animals are well-known for their remarkable ability to survive extreme environmental conditions, including desiccation, high and low temperature and pressure, and exposure to radiation (79). In the context of the recent report of a proposed physiological role of VP/OT-type signaling in dessication

resistance in ants (69), it would be interesting to determine if VP/OT-type signaling has a similar role in tardigrades. Another ecdysozoan species in which genes encoding VP/OT-type precursors/receptors have been identified is the penis worm *Priapulus caudatus* (phylum Priapulida) (47). However, as with tardigrades, nothing is known yet about the expression and function of these genes in this species.

## Spiralia

### Mollusca

Use of antibodies to VP/OT-type peptides provided early evidence that this neuropeptide family occurs in molluscs. For example, use of immunohistochemical methods enabled identification of neurons that contain VP/OT-like peptides in the pond snail *L. stagnalis* (34) and use of radioimmunoassays and high performance liquid chromatography enabled detection of a VP-like immunoreactive peptide in extracts of ganglia from the sea slug *Aplysia californica* (33) and other gastropod species (80). However, it was not until 1987 that the molecular structure of molluscan VP/OT-type peptides was determined with the purification and sequencing of Lys-conopressin G and Arg-conopressin S from venom of the cone snails *Conus geographus* and *Conus striatus,* respectively (35).

Molecular characterization of VP/OT-type neuropeptide signaling in *L. stagnalis* was accomplished by purification of the peptide CFIRNCPKG-$NH_2$ (**Table 1**) from neural ganglion extracts, cloning and sequencing of a cDNA and gene encoding this peptide and identification of the receptor that mediates its effects (37, 81, 82). Mapping of the expression of the VP/OT-type precursor in *L. stagnalis* revealed expression in neurons located in the anterior lobes of the cerebral ganglion in a position consistent with neurons that project into the penis nerve to innervate the penis complex and vas deferens. Accordingly, immunohistochemical analysis using antibodies to the VP/OT-type neuropeptide associated protein neurophysin revealed immunostained fibers in the penis nerve and vas deferens. Consistent with this pattern of expression, the *Lymnaea* VP/OT-type neuropeptide induces membrane depolarisation and rhythmic spiking of muscle cells isolated from the vas deferens (38) and triggers rhythmic contractions of *in vitro* preparations of the vas deferens (37). These findings indicate that VP/OT-type neuropeptide signaling has a physiological role in control of ejaculation in *L. stagnalis*. Interestingly, neurophysin-immunoreactivity was also observed in close proximity to the axons of neuroendocrine caudodorsal cells (CDC) in the cerebral ganglia, which release egg-laying peptides and control egg production and associated female reproductive behavior in *L. stagnalis*. Accordingly, *in vitro* electrophysiology experiments revealed that the *Lymnaea* VP/OT-type neuropeptide causes inhibition of CDCs by causing membrane hyperpolarisation and a reduction in spiking frequency (37). Collectively, these findings are of interest in the context of the simultaneous hermaphroditic mode of reproduction in *L. stagnalis* (83), indicating that VP/OT-type signaling in this species acts to promote male-type reproductive behavior whilst inhibiting female-type reproductive behavior. Furthermore, these findings provided some of the first evidence that VP/OT-type neuropeptides are evolutionarily ancient regulators of reproductive physiology and behavior in the Bilateria. However, VP/OT-type neuropeptide signaling is not solely involved in regulation of reproductive physiology and behavior in *L. stagnalis*. Thus, analysis of the expression of the VP/OT-type receptor in this species revealed that it is also present in the neuroendocrine light green cells and the *Lymnaea* VP/OT-type neuropeptide triggers depolarisation and spiking of these cells. Because the light green cells secrete insulin-related peptides that control somatic growth and metabolism in *L. stagnalis*, it was concluded that VP/OT-type signaling, in addition to reproductive functions, is also involved in regulation of metabolic processes in this gastropod mollusc (38).

The physiological roles of VP/OT-type neuropeptide signaling have also been investigated in detail in another gastropod—the sea slug *A. californica*. Immunohistochemical analysis of the central nervous system of this species revealed that VP-like immunoreactivity is restricted to a single neuron in the abdominal ganglion and two small neurons located bilaterally in each pedal ganglion (84). The abdominal ganglion contains neurons that control the gill-withdrawal reflex in *A. californica* and *in vitro* pharmacological studies revealed that VP/OT-type neuropeptides modulate the electrophysiological activity of identified neurons in the abdominal ganglia. For example, VP/OT-type neuropeptides decrease the spiking frequency of the gill motor neuron L7 and accordingly inhibit the gill-withdrawal reflex (33, 85–87). Furthermore, a more recent study has revealed that whilst the molluscan VP/OT-type neuropeptide conopressin-G inhibits gill motor neuron activity and gill withdrawal, it also increases the frequency of spontaneous gill movements. Interestingly, this combination of effects resembles activities associated with a food-aroused state in intact *A. californica* (85). Another reported effect of VP/OT-type neuropeptides on *A. californica* is to increase the spiking frequency of the R15 neuron (33). This neuron activates respiratory pumping and peristaltic movements of the large hermaphroditic duct during egg-laying behavior (88, 89) Thus, there is indirect evidence that VP/OT-type signaling may regulate reproductive processes in *A. californica*, although further studies are needed to specifically investigate this potential role.

VP/OT-type neuropeptide signaling has also been characterized in cephalopod molluscs, with identification of the peptide "cephalotocin" (CYFRNCPIG-$NH_2$) in *Octopus vulgaris* being an important first advance (90) (**Table 1**). Subsequently, a second VP/OT-type peptide ("octopressin"; CFWTSCPIG-$NH_2$) was identified in this species (91) (**Table 1**) and the genes encoding these peptides were sequenced (92). Interestingly both genes comprise single protein-coding exons, which contrasts with VP/OT-type genes in other taxa that comprise three protein-coding exons (92). The occurrence of two VP/OT-type genes/peptides in *O. vulgaris* and in other cephalopods (see below) is atypical of invertebrate species, which typically have a single VP/OT-type gene/peptide. Therefore, clade-specific gene duplication gave rise to the occurrence of two VP/OT-type genes in the cephalopod lineage, paralleling the evolution of VP and OT in the vertebrate lineage. Furthermore, two receptors that mediate effects of cephalotocin (CTR-1,

CTR-2) and one receptor that mediates the effects of octopressin (OPR) have been identified in *O. vulgaris* (93, 94).

Analysis of the expression of the octopressin gene in *O. vulgaris* revealed that it is expressed in many lobes of the supraoesophageal and suboesophageal brains and in the buccal and gastric ganglia (91). Accordingly, the octopressin receptor is expressed in the brain and central administration of octopressin has been reported to evoke hyperactivity of chromatophore cells, rapid respiration and jetting of water from the siphon similar to that seen in escape behavior (43). Additionally, expression of the octopressin receptor in the buccal and gastric ganglia is indicative of a physiological role in regulation of feeding and digestion (43). The octopressin receptor is also widely expressed peripherally, including in the rectum, oviduct and efferent branchial vessel (94) and accordingly *in vitro* pharmacological tests have revealed that octopressin induces tonic contraction of preparations of the rectum, oviduct and efferent branchial vessel and rhythmic contractions of the spermatophoric gland (91). Furthermore, investigation of potential osmoregulatory roles of VP/OT-type neuropeptides in the congeneric species *Octopus ocellatus* have revealed that octopressin, but not cephalotocin, decreases hemolymph osmolality and $Ca^{2+}$ ion concentrations, as well as urinary $Na^{+}$ ion concentrations (95).

Analysis of the expression of the cephalotocin gene in *O. vulgaris* revealed that it is largely associated with the ventral median vasomotor lobe of the suboesophageal brain, which contains neurons that are the source of the neurosecretory system of the vena cava (91). Comparison of the expression of the cephalotocin receptor genes in *O. vulgaris* revealed that CTR1 expression is predominantly associated with the central nervous system, whereas CTR2 expression is predominantly associated with peripheral organs, including the rectum, heart, vas deferens, oviduct and branchia (94). However, cephalotocin was found to have no effect when tested on *in vitro* preparations of the rectum, oviduct, efferent branchial vessel and the spermatophoric gland (91). Further studies are now needed to gain insights into the physiological roles of cephalotocin in *O. vulgaris* and the functional significance of the occurrence of the two cephalotocin receptors.

The occurrence of two genes encoding VP/OT-type peptides has also been reported in another cephalopod species, the cuttlefish *Sepia officinalis*. Thus, one gene encodes a peptide named sepiatocin (CFWTTCPIG-$NH_2$), which is most closely related to octopressin, and the second gene encodes a peptide named pro-sepiatocin (CFFRNCPPG-$NH_2$), which is most closely related to cephalotocin (96) (**Table 1**). Immunohistochemical methods have also been used to analyse the expression of VP/OT-type neuropeptides in *S. officinalis*, but interpretation of the findings are complicated by the use of antibodies to mammalian peptides (OT and VP) and the occurrence of two VP/OT-type neuropeptides in this species (97). A more specific analysis of the expression of the sepiatocin and pro-sepiatocin genes in *S. officinalis* revealed that sepiatocin has a widespread pattern of expression in the central nervous system, consistent with the expression pattern of octopressin in *O. vulgaris*. Conversely, pro-sepiatocin expression is restricted to the supraoesophageal and suboesophageal masses of the brain, consistent with the more restricted expression of cephalotocin in *O. vulgaris* (96).

Investigation of the pharmacological actions of sepiatocin in *S. officinalis* revealed that it causes tonic contraction of the oviduct, penis and vena cava (96), consistent with the previously reported myotropic effects of octopressin in *O. vulgaris* (91). Conversely, pro-sepiatocin lacked myotropic activity in *S. officinalis* (96), consistent with previous findings for cephalotocin in *O. vulgaris* (91). Pro-sepiatocin is, however, detected in the hemolymph of *S. officinalis*, indicating that it acts as a neurohormone, with a suggested role as a regulator of reproductive processes (96). Prior to the discovery of sepiatocin and pro-sepiatocin, the homologous peptides from *O. vulgaris* (octopressin and cephalotocin) were tested for effects on long-term memory (LTM) formation of a passive avoidance task in *S. officinalis*. Cephalotocin enhanced LTM at several doses tested whereas octopressin enhanced LTM at the lowest dose tested and attenuated LTM at the highest dose tested (98). These findings provided the first evidence that VP/OT-type neuropeptides are involved in learning and memory in an invertebrate species, but further studies are now needed to gain deeper insights into the underlying neural mechanisms.

In conclusion, insights into the physiological roles of VP/OT-type neuropeptide signaling in molluscs have thus far largely been obtained from experimental studies on the gastropods *L. stagnalis* and *A. californica* and the cephalopods *O. vulgaris* and *S. officinalis*. However, analysis of transcriptome/genome sequence data has also enabled identification of genes encoding VP/OT-type precursors and/or receptors in bivalve molluscs (99, 100). For example, the precursor of a VP/OT-type neuropeptide with the sequence CFIRNCPPG-$NH_2$, which is structurally very similar to pro-sepiatocin, has been identified in the scallop *Mizuhopecten yessoensis* (GenBank: OWF51696.1) (**Table 1**). In view of the economic importance of some bivalves as foodstuffs, investigation of the physiological roles of VP/OT-type neuropeptide signaling in these species could provide a basis for potential applications in aquaculture. With this objective in mind the physiological roles of other neuropeptides in bivalves have been investigated (101, 102), but to the best of our knowledge nothing is known about VP/OT-type neuropeptide function in bivalves.

## Annelida

The first paper to report the identification of a VP/OT-type neuropeptide in an annelid was published in 1993. A peptide with amino acid sequence CFIRNCPKG-$NH_2$ (**Table 1**) was purified from extracts of the central nervous system of the leech *Erpobdella octoculata* based on its immunoreactivity with antibodies to VP. Furthermore, injection of the synthetic peptide was found to cause a reduction in body mass in *E. octoculata* and it was concluded that this is due to stimulation of diuresis (103). Subsequently, the VP/OT-type peptide "annetocin" (CFVRNCPTG-$NH_2$) was isolated from the earthworm *Eisenia fetida* (**Table 1**); however, this peptide was not isolated using antibodies but on account of its myoexcitatory activity on *in vitro* preparations of the gut (crop-gizzard complex) and nephridia (excretory organs). Thus, annetocin

potentiates spontaneous rhythmic contractions of the isolated gut, potentiates pulsatile contractions of nephridia and induces pulsatile contractions in quiescent nephridia (104). Furthermore, investigation of the *in vivo* actions of annetocin in *E. fetida* revealed that it induces egg-laying related behaviors that include rotary movements, changes in body shape, mucous secretion from the clitellum and in some animals egg-laying (105). Consistent with these behavioral effects of annetocin, analysis of the expression of the annetocin precursor using mRNA *in situ* hybridization revealed that it is expressed by neurons located in the suboesophageal ganglion, which is known to be involved in regulation of reproductive behavior (106). Accordingly, immunohistochemical analysis of annetocin expression revealed a population of annetocin-immunoreactive neuronal somata in the suboesophageal ganglion and four immunoreactive neuronal somata in the cerebral ganglion. Analysis of the distribution of immunoreactive fibers revealed that they extend into the ventral nerve cord between the fourth and thirtieth segments, including the clitellum, but with a gradual reduction in the number of stained fibers proceeding posteriorly (107). Having identified and pharmacologically characterized the annetocin receptor, the expression of this receptor in *E. fetida* was investigated using mRNA *in situ* hybridization. Consistent with the effects of annetocin in inducing egg-laying related behaviors, annetocin receptor expression was found to be specifically associated with nephridia located in the clitellum region (108). Collectively, these findings indicate that a sub-population of suboesophageal neurons release annetocin from their processes in the ventral nerve cord and then annetocin binds to receptors on clitellum nephridia, which regulate production of a cocoon (108).

Interestingly, annetocin also induces egg-laying like behaviors in the leech *Whitmania pigra* (104). Accordingly, a subsequent study has shown that VP/OT-type neuropeptides induce a series of behaviors in the medicinal leech *Hirudo verbena* that closely resemble natural reproductive behavior, including twisting that aligns gonopores in preparation for copulation. Furthermore, the central pattern generator that controls this behavior was specifically localized to ganglia (M5 and M6) located in the reproductive segments of the leech (109). Annetocin was found not to induce egg-laying in the polychaete *Perinereis vancaurica* (105) but interestingly it has been reported that injection of a VP/OT-type neuropeptide in males of the polychaete *Nereis succinea* triggers swimming in tight circles and spawning (109). Analysis of the expression of the gene encoding the VP/OT-type neuropeptide CFVRNCPPG-$NH_2$ in the polychaete *Platynereis dumerilii* (**Table 1**) revealed that it is expressed in a pair of cells adjacent to the large cilia of deep brain extraocular photoreceptor cells. Furthermore, cells expressing a ciliary-type opsin were found to be located in the same position and connected to the same axons as the cells expressing the VP/OT-type precursor. Based on these findings it was concluded that the pair of cells expressing the VP/OT-type precursor in the brain of *P. dumerilii* are extraocular photoreceptor cells. Accordingly, it was concluded that these cells may coordinate reproductive behavior with seasonal changes in light cycles (110). More recently, two receptors that are activated by the VP/OT-type neuropeptide in *P. dumerilii* have been identified (111, 112) and it would be interesting to determine the spatial and temporal patterns of expression of these receptors to gain further insights into the physiological mechanisms of VP/OT-type signaling in this important new model system in neurobiology (113).

### Other Spiralians

A recent survey of the phylogenetic distribution of VP/OT-type signaling, based on an analysis of genome sequence data, reports that this signaling system is present in brachiopods (lamp shells) and gnathostomulids (jaw worms) but has been lost in rotifers (wheel animals) and platyhelminths (47). The loss of VP/OT-type signaling in the phylum Platyhelminthes is noteworthy given the biomedical importance of parasitic platyhelminths and extensive efforts to characterize other neuropeptide signaling systems in these animals (114–117).

## Invertebrate Deuterostomes

### Urochordata

VP/OT-type peptides that have been identified in urochordates have a characteristic feature of this neuropeptide family in having a conserved pair of cysteines that form a disulphide bridge in the mature peptide. However, unlike the majority of VP/OT-type neuropeptides that are amidated nonapeptides, the urochordate peptides are C-terminally extended. Thus, the VP/OT-type peptide in the sea-squirt *Styela plicata* is a C-terminally amidated peptide comprising 14 residues (118) and the VP/OT-type peptide in the sea-squirt *Ciona intestinalis* (CiVP) is 13 residue peptide without C-terminal amidation (119) (**Table 1**). Furthermore, the receptor that mediates effects of the VP/OT-type peptide in *C. intestinalis* has been identified (119).

Using mRNA *in situ* hybridization and immunohistochemical methods, respectively, expression of the gene encoding the VP/OT-type precursor and the mature peptide was revealed in the cerebral ganglion of *S. plicata*. Interestingly, VP/OT-type precursor gene expression in the cerebral ganglion is upregulated in animals maintained in diluted seawater (60%), by comparison with animals maintained in normal (100%) or concentrated (130%) seawater. Furthermore, animals maintained in diluted seawater closed their inhalent and exhalent siphons, whereas animals in normal or concentrated seawater kept their siphons open. Consistent with these behavioral observations, *in vitro* pharmacological tests revealed that the VP/OT-type peptide causes tonic contraction of siphon muscles. Collectively, these findings indicate that the VP/OT-type peptide in *S. plicata* has a VP-like role in osmoregulatory processes, regulating water intake when animals are exposed to hypotonic seawater (118).

Consistent with findings from *S. plicata*, the VP/OT-type neuropeptide gene in *C. intestinalis* is expressed in the cerebral ganglion. Furthermore, analysis of the *C. intestinalis* VP/OT-type receptor revealed that it is expressed in the cerebral ganglion and in peripheral organs, including the digestive system, endostyle, branchia sac, heart and gonad (119). Recently, detailed functional characterization of VP/OT-type signaling in *C. intestinalis* has revealed a role in physiological mechanisms of oocyte maturation via germinal vesicle breakdown. Expression of the VP/OT-type receptor gene *CiVpr* in ovarian follicles was found to increase prior to ovulation. Furthermore, exposure of developing

ovarian follicles to CiVP caused an increase in germinal vesicle breakdown and ovulation. *CiVpr* is expressed by oocytes and investigation of the mechanisms of CiVP action indicates that it causes CiVpr-mediated upregulation of the phosphorylation of extracellular signal-regulated kinase (CiErk1/2) and activation of a maturation-promoting factor, leading to oocyte maturation via germinal vesicle breakdown. Activated CiErk1/2 also induces oocyte expression of a matrix metalloproteinase (CiMMP2/9/13) and it is hypothesized that secretion of this enzyme causes digestion of collagens in the outer follicular cell layer, leading to the rupture of the follicular layer and ovulation (120).

## Cephalochordata

Analysis of the genome sequence of the cephalochordate *Branchiostoma floridae* has enabled identification of a gene encoding a 167-residue VP/OT-type precursor (Brafl1-84802) and the neuropeptide derived from this precursor has the predicted structure CYIINCPRG-$NH_2$ (22, 121) (**Table 1**). Two genes (154241, 134295) encoding candidate receptors for this neuropeptide have also been identified in *B. floridae* (122), but the pharmacological properties of these receptors have yet to be investigated experimentally. Furthermore, nothing is known about the expression of genes encoding the VP/OT-type precursor and receptors in *B. floridae*. This will be of interest to gain insights into the evolution of neuropeptide signaling in the phylum Chordata and comparison with VP/OT-type neuropeptide function in urochordates and vertebrates.

## Hemichordata

Analysis of the genome sequence of the hemichordate *Saccoglossus kowalevskii* (123) has enabled identification of a gene on genomic contig 42727 that encodes a VP/OT-type precursor and the neuropeptide derived from this precursor has the predicted structure CFISDCARG-$NH_2$ (121). An unusual feature of this peptide is the presence of an alanine residue at position seven, which is more typically occupied by a proline residue in members of this neuropeptide family (**Table 1**). Furthermore, a gene (g16853) encoding a candidate receptor for this neuropeptide has also been identified in *S. kowalevskii* (122), but the pharmacological properties of this receptor have yet to be investigated experimentally. Likewise, nothing is known about the physiological roles of VP/OT-type signaling in hemichordates, which is reflective of a general paucity of information on neuropeptide expression and function in this phylum. This will surely be a fruitful research theme for future work, both in gaining understanding of the evolution of neuropeptide function in the ambulacrarian clade of the animal kingdom and in providing insights into the functional neuroarchitecture of hemichordate nervous systems (124, 125).

## Echinodermata

Sequencing of the genome of the sea urchin *Strongylocentrotus purpuratus* enabled the first identification of genes encoding a VP/OT-type precursor and a VP/OT-type receptor in an echinoderm (126, 127). The *S. purpuratus* VP/OT-type precursor comprises a neuropeptide (echinotocin) with the predicted structure CFISNCPKG-$NH_2$ (**Table 1**). Echinotocin has been synthesized and tested on *in vitro* preparations of the esophagus and tube feet from the sea urchin *Echinus esculentus* and found to cause dose-dependent contraction of these neuromuscular organs (40). This effect of echinotocin in sea urchins is consistent with the myotropic actions of VP/OT-type neuropeptides in vertebrates and protostomian invertebrates (see above). However, the physiological/behavioral significance of the *in vitro* myotropic actions of echinotocin in adult sea urchins is not known. Expression of the echinotocin precursor during embryonic and larval development in *S. purpuratus* has also been investigated and interestingly peak expression (∼200 transcripts per embryo) is observed at 48 h post fertilization (hpf), which corresponds to the gastrula stage when the first neuronal precursor cells are detected. However, by the early larval stage (70 hpf), when a simple nervous system is present, the level of echinotocin precursor expression is lower (∼60 transcripts per larva) (128). Further studies are now needed to investigate the roles of echinotocin signaling during sea urchin development and this will be facilitated by use of gene-knockout techniques (129).

Sequencing of the neural transcriptome of the starfish *A. rubens* enabled identification of a transcript encoding a VP/OT-type precursor protein comprising the neuropeptide "asterotocin" (130) and the structure of this peptide (CLVQDCPEG-$NH_2$) has been confirmed using mass spectrometry (131). An unusual characteristic of asterotocin is that it has acidic residues at positions 5 and 8, which are usually occupied by a basic or hydrophobic residue in other VP/OT-type peptides (**Table 1**). The asterotocin receptor has also been identified in *A. rubens* and shown to be selectively activated by asterotocin and not by mammalian VP or OT (131). To gain insights into the physiological roles of asterotocin signaling in starfish, the expression of asterotocin and the asterotocin receptor in *A. rubens* has been examined using both mRNA *in situ* hybridization and immunohistochemical techniques. Asterotocin-expressing cells and fibers are present in the central nervous system (radial nerve cords and circumoral nerve ring), the locomotory organs (tube feet), several regions of the digestive system (including the cardiac stomach) and in the body wall and associated appendages (e.g., papulae that mediate gas exchange). Asterotocin receptor-expressing cells were found to be less abundant than asterotocin-expressing cells but double labeling of asterotocin and the asterotocin receptor revealed complementary patterns of expression. Investigation of the *in vitro* pharmacological actions of asterotocin revealed that it acts as a muscle relaxant in *A. rubens*, contrasting with myotropic actions of VP/OT-type neuropeptides that have been reported in vertebrates and in other invertebrates. For example, asterotocin was found to be a potent relaxant of cardiac stomach preparations from *A. rubens*. This effect of asterotocin *in vitro* was of interest because relaxation of the cardiac stomach occurs when this organ is everted out of the mouth over the soft tissues of prey when starfish feed. Therefore, the *in vivo* effects of asterotocin were also investigated and this revealed that asterotocin is a powerful inducer of cardiac stomach eversion. Furthermore, injection of asterotocin also caused arm flexion and adoption of a body posture similar to that adopted when starfish feed on prey such as mussels. Moreover, the effect of

asterotocin on body posture was so powerful that it adversely affected the ability of starfish to right themselves when upturned. Collectively, the findings of this study indicate that asterotocin signaling has an important physiological/behavioral role in triggering the unusual extra-oral feeding behavior of the starfish *A. rubens* (131). Expression of the asterotocin precursor has also been investigated in the free-swimming larvae of *A. rubens*, with transcripts detected during the brachiolaria larval stage in cells located at the tips of the brachia and associated with the adhesive disk—structures that mediate larval attachment to the substratum prior to metamorphosis (132). Therefore, it would be interesting to investigate a potential physiological role of asterotocin signaling as a regulator of larval attachment in starfish. It will also be of interest to investigate the physiological roles of asterotocin in other starfish species. Relevant to this, a gene encoding an asterotocin precursor has been identified in the crown-of-thorns starfish (COTS) *Acanthaster planci* (133) but further studies are now needed to gain insights into the physiological roles of asterotocin in this species.

Transcripts encoding VP/OT-type precursors have also been identified in the brittle star species *Ophionotus victoriae, Amphiura filiformis, and Ophiopsila aranea* and interestingly, in common with starfish, the predicted neuropeptide derived from these precursors has an acidic residue (glutamate) in the eighth position (134). For example, the VP/OT-type peptide in *O. victoriae* has the predicted structure CLVSDCPEG-$NH_2$, which is identical to the *A. rubens* peptide asterotocin in all but the fourth residue (**Table 1**). It is noteworthy, therefore, that the unusual property of asterotocin in having acidic residues at both positions five and eight is also a feature of VP/OT-type neuropeptides in brittle stars. It is not, however, a feature of the VP/OT-type neuropeptides in other echinoderm classes (see below) and therefore it can be deduced that this structural feature evolved in a common ancestor of the asterozoan clade of extant echinoderms (135), which comprises the Asteroidea (starfish) and the Ophiuroidea (brittle stars).

The other classes of echinoderms that we have yet to consider are the Holothuroidea (sea cucumbers), which are a sister class to the Echinoidea (sea urchins) in the Echinozoan clade of extant echinoderms, and the Crinoidea (e.g., featherstars) that occupy a basal position with respect to the other extant echinoderm classes (135). VP/OT-type precursor transcript sequences have been identified in several sea cucumber species, including *Apostichopus japonicus*, *Holothuria glaberrima,* and *Holothuria scabra* (136, 137). The neuropeptides derived from these precursors (e.g., CFITNCPLG-$NH_2$ in *A. japonicus*; "holotocin") share more sequence similarity with the sea urchin peptide echinotocin (CFISNCPKG-$NH_2$) than with asterozoan VP/OT-type neuropeptides (e.g., asterotocin) (**Table 1**), consistent with the sister group status of sea cucumbers (Holothuroidea) and sea urchins (Echinoidea) in the echinozoan clade of the phylum Echinodermata. Currently, nothing is known about the physiological roles of VP/OT-type neuropeptides in sea cucumbers and so this is an important objective for future studies, particularly in the context of potential applications in aquaculture of economically valuable edible species such as *A. stichopus* and *H. scabra* (136, 137).

### Xenacoelomorpha

The phylum Xenacoelomorpha comprises three sub-phyla—xenoturbellids (e.g., *Xenoturbella bocki),* nemertodermatids (e.g., *Nemertoderma westbladi*), and acoels (e.g., *Symsagittifera roscoffensis)*—all of which have a simple body plan without a through-gut (138–140). However, there is controversy regarding the phylogenetic position of the Xenacoelomorpha in the animal kingdom. The Nephrozoa hypothesis places xenacoelomorphs as a sister group to all other bilaterian animals (141, 142) whereas the Xenambulacraria hypothesis places xenacoelomorphs within the deuterostome clade of the Bilateria as a sister group to the Ambulacraria (Hemichordates and Echinoderms) (143). In discussing this phylum last, but after the Ambulacraria, both phylogenetic positions are compatible with the structure of this review article. Nevertheless, we note that the most recent phylogenetic analysis is supportive of the Xenambulacraria hypothesis (144) and therefore in **Figure 1** the phylum Xenacoelomorpha is positioned as a sister group to the Ambulacraria.

A detailed survey of the occurrence of neuropeptide signaling systems in xenacoelomorphs has been reported recently, based on analysis of genome/transcriptome sequence data from species belonging to the three sub-phyla of the phylum (145). Genes encoding VP/OT-type precursors and receptors were identified in xenoturbellids (*X. bocki* and *Xenoturbella profunda*) and a nemertodermatid (*Ascoparia* sp.) but not in acoels. Therefore, based on the data currently available it appears that VP/OT-type signaling may have been lost in acoels. Analysis of the occurrence of other neuropeptide signaling systems in xenacoelomorphs has also revealed evidence of loss in acoels and so in this respect the absence of VP/OT-type signaling in acoels may be reflective of a more widespread simplification of neuropeptide signaling in this sub-phylum. Nevertheless, analysis of complete genome sequences in a variety of acoel species may be necessary to draw definitive conclusions regarding the loss of VP/OT-type signaling.

The sequences of the predicted neuropeptides derived from VP/OT precursors in xenoturbellids and a nemertodermatid are CLVQGCPIG-$NH_2$ (in *X. bocki* and *X. profunda*) and CVIVACPRG-$NH_2$ (in *Ascoparia* sp.), respectively (145) (**Table 1**). However, nothing is known about the expression and actions of these neuropeptides and therefore this is an important and fascinating objective for future studies.

## GENERAL CONCLUSIONS AND DIRECTIONS FOR FUTURE RESEARCH

The evolutionary origin of VP/OT-type neuropeptide signaling in a common ancestor of the Bilateria was first determined in 1987 with the sequencing of VP/OT-type peptides in protostome invertebrate species (32, 35). This finding has been consolidated over the last two decades by genome/transcriptome sequencing and the discovery of genes/transcripts encoding VP/OT-type precursor proteins and VP/OT-type receptors in a wide range of bilaterian phyla (47, 122, 146). Invertebrates typically have one gene encoding a VP/OT-type precursor and one gene encoding

a VP/OT-type receptor and it can be inferred, therefore, that this reflects the ancestral condition in Urbilateria. However, gene duplication has subsequently given rise to duplicated precursor genes and/or receptor genes in several lineages, including, for example, cephalopod molluscs, nematodes, tardigrades and vertebrates (23, 24, 41–43, 78). Furthermore, genome duplication in the vertebrate lineage has given rise to expanded families of genes encoding VP/OT-type receptors (23, 24). Conversely, in some lineages the VP/OT-type signaling system has been lost. This includes apparent instances of loss in an entire phylum (e.g., Platyhelminthes, Rotifera), although definitive proof may have to await the availability of genome sequences from extant species that are representative of all the major taxonomic branches of each phylum. Accordingly, in some phyla there are instances of clade specific loss of VP/OT-type signaling; the arthropods are example of this with loss of VP/OT-signaling in several insect orders and spiders (47). Loss of VP/OT-type signaling within specific clades of a phylum is interesting because it may ultimately provide insights into the functional significance and physiological importance of VP/OT-type signaling. Thus, if we know the physiological roles of VP/OT-type signaling in one or more clades of a phylum, we can then consider how those physiological roles of VP/OT-type signaling could be dispensed with as a consequence of gene loss in other clades. With this perspective in mind, we move on now to consider what general principles emerge from our comparative survey of what is known about VP/OT-type neuropeptide function in invertebrates.

Evidence that VP/OT-type neuropeptides regulate reproductive processes and behaviors is a prevalent finding, consistent with the roles of OT, and to a lesser extent VP, in vertebrates. Perhaps most notable are the actions of annetocin in inducing egg-laying related behaviors in annelids (105, 109), evidence that nematocin signaling is required for normal male mating behavior in the nematode *C. elegans* (42) and the recent discovery that VP/OT-type signaling regulates oocyte maturation and ovulation in the urochordate *C. intestinalis* (120). Furthermore, the discovery that VP/OT-type signaling acts differentially to control male and female reproductive systems in the mollusc *L. stagnalis* illustrates how an ancient signaling system has been adapted to regulate reproduction in the context of hermaphroditism (37, 38). In the context of these findings from a variety of phyla, it is reasonable to infer that the role of VP/OT-type signaling as a regulator of reproductive processes originated in Urbilateria. Accordingly, does VP/OT signaling regulate of reproductive processes in all extant Bilateria? In the many phyla where nothing is known about VP/OT-type neuropeptide function, this question remains to be addressed. It is noteworthy that thus far no evidence of a reproductive role has been reported in insects, although it has been reported that a VP/OT-type neuropeptide inhibits steroid release from the ovary in the blue swimming crab *P. pelagicus* (75). Therefore, further investigation of potential reproductive roles of VP/OT-type signaling in arthropods would be interesting. Furthermore, our recent report that the VP/OT-type neuropeptide asterotocin triggers fictive feeding in the starfish *A. rubens* (Phylum Echinodermata) (131) may also be relevant to reproductive behavior. Spawning in starfish is triggered by a relaxin-type neuropeptide and when starfish spawn they adopt a "humped" posture similar to that adopted when they feed on prey (147–150). Therefore, whilst there is no evidence that asterotocin can trigger the release of gametes in starfish, it is possible that asterotocin signaling is involved in neural control of postural changes associated with spawning in starfish. In this context, it would be interesting to investigate if VP/OT-type neuropeptides have effects on behavior associated with reproduction in other echinoderms (e.g., sea urchins, sea cucumbers).

A second general theme that emerges from comparative analysis of VP/OT-type neuropeptide function in invertebrates are roles associated with water/salt homeostasis and excretory organs, consistent with the actions of VP as an antidiuretic in vertebrates. This includes evidence from the urochordate *S. plicata*, where VP/OT-type signaling appears to be involved in preventing intake of dilute seawater by causing closure of inhalent/exhalent siphons (118). Furthermore, it is noteworthy that annetocin has myoexcitatory effects on excretory organs (nephridia) in the earthworm *E. foetida*, which is important in a reproductive context where specialized nephridia in the clittelar region secrete a cocoon during egg-laying (104). However, this may be reflective of a more general role of VP/OT-type signaling in regulation of nephridial function in annelids and therefore it would be interesting to investigate the actions of VP/OT-type neuropeptides on species from other annelid clades [e.g., in the ragworm *P. dumerilii*; (151)]. Investigation of VP/OT-type neuropeptide function in insects has revealed what appear to conflicting actions with respect to water preservation. Thus, experimental studies on locusts and beetles have revealed a diuretic effect (i.e., water loss) (54, 64) whilst more recent studies on ants have provided evidence that the VP/OT-type neuropeptide inotocin promotes dessication resistance (i.e., water preservation) by stimulating production of cuticular hydrocarbons by oenocytes (69). In this context, it would be timely to revisit investigation VP/OT-type neuropeptide function in the locust *L. migratoria,* which as a large insect is an attractive experimental model for physiological studies. Thus, building upon previous anatomical and physiological studies (52, 56), it would be interesting to determine where the inotocin receptor is expressed in locusts as this would provide a basis for an integrative approach to analysis of the physiological roles of VP/OT-type neuropeptide signaling in this species. Furthermore, one of the most intriguing characteristics of VP/OT-type neuropeptide signaling in insects is the highly conserved expression of inotocin in a pair of neurons located in the suboesophageal ganglion, although in some species additional inotocin-expressing neurons are present in the brain (53, 64, 68, 69). This highly restricted pattern of neuronal expression may, at least in part, explain why loss of VP/OT-type neuropeptide signaling has occurred in some insect lineages (including *Drosophila*). If further advances are made in our understanding of the physiological roles of VP/OT-type signaling in insects that have retained this system, it may be possible to provide evolutionary, ecological and physiological explanations for its loss in some insect taxa.

A third emergent theme is an association of VP/OT-type neuropeptide signaling with extra-ocular photoreceptive systems and therefore by inference a potential role in circadian and/or circalunar physiological/behavioral rhythmicity. Evidence of this from invertebrates was first obtained from experimental characterization of neural inputs to suboesophageal VPLI neurons in the locust *L. migratoria* (56, 57). However, evidence of a phylogenetically more widespread association with extra-ocular photoreceptors was obtained subsequently from the annelid *P. dumerilii*, where it was concluded that suboesophageal neurons expressing the VP/OT-type neuropeptide in this species are in fact extraocular receptors themselves (110). The association of VP/OT-type neuropeptide expressing neurons with extraocular photoreception in these two invertebrate species is consistent with findings from mammals, where elevated pituitary release of OT and VP occurs during the hours of sleep (152). However, further experimental studies are now needed to investigate in more detail a proposed association of VP/OT-type signaling with circadian and/or circalunar physiological/behavioral rhythmicity in insects, annelids and/or other invertebrates.

A fourth theme that emerges from comparative studies is the role of VP/OT-type signaling in regulation of feeding, with the inhibitory effect of OT-type neuropeptides on food intake in mammals and other vertebrates (20, 27) providing a basis for investigation of feeding-related roles in invertebrates. Our discovery that asterotocin triggers fictive feeding in the starfish *A. rubens* is perhaps the most striking evidence obtained from invertebrates, although it is of course noteworthy that actions of asterotocin are indicative of an orexigenic action that contrasts with the anorexigenic action of oxytocin in mammals. Indirect evidence of VP/OT-type signaling regulating feeding and/or digestion includes the expression of VP/OT-type receptors in the digestive system and regions of nervous system that control feeding and/or digestive processes in molluscs (91), annelids (104), nematodes (41) and crustaceans (73, 74) and changes in inotocin precursor expression following a starvation period in ants (68). However, further studies are now needed to investigate in more detail the roles of VP/OT-type neuropeptides in feeding/digestion in these and other invertebrate taxa.

In summary, it is clear that in the context of comparative analysis of neuropeptide function in the Bilateria, research on VP/OT-type neuropeptide signaling has been remarkably fruitful. This may in part reflect the high level of evolutionary conservation of VP/OT-type neuropeptide/precursor structure in the Bilateria and the ease with which VP/OT-type neuropeptides can be identified in different taxa. However, it is also noteworthy that VP/OT-type neuropeptides have quite striking effects of reproduction and/or feeding related behaviors in invertebrates, which has facilitated and encouraged detailed experimental analysis of VP/OT-type signaling in a wide range of taxa. But there is still much to be learnt about the physiological roles of VP/OT-type neuropeptide signaling invertebrates if we are to achieve a level of understanding that would facilitate reconstruction of the evolution of VP/OT-type neuropeptide function in the Bilateria. Neuropeptide signaling systems operate in the context of neuronal circuits and adaptive evolutionary changes in the configuration of those circuits (153). To understand the evolution of VP/OT-type neuropeptide function it may therefore be necessary to not only determine the actions of the peptides but also to characterize the transcriptomic/proteomic/metabolomic profiles of cells expressing VP/OT-type precursors and/or VP/OT-type receptors within the framework of anatomically and functionally identified neuronal connectivities. This will not be feasible in all taxa but this should not deter comparative physiologists from taxa that were intractable prior to the genome/transcriptome-sequencing era. Investigation of the effects VP/OT-type neuropeptides in the huge variety of invertebrates that remain to be studied is sure to reveal actions that are both amazing and informative. So let this be a mission that could be celebrated in 2055, the 100th anniversary of the Nobel Prize award for chemistry to Vincent du Vigneaud.

## AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

## FUNDING

EO was supported by a Ph.D. studentship awarded by the Society for Experimental Biology. ME was supported by grants from the Biotechnology and Biological Sciences Research Council (BB/M001644/1) and the Leverhulme Trust (RPG-2018-200).

## REFERENCES

1. Oliver G, Schafer EA. On the physiological action of extracts of pituitary body and certain other glandular organs: preliminary communication. *J Physiol*. (1895) 18:277–9. doi: 10.1113/jphysiol.1895.sp000565
2. Howell WH. The physiological effects of extracts of the hypophysis cerebri and infundibular body. *J Exp Med*. (1898) 3:245–58. doi: 10.1084/jem.3.2.245
3. Dale HH. On some physiological actions of ergot. *J Physiol*. (1906) 34:163–206. doi: 10.1113/jphysiol.1906.sp001148
4. Dale HH. The action of extracts of the pituitary body. *Biochem J*. (1909) 4:427–47. doi: 10.1042/bj0040427
5. Ott I, Scott JC. The action of infundibulum upon mammary secretion. *Proc Soc Exp Biol*. (1910) 8:48–9. doi: 10.3181/00379727-8-27
6. Schafer EA, Mackenzie K. The action of animal extracts on milk secretion. *Proc R Soc Lond B*. (1911) 84:16–22. doi: 10.1098/rspb.1911.0042
7. Farini F. Diabete insipido ed opoterapia. *Gazz Osped Clin*. (1913) 34:1135–9.
8. Kamm O, Aldrich TB, Grote IW, Rowe LW, Bugbee EP. The active principles of the posterior lobe of the pituitary gland. I. the demonstration of the presence of two active principles. II. the separation of the two principles and their concentration in the form of potent solid preparations. *J Am Chem Soc*. (1928) 50:573–601. doi: 10.1021/ja01389a050
9. Turner RA, Pierce JG, Du Vigneaud V. The purification and the amino acid content of vasopressin preparations. *J Biol Chem*. (1951) 191:21–8.
10. Du Vigneaud V, Lawler HC, Popenoe EA. Enzymatic cleavage of glycinamide from vasopressin and a proposed structure for this pressor-antidiuretic hormone of the posterior pituitary. *J Am Chem Soc*. (1953) 75:4880–1. doi: 10.1021/ja01115a554

11. Acher R, Chauvet J. La structure de la vasopressine de boeuf. *Biochim Biophys Acta*. (1954) 14:421–9. doi: 10.1016/0006-3002(54)90202-4
12. Tuppy H. The amino-acid sequence in oxytocin. *Biochim Biophys Acta*. (1953) 11:449–50. doi: 10.1016/0006-3002(53)90071-7
13. Du Vigneaud V, Ressler C, Trippett S. The sequence of amino acids in oxytocin, with a proposal for the structure of oxytocin. *J Biol Chem*. (1953) 205:949–57.
14. Du Vigneaud V, Gish DT, Katsoyannis PG. A synthetic preparation possessing biological properties associated with arginine vasopressin. *J Am Chem Soc*. (1954) 76:4751–2. doi: 10.1021/ja01647a089
15. Du Vigneaud V, Ressler C, Swan JM, Roberts CW, Katsoyannis PG. The synthesis of oxytocin. *J Am Chem Soc*. (1954) 76:3115–21. doi: 10.1021/ja01641a004
16. de Bree FM, Burbach JP. Structure-function relationships of the vasopressin prohormone domains. *Cell Mol Neurobiol*. (1998) 18:173–91. doi: 10.1023/A:1022564803093
17. Bichet D, Bouvier M, Chini B, Gimpl G, Guillon G, Kimura T, et al. *Vasopressin and oxytocin receptors (version* 2019.4) in the IUPHAR/BPS Guide to Pharmacology Database. *IUPHAR/BPS Guide Pharmacol*. (2019) 2019:1–19. doi: 10.2218/gtopdb/F66/2019.4
18. Barberis C, Mouillac B, Durroux T. Structural bases of vasopressin/oxytocin receptor function. *J Endocrinol*. (1998) 156:223–29. doi: 10.1677/joe.0.1560223
19. Donaldson ZR, Young LJ. Oxytocin, vasopressin, and the neurogenetics of sociality. *Science*. (2008) 322:900–4. doi: 10.1126/science.1158668
20. Arletti R, Benelli A, Bertolini A. Oxytocin inhibits food and fluid intake in rats. *Physiol Behav*. (1990) 48:825–30. doi: 10.1016/0031-9384(90)90234-U
21. Leng G, Russell JA. The osmoresponsiveness of oxytocin and vasopressin neurones: mechanisms, allostasis and evolution. *J Neuroendocrinol*. (2019) 31:e12662. doi: 10.1111/jne.12662
22. Gwee P-C, Tay B-H, Brenner S, Venkatesh B. Characterization of the neurohypophysial hormone gene loci in elephant shark and the japanese lamprey: origin of the vertebrate neurohypophysial hormone genes. *BMC Evol Biol*. (2009) 9:47. doi: 10.1186/1471-2148-9-47
23. Lagman D, Ocampo Daza D, Widmark J, Abalo XM, Sundström G, Larhammar D. The vertebrate ancestral repertoire of visual opsins, transducin alpha subunits and oxytocin/vasopressin receptors was established by duplication of their shared genomic region in the two rounds of early vertebrate genome duplications. *BMC Evol Biol*. (2013) 13:238. doi: 10.1186/1471-2148-13-238
24. Ocampo Daza D, Lewicka M, Larhammar D. The oxytocin/vasopressin receptor family has at least five members in the gnathostome lineage, inclucing two distinct V2 subtypes. *Gen Comp Endocrinol*. (2012) 175:135–43. doi: 10.1016/j.ygcen.2011.10.011
25. Sawyer WH. Evolution of neurohypophyseal hormones and their receptors. *Fed Proc*. (1977) 36:1842–7.
26. Balment RJ, Lu W, Weybourne E, Warne JM. Arginine vasotocin a key hormone in fish physiology and behaviour: a review with insights from mammalian models. *Gen Comp Endocrinol*. (2006) 147:9–16. doi: 10.1016/j.ygcen.2005.12.022
27. Mennigen JA, Volkoff H, Chang JP, Trudeau VL. The nonapeptide isotocin in goldfish: evidence for serotonergic regulation and functional roles in the control of food intake and pituitary hormone release. *Gen Comp Endocrinol*. (2017) 254:38–49. doi: 10.1016/j.ygcen.2017.09.008
28. Acher R. Molecular evolution of fish neurohypophysial hormones: neutral and selective evolutionary mechanisms. *Gen Comp Endocrinol*. (1996) 102:157–72. doi: 10.1006/gcen.1996.0057
29. Acher R, Chauvet J, Chauvet M-T, Rouille Y. Unique evolution of neurohypophysial hormones in cartilaginous fishes: possible implications for urea-based osmoregulation. *J Exp Zool*. (1999) 284:475–84. doi: 10.1002/(SICI)1097-010X(19991001)284:5<475::AID-JEZ2>3.0.CO;2-9
30. Banerjee P, Joy KP, Chaube R. Structural and functional diversity of nonapeptide hormones from an evolutionary perspective: a review. *Gen Comp Endocrinol*. (2017) 241:4–23. doi: 10.1016/j.ygcen.2016.04.025
31. Rémy C, Girardie J, Dubois MP. Vertebrate neuropeptide-like substances in the suboesophageal ganglion of two insects: *Locusta migratoria* R. and F. (orthoptera) and *Bombyx mori* L. (lepidoptera) immunocytological investigation. *Gen Comp Endocrinol*. (1979) 37:93–100. doi: 10.1016/0016-6480(79)90050-9
32. Schooley DA, Miller CA, Proux JP. Isolation of two arginine vasopressin-like factors from ganglia of *Locusta migratoria*. *Arch Insect Biochem Physiol*. (1987) 5:157–66. doi: 10.1002/arch.940050302
33. Moore GJ, Thornhill JA, Gill V, Lederis K, Lukowiak K. An arginine vasotocin-like neuropeptide is present in the nervous system of the marine mollusc *Aplysia californica*. *Brain Res*. (1981) 206:213–8. doi: 10.1016/0006-8993(81)90119-0
34. Schot LPC, Boer HH, Swaab DF, Van Noorden S. Immunocytochemical demonstration of peptidergic neurons in the central nervous system of the pond snail *Lymnaea stagnalis* with antisera raised to biologically active peptides of vertebrates. *Cell Tissue Res*. (1981) 216:273–91. doi: 10.1007/BF00233620
35. Cruz LJ, de Santos V, Zafaralla GC, Ramilo CA, Zeikus R, Gray WR, et al. Invertebrate vasopressin/oxytocin homologs. Characterization of peptides from *Conus geographus* and *Conus straitus* venoms. *J Biol Chem*. (1987) 262:15821–4.
36. Mizuno J, Takeda N. Phylogenetic study of the arginine-vasotocin/arginine-vasopressin-like immunoreactive system in invertebrates. *Comp Biochem Physiol A Comp Physiol*. (1988) 91:739–47. doi: 10.1016/0300-9629(88)90959-0
37. Van Kesteren RE, Smit AB, De Lange RP, Kits KS, Van Golen FA, Van Der Schors RC, et al. Structural and functional evolution of the vasopressin/oxytocin superfamily: vasopressin-related conopressin is the only member present in Lymnaea, and is involved in the control of sexual behavior. *J Neurosci*. (1995) 15:5898–998. doi: 10.1523/JNEUROSCI.15-09-05989.1995
38. Van Kesteren RE, Tensen CP, Smit AB, Van Minnen J, Van Soest PF, Kits KS, et al. A novel G protein-coupled receptor mediating both vasopressin- and oxytocin-like functions of lys-conopressin in *Lymnaea stagnalis*. *Neuron*. (1995) 15:897–908. doi: 10.1016/0896-6273(95)90180-9
39. *C. elegans* Sequencing Consortium, Genome sequence of the nematode: a platform for investigating biology. *Science*. (1998) 282: 2012–8. doi: 10.1126/science.282.5396.2012
40. Elphick MR, Rowe ML. NGFFFamide and echinotocin: structurally unrelated myoactive neuropeptides derived from neurophysin-containing precursors in sea urchins. *J Exp Biol*. (2009) 212:1067–77. doi: 10.1242/jeb.027599
41. Beets I, Janssen T, Meelkop E, Temmerman L, Suetens N, Rademakers S, et al. Vasopressin/oxytocin-related signaling regulates gustatory associative learning in *C. elegans*. *Science*. (2012) 338:543–5. doi: 10.1126/science.1226860
42. Garrison JL, Macosko EZ, Bernstein S, Pokala N, Albrecht DR, Bargmann CI. Oxytocin/vasopressin-related peptides have an ancient role in reproductive behavior. *Science*. (2012) 338:540–3. doi: 10.1126/science.1226201
43. Minakata H. Oxytocin/vasopressin and gonadotropin-releasing hormone from cephalopods to vertebrates. *Ann N Y Acad Sci*. (2010) 33–42. doi: 10.1111/j.1749-6632.2010.05569.x
44. Koehbach J, Stockner T, Bergmayr C, Muttenthaler M, Gruber CW. Insights into the molecular evolution of oxytocin receptor ligand binding. *Biochem Soc Trans*. (2013) 41:197–204. doi: 10.1042/BST20120256
45. Beets I, Temmerman L, Janssen T, Schoofs L. Ancient neuromodulation by vasopressin/oxytocin-related peptides. *Worm*. (2013) 2:e24246. doi: 10.4161/worm.24246
46. Gruber CW. Physiology of invertebrate oxytocin and vasopressin neuropeptides. *Exp physiol*. (2014) 99:55–61. doi: 10.1113/expphysiol.2013.072561
47. Lockard MA, Ebert MS, Bargmann CI. Oxytocin mediated behavior in invertebrates: an evolutionary perspective. *Dev Neurobiol*. (2017) 77:128–42. doi: 10.1002/dneu.22466
48. Halanych KM. The new view of animal phylogeny. *Ann Rev Ecol Evol Syst*. (2004) 35:229–56. doi: 10.1146/annurev.ecolsys.35.112202.130124
49. Proux JP, Miller CA, Li JP, Carney RL, Girardie A, Delaage M, et al. Identification of an arginine vasopressin-like diuretic hormone from *Locusta migratoria*. *Biochem Biophys Res Commun*. (1987) 149:180–6. doi: 10.1016/0006-291X(87)91621-4

50. Stafflinger E, Hansen KK, Hauser F, Schneider M, Cazzamali G, Williamson M, et al. Cloning and identification of an oxytocin/vasopressin-like receptor and its ligand from insects. *Proc Natl Acad Sci USA*. (2008) 105:3262–7. doi: 10.1073/pnas.0710897105
51. Proux JP, Rougon-Rapuzzi G. Evidence for vasopressin-like molecule in migratory locust. Radioimmunological measurements in different tissues: correlation with various states of hydration. *Gen Comp Endocrinol*. (1980) 42:378–83. doi: 10.1016/0016-6480(80)90168-9
52. Thompson KS, Tyrer NM, May ST, Bacon JP. The vasopressin-like immunoreactive (VPLI) neurons of the locust, *Locusta migratoria*. I. anatomy. *J Comp Physiol A*. (1991) 168:605–17. doi: 10.1007/BF00215083
53. Tyrer NM, Davis NT, Arbas EA, Thompson KS, Bacon JP. Morphology of the vasopressin-like immunoreactive (VPLI) neurons in many species of grasshopper. *J Comp Neurol*. (1993) 329:385–401. doi: 10.1002/cne.903290309
54. Proux JP, Rougon G, Cupo A. Enhancement of excretion across locust malpighian tubules by a diuretic vasopressin-like hormone. *Gen Comp Endocrinol*. (1982) 47:449–57. doi: 10.1016/0016-6480(82)90123-X
55. Coast GM, Rayne RC, Hayes TK, Mallet AI, Thompson KS, Bacon JP. A comparison of the effects of two putative diuretic hormones from *Locusta migratoria* on isolated locust malpighian tubules. *J Exp Biol*. (1993) 175:1–14.
56. Thompson KS, Bacon JP. The vasopressin-like immunoreactive (VPLI) neurons of the locust, *Locusta migratoria*. II. physiology. *J Comp Physiol A*. (1991) 168:619–30. doi: 10.1007/BF00215084
57. Baines RA, Bacon JP. Pharmacological analysis of the cholinergic input to the locust VPLI neuron from an extraocular photoreceptor system. *J Neurophysiol*. (1994) 72:2864–74. doi: 10.1152/jn.1994.72.6.2864
58. Baines RA, Thompson KS, Rayne RC, Bacon JP. Analysis of the peptide content of the locust vasopressin-like immunoreactive (VPLI) neurons. *Peptides*. (1995) 16:799–807. doi: 10.1016/0196-9781(95)00038-L
59. Veenstra JA. The contribution of the genomes of a termite and a locust to our understanding of insect neuropeptides and neurohormones. *Front Physiol*. (2014) 5:454. doi: 10.3389/fphys.2014.00454
60. Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG, et al. The genome sequence of *Drosophila melanogaster*. *Science*. (2000) 287: 2185–95. doi: 10.1126/science.287.5461.2185
61. Richards S, Gibbs RA, Weinstock GM, Brown SJ, Denell R, Beeman RW, et al. The genome of the model beetle and pest *Tribolium castaneum*. *Nature*. (2008) 452:949–55. doi: 10.1038/nature06784
62. Li B, Predel R, Neupert S, Hauser F, Tanaka Y, Cazzamali G, et al. Genomics, transcriptomics, and peptidomics of neuropeptides and protein hormones in the red flour beetle *Tribolium castaneum*. *Genome Res*. (2008) 18:113–22. doi: 10.1101/gr.6714008
63. Amare A, Sweedler JV. Neuropeptide precursors in *Tribolium castaneum*. *Peptides*. (2007) 28:1282–91. doi: 10.1016/j.peptides.2007.04.014
64. Aikins MJ, Schooley DA, Begum K, Detheux M, Beeman RW, Park Y. Vasopressin-like peptide and its receptor function in an indirect diuretic signaling pathway in the red flour beetle. *Insect Biochem Mol Biol*. (2008) 38:740–8. doi: 10.1016/j.ibmb.2008.04.006
65. Gruber CW, Muttenthaler M. Discovery of defense- and neuropeptides in social ants by genome-mining. *PLoS ONE*. (2012) 7:e32559. doi: 10.1371/journal.pone.0032559
66. Di Giglio MG, Muttenthaler M, Harpsøe K, Liutkeviciute Z, Keov P, Eder T, et al. Development of a human vasopressin V(1a)-receptor antagonist from an evolutionary-related insect neuropeptide. *Sci Rep*. (2017) 7:41002. doi: 10.1038/srep41002
67. Chérasse S, Aron S. Measuring inotocin receptor gene expression in chronological order in ant queens. *Horm Behav*. (2017) 96:116–21. doi: 10.1016/j.yhbeh.2017.09.009
68. Liutkevičiute Z, Gil-Mansilla E, Eder T, Casillas-Pérez B, Di Giglio MG, Muratspahić E, et al. Oxytocin-like signaling in ants influences metabolic gene expression and locomotor activity. *FASEB J*. (2018) 32:6808–21. doi: 10.1096/fj.201800443
69. Koto A, Motoyama N, Tahara H, McGregor S, Moriyama M, Okabe T, et al. Oxytocin/vasopressin-like peptide inotocin regulates cuticular hydrocarbon synthesis and water balancing in ants. *Proc Natl Acad Sci USA*. (2019) 116:5597–606. doi: 10.1073/pnas.1817788116
70. Makki R, Cinnamon E, Gould AP. The development and functions of oenocytes. *Ann Rev Entomol*. (2014) 59:405–25. doi: 10.1146/annurev-ento-011613-162056
71. Liutkeviciute Z, Koehbach J, Eder T, Gil-Mansilla E, Gruber CW. Global map of oxytocin/vasopressin-like neuropeptide signalling in insects. *Sci Rep*. (2016) 6:39177. doi: 10.1038/srep39177
72. Marder E, Bucher D. Understanding circuit dynamics using the stomatogastric nervous system of lobsters and crabs. *Ann Rev Physiol*. (2007) 69:291–316. doi: 10.1146/annurev.physiol.69.031905.161516
73. Dickinson PS, Hull JJ, Miller A, Oleisky ER, Christie AE. To what extent may peptide receptor gene diversity/complement contribute to functional flexibility in a simple pattern-generating neural network? *Comp Biochem Physiol Part D Genomics Proteomics*. (2019) 30:262–82. doi: 10.1016/j.cbd.2019.03.002
74. Christie AE, Pascual MG. Peptidergic signaling in the crab *Cancer borealis*: tapping the power of transcriptomics for neuropeptidome expansion. *Gen Comp Endocrinol*. (2016) 237:53–67. doi: 10.1016/j.ygcen.2016.08.002
75. Saetan J, Kruangkum T, Phanthong P, Tipbunjong C, Udomuksorn W, Sobhon P, et al. Molecular cloning and distribution of oxytocin/vasopressin-like mRNA in the blue swimming crab, *Portunus pelagicus,* and its inhibitory effect on ovarian steroid release. *Comp Biochem Physiol A Mol Integr Physiol*. (2018) 218:46–55. doi: 10.1016/j.cbpa.2018.01.012
76. Oliphant A, Alexander JL, Swain MT, Webster SG, Wilcockson DC. Transcriptomic analysis of crustacean neuropeptide signaling during the moult cycle in the green shore crab, *Carcinus maenas*. *BMC Genomics*. (2018) 19:711. doi: 10.1186/s12864-018-5057-3
77. Scott E, Hudson A, Feist E, Calahorro F, Dillon J, De Freitas R, et al. An oxytocin-dependent social interaction between larvae and adult *C. elegans*. *Sci Rep*. (2017) 7:10122. doi: 10.1038/s41598-017-09350-7
78. Koziol U. Precursors of neuropeptides and peptide hormones in the genomes of tardigrades. *Gen Comp Endocrinol*. (2018) 267:116–27. doi: 10.1016/j.ygcen.2018.06.012
79. Jönsson KI, Holm I, Tassidis H. Cell biology of the tardigrades: current knowledge and perspectives. In: Tworzydlo W, Bilinski S, editors. *Evo-Devo: Non-model Species in Cell and Developmental Biology. Results and Problems in Cell Differentiation*. Cham: Springer (2019). p. 231–49. doi: 10.1007/978-3-030-23459-1_10
80. Sawyer WH, Deyrup-Olsen I, Martin AW. Immunological and biological characteristics of the vasotocin-like activity in the head ganglia of gastropod molluscs. *Gen Comp Endocrinol*. (1984) 54:97–108. doi: 10.1016/0016-6480(84)90204-1
81. Van Kesteren RE, Smit AB, Dirks RW, De With ND, Geraerts WP, Joosse J. Evolution of the vasopressin/oxytocin superfamily: characterization of a cDNA encoding a vasopressin-related precursor, preproconopressin, from the mollusc *Lymnaea stagnalis*. *Proc Natl Acad Sci USA*. (1992) 89:4593–7. doi: 10.1073/pnas.89.10.4593
82. Van Kesteren RE, Smit AB, de With ND, Van Minnen J, Dirks RW, Van der Schors RC, et al. A vasopressin-related peptide in the mollusc *Lymnaea stagnalis*: peptide structure, prohormone organization, evolutionary and functional aspects of Lymnaea conopressin. In: Joosse J, Buijs RM, Tilders FJ, editors. *Progress in Brain Research*. Elsevier (1992). p. 47–57. doi: 10.1016/S0079-6123(08)61164-4
83. Koene J. Neuro-endocrine control of reproduction in hermaphroditic freshwater snails: mechanisms and evolution. *Front Behav Neurosci*. (2010) 4:167. doi: 10.3389/fnbeh.2010.00167
84. Martinez-Padrón M, González GC, Lederis K, Lukowiak K. Localization of vasopressin-like immunoreactivity in the CNS of *Aplysia californica*. *J Neurocytol*. (1992) 21:672–8. doi: 10.1007/BF01191728
85. Martínez-Padrón M, Gray WR, Lukowiak K. Conopressin G, a molluscan vasopressin-like peptide, alters gill behaviors in *Aplysia*. *Can J Physiol Pharmacol*. (1992) 70:259–67. doi: 10.1139/y92-032
86. Thornhill JA, Lukowiak K, Cooper KE, Veale WL, Edstrom JP. Arginine vasotocin, an endogenous neuropeptide of *Aplysia*, suppresses the gill withdrawal reflex and reduces the evoked synaptic input to central gill motor neurons. *J Neurobiol*. (1981) 12:533–44. doi: 10.1002/neu.480120603
87. Lukowiak K, Thornhill JA, Cooper KE, Veale WL. Vasopressin increases the central nervous system suppressive control over gill reflex behaviours

and associated neural activity in *Aplysia*. *Can J Physiol Pharmacol*. (1980) 58:583–7. doi: 10.1139/y80-098
88. Alevizos A, Weiss KR, Koester J. Synaptic actions of identified peptidergic neuron R15 in Aplysia. I. activation of respiratory pumping. *J Neurosci*. (1991) 11:1263–74. doi: 10.1523/JNEUROSCI.11-05-01263.1991
89. Alevizos A, Weiss KR, Koester J. Synaptic actions of identified peptidergic neuron R15 in *Aplysia*. III. activation of the large hermaphroditic duct. *J Neurosci*. (1991) 11: 1282–90. doi: 10.1523/JNEUROSCI.11-05-01282.1991
90. Reich G. A new peptide of the oxytocin/vasopressin family isolated from nerves of the cephalopod *Octopus vulgaris*. *Neurosci Lett*. (1992) 134:191–4. doi: 10.1016/0304-3940(92)90514-8
91. Takuwa-Kuroda K, Iwakoshi-Ukena E, Kanda A, Minakata H. Octopus, which owns the most advanced brain in invertebrates, has two members of vasopressin/oxytocin superfamily as in vertebrates. *Regul pept*. (2003) 115:139–49. doi: 10.1016/S0167-0115(03)00151-4
92. Kanda A, Takuwa-Kuroda K, Iwakoshi-Ukena E, Minakata H. Single exon structures of the oxytocin/vasopressin superfamily peptides of octopus. *Biochem Biophys Res Commun*. (2003) 309:743–8. doi: 10.1016/j.bbrc.2003.08.061
93. Kanda A, Takuwa-Kuroda K, Iwakoshi-Ukena E, Furukawa Y, Matsushima O, Minakata H. Cloning of *Octopus* cephalotocin receptor, a member of the oxytocin/vasopressin superfamily. *J Endocrinol*. (2003) 179:281–91. doi: 10.1677/joe.0.1790281
94. Kanda A, Satake H, Kawada T, Minakata H. Novel evolutionary lineages of the invertebrate oxytocin/vasopressin superfamily peptides and their receptors in the common octopus (*Octopus vulgaris*). *Biochem J*. (2005) 387:85–91. doi: 10.1042/BJ20041230
95. Sakamoto T, Ogawa S, Nishiyama Y, Akada C, Takahashi H, Watanabe T, et al. Osmotic/ionic status of body fluids in the euryhaline cephalopod suggest possible parallel evolution of osmoregulation. *Sci Rep*. (2015) 5:14469. doi: 10.1038/srep14469
96. Henry J, Cornet V, Bernay B, Zatylny-Gaudin C. Identification and expression of two oxytocin/vasopressin-related peptides in the cuttlefish *Sepia officinalis*. *Peptides*. (2013) 46:159–66. doi: 10.1016/j.peptides.2013.05.004
97. Bardou I, Maubert E, Leprince J, Chichery R, Cocquerelle C, Launay S, et al. Distribution of oxytocin-like and vasopressin-like immunoreactivities within the central nervous system of the cuttlefish, *Sepia officinalis*. *Cell Tissue Res*. (2009) 336:249–66. doi: 10.1007/s00441-009-0763-4
98. Bardou I, Leprince J, Chichery R, Vaudry H, Agin V. Vasopressin/oxytocin-related peptides influence long-term memory of a passive avoidance task in the cuttlefish, *Sepia officinalis*. *Neurobiol Learn Mem*. (2010) 93:240–7. doi: 10.1016/j.nlm.2009.10.004
99. De Oliveira AL, Calcino A, Wanninger A. Extensive conservation of the proneuropeptide and peptide prohormone complement in mollusks. *Sci Rep*. (2019) 9:4846. doi: 10.1038/s41598-019-40949-0
100. Stewart MJ, Favrel P, Rotgans BA, Wang T, Zhao M, Sohail M, et al. Neuropeptides encoded by the genomes of the akoya pearl oyster *Pinctata fucata* and pacific oyster *Crassostrea gigas*: a bioinformatic and peptidomic survey. *BMC Genomics*. (2014) 15:840. doi: 10.1186/1471-2164-15-840
101. In VV, Ntalamagka N, O'Connor W, Wang T, Powell D, Cummins SF, et al. Reproductive neuropeptides that stimulate spawning in the Sydney rock oyster (*Saccostrea glomerata*). *Peptides*. (2016) 82:109–19. doi: 10.1016/j.peptides.2016.06.007
102. Dubos M-P, Zels S, Schwartz J, Pasquier J, Schoofs L, Favrel P. Characterization of a tachykinin signalling system in the bivalve mollusc *Crassostrea gigas*. *Gen Comp Endocrinol*. (2018) 266:110–8. doi: 10.1016/j.ygcen.2018.05.003
103. Salzet M, Bulet P, Van Dorsselaer A, Malecha J. Isolation, structural characterization and biological function of a lysine-conopressin in the central nervous system of the pharyngobdellid leech *Erpobdella octoculata*. *Eur J Biochem*. (1993) 217:897–903. doi: 10.1111/j.1432-1033.1993.tb18319.x
104. Oumi T, Ukena K, Matsushima O, Ikeda T, Fujita T, Minakata H, et al. Annetocin: an oxytocin-related peptide isolated from the earthworm, *Eisenia foetida*. *Biochem Biophys Res Commun*. (1994) 198:393–9. doi: 10.1006/bbrc.1994.1055
105. Oumi T, Ukena K, Matsushima O, Ikeda T, Fujita T, Minakata H, et al. Annetocin, an annelid oxytocin-related peptide, induces egg-laying behavior in the earthworm, *Eisenia foetida*. *J Exp Zool*. (1996) 276:151–6. doi: 10.1002/(SICI)1097-010X(19961001)276:2<151::AID-JEZ8>3.0.CO;2-N
106. Satake H, Takuwa K, Minakata H, Matsushima O. Evidence for conservation of the vasopressin/oxytocin superfamily in annelida. *Biol Chem*. (1999) 274:5605–11. doi: 10.1074/jbc.274.9.5605
107. Takahama H, Haibara K, Oumi T, Ukena K, Morishita F, Furukawa Y, et al. Immunohistochemical localization of annetocin, an earthworm oxytocin-related peptide, and identification and ultrastructural characteristics of the annetocin-secretory cells in the oligochaete earthworm *Eisenia foetida*. *Zool Sci*. (1998) 15:381–8. doi: 10.2108/zsj.15.381
108. Kawada T, Kanda A, Minakata H, Matsushima O, Satake H. Identification of a novel receptor for an invertebrate oxytocin/vasopressin superfamily peptide: molecular and functional evolution of the oxytocin/vasopressin superfamily. *Biochem J*. (2004) 382:231–7. doi: 10.1042/BJ20040555
109. Wagenaar DA, Hamilton MS, Huang T, Kristan WB, French KA. A hormone-activated central pattern generator for courtship. *Curr Biol*. (2010) 20:487–95. doi: 10.1016/j.cub.2010.02.027
110. Tessmar-Raible K, Raible F, Christodoulou F, Guy K, Rembold M, Hausen H, et al. Conserved sensory-neurosecretory cell types in annelid and fish forebrain: insights into hypothalamus evolution. *Cell*. (2007) 129:1389–400. doi: 10.1016/j.cell.2007.04.041
111. Bauknecht P, Jékely G. Large-scale combinatorial deorphanization of *Platynereis* neuropeptide GPCRs. *Cell Rep*. (2015) 12:684–93. doi: 10.1016/j.celrep.2015.06.052
112. Williams EA, Verasztó C, Jasek S, Conzelmann M, Shahidi R, Bauknecht P, et al. Synaptic and peptidergic connectome of a neurosecretory center in the annelid brain. *eLife*. (2017) 6:e26349. doi: 10.7554/eLife.26349
113. Williams EA, Jékely G. Neuronal cell types in the annelid *Platynereis dumerilii*. *Curr Opin Neurobiol*. (2019) 56:106–16. doi: 10.1016/j.conb.2018.12.008
114. McVeigh P, Atkinson L, Marks NJ, Mousley A, Dalzell JJ, Sluder A, et al. Parasite neuropeptide biology: seeding rational drug target selection? *Int J Parasitol Drugs Drug Resist*. (2012) 2:76–91. doi: 10.1016/j.ijpddr.2011.10.004
115. Marks NJ, Maule AG. Neuropeptides in helminths: occurrence and distribution. In: Geary TG, Maule AG, editors. *Neuropeptide Systems as Targets for Parasite and Pest Control. Advances in Experimental Medicine and Biology*. Boston, MA: Springer (2010) 692:49–77. doi: 10.1007/978-1-4419-6902-6_4
116. Mousley A, Novozhilova E, Kimber MJ, Day TA, Maule AG. Neuropeptide physiology in helminths. In: Geary TG, Maule AG, editors. *Neuropeptide Systems as Targets for Parasite and Pest Control. Advances in Experimental Medicine and Biology*. Boston, MA: Springer (2010) 692:78–97. doi: 10.1007/978-1-4419-6902-6_5
117. McVeigh P, Kimber MJ, Novozhilova E, Day TA. Neuropeptide signalling systems in flatworms. *Parasitology*. (2005) 131:S41–55. doi: 10.1017/S0031182005008851
118. Ukena K, Iwakoshi-Ukena E, Hikosaka A. Unique form and osmoregulatory function of a neurohypophysial hormone in a urochordate. *Endocrinology*. (2008) 149:5254–61. doi: 10.1210/en.2008-0607
119. Kawada T, Sekiguchi T, Itoh Y, Ogasawara M, Satake H. Characterization of a novel vasopressin/oxytocin superfamily peptide and its receptor from an ascidian, *Ciona intestinalis*. *Peptides*. (2008) 29:1672–8. doi: 10.1016/j.peptides.2008.05.030
120. Matsubara S, Shiraishi A, Osugi T, Kawada T, Satake H. The regulation of oocyte maturation and ovulation in the closest sister group of vertebrates. *eLife*. (2019) 8:e49062. doi: 10.7554/eLife.49062
121. Elphick MR. NG peptides: a novel family of neurophysin-associated neuropeptides. *Gene*. (2010) 458:20–6. doi: 10.1016/j.gene.2010.03.004
122. Mirabeau O, Joly JS. Molecular evolution of peptidergic signaling systems in bilaterians. *Proc Natl Acad Sci USA*. (2013) 110:E2028–37. doi: 10.1073/pnas.1219956110
123. Simakov O, Kawashima T, Marlétaz F, Jenkins J, Koyanagi R, Mitros T, et al. Hemichordate genomes and deuterostome origins. *Nature*. (2015) 527:459–65. doi: 10.1038/nature16150

124. Pani AM, Mullarkey EE, Aronowicz J, Assimacopoulos S, Grove EA, Lowe CJ. Ancient deuterostome origins of vertebrate brain signalling centres. *Nature*. (2012) 483:289–94. doi: 10.1038/nature10838
125. Yao Y, Minor PJ, Zhao Y-T, Jeong Y, Pani AM, King AN, et al. Cis-regulatory architecture of a brain signaling center predates the origin of chordates. *Nat Genet*. (2016) 48:575–80. doi: 10.1038/ng.3542
126. Sodergren E, Weinstock GM, Davidson EH, Cameron RA, Gibbs RA, Angerer RC, et al. The genome of the sea urchin *Strongylocentrotus purpuratus*. *Science*. (2006) 314:941–52. doi: 10.1126/science.1133609
127. Burke RD, Angerer LM, Elphick MR, Humphrey GW, Yaguchi S, Kiyama T, et al. A genomic view of the sea urchin nervous system. *Dev Biol*. (2006) 300:434–60. doi: 10.1016/j.ydbio.2006.08.007
128. Wood NJ, Mattiello T, Rowe ML, Ward L, Perillo M, Arnone MI, et al. Neuropeptidergic systems in pluteus larvae of the sea urchin *Strongylocentrotus purpuratus*: neurochemical complexity in a "simple" nervous system. *Front Endocrinol*. (2018) 9:628. doi: 10.3389/fendo.2018.00628
129. Lin C-Y, Oulhen N, Wessel G, Su Y-H. Chapter 11 - CRISPR/Cas9-mediated genome editing in sea urchins. In: Hamdoun A, Foltz KR, editors. *Methods in Cell Biology*. Academic Press (2019). p. 305–21. doi: 10.1016/bs.mcb.2018.10.004
130. Semmens DC, Mirabeau O, Moghul I, Pancholi MR, Wurm Y, Elphick MR. Transcriptomic identification of starfish neuropeptide precursors yields new insights into neuropeptide evolution. *Open Biol*. (2016) 6:15022. doi: 10.1098/rsob.150224
131. Odekunle EA, Semmens DC, Martynyuk N, Tinoco AB, Garewal AK, Patel RR, et al. Ancient role of vasopressin/oxytocin-type neuropeptides as regulators of feeding revealed in an echinoderm. *BMC Biol*. (2019) 17:60. doi: 10.1186/s12915-019-0680-2
132. Mayorova TD, Tian S, Cai W, Semmens DC, Odekunle EA, Zandawala M, et al. Localization of neuropeptide gene expression in larvae of an echinoderm, the starfish *Asterias rubens*. *Front Neurosci*. (2016) 10:553. doi: 10.3389/fnins.2016.00553
133. Smith MK, Wang T, Suwansa-ard S, Motti CA, Elizur A, Zhao M, et al. The neuropeptidome of the crown-of-thorns starfish, acanthaster planci. *J Proteomics*. (2017) 165:61–8. doi: 10.1016/j.jprot.2017.05.026
134. Zandawala M, Moghul I, Yañez Guerra LA, Delroisse J, Abylkassimova N, Hugall AF, et al. Discovery of novel representatives of bilaterian neuropeptide families and reconstruction of neuropeptide precursor evolution in ophiuroid echinoderms. *Open Biol*. (2017) 7:170129. doi: 10.1098/rsob.170129
135. Telford MJ, Lowe CJ, Cameron CB, Ortega-Martinez O, Aronowicz J, Oliveri P, et al. Phylogenomic analysis of echinoderm class relationships supports asterozoa. *Proc Biol Sci*. (2014) 281:20140479. doi: 10.1098/rspb.2014.0479
136. Suwansa-Ard S, Chaiyamoon A, Talarovicova A, Tinikul R, Tinikul Y, Poomtong T, et al. Transcriptomic discovery and comparative analysis of neuropeptide precursors in sea cucumbers (Holothuroidea). *Peptides*. (2018) 99:231–40. doi: 10.1016/j.peptides.2017.10.008
137. Chen M, Talarovicova A, Zheng Y, Storey KB, Elphick MR. Neuropeptide precursors and neuropeptides in the sea cucumber *Apostichopus japonicus*: a genomic, transcriptomic and proteomic analysis. *Sci Rep*. (2019) 9:8829. doi: 10.1038/s41598-019-45271-3
138. Gavilán B, Sprecher SG, Hartenstein V, Martinez P. The digestive system of xenacoelomorphs. *Cell Tissue Res*. (2019) 377:369–82. doi: 10.1007/s00441-019-03038-2
139. Hejnol A, Pang K. Xenacoelomorpha's significance for understanding bilaterian evolution. *Curr Opin Genet Dev*. (2016) 39:48–54. doi: 10.1016/j.gde.2016.05.019
140. Telford MJ, Copley RR. Zoology: war of the worms. *Curr Biol*. (2016) 26:R335–7. doi: 10.1016/j.cub.2016.03.015
141. Hejnol A, Obst M, Stamatakis A, Ott M, Rouse GW, Edgecombe GD, et al. Assessing the root of bilaterian animals with scalable phylogenomic methods. *Proc R Soc*. (2009) 276:4261–70. doi: 10.1098/rspb.2009.0896
142. Cannon JT, Vellutini BC, Smith J, Ronquist F, Jondelius U, Hejnol A. Xenacoelomorpha is the sister group to nephrozoa. *Nature*. (2016) 530:89–93. doi: 10.1038/nature16520
143. Philippe H, Brinkmann H, Copley RR, Moroz LL, Nakano H, Poustka AJ, et al. Acoelomorph flatworms are deuterostomes related to *Xenoturbella*. *Nature*. (2011) 470:255–8. doi: 10.1038/nature09676
144. Philippe H, Poustka AJ, Chiodin M, Hoff KJ, Dessimoz C, Tomiczek B, et al. Mitigating anticipated effects of systematic errors supports sister-group relationship between xenacoelomorpha and ambulacraria. *Curr Biol*. (2019) 29:1818–26.e6. doi: 10.1016/j.cub.2019.04.009
145. Thiel D, Franz-Wachtel M, Aguilera F, Hejnol A. Xenacoelomorph neuropeptidomes reveal a major expansion of neuropeptide systems during early bilaterian evolution. *Mol Biol Evol*. (2018) 35:2528–43. doi: 10.1093/molbev/msy160
146. Jekely G. Global view of the evolution diversity of metazoan neuropeptide signaling. *Proc Natl Acad Sci USA*. (2013) 110:8702–7. doi: 10.1073/pnas.1221833110
147. Mita M, Yoshikuni M, Ohno K, Shibata Y, Paul-Prasanth B, Pitchayawasin S, et al. A relaxin-like peptide purified from radial nerves induces oocyte maturation and ovulation in the starfish, *Asterina pectinifera*. *Proc Natl Acad Sci USA*. (2009) 106:9507–12. doi: 10.1073/pnas.0900243106
148. Lin M, Mita M, Egertová M, Zampronio CG, Jones AM, Elphick MR. Cellular localization of relaxin-like gonad-stimulating peptide expression in *Asterias rubens*: new insights into neurohormonal control of spawning in starfish. *J Comp Neurol*. (2017) 525:1599–617. doi: 10.1002/cne.24141
149. Lavoie ME. How sea stars open bivalves. *Biol Bull*. (1956) 111:114–22. doi: 10.2307/1539188
150. Kanatani H. Hormones in echinoderms. In: Barrington E, editor. *Hormones and Evolution*. London: Academic Press. (1979). p. 273–307.
151. Hasse C, Rebscher N, Reiher W, Sobjinski K, Moerschel E, Beck L, et al. Three consecutive generations of nephridia occur during development of *Platynereis dumerilii* (Annelida, Polychaeta). *Dev Dyn*. (2010) 239:1967–76. doi: 10.1002/dvdy.22331
152. Forsling ML. Diurnal rhythms in neurohypophysial function. *Exp Physiol*. (2000) 85:179S—86. doi: 10.1111/j.1469-445X.2000.tb00022.x
153. Jekely G, Melzer S, Beets I, I.Kadow CG, Koene J, Haddad S, et al. The long and the short of it - a perspective on peptidergic regulation of circuits and behaviour. *J Exp Biol*. (2018) 221:1–14. doi: 10.1242/jeb.166710


**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 20 May 2020
doi: 10.3389/fnins.2020.00366

# Tracing the Origins of the Pituitary Adenylate-Cyclase Activating Polypeptide (PACAP)

*João C. R. Cardoso*, Manuel G. Garcia and Deborah M. Power**

*Comparative Molecular and Integrative Biology, Centre of Marine Sciences, University of Algarve, Faro, Portugal*

Pituitary adenylate cyclase activating polypeptide (PACAP) is a well-conserved neuropeptide characteristic of vertebrates. This pluripotent hypothalamic neuropeptide regulates neurotransmitter release, intestinal motility, metabolism, cell division/differentiation, and immunity. In vertebrates, PACAP has a specific receptor ($PAC_1$) but it can also activate the Vasoactive Intestinal Peptide receptors ($VPAC_1$ and $VPAC_2$). The evolution of the vertebrate PACAP ligand – receptor pair has been well-described. In contrast, the situation in invertebrates is much less clear. The PACAP ligand – receptor pair in invertebrates has mainly been studied using heterologous antibodies raised against mammalian peptides. A few partial PACAP cDNA clones sharing >87% aa identity with vertebrate PACAP have been isolated from a cnidarian, several protostomes and tunicates but no gene has been reported. Moreover, current evolutionary models of the peptide and receptors using molecular data from phylogenetically distinct invertebrate species (mostly nematodes and arthropods) suggests the PACAP ligand and receptors are exclusive to vertebrate genomes. A basal deuterostome, the cephalochordate amphioxus (*Branchiostoma floridae*), is the only invertebrate in which elements of a PACAP-like system exists but the peptides and receptor share relatively low sequence conservation with the vertebrate homolog system and are a hybrid with the vertebrate glucagon system. In this study, the evolution of the PACAP system is revisited taking advantage of the burgeoning sequence data (genome and transcriptomes) available for invertebrates to uncover clues about when it first appeared. The results suggest that elements of the PACAP system are absent from protozoans, non-bilaterians, and protostomes and they only emerged after the protostome-deuterostome divergence. PACAP and its receptors appeared in vertebrate genomes and they probably shared a common ancestral origin with the cephalochordate PACAP/GCG-like system which after the genome tetraploidization events that preceded the vertebrate radiation generated the PACAP ligand and receptor pair and also the other members of the Secretin family peptides and their receptors.

**Keywords: deuterostomes, early metazoan, evolution, protostomes, neuropeptide, receptor**

OPEN ACCESS

***Edited by:***
*Giovanne B. Diniz,*
*Yale University, United States*

***Reviewed by:***
*Lars Edvinsson,*
*Lund University, Sweden*
*Dora Reglodi,*
*University of Pécs, Hungary*

****Correspondence:***
*João C. R. Cardoso*
*jccardo@ualg.pt*
*Deborah M. Power*
*dpower@ualg.pt*

***Specialty section:***
*This article was submitted to Neuroendocrine Science, a section of the journal Frontiers in Neuroscience*

***Received:*** *09 January 2020*
***Accepted:*** *25 March 2020*
***Published:*** *20 May 2020*

***Citation:***
*Cardoso JCR, Garcia MG and Power DM (2020) Tracing the Origins of the Pituitary Adenylate-Cyclase Activating Polypeptide (PACAP). Front. Neurosci. 14:366. doi: 10.3389/fnins.2020.00366*

## INTRODUCTION

The pituitary adenylate cyclase-activating polypeptide (PACAP) is one of the most extensively studied neuropeptides due to its biomedical interest and its well-conserved functions in vertebrates. The first description of PACAP was over 30 years ago when it was identified in extracts of ovine hypothalamus as a factor that stimulated cAMP production in anterior pituitary cells (Miyata et al., 1989, 1990). Since then PACAP has been isolated and characterized in representatives of most of the major vertebrate phyla and has a diversity of functions including the regulation of neurotransmission, vasodilation, intestinal motility, cell proliferation and differentiation and immunity (Sherwood et al., 2000; Vaudry et al., 2000, 2009). Recently PACAP was also described as a potent antimicrobial peptide (AMP) (Starr et al., 2018).

Pituitary adenylate cyclase-activating polypeptide has been linked to several clinical disorders that have highlighted its important role as a neurotransmitter and also its neuroprotective actions (Shioda and Nakamachi, 2015; Maasz et al., 2017; Denes et al., 2019). In the mammalian brain PACAP is most abundant in the hypothalamus from where it is secreted to the pituitary gland but it is also detected in other brain regions such as the telencephalon, cerebellum, and brainstem (Arimura et al., 1991; Ghatei et al., 1993; Hirabayashi et al., 2018; Warfvinge and Edvinsson, 2019). In rat the distribution of the PACAP system is relatively well-characterized and PACAP maps to the parvo- and magnocellular neurons of the paraventricular (PVN) and supraoptic (SON) nuclei of the hypothalamus (Vaudry et al., 2009; Warfvinge and Edvinsson, 2019).

Pituitary adenylate cyclase-activating polypeptide is a member of the Secretin brain-gut peptide superfamily. Secretin (SCT) was identified more than 100 years ago by Bayliss and Starling who demonstrated its role in the regulation of pancreatic secretion and "coined the term" *hormone* to describe the mode of action of the factor (Bayliss and Starling, 1902). The SCT superfamily is a group of small peptides that share similarity at the level of their amino acid sequence and structure. In humans the PACAP-like superfamily includes SCT, Vasoactive Intestinal Peptide (VIP), Peptide Histidine Isoleucine (PHI), PACAP-Related Peptide (PRP) and Growth Hormone-Releasing Hormone (GHRH). The glucagon-like peptides (Glucagon, GCG; Glucagon-Like Peptide 1 and 2, GLP 1 and 2; and Glucose-dependent Insulinotropic Peptide; GIP) are also members of the SCT superfamily but they diverged earlier than the other peptide members. All the peptides of the SCT superfamily are proposed to have arisen from a common ancestral gene by exon duplication followed by local duplication and expansion during the two genome tetraploidization events prior to the vertebrate radiation (**Figure 1**) (Sherwood et al., 2000; Cardoso et al., 2007a, 2010; Ng et al., 2012; Hwang et al., 2013). Vertebrate PRP is encoded by the same gene as PACAP and the peptide VIP is encoded in the same gene precursor as PHI but SCT and GHRH are encoded by specific genes. Within the SCT family of peptides, PACAP shares the highest amino acid sequence resemblance (up to 68%) with VIP, a peptide that was first described as a potent vasodilator in the pig small intestine by Said and Mutt (1970).

In mammals, PACAP is encoded by the *ADCYAP1* gene, which has four exons. Exon 3 of the gene encodes the peptide PRP and exon 4 encodes PACAP and originates two biologically active peptide isoforms (Sherwood et al., 2000; Vaudry et al., 2000, 2009). PACAP-38 is the predominant form and the shorter form, PACAP-27, arises by post-translational processing of PACAP-38 and shares the same N-terminal amino acid (aa) sequence but has a shorter C-terminus (Miyata et al., 1989, 1990; Cox, 1992; Arimura and Shioda, 1995; Sherwood et al., 2000; Vaudry et al., 2000, 2009). In the genomes of mammals and other tetrapods a single *ADCYAP1* gene exists and the peptide it encodes has high sequence conservation between species. In fish a single gene encoding the PACAP peptide precursor that shares high sequence similarity and organization with the tetrapod homolog was identified in the genomes of lamprey, elephant shark, spotted gar, and coelacanth. In contrast, in the teleosts, two PACAP precursor genes *adcyap1a* (protein; Pacapa) and *adcyap1b* (protein; Pacapb), have been isolated and arose from the teleost specific genome duplication event and four mature PACAP peptides (two PACAP-38 and two PACAP-27) are considered to be produced (Cardoso et al., 2007a, 2015; Ng et al., 2012). Analysis across the vertebrates of the genes flanking *ADCYAP1* reveals syntenic genome regions in fish and tetrapods and supports a common evolutionary origin for PACAP (Cardoso et al., 2007b, 2015).

In amphibian and fish (teleost and cartilaginous fish) brain PACAP is abundant and has a similar distribution to that in mammals and is predominantly expressed in the hypothalamic nuclei but also in other brain regions (Valiante et al., 2006; Vaudry et al., 2009). In the zebrafish that has duplicated *adcyap1* genes, transcripts for *adcyap1a* are most expressed in the brain stem and diencephalon, while *adcyap1b* gene transcripts are abundant in the telencephalon and diencephalon (Nakamachi et al., 2019). Although, the distribution of PACAP is relatively well-characterized in fish, few studies have characterized the function of the teleost duplicate PACAPs. The outcome of the studies that exist suggest that the duplicate PACAPs possess similar functions to the mammalian homolog although teleost specific functions are proposed to have also been acquired (Cardoso et al., 2010, 2015).

Homologs of the PACAP system have been predicted in invertebrates and the identified peptides share high sequence conservation with vertebrate PACAP (McRory and Sherwood, 1997; Cardoso et al., 2007a, 2010; Kiss and Pirger, 2013; Lugo et al., 2013; Pirger et al., 2016) (**Figure 2**). This proposal is based on the isolation of partial cDNAs encoding a PACAP-like

**Abbreviation:** ADCYAP1, pituitary adenylate cyclase activating polypeptide gene; ADCYAP1R1, pituitary adenylate cyclase activating polypeptide receptor gene; CAL, calcitonin; CALCR, calcitonin receptor; CRH, corticotrophin releasing hormone; GCG, glucagon; GCGR, glucagon receptor; GPCRs, G-protein coupled receptors; PACAP, pituitary adenylate cyclase activating polypeptide; $PAC_1$, pituitary adenylate cyclase activating peptide receptor; PDF, pigment dispersing factor; PDFR, pigment dispersing factor receptor; PTH, parathyroid hormone; PTHR, parathyroid hormone receptor; SCT, secretin; TSA, transcriptome shotgun assembly sequence; VIP, vasoactive intestinal peptide; $VPAC_1$, vasoactive intestinal peptide receptor 1; $VPAC_2$, vasoactive intestinal peptide receptor 2; VIPR1, vasoactive intestinal peptide receptor 1 gene; VIPR2, vasoactive intestinal peptide receptor 2 gene; WGS, whole genome shotgun.

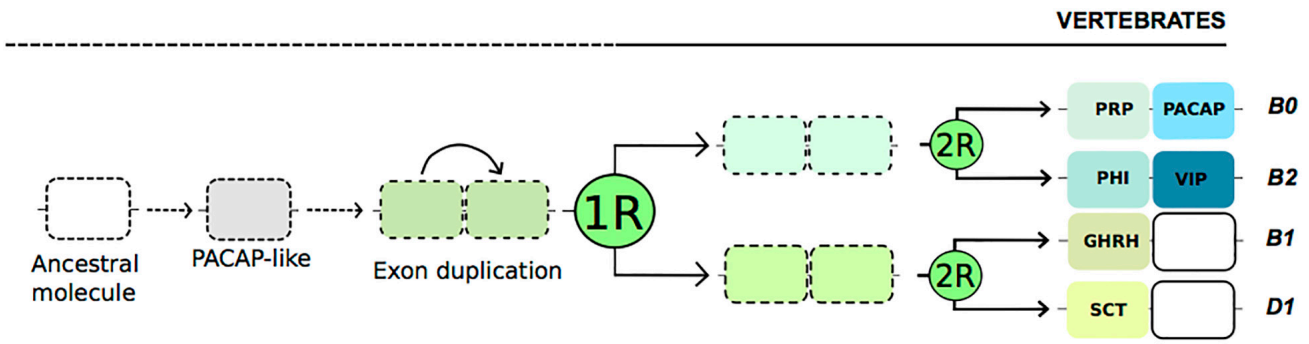


**FIGURE 1 |** Current evolutionary model for Secretin peptide family members. The exons of the peptide coding region are represented. PACAP-like family members are proposed to have evolved from a common ancestor exon early in the chordate radiation via exon and gene/chromosome duplication events (Sherwood et al., 2000; Cardoso et al., 2007a, 2010; Ng et al., 2012; Hwang et al., 2013). The same ancestral molecule is also suggested to originate the glucagon(GCG)-like peptides but for simplicity this is not represented in the figure. Boxes represent exons and lines introns and the peptide coding exons are indicated by the peptide abbreviation. Dashed lines indicate undefined evolutionary pathways. PACAP and PRP and VIP and PHI share the same gene precursor and GHRH and SCT are encoded by a single exon. The predicted position of each gene block in the Gnathostome Ancestral Genome (GAC) linkage groups are indicated (Hwang et al., 2013). The genome tetraploidization events (1R and 2R) are represented.

| | | % ID | | |
|---|---|---|---|---|
| | | H | T | GT |
| Human (H) | -HSDGIFTDSYSRYRKQMAVKKYLAAVL**GKR**YKQRVKNK | 100 | 89 | 84 |
| Trout (T) | -HSDGIFTDSYSRYRKQMAVKKYLAAVL**GKR**YRQRYRSK | 89 | 100 | 97 |
| Giant trout (GT) | -HSDGIFTDSYSRYRKQMAVKKYLAAVL**GKR**YRQRYRNK | 92 | 97 | 100 |
| *Chelyosoma_1* | -HSDGIFTDSYSRYRNQMAVKKYLAAVL---------- | 97 | 96 | 89 |
| *Chelyosoma_2* | -HSDGIFTDSYSRYRNQMAVKKYINALL---------- | 85 | 85 | 85 |
| *Halocynthia* | -HSDGIFTDSYSRYRKQMAVKKYLAAVL**GKR**YRQRYRNE | 89 | 89 | 89 |
| *Branchiostoma_a* | QLGDQSITSEMSVRLREAEARRLLQSLMAKQG------- | 17 | 17 | 13 |
| *Branchiostoma_b* | ALGDQGFTSDLASKLSEAEARRMIQTLMAQAIG------ | 10 | 9 | 13 |
| *Branchiostoma_c* | QLGDQGVTSALAARLEQAEARQYIKDLLAQAVG------ | 16 | 16 | 16 |
| Coakroach | -HSDGIFTDSYSRYRKQMAVKKYLAAVL**GKR**YRQRYRSK | 89 | 100 | 95 |
| Shrimp | -HSDGIFTDSYSRYRKQMAVKKYLAAVL**GRR**YRQRYRNE | 87 | 92 | 95 |
| Crab | -HSDGIFTDSYSRYREQMAVKKYLAAVL**GKR**YRQRYRNK | 92 | 97 | 100 |
| Squid | -HSDGIFTDSYSRYRKQMAVKKYLAAVL**GKR**YRQRYRNK | 89 | 95 | 97 |
| Planaria | -HSDGIFTDSYSRYRKQMAVKKYLAAVL**GKR**YRQRYRNK | 92 | 97 | 100 |
| Hydra | -HSDGIFTDSYSRYRKQMAVKKYLAAVL**GKR**YRQRYRNK | 92 | 97 | 100 |
| | .* .*. : : .:: : :: | | | |

**FIGURE 2 |** Sequence conservation of the invertebrate PACAP mature peptides. The mature sequence of the invertebrate peptides were extracted and compared with PACAP from the human (P18509) and two Salmonidae fish the river trout (*Salmo trutta,* XM_029756051.1) and the Siberian giant trout (*Hucho taimen*, HAGJ01147357.1) peptide homologs. The post-translational internal cleavage–amidation site ($Gly^{28}$-$Lys^{29}$-$Arg^{30}$) that generates the shortest peptide isoform (PACAP-27) in human, which is predicted in the other peptide sequences is underlined. The percentage of amino acid sequence identity (%ID) for the human (H), river trout (T), and Siberian giant trout (GT) mature PACAP-38 peptides is given. The vertebrate, hydra, protostome and invertebrate deuterostome mature peptide sequences were used to interrogate the protozoan, non-bilaterian, protostome and invertebrate deuterostome genomes and transcriptomes for homologs. Accession numbers of the non-vertebrate peptides are: cockroach (*Periplaneta americana*, AB083652), crab (*Eriocheir japonica*, AB121765), squid (*Sepioteuthis lessoniana*, AB083651), planarian (*Dugesia japonica*, AB083649), and hydra (*Hydra magnipapillata*, AB083650). The tunicate *Chelyosoma productum* (Chelyosoma-1 and Chelyosoma_2, were obtained from McRory and Sherwood, 1997) and the shrimp (*Litopenaeus vannamei*) from Lugo et al. (2013). Complete amino acid conservation is annotated with "*", partial conservation with "." and the position of the consensus amino acids conserved in the greatest number of sequences is indicated with ":".

peptide from a cnidarian, a few protostomes and tunicates and immunohistochemical studies with heterologous antisera in the 1980 and 1990s (Pirger et al., 2016). Surprisingly, data mining of publicly available invertebrate genomes in the late 2000s, failed to identify sequence homologs of the vertebrate genes (Cardoso et al., 2007a, 2010). Recently, in a cephalochordate a PACAP/GCG-like peptide and receptor were described and their function characterized (On et al., 2015). The predicted cephalochordate peptides and receptor shared low sequence conservation (<17% peptide, <40% receptor) with the vertebrate homologs, which is at odds with the high sequence conservation of the putative peptides of other invertebrate and raises questions about the existing models for peptide/receptor evolution. The advent of next generation sequencing (NGS) and the massive increase in publicly available invertebrate genomes and transcriptomes provides a unique opportunity to trace the evolution of gene families. In this study we revisited the evolution of PACAP by examining past studies and exploring current publicly available genome and transcriptome data for representatives of major non-vertebrate phyla (Protozoans, non-bilaterians, Ecdysozoans, Lophotrochozoan, and invertebrate deuterostomes). The nomenclature for PACAP and its receptors used in this review for the vertebrate species follows the guidelines established by the International Union of Pharmacology (IUPHAR) (Harmar et al., 1998) and the zebrafish nomenclature convention for fish[1] (**Table 1**). Nomenclature for urochordate, cephalochordate and protostome follow (McRory and Sherwood, 1997; On et al., 2015; Pirger et al., 2016).

**TABLE 1 |** Nomenclature for PACAP and its receptors.

| PACAP | | PACAP Receptor | |
|---|---|---|---|
| **Gene/transcripts** | **Peptide** | **Gene/transcripts** | **Protein** |
| **Primate** | | | |
| *ADCYAP1* | PACAP | *ADCYAP1R1* | $PAC_1$ |
| | | *VIPR1* | $VPAC_1$ |
| | | *VIPR2* | $VPAC_2$ |
| **Mammalian (non-primate)** | | | |
| *Adcyap1* | PACAP | *Adcyap1R1* | $PAC_1$ |
| | | *Vipr1* | $VPAC_1$ |
| | | *Vipr2* | $VPAC_2$ |
| **Aves** | | | |
| *ADCYAP1* | PACAP | *ADCYAP1R1* | $PAC_1$ |
| | | *VIPR1* | $VPAC_1$ |
| | | *VIPR2* | $VPAC_2$ |
| **Actinopterygii (non-teleost)** | | | |
| *adcyap1* | Pacap | *adcyap1r1* | $Pac_1$ |
| | | *vipr1* | $Vpac_1$ |
| | | *vipr2* | $Vpac_2$ |
| **Teleost** | | | |
| *adcyap1a* | Pacapa | *adcyap1r1a* | $Pac_1a$ |
| *adcyap1b* | Pacapb | *adcyap1r1b* | $Pac_1b$ |
| | | *vipr1a* | $Vpac_1a$ |
| | | *vipr1b* | $Vpac_1b$ |
| | | *vipr2a* | $Vpac_2a$ |
| | | *vipr2b* | $Vpac_2b$ |
| **Agnathan** | | | |
| *ADCYAP1* | PACAP | *VIPR* | VPAC |
| **Urochordate** | | | |
| *pacap1/pacap2* | PACAP | ni | ni |
| **Cephalochordate** | | | |
| *PACAP/GCG* | PACAP/GCG | *PACAP/GCGR* | PACAP/GCGR |
| **Protostomes** | | | |
| ni | PACAP | ni | ni |
| **Cnidaria** | | | |
| ni | PACAP | ni | ni |

*Receptor nomenclature is according to IUPHAR (Harmar et al., 1998) and the ZFIN Zebrafish Nomenclature Convention (https://wiki.zfin.org/). Nomenclature for urochordate, cephalochordate, and protostome follow (On et al., 2015; Pirger et al., 2016). ni, not identified.*

## PACAP RECEPTORS IN VERTEBRATES

The discovery of PACAP in vertebrates was soon followed by the identification of its specific receptor $PAC_1$. PACAP also stimulates the activity of the VIP receptors (VPAC1 and VPAC2) and this explains why the two peptides have an overlapping spectrum of physiological activities. The receptors for PACAP have a widespread distribution in the CNS of vertebrates (reviewed by Harmar et al., 2012; Hirabayashi et al., 2018). In the rat brain $PAC_1$ is most abundant and has a similar distribution to PACAP and is present in the hypothalamus and non-hypothalamic brain regions (e.g., cerebellum and spinal cord). The expression of the other PACAP receptors is sparse and $VPAC_1$ is found in the cerebral cortex and hippocampus and $VPAC_2$ in the amygdala, hippocampus, thalamus, and hypothalamus (Hirabayashi et al., 2018). Overall the PACAP system maps to brain regions associated with the stress response, reward seeking and aversive responses (Warfvinge and Edvinsson, 2019).

Pituitary adenylate cyclase activating polypeptide receptors and the receptors for other SCT-like peptides [e.g., parathyroid hormone/parathyroid hormone-related peptide (PTH/PTHrP), calcitonin/calcitonin gene-related peptide (CAL/CGRP), and corticotrophin-releasing hormone (CRH)] form a large family of receptor proteins that belong to the GPCR family B1, also known as Secretin-GPCRs or class II. Family B1 GPCRs possess seven transmembrane domains and a relatively long N-terminus (∼120 amino acids) with six conserved cysteine residue that form three disulphide bridges and create a ligand binding pocket (Harmar, 2001; Couvineau et al., 2012; Harmar et al., 2012). Family B1 receptors share a common origin and emerged prior to the protostome–deuterostome divergence since homologs of the vertebrate members exist in invertebrates (Cardoso et al., 2006, 2014; Hwang et al., 2013). B1 family receptors are suggested to share the same ancestral precursor gene as the adhesion-GPCRs, a group of receptors involved in cell growth, differentiation, and immunity (Springer, 1990; Rosales et al., 1995; Nordström et al., 2009).

[1] https://wiki.zfin.org/

Upon receptor binding the PACAP peptide triggers intracellular signal transduction and a biological response. Signaling involves trimeric G-protein complexes that when coupled to the receptor C-terminal domain, stimulate a series of intracellular signaling pathway, which predominately involve, (a) the production of cyclic adenosine monophosphate (cAMP) via the adenylate-cyclase (AC) pathway or (b) the mobilization of the calcium ion ($Ca^{2+}$) pathway involving phospholipase C and inositol 1,4,5-triphosphate (IP3) activity (Rawlings, 1994; Laburthe and Couvineau, 2002; Harmar et al., 2012; Langer, 2012). In teleosts, PACAP receptor number is duplicated in comparison to tetrapods. Six putative PACAP receptor genes (two $Pac_1$, *adcyap1r1a* and *adcyap1r1b*; two $Vpac_1$, *vipr1a* and *vipr1b* and two $Vpac_2$, *vipr2a,* and *vipr2b*) have been characterized and receptor activation triggers similar signaling pathways to those in mammals, birds, and amphibians. The teleost PACAP receptor gene paralogs stimulate similar functions to those of other vertebrates but also acquired specialized functions during the teleost radiation (Cardoso et al., 2004, 2007b, 2015; Roch et al., 2009). In common with the human receptors, interaction of teleost family B1 GPCRs with receptor activity-modifying proteins (RAMPs) (Christopoulos et al., 2003; Archbold et al., 2011; Couvineau and Laburthe, 2011), a class of membrane accessory proteins, can modulate their activity by changing receptor pharmacology (Cardoso et al., 2015).

## THE PUZZLING EXISTENCE OF A PACAP PRECURSOR IN INVERTEBRATES

Pituitary adenylate cyclase activating polypeptide is proposed to be one of the most well-conserved neuropeptides in the animal kingdom since it has been reported to exist from invertebrates to vertebrates. In invertebrates, putative PACAP-like peptides sharing high sequence identity (>87% aa identity) with the human homolog or identical to the teleost fish peptides have been described (**Figure 2**). For example, in tunicates, the closest relative to vertebrates (Delsuc et al., 2006), two full length cDNAs (*pacap1* and *pacap2*) encoding PACAP peptides were isolated from the marine disk-top tunicate (*Chelyosoma productum*) (McRory and Sherwood, 1997) and a putative partial PACAP cDNA was isolated from the sea pineapple (*Halocynthia roretzi*) and their deduced peptides were highly identical in sequence to the vertebrate PACAP (**Figure 2**). In the cnidarian (*Hydra magnipapillata*) and several protostomes partial PACAP cDNA sequences have also been reported. Three arthropods, the crab (*Eriocheir japonica*), the white shrimp (*Litopenaeus vannamei*) and the cockroach (*Periplaneta americana*), a mollusc, the squid (*Sepioteuthis lessoniana*), and the planarian (*Dugesia japonica*) are reported to possess cDNA encoding putative peptides highly identical (>89% aa identity) to the human (Kiss and Pirger, 2013; Lugo et al., 2013) and the trout peptides (>95% aa identity) (**Figure 2**). A PACAP gene or transcript homolog from invertebrate species with a sequenced genome remains to be convincingly demonstrated (Cardoso et al., 2010).

Furthermore, searches performed in the sea squirt [*Ciona intestinalis*, a.k.a. *Ciona intestinalis type A* (*Ciona robusta*)] genome which has evolutionary proximity with the disk-top tunicate (*Chelyosoma productum*) and sea pineapple (*Halocynthia roretzi*) failed to retrieve a homolog peptide encoding gene (Cardoso et al., 2007a, 2010). Similarly, searches in the genome of an echinoderm, the sea urchin (*Strongylocentrotus purpuratus*) also failed to retrieve a homolog of the vertebrate PACAP gene (Cardoso et al., 2010). Recently, a PACAP/GCG-like peptide gene encoding three putative mature peptides was identified in the genome of the cephalochordate, amphioxus (*Branchiostoma floridae*), the closest extant organism to tunicates. However, in contrast to PACAP in other invertebrates, the amphioxus PACAP/GCG-like peptide had extremely low sequence conservation (<17% aa identity) with the vertebrate PACAP and only a few functionally important amino acid residues for the peptide bioactivity were found (Mirabeau and Joly, 2013; On et al., 2015) (**Figure 2**).

In the present study we took advantage of the recently released genomes (whole genome shotgun assemblies, WGS), and transcriptomes [Transcriptome Shotgun Assembly (TSA), computationally assembled mRNA sequences from ESTs and raw sequence reads, **Supplementary Tables 1, 2**] to search for molecular evidence that PACAP emerged early during evolution and was highly conserved from single-celled organisms (protozoans) to invertebrates (non-bilaterian animals and from protostomes and invertebrate deuterostomes) and vertebrates. The conserved mature human and invertebrate PACAP-like peptides were used to screen nucleotide databases for sequence homologs (**Figure 2**). Despite the limitations caused by the small size of the metazoan PACAP mature peptides (non-specific sequence matches tend to be high), we reasoned that the high degree of sequence conservation between the human and the previously described cnidarian, protostome, and tunicate sequences means that if a homolog exists it should be found using sequence identity searches. Searches for PACAP homologs also included the PACAP/GCG-like peptides recently described in the cephalochordate (On et al., 2015). To favor the identification of short peptide hits with strong similarities the BLAST algorithm was automatically adjusted. The state of the art about PACAP or other SCT-like peptides in single cell organisms to invertebrate deuterostomes was updated in the current study. It was reasoned that characterization of PACAP across phylogenetically distinct non-vertebrate species should reveal the origin and evolution of this important neuropeptide and give clues about function that can contribute to understanding the acquisition of its pleotropic actions in vertebrates. The evolution and phylogeny of the PACAP receptor is also briefly considered as an adjunct to understanding ligand evolution (see section "PACAP Receptors in Invertebrates").

## PACAP PRECURSOR IN PROTOZOANS

Protozoans are a group of free-living single-celled organisms or parasitic microorganisms. The first description of PACAP signaling in a single celled eukaryote was obtained from the free-living ciliate protozoan *Tetrahymena thermophila* a biological and biomedical model commonly used to study avoidance behavior

(Hassenzahl et al., 2001; Lucas et al., 2004). *T. thermophila* was described to be repelled by human PACAP-38 leading to the suggestion that a receptor for PACAP exists in *T. thermophila*. Although no putative protozoan PACAP receptor has been identified peptide signaling is proposed to be similar to that in vertebrates and involve the activation of an intracellular G-protein complex which stimulates both the Adenylyl Cyclase and phospholipase C pathways but also the NO/cGMP pathway (Hassenzahl et al., 2001; Lucas et al., 2004). In addition, the pharmacological profile of the putative *T. thermophila* PACAP receptor is suggested to be distinct from the vertebrate homolog as it was activated by the vertebrate receptor antagonists (PACAP 6-27 and PACAP 6-38) (Keedy et al., 2003).

The genome of *T. thermophila* is available (wgs projects: AAGF, AFSS) (Stover, 2006) as is the genome and transcriptome for other members of the superphylum Alveolata (taxid: 33630). The Alveolata are a monophyletic group that include the Ciliophora – aveolates that have short hair-like cilia such as is found in *T. thermophila* and the phylum Apicomplexa – a large phylum of parasitic alveolates that includes the malaria parasite. The results of searches of the genomes and transcriptomes of species from this superphylum in the present study failed to retrieve putative peptide transcripts or genes highly related in sequence to the metazoan PACAP.

## PACAP PRECURSOR IN NON-BILATERIANS

The sponges (Porifera), the oldest animal phylum, the comb jellies (Ctenophora), jellyfish and corals (Cnidaria) and plate animals (Placozoa) are four evolutionarily ancient phyla of non-bilaterian animals (**Figure 3**). They together form the non-bilaterian animals that diverged more than 600 million years ago from the metazoan lineage (Pisani et al., 2015; King and Rokas, 2017). A partial cDNA encoding for a PACAP-like peptide that shares 92% aa identity with human PACAP 38 was isolated and deposited in the NCBI database from the fresh-water polyp *Hydra magnipapillata* (a.k.a. *Hydra vulgaris* or *Hydra attenuata*) (**Figure 2**) (reviewed in Cardoso et al., 2010; Pirger et al., 2016) and this is the only non-bilaterian PACAP described. No functional or expression data in cnidaria or in any other non-bilaterian animal exists (**Figure 3**).

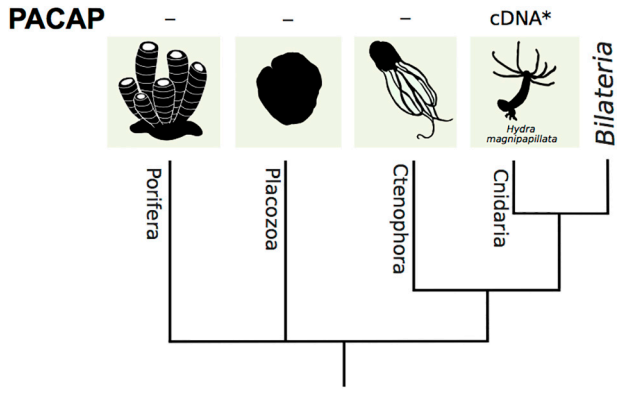


**FIGURE 3 |** A dendrogram representing the phylogenetic relationship of non-bilaterian phyla. A partial PACAP cDNA was isolated from the *Hydra magnipapillata*. In other non-bilaterian phyla the existence of PACAP remains to be established (-). The evolutionary relationship between the species was based on (Brunet and King, 2017).

Whole body transcriptome data for seven Porifera species, *Amphimedon queenslandica* (GBXN), *Corticium candelabrum* (GAQT), *Cymbastela stipitate* (GHWA); *Haliclona tubifera* (GFAV); *Halisarca caerulea* (GFSI/GFTO/GFTP/GFTQ), *Halisarca dujardini* (HADA), *Sycon coactum* (GAQU) and genome data for two dermosponges (*Amphimedon queenslandica*, ACUQ; *Aplysina aerophoba*, OIVB/OIVD/OIVE/OIVF/OIVG) are available (**Supplementary Tables 1, 2**). Genome (ABGP) and transcriptome (GFSG) data for the Placozoa *Trichoplax adhaerens* and molecular data for four Ctenophora species, *Beroe ovata* (genome-UOYG), *Mnemiopsis leidyi* (genome-AGCP/TSA project-GFSG), *Pleurobrachia bachei* (genome-AVPN, *Hormiphora californensis*, GGLO) also exist and for Cnidarians an even larger dataset is available including transcriptome and genome assemblies for *Hydra vulgaris* (genome – ACZU/ABRM; TSA projects-GEVZ/GANC/GAOL/GHHG/HAAC/HAAD/GGKF/ GGKH) (Chapman et al., 2010). Searches for a potential non-bilaterian PACAP in the present study using the mature sequence of the bilaterian PACAP peptides and cnidarian (peptide and nucleotide sequence) homologs failed to identify a putative transcript or gene.

## PACAP PRECURSOR IN PROTOSTOMES

The protostomes are the most diverse group of animals and they diverged from the deuterostome lineage approximately 600 million years ago prior to the Cambrian period (Ayala et al., 1998). Two major sister monophyletic protostomian clades that diverged early in evolution exist: (1) the Ecdysozoans and (2) the Lophotrochozoans (Erwin et al., 2011). Studies directed at identifying a homolog of the vertebrate PACAP system are available for both clades.

### Ecdysozoans

The Ecdysozoa are the largest superphylum of the animal kingdom and include all the arthropods (insects, spiders, and crustaceans), the most diverse and specious animal phyla, the nematodes and several other smaller phyla. This is a morphologically heterogeneous group and includes animals that have a cuticle and grow by molting and over a million species have been described (Telford et al., 2008). Despite their large biodiversity, the basic body plan of Ecdysozoans has been largely conserved and they are either insect-like with a segmented body and jointed appendages or worm-like with an anterior circum-oesophageal nerve ring and a terminal mouth usually found on an introvert (Telford et al., 2008). Ecdysozoans play a central role in the understanding of invertebrate physiology and the nematode *Caenorhabditis elegans* (Consortium, 1998) and the

fruit fly *Drosophila melanogaster* genomes (Adams et al., 2000) were the first published animal genomes. In the nematodes there are no reports about the isolation and expression of a PACAP-like system but homologs of the vertebrate VIP and PHI peptides have been detected by dot blot analysis in the excretions/secretions of three parasitic nematodes (*Ascaridia galli*, *Nematodirus battus*, *Nippostrongylus brasiliensis*) (**Figure 4**) (Foster and Lee, 1996).

In arthropods a PACAP system similar to the vertebrates has been proposed. Antibodies raised against human PACAP-38 were used to detect a peptide homolog in the *Drosophila* central and peripheral nervous systems (Zhong and Peña, 1995). Exposure of larval muscle to the human PACAP-38 peptide modified calcium ion transport (Bhattacharya et al., 2004). Nonetheless, no peptide, transcript or gene for the fruit-fly PACAP or its receptor has been identified, although Western Blot analysis with heterologous antisera led to the suggestion that the putative insect peptide (5.4 kDa) had a similar size to the mammalian homolog (4.5 kDa) (Zhong and Peña, 1995). Subsequently, in *Drosophila* amnesiac peptide was proposed to be the functional homolog of vertebrate PACAP. Sequence similarity between amnesiac and vertebrate PACAP is low but they are proposed to share conserved functions in learning and memory (Feany and Quinn, 1995; Hashimoto et al., 2002). In the insect the American cockroach (*Periplaneta americana*) and in two crustaceans, the crab (*Eriocheir japonica*) and the white shrimp (*Litopenaeus vannamei*), partial PACAP cDNAs have been isolated and the deduced mature peptides are highly identical to the vertebrate peptide (**Figures 2**, **4**) (Lugo et al., 2013; Pirger et al., 2016). No functional studies of PACAP in the cockroach or crab have been described although in the shrimp, innate immunity is boosted (e.g., increased hemocyte number, superoxide dismutase activity, etc.) after bacterial infection in specimens given catfish (*Clarias gariepinus*) recombinant PACAP (Lugo et al., 2013). In arthropods, homologs of other SCT-peptide family members are also suggested to exist and the insect AdipoKinetic Hormones (AKH) are the sequence and function homologs of mammalian GCG and they are also involved in the regulation of food metabolism (sugar homeostasis and mobilization of sugars and lipids from the fat body) (Clynen et al., 2004). In the brain and in the gut muscle layer of the American cockroach immunoreactivity for VIP and PHI has also been detected with heterologous antisera (**Figure 4**) (Fujita et al., 1981; Iwanaga et al., 1981; Kuramoto et al., 1985).

The Ecdysozoans are the protostome subphylum where the greatest amount of molecular data exists (**Supplementary Tables 1, 2**). Currently transcriptome assembly data (TSA) for 35 nematodes (free-living and parasitic) and 1410 arthropods (136 Chelicerata, 170 Crustacea, and 1104 Insecta) have been deposited in NCBI (**Supplementary Table 1**) and were searched in the present study. Whole genome assemblies (wgs) are also available for several representatives of the Nematoda (taxid: 6231) and Arthropoda (taxid: 6656) phyla (**Supplementary Table 2**). The transcriptome (midgut, GEIF; CNS, GFCQ; whole body, GAWS; testis, GBJC) and genome (PGRX) is available for the American cockroach (*Periplaneta americana*) as is the genome (LQIF) and transcriptome [from precocious and normal juvenile stages, GEFT; eyestalk, Y-organ, hepatopancreas (HAAX, GBZW), fertilized eggs and larvae, GGQO] of a crustacean, the crab (*Eriocheir sinensis*).

Searches in wgs and transcriptomes in the present study failed to yield a homolog gene or transcript of the mature PACAP in representatives of the phylum Arthropoda, subphylum Hexapoda (taxid: 6960, which includes the Insecta), Crustacea (taxid: 6657), Chelicerata (taxid: 6843, which includes the Arachnida) or in the phylum Nematoda (taxid: 6231). In a few species,

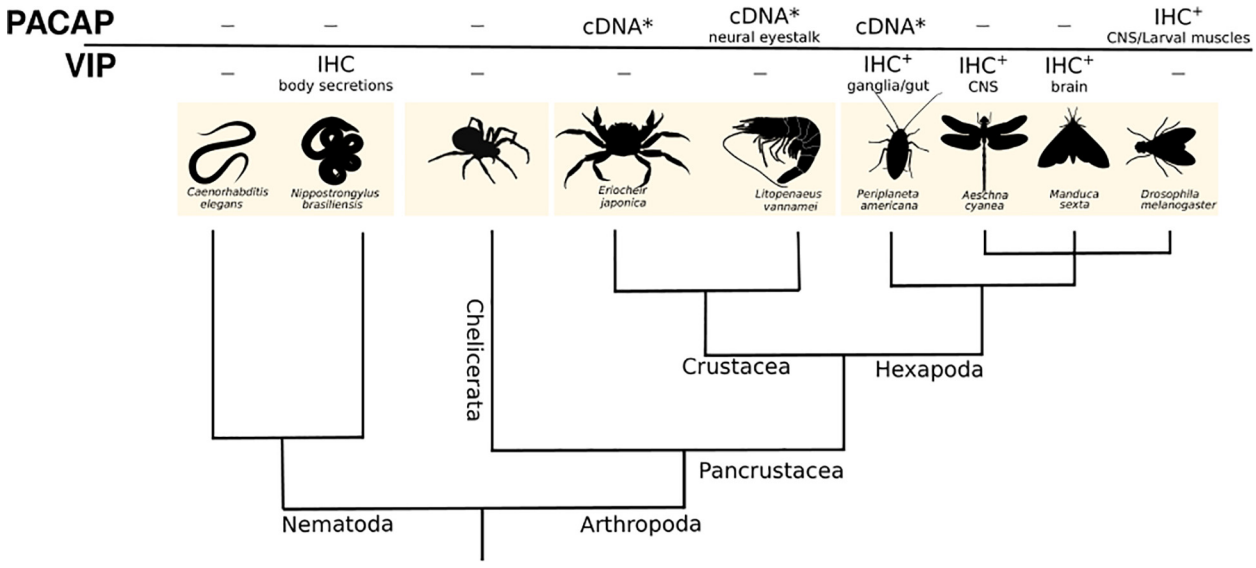


**FIGURE 4 |** A dendrogram representing the phylogenetic relationships of the main Ecdysozoan phyla. Molecular and expression data (IHC) available for putative PACAP and VIP is presented (El-Salhy et al., 1983; Andries et al., 1984; Andriès et al., 1991). The tissues where putative PACAP or VIP were found are also indicated. - no data available; * partial cDNA.

short positive sequence matches for PACAP were found as part of genes predicted to encode much larger non-PACAP proteins. Interrogation of vertebrate databases with the identified arthropod genes failed to retrieve the PACAP gene or any other homologs of the SCT superfamily suggesting the isolated gene fragments are unlikely to be authentic Ecdysozoan genes for PACAP or other SCT family members.

## Lophotrochozoans

The superphylum Lophotrochozoa are the largest group of marine invertebrates and include the greatest number of animal phyla such as molluscs (the second most diverse specious group after the arthropods), annelids and flatworms amongst others (**Figure 5**) (Kocot, 2016). Far fewer Lophotrochozoan genomes are currently available compared to their sister protostome group, the Ecdysozoans. Nonetheless, comparative genome analysis has revealed that unlike the Ecdysozoans the Lophotrochozoans possess a more similar gene complement and genome organization to deuterostomes (Raible et al., 2005; Miller and Ball, 2009; Takahashi et al., 2009; Simakov et al., 2013). For this reason, it is considered that Lophotrochozoan can contribute to understanding metazoan genome and gene family evolution by creating a link between Ecdysozoans and deuterostomes.

Evidence of a PACAP-like system similar to what exists in vertebrates has been described in lophotrochozoans. In the planarian (*Dugesia japonica*) and in the mollusc squid (*Sepioteuthis lessoniana*) partial cDNAs encoding PACAP-like peptides have been isolated and deposited in the NCBI database and the deduced mature peptides are 92% identical in amino acid sequence to the human peptide (**Figure 2**) (Cardoso et al., 2007a, 2010; Pirger et al., 2016). In annelids (oligochaeta) and in two gastropod molluscs, the garden snail (*Helix pomatia*) and the pond snail (*Lymnaea stagnalis*) no PACAP precursor has been isolated but results from a series of expression studies have led to suggestions that an active PACAP-like peptide and specific receptor exist (**Figure 5**) (Pirger et al., 2010a,

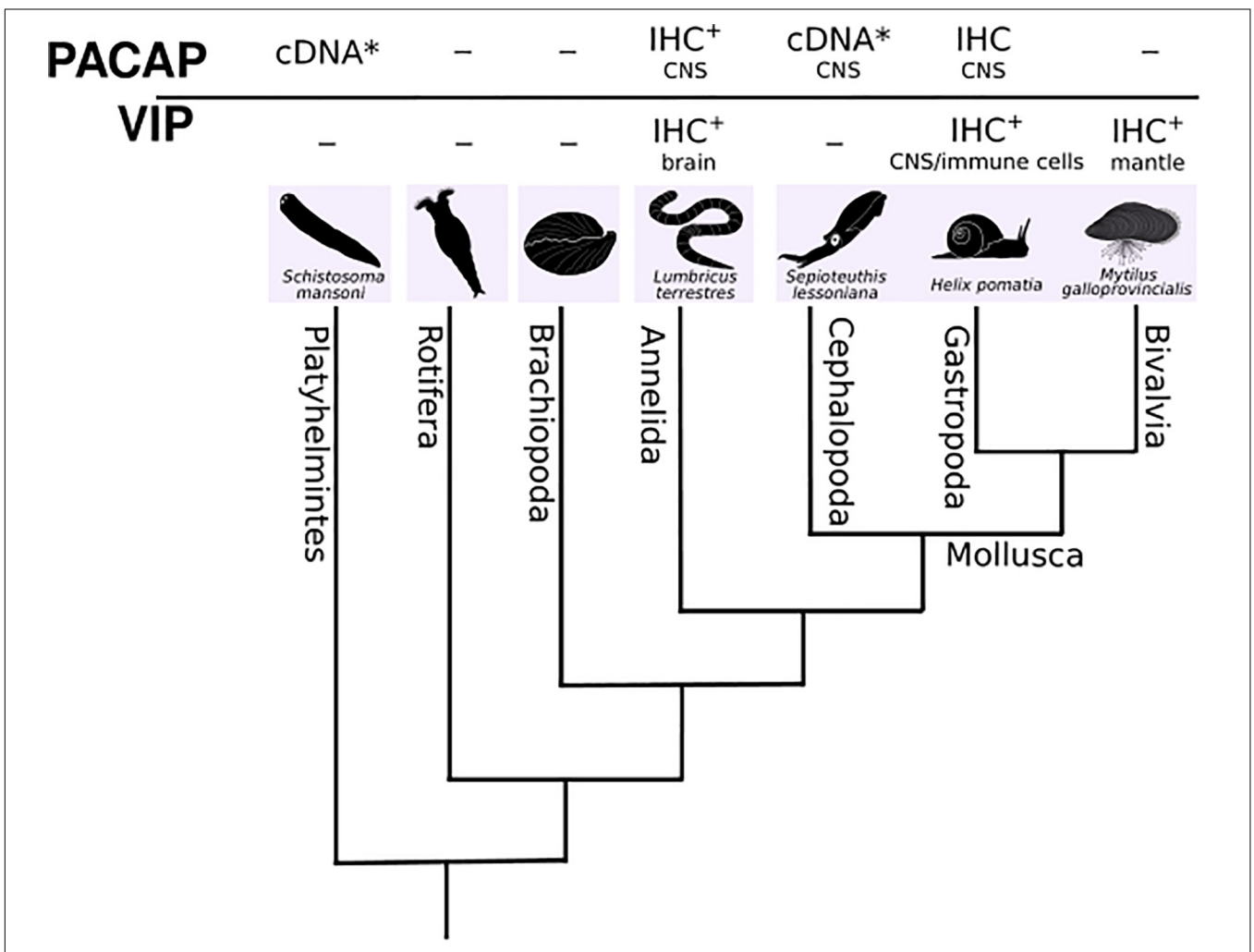


**FIGURE 5 |** A dendrogram representing the phylogenetic relationships of the main Lophotrochozoan phyla. Available molecular and expression data (IHC) for putative PACAP or VIP is presented (Reglödi et al., 2000; Somogyvári-Vigh et al., 2000; Hernádi et al., 2008; Boros et al., 2010; Pirger et al., 2010a). The tissues where putative PACAP or VIP were found are also indicated. - no data available; * partial cDNA; + heterologous antibody.

2016; Kiss and Pirger, 2013). Using antibodies specific for mammalian PACAP, positive immunoreactivity was detected in the CNS and peripheral organs of adults of three species of annelids (*Lumbricus terrestris*, *Eisenia fetida*, *Lumbricus polyphemus*) and during the embryonic development of the earthworm, *Eisenia fetida* (Reglödi et al., 2000; Boros et al., 2010). PACAP immunoreactivity was also detected in the cerebral ganglia and lip sensory epithelium of the pond snail (Pirger et al., 2010a). In the garden snail, radioimmunoassay using antisera raised against mammalian PACAPs, revealed PACAP-27 and 38 in nervous tissue and peripheral organs and peptide abundance was associated with increased activity of the animals (Hernádi et al., 2008). In both the garden snail and pond snail, PACAP-like peptide fragments were isolated from the brain and two peptide isoforms similar to the vertebrate PACAP-27 and PACAP-38 are proposed to arise from a putative gastropod PACAP gene (Hernádi et al., 2008; Pirger et al., 2010c). Stimulation by mammalian PACAP of cAMP production by pond snail cerebral ganglia homogenates has been proposed to support the existence of a functional PACAP receptor in gastropods, although it has not yet been isolated (Pirger et al., 2010a). In common with mammals, PACAP in snails is proposed to regulate cell proliferation and differentiation and be a neuroendocrine regulator in associative memory (Reglödi et al., 2000; Pirger et al., 2010b, 2016; Kiss and Pirger, 2013; Krajcs et al., 2015) and in gastropods and earthworm PACAP-27 is proposed to be the most abundant form in the CNS (Reglödi et al., 2000; Somogyvári-Vigh et al., 2000; Boros et al., 2008; Hernádi et al., 2008; Pirger et al., 2010a).

Other homologs of the vertebrate SCT peptide family have also been detected in the lophotrochozoans (annelids, molluscs, and platyhelminths) by immunohistochemistry using heterologous antisera (**Figure 5**) (Cardoso et al., 2010). In annelids, a sister clade of molluscs, immunoreactive PACAP-like peptides and receptors were identified in the CNS and peripheral nervous system (PNS) (Reglödi et al., 2000; Molnar et al., 2006; Boros et al., 2008, 2010; Varhalmi et al., 2008). Similarly, in the bivalve mollusc the Mediterranean mussel (*Mytilus galloprovincialis*), immunoreactivity for VIP was detected in the mantle and in gastropod molluscs, VIP-like molecules were detected in the nervous system of the sea hare (*Aplysia kurodai*) and land snail (*Helix pomatia*) and in innate immune cells of two freshwater snails (*Planorbarius corneus* and *Viviparus ater*) (Kuramoto et al., 1985; Ottavianil and Cossarizza, 1990; Ottaviani et al., 1992; Kaufmann et al., 1995; Licata et al., 2003). In the nervous system of the pond snail (*Lymnaea stagnalis*) two VIP immunoreactive neurons were detected (Schot et al., 1981). Immunoreactive GCG/GLP and SCT was also detected in immune cells of the two freshwater snails (Ottavianil and Cossarizza, 1990; Ottaviani et al., 1992). In annelids, VIP-like positive cells were detected in the CNS of the leech (*Hirudo medicinalis*), earthworm (*Lumbricus terrestres*), oligochaete (*Nereis diversicolor*) (Sundler et al., 1977; Osborne et al., 1982), and planarian (*Schistosoma mansoni*) (Gustafsson, 1987).

For the major lophotrochozoan phyla assembled tissue transcriptomes (TSA) and genomes (wgs, Annelida, taxid: 6340; Mollusca, taxid: 6447 (Bivalvia, taxid: 6544; Gastropoda, taxid: 6448; Cephalopoda, taxid: 6605); Rotifera, taxid: 10190; Brachiopoda, taxid: 7568; Platyhelminthes, taxid: 6157) are available (**Supplementary Tables 1, 2**). These include a transcriptome and genome for the planarian *Dugesia japonica* (transcriptomes – GFJY, GALW, IAAB, genome- MQRL) and a transcriptome of the sucker ring tissue of the squid *Sepioteuthis lessoniana* (transcriptome – GBGT). Assembled transcriptomes of 21 Annelids, 10 Brachiopods, 41 Platyhelminthes, 9 Rotifers, and 138 molluscs (68 gastropods, 44 bivalves, 26 cephalopods) are also available.

Searches in wgs and transcriptomes of lophotrochozoa in the present study using the mature conserved metazoan and the cephalochordate PACAPs as bait failed to retrieve sequence matches in the majority of the transcriptomes and genomes analyzed, including the planarian and squid, which were previously proposed to possess PACAP. The short sequence hits identified were further analyzed by using them to search the human genome. However, they failed to retrieve PACAP or any other SCT family member. In summary, searches in lophotrochozoan transcriptomes and genomes failed to identify transcripts or genes that shared high sequence identity with human PACAP or other members of the SCT superfamily.

## PACAP PRECURSOR IN INVERTEBRATE DEUTEROSTOMES

The invertebrate deuterostomes include the echinoderms, hemichordates, and chordates (urochordates and cephalochordates) and they are all marine animals (**Figure 6**). Of all the invertebrates, the invertebrate deuterostomes genomes are proposed to be most like the vertebrate genomes. Furthermore, since the invertebrate deuterostomes did not experience a genome tetraploidization they possess a single copy of the gene homologs present as multiple gene copies in vertebrates (Nakatani et al., 2007; Putnam et al., 2008). The invertebrate deuterostomes are regarded as an important link between the protostome–deuterostome ancestor and vertebrates and can provide relevant insight into vertebrate gene family origin and evolution. PACAP precursors have been isolated in tunicates and cephalochordates and the tunicate deduced mature peptides are highly similar to the vertebrate peptides while the cephalochordate peptides are poorly conserved (**Figure 2**). Other members of the SCT family of peptides have been identified by immunohistochemistry (IHC) in the cerebral ganglion and digestive system of two tunicates (*Ciona intestinalis* and *Styela plicata*) (Pestarino, 1990) and in the digestive tract of a cephalochordate, the common lancelet (*Branchiostomata lanceolatum*) (Reinecke, 1981) (**Figure 7**). This suggests that a similar gene repertoire to the human SCT family is present in invertebrate deuterostomes and this supports the notion that the gene family may have emerged just prior to the vertebrate radiation.

Two *pacap* transcripts from the disk-top tunicate (*Chelyosoma productum*) encode two PACAP-27 mature peptides, that share 96 and 85% aa sequence identity with human PACAP

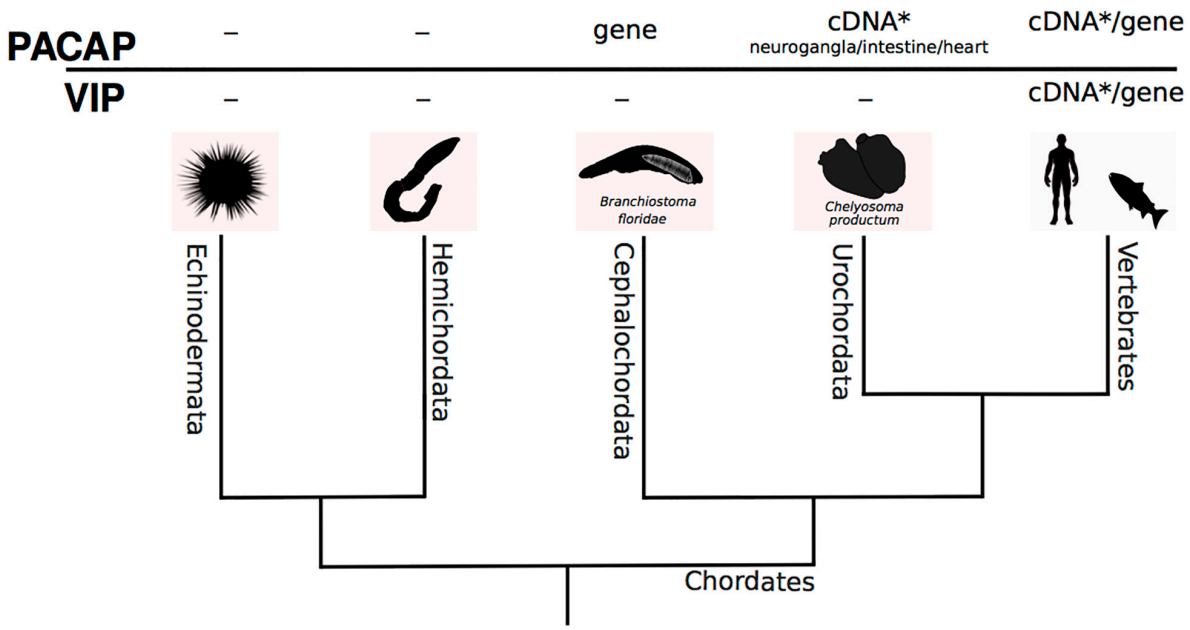


**FIGURE 6 |** A dendrogram representing the phylogenetic relationship of the invertebrate deuterostome phyla. Available molecular and expression data (PCR and IHC) for putative PACAP and VIP is presented (McRory and Sherwood, 1997). The tissues where putative PACAP or VIP were found are indicated. - no data available; * full-length cDNA.

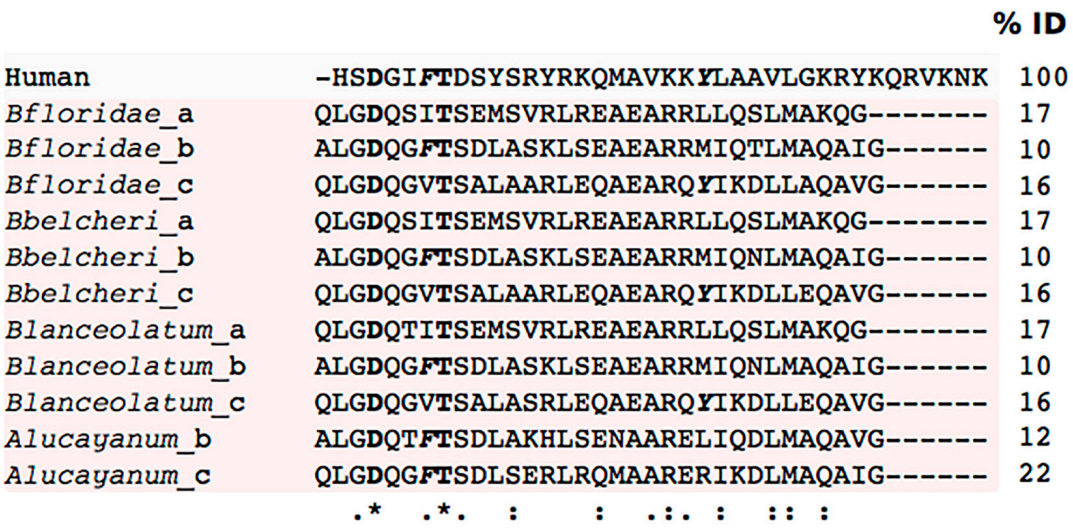


**FIGURE 7 |** Multiple sequence alignment of the cephalochordate PACAP-like peptides. Comparison of the human mature peptide sequence with the predicted cephalochordate PACAP-like peptides. The percentage of identity (%ID) with the human mature PACAP-38 peptide is given. Complete residue conservation is annotated with a "*" and highlighted in bold, partially conserved residues are denoted by "." and the position of the consensus amino acids present in the greatest number of sequences are indicated with ":". Residues in the vertebrate peptides important for receptor activation are annotated in bold and italics.

(McRory and Sherwood, 1997). Tunicate *pacap1* is highly expressed in the neuro ganglia and *pacap2* has a widespread tissue distribution and is also present in the gonad/digestive gland and intestine (McRory and Sherwood, 1997). In the cephalochordate amphioxus (*Branchiostoma floridae*) a single PACAP-like gene and a functional receptor have been characterized (Mirabeau and Joly, 2013; On et al., 2015). The organization of the cephalochordate peptide precursor differs from that of the tunicate and it produces three PACAP-like peptides, bfPACAP/GCGa, bfPACAP/GCGb, bfPACAP/GCGc (32 to 33 aa in length) that share low sequence identity (10–15% aa) with human PACAP and are proposed to have co-evolved with the GCG-peptides (Mirabeau and Joly, 2013). The cephalochordate PACAP/GCG receptor (bf95) also shares

relatively low sequence similarity with the human $PAC_1$ (On et al., 2015).

Transcriptome (TSA) data available for four tunicates (*Ciona intestinalis*, GBKV; *Ciona savignyi*, GGEI; *Oikopleura* dioica, GCJN and *Salpa thompsoni*, GFCC), two cephalochordates (*Asymmetron lucayanum*, GESY/GETC and *Branchiostoma floridae*, GESZ/GETA/GAMX), a hemichordate (*Ptychodera flava*, GDGM) and 40 echinoderms was investigated in the present study (**Supplementary Table 1**). Searches in tunicate and hemichordate transcriptomes for PACAP using as the bait the protostome and deuterostome mature PACAPs failed to identify homologs in tunicates, hemichordates, or echinoderms. PACAP homolog sequences were found in searches of cephalochordate transcriptomes. In the Bahama lancelet (*Asymmetron lucayanum*) four potential PACAP homologs with an identical sequence (GESY01044927.1, GESY01044926.1 GESY01044925.1 GESY01044923.1) were retrieved. Putative PACAP homolog sequences were also retrieved from the Florida lancelet (*Branchiostoma floridae*). Further analysis of the isolated transcripts revealed that in both species they correspond to an alternative splice form of the previously published PACAP/GCG transcript (XP_002608413.1). The transcripts identified in the present study lacked the putative bfPACAP/GCGa peptide predicted to occur at the N-terminus of the protein precursor (On et al., 2015). The Florida lancelet transcripts are 100% identical to the PACAP/GCGb and PACAP/GCGc previously described (Mirabeau and Joly, 2013; On et al., 2015) and share 72 and 66% sequence identity, respectively with the predicted PACAP peptides in the Bahama lancelet (**Figure 7**).

To identify the putative origin of the basal deuterostome PACAP, whole genome assemblies (wgs) for tunicates, hemichordates, echinoderms, and cephalochordates were screened (**Supplementary Table 2**). In Tunicata (taxid: 7712), Hemichordata (taxid: 10219), and Echinodermata (taxid: 7586) genomes no gene homolog of human PACAP was identified. In cephalochordates (Cephalochordata, taxid: 7735) the PACAP/GCG-like gene was retrieved from the Florida lancelet, Belcher's lancelet (*Branchiostoma belcheri,* three genome assemblies), common lancelet (*Branchiostoma lanceolatum)* and Bahama lancelet. The deduced cephalochordate peptides shared very low sequence identity with human PACAP and considering all isoforms (a-c) only two amino acid residues were conserved (**Figure 7**). However, amino acid residues important for receptor binding in human PACAP, $Phe^6$ and $Tyr^{22}$ (Sun et al., 2007; Bourgault et al., 2009; Dejda et al., 2011) were conserved in the cephalochordate PACAP/GCG_b and PACAP/GCG_c peptides, respectively (**Figure 7**). Overall our searches in invertebrate deuterostomes

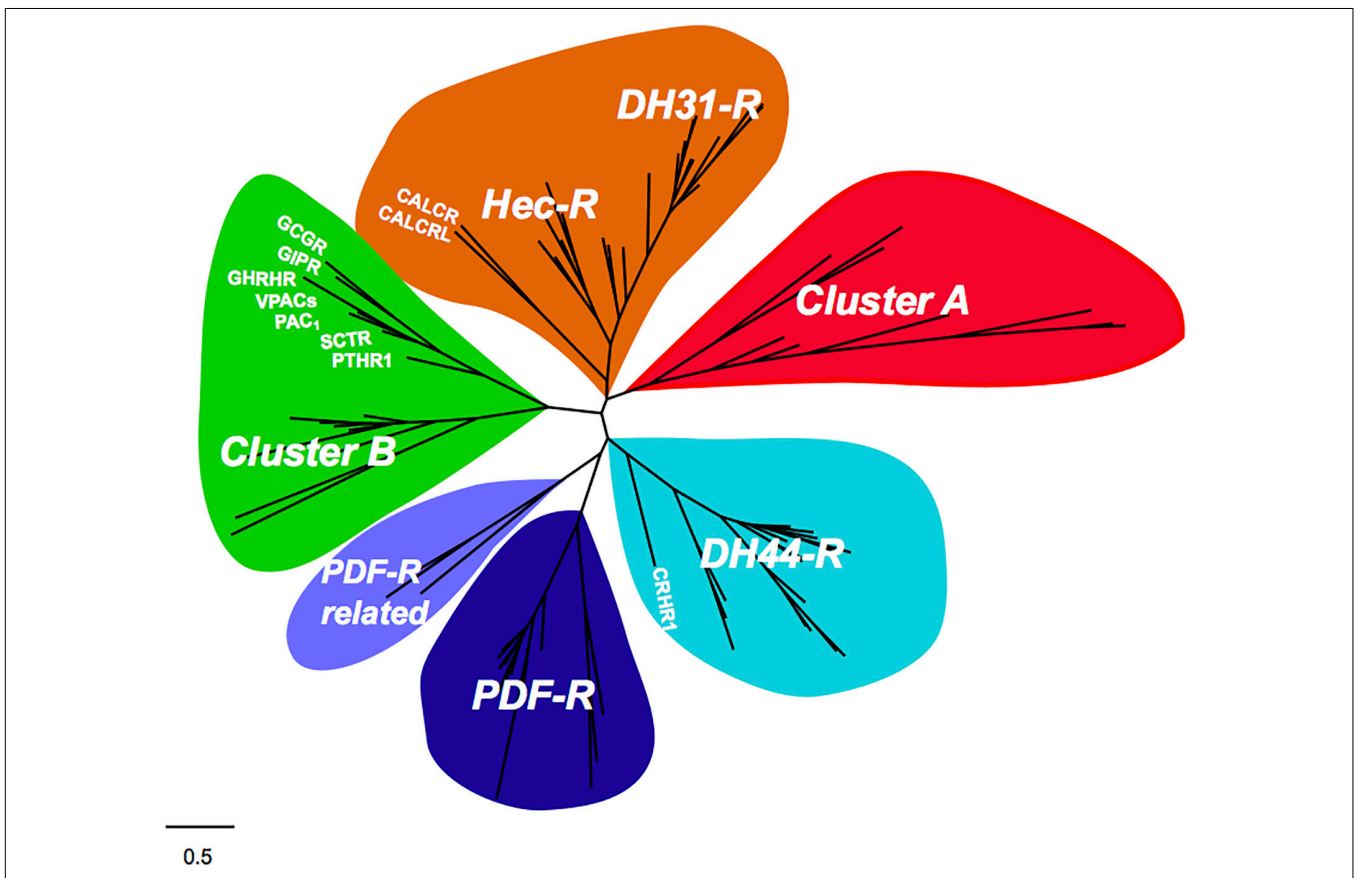


**FIGURE 8 |** Evolutionary relationships of the nematode and arthropod family B1 GPCRs with the human homologs. The six nematode and arthropod receptor subfamilies are represented in different colors and the human gene homologs are indicated. The phylogenetic tree was modified from Cardoso et al. (2014).

failed to identify sequence homologs of either the human or previously reported tunicate PACAP (McRory and Sherwood, 1997). However, transcripts and genes for cephalochordate PACAP/GCG-like peptides were identified, suggesting that a PACAP-like gene emerged in the lineage giving origin to cephalochordates.

## PACAP RECEPTORS IN INVERTEBRATES

Searches for SCT family peptides in invertebrates has frequently been accompanied by searches for the cognate receptors as additional proof that the system exists. In vertebrates' receptors that are activated by PACAP are member of family B1 GPCRs. To date the only invertebrate PACAP/GCG-like receptor (bf95) isolated and functionally characterized is from the cephalochordate, amphioxus (*Branchiostoma floridae*) but this receptor shares poor sequence similarity (37% aa identity) with the vertebrate $PAC_1$ but when it is activated it triggers intracellular signaling processes similar to the mammalian homolog (On et al., 2015). To provide evidence supporting the existence of a non-vertebrate homolog of vertebrate $PAC_1$ and to further understand how the peptide-receptor system emerged we searched for putative receptor sequence homologs in representative species of major invertebrate phyla and compared receptor evolution to that of other family B1 GPCRs.

In protostomes, a PACAP-like receptor similar to that in vertebrates was previously predicted in an annelid, *Eisenia fetida* and in the gastropods, *Lymnaea stagnalis* and *Helix pomatia* based on protein detection with heterologous antisera raised against the homolog mammalian receptor. In the earthworm (*Eisenia fetida*) $PAC_1$-like immunoreactivity was detected in adult CNS and embryos and the protein was estimated to be 50 kDa and to have a similar organization to the vertebrate homolog (Boros et al., 2010). In the gastropod, abundant $PAC_1$-like immunoreactivity was found in both the CNS and peripheral nervous system of the snail *Helix pomatia* and human PACAP-27 and PACAP-38 shown to elicit a response in neurons expressing the receptors (Hernádi et al., 2008). In the nervous system of the snail, *Lymnaea stagnalis*, human PACAP-38 increased cAMP levels (Pirger et al., 2010b). In insects a PACAP receptor has not been isolated although there are studies that suggest a functional receptor may exist.

A study aimed at characterizing G-protein-coupled neurotransmission using an insect "learning model" rutabaga-type *Drosophila* mutants that lack the type I Ca(2+)/CaM-dependent adenylyl cyclase (AC) gene revealed that vertebrate PACAP-38 stimulates synaptic currents through the coactivation of the Ras/Raf and Rutabaga-adenylyl cyclase

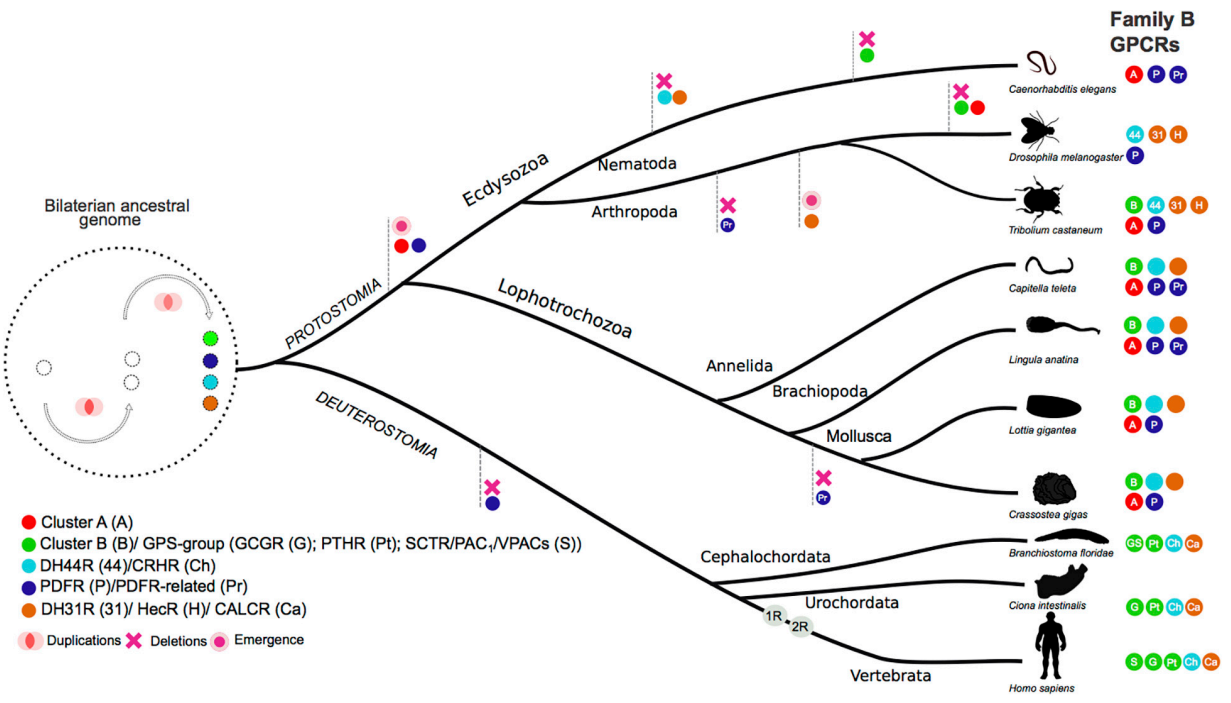


**FIGURE 9 |** Proposed evolutionary model for the Ecdysozoans and Lophotrochozoan family B1 GPCRs. The main metazoan receptor subfamily gene clusters are represented by full colored circles according to their proposed common origin in the bilaterian ancestral genome. Four genes precursors for family B GPCR subfamilies arouse from gene duplication events in the bilaterian ancestral genome. These genes subsequently evolved under distinct evolutionary pressures in the protostome and deuterostome lineages (for more details see Cardoso et al., 2014). Species-specific gene duplications/deletions within each receptor family are not represented. The two rounds of genome duplication (1R and 2R) in the deuterostome radiation are represented. The phylogeny of Cluster B is represented in **Supplementary Figure 1**. The rest of the data was obtained from Cardoso et al. (2006; 2010, unpublished; On et al., 2015). The figure was adapted from Cardoso et al. (2014) and is not drawn to scale.

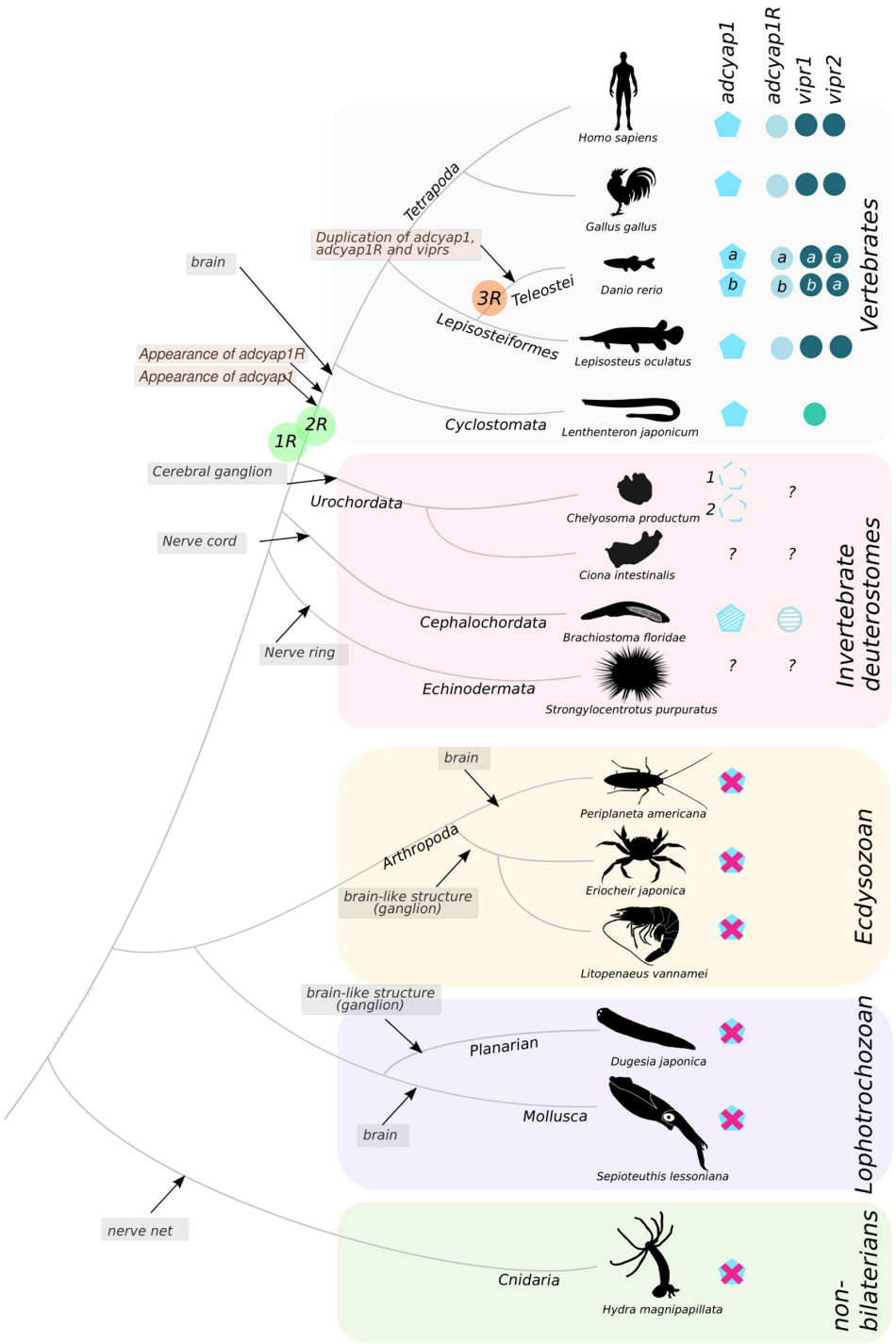


**FIGURE 10 |** Molecular data for PACAP and PACAP receptors in metazoan. The peptide is represented by pentagons and the receptors by circles. The major events associated with gene family evolution are indicated. In lamprey, a *adcyap1* peptide gene precursor and *adcyap1R1/vipr1* receptor gene were isolated (Ng et al., 2012). In the tunicate *Chelyosoma productum* two PACAP precursors (*pacap1* and *pacap2*) were isolated but their existence remains to be confirmed, as do the PACAP precursors in the related species, *Ciona intestinalis* and in the scheme are represented by a dashed lined pentagon (McRory and Sherwood, 1997; Cardoso et al., 2010). The cephalochordate PACAP-like peptides and receptors are represented by striped pentagons and circles, respectively, as they are hybrids of the vertebrate PACAP/GCG peptide and receptor system (On et al., 2015). The genome duplication events that occurred earlier during the vertebrate radiation (1R, 2R) and the teleost specific (3R) are also annotated. The teleost peptide and receptor genes that resulted from 3R are represented by a and b. The major evolutionary events that explain existing molecular data are mapped with boxes with a flesh colored background. The main events associated with the appearance of the metazoan nervous system are also represented. The evolutionary relationship between the species represented was based on (Brunet and King, 2017). A cross (X) means not found in molecular databases (TSA and WGS) and likely to be absent in the species represented. ?- unknown existence.

pathways (Zhong, 1995). Furthermore, human PACAP activates the receptor for insect PDF when co-expressed with Neurofibromatosis 1 (NF1) protein that potentiates PDF action by coupling to AC (Mertens et al., 2005). PDFR and $PAC_1$ are both members of family B1 GPCRs and share some structural resemblance although PDFR is exclusively found in invertebrates (Cardoso et al., 2014). The insect *maxadilan* peptide, which is abundant in the saliva of the sand fly (*Lutzomyia longipalpis*), the vector of leishmaniasis, activates human $PAC_1$. Despite low sequence similarity *maxadilan* shares similar functions to PACAP and in vertebrates it elicits vasodilation and modifies the secretion of pro-inflammatory cytokines by macrophages (Bozza et al., 1998; Soares et al., 1998). This suggests that in insects there is promiscuity between peptide ligand–receptor pairs and that this plasticity has been used by organisms to advantageously modulate host physiology via the ancient GPCR system. This idea is reinforced by the recent demonstration that human and fish SCT family peptides can modulate the physiology of the mosquito vector of malaria when provided to animals in an artificial meal (Marques et al., 2018). Furthermore, exposure to human GLP 2 peptide (GCG-peptide member) significantly increased vitellogenin expression, mosquito egg production and offspring fitness although if this was due to the activation of mosquito GPCRs was not established (Marques et al., 2018).

In arthropods, DH31R and DH44R are the sequence homologs of the vertebrate CALCR and CRHRs but invertebrate genomes contain a larger family B1 GPCR receptor gene repertoire most of which are orphans. Recently searches in nematode and arthropod genomes revealed that six main B1 GPCRs subfamilies exist and that they evolved under lineage and species-specific pressure (Cardoso et al., 2014) (**Figure 8**). In addition to the DH31R and DH44R subfamilies the nematode and arthropod genomes also possess receptors for the peptide PDF (PDFR) and for the recently identified PDFR-related, Cluster A and Cluster B (Cardoso et al., 2014). The PDFR, the PDFR-related and Cluster A have no homolog genes in vertebrates but Cluster B receptors are considered to be the orthologous of vertebrate $PAC_1$/$VPAC_1$, GCGR, and PTHR with which they probably share a common evolutionary origin (Cardoso et al., 2014) (**Figure 9**). Receptor members of Cluster B, which are most similar to the vertebrate $PAC_1$, have been lost from the genomes of Diptera (*Drosophila melanogaster* and mosquitoes) and from the nematode, *C. elegans,* although the receptor genes persisted in other arthropods as well as in the genome of a parasitic nematode suggesting that they evolved under distinct pressures potentially driven by specific chromosome rearrangements (Cardoso et al., 2014). Considering the nematode and arthropod receptors it has been hypothesized that at least four ancestral family B1 GPCR genes emerged early in the metazoan radiation and underwent distinct evolutionary trajectories after the protostome-deuterostome split (**Figure 9**).

Nonetheless, in other protostomes, family B1 GPCR genes are mostly unknown in lophotrochozoan which are suggested to have a more similar gene repertoire to vertebrates than arthropods and nematodes (Simakov et al., 2013; Cardoso et al., 2016). Searches for putative sequences related to the vertebrate $PAC_1$ in Lophotrochozoan genomes (Molluscs, Annelids, Brachiopod) only retrieved members of the Cluster B subfamily and no direct sequence homologs of vertebrate $PAC_1$ were found (**Figure 9** and **Supplementary Figure 1**). Our phylogenetic tree topology confirms that the protostome receptors of Cluster B are the most similar in sequence to the vertebrate GCG/PTH/PACAP receptors. In addition, homologs of the previously identified nematode and arthropod B1 GPCR subfamilies were found and the previously proposed evolutionary model was confirmed. The cephalochordate PACAP/GCGR-like gene diverged prior to the vertebrate *ADCYAP1R1* and *GCGR* genes confirming its identity but no gene homolog was identified in tunicate genomes. In contrast, in the *Ciona* genome two sequence homologs of the vertebrate GCGR-subfamily were retrieved, but a putative PACAP receptor gene was absent (**Supplementary Figure 1**).

## CONCLUSION

No evidence for a highly conserved PACAP system or any other member of the SCT superfamily outside the vertebrate clade was found in our study. Molecular searches of numerous representatives of major non-vertebrate phyla failed to identify a putative gene or transcript of the PACAP peptide precursors previously reported in the tunicate (*Chelyosoma productum* and *Halocynthia roretzi*), crab (*Eriocheir japonica*), shrimp (*Litopenaeus vannamei*), cockroach (*Periplaneta americana*), squid (*Sepioteuthis lessoniana*), planarian (*Dugesia japonica*), or cnidarian (*Hydra vulgaris*) (**Figure 10**). Our searches only revealed genes and transcript homologs of the cephalochordate (*Branchiostoma floridae*) PACAP/GCG-like precursor in other related cephalochordate species. Similarly, in protostomes only Cluster B receptor genes were identified and phylogenetic analysis indicates they probably shared a common ancestral origin with the vertebrate *ADCYAP1R1* gene but also with *GCGR* and *PTHR* subfamily genes.

The origin of the previously reported protostome cDNAs encoding a peptide highly similar to vertebrate PACAP is difficult to explain. The peptide and nucleotide sequences of protostome PACAP overlap totally (100% aa and nucleotide identity) with PACAP-family precursors from salmoniformes (*Salmo trutta*, *Oncorhynchus nerka,* and *Oncorhynchus tshawytscha*) (**Figure 2** and **Supplementary Table 3**) suggesting that they may be artifacts. The existence of PACAP genes/transcripts in urochordates remains unresolved as sequence homologs (peptides or nucleotides) of the isolated *Chelyosoma productum* PACAP precursors (McRory and Sherwood, 1997) were not identified in other urochordates or invertebrate deuterostomes. We propose a checklist for establishing the validity of cDNA encoding PACAP or other regulatory peptides. Specifically, proof should be provided at the level of gene, protein and function for confirming the veracity of gene/cDNA/peptide identity. If a genome is available the cDNA should be mapped and the gene confirmed by PCR; codon usage bias should be considered and clustering in phylogenetic analysis is expected generally to follow accepted models for phyla relatedness; independent confirmation from a lab focusing only on non-vertebrates would be encouraged. At the level of proteins, it should be possible to

isolate PACAP-like peptides and then confirm by *de novo* peptide sequencing the identity. Finally, at the level of function vertebrate peptides may be used in invertebrates but proof from CRISPR-Cas9 or interfering RNA of activity ablation is essential along with complimentary experiments with the receptors in the case of neuropeptides. While it may not always be possible to gather strong functional proof in non-model organisms at least robust analysis at the level of nucleic acids and proteins is an important step for identification.

Receptors for family B1 GPCRs emerged much earlier than the ligands of the SCT superfamily of peptides as we identified homologs of the vertebrate CALCR and CRHR in protostomes. The Cluster B receptors are closest to the vertebrate PACAP receptors but based on the molecular evidence gathered we propose that the *ADCYAP1R1* gene only appeared during the vertebrate radiation at a similar time to the *ADCYAP1* gene (**Figure 10**). In the most ancient extant vertebrate representative, the lamprey and hagfish (cyclostomes) an *ADCYAP1* gene and two *VIPR* genes were identified but an *ADCYAP1R1* gene was absent. Nonetheless, it is not possible to rule out that it never existed in their genomes, which are highly modified due to the independent gene duplications/deletions and genome rearrangements that occurred after the gnathostome divergence (Ng et al., 2012). The results of the previous protostome IHC studies using heterologous antibodies raised against human peptides and receptors are puzzling and the positive signals obtained may be due to interactions with other related GPCRs or unrelated proteins that bear some similarity with human PACAP receptors. We propose that strict procedures should be followed for IHC with heterologous antisera to minimize cross-reactivity or low specificity interactions. This should involve, (a) the identification of the protein/peptide sequence used to raise homologous antisera and confirmation of peptide existence in the experimental model by searches against genome/transcriptome data, (b) a battery of control assays to check for antisera specificity using recombinant proteins (if they exist) or peptides used to raise the antisera, antisera pre-absorption studies, staining reactions on multiple individuals and sections to confirm if the general staining pattern is conserved and Western blots using protein extracts of the species being studied, with and without peptide pre-absorption of antisera to check if one principle protein of the predicted size is detected.

The explanation for the effects caused by exposure of protostome tissues to human PACAP peptides is uncertain but may be due to low specificity ligand–receptor interactions that have previously been reported for heterologous peptides (Marques et al., 2018). In summary, based on our in-depth molecular study we propose that the *ADCYAP1* gene appeared in vertebrates and probably shared a common origin with the cephalochordate PACAP/GCG-like gene. The ancestral PACAP/GCG-like gene probably expanded during the tetraploidization events preceding the vertebrate radiation (1R and 2R) and generated the *ADCYAP1* gene and other members of the SCT-family peptides (**Figure 10**).

Comparative endocrinology is crucial for clinical/pharmacology research and identification of homolog systems is important to understand the endocrine peptide-receptor function and regulation (Cardoso and Larhammar, 2014). In addition, it also provides insights into how function has drifted and changed during evolution and facilitates the discovery of novel ligand–receptor pairs. In the future we proposed that characterization of the Cluster B receptor–ligand pair may provide important clues about the function of the PACAP-like system in metazoans and how this is linked with the acquisition of a nervous system and neuropeptide signaling. Given the multifunctional role of PACAP in vertebrates, characterization of such molecules would provide novel insights into the regulatory role of the PACAP-system and studies of much broader scope particularly in under represented phyla would contribute to the development of more robust evolutionary models to explain the emergence and persistence of PACAP and how its pleotropic role was acquired during evolution.

## DATA AVAILABILITY STATEMENT

The datasets analyzed in this study can be found at https://www.ncbi.nlm.nih.gov. All TSA and WGS databases enquired (accession numbers and bioprojects) are listed in **Supplementary Tables 1** (TSA) and **2** (WGS). Accession numbers for all sequences used are cited in the paper and when not available the study where it was described is indicated in the text.

## AUTHOR CONTRIBUTIONS

DP and JC conceived and planned the study, analyzed and integrated the datasets, and wrote the manuscript. JC and MG performed the bioinformatic analysis searches and prepared the figures. All authors critically read the manuscript.

## FUNDING

This study received Portuguese national funds from FCT – Foundation for Science and Technology through project UIDB/04326/2020 and from the operational programmes CRESC Algarve 2020 and COMPETE 2020 through project EMBRC.PT ALG-01-0145-FEDER-022121. JC was supported by FCT project UIDB/04326/2020. MG was supported by the ERASMUS Plus programme.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins.2020.00366/full#supplementary-material

**FIGURE S1 |** Phylogenetic trees of the invertebrate Cluster B receptors and the deuterostomes $PAC_1$, GCGR, and PTHR members. The consensus phylogenetic trees were obtained **(A)** according to the Bayesian inference (BI) and **(B)** the maximum likelihood (ML) methods using has input an edited multiple sequence alignment of the predicted receptor protein sequences obtained using the MUSCLE algorithm (Edgar, 2004) available from Aliview platform 1.18 (Larsson, 2014). BI phylogenetic reconstruction was performed using MrBayes 3.2.6 (Ronquist et al., 2012) in the CIPRES Science Gateway V. 3.3 (http://www.phylo.org) and ML tree was performed using the PhyML3.0 program available from the

ATGC platform (http://www.atgc-montpellier.fr/phyml/). The BI and ML trees were constructed with an LG substitution model according to the Akaike information criterion (Lefort et al., 2017) and BI used 1,000,000 generation sampling probability values to support tree branching and ML statistical branch support was 100 bootstrap replicates and >50 are mapped. The cephalochordate sequences are highlighted in pink and were obtained from On et al. (2015). The Arthropod receptor sequences are highlighted in yellow and were obtained from Cardoso et al. (2014) and the Lophotrochozoan sequences (highlighted in purple) were obtained by querying the metazoan ENSEMBLE GENOMES (http://metazoa.ensembl.org/index.html) database using the human receptors. Accession numbers of the invertebrate sequences are indicated in the tree. Accession numbers of human receptors are: CALCR, NP_001733.1; CALCRL, NP_005786.1; CRHR2, NP_001874.2; CRHR1, NP_004373.2; PTH1R, NP_000307.1; PTH2R, NP_005039.1; $VPAC_1$, NP_004615.2; $PAC_1$, NP_001109.2; VPAC2, NP_003373.2; GCGR, NP_000151.1; GLP1R, NP_002053.3. The tree was rooted using the human CALC/CALCRL and CRHR1/CRHR2 branches.

**TABLE S1 |** List of non-vertebrate transcriptome databases and their tissue of origin that were searched for PACAP transcripts. Available data from NCBI, November 2019. TSA databases that were available to search at NCBI are listed. Bioproject numbers are indicate.

**TABLE S2 |** List of whole genome assemblies that were searched for non-vertebrate PACAP genes. Available data from NCBI, November 2019. Bioproject numbers are indicate.

**TABLE S3 |** List of the hydra, protostome and tunicate PACAP nucleotide top five hits against the NCBI database **(A)** and Salmoniformes (taxid: 8006) transcriptomes (TSA) **(B)**. The e-values (e value) and percent of identity (%ID) are given * Nucleotide sequence not available.

## REFERENCES


Adams, M. D., Celniker, S. E., Holt, R. A., Evans, C. A., Gocayne, J. D., Amanatides, P. G., et al. (2000). The genome sequence of *Drosophila melanogaster*. *Science* 287, 2185–2195. doi: 10.1126/science.287.5461.2185

Andries, J. C., Belemtougri, G., and Tramu, G. (1984). Immunohistochemical identification of growth hormone-releasing factor-like material in the nervous system of an insect, *Aeshna cyanea* (Odonata). *Neuropeptides*. 4, 519–528. doi: 10.1016/0143-4179(84)90095-7

Andriès, J. C., Belemtougri, G., and Tramu, G. (1991). Multiple peptide immunoreactivities in the nervous system of *Aeschna cyanea* (Insecta, Odonata) – an immunohistochemical study using antisera to cholecystokinin octapeptide, somatoliberin, vasoactive intestinal peptide, motilin and proctolin. *Histochemistry* 96, 139–148. doi: 10.1007/BF00315984

Archbold, J. K., Flanagan, J. U., Watkins, H. A., Gingell, J. J., and Hay, D. L. (2011). Structural insights into RAMP modification of secretin family G protein-coupled receptors: implications for drug development. *Trends Pharmacol. Sci.* 32, 591–600. doi: 10.1016/j.tips.2011.05.007

Arimura, A., and Shioda, S. (1995). Pituitary adenylate cyclase activating polypeptide (PACAP) and its receptors: neuroendocrine and endocrine interaction. *Front. Neuroendocrinol.* 16:1003. doi: 10.1006/frne.1995.1003

Arimura, A., Somogyvàri-Vigh, A., Miyata, A., Mizuno, K., Coy, D. H., and Kitada, C. (1991). Tissue distribution of pacap as determined by ria: highly abundant in the rat brain and testes. *Endocrinology* 129, 2787–2789. doi: 10.1210/endo-129-5-2787

Ayala, F. J., Rzhetsky, A., and Ayala, F. J. (1998). Origin of the metazoan phyla: molecular clocks confirm paleontological estimates. *Proc. Natl. Acad. Sci. U.S.A.* 95, 606–611. doi: 10.1073/pnas.95.2.606

Bayliss, W. M., and Starling, E. H. (1902). The mechanism of pancreatic secretion. *J. Physiol.* 28, 325–353. doi: 10.1113/jphysiol.1902.sp000920

Bhattacharya, A., Lakhman, S. S., and Singh, S. (2004). Modulation of L-type calcium channels in *Drosophila* via a pituitary adenylyl cyclase-activating polypeptide (PACAP)-mediated pathway. *J. Biol. Chem.* 279, 37291–37297. doi: 10.1074/jbc.M403819200

Boros, A., Reglodi, D., Herbert, Z., Kiszler, G., Nemeth, J., Lubics, A., et al. (2008). Changes in the expression of PACAP-like compounds during the embryonic development of the earthworm *Eisenia fetida*. *J. Mol. Neurosci.* 36, 157–165. doi: 10.1007/s12031-008-9102-6

Boros, Á, Somogyi, I., Engelmann, P., Lubics, A., Reglodi, D., Pollák, E., et al. (2010). Pituitary adenylate cyclase-activating polypeptide type 1 (PAC1) receptor is expressed during embryonic development of the earthworm. *Cell Tissue Res.* 339, 649–653. doi: 10.1007/s00441-009-0909-4

Bourgault, S., Vaudry, D., Ségalas-Milazzo, I., Guilhaudis, L., Couvineau, A., Laburthe, M., et al. (2009). Molecular and conformational determinants of pituitary adenylate cyclase-activating polypeptide (PACAP) for activation of the PAC1 receptor. *J. Med. Chem.* 52, 3308–3316. doi: 10.1021/jm900291j

Bozza, M., Soares, M. B. P., Bozza, P. T., Satoskar, A. R., Diacovo, T. G., Brombacher, F., et al. (1998). The PACAP-type I receptor agonist maxadilan from sand fly saliva protects mice against lethal endotoxemia by a mechanism partially dependent on IL-10. *Eur. J. Immunol.* 28, 3120–3127. doi: 10.1002/(SICI)1521-4141(199810)28:10<3120::AID-IMMU3120<3.0.CO;2-3

Brunet, T., and King, N. (2017). The origin of animal multicellularity and cell differentiation. *Dev. Cell* 43, 124–140. doi: 10.1016/j.devcel.2017.09.016

Cardoso, J. C. R., De Vet, E. C. J. M., Louro, B., Elgar, G., Clark, M. S., and Power, D. M. (2007a). Persistence of duplicated PAC1receptors in the teleost, *Sparus auratus*. *BMC Evol. Biol.* 7:221. doi: 10.1186/1471-2148-7-221

Cardoso, J. C. R., Félix, R. C., Björnmark, N., and Power, D. M. (2016). Allatostatin-type A, kisspeptin and galanin GPCRs and putative ligands as candidate regulatory factors of mantle function. *Mar. Genomics* 27, 25–35. doi: 10.1016/j.margen.2015.12.003

Cardoso, J. C. R., Félix, R. C., Martins, R. S. T., Trindade, M., Fonseca, V. G., Fuentes, J., et al. (2015). PACAP system evolution and its role in melanophore function in teleost fish skin. *Mol. Cell. Endocrinol.* 411, 130–145. doi: 10.1016/j.mce.2015.04.020

Cardoso, J. C. R., Félix, R. C., and Power, D. M. (2014). Nematode and arthropod genomes provide new insights into the evolution of class 2 B1 GPCRs. *PLoS One* 9:e92220. doi: 10.1371/journal.pone.0092220

Cardoso, J. C. R., and Larhammar, D. (2014). Comparative evolution of peptide hormone-binding GPCRs: a route to understanding functional complexity. *Gen. Comp. Endocrinol.* 209, 1–2. doi: 10.1016/j.ygcen.2014.10.001

Cardoso, J. C. R., Pinto, V. C., Vieira, F. A., Clark, M. S., and Power, D. M. (2006). Evolution of secretin family GPCR members in the metazoa. *BMC Evol. Biol.* 6:108. doi: 10.1186/1471-2148-6-108

Cardoso, J. C. R., Power, D. M., Elgar, G., and Clark, M. S. (2004). Duplicated receptors for VIP and PACAP (VPAC1R and PAC1R) in a teleost fish, *Fugu rubripes*. *J. Mol. Endocrinol.* 33, 411–428. doi: 10.1677/jme.1.01575

Cardoso, J. C. R., Vieira, F. A., Gomes, A. S., and Power, D. M. (2007b). PACAP, VIP and their receptors in the metazoa: insights about the origin and evolution of the ligand-receptor pair. *Peptides* 28, 1902–1919. doi: 10.1016/j.peptides.2007.05.016

Cardoso, J. C. R., Vieira, F. A., Gomes, A. S., and Power, D. M. (2010). The serendipitous origin of chordate secretin peptide family members. *BMC Evol. Biol.* 10:135. doi: 10.1186/1471-2148-10-135

Chapman, J. A., Kirkness, E. F., Simakov, O., Hampson, S. E., Mitros, T., Weinmaier, T., et al. (2010). The dynamic genome of Hydra. *Nature* 464, 592–596. doi: 10.1038/nature08830

Christopoulos, A., Christopoulos, G., Morfis, M., Udawela, M., Laburthe, M., Couvineau, A., et al. (2003). Novel receptor partners and function of receptor activity-modifying proteins. *J. Biol. Chem.* 278, 3293–3297. doi: 10.1074/jbc.C200629200

Clynen, E., De Loof, A., and Schoofs, L. (2004). New insights into the evolution of the GRF superfamily based on sequence similarity between the locust APRPs and human GRF. *Gen. Comp. Endocrinol.* 139, 173–178. doi: 10.1016/j.ygcen.2004.07.006

Consortium, T. C. E. S. (1998). Genome sequence of the nematode C. elegans: a platform for investigating biology. *Science* 282, 2012–2018. doi: 10.1126/science.282.5396.2012

Couvineau, A., Ceraudo, E., Tan, Y. V., Nicole, P., and Laburthe, M. (2012). The VPAC1 receptor: structure and function of a class B GPCR prototype. *Front. Endocrinol. (Lausanne).* 3:139. doi: 10.3389/fendo.2012.00139

Couvineau, A., and Laburthe, M. (2011). The family B1 GPCR: 'eins. *Curr. Drug Targets* 13, 103–115. doi: 10.2174/138945012798868434

Cox, H. M. (1992). Pituitary adenylate cyclase activating polypeptides, PACAP−27 and PACAP−38: stimulators of electrogenic ion secretion in the rat small intestine. *Br. J. Pharmacol.* 106, 498–502. doi: 10.1111/j.1476-5381.1992.tb14363.x

Dejda, A., Bourgault, S., Doan, N. D., Létourneau, M., Couvineau, A., Vaudry, H., et al. (2011). Identification by photoaffinity labeling of the extracellular N-terminal domain of PAC1 receptor as the major binding site for PACAP. *Biochimie* 93, 669–677. doi: 10.1016/j.biochi.2010.12.010

Delsuc, F., Brinkmann, H., Chourrout, D., and Philippe, H. (2006). Tunicates and not cephalochordates are the closest living relatives of vertebrates. *Nature* 439, 965–968. doi: 10.1038/nature04336

Denes, V., Geck, P., Mester, A., and Gabriel, R. (2019). Pituitary adenylate cyclase-activating polypeptide: 30 years in research spotlight and 600 million years in service. *J. Clin. Med.* 8:E1488. doi: 10.3390/jcm8091488

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340

El-Salhy, M., Falkmer, S., Kramer, K. J., and Speirs, R. D. (1983). Immunohistochemical investigations of neuropeptides in the brain, corpora cardiaca, and corpora allata of an adult lepidopteran insect, *Manduca sexta* (L). *Cell Tissue Res.* 232, 295–317. doi: 10.1007/BF00213788

Erwin, D. H., Laflamme, M., Tweedt, S. M., Sperling, E. A., Pisani, D., and Peterson, K. J. (2011). The Cambrian conundrum: early divergence and later ecological success in the early history of animals. *Science* 334, 1091–1097. doi: 10.1126/science.1206375

Feany, M. B., and Quinn, W. G. (1995). A neuropeptide gene defined by the *Drosophila* memory mutant amnesiac. *Science* 268, 869–873. doi: 10.1126/science.7754370

Foster, N., and Lee, D. L. (1996). Vasoactive intestinal polypeptide-like and peptide histidine isoleucine-like proteins excreted/secreted by *Nippostrongylus brasiliensis*, *Nematodirus battus* and *Ascaridia galli*. *Parasitology* 113, 287–292. doi: 10.1017/s0031182000082056

Fujita, T., Yui, R., Iwanaga, T., Nishitsutsuji-Uwo, J., Endo, Y., and Yanaihara, N. (1981). Evolutionary aspects of "brain-gut peptides": an immunohistochemical study. *Peptides* 2(Suppl. 2), 123–131. doi: 10.1016/0196-9781(81)90023-1

Ghatei, M. A., Takahashi, K., Suzuki, Y., Gardiner, J., Jones, P. M., and Bloom, S. R. (1993). Distribution, molecular characterization of pituitary adenylate cyclase-activating polypeptide and its precursor encoding messenger RNA in human and rat tissues. *J. Endocrinol.* 136, 159–166. doi: 10.1677/joe.0.1360159

Gustafsson, M. K. S. (1987). Immunocytochemical demonstration of neuropeptides and serotonin in the nervous system of adult *Schistosoma mansoni*. *Parasitol. Res.* 74, 168–174. doi: 10.1007/BF00536029

Harmar, A. J. (2001). Family-B G-protein-coupled receptors. *Genome Biol.* 2:reviews3013. doi: 10.1186/gb-2001-2-12-reviews3013

Harmar, A. J., Arimura, A., Gozes, I., Journot, L., Laburthe, M., Pisegna, J. R., et al. (1998). International union of pharmacology. XVIII. nomenclature of receptors for vasoactive intestinal peptide and pituitary adenylate cyclase-activating polypeptide. *Pharmacol. Rev.* 50, 265–270.

Harmar, A. J., Fahrenkrug, J., Gozes, I., Laburthe, M., May, V., Pisegna, J. R., et al. (2012). Pharmacology and functions of receptors for vasoactive intestinal peptide and pituitary adenylate cyclase-activating polypeptide: IUPHAR review 1. *Br. J. Pharmacol.* 166, 4–17. doi: 10.1111/j.1476-5381.2012.01871.x

Hashimoto, H., Shintani, N., and Baba, A. (2002). Higher brain functions of PACAP and a homologous *Drosophila* memory gene amnesiac: insights from knockouts and mutants. *Biochem. Biophys. Res. Commun.* 297, 427–432. doi: 10.1016/S0006-291X(02)02144-7

Hassenzahl, D. L., Yorgey, N. K., Keedy, M. D., Price, A. R., Hall, J. A., Myzcka, C. C., et al. (2001). Chemorepellent signaling through the PACAP/lysozyme receptor is mediated through cAMP and PKC in *Tetrahymena thermophila*. *J. Comp. Physiol. A* 187, 171–176. doi: 10.1007/s003590100185

Hernádi, L., Pirger, Z., Kiss, T., Németh, J., Mark, L., Kiss, P., et al. (2008). The presence and distribution of pituitary adenylate cyclase activating polypeptide and its receptor in the snail *Helix pomatia*. *Neuroscience* 155, 387–402. doi: 10.1016/j.neuroscience.2008.05.003

Hirabayashi, T., Nakamachi, T., and Shioda, S. (2018). Discovery of PACAP and its receptors in the brain. *J. Headache Pain* 19:28. doi: 10.1186/s10194-018-0855-1

Hwang, J. I., Moon, M. J., Park, S., Kim, D. K., Cho, E. B., Ha, N., et al. (2013). Expansion of secretin-like G protein-coupled receptors and their peptide ligands via local duplications before and after two rounds of whole-genome duplication. *Mol. Biol. Evol.* 30, 1119–1130. doi: 10.1093/molbev/mst031

Iwanaga, T., Fujita, T., Nishiitsutsuji-Uwo, J., and Endo, Y. (1981). Immunohistochemical demonstration of pp-, somatostatin-, enteroglucagon- and vip-like immunoreactivities in the cockroach midgut. *Biomed. Res.* 2, 201–207. doi: 10.2220/biomedres.2.202

Kaufmann, W., Kerschbaum, H. H., Hauser-Kronberger, C., Hacker, G. W., and Hermann, A. (1995). Distribution and seasonal variation of vasoactive intestinal (VIP)-like peptides in the nervous system of *Helix pomatia*. *Brain Res.* 695, 125–136. doi: 10.1016/0006-8993(95)00659-E

Keedy, M., Yorgey, N., Hilty, J., Price, A., Hassenzahl, D., and Kuruvilla, H. (2003). Pharmacological evidence suggests that the lysozyme/PACAP receptor of *Tetrahymena thermophila* is a polycation receptor. *Acta Protozool.* 42, 11–17.

King, N., and Rokas, A. (2017). Embracing uncertainty in reconstructing early animal evolution. *Curr. Biol.* 27, R1081–R1088. doi: 10.1016/j.cub.2017.08.054

Kiss, T., and Pirger, Z. (2013). Multifunctional role of pacap-like peptides in molluscs. *Protein Pept. Lett.* 6, 628–635. doi: 10.2174/0929866511320060003

Kocot, K. M. (2016). On 20 years of Lophotrochozoa. *Org. Divers. Evol.* 16, 329–343. doi: 10.1007/s13127-015-0261-3

Krajcs, N., Hernádi, L., Pirger, Z., Reglõdi, D., Tóth, G., and Kiss, T. (2015). PACAP modulates acetylcholine-elicited contractions at nicotinic neuromuscular contacts of the land snail. *J. Mol. Neurosci.* 57, 492–500. doi: 10.1007/s12031-015-0605-7

Kuramoto, H., Yui, R., Iwanaga, T., Fujita, T., and Yanaihara, N. (1985). PHI-like immunoreactivity in the nervous system of the cockroach (Insect) and *Aplysia* (Mollusc) with special reference to its relationship to VIP-like immunoreactivity. *Arch. Histol. Jpn.* 48, 427–433. doi: 10.1679/aohc.48.427

Laburthe, M., and Couvineau, A. (2002). Molecular pharmacology and structure of VPAC receptors for VIP and PACAP. *Regul. Pept.* 108, 165–173. doi: 10.1016/S0167-0115(02)00099-X

Langer, I. (2012). Mechanisms involved in VPAC receptors activation and regulation: lessons from pharmacological and mutagenesis studies. *Front. Endocrinol. (Lausanne).* 3:129. doi: 10.3389/fendo.2012.00129

Larsson, A. (2014). AliView: a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics* 30, 3276–3278. doi: 10.1093/bioinformatics/btu531

Lefort, V., Longueville, J. E., and Gascuel, O. (2017). SMS: smart model selection in PhyML. *Mol. Biol. Evol.* 34, 2422–2424. doi: 10.1093/molbev/msx149

Licata, A., Ainis, L., Martella, S., Ricca, M. B., Licata, P., Pergolizzi, S., et al. (2003). Immunohistochemical localization of nNOS and VIP in the mantle integument of the mussel, *Mytilus galloprovincialis*. *Acta Histochem.* 105, 143–149. doi: 10.1078/0065-1281-00700

Lucas, J., Riddle, M., Bartholomew, J., Thomas, B., Forni, J., Nickerson, L., et al. (2004). PACAP-38 signaling in *Tetrahymena thermophila* involves NO and cGMP. *Acta Protozool.* 43, 15–20.

Lugo, J. M., Carpio, Y., Morales, R., Rodríguez-Ramos, T., Ramos, L., and Estrada, M. P. (2013). First report of the pituitary adenylate cyclase activating polypeptide (PACAP) in crustaceans: conservation of its functions as growth promoting factor and immunomodulator in the white shrimp *Litopenaeus vannamei*. *Fish Shellfish Immunol.* 35, 1788–1796. doi: 10.1016/j.fsi.2013.08.028

Maasz, G., Zrinyi, Z., Reglodi, D., Petrovics, D., Rivnyak, A., Kiss, T., et al. (2017). Pituitary adenylate cyclase-activating polypeptide (PACAP) has a neuroprotective function in dopamine-based neurodegeneration in rat and snail parkinsonian models. *DMM Dis. Model. Mech.* 10, 127–139. doi: 10.1242/dmm.027185

Marques, J., Cardoso, J. C. R., Felix, R. C., Santana, R. A. G., Guerra, M. D. G. B., Power, D., et al. (2018). Fresh-blood-free diet for rearing malaria mosquito vectors. *Sci. Rep.* 8:17807. doi: 10.1038/s41598-018-35886-3

McRory, J., and Sherwood, N. M. (1997). Two protochordate genes encode pituitary adenylate cyclase-activating polypeptide and related family members. *Endocrinology* 138, 2380–2390. doi: 10.1210/endo.138.6.5167

Mertens, I., Vandingenen, A., Johnson, E. C., Shafer, O. T., Li, W., Trigg, J. S., et al. (2005). PDF receptor signaling in *Drosophila* contributes to both circadian and geotactic behaviors. *Neuron* 48, 213–219. doi: 10.1016/j.neuron.2005.09.009

Miller, D. J., and Ball, E. E. (2009). The gene complement of the ancestral bilaterian-was Urbilateria a monster? *J. Biol.* 8:89. doi: 10.1186/jbiol192

Mirabeau, O., and Joly, J. S. (2013). Molecular evolution of peptidergic signaling systems in bilaterians. *Proc. Natl. Acad. Sci. U.S.A.* 110, E2028–E2037. doi: 10.1073/pnas.1219956110

Miyata, A., Arimura, A., Dahl, R. R., Minamino, N., Uehara, A., Jiang, L., et al. (1989). Isolation of a novel 38 residue-hypothalamic polypeptide which

stimulates adenylate cyclase in pituitary cells. *Biochem. Biophys. Res. Commun.* 164, 567–574. doi: 10.1016/0006-291X(89)91757-9

Miyata, A., Jiang, L., Dahl, R. D., Kitada, C., Kubo, K., Fujino, M., et al. (1990). Isolation of a neuropeptide corresponding to the N-terminal 27 residues of the pituitary adenylate cyclase activating polypeptide with 38 residues (PACAP38). *Biochem. Biophys. Res. Commun.* 170, 643–648. doi: 10.1016/0006-291X(90)92140-U

Molnar, L., Pollak, E., Boros, A., Reglödi, D., Tamás, A., Lengvari, I., et al. (2006). Comparative anatomy of PACAP-immunoreactive structures in the ventral nerve cord ganglia of lumbricid oligochaetes. *Ann. N. Y. Acad. Sci.* 1070, 427–430. doi: 10.1196/annals.1317.056

Nakamachi, T., Tanigawa, A., Konno, N., Shioda, S., and Matsuda, K. (2019). Expression patterns of PACAP and PAC1R genes and anorexigenic action of PACAP1 and PACAP2 in zebrafish. *Front. Endocrinol. (Lausanne)* 10:227. doi: 10.3389/fendo.2019.00227

Nakatani, Y., Takeda, H., Kohara, Y., and Morishita, S. (2007). Reconstruction of the vertebrate ancestral genome reveals dynamic genome reorganization in early vertebrates. *Genome Res.* 17, 1254–1265. doi: 10.1101/gr.6316407

Ng, S. Y. L., Chow, B. K. C., Kasamatsu, J., Kasahara, M., and Lee, L. T. O. (2012). Agnathan VIP, PACAP and their receptors: ancestral origins of today's highly diversified forms. *PLoS One* 7:e44691. doi: 10.1371/journal.pone.0044691

Nordström, K. J. V., Lagerström, M. C., Wallér, L. M. J., Fredriksson, R., and Schiöth, H. B. (2009). The secretin GPCRs descended from the family of Adhesion GPCRs. *Mol. Biol. Evol.* 26, 71–84. doi: 10.1093/molbev/msn228

On, J. S. W., Duan, C., Chow, B. K. C., and Lee, L. T. O. (2015). Functional pairing of class B1 ligand-GPCR in cephalochordate provides evidence of the origin of PTH and PACAP/glucagon receptor family. *Mol. Biol. Evol.* 32, 2048–2059. doi: 10.1093/molbev/msv087

Osborne, N. N., Patel, S., and Dockray, G. (1982). Histochemistry immunohistochemical demonstration of peptides, serotonin and dopamine-fl-hydroxylase-like material in the nervous system of the leech *Hirudo medicinalis*. *Histochemistry* 75, 573–583.

Ottaviani, E., Franchini, A., and Fontanili, P. (1992). The presence of immunoreactive vertebrate bioactive peptide substances in hemocytes of the freshwater snail *Viviparus ater* (gastropoda, prosobranchia). *Cell. Mol. Neurobiol.* 12, 455–462. doi: 10.1007/BF00711546

Ottavianil, E., and Cossarizza, A. (1990). Immunocytochemical evidence of vertebrate bioactive peptide-like molecules in the immuno cell types of the freshwater snail *Planorbarius corneus* (L.) (Gastropoda, Pulmonata). *FEBS Lett.* 267, 250–252. doi: 10.1016/0014-5793(90)80937-e

Pestarino, M. (1990). Immunocytochemical localization of glucagon—related peptides in the cerebral ganglion of the protochordate ascidian, *Styela plicata*. *J. Neuroendocrinol.* 2, 641–645. doi: 10.1111/j.1365-2826.1990.tb00459.x

Pirger, Z., Krajcs, N., and Kiss, T. (2016). "Occurrence, distribution, and physiological function of pituitary adenylyl cyclase-activating polypeptide in invertebrate species," in *Pituitary Adenylate Cyclase Activating Polypeptide – PACAP. Current Topics in Neurotoxicity*, Vol. 11, eds D. Reglodi, and A. Tamas (Cham: Springer), doi: 10.1007/978-3-319-35135-3_2

Pirger, Z., Laszlo, Z., Hiripi, L., Hernadi, L., Toth, G., Lubics, A., et al. (2010a). Pituitary adenylate cyclase activating polypeptide (PACAP) and its receptors are present and biochemically active in the central nervous system of the pond snail *Lymnaea stagnalis*. *J. Mol. Neurosci.* 42, 464–471. doi: 10.1007/s12031-010-9361-x

Pirger, Z., László, Z., Kemenes, I., Tóth, G., Reglodi, D., and Kemenes, G. (2010b). A homolog of the vertebrate pituitary adenylate cyclase-activating polypeptide is both necessary and instructive for the rapid formation of associative memory in an invertebrate. *J. Neurosci.* 30, 13766–13773. doi: 10.1523/JNEUROSCI.2577-10.2010

Pirger, Z., Lubics, A., Reglodi, D., Laszlo, Z., Mark, L., and Kiss, T. (2010c). Mass spectrometric analysis of activity-dependent changes of neuropeptide profile in the snail, *Helix pomatia*. *Neuropeptides* 44, 475–483. doi: 10.1016/j.npep.2010.07.003

Pisani, D., Pett, W., Dohrmann, M., Feuda, R., Rota-Stabelli, O., Philippe, H., et al. (2015). Genomic data do not support comb jellies as the sister group to all other animals. *Proc. Natl. Acad. Sci. U.S.A.* 112, 15402–15407. doi: 10.1073/pnas.1518127112

Putnam, N. H., Butts, T., Ferrier, D. E. K., Furlong, R. F., Hellsten, U., Kawashima, T., et al. (2008). The amphioxus genome and the evolution of the chordate karyotype. *Nature* 453, 1064–1071. doi: 10.1038/nature06967

Raible, F., Tessmar-Raible, K., Osoegawa, K., Wincker, P., Jubin, C., Balavoine, G., et al. (2005). Vertebrate-type intron-rich genes in the marine annelid *Platynereis dumerilii*. *Science* 310, 1325–1326. doi: 10.1126/science.1072104

Rawlings, S. R. (1994). PACAP, PACAP receptors, and intracellular signalling. *Mol. Cell. Endocrinol.* 101, C5–C9. doi: 10.1016/0303-7207(94)90212-7

Reglödi, D., Lengvari, I., Szelier, M., Vigh, S., and Arimura, A. (2000). Distribution of PACAP-like immunoreactivity in the nervous system of oligochaeta. *Peptides* 21, 183–188. doi: 10.1016/S0196-9781(99)00201-6

Reinecke, M. (1981). Immunohistochemical localization of polypeptide hormones in endocrine cells of the digestive tract of *Branchiostoma lanceolatum*. *Cell Tissue Res.* 219, 445–456. doi: 10.1007/BF00209985

Roch, G. J., Wu, S., and Sherwood, N. M. (2009). Hormones and receptors in fish: do duplicates matter? *Gen. Comp. Endocrinol.* 161, 3–12. doi: 10.1016/j.ygcen.2008.10.017

Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D. L., Darling, A., Höhna, S., et al. (2012). Mrbayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61, 539–542. doi: 10.1093/sysbio/sys029

Rosales, C., O'Brien, V., Kornberg, L., and Juliano, R. (1995). Signal transduction by cell adhesion receptors. *Biochim. Biophys. Acta* 1242, 77–98.

Said, S. I., and Mutt, V. (1970). Polypeptide with broad biological activity: isolation from small intestine. *Science* 169, 1217–1218. doi: 10.1126/science.169.3951.1217

Schot, L. P. C., Boer, H. H., Swaab, D. F., and Van Noorden, S. (1981). Immunocytochemical demonstration of peptidergic neurons in the central nervous system of the pond snail *Lymnaea stagnalis* with antisera raised to biologically active peptides of vertebrates. *Cell Tissue Res.* 216, 273–291. doi: 10.1007/BF00233620

Sherwood, N. M., Krueckl, S. L., and McRory, J. E. (2000). The origin and function of the pituitary adenylate cyclase-activating polypeptide (PACAP)/glucagon superfamily. *Endocr. Rev.* 21, 619–670. doi: 10.1210/er.21.6.619

Shioda, S., and Nakamachi, T. (2015). PACAP as a neuroprotective factor in ischemic neuronal injuries. *Peptides* 72, 202–207. doi: 10.1016/j.peptides.2015.08.006

Simakov, O., Marletaz, F., Cho, S. J., Edsinger-Gonzales, E., Havlak, P., Hellsten, U., et al. (2013). Insights into bilaterian evolution from three spiralian genomes. *Nature* 493, 526–531. doi: 10.1038/nature11696

Soares, M. B., Titus, R. G., Shoemaker, C. B., David, J. R., and Bozza, M. (1998). The vasoactive peptide maxadilan from sand fly saliva inhibits TNF-alpha and induces IL-6 by mouse macrophages through interaction with the pituitary adenylate cyclase-activating polypeptide (PACAP) receptor. *J. Immunol.* 160, 1811–1816.

Somogyvári-Vigh, A., Reglödi, D., Li, M., Lengvári, I., Vigh, S., and Arimura, A. (2000). Tissue distribution of PACAP27 and -38 in oligochaeta: PACAP27 is the predominant form in the nervous system of *Lumbricus polyphemus*. *Peptides* 21, 1185–1191. doi: 10.1016/S0196-9781(00)00258-8

Springer, T. A. (1990). Adhesion receptors of the immune system. *Nature* 346, 425–434. doi: 10.1038/346425a0

Starr, C. G., Maderdrut, J. L., He, J., Coy, D. H., and Wimley, W. C. (2018). Pituitary adenylate cyclase-activating polypeptide is a potent broad-spectrum antimicrobial peptide: structure-activity relationships. *Peptides* 104, 35–40. doi: 10.1016/j.peptides.2018.04.006

Stover, N. A. (2006). *Tetrahymena* Genome Database (TGD): a new genomic resource for *Tetrahymena thermophila* research. *Nucleic Acids Res.* 34, D500–D503. doi: 10.1093/nar/gkj054

Sun, C., Song, D., Davis-Taber, R. A., Barrett, L. W., Scott, V. E., Richardson, P. L., et al. (2007). Solution structure and mutational analysis of pituitary adenylate cyclase-activating polypeptide binding to the extracellular domain of PAC1-RS. *Proc. Natl. Acad. Sci. U.S.A* 104, 7875–7880. doi: 10.1073/pnas.0611397104

Sundler, F., Håkanson, R., Alumets, J., and Walles, B. (1977). Neuronal localization of pancreatic polypeptide (PP) and vasoactive intestinal peptide (VIP) immunoreactivity in the earthworm (*Lumbricus terrestris*). *Brain Res. Bull.* 2, 61–65. doi: 10.1016/0361-9230(77)90027-2

Takahashi, T., McDougall, C., Troscianko, J., Chen, W. C., Jayaraman-Nagarajan, A., Shimeld, S. M., et al. (2009). An EST screen from the annelid *Pomatoceros lamarckii* reveals patterns of gene loss and gain in animals. *BMC Evol. Biol.* 9:240. doi: 10.1186/1471-2148-9-240

Telford, M. J., Bourlat, S. J., Economou, A., Papillon, D., and Rota-Stabelli, O. (2008). The evolution of the Ecdysozoa. *Philos. Trans. R. Soc. B Biol. Sci.* 363, 1529–1537. doi: 10.1098/rstb.2007.2243

Valiante, S., Prisco, M., Ricchiari, L., Laforgia, V., Varano, L., and Andreuccetti, P. (2006). Distribution of PACAP in the brain of the cartilaginous fish *Torpedo Marmorata*. *Ann. N. Y. Acad. Sci.* 1070, 591–596. doi: 10.1196/annals.1317.087

Varhalmi, E., Somogyi, I., Kiszler, G., Nemeth, J., Reglodi, D., Lubics, A., et al. (2008). Expression of PACAP-like compounds during the caudal regeneration of the earthworm *Eisenia fetida*. *J. Mol. Neurosci.* 36, 166–174. doi: 10.1007/s12031-008-9125-z

Vaudry, D., Falluel-Morel, A., Bourgault, S., Basille, M., Burel, D., Wurtz, O., et al. (2009). Pituitary adenylate cyclase-activating polypeptide and its receptors: 20 Years after the discovery. *Pharmacol. Rev.* 61, 283–357. doi: 10.1124/pr.109.001370

Vaudry, D., Gonzalez, B. J., Basille, M., Yon, L., Fournier, A., and Vaudry, H. (2000). Pituitary adenylate cyclase-activating polypeptide and its receptors: from structure to functions. *Pharmacol. Rev.* 52, 269–324.

Warfvinge, K., and Edvinsson, L. (2019). Cellular distribution of PACAP-38 and PACAP receptors in the rat brain: relation to migraine activated regions. *Cephalalgia* doi: 10.1177/0333102419893962 [Epub ahead of print].,

Zhong, Y. (1995). Mediation of PACAP-like neuropeptide transmission by coactivation of Ras/Raf and cAMP signal transduction pathways in *Drosophila*. *Nature* 375, 588–592. doi: 10.1038/375588a0

Zhong, Y., and Peña, L. A. (1995). A novel synaptic transmission mediated by a PACAP-like neuropeptide in *Drosophila*. *Neuron* 14, 527–536. doi: 10.1016/0896-6273(95)90309-7

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 03 June 2020
doi: 10.3389/fnins.2020.00525

# Bsx Is Essential for Differentiation of Multiple Neuromodulatory Cell Populations in the Secondary Prosencephalon

*Theresa Schredelseker[1,2,3]*, Florian Veit[1], Richard I. Dorsky[4] and Wolfgang Driever[1,2,3]**

*[1] Developmental Biology, Institute Biology 1, Faculty of Biology, Albert Ludwig University of Freiburg, Freiburg im Breisgau, Germany, [2] BIOSS Centre for Biological Signalling Studies, University of Freiburg, Freiburg im Breisgau, Germany, [3] CIBSS – Centre for Integrative Biological Signalling Studies, University of Freiburg, Freiburg im Breisgau, Germany, [4] Department of Neurobiology & Anatomy, University of Utah, Salt Lake City, UT, United States*

The hypothalamus is characterized by great neuronal diversity, with many neuropeptides and other neuromodulators being expressed within its multiple anatomical domains. The regulatory networks directing hypothalamic development have been studied in detail, but, for many neuron types, control of differentiation is still not understood. The highly conserved Brain-specific homeobox (Bsx) transcription factor has previously been described in regulating *Agrp* and *Npy* expression in the hypothalamic arcuate nucleus (ARC) in mice. While *Bsx* is expressed in many more subregions of both tuberal and mamillary hypothalamus, the functions therein are not known. Using genetic analyses in zebrafish, we show that most *bsx* expression domains are dependent on Nkx2.1 and Nkx2.4 homeodomain transcription factors, while a subset depends on Otp. We show that the anatomical pattern of the ventral forebrain appears normal in *bsx* mutants, but that Bsx is necessary for the expression of many neuropeptide encoding genes, including *agrp, penka*, *vip*, *trh*, *npb,* and *nts*, in distinct hypothalamic anatomical domains. We also found Bsx to be critical for normal expression of two Crh family members, *crhb* and *uts1*, as well as *crhbp*, in the hypothalamus and the telencephalic septal region. Furthermore, we demonstrate a crucial role for Bsx in serotonergic, histaminergic and nitrergic neuron development in the hypothalamus. We conclude that Bsx is critical for the terminal differentiation of multiple neuromodulatory cell types in the forebrain.

**Keywords: neuromodulators, evolutionary conservation, brain-specific homeobox (bsx), zebrafish, basal hypothalamus, neuropeptides, aminergic neurons, nitrergic neurons**

OPEN ACCESS

***Edited by:***
*Jackson Cioni Bittencourt, University of São Paulo, Brazil*

***Reviewed by:***
*David Deitcher, Cornell University, United States*
*Girardet Clémence, Sorbonne Université, France*
*Alex C. Manhães, Universidade Federal do Estado do Rio de Janeiro, Brazil*

****Correspondence:***
*Theresa Schredelseker*
*schredet@tcd.ie*
*Wolfgang Driever*
*driever@biologie.uni-freiburg.de*

***Specialty section:***
*This article was submitted to Neuroendocrine Science, a section of the journal Frontiers in Neuroscience*

***Received:*** *29 November 2019*
***Accepted:*** *28 April 2020*
***Published:*** *03 June 2020*

***Citation:***
*Schredelseker T, Veit F, Dorsky RI and Driever W (2020) Bsx Is Essential for Differentiation of Multiple Neuromodulatory Cell Populations in the Secondary Prosencephalon. Front. Neurosci. 14:525. doi: 10.3389/fnins.2020.00525*

**Abbreviations:** 5-HT, 5-hydroxytryptamine/serotonin; ABas, anterobasal area; AC, anterior commissure; aHyp, alar hypothalamus; ARC arcuate nucleus; bHyp, basal hypothalamus; CSF-c, cerebrospinal fluid-contacting; DC, diencephalic cluster; dpf, days post fertilization; GWAS, genome-wide association study; Hpf, hours post fertilization; LRR, lateral recess region; Mam, mamillary region; OC, optic commissure; ORR, optic recess region; PBas, posterobasal area; PC, posterior commissure; Pit, pituitary; PM, perimamillary domain; PRM, periretromamillary domain; PRR, posterior recess region; "PTv", posterior tuberculum, ventral part; PVN, paraventricular nucleus; RM, retromamillary domain; RTu, retrotuberal domain; RTuD, retrotuberal domain, dorsal part; RTuI, retrotuberal domain, intermediate part; RTuV, retrotuberal domain, ventral part; TALEN, transcription activator-like effector nuclease; TCPT, tract of the commissure of the posterior tuberculum; TelSep, telencephalic septum; Tub, tuberal region; TuD, tuberal domain, dorsal part; TuI, tuberal domain, intermediate part; TuV, tuberal domain, ventral part; VMH, ventromedial hypothalamic nucleus region.

## INTRODUCTION

Together with the pituitary gland, the hypothalamus represents the main neuroendocrine center of the body, regulating core physiological parameters such as growth, body temperature, fluid balance, feeding and energy expenditure, stress and arousal, fatigue and sleep, maternal bonding, and lactation. This functional diversity is reflected by the cellular heterogeneity of the hypothalamus, which consists of multiple distinct nuclei harboring a great variety of cell types (Bedont et al., 2015). Beyond truly neuroendocrine cells secreting neurohormones, a great variety of neuromodulators are expressed in the hypothalamus (Romanov et al., 2019). These neuromodulators influence the response behavior of numerous neural circuits, both within and beyond the hypothalamus (Lee and Dan, 2012; Nadim and Bucher, 2014). Biochemically, neuromodulators range from small molecules, such as nitric oxide or monoamines, to peptides, with more than 100 different neuropeptides being known in the human brain (Wang et al., 2015; Russo, 2017).

Thus, the hypothalamus represents an excellent model to study how extraordinary cellular heterogeneity arises from a limited number of progenitor cells. Previous work already suggested that multiple signaling pathways and a highly combinatorial network of transcription factors are required to generate the great diversity of neuromodulatory cell types in the hypothalamus (reviewed in Bedont et al., 2015; Xie and Dorsky, 2017; Alvarez-Bolado, 2019). Deciphering these signaling pathways and transcription factors will deepen our understanding of how this complex brain structure develops, and ultimately may facilitate the establishment of *in vitro* differentiation protocols for use in regenerative approaches to treat neuroendocrine disorders (Suga, 2019). Extensive evolutionary conservation of hypothalamus anatomy, development, and function facilitates comparisons between different animal models from fish to mammals (Löhr and Hammerschmidt, 2011; Machluf et al., 2011; Dominguez et al., 2015; Puelles and Rubenstein, 2015; Santos-Duran et al., 2015; Xie and Dorsky, 2017; Alié et al., 2018; Schredelseker and Driever, 2020).

The highly conserved Brain-specific Homeobox (Bsx) transcription factor has first been reported in *Drosophila* (Jones and McGinnis, 1993). Bsx expression in several subregions of the hypothalamus, the pineal gland, and the telencephalic septum (TelSep) has been described for mice (Cremona et al., 2004) and zebrafish (Schredelseker and Driever, 2018). While recently Bsx functions in the development of the epithalamus have been elucidated (D'Autilia et al., 2010; Schredelseker and Driever, 2018; Mano et al., 2019), no data exist on the role of Bsx in other forebrain regions of teleosts.

In the mouse ARC, *Bsx* has been shown to be coexpressed with *Agouti-related neuropeptide* (*Agrp*) and *Neuropeptide Y* (*Npy*) (Sakkou et al., 2007). Moreover, *Agrp* and *Npy* expression has been found to be strongly reduced during mouse embryonic development in *Bsx* mutants (Sakkou et al., 2007). Follow-up studies showed that upon activation by Ghrelin (Nogueiras et al., 2008) Bsx directly binds the promoter regions of *Agrp* and *Npy* (Lee et al., 2013). Bsx has been discussed nearly exclusively as a regulator of orexigenic peptide expression in the ARC (Burbridge et al., 2016; Alvarez-Bolado, 2019). Bsx functions beyond the regulation of orexigenic factors in cells of the melanocortin system have been scarcely explored, but lactation deficiencies in *bsx* mutant mice have been reported (McArthur and Ohtoshi, 2007). Given the much broader expression domains of *Bsx*, however, we hypothesized that Bsx has additional functions in the development of other cell types in the hypothalamus. Here, we assessed Bsx functions in the secondary prosencephalon with a focus on the bHyp, where *bsx* is broadly expressed in domains that we recently characterized in detail (Schredelseker and Driever, 2020). We identified transcription factors that regulate the expression of *bsx* in the hypothalamus. In *bsx* mutant embryos, we found patterning in the secondary prosencephalon to be normal. To identify Bsx roles in neuronal differentiation, we focused on peptidergic and aminergic neuromodulators. Comparing wildtype and *bsx* mutant zebrafish embryos, we analyzed the expression of genes encoding zebrafish homologs of the neuropeptides assessed by Díaz et al. (2014), with the exception of *oxytocin* (*oxt*), which in the embryonic hypothalamus of both zebrafish and mouse is only expressed in alar regions (Eaton and Glasgow, 2007; Díaz et al., 2014), where *bsx* is not expressed. We extended our analysis to additional markers for peptidergic, nitrergic and monoaminergic neurons. For 13 of the 26 markers analyzed, we detected absent or strongly reduced expression in defined bHyp subregions of *bsx* mutant embryos, demonstrating that Bsx exerts functions beyond the specification of orexigenic neurons in the ARC. In addition, we found Bsx to be required for *uts1* expression in the TelSep.

Notably, we found that Bsx functions are not restricted to a single hypothalamic nucleus, and that Bsx is also not selectively required for expression of a particular gene specific to a certain neuromodulatory cell type. Instead, Bsx appears crucial for expression of multiple genes in distinct clusters distributed over several distinct hypothalamic areas, while the same genes are expressed independently of Bsx in other areas. This supports the idea that the development of the numerous neuromodulatory cell types in the hypothalamus is controlled by transcription factors in a highly combinatorial manner. By demonstrating that Bsx is a determinant of a surprisingly large number of hypothalamic and septal neuromodulatory cell populations, we propose to replace the notion of Bsx as a transcriptional regulator in a single neuron type by a model that presents Bsx as a major developmental factor in many neuromodulatory cell types within and beyond the hypothalamus. Bsx is a crucial component of a so far not fully understood complex combinatorial code for neuromodulatory neuron differentiation.

## RESULTS

### *bsx* Expression Is Differentially Regulated by Homeobox Transcription Factors in Different Regions of the Secondary Prosencephalon

While *bsx* expression domains in the bHyp have recently been characterized in detail (Schredelseker and Driever, 2020), no data exist on the upstream regulation of *bsx* expression

in the hypothalamus. Severe hypoplasia and deformities in the bHyp were described for *Nkx2.1* mutant mice (Kimura et al., 1996) and critical functions of Nkx-homeodomain factors Nkx2.1, Nkx2.4a, and Nkx2.4b were revealed in zebrafish hypothalamus development (Manoli and Driever, 2014). To asses if hypothalamic *bsx* expression depends on the activity of early acting Nkx-homeodomain transcription factors, we used TALENs to generate loss-of-function alleles for *nkx2.1*, *nkx2.4a* and *nkx2.4b* (**Supplementary Figure S1**). We analyzed *bsx* expression in single and compound mutants by *in situ* hybridization, and found it to be lost in *nkx2.1*, *nkx2.4a,* and *nkx2.4b* triple mutants (here abbreviated *nkx2.1/2.4a/b*) in all hypothalamic areas of embryos 4 dpf (**Figures 1a–d**). However, expression of *bsx* is preserved in *nkx2.1/2.4a/b* triple mutants in two ventral forebrain areas. The first one might correspond to an expression domain which in wildtype embryos is located at the border between the alar and basal plate at the rostrocaudal level of the border between diencephalon and secondary prosencephalon (**Figures 1a–d**, arrow), where *bsx* is expressed in close proximity to the *th* expressing dopaminergic diencephalic cluster 2 (DC2) cells in the "ventral posterior tuberculum" ["PTv"; **Figures 2c–c″** and (Schredelseker and Driever, 2020)]. However, this putative "PTv" domain seem to be reduced in all triple mutants analyzed, suggesting that also this domain is not fully independent of Nkx2.1/2.4a/b factors. Second, *bsx* expression in the TelSep region, which in mice is fused at the midline upon Nkx2.1 loss (Kimura et al., 1996), is normal in *nkx2.1/2.4a/b* triple mutants (**Figures 1a–d**, arrowhead).

Single *nkx2.4b* mutants or *nkx2.1*, *nkx2.4a* double mutants show reduced expression of *bsx* in the bHyp (**Figures 1b,c**), which is, however, not nearly as severe as in the triple mutants (**Figure 1d**). These data are consistent with previous studies showing partial redundancy of Nkx2.1 and Nkx2.4 transcription factors in hypothalamus patterning (Manoli and Driever, 2014). As expected, *bsx* expression in the pineal complex is unaffected in *nkx2.1/2.4a/b* triple mutants (**Figures 1a–d**).

We recently showed that *bsx* is coexpressed with *orthopedia homeobox a* (*otpa*) in mamillary regions as well as in the terminal tuberal hypothalamus (Schredelseker and Driever, 2020). The two paralogs *otpa* and *otpb* function partially redundantly in the ventral forebrain (Ryu et al., 2007; Fernandes et al., 2013). For *Otp* mutant mice, reduced *Bsx* expression has been reported only for the ARC (Lee et al., 2018). In zebrafish embryos, however, we found *bsx* expression to be lost in *otpa*/*otpb* double mutants not only in the ARC but also in all other regions in which *bsx* colocalizes with *otpa* (3 dpf; **Figures 1e–h** and **Supplementary Figure S2**). We concluded that in all $bsx^{+}$/$otpa^{+}$ regions, i.e., the "PTv", parts of the Mam as well as the ARC and the dorsotuberal/anterobasal (TuD/ABas) region, *bsx* expression is strictly dependent on Otp transcription factors. To investigate whether Otpa is also sufficient to induce ectopic *bsx* expression, we assessed *Tg(hsp70l:otpa-ires-egfp-caax)* embryos, which, upon heat shock, express *otpa* broadly throughout the brain (**Supplementary Figures S3A,B**). However, *bsx* expression was normal in embryos 2 dpf even after three waves of *otpa* overexpression induced by heat shocks performed at 12, 24, and 30 hpf (**Supplementary Figures S3C,D**), indicating that Otpa is not sufficient to activate ectopic *bsx* expression.

We further assessed *bsx* expression in *bsx* mutants (Schredelseker and Driever, 2018) to investigate potential autoregulative Bsx functions. We found *bsx* to be expressed normally in the pineal complex and hypothalamus but to be upregulated in the TelSep area in *bsx* mutant embryos (2 dpf: **Figures 1i,j**; 3 dpf: **Figures 1k,l**), suggesting negative autoregulation in this area. We also found few dispersed *bsx* expressing cells around the di-mesencephalic boundary in all *bsx* mutant embryos (**Figures 1j,l**, asterisk) but never in wildtypes. We summarized the upstream regulation of *bsx* in different domains of the forebrain in **Figure 1m**.

## *bsx* Is Expressed in Differentiating Neurons Rather Than in Neural Stem Cells

The hypothalamus, like many brain regions, displays a pronounced centrifugal ventricular to mantle gradient of neurogenesis and cell maturation (Affaticati et al., 2015; Alvarez-Bolado, 2019). By double-fluorescent *in situ* hybridization, we compared *bsx* expression along the apical to basal axis of the hypothalamic neuroepithelium with expression of genes characteristic for proliferating neural stem cells or postmitotic differentiated neurons. Expression of the neural stem and radial glia markers *emx2* and *fezf2* (Cecchi, 2002; Yang et al., 2012; Alvarez-Bolado, 2019) are restricted to areas close to the ventricle, while *bsx* expression was found to be located further away from the ventricle in mantle zones (**Figures 2a–b″**).

We then compared *bsx* expression to markers of differentiated neurons. We detected *bsx* expression in dopaminergic DC5 and 6 cells in the mamillary hypothalamus (**Figures 2c–c″**). In the mamillary hypothalamus, we also found cells expressing both *bsx* and *vasoactive intestinal peptide* (*vip*) (**Figures 2d–d″**). The location of those $vip^{+}$ cells within the $otpa^{+}$ positive parts of the mamillary hypothalamus has been shown by others (Wolf and Ryu, 2013). We recently found medial regions of this brain area to correspond to mamillary hypothalamus, while more lateral regions express markers characteristic for the tuberal hypothalamus (Schredelseker and Driever, 2020). Within the tuberal hypothalamus, lateral to the *vip* expressing cells, as well as in the TuD/ABas region, we found cells expressing both *bsx* and the endogenous opioid preproprotein *proenkephalin a* (*penka*) (**Figures 2e–e″**). In an adjacent region of the terminal tuberal hypothalamus, corresponding to the ARC region (Schredelseker and Driever, 2020), we also detected *agrp* cells within the *bsx* expressing territories (**Figures 2f–f″**). Notably, in the border region between alar and basal plate at the rostrocaudal level of the border between hypothalamus and diencephalon, in which also the *th* expressing dopaminergic DC2 cluster is located, we found *crhb* cells to express *bsx* (**Figures 2g–g″**). Taken together, these data suggest that *bsx* is expressed in mantle postmitotic

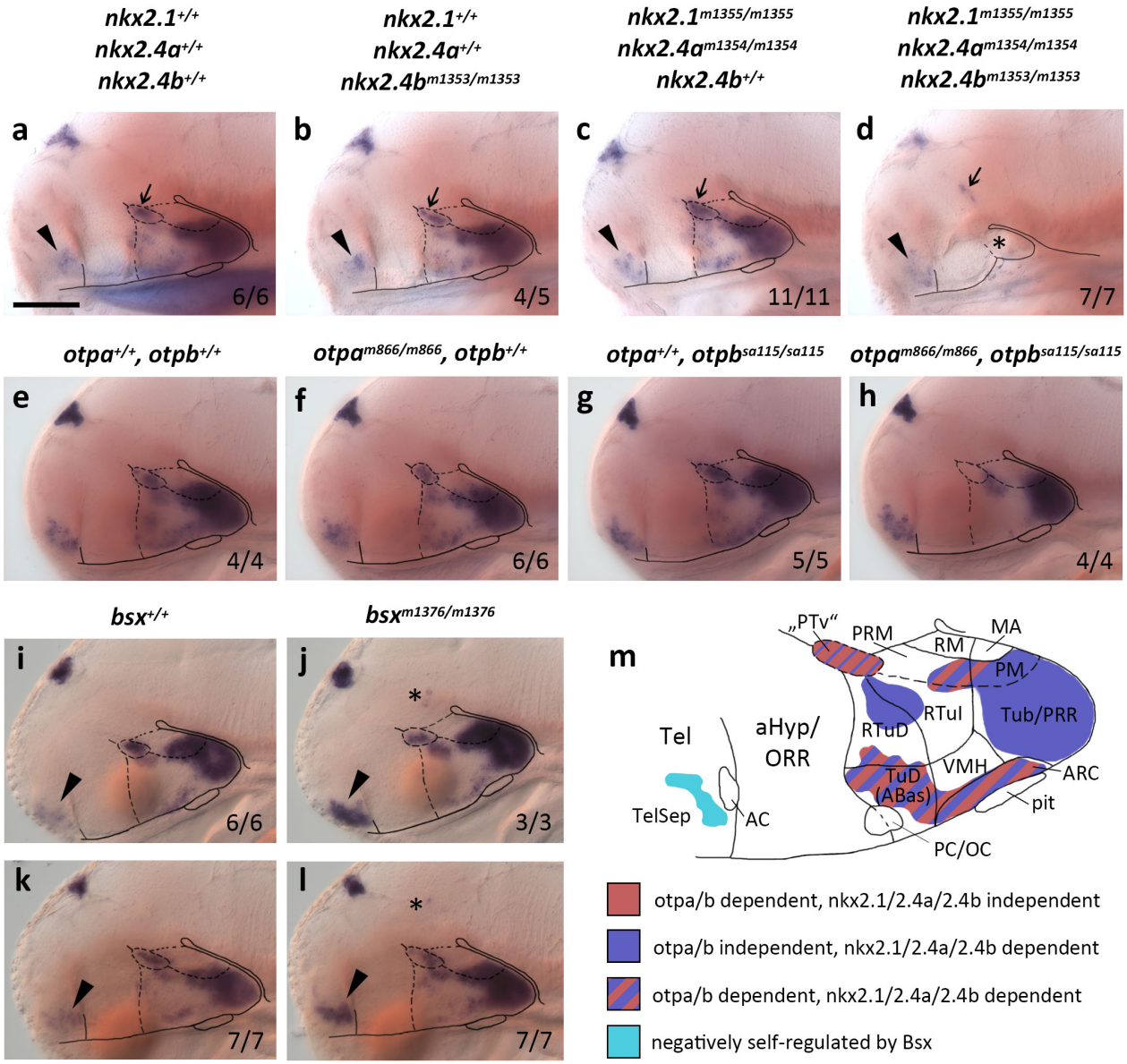


**FIGURE 1 |** *bsx* expression depends on Nkx2.1, Nkx2.4, and Otp transcription factors. **(a–l)** Lateral views of wildtype and mutant (genotypes as indicated above panels) 4 dpf **(a–d)**, 3 dpf **(e–h,k,l)** or 2 dpf **(i,j)** embryos, stained by *in situ* hybridization for *bsx* expression. All images are minimum intensity projections of 50 brightfield focal planes (distance 1 μm). Anterior is to the left. Scale bar **(a–l)**: 100 μm. The ratio of observed phenotype and total analyzed embryos per genotype [*n* (phenotype)/*n* (total)] is indicated in right bottom corner of each panel. Arrowheads indicates *bsx* expression in the telencephalic septal region; small arrows indicate *bsx* expression which might correspond to "PTv" region; asterisk in **(d)** indicates tissue which might correspond to residual basal hypothalamus (bHyp) but which was not confirmed through marker gene expression in *nkx2.1*, *nkx2.4a*, *nkx2.4b* triple mutants; asterisks in **(j,l)** indicate ectopic *bsx* expressing cells. **(m)** Schematic drawing of a lateral view of *bsx* expressing forebrain regions of a zebrafish embryo 3–4 dpf illustrating which of the *bsx* expression domains depend on Otp or Nkx2.1 and Nkx2.4 transcription factors as indicated by color code. Anterior at left. For abbreviations see list.

differentiating neurons rather than in ventricular neural stem or progenitor cells.

## Forebrain Patterning Is Normal in *bsx* Mutants

To investigate a potential role of Bsx in hypothalamus patterning, we compared *bsx* mutants with wildtype siblings for gene expression patterns which define specific progenitor domains in the embryonic hypothalamus. *sonic hedgehog a* (*shha*) expression in the secondary prosencephalon marks the alar-basal boundary [(Puelles et al., 2012); **Figure 3a**, red dashed line; for regional organization of the hypothalamus see **Figure 1m**]. Expression of *shha* being unaffected in *bsx* mutants suggests a normal course of the alar-basal boundary in *bsx* mutants (**Figure 3a**). Expression of *pax6a* and *pax7a* is unaffected in *bsx* mutants compared to

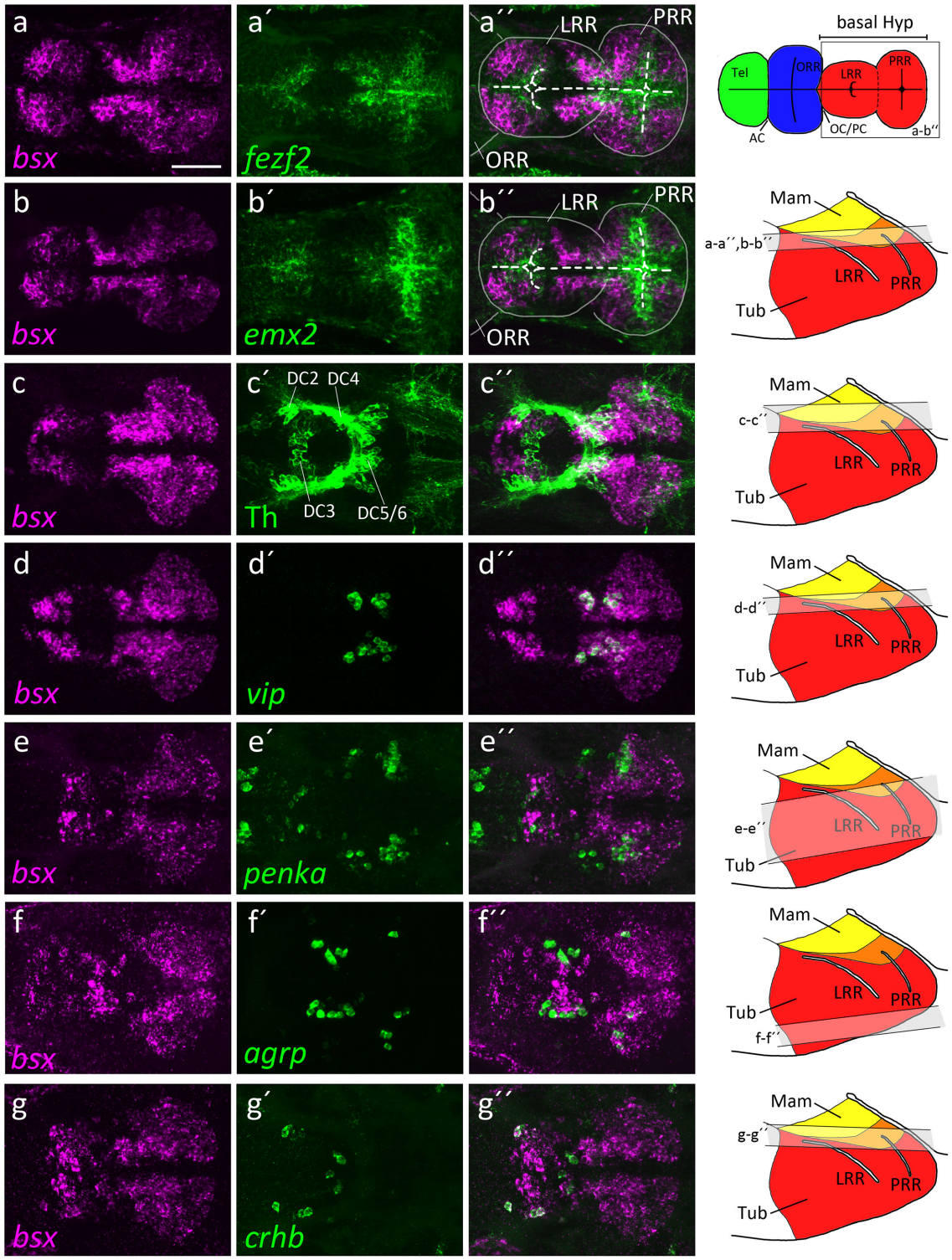


**FIGURE 2 |** *bsx* expression within the bHyp in relation to expression of neural progenitor markers and neuropeptidergic genes. Dorsal view of confocal sections of zebrafish embryos at 2 dpf **(a–b′′)**, 3 dpf **(c–e′′)** or 4 dpf **(f–g′′)** after double-fluorescent whole-mount *in situ* hybridization using probes as indicated. Maximum intensity *Z*-projections of 30 **(a–b′′,d–d′′)**, 40 **(f–g′′)**, 50 **(c–c′′)**, or 70 **(e–e′′)** single confocal planes (1 μm steps) are shown. Schematics on the right show lateral view of the bHyp for 2–4 dpf embryos indicating which planes were selected for *Z*-projections in the panels indicated. Expression of *bsx* is detected further away from the ventricle [white dashed line in **(a–b′′)**] than expression of neuronal progenitor markers **(a–b′′)**. *bsx* expression colocalizes in neurons expressing Th [anti-Th immunostain; **(c–c′′)**] or neuropeptidergic transcripts **(d–g′′)**. Scale bar in **(a)** for all images: 50 μm. Schematics in the top right represents a model of the zebrafish forebrain highlighting the ventricular recesses. Anterior at left. For abbreviations see list.

wildtype (**Figures 3b,c**, arrowhead), indicating regular formation of alar and basal parts of prosomer 3, respectively. Unchanged expression of *pax7a* in the pituitary further indicates proper formation of the pituitary intermediate lobe (**Figure 3c**, arrow). *pax6a* in the telencephalon and in the progenitor cells lining the optic recess (**Figure 3b**, arrow) is also expressed identically to wildtype siblings in *bsx* mutants.

*otpa* expression is normal in the "PTv" (**Figure 3d**, black arrowhead), perimamillary (PM) region (**Figure 3d**, white arrow), TuD/ABas area (**Figure 3d**, black arrow) as well as in the aHyp/ORR (**Figure 3d**, white arrowhead). Similarly, *lhx5* expression is unchanged in mamillary regions (**Figure 3e**, black arrowhead), TuD/ABas (**Figure 3e**, white arrow), aHyp/ORR (**Figure 3e**, black arrow) and the telencephalon (**Figure 3e**, white arrowhead). Normal *foxb1a* expression indicates the Mam to develop properly (**Figure 3f**, black arrowhead). *lhx6* expression is unchanged in the pallidum (**Figure 3g**, white arrowhead), in the aHyp/ORR (**Figure 3g**, black arrow), as well as in the bHyp (**Figure 3g**, black arrowhead) where in mouse it has been demonstrated to selectively mark RTuV/TuV (Puelles et al., 2012).

Also other tuberal regions appear to be well developed in *bsx* mutants, as judged from unaltered expression of *dlx5a* (**Figure 3h**, black arrowhead), which marks all tuberal regions except the core region of the ventromedial hypothalamus (VMH) (Puelles et al., 2012; Morales-Delgado et al., 2014), and of *isl1* (**Figure 3i**, black arrowheads), which marks all tuberal regions (Puelles et al., 2012). Expression of *dlx5a* and *isl1* is also unchanged in the telencephalon (**Figures 3h,i**, white arrow) and aHyp/ORR (**Figures 3h,i**, white arrowhead). *nkx2.1* expression is normal in both mamillary and tuberal regions (**Figure 3j**, black and white arrowheads, respectively) as well as in the telencephalon and aHyp/ORR (**Figure 3j**, black and white arrow, respectively). The gap in *nkx2.1* expression in the intermediate and dorsal retrotuberal (RTuI/RTuD) regions (**Figure 3j**, white asterisk) was observed previously (Rohr et al., 2001; Manoli and Driever, 2014), and is also unchanged in *bsx* mutants. In this *nkx2.1* negative territory, *nkx2.2a* is expressed (**Figure 3k**, white asterisk). *nkx2.2a* has been described to be expressed in both the liminal and subliminal band along the alar-basal boundary (Puelles et al., 2012; **Figure 3k**, red dashed line), and is expressed identically to wildtype in *bsx* mutants. *lhx9* expression indicates normal development of the RTuD/PBas region (Puelles et al., 2012; **Figure 3l**, black arrowhead). *nr5a1a* and *nr5a2* are expressed in equal patterns in wildtype and *bsx* mutant embryos, suggesting proper formation of TuD/ABas, VMH and the ARC region (Kurrasch et al., 2007; Bedont et al., 2015; Xie and Dorsky, 2017) (**Figures 3m,n**, black arrowheads). The PRR also appears to develop orderly, as observed through *lef1* and *fezf2* expression (Levkowitz et al., 2003; Wang et al., 2012; **Figures 3o,p**, black arrowheads). A small expression domain of *lef1* in the RTu region (**Figure 3o**, black arrow), which has previously been described in mice (Ferran et al., 2015), also appears to be unchanged in *bsx* mutants. *fezf2* expression is normal along the optic recess and in the telencephalon (**Figure 3p**, black and white arrow, respectively). In conclusion, we could not detect any difference in patterning of the secondary prosencephalon between wildtypes and *bsx* mutants.

## Expression of *agrp* and *penka*, but Not *pomca* or *npy,* in Terminal Regions of the Tuberal Hypothalamus Depends on Bsx

In the mouse ARC, Bsx has been shown to regulate expression of *Agrp* and *Npy,* but not *Pomc* (Sakkou et al., 2007). We found *agrp* expression in the putative ARC area to be lost in *bsx* mutants at 3 dpf (**Figures 4a–d**, arrowheads) and 4 dpf (**Supplementary Figures S4A–D**, arrowheads), while *pomca* is expressed normally at both 3 dpf (**Supplementary Figure S8I**) and 4 dpf (**Supplementary Figure S9I**), indicating that Bsx functions are conserved in zebrafish. We observed a few dispersed *agrp* expressing cells in peduncular RTu regions which are unaffected by loss of Bsx (**Figures 4a–d**). We do not know whether these cells represent a distinct cluster which remains anatomically and potentially functionally disjunct from the ARC population, whether those cells are migratory, and/or whether those cells express *agrp* only transiently during embryonic development. Notably, *npy* is not expressed in the ARC area in zebrafish larvae, and we could not detect altered *npy* expression in any brain region of *bsx* mutant embryos at 3 dpf (**Supplementary Figure S8H**) or 4 dpf (**Supplementary Figure S9H**).

Given that *bsx*, in mice and in teleosts, is expressed not only in the ARC, but also in multiple other regions of the bHyp (Cremona et al., 2004; Schredelseker and Driever, 2020; **Figures 1**, **2**), we analyzed several other neuropeptidergic cell populations therein. Since we demonstrated several clusters of *penka* expressing neurons to be located within the *bsx* expression domain in both the LRR and PRR (**Figures 2e–e″**), we analyzed *penka* expression in *bsx* mutant embryos. We found *penka* expression in one cell cluster within the LRR to be lost in *bsx* mutants at 3 dpf (**Figures 4e–h**, arrowheads) and 4 dpf (**Supplementary Figures S4E–H**, arrowheads). In prosomeric anatomical terms this cluster is located in the terminal TuD/ABas region (Schredelseker and Driever, 2020). We concluded that in addition to the role of Bsx on *agrp* neuron differentiation, which has been shown previously in mice (Sakkou et al., 2007), Bsx also functions in differentiation of *penka* neurons in the terminal tuberal hypothalamus in zebrafish.

## Expression of *vip*, *npb,* and *trh* in Subregions of the Mamillary Region Is Dependent on Bsx

In mice, *vip* expressing cells originate in the aHyp (Díaz et al., 2014) and play a crucial role in circadian rhythm regulation (Maywood et al., 2006). In contrast, during zebrafish development *vip* is expressed not only in the aHyp/ORR but also in mamillary regions of the bHyp, where *vip* expression is dependent on Otp transcription factors (Wolf and Ryu, 2013). We found the mamillary *vip* cluster to be absent in *bsx* mutants at 3 dpf (**Figures 5a–d**, arrowheads) and 4 dpf (**Supplementary Figures S5A–D**, arrowheads). Taken together with the observation that *bsx* expression in the Mam is reduced in *otpa*/*otpb* double mutants (**Figures 1e–h**), we hypothesized that Bsx acts downstream of Otp factors in the differentiation of *vip* neurons.

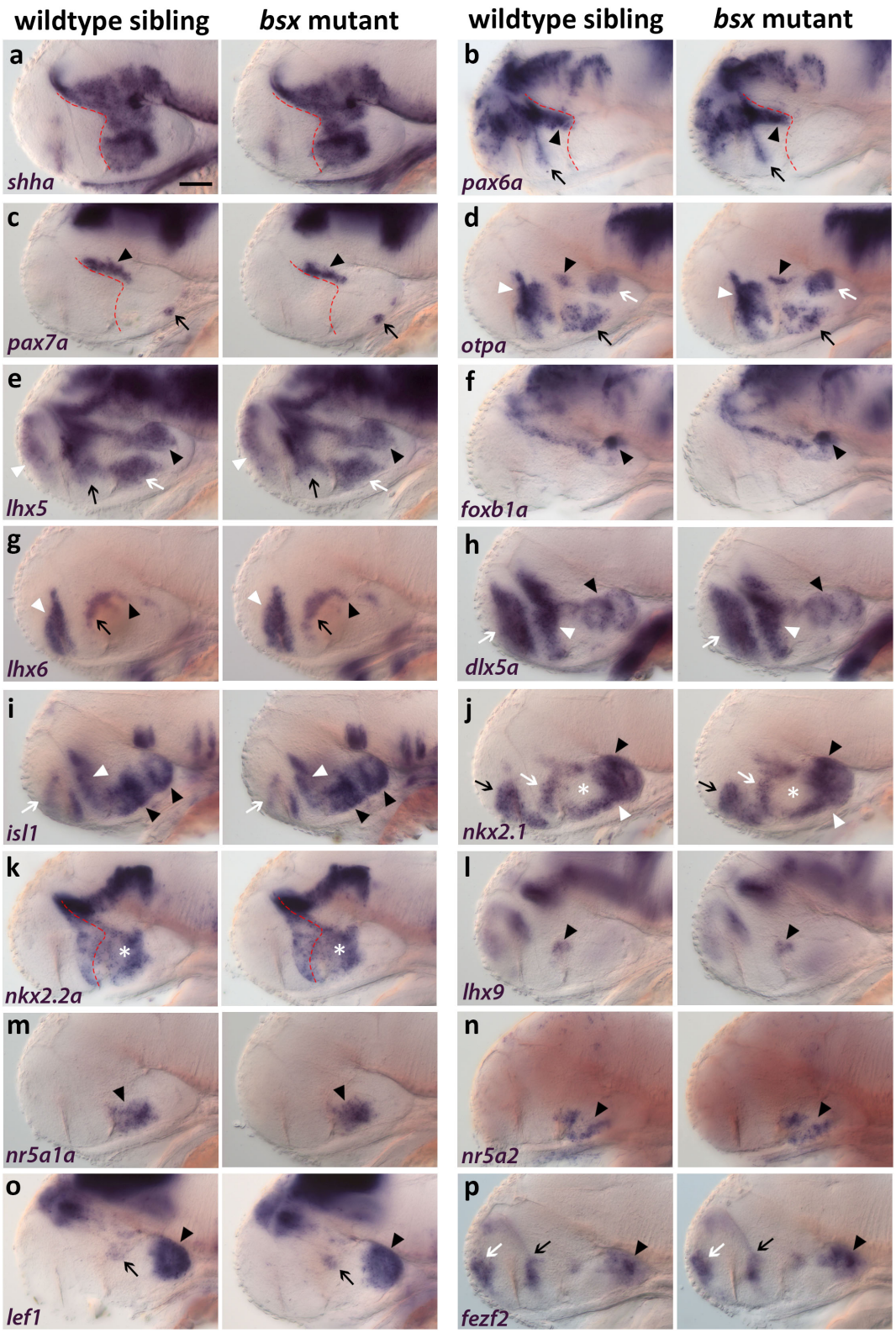


**FIGURE 3 |** Forebrain patterning is normal in *bsx* mutants. **(a–p)** Lateral views of 2 dpf zebrafish embryo heads stained by *in situ* hybridization using probes as indicated. No differences in expression of any marker was observed between wildtype and *bsx* mutant embryos. All images are minimum intensity projections of 40 brightfield focal planes (distance 1 μm). Anterior is to the left. Scale bar in **(a)** for all images: 50 μm. Anterior at left. Number *n* of embryos analyzed (wildtype) = 3 **(a,e,h,i)**; 5 **(b,m)**; 6 **(c,p)**; 10 **(d)**; 9 **(f,n)**; 14 **(g)**; 8 **(j,k,o)**; 12 **(l)**. *n* (*bsx* mutants) = 3 **(a)**; 11 **(b,g)**; 4 **(c,j)**; 12 **(d)**; 5 **(e,l,o)**; 9 **(f,n,p)**; 8 **(h,i)**; 10 **(k)**; 6 **(m)**.

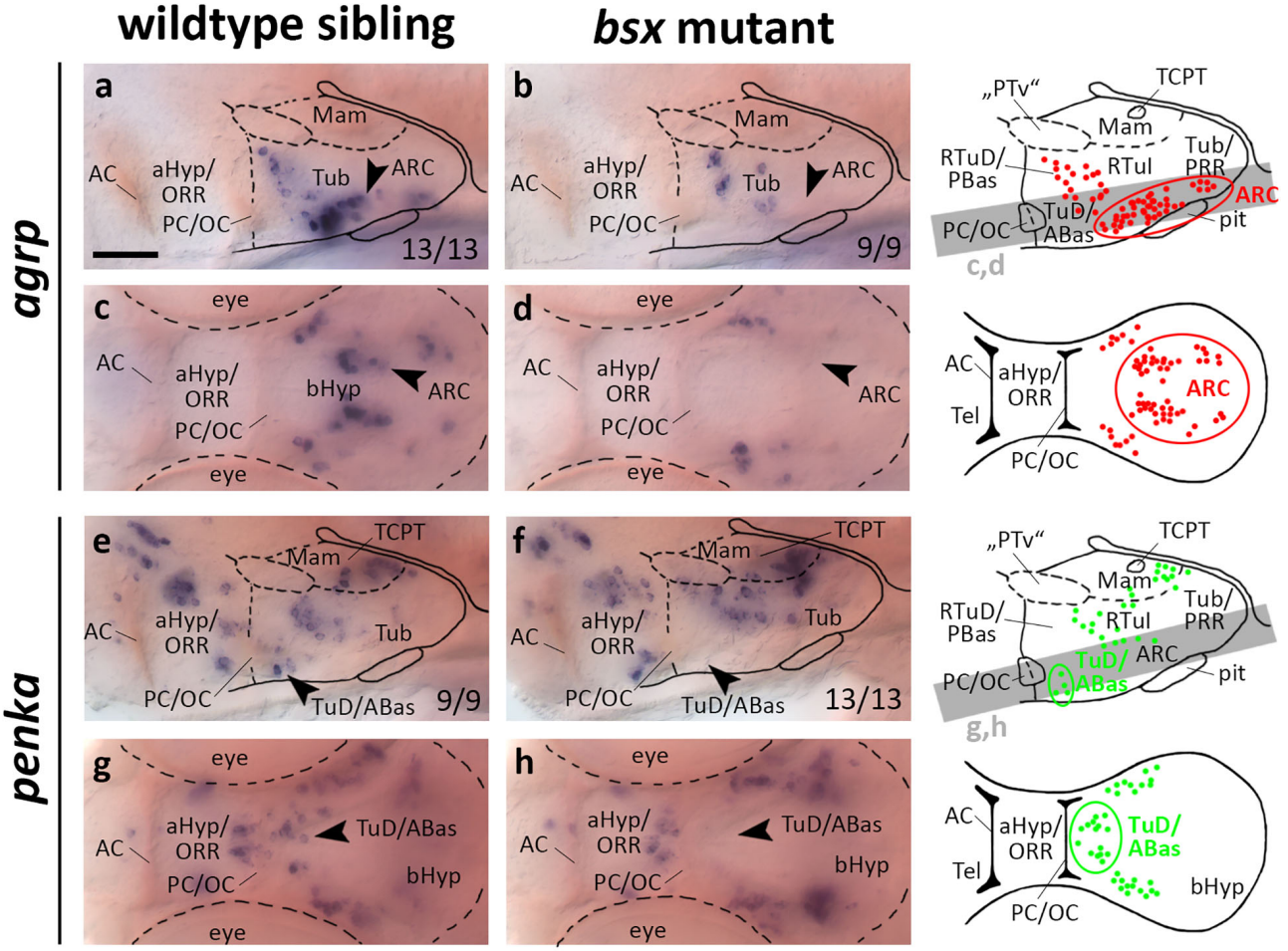


**FIGURE 4 |** *agrp* and *penka* expression domains in the terminal tuberal hypothalamus are lost in *bsx* mutant embryos. Lateral **(a,b,e,f)** and dorsal views **(c,d,g,h)** of the ventral forebrain in 3 dpf embryos stained by *in situ* hybridization using probes as indicated. All images are minimum intensity projections of 40 brightfield focal planes (distance 1 μm). Scale bar in **(a)** for all images: 50 μm. Anterior at left. Schematics to the right show location of *agrp* (red) or *penka* (green) expressing cells as observed by stainings and indicate the focal planes which are shown in dorsal view pictures and schematics. Cells with Bsx-dependent expression are circled. Numbers *n* (phenotype shown)/*n* (total analyzed) as indicated.

In a neighboring domain within the Mam, in close proximity to the tract of the commissure of the caudal tuberculum (TCPT; Wilson et al., 1990), we found expression of *npb* to be strongly reduced upon Bsx loss both at 3 dpf (**Figures 5e–h**, arrowheads) and 4 dpf (**Supplementary Figures S5E–H**, arrowheads). In mouse, *Npb* is expressed in multiple nuclei within the telencephalon, aHyp and tuberal hypothalamus (Takenoya et al., 2013), and has been shown to be involved in pain processing, circadian rhythm, sleep/wake regulation, and feeding behavior (Dvorakova, 2018). To our knowledge, thus far, no reports exist on *npb* expression in mamillary regions, or on *npb* functions in zebrafish.

Thyrotropin-releasing hormone (Trh) neurons have been implicated in energy homeostasis, arousal and cold response (Hollenberg, 2008; Hara et al., 2009; Zhang et al., 2018). We found a cluster of *trh* expressing cells that is absent in the Mam of *bsx* mutants at 3 dpf (**Figures 5i–l**, arrowheads) and 4 dpf (**Supplementary Figures S5I–L**, arrowheads). We conclude that Bsx is essential for differentiation of neuropeptidergic cell clusters not only within the developing ARC but also in multiple other regions within the bHyp.

## Expression of *hcrt*, *cart4*, *sst1.1*, *pdyn*, *galn*, *avp*, *pmch*, *pmchl*, *npvf*, and *nmu* Is Normal in *bsx* Mutant Embryos

Of the 16 neuropeptide encoding genes analyzed by Díaz et al. (2014), for 11 expression has been reported to originate in the bHyp in mouse. Except for *growth hormone-related hormone* (*ghrh*), which we could not detect in the zebrafish hypothalamus at 3 or 4 dpf (data not shown), and *oxytocin*, which in zebrafish larvae as in mice is restricted to aHyp regions (Unger and Glasgow, 2003; Díaz et al., 2014), we found zebrafish homologs for all other genes to be expressed in at least few cells within the bHyp. In addition to *npy* and *pomca* which we previously mentioned, we detected the expression of

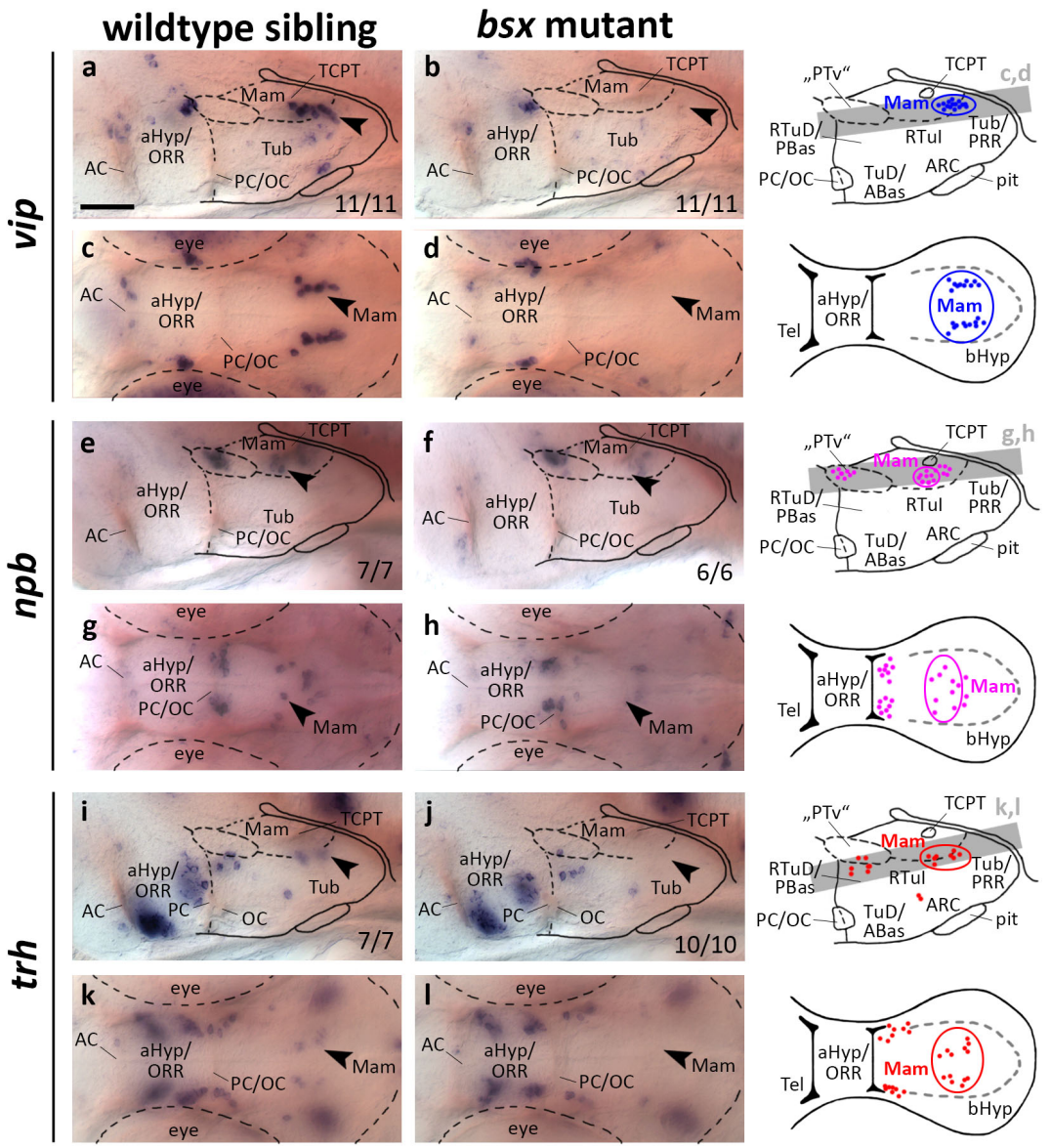


**FIGURE 5 |** *vip*, *npb*, and *trh* expression domains in mamillary hypothalamic regions are lost in *bsx* mutant embryos. Lateral **(a,b,e,f,i,j)** and dorsal views **(c,d,g,h,k,l)** of the ventral forebrain in 3 dpf embryos stained by *in situ* hybridization using probes as indicated. All images are minimum intensity projections of 40 brightfield focal planes (distance 1 μm). Scale bar in **(a)** for all images: 50 μm. Anterior at left. Schematics to the right show location of *vip* (blue), *npb* (magenta), or *trh* (red) expressing cells as observed by stainings and indicate the focal planes which are shown in dorsal view pictures and schematics. Cells with Bsx-dependent expression are circled. Numbers *n* (phenotype shown)/*n* (total analyzed) as indicated.

*hypocretin* (*hcrt*), *cocaine- and amphetamine-regulated transcript 4* (*cart4*), *somatostatin 1.1* (*sst1.1*), *prodynorphin* (*pdyn*), *galanin* (*galn*), *arginine vasopressin* (*avp*), *pro-melanin-concentrating hormone* (*pmch*), *pro-melanin-concentrating hormone, like* (*pmchl*), *neuropeptide VF precursor* (*npvf*), and *neuromedin U* (*nmu*) in *bsx* mutants to be similar to wildtype both at 3 dpf (**Supplementary Figure S8**) and 4 dpf (**Supplementary Figure S9**). Therefore, many neuropeptide encoding genes are normally expressed in *bsx* mutants, and Bsx appears to be specifically required only for expression of a defined group of neuropeptide precursor genes.

## Bsx Is Required for Normal Development of the Crh System and for *uts1* Expression in the Telencephalic Septal Region

In teleosts, four genes encode the corticotropin-releasing hormone (Crh) family of ligands: *crha*, *crhb*, *uts1*, and *urocortin 3 like* (Cardoso et al., 2016). Their binding to the Crh receptors is additionally regulated through presence and concentration of a Crh binding protein (Crhbp) (Ketchesin et al., 2017). The Crh system is a core component of the stress response acting

through the hypothalamic-pituitary-adrenal axis (Denver, 2009), but has also been implicated in energy homeostasis and appetite regulation (Matsuda, 2013).

In the embryonic zebrafish brain *crhb* expressing cells in the "PTv" region have been found not only to be located in close proximity to the *th* expressing DC2 dopaminergic neurons, but also to be expressed under control of shared transcriptional regulation (Löhr et al., 2009). For instance, "PTv" *crhb* neurons also depend on the Otp transcription factors (Löhr et al., 2009). When we assessed *crhb* expression in the "PTv" region in *bsx* mutants, we found this cell cluster to be absent (**Figures 6a–d**, arrowheads, and **Supplementary Figures S6A–D**, arrowheads). Taken together with our observation that *bsx* expression in this region is dependent on Otp (**Figures 1e–h**), we conclude that Otp may contribute to differentiation of *crhb* neurons in the "PTv" area by activation of *bsx* expression, or that both factors may act in a combinatorial manner in *crhb* differentiation.

*crhbp* in the embryonic zebrafish is expressed in the telencephalon, aHyp/ORR, "PTv" area, and the border region between tuberal and mamillary regions of the bHyp (**Figures 6e–j** and **Supplementary Figures S6E–H**, arrowheads). Additionally, a cluster of *crhbp* expressing cells was found in the terminal tuberal hypothalamus reaching from the TuD/ABas region to the ARC (**Figures 6e–h**, black arrowheads). This cluster was not detected in *bsx* mutants (**Figures 6f,h**, arrowheads, and **Supplementary Figures S6E–H**, arrowheads), indicating a differentiation defect of those neurons upon loss of Bsx function. In the border region between mamillary and tuberal regions we found a far laterally located cluster of *crhbp* expressing cells to be absent in *bsx* mutants (**Figures 6i,j**, arrowheads), while more medially located *crhbp* cells appear to be unaffected. Based on our previous studies (Schredelseker and Driever, 2020), within the PRR we associated the Bsx-dependent lateral *crhbp* cells with tuberal regions and the Bsx-independent cells with mamillary regions.

*urotensin 1* (*uts1*) expression in the zebrafish Mam has been demonstrated to be dependent on Otp transcription factors (Wolf and Ryu, 2013). A member of the mammalian homologs to Urotensin, Urocortin 1, has been detected in the lateral telencephalic septum (Kozicz et al., 1998), which is of pallidal origin in mice (Medina and Abellán, 2012). We detected a few *uts1* expressing cells adjacent to the AC in an area hypothesized to be homologous to the mammalian TelSep region (Schredelseker and Driever, 2020; **Figure 6k**, arrowhead). We could not detect any *uts1* expressing neurons in this region in *bsx* mutant embryos 3 dpf (**Figures 6k–n**, arrowheads) or 4 dpf (**Supplementary Figures S6I–L**, arrowheads). Since we have demonstrated the pallidum to develop normally as inferred from normal expression of *nkx2.1* and *lhx6* (**Figure 3**), we conclude that Bsx is critical for differentiation of *uts1* neurons, but not for patterning of the pallidal domain.

## Bsx Functions in Differentiation of Monoaminergic Cells in the Posterior Recess Region

For the PRR we recently proposed homology of ventral and dorsolateral parts to tuberal regions, and of dorsomedial parts to mamillary regions, based on gene expression data (Schredelseker and Driever, 2020). Still, the phylogenetic relationship of the PRR remains particularly elusive as most phyla lack a posterior hypothalamic recess (Vernier, 2017; Yamamoto et al., 2017). Cell types within the teleost-specific PRR are well characterized and include many monoaminergic cerebrospinal fluid contacting (CSF-c) cells (Eriksson et al., 1998; Filippi et al., 2010; Xavier et al., 2017; Xie et al., 2017).

Since *bsx* is expressed in the PRR, we aimed to assess the differentiation of monoaminergic cells therein. We made use of *vmat2* as a marker for all monoaminergic cell types, of *hdc* as a specific marker for histaminergic cells, and of *th* as well as its paralog *th2* to label dopaminergic cells. Since *tryptophan hydroxylase* expression is low in hypothalamic serotonergic cells (Bellipanni et al., 2002), we made use of an antibody recognizing a 5-HT/paraformaldehyde conjugate (Bosco et al., 2013). We found the expression of *vmat2* to be strongly reduced in the PRR (**Figures 7a–d**, arrowheads, and **Supplementary Figures S7A–D**, arrowheads), indicating that monoaminergic cells develop abnormally in *bsx* mutants. The presence of *hdc* expressing neurons in the PRR supports our hypothesis that parts of the PRR are homologous to the RTuV/TuV territory as defined by the prosomeric model, in which RTuV/TuV has been proposed to be the only source of histaminergic neurons in the brain (Puelles et al., 2012). We detected no *hdc* expressing cells in *bsx* mutants 3 dpf (**Figures 7e–h**, arrowheads) and 4 dpf (**Supplementary Figures S7E–H**, arrowheads), and concluded that all histaminergic cells present in the brain at these embryonic stages are strictly Bsx-dependent. We found both the expression of *th1* (DC7; **Supplementary Figures S10A–H**) and *th2* (**Supplementary Figures S10I–L**) to be normal in the *bsx* mutant PRR, suggesting dopaminergic cells to develop in a Bsx-independent manner. The *th* expressing cell clusters DC2, 4, 5, and 6 have previously been shown to be dependent on Otp (Ryu et al., 2007). Since we found multiple Otp-dependent expression domains of neuropeptidergic genes also to be affected in *bsx* mutants, we were surprised to find *th* expression in these clusters to be unaffected by Bsx loss (**Supplementary Figures S10A–H**). This indicates that Otp transcription factors function independently of Bsx in dopaminergic differentiation.

In *bsx* mutants, 5-HT immunoreactivity is strongly reduced not only in the PRR (**Figures 7i,j**, white outline arrowheads) but also in the other two hypothalamic serotonergic cell clusters (**Figures 7i,j**, filled white arrowheads and arrows). We conclude that Bsx is crucial for the terminal differentiation of histaminergic and serotonergic, but not dopaminergic cells in the hypothalamus.

## Bsx Is Required for Normal *nts* and *nos1* Expression in the Posterior Recess Region

Neurotensin attracted much attention as an important neuromodulator of multiple, and most notably dopaminergic, neuronal circuits (Dobner and Carraway, 2013). We detected *nts* expression in the PRR of wildtype but not *bsx* mutant embryos 3 dpf (**Figures 8a–d**, arrowheads). At 4 dpf we detected few

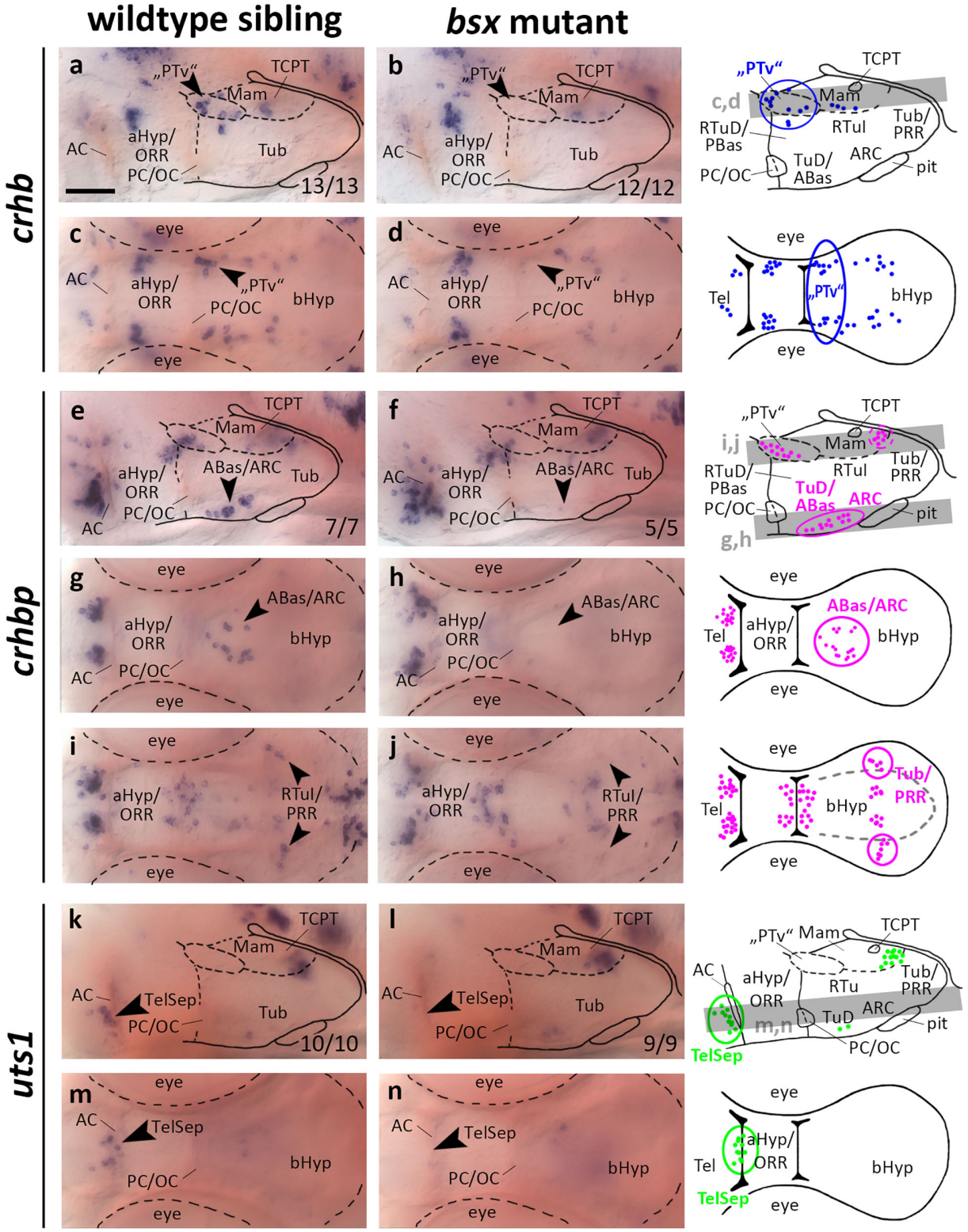


**FIGURE 6 |** *crhb, crhbp* and *uts1* expression domains in the secondary prosencephalon are lost in *bsx* mutant embryos. Lateral **(a,b,e,f,k,l)** and dorsal views **(c,d,g–j,m,n)** of the ventral forebrain in 3 dpf embryos stained by *in situ* hybridization using probes as indicated. All images are minimum intensity projections of 40 brightfield focal planes (distance 1 μm). Scale bar in **(a)** for all images: 50 μm. Anterior at left. Schematics to the right show location of *crhb* (blue), *crhbp* (magenta), or *uts1* (green) expressing cells as observed by stainings and indicate the focal planes which are shown in dorsal view pictures and schematics. Cells with Bsx-dependent expression are circled. Numbers *n* (phenotype shown)/*n* (total analyzed) as indicated.

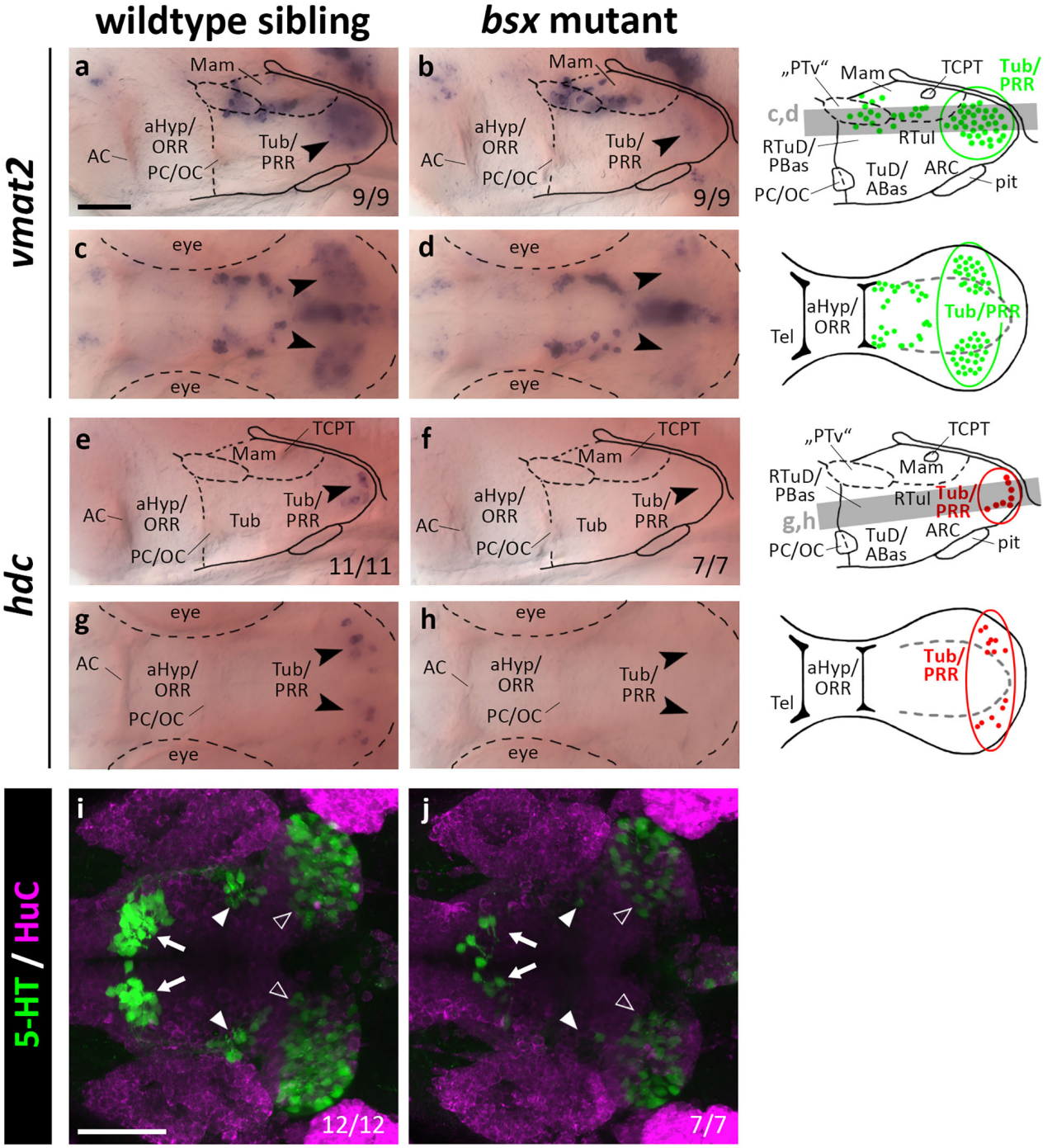


**FIGURE 7 |** *vmat2* expression, *hdc* expression and 5-HT immunoreactivity in the posterior recess region are reduced or absent in *bsx* mutant embryos. Lateral **(a,b,e,f)** and dorsal views **(c,d,g,h)** of the ventral forebrain in 3 dpf embryos stained by *in situ* hybridization using probes as indicated. **(i,j)** Dorsal view of the ventral forebrain in 3 dpf embryos immunostained using 5-HT (green) and HuC antibody (magenta). Images show minimum intensity projections of 40 brightfield focal planes (**a–h**; 1 μm steps) or maximum intensity projections of 40 confocal planes (**i,j**; 1 μm steps). Scale bars in **(a)** for **(a–h)** and in **(i)** for **(i)** and **(j)**: 50 μm. Anterior at left. Schematics to the right show location of *vmat2* (green) and *hdc* (red) expressing cells as observed in stainings and indicate the focal planes which are shown in dorsal view pictures and schematics. Cells with Bsx-dependent expression are circled. Numbers *n* (phenotype shown)/*n* (total analyzed) as indicated.

*nts* expressing cells in the PRR of *bsx* mutants, while wildtype embryos at that stage showed a stronger signal (**Supplementary Figures S11A–D**, arrowheads).

*nos1* has been demonstrated to be coexpressed in *th* expressing neurons in *Xenopus* (Lopez et al., 2005) and zebrafish embryos (Poon et al., 2003). Functionally, hypothalamic *nos1* has been

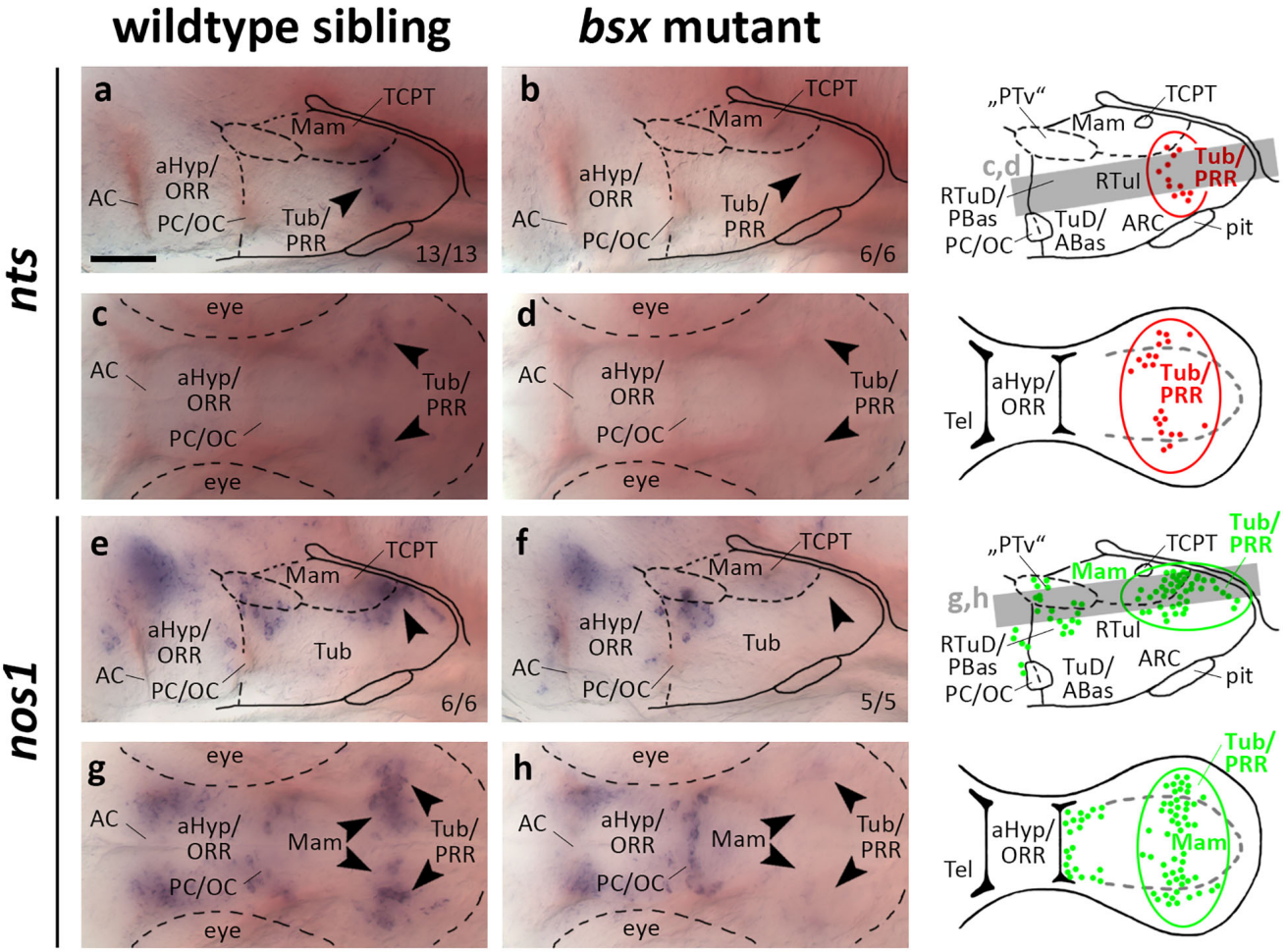


**FIGURE 8 |** *nts* and *nos1* expression in the posterior recess region is lost in *bsx* mutant embryos. Lateral **(a,b,e,f)** and dorsal views **(c,d,g,h)** of the ventral forebrain in 3 dpf embryos stained by *in situ* hybridization using probes as indicated. All images are minimum intensity projections of 40 brightfield focal planes (distance 1 μm). Scale bar in **(a)** for all images: 50 μm. Anterior at left. Schematics to the right show location of *nts* (red) or *nos1* (green) expressing cells as observed in stainings and indicate the focal planes which are shown in dorsal view pictures and schematics. Cells with Bsx-dependent expression are circled. Numbers *n* (phenotype shown)/*n* (total analyzed) as indicated.

implicated in energy homeostasis and feeding regulation as well as fertility and reproduction (Leshan et al., 2012; Sutton et al., 2014; Chachlaki et al., 2017). We detected widespread expression of *nos1* in the PRR reaching from medial mamillary regions to lateral Tub regions, which, however, is absent in *bsx* mutants at 3 dpf (**Figures 8e–h**, arrowheads). At 4 dpf we found few *nos1* expressing cells in the more laterally located tuberal regions of *bsx* mutants while the more medial cells in mamillary regions remained absent (**Supplementary Figures S11E–H**, arrowheads). Another *nos1* expression domain, consisting of cells spread along the "PTv" and RTuD/PBas region, was unaffected by Bsx loss both at 3 dpf (**Figures 8e–h**, arrowheads) and 4 dpf (**Supplementary Figures S11E–H**, arrowheads), indicating that *nos1* expression in the PRR, but not the "PTv" or RTuD/PBas region, strictly depends on Bsx. We conclude that Bsx functions in the differentiation of the nitric oxide and neurotensin neuromodulatory systems in the PRR.

## DISCUSSION

In wildtype and *bsx* mutant embryos, we analyzed the expression of different monoamine pathway components and multiple neuropeptidergic precursor genes, for many of which the expression in the zebrafish forebrain has not previously been described in detail. We demonstrate that Bsx is necessary not only for normal expression of a surprisingly large number of peptidergic neuromodulators, but also for the expression of enzymes involved in the synthesis of small neuromodulatory molecules, such as histamine and nitric oxide, and for normal serotonin levels in the hypothalamus (**Figure 9**). Even though the patterning and gross anatomical structure of the secondary prosencephalon is unchanged in *bsx* mutants, they do not develop the full complement of neuromodulatory molecules in the forebrain, which might have strong implications on physiology and behavior of animals devoid of functional Bsx.

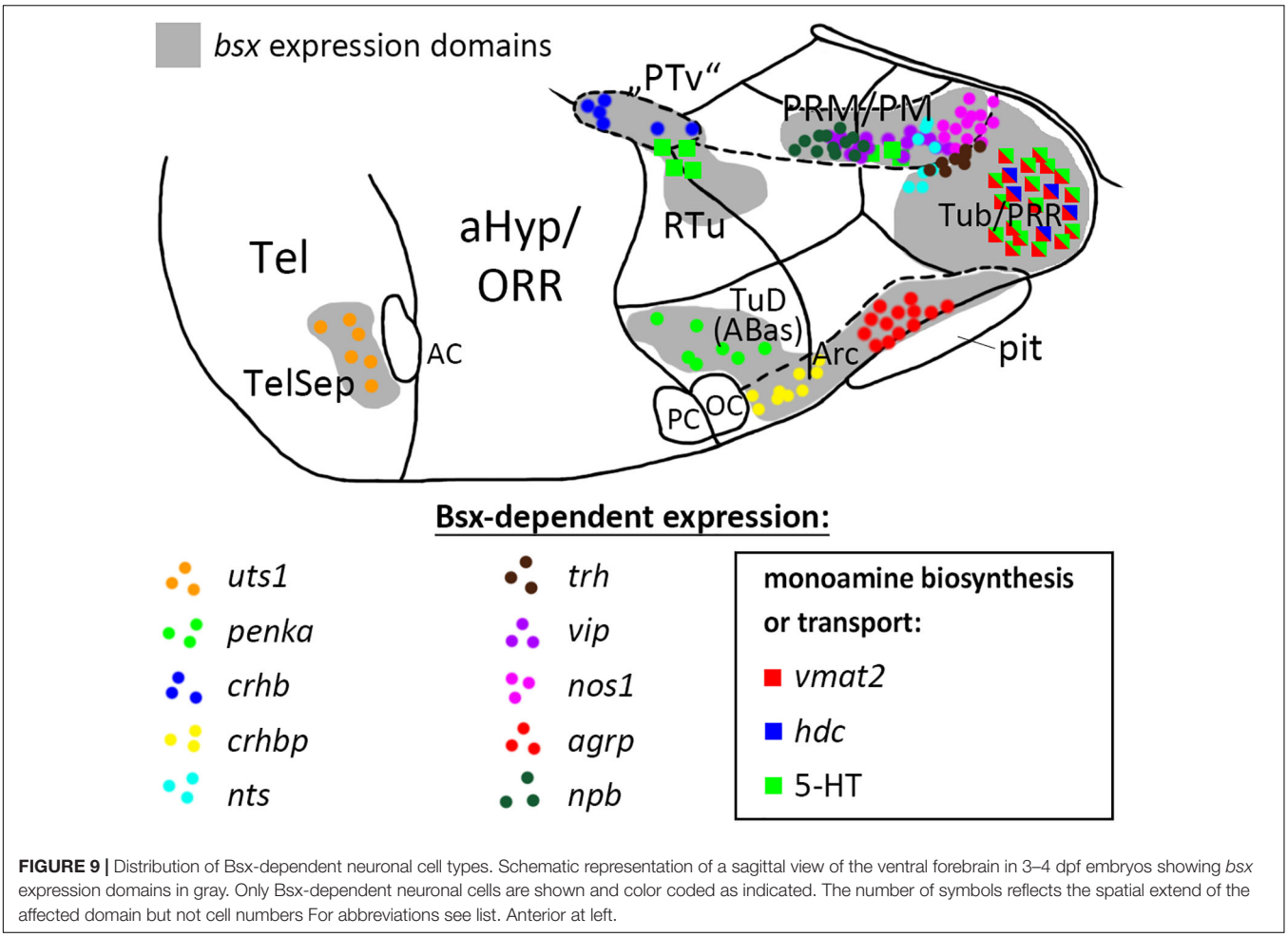


**FIGURE 9 |** Distribution of Bsx-dependent neuronal cell types. Schematic representation of a sagittal view of the ventral forebrain in 3–4 dpf embryos showing *bsx* expression domains in gray. Only Bsx-dependent neuronal cells are shown and color coded as indicated. The number of symbols reflects the spatial extend of the affected domain but not cell numbers For abbreviations see list. Anterior at left.

Our data reveal that all *bsx* expression in the basal prosencephalon depends on combined activity of *nkx2.1* and *nkx2.4a/b* genes in zebrafish. Expression of these three genes is maintained during all phases of *bsx* expression studied here (Manoli and Driever, 2014), thus it is possible that Nkx2.1 and 2.4 transcription factors may directly control *bsx* expression. However, embryos devoid of *nkx2.*1 and *nkx2.4a/b* develop severe patterning abnormalities in the hypothalamus with lack of most or all hypothalamic regions in which *bsx* will be expressed (Manoli and Driever, 2014). Therefore, also indirect regulation is possible. Downregulation of *bsx* expression in the ARC upon loss of the Otp transcription factor has recently been shown in mice (Lee et al., 2018), and our data reveal conserved Otp regulation of *bsx* in teleosts. It is unclear how well conserved this gene regulatory network is beyond the ARC, as for *Otp* knock-out mice no reports exist on *Bsx* expression in those areas of the mouse bHyp for which we demonstrate dependency on Otp factors in zebrafish. Other functions of Otp, in contrast, appear to be Bsx-independent, since expression of *th* and *avp* is reduced in *otp* mutants (Ryu et al., 2007; Fernandes et al., 2013), but not in *bsx* mutants.

We recently demonstrated prosomeric organization of the bHyp to be highly conserved in mammalian and teleost embryos (Schredelseker and Driever, 2020). The homology relations of hypothalamic subregions that we proposed in this model are also supported by the expression domains of many neuropeptidergic and aminergic genes that were assessed in the present study. This combined information may also help to understand additional anatomical features. For instance, the location of commissures, visible in stained embryos by mere tissue contrast, is in line with our anatomical analysis and shall be illustrated by one example. We proposed large parts of what was previously suggested to belong to the "PTv" (thus being a basal part of prosomer 3), to actually correspond to mamillary hypothalamic regions (Schredelseker and Driever, 2020). We should thus reconsider the anatomical status (and the name) of the TCPT (Wilson et al., 1990), and, based on its location, hypothesize this commissure to correspond to the retromamillary commissure in mouse (Puelles et al., 2012).

The high degree of conservation in both protein structure and expression pattern of Bsx suggests that many of the herein shown Bsx functions in neuronal differentiation might be conserved in other phyla. However, in demonstrating the loss of monoaminergic markers in the PRR, we also describe Bsx functions which might have undergone substantial evolutionary divergence. The special phylogenetic status of the PRR has

been discussed previously (Vernier, 2017; Xavier et al., 2017; Schredelseker and Driever, 2020). The dependence of the monoaminergic CSF-c cells that densely populate the PRR on Bsx might help elucidating the evolutionary relationship of these neurons across vertebrates.

With respect to neuropeptides, we find conservation of *agrp* regulation by Bsx in zebrafish, as reported for mice (Sakkou et al., 2007). As of yet, no reports exist on Bsx functions in other subregions of the developing hypothalamus beyond the ARC. Here, we present Bsx functions in the differentiation of monoaminergic and neuropeptidergic cell types within several subregions of both, the tuberal and mamillary hypothalamus, and the telencephalic septum (summarized in **Figure 9**, only cell types affected in *bsx* mutants are shown). It should be noted that for all regions in which we found a neuronal differentiation gene not to be expressed in *bsx* mutants, we found other marker genes to be unaffected by Bsx loss. Moreover, for each expressed neuronal differentiation gene affected by Bsx loss in specific regions, we found expression of the same neuronal differentiation gene being not regulated by Bsx in other regions. Together, these findings support the conclusion that the differentiation program for the hypothalamus is regulated in a highly combinatorial manner, and that Bsx functions are context-dependent, rather than Bsx being a specific determinant for a defined cell type. In the case of *penka*, *pmchl, avp*, *vip* and *npb*, we observed, in some embryos, ectopic expression and/or a seemingly denser *in situ* hybridization staining intensity in certain Bsx-independent clusters, which, however, we did not quantify. If confirmed in future studies, this would indicate that Bsx regulates fate switches in the differentiation programs of neuromodulatory cells in the hypothalamus, such that loss of one cell type may correspond to more cells of another type. Bsx function appears to be limited to differentiation rather than patterning mechanisms, as the gross anatomy of the bHyp is unchanged in *bsx* mutants. Notably, Jones and McGinnis previously hypothesized in 1993 that the *Drosophila bsx* homolog *bsh* "is not required for the actual morphology of brain cells, the axonal pathways, etc., but instead regulates expression of neurotransmitters, receptors, channels, synaptic specializations, or other effector molecules that are critical to the physiologic function of particular neural cells" (Jones and McGinnis, 1993).

We were surprised by the strong effects of Bsx loss on histaminergic and serotonergic differentiation in the hypothalamus. Two different Ets domain transcription factors, Pet1 and Etv5b, specify Raphe nucleus and hypothalamic serotonergic neurons, respectively (Hendricks et al., 2003; Bosco et al., 2013). Here, we show that Bsx is an additional component of the transcriptional network that specifies serotonergic differentiation in the hypothalamus, but not the Raphe nucleus. Since in mammals serotonergic somata became restricted to the Raphe, functional characterization of the hypothalamic serotonergic system is sparse. Recent work suggests that the activity of hypothalamic serotonergic cells correlates with hunger states in zebrafish (Wee et al., 2019). The histaminergic system has been well described in zebrafish, but transcriptional regulators of its development have not been reported so far (Chen et al., 2016).

While *bsx* is also expressed in a subset of dopaminergic neurons in the mamillary hypothalamus, the expression of the dopaminergic marker genes *th* and *th2* is normal in *bsx* mutants. DC2, 4, 5, and 6 dopaminergic cells and adjacent neuropeptidergic cell types, including *crhb* expressing neurons, are specified by a shared transcriptional network (Löhr et al., 2009). So far, no transcription factors specific for only one of those two lineages have been found. In demonstrating that *th* expression is unaffected by Bsx loss, while adjacent *crhb* expression is lost, we identify Bsx as a factor that separates those lineages. A model has been put forward that neurons, particularly in the hypothalamus, may frequently express more than one neurotransmitter type, including neuropeptide precursor peptides (Romanov et al., 2019). Therefore, it is tempting to speculate that Bsx may be required in dopaminergic neurons to establish the full complement of neuromodulatory factors which they might express in addition to *th*.

It has recently been suggested that the functional homolog of the *crhb* expressing neurons in the mammalian PVN, which stimulate corticotropes in the anterior pituitary, are located in the neurosecretory preoptic area (Herget et al., 2014), a well characterized region densely populated by various neuropeptidergic cell types within the aHyp/ORR. In contrast, the functions of the Bsx-dependent clusters expressing *crhb*, *crhbp* and *uts1* in the "PTv," PRR and TelSep region, respectively, are poorly understood. *uts1* was the only gene for which we found TelSep expression to be Bsx-dependent, providing a first description of a Bsx function in the telencephalon. However, we also found several other genes encoding neuropeptidergic precursors to be expressed in this region, suggesting that the TelSep might act as a neuromodulatory center in zebrafish.

Even though a function in feeding regulation has been suggested for many of the analyzed neuropeptides, they are known to exert multiple functions beyond energy homeostasis, ranging from stress response behavior to the regulation of sleep and arousal. While so far phenotypes of *Bsx* knock-out animals, such as high body fat and reduced locomotion (Sakkou et al., 2007), were discussed only in the light of Bsx as a regulator of *Agrp* and *Npy* expression, alternative explanations should be considered given the breadth of neuropeptides affected.

While Bsx in mouse has been shown to directly bind to the promoters of the *Agrp* and *Npy* genes, we do not know whether this is the case in zebrafish for *agrp, npy,* or any of the other genes for which we found expression to depend on Bsx. It might well be that expression of some of the affected genes in the expected areas was not detected due to a failure of precursors or neurons to migrate or survive in *bsx* mutants. Such roles have recently been described for the Otp transcription factor, which shall be used as an example to illustrate the highly combinatorial action of transcription factors in hypothalamus development. In *Otp* knock-out mice, *Trh*, *Crh*, *Avp,* and *Oxt* and *Sst* expression is lost in specific subregions of the aHyp (Acampora et al., 1999). In the ARC of *Otp* mutant mice, *Agrp*, *Npy* and *Sst* expression is absent while *Ghrh* and *Pomc* expression is normal (Acampora et al., 1999; Lee et al., 2018). A careful re-examination of *Otp* mutant mice revealed that all *Ghrh* expressing cells in the bHyp which originate in alar

territories, but migrate to the VMH region, are absent upon Otp loss, while the intrinsically basal *Ghrh* expression in the ARC is unaffected (Morales-Delgado et al., 2014). Unfortunately, we could not analyze *ghrh* expression in *bsx* mutant zebrafish since we could not detect any signal at the embryonic stages analyzed when using a *ghrh* probe.

Since serotonin, Vip and Trh have been implicated in lactation in mammals (Kato et al., 1978; Crowley, 2014), the defects in these systems might contribute to the nursing defects which were described for *Bsx* mutant female mice (McArthur and Ohtoshi, 2007). The hypothalamic nitric oxide system has been linked to ovarian cyclicity and the onset of puberty (Chachlaki et al., 2017) and, notably, *BSX* was identified as a locus for age at menarche in GWAS (Elks et al., 2010; Demerath et al., 2013). Moreover, our data might help to reevaluate endocrine features of patients suffering from Jacobsen syndrome, a disease in which chromosome 11 deletions which encompass the *BSX* locus, were implicated (Haghi et al., 2004; Coldren et al., 2009).

# MATERIALS AND METHODS

## Zebrafish Strains, Maintenance and Heat Shock Treatment

Zebrafish embryos of the ABTL strain were obtained through natural breeding, incubated in constant darkness at 28°C, and staged according to Kimmel et al. (1995). The following transgenic or mutant lines were used: *Tg(hsp70l:otpa-IRES-gfp-caax)*$^{m1178}$ (generated by Heiko Löhr, Driever Lab), *bsx*$^{m1376}$ (Schredelseker and Driever, 2018), *otpa*$^{m866}$ (Ryu et al., 2007), *otpb*$^{sa115}$ (Fernandes et al., 2012), *nkx2.1*$^{m1355}$, *nkx2.4a*$^{m1354}$, *nkx2.4b*$^{m1353}$ (this study). All mutant embryos were generated by incrossing heterozygous parent animals. Heat shock was performed by keeping petri dishes with zebrafish embryos in a 39°C water bath for 1 h. Heat shock transgenic embryos were identified through GFP signal by brightfield fluorescent microscopy. All experiments were carried out in accordance with the German Animal Welfare Act.

## Generation of the Heat Shock-Inducible *otpa* Overexpressing Line

The open reading frame of the *otpa* gene (NCBI RefSeq NM_001128703.1) was PCR-amplified from a *pCS2-otpa* plasmid (Ryu et al., 2007) as a template with the following primers containing attB1/2 sites (in bold):

5′**GGGGACAAGTTTGTACAAAAAAGCAGGCT**CGATGCTTTCGCATGCCGACCTGCTG3′,

5′**GGGGACCACTTTGTACAAGAAAGCTGGGT**ATTAGGTGAAGCTCATGGACACTGTG3′

(start and stop codons underlined). PCR-products were combined with a *pDONR221* (Kwan et al., 2007) donor vector in a Gateway BP recombination reaction (Thermo Fisher) to generate the middle entry clone *pME-otpa* which was then recombined with a *p5E-hsp70l* and *p3E-IRES-egfp-caax-pA* onto a *pDestTol2CG2* [for all three plasmids see Kwan et al. (2007)] in a Gateway LR reaction (Thermo Fisher). The resulting *hsp70l-otpa-IRES-egfp-caax-pA-CG2* construct containing Tol2 recombination sites was injected into single-cell-staged embryos together with *transposase* mRNA, which was transcribed *in vitro* from *pCS2FA-transposase* (Kwan et al., 2007) using the mMessage mMachine Sp6 Transcription Kit (Thermo Fisher). The *cmlc2:gfp* cassette (CG2) present on the construct allowed easy screening for germline line transgenic animals through GFP signal in the heart.

## Targeted Mutagenesis of *nkx2.1* and *nkx2.4a/b* Alleles

TALEN target sites were identified for Exon 1 of *nkx2.1* (NCBI Gene ID: 81883), *nkx2.4a* (NCBI Gene ID: 562300) or *nkx2.4b* (NCBI Gene ID: 58112) gene using the Mojo Hand design tool, Version 1.5 (Neff et al., 2013). The following TALEN plasmids were assembled using the Golden Gate TALEN and TAL Effector Kit 2.0 onto an RCI-Script-GoldyTALEN backbone (addgene.org), (Cermak et al., 2011): *nkx2.1* left – GGAGTTTCCCAGCTGTC; *nkx2.1* right – GGGCTACTGTAACGGG (binding strand: CCC GTTACAGTAGCCC); *nkx2.4a* left: GGAAACCTCACATCTCC; *nkx2.4a* right – GCAACCTCAGGTGTCCC (binding strand: GGGACACCTGAGGTTGC); *nkx2.4b* left – GAGAAACGGCGCCACGG; *nkx2.4b* right – GCTCGAA CCCGGAGCCG (binding strand: CGGCTCCGGGT TCGAGC). TALEN assembly products were verified by DNA sequencing and restriction digest with BamHI (NEB) and BsaI (NEB). Purified products were used for *in vitro* transcription of mRNA using the mMessage mMachine T3 Transcription Kit (Thermo Fisher Scientific). For *nkx2.4a* or *nkx2.4b* 200-300 pg of each TALEN mRNA were injected into wildtype one-cell stage embryos. Efficiency of TALEN-mediated indel generation was assayed through PCR and restriction enzyme digest on pooled injected embryos 3 dpf. Potential mutant alleles of F1 animals were sequenced and frameshift indels identified. *nkx2.4a*$^{m1354}$, *nkx2.4b*$^{m1353}$ double heterozygous F2 animals were raised and identified through genotyping. Since *nkx2.4a* and *nkx2.1* are both located on chromosome 17, we introduced the *nkx2.1* mutation *de novo* in the *nkx2.4a*$^{m1354}$, *nkx2.4b*$^{m1353}$ fish. 200–300 pg *nkx2.1* TALEN mRNA was injected into one cell stage embryos from crosses of *nkx2.4a*$^{m1354}$, *nkx2.4b*$^{m1353}$ fish crossed with ABTL fish. A new mutant *nkx2.4b*$^{m1353}$ allele with a small deletion generating a frameshift was identified by sequencing in F2 animals. Triple heterozygous *nkx2.1*$^{m1355}$, *nkx2.4a*$^{m1354}$, *nkx2.4b*$^{m1353}$ animals were genotyped and crossed to generate *nkx2.1, nkx2.4a, nkx2.4b* triple homozygous embryos.

## Genotyping

All breeding adult animals and experimental embryos were genotyped through PCR, followed by restriction digest. Fin or tail biopsies were lysed in 50 mM NaOH for 45 min at 95°C, then neutralized with 1/10 volume of 1M Tris-HCl pH 7.5 before being used as template. Primer sequences, amplicon length, restriction enzymes and length of digestion products are given in **Supplementary Table S2**.

**TABLE 1 |** Abbreviations, full names, and zfin.org Gene ID of genes in this study.

| Abbreviation | Full gene name | zfin.org Gene ID |
|---|---|---|
| *agrp* | *agouti related neuropeptide* | ZDB-GENE-040817-1 |
| *avp* | *arginine vasopressin* | ZDB-GENE-030407-2 |
| *bsx* | *brain-specific homeobox* | ZDB-GENE-040628-4 |
| *cart4* | *cocaine- and amphetamine-regulated transcript 4* | ZDB-GENE-060503-863 |
| *crh* | *corticotropin releasing hormone b* | ZDB-GENE-041114-75 |
| *crhbp* | *corticotropin releasing hormone binding protein* | ZDB-GENE-040801-196 |
| *dlx5a* | *distal-less homeobox 5a* | ZDB-GENE-990415-49 |
| *emx2* | *empty spiracles homeobox 2* | ZDB-GENE-990415-54 |
| *fezf2* | *FEZ family zinc finger 2* | ZDB-GENE-001103-3 |
| *foxb1a* | *forkhead box B1a* | ZDB-GENE-990616-47 |
| *galn* | *galanin/GMAP prepropeptide* | ZDB-GENE-111117-2 |
| *ghrh* | *growth hormone releasing hormone* | ZDB-GENE-070426-1 |
| *hcrt* | *hypocretin (orexin) neuropeptide precursor* | ZDB-GENE-040324-1 |
| *hdc* | *histidine decarboxylase* | ZDB-GENE-080102-5 |
| *isl1* | *ISL LIM homeobox 1* | ZDB-GENE-980526-112 |
| *lef1* | *lymphoid enhancer-binding factor 1* | ZDB-GENE-990714-26 |
| *lhx5* | *LIM homeobox 5* | ZDB-GENE-980526-484 |
| *lhx6* | *LIM homeobox 6* | ZDB-GENE-041025-1 |
| *lhx9* | *LIM homeobox 9* | ZDB-GENE-050417-210 |
| *nkx2.1* | *NK2 homeobox 1* | ZDB-GENE-010404-1 |
| *nkx2.4a* | *NK2 homeobox 4a* | ZDB-GENE-030131-6336 |
| *nkx2.4b* | *NK2 homeobox 4b* | ZDB-GENE-000830-1 |
| *nmu* | *neuromedin U* | ZDB-GENE-041001-111 |
| *nos1* | *nitric oxide synthase 1 (neuronal)* | ZDB-GENE-001101-1 |
| *npb* | *neuropeptide B* | ZDB-GENE-040107-40 |
| *npvf* | *neuropeptide VF precursor* | ZDB-GENE-070424-226 |
| *npy* | *neuropeptide y* | ZDB-GENE-980526-438 |
| *nts* | *neurotensin* | ZDB-GENE-091204-433 |
| *nr5a1a* | *nuclear receptor subfamily 5, group A, member 1a* | ZDB-GENE-010504-1 |
| *nr5a2* | *nuclear receptor subfamily 5, group A, member 2* | ZDB-GENE-990415-79 |
| *otpa* | *orthopedia homeobox a* | ZDB-GENE-070216-1 |
| *otpb* | *orthopedia homeobox b* | ZDB-GENE-990708-7 |
| *pax6a* | *paired box 6a* | ZDB-GENE-990415-200 |
| *pax7a* | *paired box 7a* | ZDB-GENE-990415-201 |
| *pdyn* | *prodynorphin* | ZDB-GENE-060417-1 |
| *penka* | *proenkephalin a* | ZDB-GENE-030729-31 |
| *pmch* | *pro-melanin-concentrating hormone* | ZDB-GENE-041210-150 |
| *pmchl* | *pro-melanin-concentrating hormone, like* | ZDB-GENE-030131-7863 |
| *pomca* | *proopiomelanocortin a* | ZDB-GENE-030513-2 |
| *shha* | *sonic hedgehog signaling molecule a* | ZDB-GENE-980526-166 |
| *sst1.1* | *somatostatin 1, tandem duplicate 1* | ZDB-GENE-030131-4743 |
| *th* | *tyrosine hydroxylase* | ZDB-GENE-990621-5 |
| *th2* | *tyrosine hydroxylase 2* | ZDB-GENE-050201-1 |
| *trh* | *thyrotropin-releasing hormone* | ZDB-GENE-020930-1 |
| *uts1* | *urotensin 1* | ZDB-GENE-041014-348 |
| *vip* | *vasoactive intestinal peptide* | ZDB-GENE-080204-3 |
| *vmat2 = slc18a2* | *vesicular monoamine transporter 2 = solute carrier family 18 member 2* | ZDB-GENE-080514-1 |

*All genes for which antisense RNA probes were used during in situ hybridization. More detail is given in* ***Supplementary Table S1****.*

## Cloning of cDNA Fragments Used as Templates for *in situ* Hybridization

Gene fragments were PCR amplified from cDNA and cloned into pCRII-TOPO (Invitrogen) as has been described (Schredelseker and Driever, 2018) using primer sequences shown in **Supplementary Table S1**. Plasmids were linearized with restriction enzymes (**Supplementary Table S1**) and used as template for *in vitro* transcription of digoxigenin (DIG)-labeled probe using T3, T7, or SP6 RNA polymerase [**Supplementary Table S1**; Schredelseker and Driever (2018)].

## *In situ* Hybridization and Immunohistochemistry

Gene IDs for all antisense probes generated are provided in **Table 1**. Whole-mount DIG-labeled RNA *in situ* hybridization based on alkaline phosphatase reaction was carried out as previously described (Schredelseker and Driever, 2018). Tyramide signal amplification (TSA) fluorescent *in situ* hybridization (FISH) was carried out as described previously (Ronneberger et al., 2012). For combined FISH/fluorescent immunohistochemistry (FIHC), after detection of a single DIG-labeled probe through TSA reaction using Alexa Fluor 488, peroxidase activity was inactivated through extensive washing in TNT [100 mM Tris-HCl pH7.5, 150 mM NaCl, 0.5% Tween20] and incubation in 1% $H_2O_2$/TNT for 30 min at room temperature (RT). Embryos were then washed in PBTD [1% DMSO/PBST (0.1% Tween20 in PBS)] and blocked for 2 h in Blocking Solution [5% goat serum (Sigma-Aldrich #G9023), 1% blocking reagent (Roche, #1096176), 1% BSA (Sigma-Aldrich #A6003) in PBTD] at RT. Embryos were incubated with primary rabbit anti 5-HT antibody (Sigma-Aldrich #S5545; 1:200), primary mouse HuC/HuD antibody (16A11, Thermo Fisher Scientific #A-21271; 1:500) or preabsorbed primary rabbit anti-Th antibody (Ryu et al., 2007; 1:500) 1:500 in Blocking Solution over night at 4°C. Embryos were washed all day in PBTD at RT, then incubated with secondary goat anti-rabbit IgG Alexa 488 or 546 antibody (Thermo Fisher Scientific, #A-11008 or #A-11035; 1:1000) and secondary goat anti-mouse Alexa 555 antibody (Thermo Fisher Scientific, #A-21425; 1:1000) in PBTD over night at 4°C. Embryos were washed several times in PBTD and PBST, then transferred to 80% glycerol/PBST and stored at 4°C until imaged. All mutant embryos came from incrosses of

heterozygous parent animals and were stained in the same tube and/or wells together with their wildtype siblings, and, after the staining procedure, but before imaging, were genotyped based on tail clip DNA. Heat shock transgenic embryos, identified by GFP signal, were presorted and stained in tubes and/or wells separate from their wildtype siblings, but the staining procedure was performed in parallel, with all solutions being applied at the same time and from the same stock.

For all previously unpublished probes we performed *in silico* alignment analysis using the BLASTN discontiguous megablast algorithm on the *Danio rerio* nucleotide collection (taxid: 7955) with default parameters and for most probe sequences found no hits except the specific transcript which we aimed to target. There are, however, three exceptions. Our *nos1* probe has some similarity to a *nos2a* transcript (XM_005165296.4). However, *nos2a* was found not to be expressed in the larval zebrafish until 5 dpf (Thisse and Thisse, 2005). For *nr5a1a* we cannot exclude hybridization of our probe to *nr5a1b* (NM_212834.1) or *nr5a2* (NM_001313729.1). Since we used *nr5a1a* staining only to demonstrate that forebrain patterning is normal, we concluded potential hybridization to those other genes to be rather inconsequential for our statement. Our *vmat2* probe also has some similarity to a *vmat1* transcript (XM_021478484.1), which, however, is not expressed in the larval zebrafish brain (Puttonen et al., 2017). We also found the expression patterns generated by all *in situ* hybridization probe sequences cloned in this study to match the expression patterns of the same gene as has been published previously (for references see **Supplementary Table S1**).

### Imaging and Figure Preparation

Alkaline phosphatase blue *in situ* hybridization stained embryos were mounted for brightfield microscopy in 80% glycerol, 20% PBST, 1 mM EDTA. Images were obtained using an AxioCam CC1 on an AxioPlan2 microscope with a PLAN-NEOFLUAR 20×/0.5 or 10×/0.3 objective and DIC optics using the AxioVs40 Software (all Zeiss). From image stacks (1 μm step size), minimum intensity projections of *z*-planes as indicated in the figure legends were generated using ImageJ. Embryos stained by fluorescent *in situ* hybridization and immunohistochemistry were recorded using a Zeiss LSM 510 or LSM 880. All figures were assembled using Adobe Photoshop CS4 or CS6. When linear adjustment of levels was made, histograms were clipped to the same values for experimental and control images.

## DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

## ETHICS STATEMENT

The animal study was reviewed and approved by Regierungspräsidium Freiburg, Freiburg, Germany.

## AUTHOR CONTRIBUTIONS

TS conceptualized and designed the study and performed the experiments, assembled the figures, and wrote the first draft of the manuscript. FV generated the *nkx2.1* and *nkx2.4a/b* mutant lines. RD communicated unpublished data and suggested hypothalamic markers. WD contributed to design, supervision and editing, and provided project administration and funding acquisition.

## FUNDING

This study was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2189 – Project ID: 390939984 [Gefördert durch die Deutsche Forschungsgemeinschaft (DFG) im Rahmen der Exzellenzstrategie des Bundes und der Länder – EXC-2189 – Projektnummer 390939984] and the Excellence Initiative of the German Federal and State Governments (BIOSS - EXC 294).

## ACKNOWLEDGMENTS

Many thanks go to Nadja Bischoff, Daniel Armbruster, Daniela Reuter-Schmitt, and Dörte Meyer for technical help with the staining procedures or genotyping. Many thanks go to Eva Carl for cloning the *fezf2* plasmid, to Jochen Holzschuh for cloning the *vmat2* plasmid and to Pooja Sant for cloning the *npb* plasmid used for probe synthesis. Many thanks go to Heiko Löhr for generating the *Tg(hsp70l:otpa-IRES-gfp-caax)*$^{m1178}$ line and for discussions. Many thanks go to Matthias Hammerschmidt for discussions and for providing the *agrp*, *pomca*, *hcrt*, *nr5a2* and *npy* plasmids. Many thanks go to Sabine Götter for excellent fish care and to the Life Imaging Center Freiburg for technical support with confocal microscopy.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins.2020.00525/full#supplementary-material

## REFERENCES

Acampora, D., Postiglione, M. P., Avantaggiato, V., Di Bonito, M., Vaccarino, F. M., Michaud, J., et al. (1999). Progressive impairment of developing neuroendocrine cell lineages in the hypothalamus of mice lacking the Orthopedia gene. *Genes Dev.* 13, 2787–2800. doi: 10.1101/gad.13.21.2787

Affaticati, P., Yamamoto, K., Rizzi, B., Bureau, C., Peyrieras, N., Pasqualini, C., et al. (2015). Identification of the optic recess region as a morphogenetic entity in the zebrafish forebrain. *Sci. Rep.* 5:8738. doi: 10.1038/srep08738

Alié, A., Devos, L., Torres-Paz, J., Prunier, L., Boulet, F., Blin, M., et al. (2018). Developmental evolution of the forebrain in cavefish, from natural variations in neuropeptides to behavior. *eLife* 7:e32808. doi: 10.7554/eLife.32808

Alvarez-Bolado, G. (2019). Development of neuroendocrine neurons in the mammalian hypothalamus. *Cell Tissue Res.* 375, 23–39. doi: 10.1007/s00441-018-2859-1

Bedont, J. L., Newman, E. A., and Blackshaw, S. (2015). Patterning, specification, and differentiation in the developing hypothalamus. *Wiley Interdiscip. Rev. Dev. Biol.* 4, 445–468. doi: 10.1002/wdev.187

Bellipanni, G., Rink, E., and Bally-Cuif, L. (2002). Cloning of two tryptophan hydroxylase genes expressed in the diencephalon of the developing zebrafish brain. *Mech. Dev.* 119(Suppl. 1), S215–S220. doi: 10.1016/s0925-4773(03)00119-9

Bosco, A., Bureau, C., Affaticati, P., Gaspar, P., Bally-Cuif, L., and Lillesaar, C. (2013). Development of hypothalamic serotoninergic neurons requires Fgf signalling via the ETS-domain transcription factor Etv5b. *Development* 140, 372–384. doi: 10.1242/dev.089094

Burbridge, S., Stewart, I., and Placzek, M. (2016). Development of the neuroendocrine hypothalamus. *Compr. Physiol.* 6, 623–643. doi: 10.1002/cphy.c150023

Cardoso, J. C. R., Bergqvist, C. A., Félix, R. C., and Larhammar, D. (2016). Corticotropin-releasing hormone family evolution: five ancestral genes remain in some lineages. *J. Mol. Endocrinol.* 57, 73–86. doi: 10.1530/JME-16-0051

Cecchi, C. (2002). Emx2: a gene responsible for cortical development, regionalization and area specification. *Gene* 291, 1–9. doi: 10.1016/s0378-1119(02)00623-6

Cermak, T., Doyle, E. L., Christian, M., Wang, L., Zhang, Y., Schmidt, C., et al. (2011). Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res.* 39:e82. doi: 10.1093/nar/gkr218

Chachlaki, K., Garthwaite, J., and Prevot, V. (2017). The gentle art of saying NO: how nitric oxide gets things done in the hypothalamus. *Nat. Rev. Endocrinol.* 13, 521–535. doi: 10.1038/nrendo.2017.69

Chen, Y.-C., Semenova, S., Rozov, S., Sundvik, M., Bonkowsky, J. L., and Panula, P. (2016). A novel developmental role for dopaminergic signaling to specify hypothalamic neurotransmitter identity. *J. Biol. Chem.* 291, 21880–21892. doi: 10.1074/jbc.M115.697466

Coldren, C. D., Lai, Z., Shragg, P., Rossi, E., Glidewell, S. C., Zuffardi, O., et al. (2009). Chromosomal microarray mapping suggests a role for BSX and Neurogranin in neurocognitive and behavioral defects in the 11q terminal deletion disorder (Jacobsen syndrome). *Neurogenetics* 10, 89–95. doi: 10.1007/s10048-008-0157-x

Cremona, M., Colombo, E., Andreazzoli, M., Cossu, G., and Broccoli, V. (2004). Bsx, an evolutionary conserved Brain Specific homeoboX gene expressed in the septum, epiphysis, mammillary bodies and arcuate nucleus. *Gene Expr. Patterns* 4, 47–51. doi: 10.1016/s1567-133x(03)00151-0

Crowley, W. R. (2014). Neuroendocrine regulation of lactation and milk production. *Compr. Physiol.* 5, 255–291. doi: 10.1002/cphy.c140029

D'Autilia, S., Broccoli, V., Barsacchi, G., and Andreazzoli, M. (2010). Xenopus Bsx links daily cell cycle rhythms and pineal photoreceptor fate. *Proc. Natl. Acad. Sci. U.S.A.* 107, 6352–6357. doi: 10.1073/pnas.1000854107

Demerath, E. W., Liu, C.-T., Franceschini, N., Chen, G., Palmer, J. R., Smith, E. N., et al. (2013). Genome-wide association study of age at menarche in African-American women. *Hum. Mol. Genet.* 22, 3329–3346. doi: 10.1093/hmg/ddt181

Denver, R. J. (2009). Structural and functional evolution of vertebrate neuroendocrine stress systems. *Ann. N. Y. Acad. Sci.* 1163, 1–16. doi: 10.1111/j.1749-6632.2009.04433.x

Díaz, C., Morales-Delgado, N., and Puelles, L. (2014). Ontogenesis of peptidergic neurons within the genoarchitectonic map of the mouse hypothalamus. *Front. Neuroanat.* 8:162. doi: 10.3389/fnana.2014.00162

Dobner, P. R., and Carraway, R. E. (2013). "Chapter 117 - Neurotensin/Neuromedin N," in *Handbook of Biologically Active Peptides*, 2nd Edn, ed. A. J. Kastin, (Boston, MA: Academic Press), 875–882.

Dominguez, L., Gonzalez, A., and Moreno, N. (2015). Patterns of hypothalamic regionalization in amphibians and reptiles: common traits revealed by a genoarchitectonic approach. *Front. Neuroanat.* 9:3. doi: 10.3389/fnana.2015.00003

Dvorakova, C. M. (2018). Distribution and function of neuropeptides W/B signaling system. *Front Physiol* 9:981. doi: 10.3389/fphys.2018.00981

Eaton, J. L., and Glasgow, E. (2007). Zebrafish orthopedia (otp) is required for isotocin cell development. *Dev. Genes Evol.* 217, 149–158. doi: 10.1007/s00427-006-0123-2

Elks, C. E., Perry, J. R. B., Sulem, P., Chasman, D. I., Franceschini, N., He, C., et al. (2010). Thirty new loci for age at menarche identified by a meta-analysis of genome-wide association studies. *Nat. Genet.* 42, 1077–1085. doi: 10.1038/ng.714

Eriksson, K. S., Peitsaro, N., Karlstedt, K., Kaslin, J., and Panula, P. (1998). Development of the histaminergic neurons and expression of histidine decarboxylase mRNA in the zebrafish brain in the absence of all peripheral histaminergic systems. *Eur. J. Neurosci.* 10, 3799–3812. doi: 10.1046/j.1460-9568.1998.00394.x

Fernandes, A. M., Beddows, E., Filippi, A., and Driever, W. (2013). Orthopedia transcription factor otpa and otpb paralogous genes function during dopaminergic and neuroendocrine cell specification in larval Zebrafish. *PLoS One* 8:e75002. doi: 10.1371/journal.pone.0075002

Fernandes, A. M., Fero, K., Arrenberg, A. B., Bergeron, S. A., Driever, W., and Burgess, H. A. (2012). Deep brain photoreceptors control light-seeking behavior in zebrafish larvae. *Curr. Biol.* 22, 2042–2047. doi: 10.1016/j.cub.2012.08.016

Ferran, J. L., Puelles, L., and Rubenstein, J. L. (2015). Molecular codes defining rostrocaudal domains in the embryonic mouse hypothalamus. *Front. Neuroanat.* 9:46. doi: 10.3389/fnana.2015.00046

Filippi, A., Mahler, J., Schweitzer, J., and Driever, W. (2010). Expression of the paralogous tyrosine hydroxylase encoding genes th1 and th2 reveals the full complement of dopaminergic and noradrenergic neurons in zebrafish larval and juvenile brain. *J. Comp. Neurol.* 518, 423–438. doi: 10.1002/cne.22213

Haghi, M., Dewan, A., Jones, K. L., Reitz, R., Jones, C., and Grossfeld, P. (2004). Endocrine abnormalities in patients with Jacobsen (11q-) syndrome. *Am. J. Med. Genet. Part A* 129A, 62–63. doi: 10.1002/ajmg.a.30248

Hara, J., Gerashchenko, D., Wisor, J. P., Sakurai, T., Xie, X., and Kilduff, T. S. (2009). Thyrotropin-releasing hormone increases behavioral arousal through modulation of hypocretin/orexin neurons. *J. Neurosci.* 29, 3705–3714. doi: 10.1523/JNEUROSCI.0431-09.2009

Hendricks, T. J., Fyodorov, D. V., Wegman, L. J., Lelutiu, N. B., Pehek, E. A., Yamamoto, B., et al. (2003). Pet-1 ETS gene plays a critical role in 5-HT neuron development and is required for normal anxiety-like and aggressive behavior. *Neuron* 37, 233–247. doi: 10.1016/s0896-6273(02)01167-4

Herget, U., Wolf, A., Wullimann, M. F., and Ryu, S. (2014). Molecular neuroanatomy and chemoarchitecture of the neurosecretory preoptic-hypothalamic area in zebrafish larvae. *J. Comp. Neurol.* 522, 1542–1564. doi: 10.1002/cne.23480

Hollenberg, A. N. (2008). The role of the thyrotropin-releasing hormone (TRH) neuron as a metabolic sensor. *Thyroid* 18, 131–139. doi: 10.1089/thy.2007.0251

Jones, B., and McGinnis, W. (1993). A new Drosophila homeobox gene, bsh, is expressed in a subset of brain cells during embryogenesis. *Development* 117, 793–806.

Kato, Y., Iwasaki, Y., Iwasaki, J., Abe, H., Yanajhara, N., and Imura, H. (1978). Prolactin release by vasoactive intestinal polypeptide in rats. *Endocrinology* 103, 554–558. doi: 10.1210/endo-103-2-554

Ketchesin, K. D., Stinnett, G. S., and Seasholtz, A. F. (2017). Corticotropin-releasing hormone-binding protein and stress: from invertebrates to humans. *Stress* 20, 449–464. doi: 10.1080/10253890.2017.1322575

Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B., and Schilling, T. F. (1995). Stages of embryonic development of the zebrafish. *Dev. Dyn.* 203, 253–310. doi: 10.1002/aja.1002030302

Kimura, S., Hara, Y., Pineau, T., Fernandez-Salguero, P., Fox, C. H., Ward, J. M., et al. (1996). The T/ebp null mouse: thyroid-specific enhancer-binding protein is essential for the organogenesis of the thyroid, lung, ventral forebrain, and pituitary. *Genes Dev.* 10, 60–69.

Kozicz, T., Yanaihara, H., and Arimura, A. (1998). Distribution of urocortin-like immunoreactivity in the central nervous system of the rat. *J. Comp. Neurol.* 391, 1–10. doi: 10.1002/(sici)1096-9861(19980202)391:1<1::aid-cne1>3.0.co;2-6

Kurrasch, D. M., Cheung, C. C., Lee, F. Y., Tran, P. V., Hata, K., and Ingraham, H. A. (2007). The neonatal ventromedial hypothalamus transcriptome reveals novel markers with spatially distinct patterning. *J. Neurosci.* 27, 13624–13634. doi: 10.1523/JNEUROSCI.2858-07.2007

Kwan, K. M., Fujimoto, E., Grabher, C., Mangum, B. D., Hardy, M. E., Campbell, D. S., et al. (2007). The Tol2kit: a multisite gateway-based construction kit for Tol2 transposon transgenesis constructs. *Dev. Dyn.* 236, 3088–3099. doi: 10.1002/dvdy.21343

Lee, B., Kim, J., An, T., Kim, S., Patel, E. M., Raber, J., et al. (2018). Dlx1/2 and Otp coordinate the production of hypothalamic GHRH- and AgRP-neurons. *Nat. Commun.* 9:2026. doi: 10.1038/s41467-018-04377-4

Lee, B., Kim, S.-G., Kim, J., Choi, K. Y., Lee, S., Lee, S.-K., et al. (2013). Brain-specific homeobox factor as a target selector for glucocorticoid receptor in energy balance. *Mol. Cell. Biol.* 33, 2650–2658. doi: 10.1128/MCB.00094-13

Lee, S.-H., and Dan, Y. (2012). Neuromodulation of Brain States. *Neuron* 76, 209–222. doi: 10.1016/j.neuron.2012.09.012

Leshan, R. L., Greenwald-Yarnell, M., Patterson, C. M., Gonzalez, I. E., and Myers, M. G. Jr. (2012). Leptin action through hypothalamic nitric oxide synthase-1-expressing neurons controls energy balance. *Nat. Med.* 18, 820–823. doi: 10.1038/nm.2724

Levkowitz, G., Zeller, J., Sirotkin, H. I., French, D., Schilbach, S., Hashimoto, H., et al. (2003). Zinc finger protein too few controls the development of monoaminergic neurons. *Nat. Neurosci.* 6, 28–33. doi: 10.1038/nn979

Löhr, H., and Hammerschmidt, M. (2011). Zebrafish in endocrine systems: recent advances and implications for human disease. *Annu. Rev. Physiol.* 73, 183–211. doi: 10.1146/annurev-physiol-012110-142320

Löhr, H., Ryu, S., and Driever, W. (2009). Zebrafish diencephalic A11-related dopaminergic neurons share a conserved transcriptional network with neuroendocrine cell lineages. *Development* 136, 1007–1017. doi: 10.1242/dev.033878

Lopez, J. M., Moreno, N., Morona, R., Munoz, M., and Gonzalez, A. (2005). Colocalization of nitric oxide synthase and monoamines in neurons of the amphibian brain. *Brain Res. Bull.* 66, 555–559. doi: 10.1016/j.brainresbull.2005.05.008

Machluf, Y., Gutnick, A., and Levkowitz, G. (2011). Development of the zebrafish hypothalamus. *Ann. N. Y. Acad. Sci.* 1220, 93–105. doi: 10.1111/j.1749-6632.2010.05945.x

Mano, H., Asaoka, Y., Kojima, D., and Fukada, Y. (2019). Brain-specific homeobox Bsx specifies identity of pineal gland between serially homologous photoreceptive organs in zebrafish. *Commun. Biol.* 2:364. doi: 10.1038/s42003-019-0613-1

Manoli, M., and Driever, W. (2014). nkx2.1 *and nkx*2.4 genes function partially redundant during development of the zebrafish hypothalamus, preoptic region, and pallidum. *Front. Neuroanat.* 8:145. doi: 10.3389/fnana.2014.00145

Matsuda, K. (2013). Regulation of feeding behavior and psychomotor activity by corticotropin-releasing hormone (CRH) in fish. *Front Neurosci* 7:91. doi: 10.3389/fnins.2013.00091

Maywood, E. S., Reddy, A. B., Wong, G. K., O'Neill, J. S., O'Brien, J. A., McMahon, D. G., et al. (2006). Synchronization and maintenance of timekeeping in suprachiasmatic circadian clock cells by neuropeptidergic signaling. *Curr. Biol.* 16, 599–605. doi: 10.1016/j.cub.2006.02.023

McArthur, T., and Ohtoshi, A. (2007). A brain-specific homeobox gene, Bsx, is essential for proper postnatal growth and nursing. *Mol. Cell. Biol.* 27, 5120–5127. doi: 10.1128/MCB.00215-07

Medina, L., and Abellán, A. (2012). "Chapter 7 - Subpallial Structures," in *The Mouse Nervous System*, eds C. Watson, G. Paxinos, and L. Puelles, (San Diego, CA: Academic Press), 173–220.

Morales-Delgado, N., Castro-Robles, B., Ferrán, J. L., Martinez-de-la-Torre, M., Puelles, L., and Díaz, C. (2014). Regionalized differentiation of CRH, TRH, and GHRH peptidergic neurons in the mouse hypothalamus. *Brain Struct. Funct.* 219, 1083–1111. doi: 10.1007/s00429-013-0554-2

Nadim, F., and Bucher, D. (2014). Neuromodulation of neurons and synapses. *Curr. Opin. Neurobiol.* 29, 48–56. doi: 10.1016/j.conb.2014.05.003

Neff, K. L., Argue, D. P., Ma, A. C., Lee, H. B., Clark, K. J., and Ekker, S. C. (2013). Mojo Hand, a TALEN design tool for genome editing applications. *BMC Bioinformatics* 14:1. doi: 10.1186/1471-2105-14-1

Nogueiras, R., López, M., Lage, R., Perez-Tilve, D., Pfluger, P., Mendieta-Zerón, H., et al. (2008). Bsx, a novel hypothalamic factor linking feeding with locomotor activity, is regulated by energy availability. *Endocrinology* 149, 3009–3015. doi: 10.1210/en.2007-1684

Poon, K. L., Richardson, M., Lam, C. S., Khoo, H. E., and Korzh, V. (2003). Expression pattern of neuronal nitric oxide synthase in embryonic zebrafish. *Gene Expr. Patterns* 3, 463–466. doi: 10.1016/s1567-133x(03)00063-2

Puelles, L., Martinez-de-la-Torre, M., Bardet, S., and Rubenstein, J. L. R. (2012). "Chapter 8 - Hypothalamus," in *The Mouse Nervous System*, eds C. Watson, G. Paxinos, and L. Puelles, (San Diego, CA: Academic Press), 221–312.

Puelles, L., and Rubenstein, J. L. (2015). A new scenario of hypothalamic organization: rationale of new hypotheses introduced in the updated prosomeric model. *Front. Neuroanat.* 9:27. doi: 10.3389/fnana.2015.00027

Puttonen, H. A. J., Semenova, S., Sundvik, M., and Panula, P. (2017). Storage of neural histamine and histaminergic neurotransmission is VMAT2 dependent in the zebrafish. *Sci. Rep.* 7:3060. doi: 10.1038/s41598-017-02981-w

Rohr, K. B., Barth, K. A., Varga, Z. M., and Wilson, S. W. (2001). The nodal pathway acts upstream of hedgehog signaling to specify ventral telencephalic identity. *Neuron* 29, 341–351. doi: 10.1016/s0896-6273(01)00210-0

Romanov, R. A., Alpár, A., Hökfelt, T., and Harkany, T. (2019). Unified classification of molecular, network, and endocrine features of hypothalamic neurons. *Annu. Rev. Neurosci.* 9, 1–26. doi: 10.1146/annurev-neuro-070918-050414

Ronneberger, O., Liu, K., Rath, M., Rueβ, D., Mueller, T., Skibbe, H., et al. (2012). ViBE-Z: a framework for 3D virtual colocalization analysis in zebrafish larval brains. *Nat. Methods* 9, 735–742. doi: 10.1038/nmeth.2076

Russo, A. F. (2017). Overview of neuropeptides: Awakening the senses? *Headache* 57, 37–46. doi: 10.1111/head.13084

Ryu, S., Mahler, J., Acampora, D., Holzschuh, J., Erhardt, S., Omodei, D., et al. (2007). Orthopedia homeodomain protein is essential for diencephalic dopaminergic neuron development. *Curr. Biol.* 17, 873–880. doi: 10.1016/j.cub.2007.04.003

Sakkou, M., Wiedmer, P., Anlag, K., Hamm, A., Seuntjens, E., Ettwiller, L., et al. (2007). A role for brain-specific homeobox factor Bsx in the control of hyperphagia and locomotory behavior. *Cell Metab.* 5, 450–463. doi: 10.1016/j.cmet.2007.05.007

Santos-Duran, G. N., Menuet, A., Lagadec, R., Mayeur, H., Ferreiro-Galve, S., Mazan, S., et al. (2015). Prosomeric organization of the hypothalamus in an elasmobranch, the catshark *Scyliorhinus canicula*. *Front. Neuroanat.* 9:37. doi: 10.3389/fnana.2015.00037

Schredelseker, T., and Driever, W. (2018). Bsx controls pineal complex development. *Development* 145:dev163477. doi: 10.1242/dev.163477

Schredelseker, T., and Driever, W. (2020). Conserved genoarchitecture of the basal hypothalamus in Zebrafish embryos. *Front. Neuroanat.* 14:3. doi: 10.3389/fnana.2020.00003

Suga, H. (2019). Application of pluripotent stem cells for treatment of human neuroendocrine disorders. *Cell Tissue Res.* 375, 267–278. doi: 10.1007/s00441-018-2880-4

Sutton, A. K., Pei, H., Burnett, K. H., Myers, M. G. Jr., Rhodes, C. J., and Olson, D. P. (2014). Control of food intake and energy expenditure by Nos1 neurons of the paraventricular hypothalamus. *J. Neurosci.* 34, 15306–15318. doi: 10.1523/JNEUROSCI.0226-14.2014

Takenoya, F., Kageyama, H., Kang, K. S., and Shioda, S. (2013). "Chapter 115 - Neuropeptide B/W," in *Handbook of Biologically Active Peptides*, Second Edn, ed. A. J. Kastin, (Boston, MA: Academic Press), 863–868.

Thisse, C., and Thisse, B. (2005). *High Throughput Expression Analysis of ZF-Models Consortium Clones. ZFIN Direct Data Submission*. Available online at at: https://zfin.org/ZDB-PUB-051025-1 (accessed May 12, 2020).

Unger, J. L., and Glasgow, E. (2003). Expression of isotocin-neurophysin mRNA in developing zebrafish. *Gene Expr. Patterns* 3, 105–108. doi: 10.1016/s1567-133x(02)00064-9

Vernier, P. (2017). "1.04 - The brains of teleost fishes," in *Evolution of Nervous Systems (Second Edition)*, ed. J. H. Kaas, (Oxford: Academic Press), 59–75.

Wang, X., Kopinke, D., Lin, J., McPherson, A. D., Duncan, R. N., Otsuna, H., et al. (2012). Wnt signaling regulates postembryonic hypothalamic progenitor differentiation. *Dev. Cell* 23, 624–636. doi: 10.1016/j.devcel.2012.07.012

Wang, Y., Wang, M., Yin, S., Jang, R., Wang, J., Xue, Z., et al. (2015). NeuroPep: a comprehensive resource of neuropeptides. *Database* 2015:bav038. doi: 10.1093/database/bav038

Wee, C. L., Song, E. Y., Johnson, R. E., Ailani, D., Randlett, O., Kim, J.-Y., et al. (2019). A bidirectional network for appetite control in larval zebrafish. *eLife* 8:e43775. doi: 10.7554/eLife.43775

Wilson, S. W., Ross, L. S., Parrett, T., and Easter, S. S. (1990). The development of a simple scaffold of axon tracts in the brain of the embryonic zebrafish, Brachydanio rerio. *Development* 108, 121–145.

Wolf, A., and Ryu, S. (2013). Specification of posterior hypothalamic neurons requires coordinated activities of Fezf2, Otp, Sim1a and Foxb1.2. *Development* 140, 1762–1773. doi: 10.1242/dev.085357

Xavier, A. L., Fontaine, R., Bloch, S., Affaticati, P., Jenett, A., Demarque, M., et al. (2017). Comparative analysis of monoaminergic cerebrospinal fluid-contacting cells in Osteichthyes (bony vertebrates). *J. Comp. Neurol.* 525, 2265–2283. doi: 10.1002/cne.24204

Xie, Y., and Dorsky, R. I. (2017). Development of the hypothalamus: conservation, modification and innovation. *Development* 144, 1588–1599. doi: 10.1242/dev.139055

Xie, Y., Kaufmann, D., Moulton, M. J., Panahi, S., Gaynes, J. A., Watters, H. N., et al. (2017). Lef1-dependent hypothalamic neurogenesis inhibits anxiety. *PLoS Biol.* 15:e2002257. doi: 10.1371/journal.pbio.2002257

Yamamoto, K., Bloch, S., and Vernier, P. (2017). New perspective on the regionalization of the anterior forebrain in Osteichthyes. *Dev. Growth Differ.* 59, 175–187. doi: 10.1111/dgd.12348

Yang, N., Dong, Z., and Guo, S. (2012). Fezf2 regulates multilineage neuronal differentiation through activating basic helix-loop-helix and homeodomain genes in the zebrafish ventral forebrain. *J. Neurosci.* 32, 10940–10948. doi: 10.1523/JNEUROSCI.2216-12.2012

Zhang, Z., Boelen, A., Kalsbeek, A., and Fliers, E. (2018). TRH neurons and thyroid hormone coordinate the hypothalamic response to cold. *Eur. Thyroid J.* 7, 279–288. doi: 10.1159/000493976

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neuroscience

REVIEW
published: 16 June 2020
doi: 10.3389/fnins.2020.00621

# Secretion and Function of Pituitary Prolactin in Evolutionary Perspective

*Arpád Dobolyi[1,2]*, Szilvia Oláh[1], Dávid Keller[2], Rashmi Kumari[1], Emese A. Fazekas[1], Vivien Csikós[1], Éva Renner[3] and Melinda Cservenák[1]*

*[1] MTA-ELTE Laboratory of Molecular and Systems Neurobiology, Department of Physiology and Neurobiology, Institute of Biology, Eötvös Loránd University, Budapest, Hungary, [2] Laboratory of Neuromorphology, Department of Anatomy, Histology and Embryology, Semmelweis University, Budapest, Hungary, [3] Human Brain Tissue Bank and Microdissection Laboratory, Semmelweis University, Budapest, Hungary*

The hypothalamo-pituitary system developed in early vertebrates. Prolactin is an ancient vertebrate hormone released from the pituitary that exerts particularly diverse functions. The purpose of the review is to take a comparative approach in the description of prolactin, its secretion from pituitary lactotrophs, and hormonal functions. Since the reproductive and osmoregulatory roles of prolactin are best established in a variety of species, these functions are the primary subjects of discussion. Different types of prolactin and prolactin receptors developed during vertebrate evolution, which will be described in this review. The signal transduction of prolactin receptors is well conserved among vertebrates enabling us to describe the whole subphylum. Then, the review focuses on the regulation of prolactin release in mammals as we have the most knowledge on this class of vertebrates. Prolactin secretion in response to different reproductive stimuli, such as estrogen-induced release, mating, pregnancy and suckling is detailed. Reproduction in birds is different from that in mammals in several aspects. Prolactin is released during incubation in avian species whose regulation and functional significance are discussed. Little information is available on prolactin in reptiles and amphibians; therefore, they are mentioned only in specific cases to explain certain evolutionary aspects. In turn, the osmoregulatory function of prolactin is well established in fish. The different types of pituitary prolactin in fish play particularly important roles in the adaptation of eutherian species to fresh water environments. To achieve this function, prolactin is released from lactotrophs in hyposmolarity, as they are directly osmosensitive in fish. In turn, the released prolactin acts on branchial epithelia, especially ionocytes of the gill to retain salt and excrete water. This review will highlight the points where comparative data give new ideas or suggest new approaches for investigation in other taxa.

**Keywords: evolution, comparative, neuroendocrinology, hypothalamus, dopamine, osmoregulation, lactation**

OPEN ACCESS

***Edited by:***
*Giovanne B. Diniz,*
*Yale University, United States*

***Reviewed by:***
*Stanko S. Stojilkovic,*
*National Institutes of Health (NIH), United States*
*Arieh Gertler,*
*Hebrew University of Jerusalem, Israel*
*Dave Grattan,*
*University of Otago, New Zealand*

****Correspondence:***
*Arpád Dobolyi*
*dobolyi.arpad@ttk.elte.hu*

***Specialty section:***
*This article was submitted to Neuroendocrine Science, a section of the journal Frontiers in Neuroscience*

***Received:*** *25 January 2020*
***Accepted:*** *19 May 2020*
***Published:*** *16 June 2020*

***Citation:***
*Dobolyi A, Oláh S, Keller D, Kumari R, Fazekas EA, Csikós V, Renner É and Cservenák M (2020) Secretion and Function of Pituitary Prolactin in Evolutionary Perspective. Front. Neurosci. 14:621. doi: 10.3389/fnins.2020.00621*

## INTRODUCTION

Prolactin is an ancient regulatory molecule with diverse regulatory functions (Freeman et al., 2000). Prolactin has been shown to be expressed in a variety of different organs, however, its expression level is highest in the pituitary (Bu et al., 2015). It was suggested that in early vertebrates, the expression of prolactin was more diverse, but even in fish, it is already predominantly expressed

in specific cells of the adenohypophysis, from which prolactin is released into the bloodstream to act as a multifunctional hormone. Based on the structure and receptor type of prolactin, it belongs to the cytokines. Thus, together with growth hormone, it forms a group of pituitary hormones, which are not a 3–51 amino acid long neuropeptides acting on G-protein coupled receptors, but possess 1 transmembrane domain cytokine receptor. In these properties, prolactin and growth hormone are also different from the structure of other peptide and glycoprotein pituitary hormones. Another unique property of prolactin among adenohypophyseal hormones is that it does not have a target endocrine gland, which would mediate its actions but rather it exerts its actions directly via prolactin receptors localized in a variety of different target organs (Grattan, 2015). Most of the major targets of prolactin are epithelial cells, on which prolactin can exert proliferative effects as well as faster gene expression and even faster molecular actions (Aoki et al., 2019). In this paper, we will review the major actions of prolactin in vertebrate taxa. Most knowledge is available in mammals and fish where lactation and osmoregulation are the most established functions of prolactin, respectively (Horseman and Gregerson, 2014). Data are also accumulating in birds where prolactin is critically important in parental behavioral control (Smiley, 2019). The different functions require diverse stimuli for the release of prolactin. Our knowledge is more limited in this aspect of prolactin regulation. An aim of this review is to compare prolactin-related regulations between different vertebrate taxa to generate new research approaches.

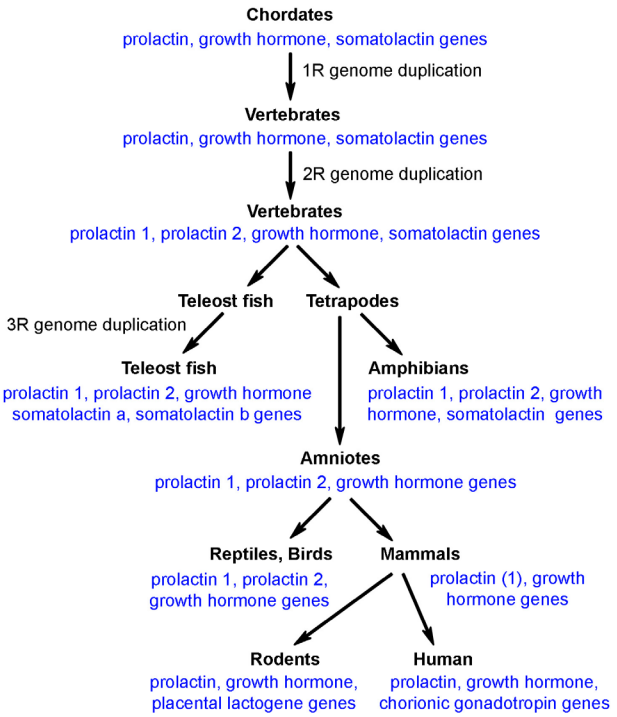


**FIGURE 1 |** The presence of prolactin/growth hormone family of genes in different clades of vertebrates.

## PROLACTIN AND THE EVOLUTION OF PROLACTIN GENES

Prolactin belongs to a gene family that comprises prolactin, growth hormone and somatolactin. The peptide sequences of these proteins exhibit approximately 20% homologies to each other in teleost species where all 3 of them are present. There are several additional versions of prolactin in specific vertebrate species, which were formed by local gene duplication. Although some reports suggested the presence of prolactin-like peptides in invertebrates, the general agreement is that the whole protein family was formed in chordates. It is known that whole genome duplication occurred 3 times during vertebrate evolution; the first round (1R) at the transition from chordates to vertebrates, the second round (2R) at the transition from agnathans to gnathostomes, and the third round (3R) after divergence of the teleost lineage. As for the prolactin gene family, however, even ancient chordate species before the 1R whole genome duplication possessed the three genes (**Figure 1**). Following 1R and 2R whole genome duplications, only one additional gene, prolactin 2 emerged while the teleost specific 3R whole genome duplication resulted in another gene, somatolactin b. In turn, somatolactin was lost in amniotes, and prolactin 2 was lost in mammals (Ocampo Daza and Larhammar, 2018a). The two different prolactins in teleost fish have similar functions. The prolactins are co-secreted from lactotrophs (Specker et al., 1993) and may bind to the same prolactin receptor albeit with different affinities (Auperin et al., 1995). The change in prolactin sequence showed an uneven speed in vertebrate evolution, which is in line with its role in major adaptive events, such as freshwater adaptation, or the development of lactation in mammals. Furthermore, new prolactin-like protein coding genes have developed, which, however, are typically not expressed in the pituitary. For example, rodents possess several placental lactogens (Bridges et al., 1996; Goffin et al., 1996). However, the present review, focuses on pituitary prolactin and mentions other prolactin-related proteins only if necessary for understanding of the functions of pituitary-derived prolactin.

The pituitary itself is an organ unique to vertebrates. In cephalochordates, the pituitary has a homolog, Hatschek's pit, an organ in the epithelial invagination of the oral cavity whose morphology and development is similar to the Rathke's pouch of the vertebrate embryo (Kubokawa et al., 2010). Cells of the Hatschek's pit express some proteins homologous to some adenohypophyseal hormones but not to prolactin (Holland et al., 2008). In vertebrates, the structure of the pituitary is similar in all classes (Sage and Bern, 1971). This structure consists of an anterior lobe, an intermediate lobe and a posterior lobe, or neurohypophysis. The anterior lobe of the pituitary is the adenohypophysis, which contains lactotrophs as well as other types of secretory cells (Bern and Nicoll, 1968). Lactotrophs actually form a more definitive mass within the rostral pars distalis of the adenohypophysis as opposed to the more intermingled localization of lactotrophs with other hormone producing cells in the lungfish and tetrapod pituitary

(Manzon, 2002). The secretory activity of adenohypophyseal cells is regulated by releasing and inhibitory factors of hypothalamic origin. Tetrapods have a hypophyseal portal system, which is reached by hypothalamic hormones in the median eminence. Fish do not have a median eminence, and the anterior part of the neurohypophysis was suggested to play a similar role in them (Sage and Bern, 1971).

## LACTOTROPH CELLS IN THE PITUITARY AND THEIR PROLACTIN RELEASE

Lactotrophs are located in the anterior lobe of the pituitary. In fish, prolactin secretion from lactotrophs seems to be affected directly by osmolarity as even cultured lactotrophs are sensitive to reduced osmolarity in their extracellular environment and respond to it by increased size and prolactin secretion (Labella et al., 1975; Weber et al., 2004). In mammals, in contrast, lactotrophs spontaneously release prolactin. It has been demonstrated that lactotrophs fired spontaneous plateau-bursting action potentials, which generated high amplitude calcium signals due to calcium influx via voltage-gated calcium channels (Van Goor et al., 2001). While different hypothalamic prolactin releasing factors have been proposed in mammals – the most compelling evidence is available for thyrotropin-releasing hormone -, it is still a general consensus that prolactin secretion from lactotrophs is only controlled by the inhibitory action of dopamine exerted by D2 dopamine receptors in mammals. The G-protein coupled D2 receptors utilize different signal transduction pathways to inhibit spontaneous prolactin release. They block voltage-gated calcium channels via pertussis toxin sensitive $G_{i/o}$ proteins while also desensitize calcium ion secretion coupling via pertussis toxin insensitive Gz proteins (Gonzalez-Iglesias et al., 2008). Accordingly, D2 receptor agonists, e.g., bromocriptine, block spontaneous electrical activity of lactotrophs as well as accompanied prolactin release (Auriemma et al., 2019). Therefore, bromocriptine can be used to cease milk production or treat hyperprolactinemia induced by pituitary tumors. In addition, mice lacking the D2 receptor are hyperprolactinemic (Saiardi et al., 1997). Dopamine reaches lactotrophs from the portal circulation of the pituitary, which contains dopamine following its release from dopaminergic neurons located in the arcuate nucleus (A12 dopaminergic cell group in mammals). Additional alternative ways of dopamine reaching lactotrophs have also been suggested, notably dopaminergic neurons located in the periventricular area (A14 dopaminergic cell group in mammals) projecting to the intermediate lobe of the pituitary. Although D2 receptors on lactotrophs play a role in the control of prolactin secretion in birds (Lv et al., 2018), they seem to possess a prolactin-releasing hormone, vasoactive intestinal peptide (VIP) (Tong et al., 1998; Christensen and Vleck, 2008)- Contrary to the lack of active prolactin-releasing mechanisms in mammals. VIP, similar to dopamine, is also released from neurons located in the infundibular/arcuate nucleus (Kosonsiriluk et al., 2008). The mechanism of negative feedback regulation for stable prolactin levels is driven by prolactin itself, which increases the neuronal activity of neuroendocrine dopaminergic neurons (Brown et al., 2012) and also their dopamine synthesis (Arbogast and Voogt, 1995) to reduce prolactin secretion. In birds, the participation of VIP in negative feedback would also make sense as an additional mechanism of feedback inhibition of prolactin release, but the available evidence is scarce (Namken et al., 2017). In mammals, VIP has also been suggested as regulator of prolactin secretion. VIP was found to be synthesized both in the hypothalamus and pituitary (Lam, 1991). *In vitro*, VIP stimulated prolactin release from lactotrophs (Bjoro et al., 1990). However, its effect was not specific as it also released other pituitary hormones (Lam, 1991). More recently it was suggested that VIP of suprachiasmatic origin could affect circadian rhythm of prolactin secretion (Egli et al., 2004; Kennett et al., 2008), or only some subtypes of lactotrophs could be affected by VIP (Christian et al., 2007), alternatively, VIP could affect proliferation of lactotrophs (Carretero et al., 2006). But most likely, VIP is not a physiological regulator of prolactin secretion in mammals (Phillipps et al., 2019). It is also not established yet if prolactin release in fish is regulated by the hypothalamus. It seems likely that not only osmolarity in the pituitary but other factors play a part in prolactin secretion. Estrogen evoked prolactin release from the marine teleost, sea bream (*Sparus aurata L.*), and this effect was inhibited by VIP (Brinca et al., 2003). Thus, VIP could be involved in the regulation of prolactin release as it is in avian species but exerts the opposite action by inhibiting the release.

## PROLACTIN RECEPTORS AND THEIR SIGNAL TRANSDUCTION

Prolactin is secreted from the pituitary to the circulation. Prolactin may bind to prolactin-binding proteins, which have been suggested in mammals but not in other taxa (Kline and Clevenger, 2001). Prolactin then exerts its actions by binding to its plasma membrane receptors.

### Evolution of Prolactin Receptors

DNA sequence comparisons as well as synthetic analysis revealed that early vertebrates possessed a common growth hormone/prolactin receptor even after the 2R tetraploidization event (even though separate growth hormone and prolactin genes were present before 1R). The separate prolactin receptor appeared soon after that by gene duplication (Ocampo Daza and Larhammar, 2018b). The teleost-specific 3R tetraploidization resulted in 2 prolactin receptor genes (PrlRa and b), which is characteristic of most teleost fish.

### Signal Transduction of Prolactin

The prolactin receptors all belong to the type I cytokine receptor family (Bole-Feysot et al., 1998). In mammals, there are long, intermediate and short prolactin receptor isoforms generated by alternative splicing (Freeman et al., 2000). All these isoforms are 1 transmembrane domain plasma membrane receptors. Signal transduction requires dimerization of the receptors. The receptors do not have enzyme activity but rather attract adaptor molecules upon prolactin binding to initiate signal transduction.

There are two major types of signal transduction pathways for the prolactin receptor. The conventional pathway uses a cytosolic tyrosine kinase, Janus kinase 2 (Campbell et al., 1994), which phosphorylates signal transducer and activator (STAT) 5 (Clevenger and Kline, 2001). Phosphorylated STAT5 (pSTAT5) then forms homodimers and acts as a transcription factor to induce the expression of various proteins, e.g., casein milk proteins. An alternative pathway is the mitogen-activated protein (MAP) kinase pathway, which generally mediates the proliferative actions of prolactin (Radhakrishnan et al., 2012).

## Expression of Prolactin Receptors in Different Tissues

Prolactin receptors are expressed in a variety of different organs based on different blotting and PCR technologies. In the fish, the osmoregulatory organs, such as the gills, kidney, and intestine contained the highest amount of prolactin receptor (Lee et al., 2006). In addition, prolactin receptors are also abundant in the brain, liver, gonads, liver, spleen, and present in the heart, muscle, bone, and skin, too (Sandra et al., 2000; Santos et al., 2001). Prolactin receptor has a similarly widespread tissue distribution in other clades, too. In mammals, the long and the short forms of prolactin receptors had similar distribution patterns with the long form dominating in most organs except for the kidney and lung (Ouhtit et al., 1993). More detailed distributional patterns of prolactin receptors within the organs expressing it have also been established, which, for example, indicated particularly high level of prolactin receptors within ionocytes of gill epithelium in fish, or alveolar cells of the mammary gland.

## Distribution of Prolactin Receptors in the Brain

Given the enormously high number of different cell types in the brain and the different functions connected to the nervous system, the distribution of prolactin receptors within the mammalian brain has been in focus for decades. Initial immunolabeling studies demonstrated neuronal expression and a topographical distribution within the brain with high level of prolactin receptor in the anteroventral periventricular nucleus, the medial preoptic area, the paraventricular, and the arcuate nuclei with even more hypothalamic sites becoming visible in lactating rats (Pi and Voogt, 2000) with weak labeling in some striatal, thalamic and cortical sites, too. This distributional pattern was confirmed by *in situ* hybridization histochemistry, which also showed that the short form of the receptor may be present in hypothalamic but also in extrahypothalamic sites (Bakowska and Morrell, 1997, 2003). Modern molecular biological techniques using prolactin receptor-Cre recombinase mice bred with green fluorescent protein reporter mice revealed the precise expression of the long form of the prolactin receptor in the same sites as also described by *in situ* hybridization histochemistry (Kokay et al., 2018). The sites of signal transduction of prolactin can also be directly examined using pSTAT5 immunohistochemistry, which resulted in the same labeling pattern following injection of exogenous prolactin (Brown et al., 2010; Sjoeholm et al., 2011) or following suckling in lactating mice (Olah et al., 2018) as described with the above techniques (**Figure 2**). Some of these methods cannot properly address the subcellular location of prolactin receptors, e.g., where they are located in relation to synapses. Therefore, techniques, which can address these type of questions, such as immunohistochemistry remain important research tools in the field in the future, too.

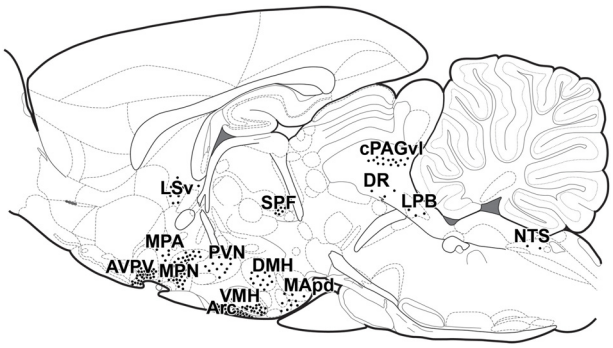


**FIGURE 2 |** Distribution of prolactin sensitive neurons in lactating mothers in schematic sagittal section. Each dot represents 20 pSTAT5-positive neurons evoked by suckling in lactating mice, with the exception of Arc, where each dot represents 85 neurons/mm$^2$ in a 40 mm thick section (Olah et al., 2018). This distribution of pSTAT5-positive cells is essentially the same as the distribution of exogenous prolactin-evoked pSTAT5 and of prolactin receptors visualized by *in situ* hybridization histochemistry (Brown et al., 2010). Arc, arcuate nucleus; AVPV, anteroventral periventricular nucleus; DMH, dorsomedial hypothalamic nucleus; DR, dorsal raphe nucleus; cPAGvl, caudal periaqueductal gray, ventrolateral subdivision; LPB, lateral parabrachial nucleus; LSv, lateral septal nucleus, ventral subdivision; MApd, medial amygdaloid nucleus, posterodorsal subdivision; MPA, medial preoptic area; MPN, medial preoptic nucleus; NTS, nucleus of the solitary tract; PVN, paraventricular hypothalamic nucleus; SPF, subparafascicular area; VMHvl, ventromedial hypothalamic nucleus, ventrolateral subdivision.

# REPRODUCTIVE ACTIONS OF PROLACTIN

The reproductive cycle of animals can be generally divided into sexual and parental phases. Despite the phylogenetic diversity in the specific regulations, it can be claimed that gonadotropins and sexual steroid hormones play pivotal roles in the control of the sexual phase while prolactin is the major regulator of the parental phase (Everett, 1964). Of course, the 2 phases are connected with each other. Indeed, prolactin is released during the luteinizing hormone surge in the estrous cycle, due to increased estrogen levels, which could be a direct pituitary action or a kisspeptin-evoked suppression of dopaminergic activity (Szawka et al., 2011; Grattan and Szawka, 2019). The role of the prolactin released during estrous is not well established to date (Phillipps et al., 2019) as it does not seem to affect female sexual behavior (Witcher and Freeman, 1985). Prolactin secretion is also induced by mating, both in males and females, and a characteristic prolactin secretory pattern appears during pregnancy. Our knowledge is highly limited in nonmammalian species regarding estrogen-

or mating-induced prolactin release. Furthermore, dopamine release under these circumstances varies even within mammals, as discussed below.

## Prolactin Released During Mammalian Pregnancy

Mating induces prolactin secretion (Exton et al., 2001). Information on mating is probably conveyed to the hypothalamus in a neuronal pathway similar to that involved in suckling (described below) as both transfer somatosensory information from the spinal cord to the same hypothalamic site, the dopaminergic neurons in the arcuate nucleus. The role of mating-induced prolactin secretion is not known (Voogt et al., 2001). In males, it could contribute to the sexual refractory period or the formation of parental motivation. In females, mating-induced release may initiate the characteristic pattern of prolactin secretion during pregnancy. This secretory pattern can be highly different depending on the species. In humans, progesterone synthesis in the corpus luteum is maintained by chorionic gonadotropin of placental origin. In contrast, prolactin maintains the corpus luteum in rodents, which requires an immediate high concentration of prolactin in the serum (Phillipps et al., 2019). To this end, prolactin is released from the pituitary twice a day (Gunnet and Freeman, 1983). This secretory pattern is probably triggered by mating and lasts for 12–13 days if the animal is pseudopregnant, e.g., following artificial vaginocervical stimulation (Gunnet and Freeman, 1984). In turn, prolactin secretion from the pituitary ceases at approximately 9–10 days of pregnancy, as it is terminated by negative feedback due to lactogens of placental origin (Goffin et al., 1996). The placenta is fully functional by that time of pregnancy, and rodents have expanded the number of genes encoding prolactin-like proteins, which are expressed in the placenta (Soares et al., 2007). Some of these placental lactogens act on the prolactin receptor, therefore, their high serum concentration inhibits prolactin release from the pituitary by activating dopaminergic neurons in the arcuate nucleus via the prolactin receptor expressed in these neurons (Goffin et al., 1996). In contrast to rodents, human prolactin levels rise gradually during pregnancy, which is likely a steroid driven process (Phillipps et al., 2019). The major role of prolactin and placental lactogens in late pregnancy is mammopoesis. This function can be well assessed in mice lacking prolactin or prolactin receptor: mammary glands do not have proper side branching and alveologenesis, of which only the former can be rescued by progesterone (Horseman and Gregerson, 2014). These actions of prolactin are complex as they are mediated by receptor activator of nuclear factor kappa-B (RANK) present in secretory cells of the alveolar epithelium. RANK ligand is released from prolactin-sensing cells in the alveolar epithelium to mediate paracrine actions (Fernandez-Valdivia et al., 2009). The pregnancy of rats lasted for 22 days. Approximately 1 day before parturition, a prolactin surge emerges probably because progesterone is reduced at this point, which helps dopaminergic neurons escape feedback stimulation, even though placental lactogen levels are still high. It was proposed that the dopamine content of dopaminergic neurons in the arcuate nucleus is reduced (Andrews and Grattan, 2003), possibly producing encephalin instead of dopamine (Yip et al., 2019). The function of this release of prolactin before parturition is not clearly established because prolactin effects required during pregnancy, such as lactogenesis, adaptation of the brain for maternal behavior, or increased insulin secretion via prolactin receptors in the beta cells of the pancreas to avoid hyperglycemia can all be performed by placental lactogens.

## Prolactin Released in the Postpartum Period in Response to Suckling in Mammals

Serum prolactin levels are generally high in the postpartum period (Crowley, 2015). Although the proliferation and hypertrophy of lactotrophs are required for their maintenance (Le Tissier et al., 2015), some processes are required to prevent the negative feedback caused by prolactin itself. The mechanisms of these alterations during lactation have not been elucidated to date. The phenotypic changes described above for late pregnancy (that is an induced enkephalin production by tuberoinfundibular dopaminergic neurons) could be involved. In addition, insulin-like growth factor-1 (IGF-1) has also been suggested to play a role (Dobolyi and Leko, 2019). Prolonged intracerebroventricular IGF-1 was shown to stimulate dopamine secretion in the mediobasal hypothalamus and inhibit prolactin secretion from the pituitary (Leko et al., 2017a). In turn, a binding protein of IGF-1, IGF-binding protein-3 (IGFBP-3) is induced dramatically during lactation specifically in the arcuate nucleus (Leko et al., 2017b). Therefore, IGFBP-3 may be able to sequester IGF-1 from around dopaminergic neurons thereby eliminating the local stimulatory effects of otherwise elevated IGF-1 on dopaminergic neurons (Leko et al., 2017b).

The above mentioned, and potentially other mechanisms preventing feedback inhibition of prolactin levels are necessary for elevated prolactin level during lactation. However, the major stimulus that induces prolactin secretion in mothers, is suckling by the pups. Suckling-induced prolactin release was measured in several species. The experimental paradigm most often used in rats includes removal of the pups for 4 h from the dam, during which her serum prolactin levels decrease to a basal minimum level (Neill and Nagy, 1994; Nagy et al., 2005). When the litter is given back to the mother, suckling starts almost immediately. Serum prolactin increases approximately 60 fold, and the maximum level is reached at 30 min following the beginning of suckling (Cservenak et al., 2013). The increased prolactin level is a consequence of reduced dopamine action from the arcuate nucleus (Phillipps et al., 2019). Therefore, tuberoinfundibular dopaminergic neurons must be inhibited by somatosensory stimulus of the nipples. Early studies suggested that this pathway runs ventromedial to the medial geniculate body because microstimulation of this brain area evoked lactogenesis (Tindal and Knaggs, 1969) (**Figure 3A**). We found neurons expressing tuberoinfundibular peptide of 39 residues (TIP39) in the same location (Dobolyi et al., 2003b) (**Figure 3B**). TIP39 belongs to the parathyroid hormone family of peptides, which is most abundantly expressed in the brain and is not known

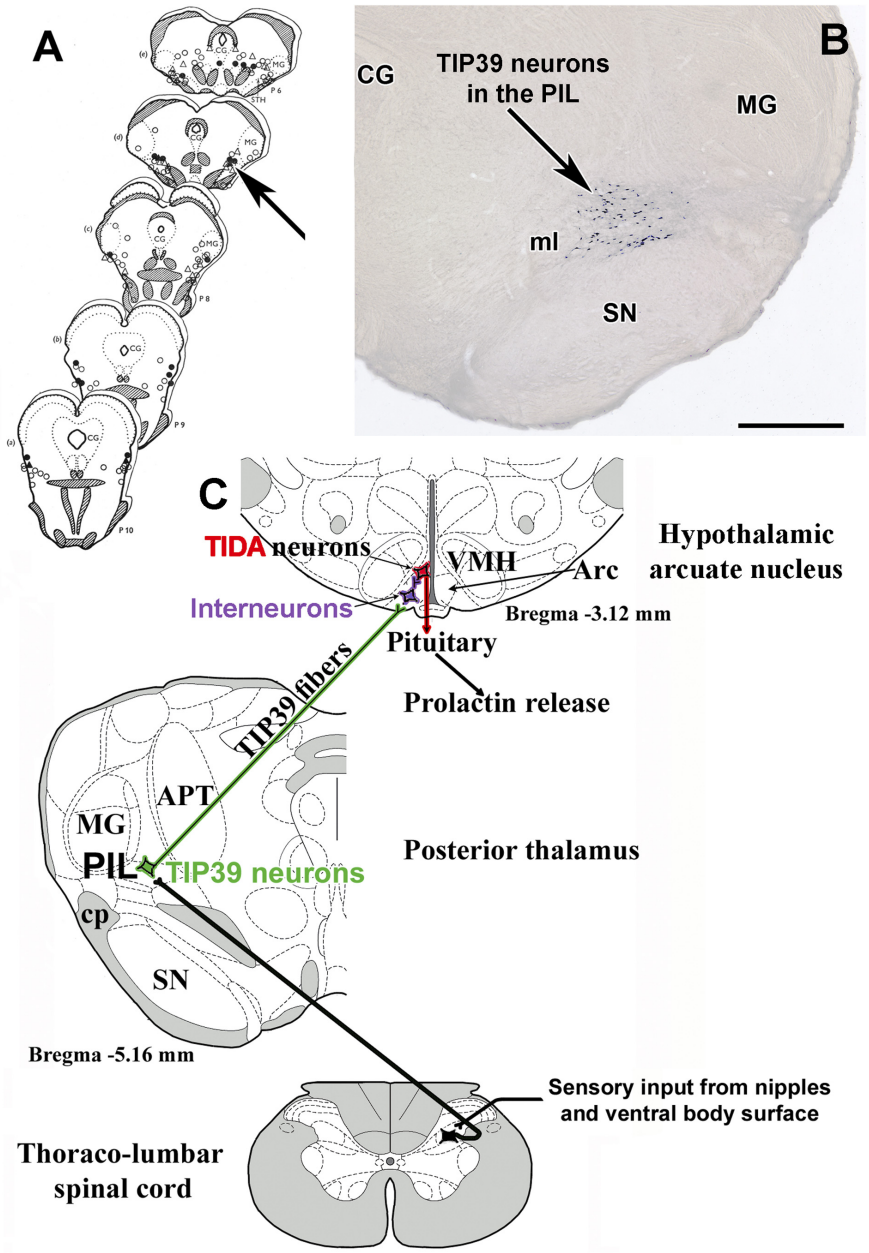


**FIGURE 3 |** Brain sites whose microstimulation can evoke lactogenesis **(A)** are compared to the position of TIP39 neurons **(B)** in the medial subdivision of the posterior intralaminar complex of the thalamus (PIL). **(C)** The proposed neuronal pathway of suckling-induced prolactin release containing TIP39 neurons in the PIL, which project to the arcuate nucleus. Arrows indicate corresponding brain areas. CG, central gray; MG, medial geniculate body; ml, medial lemniscus; SN, substantia nigra. Scale bar = 1 mm.

to play any role in calcium homeostasis as other members of the peptide family (Suarez-Bregua et al., 2017). Based on previous topographical characterization of the part of the brain expressing TIP39 (Ledoux et al., 1987), we called the position of TIP39 neurons in the medial subdivision of the posterior intralaminar complex of the thalamus (PIL) (Dobolyi et al., 2010). These neurons were activated in response to suckling and markedly increased their TIP39 expression during lactation (Cservenak et al., 2010). Furthermore, the neurons were shown to project to the arcuate nucleus (Dobolyi et al., 2003a), which contains the receptor of TIP39, the parathyroid hormone 2 receptor (PTH2R) (Usdin et al., 2003) in mice (Faber et al., 2007), rat (Dobolyi et al., 2006), and humans (Bago et al., 2009). Injecting a PTH2R antagonist into the lateral ventricle or expressing it with a virus infecting neurons in the vicinity of the arcuate nucleus markedly reduced suckling induced prolactin release suggesting the prolactin-inducing action of TIP39 (Cservenak et al., 2013) (**Figure 3C**). TIP39 neurons are likely glutamatergic based on

electron microscopic and double labeling studies (Cservenak et al., 2017). Glutamate released from these neurons could also be involved in mediating prolactin release. The role of these thalamo-hypothalamic neurons is also consistent with the finding that the pathway, along which projections of TIP39 neurons reach the hypothalamus within the zona incerta (Palkovits et al., 2010), overlaps with the locations of microstimuli evoking lactogenesis rostral to the area ventromedial to the medial geniculate body (Tindal and Knaggs, 1972). Interestingly, stimulation of the preoptic area of the hypothalamus also elicited prolactin secretion. It was interpreted that olfactory information could influence prolactin secretion by that route (Tindal and Knaggs, 1977). In turn, it is also possible that retrograde activation of TIP39 neurons, whose major target is the preoptic area (Cservenak et al., 2017), could also contribute to the stimulatory effect of the preoptic area on prolactin secretion.

The pathway relaying in the PIL (**Figure 3C**) may also convey the effects of suckling to forebrain sites other than the arcuate nucleus to release prolactin, as suckling induced c-fos expression not only in the PIL (Lin et al., 1998) but also in a variety of different brain regions where PIL neurons project (Li et al., 1999; Lonstein and Stern, 1999). Indeed, the PTH2 receptor was identified in the preoptic area, the paraventricular and dorsomedial hypothalamic nuclei, and the lateral septum (Dobolyi et al., 2006, 2012) and TIP39 terminals were shown to innervate oxytocin neurons in the paraventricular nucleus (Dobolyi et al., 2018) and galanin neurons in the preoptic area (Cservenak et al., 2017) known to control maternal behaviors (Wu et al., 2014).

The functions of prolactin in the postpartum period are numerous. The most well established one is maintaining lactation by acting on mammary epithelial cells. Thereby, prolactin released from a suckling bout enables the mammary gland to further maintain milk production for the next suckling bout via effects of the released prolactin, so in a sense, the pups order their next meal via suckling-induced prolactin release (Phillipps et al., 2019). In addition to lactation, prolactin exerts a variety of different actions in mothers (Bridges and Grattan, 2019) including stimulation of the immune system (Borba et al., 2019) and important effects in the brain by penetrating through the blood-brain barrier (Brown et al., 2016) to reach its multiple targets expressing prolactin receptors in the brain (Bakowska and Morrell, 1997; Kokay et al., 2018) (**Figure 2**). Prolactin contributes to the increased maternal food intake (Sauve and Woodside, 1996; Naef and Woodside, 2007), lactational anestrus (Grattan and Szawka, 2019), and the induction of maternal behaviors (Brown et al., 2017). These actions are conveyed by the prolactin receptor, although it is established only in some cases which prolactin action is mediated by which location (**Table 1**). While prolactin is the major maternal hormone affecting the brain in mammals, the brain functions of mothers are also affected by incoming sensory inputs, primarily from the pups. Pups are known to activate a variety of different brain centers, which can be identified at the cellular level using the c-fos technique. The hormonal and neuronal inputs have to support each other to form the proper adaptive responses including maternal behaviors. Their interaction was addressed by double labeling comparing prolactin activated (pSTAT5-positive) versus directly suckling activated (c-fos-positive) brain areas and neurons. Surprisingly, only a relatively small portion of neurons were affected by both stimuli (Olah et al., 2018), suggesting that prolactin provides different types of information for the maternal adaptation of the brain than direct neuronal inputs arriving primarily from the suckling stimulus.

Suckling-induced prolactin release can be prevented by dehydration of the mother. Drinking high salt (2.5%) water for a day reduced suckling-induced prolactin release the following day (Nagy et al., 1992). Acute hyperosmolarity evoked by intraventricular injection of 0.5 ml 10% saline within a suckling bout immediately blocked prolactin release. The blockade of suckling in a hyperosmotic state makes sense as a lactating mother loses a large amount of water during milk production. It is not known whether lactotrophs themselves would be osmoreceptors in mammals; therefore, it is more likely that osmoreceptor cells in the preoptic area convey information on serum osmolarity and blood volume to them, probably via dopaminergic neurons, as both acute and chronic effects of high saline concentrations could be blocked with dopamine receptor antagonists (Nagy et al., 1992). When released, prolactin has anti-diuretic action to replace lost salt and water by nursing. Prolactin can directly act on thirst centers to increase water intake and on renal tubule cells in the kidney to increase salt and water retention (Morrissey et al., 2001), and its indirect action via anti-diuretic hormone (ADH) was also demonstrated (Walker et al., 2001). In addition, ADH could also be increased via the suckling stimulus independent of prolactin (Suzuki et al., 2000).

## Prolactin Released During Incubation in Birds

The serum prolactin level is increased in birds during brooding (Kuwayama et al., 1992), and its level correlated with brooding behavior (Smiley and Adkins-Regan, 2016). In fact, an early prolactin surge is responsible for the formation of a brood patch, a defeathered area on the belly skin, which has an abundant blood supply to effectively transfer heat toward the eggs (Ohkubo, 2017). The role of elevated prolactin is likely to be related to incubation behavior (Smiley, 2019), as prolactin administration can induce brooding behavior (Youngren et al., 1991), while reducing prolactin, e.g., by immunization against it, leads to the cessation of incubation (March et al., 1994). The incubation-promoting effects of prolactin are likely to be mediated via prolactin receptors in the preoptic area (Youngren et al., 1989). In addition, prolactin may also contribute to the increased aggressive and defensive behaviors of incubating birds (Romanov et al., 2002) and to the decline in gonadal function by inhibiting gonadotropin-releasing hormone-producing neurons (Rozenboim et al., 1993). Our knowledge of the time course of changes in prolactin levels and actions in the incubation period is not good. It would be interesting to know how fast the prolactin level decreases if the parent is removed from the nest and how fast prolactin rises when allowed to incubate again.

The mechanism by which prolactin is induced by incubation or for incubation is not fully understood. In birds, dopamine

**TABLE 1 |** The proposed brain functions of prolactin mediated by prolactin receptors localized in different brain regions.

| Brain area | Function | Experimental evidence or *suggested function* |
|---|---|---|
| Lateral septal nucleus, ventral (LSv) | Prolactin in LSv may be related to maternal aggression. | *Cabrera-Reyes et al. (2017); Salais-Lopez et al. (2017)* |
| Anteroventral periventricular nucleus (AVPV) | Dopaminergic neurons expressing pSTAT5 in response to lactation promote maternal care and oxytocin secretion. | Brown et al. (2015); Higo et al. (2015), Scott et al. (2015) |
| | Lactation induced rapid modulations of kisspeptin are mediated by prolactin. | |
| Medial preoptic area (MPOA) | Prolactin is necessary for onset of maternal behavior and stimulates maternal care. | Bridges et al. (1990); Wu et al. (2014), Brown et al. (2017); Cservenak et al. (2017) |
| | Galanin expressing neurons implicated in maternal behavior contain pSTAT5. | |
| Paraventricular hypothalamic nucleus (PVN) | Oxytocin neurons, involved in lactation and maternal behaviors, express pSTAT5 in lactating rats. Prolactin enhances oxytocin release. | Parker et al. (1991); Nishimori et al. (1996), Fodor et al. (2012); Augustine et al. (2016, 2018) |
| | Prolactin inhibits vasopressin neurons in lactating rats, which play a part in the development of maternal behavior. | |
| Arcuate nucleus (Arc) | Dopamine neurons inhibiting prolactin release contain pSTAT5. | Cave et al. (2001); Sapsford et al. (2012), Romano et al. (2013); Araujo-Lopes et al. (2014) |
| | Prolactin regulates kisspeptin neurons to suppress LH secretion. | |
| Ventromedial nucleus, ventrolateral (VMHvl) | Activation by prolactin in VMH may be involved in regulation of increased feeding behavior in lactating rats. | *Pi and Grattan (1999)* |
| Dorsomedial hypothalamic nucleus (DM) | DM regulates food intake and energy balance. | Augustine et al. (2008); Nagaishi et al. (2014), Lopez-Vicchi et al. (2020) |
| | Prolactin affects DM neurons, which play a role in the metabolic changes triggered by pregnancy and lactation. | |
| Medial amygdaloid nucleus, posterodorsal (MApd) | Prolactin in MApd may be related to maternal aggression and modulation of the neuroendocrine stress axis. | *Cabrera-Reyes et al. (2017); Salais-Lopez et al. (2017)* |
| Subparafascicular area (SPF) | not known | |
| Periaqueductal gray, caudal, ventrolateral (PAGvl) | PAG is critical for suckling induced kyphosis, prolactin may promote it. | *Lonstein and Stern (1997)* |
| Dorsal raphe nucleus (DR) | Serotonin neurons project to GnRH neurons located in the preoptic area. Prolactin may act on DR serotonin neurons to suppress the activity of GnRH neurons in lactating dams. | *Brown et al. (2011)* |
| Lateral parabrachial nucleus (LPB) | Not known | |
| Nucleus of the solitary tract (NTS) | Prolactin plays a role in the metabolic changes triggered by pregnancy and lactation possibly via the NTS, too. | *Brown et al. (2011); Nagaishi et al. (2014)* |
| | Noradrenergic neurons of NTS origin might mediate the suppression of GnRH neuronal activity. | |

*The references written in italics represent only presumed or suggested functions.*

can inhibit spontaneous prolactin release from lactotrophs via D2 receptors (Christensen and Vleck, 2008), similar to mammals. However, dopamine also has a stimulatory action via D1 receptors best established in the turkey (Bhatt et al., 2003; Chaiseha et al., 2003) but also in other avian species (Xu et al., 2010). Action via D1 receptors can stimulate hypothalamic neurons expressing VIP (Sartsoongnoen et al., 2008), which acts as a prolactin-releasing hormone in avian species (Proudman and Opel, 1988; Kosonsiriluk et al., 2008). It is unlikely that sexual hormones elicit prolactin release during incubation because increased prolactin levels are found even in ovariectomized birds during incubation. Thus, it is plausible that somatosensory input from the eggs itself can evoke prolactin secretion via neuronal pathways (Massaro et al., 2007). Indeed, swapping eggs from the parents leads to reduced serum prolactin levels (Sinpru et al., 2018). The potential neuronal pathways involved have not been revealed to date. It is not possible that TIP39 plays a role in prolactin release in avian species as it does in mammals because the gene encoding this peptide is missing in birds, even though it is present in all other vertebrate taxa (On et al., 2015), but a homologous neuronal pathway could be involved. Furthermore, it would be interesting to learn whether osmolarity affects prolactin release in birds during incubation as it does during lactation in mammals. However, this question has not been addressed to date, even though it is reasonable that dehydration contributes to the cessation of actual incubation driven by both thirst and a reduced parental motivation due to reduced prolactin level.

In addition to brooding behavior, many avian species, the altricial birds, also show parental behavior in the form of feeding the nestlings (regurgitation) in the posthatching period. Since prolactin level is not particularly high in the posthatching

period, feeding behavior may not depend on prolactin in most avian species even though the reduction of prolactin following hatching is slower in altricial as compared to precocial birds (Lea and Sharp, 1991) and prolactin level correlated with individual differences in parental behavior in the passerine zebra finch (Smiley and Adkins-Regan, 2016). In turn, rearing behavior in the posthatching period may be induced by tactile stimulation by nestlings in combination with visual or auditory inputs (Richard-Yris et al., 1998). These inputs are likely mediated toward lactotrophs for prolactin secretion by VIP neurons in the arcuate nucleus as during incubation (Buntin et al., 1991). Tactile, visual and auditory inputs lead to direct neuronal activation of parental brain centers in zebra finch (Fazekas et al., 2020). To date, the interconnection of these activated cells with those sensitive to prolactin has not been established in birds as in mice (Olah et al., 2018), despite the availability of pSTAT5 immunolabeling in birds (Buntin and Buntin, 2014). In addition to a presumed role of prolactin in parenting behavior in birds (Buntin et al., 1991), the hormone may have an additional, specific role in avian species that produce crop milk (Wan et al., 2019). In fact, prolactin was originally discovered by its ability to induce crop milk production in the pigeons. Still, the mechanism of prolactin secretion for crop milk production has not yet been established. Microstimulation experiments suggested that the preoptic area may be involved in crop milk production while direct stimulation of the mediobasal hypothalamus including the median eminence, had no effect (Kanematsu, 1980).

## The Roles of Prolactin in Fish Parenting

Parental behaviors appear in about 30% of teleost fish (Royle et al., 2014). Interestingly, in these species, the males provide parental care somewhat more often than the females (Reynolds et al., 2002). Although parental behaviors have various forms in fish, the most common forms include nest building and egg (or embryo) attendance. The latter not only protects the eggs from predators but fanning also aerates the eggs and frees them from debris (Rosenblatt, 2003). Although the available knowledge on the control of these behaviors is scarce, some recent data indicated the potential involvement of prolactin (Whittington and Wilson, 2013). Prolactin implants increased nurturing behavior in bluegill (*Lepomis macrochirus*) (Cunha et al., 2019) and also in three-spined stickleback (*Gasterosteus aculeatus*) (de Ruiter et al., 1986). Prolactin level may rise during spawning and remains elevated during parental care in cichlid fish (*Oreochromis niloticus)* (Tacon et al., 2000). The mechanism how prolactin in induced in unknown at present although simple cues as sensory inputs from the eggs are a likely candidate for fanning (Dulac et al., 2014). It should also be mentioned that data arguing against a role of prolactin in fish parenting are also available (Bender et al., 2008).

# OSMOREGULATORY ACTIONS OF PROLACTIN

Osmoregulation is a complex process that includes a variety of different regulatory hormones and systems including prolactin

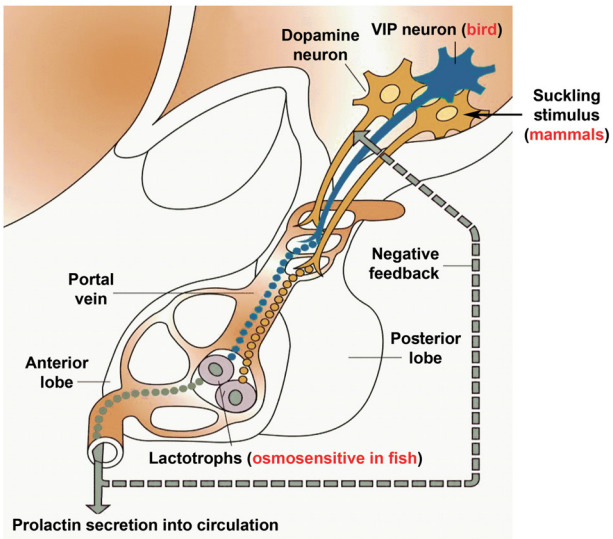


**FIGURE 4 |** Schematics on the control of prolactin secretion. Prolactin releasing factor, which is truly established only in birds to be vasoactive intestinal polypeptide (VIP), reaches the anterior lobe of the pituitary via the portal circulation to evoke prolactin secretion from the prolactin-producing lactotrophs located in the pituitary. In contrast, prolactin release in fish is most profoundly driven by hyposmolarity sensed by the lactotrophs themselves. In contrast, mammalian lactotrophs release prolactin spontaneously, and the regulation takes place by inhibition, which is carried out by dopamine released from dopaminergic neurons located in the mediobasal hypothalamus. The inhibition of prolactin secretion by dopamine is likely present in all vertebrate taxa. In turn, suckling-induced prolactin secretion is characteristic of mammals only, in which suckling stimulus of the pups is conveyed to dopaminergic neurons in a classic neuroendocrine reflex to stimulate prolactin release according to the need of the pups. The clade-specific features were added to a modified version of the basic scheme (Loìpez et al., 2003).

but also renin-angiotensin, anti-diuretic hormone, aldosterone, and atrial natriuretic peptide (Takei et al., 2014). These endocrine components of the regulatory system are relatively conserved in vertebrates despite the differing needs of various animals (the major difference between fish and mammals being the lack of involvement of aldosterone in fish), while the neuronal components of osmoregulation show remarkable differences. Notably, in mammals, forebrain circumventricular organs, such as the vascular organ of the lamina terminalis and the subfornical organ play pivotal roles in sensing both plasma osmolarity and plasma hormone content, most importantly angiotensin II levels (McKinley and Johnson, 2004). These organs convey this information toward thirst centers of the cerebral cortex as well as toward neurohypophyseal neurons synthesizing antidiuretic hormone (ADH). ADH neurons are themselves osmosensors, which secrete ADH in response to hyperosmolarity (Bourque and Oliet, 1997). Teleost fish do not have a subfornical organ and the role of the vascular organ of the lamina terminalis is not known in osmoregulation (Katayama et al., 2018). They may be able to respond to angiotensin II detected through the area a postrema (Nobata et al., 2013), a circumventricular organ in the hindbrain. However, the major regulator of their drinking in vagal input from peripheral receptors

(Mayer-Gostan and Hirano, 1976). These differences between clades may simply is a consequence of forebrain development in mammals, which better allows reconciliation of drinking with other behaviors. Alternatively, the different requirement associated terrestrial environment created the new type of osmoregulatory systems. Indeed, terrestrial animals need to retain water while fish have to deal with the difference in their ionic composition from their environment with which they are in direct contact. Generally, the environment of fish can be freshwater or saltwater, which requires very different regulations. Furthermore, there are euryhaline species that can live in both fresh and seawater, as they migrate or live in brackish water. The major difference between fresh and seawater adaptation of fish is that in the gills, they actively excrete salt in seawater while actively take up salt in fresh water. In addition, fish drink much more in seawater than in fresh water (Takei et al., 2014). Prolactin was shown to play a role in freshwater adaptation as first demonstrated in killifish (*Fundulus heteroclitus*), a species that could survive following hypophysectomy in fresh water only in the presence of external prolactin (Pickford and Phillips, 1959). More recently, strong evidence came from zebrafish models lacking the prolactin gene: the larvae survived to adulthood in brackish but not in egg water (Shu et al., 2016). This role of prolactin is widespread in euryhaline fish but not ubiquitous, as catfish and salmonids can survive in fresh water following hypophysectomy without prolactin (Hirano, 1986). Although it is possible that the latter species produce prolactin outside of the pituitary as lactotrophs could be located in other locations in relatively primitive eels (Sakamoto and McCormick, 2006). Nevertheless, all teleost species have pituitary with lactotrophs in their anterior lobe; therefore, it is more likely that other osmoregulatory systems can provide freshwater adaptability in these species.

Prolactin exerts its regulatory function in all branchial epithelia, including the gill, kidney, urinary bladder, and gastrointestinal tract (Takei et al., 2014). In fish, the major osmoregulatory organ is the gill where ionocytes are located for water and ion transport while the kidney has the most important osmoregulatory function in mammals (Manzon, 2002). Prolactin may be able induce the proliferation of specific types of ionocytes, which can remove water and take up ions (Hiroi and McCormick, 2012). While prolactin is among the osmoregulatory hormones with relatively slower actions in general, it is also able to induce and stimulate ion transporters, such as $Na^+/Cl^-$ cotransporter and the $Na^+/K^+$ pump (Breves et al., 2014) via its receptors expressed in ionocytes (Santos et al., 2001). It is also under investigation how the 2 types of prolactin receptors (a and b) both expressed in osmoregulatory epithelia of several teleost species support each other's actions. Less evidence is available but it is still likely that prolactin can inhibit aquaporins, e.g., aquaporin 3, which is known to be involved in freshwater adaptation (Lignot et al., 2002; Breves et al., 2016), and induce tight junction forming cadherins to reduce water uptake in the gills. Our knowledge is more limited on potential action of prolactin on teleost kidney, bladder or inhibition of drinking, which are all potential sites of action. Furthermore, the osmoregulatory function of prolactin is less pronounced in mammals unless we consider its effect on milk production as an osmoregulatory action. Milk contains high amount of fluid as well as sodium ion. Old literature suggested water retention ability of prolactin, however, it was shown to be caused by contamination with ADH (Keeler and Wilson, 1976). More recent experiments demonstrated increased sodium and chloride retention by prolactin (Greenlee et al., 2015) while natriuretic effect in mammals by acting on the $Na^+/K^+$ pump via local dopaminergic system in the proximal renal tubules has also been reported (Ibarra et al., 2005). The effect of prolactin on the kidney function of birds is also not known. However, an interesting stimulatory effect of prolactin has been reported on the duck nasal salt gland, an important avian salt excretory organ (Peaker et al., 1970).

The secretion of prolactin in response to changes in osmolarity is less well studied than its osmoregulatory actions. Nevertheless, it has been demonstrated that hyposmolarity evokes prolactin release from the pituitary of teleost fish and an increase in prolactin gene expression also takes place (Lee et al., 2006; Fuentes et al., 2010). It seems likely that lactotrophs themselves are osmosensitive in fish and release prolactin in response to the hypoosmotic local environment (Kwong et al., 2009; Watanabe et al., 2009; Seale et al., 2012). Furthermore, the prolactin secretion as an osmotic response is not affected by pharmacological blockade of dopamine receptors (Liu et al., 2006). Evidence is available that cultured lactotrophs react to reduced osmolarity by prolactin secretion and that it is accompanied by increased cell size of the lactotrophs (Weber et al., 2004). In lactotrophs from Tilapia, the increased cell volume was blocked by aquaporin inhibitors (Watanabe et al., 2009). Given that aquaporin 3 is present in Tilapia lactotrophs (Watanabe et al., 2005), its involvement in water intake of the cells is likely. The increased volume may activate the transient receptor potential-vanilloid (TRPV) 4 receptor, a stretch-activated calcium channel (Watanabe et al., 2002), and the resulting elevated intracellular calcium level leads to prolactin release (Seale et al., 2012). A positive feedback of prolactin has also been suggested to increase the response (Yamaguchi et al., 2016). For long-term adaptation to fresh water, prolactin expression in tremendously increased while aquaporin 3 and TRPV 4 expression are reduced for sensitization of prolactin secretion (Seale et al., 2012). In contrast to fish, there is no compelling evidence on the direct osmosensitivity of lactotrophs in mammals or in other vertebrate taxa. The mammals have osmosensitive cells in the vascular organ of the lamina terminalis and the subfornical organ (McKinley and Johnson, 2004). Neuronal output from these preoptic hypothalamic regions can reach vasopressin neurons in the paraventricular nucleus and thirst centers of the brain. It seems likely that dopaminergic neurons in the mediobasal hypothalamus regulating prolactin secretion from lactotrophs also receive information on osmolarity from the same preoptic receptor cells as hyperosmolarity prevented suckling-induced prolactin release but dopamine receptor antagonists could block the effect of hypoosmolarity as discussed above (Nagy et al., 1992). On the other hand, it has not been addressed in detail whether other types of prolactin release, e.g., during pregnancy or stress can also be inhibited by hyperosmolarity in the pituitary

in tetrapods although different effects of osmotic inhibition on differently elicited prolactin release have been reported (Dohanics et al., 1994). It is also not known if prolactin release from mammalian lactotrophs are influenced by osmolality of its environment physiologically. Some early results indicated direct osmoreceptive lactotrophs in the mammalian pituitary (Labella et al., 1975; Lorenson and Jacobs, 1987). More recently, it was found that sub-physiological hypotonicity elicited transient release followed by sustained depression of prolactin release from perfused rat lactotrophs (Jorgacevski et al., 2008). The role of osmolality in prolactin secretion of birds is also not established even though it would also be interesting to address whether prolactin secretion during incubation in birds can be prevented by hyperosmolarity. It would make sense if the bird parent would stop incubating and go drinking when hyperosmolarity occurs.

## CONCLUSION

Prolactin is a vertebrate-specific hormone whose functions have been studied for a long time. Indeed, immense knowledge has accumulated on prolactin secretion and function in a variety of different species. Thus, prolactin represents an exciting opportunity for evolutionary neuroendocrinology as its functions are compared between the different species and even different classes of vertebrates. For example, prolactin, a major osmoregulator in fish turned into a hormone that regulates lactation in mammals. In fact, both of these functions require the action of prolactin on epithelial cells, both as far as their proliferation and the control of their transport processes. Another intriguing change is the parental behavioral action of prolactin, which already appears in fish (Cunha et al., 2019) but becomes prominent in birds and mammals. In both classes, parental care has different forms, such as brooding and nursing behaviors. Other effects of prolactin, such as inhibiting gonadotropins is also maintained in a variety of different taxa. Evolutionary comparison of the regulation of prolactin secretion is also instructive. Direct osmosensitivity of lactotrophs is characteristic only of fish; however, hyperosmolarity also inhibits prolactin release in mammals, and research investigating this question is also proposed in birds. The inhibitory influence of dopamine of hypothalamic origin on prolactin secretion seems to be present in all taxa although its role is most important in mammals in which a regulatory releasing mechanism has not been identified to date (**Figure 4**).

## AUTHOR CONTRIBUTIONS

All authors participated in the design of literature review, reading of relevant manuscripts, interpretation of the literature, and writing of the manuscript.

## FUNDING

This work was supported by the Hungarian National Research, Development and Innovation Office OTKA K116538, OTKA K134221, NKFIH-4300-1/2017-NKP_17, NKFIH-2920-1/2016-VEKOP-2.3.-15, and NKFIH-6785-1/2016-VEKOP-2.3.3-15 research grants, and Semmelweis University (783-3/2018/FEKUTSRAT) and Eötvös Loránd University Institutional Excellence Programs (NKFIH-1157-8/2019-DT). MC was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences and the ÚNKP-19-4-ELTE-180 and ÚNKP-19-3-I-SE-93 New National Excellence Program of the Ministry for Innovation and Technology, Hungary.

## REFERENCES

Andrews, Z. B., and Grattan, D. R. (2003). Opioid receptor subtypes involved in the regulation of prolactin secretion during pregnancy and lactation. *J. Neuroendocrinol.* 15, 227–236. doi: 10.1046/j.1365-2826.2003.00975.x

Aoki, M., Wartenberg, P., Grunewald, R., Phillipps, H. R., Wyatt, A., Grattan, D. R., et al. (2019). Widespread cell-specific prolactin receptor expression in multiple murine organs. *Endocrinology* 160, 2587–2599. doi: 10.1210/en.2019-00234

Araujo-Lopes, R., Crampton, J. R., Aquino, N. S., Miranda, R. M., Kokay, I. C., Reis, A. M., et al. (2014). Prolactin regulates kisspeptin neurons in the arcuate nucleus to suppress lh secretion in female rats. *Endocrinology* 155, 1010–1020. doi: 10.1210/en.2013-1889

Arbogast, L. A., and Voogt, J. L. (1995). Hypoprolactinemia decreases tyrosine hydroxylase activity in the tuberoinfundibular dopaminergic neurons acutely by protein dephosphorylation and chronically by changes in gene expression. *Endocrine* 3, 801–806. doi: 10.1007/bf02935684

Augustine, R. A., Bouwer, G. T., Seymour, A. J., Grattan, D. R., and Brown, C. H. (2016). Reproductive regulation of gene expression in the hypothalamic supraoptic and paraventricular nuclei. *J. Neuroendocrinol.* 28, 1–12.

Augustine, R. A., Ladyman, S. R., and Grattan, D. R. (2008). From feeding one to feeding many: hormone-induced changes in bodyweight homeostasis during pregnancy. *J. Physiol.* 586, 387–397. doi: 10.1113/jphysiol.2007.146316

Augustine, R. A., Seymour, A. J., Campbell, R. E., Grattan, D. R., and Brown, C. H. (2018). Integrative neuro-humoral regulation of oxytocin neuron activity in pregnancy and lactation. *J. Neuroendocrinol.* 30, 1–15.

Auperin, B., Rentier-Delrue, F., Martial, J. A., and Prunet, P. (1995). Regulation of gill prolactin receptors in tilapia (*Oreochromis niloticus*) after a change in salinity or hypophysectomy. *J. Endocrinol.* 145, 213–220. doi: 10.1677/joe.0.1450213

Auriemma, R. S., Pirchio, R., De Alcubierre, D., Pivonello, R., and Colao, A. (2019). Dopamine agonists: from the 1970s to today. *Neuroendocrinology* 109, 34–41. doi: 10.1159/000499470

Bago, A. G., Dimitrov, E., Saunders, R., Seress, L., Palkovits, M., Usdin, T. B., et al. (2009). Parathyroid hormone 2 receptor and its endogenous ligand tuberoinfundibular peptide of 39 residues are concentrated in endocrine, viscerosensory and auditory brain regions in macaque and human. *Neuroscience* 162, 128–147. doi: 10.1016/j.neuroscience.2009.04.054

Bakowska, J. C., and Morrell, J. I. (1997). Atlas of the neurons that express mrna for the long form of the prolactin receptor in the forebrain of the female rat. *J. Comp. Neurol.* 386, 161–177. doi: 10.1002/(sici)1096-9861(19970922)386:2<161::aid-cne1>3.0.co;2-#

Bakowska, J. C., and Morrell, J. I. (2003). The distribution of mrna for the short form of the prolactin receptor in the forebrain of the female rat. *Brain Res. Mol. Brain Res.* 116, 50–58. doi: 10.1016/s0169-328x(03)00213-4

Bender, N., Taborsky, M., and Power, D. M. (2008). The role of prolactin in the regulation of brood care in the cooperatively breeding fish neolamprologus pulcher. *J. Exp. Zool. A Ecol. Genet. Physiol.* 309, 515–524. doi: 10.1002/jez.482

Bern, H. A., and Nicoll, C. S. (1968). The comparative endocrinology of prolactin. *Recent Prog Horm Res* 24, 681–720.

Bhatt, R., Youngren, O., Kang, S., and El Halawani, M. (2003). Dopamine infusion into the third ventricle increases gene expression of hypothalamic vasoactive intestinal peptide and pituitary prolactin and luteinizing hormone beta subunit in the turkey. *Gen. Comp. Endocrinol.* 130, 41–47. doi: 10.1016/s0016-6480(02)00533-6

Bjoro, T., Sand, O., Ostberg, B. C., Gordeladze, J. O., Torjesen, P., Gautvik, K. M., et al. (1990). The mechanisms by which vasoactive intestinal peptide (vip) and thyrotropin releasing hormone (trh) stimulate prolactin release from pituitary cells. *Biosci. Rep.* 10, 189–199. doi: 10.1007/bf01116578

Bole-Feysot, C., Goffin, V., Edery, M., Binart, N., and Kelly, P. A. (1998). Prolactin (prl) and its receptor: actions, signal transduction pathways and phenotypes observed in prl receptor knockout mice. *Endocr. Rev.* 19, 225–268. doi: 10.1210/edrv.19.3.0334

Borba, V. V., Zandman-Goddard, G., and Shoenfeld, Y. (2019). Prolactin and autoimmunity: the hormone as an inflammatory cytokine. *Best Pract. Res. Clin. Endocrinol. Metab.* 33:101324. doi: 10.1016/j.beem.2019.101324

Bourque, C. W., and Oliet, S. H. (1997). Osmoreceptors in the central nervous system. *Annu. Rev. Physiol.* 59, 601–619. doi: 10.1146/annurev.physiol.59.1.601

Breves, J. P., Inokuchi, M., Yamaguchi, Y., Seale, A. P., Hunt, B. L., Watanabe, S., et al. (2016). Hormonal regulation of aquaporin 3: opposing actions of prolactin and cortisol in tilapia gill. *J. Endocrinol.* 230, 325–337. doi: 10.1530/joe-16-0162

Breves, J. P., McCormick, S. D., and Karlstrom, R. O. (2014). Prolactin and teleost ionocytes: new insights into cellular and molecular targets of prolactin in vertebrate epithelia. *Gen. Comp. Endocrinol.* 203, 21–28. doi: 10.1016/j.ygcen.2013.12.014

Bridges, R. S., and Grattan, D. R. (2019). 30 years after: CNS actions of prolactin: sources, mechanisms and physiological significance. *J. Neuroendocrinol.* 31:e12669. doi: 10.1111/jne.12669

Bridges, R. S., Numan, M., Ronsheim, P. M., Mann, P. E., and Lupini, C. E. (1990). Central prolactin infusions stimulate maternal behavior in steroid-treated, nulliparous female rats. *Proc. Natl. Acad. Sci. U.S.A.* 87, 8003–8007. doi: 10.1073/pnas.87.20.8003

Bridges, R. S., Robertson, M. C., Shiu, R. P. C., Friesen, H. G., Stuer, A. M., and Mann, P. E. (1996). Endocrine communication between conceptus and mother: placental lactogen stimulation of maternal behavior. *Neuroendocrinology* 64, 57–64. doi: 10.1159/000127098

Brinca, L., Fuentes, J., and Power, D. M. (2003). The regulatory action of estrogen and vasoactive intestinal peptide on prolactin secretion in sea bream (sparus aurata, l.). *Gen. Comp. Endocrinol.* 131, 117–125. doi: 10.1016/s0016-6480(02)00628-7

Brown, R. S., Herbison, A. E., and Grattan, D. R. (2011). Differential changes in responses of hypothalamic and brainstem neuronal populations to prolactin during lactation in the mouse. *Biol. Reprod.* 84, 826–836. doi: 10.1095/biolreprod.110.089185

Brown, R. S., Herbison, A. E., and Grattan, D. R. (2015). Effects of prolactin and lactation on a15 dopamine neurones in the rostral preoptic area of female mice. *J. Neuroendocrinol.* 27, 708–717. doi: 10.1111/jne.12297

Brown, R. S., Kokay, I. C., Herbison, A. E., and Grattan, D. R. (2010). Distribution of prolactin-responsive neurons in the mouse forebrain. *J. Comp. Neurol.* 518, 92–102. doi: 10.1002/cne.22208

Brown, R. S., Piet, R., Herbison, A. E., and Grattan, D. R. (2012). Differential actions of prolactin on electrical activity and intracellular signal transduction in hypothalamic neurons. *Endocrinology* 153, 2375–2384. doi: 10.1210/en.2011-2005

Brown, R. S., Wyatt, A. K., Herbison, R. E., Knowles, P. J., Ladyman, S. R., Binart, N., et al. (2016). Prolactin transport into mouse brain is independent of prolactin receptor. *FASEB J.* 30, 1202–1210.

Brown, R. S. E., Aoki, M., Ladyman, S. R., Phillipps, H. R., Wyatt, A., Boehm, U., et al. (2017). Prolactin action in the medial preoptic area is necessary for postpartum maternal nursing behavior. *Proc. Natl. Acad. Sci. U.S.A.* 114, 10779–10784. doi: 10.1073/pnas.1708025114

Bu, G., Liang, X., Li, J., and Wang, Y. (2015). Extra-pituitary prolactin (prl) and prolactin-like protein (prl-l) in chickens and zebrafish. *Gen. Comp. Endocrinol.* 220, 143–153. doi: 10.1016/j.ygcen.2015.02.001

Buntin, J. D., Becker, G. M., and Ruzycki, E. (1991). Facilitation of parental behavior in ring doves by systemic or intracranial injections of prolactin. *Horm. Behav.* 25, 424–444. doi: 10.1016/0018-506x(91)90012-7

Buntin, J. D., and Buntin, L. (2014). Increased stat5 signaling in the ring dove brain in response to prolactin administration and spontaneous elevations in prolactin during the breeding cycle. *Gen. Comp. Endocrinol.* 200, 1–9. doi: 10.1016/j.ygcen.2014.02.006

Cabrera-Reyes, E. A., Limon-Morales, O., Rivero-Segura, N. A., Camacho-Arroyo, I., and Cerbon, M. (2017). Prolactin function and putative expression in the brain. *Endocrine* 57, 199–213. doi: 10.1007/s12020-017-1346-x

Campbell, G. S., Argetsinger, L. S., Ihle, J. N., Kelly, P. A., Rillema, J. A., and Carter-Su, C. (1994). Activation of jak2 tyrosine kinase by prolactin receptors in nb2 cells and mouse mammary gland explants. *Proc. Natl. Acad. Sci. U.S.A.* 91, 5232–5236. doi: 10.1073/pnas.91.12.5232

Carretero, J., Angoso, M., Rubio, M., Blanco, E. J., Sierra, E., Herrero, J. J., et al. (2006). In vitro immunoblockade of vip inhibits the proliferation of pituitary prolactin cells. *Anat. Embryol. (Berl.)* 211, 11–18. doi: 10.1007/s00429-005-0058-9

Cave, B. J., Wakerley, J. B., Luckman, S. M., and Tortonese, D. J. (2001). Hypothalamic targets for prolactin: assessment of c-fos induction in tyrosine hydroxylase- and proopiomelanocortin-containing neurones in the rat arcuate nucleus following acute central prolactin administration. *Neuroendocrinology* 74, 386–395. doi: 10.1159/000054705

Chaiseha, Y., Youngren, O., Al-Zailaie, K., and El Halawani, M. (2003). Expression of d1 and d2 dopamine receptors in the hypothalamus and pituitary during the turkey reproductive cycle: colocalization with vasoactive intestinal peptide. *Neuroendocrinology* 77, 105–118. doi: 10.1159/000068649

Christensen, D., and Vleck, C. M. (2008). Prolactin release and response to vasoactive intestinal peptide in an opportunistic breeder, the zebra finch (*Taeniopygia guttata*). *Gen. Comp. Endocrinol.* 157, 91–98. doi: 10.1016/j.ygcen.2008.04.013

Christian, H. C., Chapman, L. P., and Morris, J. F. (2007). Thyrotrophin-releasing hormone, vasoactive intestinal peptide, prolactin-releasing peptide and dopamine regulation of prolactin secretion by different lactotroph morphological subtypes in the rat. *J. Neuroendocrinol.* 19, 605–613. doi: 10.1111/j.1365-2826.2007.01567.x

Clevenger, C. V., and Kline, J. B. (2001). Prolactin receptor signal transduction. *Lupus* 10, 706–718. doi: 10.1191/096120301717164949

Crowley, W. R. (2015). Neuroendocrine regulation of lactation and milk production. *Compr. Physiol.* 5, 255–291. doi: 10.1002/cphy.c140029

Cservenak, M., Bodnar, I., Usdin, T. B., Palkovits, M., Nagy, G. M., and Dobolyi, A. (2010). Tuberoinfundibular peptide of 39 residues is activated during lactation and participates in the suckling-induced prolactin release in rat. *Endocrinology* 151, 5830–5840. doi: 10.1210/en.2010-0767

Cservenak, M., Kis, V., Keller, D., Dimen, D., Menyhart, L., Olah, S., et al. (2017). Maternally involved galanin neurons in the preoptic area of the rat. *Brain Struct. Funct.* 222, 781–798. doi: 10.1007/s00429-016-1246-5

Cservenak, M., Szabo, E. R., Bodnar, I., Leko, A., Palkovits, M., Nagy, G. M., et al. (2013). Thalamic neuropeptide mediating the effects of nursing on lactation and maternal motivation. *Psychoneuroendocrinology* 38, 3070–3084. doi: 10.1016/j.psyneuen.2013.09.004

Cunha, A. A. P., Partridge, C. G., Knapp, R., and Neff, B. D. (2019). Androgen and prolactin manipulation induces changes in aggressive and nurturing behavior in a fish with male parental care. *Horm. Behav.* 116:104582. doi: 10.1016/j.yhbeh.2019.104582

de Ruiter, A. J., Wendelaar Bonga, S. E., Slijkhuis, H., and Baggerman, B. (1986). The effect of prolactin on fanning behavior in the male three-spined stickleback, gasterosteus aculeatus l. *Gen. Comp. Endocrinol.* 64, 273–283. doi: 10.1016/0016-6480(86)90014-6

Dobolyi, A., Cservenak, M., and Young, L. J. (2018). Thalamic integration of social stimuli regulating parental behavior and the oxytocin system. *Front. Neuroendocrinol.* 51:2. doi: 10.1016/j.yfrne.2018.05.002

Dobolyi, A., Dimitrov, E., Palkovits, M., and Usdin, T. B. (2012). The neuroendocrine functions of the parathyroid hormone 2 receptor. *Front. Endocrinol. (Lausanne)* 3:121. doi: 10.3389/fendo.2012.00121

Dobolyi, A., Irwin, S., Wang, J., and Usdin, T. B. (2006). The distribution and neurochemistry of the parathyroid hormone 2 receptor in the rat hypothalamus. *Neurochem. Res.* 31, 227–236. doi: 10.1007/s11064-005-9011-9

Dobolyi, A., and Leko, A. H. (2019). The insulin-like growth factor-1 system in the adult mammalian brain and its implications in central maternal adaptation. *Front. Neuroendocrinol.* 52:2. doi: 10.1016/j.yfrne.2018.12.002

Dobolyi, A., Palkovits, M., Bodnar, I., and Usdin, T. B. (2003a). Neurons containing tuberoinfundibular peptide of 39 residues project to limbic, endocrine, auditory and spinal areas in rat. *Neuroscience* 122, 1093–1105. doi: 10.1016/j.neuroscience.2003.08.034

Dobolyi, A., Palkovits, M., and Usdin, T. B. (2003b). Expression and distribution of tuberoinfundibular peptide of 39 residues in the rat central nervous system. *J. Comp. Neurol.* 455, 547–566. doi: 10.1002/cne.10515

Dobolyi, A., Palkovits, M., and Usdin, T. B. (2010). The tip39-pth2 receptor system: unique peptidergic cell groups in the brainstem and their interactions with central regulatory mechanisms. *Prog. Neurobiol.* 90, 29–59. doi: 10.1016/j.pneurobio.2009.10.017

Dohanics, J., Smith, M. S., Blackburn, R. E., and Verbalis, J. G. (1994). Osmotic inhibition of prolactin secretion in rats. *J. Neuroendocrinol.* 6, 291–298. doi: 10.1111/j.1365-2826.1994.tb00585.x

Dulac, C., O'Connell, L. A., and Wu, Z. (2014). Neural control of maternal and paternal behaviors. *Science* 345, 765–770. doi: 10.1126/science.1253291

Egli, M., Bertram, R., Sellix, M. T., and Freeman, M. E. (2004). Rhythmic secretion of prolactin in rats: action of oxytocin coordinated by vasoactive intestinal polypeptide of suprachiasmatic nucleus origin. *Endocrinology* 145, 3386–3394. doi: 10.1210/en.2003-1710

Everett, J. W. (1964). Central neural control of reproductive functions of the adenohypophysis. *Physiol. Rev.* 44, 373–431. doi: 10.1152/physrev.1964.44.3.373

Exton, M. S., Kruger, T. H., Koch, M., Paulson, E., Knapp, W., Hartmann, U., et al. (2001). Coitus-induced orgasm stimulates prolactin secretion in healthy subjects. *Psychoneuroendocrinology* 26, 287–294. doi: 10.1016/s0306-4530(00)00053-6

Faber, C. A., Dobolyi, A., Sleeman, M., and Usdin, T. B. (2007). Distribution of tuberoinfundibular peptide of 39 residues and its receptor, parathyroid hormone 2 receptor, in the mouse brain. *J. Comp. Neurol.* 502, 563–583. doi: 10.1002/cne.21330

Fazekas, E. A., Morvai, B., Zachar, G., Dora, F., Szekely, T., Pogany, A., et al. (2020). Neuronal activation in zebra finch parents associated with reintroduction of nestlings. *J. Comp. Neurol.* 528, 363–379. doi: 10.1002/cne.24761

Fernandez-Valdivia, R., Mukherjee, A., Ying, Y., Li, J., Paquet, M., DeMayo, F. J., et al. (2009). The rankl signaling axis is sufficient to elicit ductal side-branching and alveologenesis in the mammary gland of the virgin mouse. *Dev. Biol.* 328, 127–139. doi: 10.1016/j.ydbio.2009.01.019

Fodor, A., Klausz, B., Pinter, O., Daviu, N., Rabasa, C., Rotllant, D., et al. (2012). Maternal neglect with reduced depressive-like behavior and blunted c-fos activation in brattleboro mothers, the role of central vasopressin. *Horm. Behav.* 62, 539–551. doi: 10.1016/j.yhbeh.2012.09.003

Freeman, M. E., Kanyicska, B., Lerant, A., and Nagy, G. (2000). Prolactin: structure, function, and regulation of secretion. *Physiol. Rev.* 80, 1523–1631. doi: 10.1152/physrev.2000.80.4.1523

Fuentes, J., Brinca, L., Guerreiro, P. M., and Power, D. M. (2010). Prl and gh synthesis and release from the sea bream (*Sparus auratus* l.) pituitary gland in vitro in response to osmotic challenge. *Gen. Comp. Endocrinol.* 168, 95–102. doi: 10.1016/j.ygcen.2010.04.005

Goffin, V., Shiverick, K. T., Kelly, P. A., and Martial, J. A. (1996). Sequence-function relationships within the expanding family of prolactin, growth hormone, placental lactogen, and related proteins in mammals. *Endocr. Rev.* 17, 385–410. doi: 10.1210/er.17.4.385

Gonzalez-Iglesias, A. E., Murano, T., Li, S., Tomic, M., and Stojilkovic, S. S. (2008). Dopamine inhibits basal prolactin release in pituitary lactotrophs through pertussis toxin-sensitive and -insensitive signaling pathways. *Endocrinology* 149, 1470–1479. doi: 10.1210/en.2007-0980

Grattan, D. R. (2015). The hypothalamo-prolactin axis. *J. Endocrinol.* 226, T101–T122.

Grattan, D. R., and Szawka, R. E. (2019). Kisspeptin and prolactin. *Semin. Reprod. Med.* 37, 93–104.

Greenlee, M. M., Mitzelfelt, J. D., Duke, B. J., Al-Khalili, O., Bao, H. F., and Eaton, D. C. (2015). Prolactin stimulates sodium and chloride ion channels in a6 renal epithelial cells. *Am. J. Physiol. Renal Physiol.* 308, F697–F705.

Gunnet, J. W., and Freeman, M. E. (1983). The mating-induced release of prolactin: a unique neuroendocrine response. *Endocr. Rev.* 4, 44–61. doi: 10.1210/edrv-4-1-44

Gunnet, J. W., and Freeman, M. E. (1984). Hypothalamic regulation of mating-induced prolactin release. Effect of electrical stimulation of the medial preoptic area in conscious female rats. *Neuroendocrinology* 38, 12–16. doi: 10.1159/000123859

Higo, S., Aikawa, S., Iijima, N., and Ozawa, H. (2015). Rapid modulation of hypothalamic kiss1 levels by the suckling stimulus in the lactating rat. *J. Endocrinol.* 227, 105–115. doi: 10.1530/joe-15-0143

Hirano, T. (1986). The spectrum of prolactin action in teleosts. *Prog. Clin. Biol. Res.* 205, 53–74.

Hiroi, J., and McCormick, S. D. (2012). New insights into gill ionocyte and ion transporter function in euryhaline and diadromous fish. *Respir. Physiol. Neurobiol.* 184, 257–268. doi: 10.1016/j.resp.2012.07.019

Holland, L. Z., Albalat, R., Azumi, K., Benito-Gutierrez, E., Blow, M. J., Bronner-Fraser, M., et al. (2008). The amphioxus genome illuminates vertebrate origins and cephalochordate biology. *Genome Res.* 18, 1100–1111.

Horseman, N. D., and Gregerson, K. A. (2014). Prolactin actions. *J. Mol. Endocrinol.* 52, R95–R106.

Ibarra, F., Crambert, S., Eklof, A. C., Lundquist, A., Hansell, P., and Holtback, U. (2005). Prolactin, a natriuretic hormone, interacting with the renal dopamine system. *Kidney Int.* 68, 1700–1707. doi: 10.1111/j.1523-1755.2005.00586.x

Jorgacevski, J., Stenovec, M., Kreft, M., Bajic, A., Rituper, B., Vardjan, N., et al. (2008). Hypotonicity and peptide discharge from a single vesicle. *Am. J. Physiol. Cell Physiol.* 295, C624–C631.

Kanematsu, S. (1980). Crop sac stimulation after electrochemical stimulation of the brain in the pigeon. *Gen. Comp. Endocrinol.* 42, 212–218. doi: 10.1016/0016-6480(80)90190-2

Katayama, Y., Sakamoto, T., Saito, K., Tsuchimochi, H., Kaiya, H., Watanabe, T., et al. (2018). Drinking by amphibious fish: convergent evolution of thirst mechanisms during vertebrate terrestrialization. *Sci. Rep.* 8:625.

Keeler, R., and Wilson, N. (1976). Vasopressin contamination as a cause of some apparent renal actions of prolactin. *Can. J. Physiol. Pharmacol.* 54, 887–890. doi: 10.1139/y76-124

Kennett, J. E., Poletini, M. O., and Freeman, M. E. (2008). Vasoactive intestinal polypeptide modulates the estradiol-induced prolactin surge by entraining oxytocin neuronal activity. *Brain Res.* 1196, 65–73. doi: 10.1016/j.brainres.2007.12.061

Kline, J. B., and Clevenger, C. V. (2001). Identification and characterization of the prolactin-binding protein in human serum and milk. *J. Biol. Chem.* 276, 24760–24766. doi: 10.1074/jbc.m011786200

Kokay, I. C., Wyatt, A., Phillipps, H. R., Aoki, M., Ectors, F., Boehm, U., et al. (2018). Analysis of prolactin receptor expression in the murine brain using a novel prolactin receptor reporter mouse. *J. Neuroendocrinol.* 30:e12634. doi: 10.1111/jne.12634

Kosonsiriluk, S., Sartsoongnoen, N., Chaiyachet, O. A., Prakobsaeng, N., Songserm, T., Rozenboim, I., et al. (2008). Vasoactive intestinal peptide and its role in continuous and seasonal reproduction in birds. *Gen. Comp. Endocrinol.* 159, 88–97. doi: 10.1016/j.ygcen.2008.07.024

Kubokawa, K., Tando, Y., and Roy, S. (2010). Evolution of the reproductive endocrine system in chordates. *Integr. Comp. Biol.* 50, 53–62. doi: 10.1093/icb/icq047

Kuwayama, T., Shimada, K., Saito, N., Ohkubo, T., Sato, K., Wada, M., et al. (1992). Effects of removal of chicks from hens on concentrations of prolactin, luteinizing hormone and oestradiol in plasma of brooding gifujidori hens. *J. Reprod. Fertil.* 95, 617–622. doi: 10.1530/jrf.0.0950617

Kwong, A. K., Ng, A. H., Leung, L. Y., Man, A. K., and Woo, N. Y. (2009). Effect of extracellular osmolality and ionic levels on pituitary prolactin release in euryhaline silver sea bream (*Sparus sarba*). *Gen. Comp. Endocrinol.* 160, 67–75. doi: 10.1016/j.ygcen.2008.10.024

Labella, F., Dular, R., Queen, G., and Vivian, S. (1975). Anterior pituitary hormone release in vitro inversely related to extracellular osmolarity. *Endocrinology* 96, 1559–1565. doi: 10.1210/endo-96-6-1559

Lam, K. S. (1991). Vasoactive intestinal peptide in the hypothalamus and pituitary. *Neuroendocrinology* 53(Suppl. 1), 45–51. doi: 10.1159/000125795

Le Tissier, P. R., Hodson, D. J., Martin, A. O., Romano, N., and Mollard, P. (2015). Plasticity of the prolactin (prl) axis: mechanisms underlying regulation of output in female mice. *Adv. Exp. Med. Biol.* 846, 139–162. doi: 10.1007/978-3-319-12114-7_6

Lea, R. W., and Sharp, P. J. (1991). Effects of presence of squabs upon plasma concentrations of prolactin and lh and length of time of incubation in ringdoves on "extended" incubatory patterns. *Horm. Behav.* 25, 275–282. doi: 10.1016/0018-506x(91)90001-x

Ledoux, J. E., Ruggiero, D. A., Forest, R., Stornetta, R., and Reis, D. J. (1987). Topographic organization of convergent projections to the thalamus from the inferior colliculus and spinal cord in the rat. *J. Comp. Neurol.* 264, 123–146. doi: 10.1002/cne.902640110

Lee, K. M., Kaneko, T., and Aida, K. (2006). Prolactin and prolactin receptor expressions in a marine teleost, pufferfish takifugu rubripes. *Gen. Comp. Endocrinol.* 146, 318–328. doi: 10.1016/j.ygcen.2005.12.003

Leko, A. H., Cservenak, M., and Dobolyi, A. (2017a). Suckling induced insulin-like growth factor-1 (igf-1) release in mother rats. *Growth Horm. IGF Res.* 37, 7–12. doi: 10.1016/j.ghir.2017.10.003

Leko, A. H., Cservenak, M., Szabo, E. R., Hanics, J., Alpar, A., and Dobolyi, A. (2017b). Insulin-like growth factor i and its binding protein-3 are regulators of lactation and maternal responsiveness. *Sci. Rep.* 7:3396.

Li, C., Chen, P., and Smith, M. S. (1999). Neural populations in the rat forebrain and brainstem activated by the suckling stimulus as demonstrated by cfos expression. *Neuroscience* 94, 117–129. doi: 10.1016/s0306-4522(99)00236-5

Lignot, J. H., Cutler, C. P., Hazon, N., and Cramb, G. (2002). Water transport and aquaporins in the european eel (*Anguilla anguilla*). *Symp. Soc. Exp. Biol.* 54, 49–59.

Lin, S. H., Miyata, S., Matsunaga, W., Kawarabayashi, T., Nakashima, T., and Kiyohara, T. (1998). Metabolic mapping of the brain in pregnant, parturient and lactating rats using fos immunohistochemistry. *Brain Res.* 787, 226–236. doi: 10.1016/s0006-8993(97)01484-4

Liu, N. A., Liu, Q., Wawrowsky, K., Yang, Z., Lin, S., and Melmed, S. (2006). Prolactin receptor signaling mediates the osmotic response of embryonic zebrafish lactotrophs. *Mol. Endocrinol.* 20, 871–880. doi: 10.1210/me.2005-0403

Lonstein, J. S., and Stern, J. M. (1997). Role of the midbrain periaqueductal gray in maternal nurturance and aggression: C-fos and electrolytic lesion studies in lactating rats. *J. Neurosci.* 17, 3364–3378. doi: 10.1523/jneurosci.17-09-03364.1997

Lonstein, J. S., and Stern, J. M. (1999). Effects of unilateral suckling on nursing behavior and c-fos activity in the caudal periaqueductal gray in rats. *Dev. Psychobiol.* 35, 264–275. doi: 10.1002/(sici)1098-2302(199912)35:4<264::aid-dev2>3.0.co;2-u

Lopez-Vicchi, F., Ladyman, S. R., Ornstein, A. M., Gustafson, P., Knowles, P., Luque, G. M., et al. (2020). Chronic high prolactin levels impact on gene expression at discrete hypothalamic nuclei involved in food intake. *FASEB J.* 34, 3902–3914. doi: 10.1096/fj.201902357r

Lorenson, M. Y., and Jacobs, L. S. (1987). Osmotic pressure regulation of prolactin and growth hormone release from bovine secretory granules. *Endocrinology* 120, 365–372. doi: 10.1210/endo-120-1-365

Loìpez, M. A. C., Rodriìguez, J. L. R., and Garciìa, M. R. (2003). "Physiological and pathological hyperprolactinemia: can we minimize errors in the clinical practice?," in *Prolactin*, eds G. M. Nagy, and B. E. Toth, (London: IntechOpen).

Lv, C., Mo, C., Liu, H., Wu, C., Li, Z., Li, J., et al. (2018). Dopamine d2-like receptors (drd2 and drd4) in chickens: tissue distribution, functional analysis, and their involvement in dopamine inhibition of pituitary prolactin expression. *Gene* 651, 33–43. doi: 10.1016/j.gene.2018.01.087

Manzon, L. A. (2002). The role of prolactin in fish osmoregulation: a review. *Gen. Comp. Endocrinol.* 125, 291–310. doi: 10.1006/gcen.2001.7746

March, J. B., Sharp, P. J., Wilson, P. W., and Sang, H. M. (1994). Effect of active immunization against recombinant-derived chicken prolactin fusion protein on the onset of broodiness and photoinduced egg laying in bantam hens. *J. Reprod. Fertil.* 101, 227–233. doi: 10.1530/jrf.0.1010227

Massaro, M., Setiawan, A. N., and Davis, L. S. (2007). Effects of artificial eggs on prolactin secretion, steroid levels, brood patch development, incubation onset and clutch size in the yellow-eyed penguin (*Megadyptes antipodes*). *Gen. Comp. Endocrinol.* 151, 220–229. doi: 10.1016/j.ygcen.2007.01.034

Mayer-Gostan, N., and Hirano, T. (1976). The effects of transecting the ixth and xth cranial nerves on hydromineral balance in the eel anguilla anguilla. *J. Exp. Biol.* 64, 461–475.

McKinley, M. J., and Johnson, A. K. (2004). The physiological regulation of thirst and fluid intake. *News Physiol. Sci.* 19, 1–6. doi: 10.1152/nips.01470.2003

Morrissey, S. E., Newth, T., Rees, R., Barr, A., Shora, F., and Laycock, J. F. (2001). Renal effects of recombinant prolactin in anaesthetized rats. *Eur. J. Endocrinol.* 145, 65–71. doi: 10.1530/eje.0.1450065

Naef, L., and Woodside, B. (2007). Prolactin/leptin interactions in the control of food intake in rats. *Endocrinology* 148, 5977–5983. doi: 10.1210/en.2007-0442

Nagaishi, V. S., Cardinali, L. I., Zampieri, T. T., Furigo, I. C., Metzger, M., and Donato, J. Jr. (2014). Possible crosstalk between leptin and prolactin during pregnancy. *Neuroscience* 259, 71–83. doi: 10.1016/j.neuroscience.2013.11.050

Nagy, G. M., Arendt, A., Banky, Z., and Halasz, B. (1992). Dehydration attenuates plasma prolactin response to suckling through a dopaminergic mechanism. *Endocrinology* 130, 819–824. doi: 10.1210/en.130.2.819

Nagy, G. M., Bodnar, I., Banky, Z., and Halasz, B. (2005). Control of prolactin secretion by excitatory amino acids. *Endocrine* 28, 303–308. doi: 10.1385/endo:28:3:303

Namken, S., Sinpru, P., Kamkrathok, B., Sartsoongnoen, N., and Chaiseha, Y. (2017). Role of vasoactive intestinal peptide during the transition from incubation behavior to rearing behavior in the female native thai chicken. *Poult. Sci.* 96, 3768–3774. doi: 10.3382/ps/pex180

Neill, J. D., and Nagy, G. M. (1994). "Prolactin secretion and its control," in *Physiology of Reproduction*, eds E. Knobil, and J. D. Neill, (New York: Raven Press), 1833–1860.

Nishimori, K., Young, L. J., Guo, Q., Wang, Z., Insel, T. R., and Matzuk, M. M. (1996). Oxytocin is required for nursing but is not essential for parturition or reproductive behavior. *Proc. Natl. Acad. Sci. U.S.A.* 93, 11699–11704. doi: 10.1073/pnas.93.21.11699

Nobata, S., Ando, M., and Takei, Y. (2013). Hormonal control of drinking behavior in teleost fishes; insights from studies using eels. *Gen. Comp. Endocrinol.* 192, 214–221. doi: 10.1016/j.ygcen.2013.05.009

Ocampo Daza, D., and Larhammar, D. (2018a). Evolution of the growth hormone, prolactin, prolactin 2 and somatolactin family. *Gen. Comp. Endocrinol.* 264, 94–112. doi: 10.1016/j.ygcen.2018.01.007

Ocampo Daza, D., and Larhammar, D. (2018b). Evolution of the receptors for growth hormone, prolactin, erythropoietin and thrombopoietin in relation to the vertebrate tetraploidizations. *Gen. Comp. Endocrinol.* 257, 143–160. doi: 10.1016/j.ygcen.2017.06.021

Ohkubo, T. (2017). Neuroendocrine control of broodiness. *Adv. Exp. Med. Biol.* 1001, 151–171. doi: 10.1007/978-981-10-3975-1_10

Olah, S., Cservenak, M., Keller, D., Fazekas, E. A., Renner, E., Low, P., et al. (2018). Prolactin-induced and neuronal activation in the brain of mother mice. *Brain Struct. Funct.* 223, 3229–3250. doi: 10.1007/s00429-018-1686-1

On, J. S., Chow, B. K., and Lee, L. T. (2015). Evolution of parathyroid hormone receptor family and their ligands in vertebrate. *Front. Endocrinol. (Lausanne)* 6:28. doi: 10.3389/fendo.2015.00028

Ouhtit, A., Morel, G., and Kelly, P. A. (1993). Visualization of gene expression of short and long forms of prolactin receptor in the rat. *Endocrinology* 133, 135–144. doi: 10.1210/endo.133.1.8319561

Palkovits, M., Usdin, T. B., Makara, G. B., and Dobolyi, A. (2010). Tuberoinfundibular peptide of 39 residues- immunoreactive fibers in the zona incerta and the supraoptic decussations terminate in the neuroendocrine hypothalamus. *Neurochem. Res.* 35, 2078–2085. doi: 10.1007/s11064-010-0292-2

Parker, S. L., Armstrong, W. E., Sladek, C. D., Grosvenor, C. E., and Crowley, W. R. (1991). Prolactin stimulates the release of oxytocin in lactating rats: evidence for a physiological role via an action at the neural lobe. *Neuroendocrinology* 53, 503–510. doi: 10.1159/000125764

Peaker, M., Phillips, J. G., and Wright, A. (1970). The effect of prolactin on the secretory activity of the nasal salt-gland of the domestic duck (*Anas platyrhynchos*). *J. Endocrinol.* 47, 123–127. doi: 10.1677/joe.0.0470123

Phillipps, H. R., Yip, S. H., and Grattan, D. R. (2019). Patterns of prolactin secretion. *Mol. Cell Endocrinol.* 502:110679. doi: 10.1016/j.mce.2019.110679

Pi, X., and Voogt, J. L. (2000). Effect of suckling on prolactin receptor immunoreactivity in the hypothalamus of the rat. *Neuroendocrinology* 71, 308–317. doi: 10.1159/000054551

Pi, X. J., and Grattan, D. R. (1999). Increased prolactin receptor immunoreactivity in the hypothalamus of lactating rats. *J. Neuroendocrinol.* 11, 693–705. doi: 10.1046/j.1365-2826.1999.00386.x

Pickford, G. E., and Phillips, J. G. (1959). Prolactin, a factor in promoting survival of hypophysectomized killifish in fresh water. *Science* 130, 454–455. doi: 10.1126/science.130.3373.454

Proudman, J. A., and Opel, H. (1988). Stimulation of prolactin secretion from turkey anterior pituitary cells in culture. *Proc. Soc. Exp. Biol. Med.* 187, 448–454. doi: 10.3181/00379727-187-42687

Radhakrishnan, A., Raju, R., Tuladhar, N., Subbannayya, T., Thomas, J. K., Goel, R., et al. (2012). A pathway map of prolactin signaling. *J. Cell Commun. Signal.* 6, 169–173.

Reynolds, J. D., Goodwin, N. B., and Freckleton, R. P. (2002). Evolutionary transitions in parental care and live bearing in vertebrates. *Philos. Trans. R Soc. Lond. B Biol. Sci.* 357, 269–281. doi: 10.1098/rstb.2001.0930

Richard-Yris, M. A., Sharp, P. J., Wauters, A. M., Guemene, D., Richard, J. P., and Foraste, M. (1998). Influence of stimuli from chicks on behavior and concentrations of plasma prolactin and luteinizing hormone in incubating hens. *Horm. Behav.* 33, 139–148. doi: 10.1006/hbeh.1998.1444

Romano, N., Yip, S. H., Hodson, D. J., Guillou, A., Parnaudeau, S., Kirk, S., et al. (2013). Plasticity of hypothalamic dopamine neurons during lactation results in dissociation of electrical activity and release. *J. Neurosci.* 33, 4424–4433. doi: 10.1523/jneurosci.4415-12.2013

Romanov, M. N., Talbot, R. T., Wilson, P. W., and Sharp, P. J. (2002). Genetic control of incubation behavior in the domestic hen. *Poult. Sci.* 81, 928–931. doi: 10.1093/ps/81.7.928

Rosenblatt, J. S. (2003). Outline of the evolution of behavioral and nonbehavioral patterns of parental care among the vertebrates: critical characteristics of mammalian and avian parental behavior. *Scand. J. Psychol.* 44, 265–271. doi: 10.1111/1467-9450.00344

Royle, N. J., Russell, A. F., and Wilson, A. J. (2014). The evolution of flexible parenting. *Science* 345, 776–781. doi: 10.1126/science.1253294

Rozenboim, I., Tabibzadeh, C., Silsby, J. L., and el Halawani, M. E. (1993). Effect of ovine prolactin administration on hypothalamic vasoactive intestinal peptide (vip), gonadotropin releasing hormone i and ii content, and anterior pituitary vip receptors in laying turkey hens. *Biol. Reprod.* 48, 1246–1250. doi: 10.1095/biolreprod48.6.1246

Sage, M., and Bern, H. A. (1971). Cytophysiology of the teleost pituitary. *Int. Rev. Cytol.* 31, 339–376. doi: 10.1016/s0074-7696(08)60062-5

Saiardi, A., Bozzi, Y., Baik, J. H., and Borrelli, E. (1997). Antiproliferative role of dopamine: loss of d2 receptors causes hormonal dysfunction and pituitary hyperplasia. *Neuron* 19, 115–126. doi: 10.1016/s0896-6273(00)80352-9

Sakamoto, T., and McCormick, S. D. (2006). Prolactin and growth hormone in fish osmoregulation. *Gen. Comp. Endocrinol.* 147, 24–30. doi: 10.1016/j.ygcen.2005.10.008

Salais-Lopez, H., Lanuza, E., Agustin-Pavon, C., and Martinez-Garcia, F. (2017). Tuning the brain for motherhood: prolactin-like central signalling in virgin, pregnant, and lactating female mice. *Brain Struct. Funct.* 222, 895–921. doi: 10.1007/s00429-016-1254-5

Sandra, O., Le Rouzic, P., Cauty, C., Edery, M., and Prunet, P. (2000). Expression of the prolactin receptor (tiprl-r) gene in tilapia oreochromis niloticus: tissue distribution and cellular localization in osmoregulatory organs. *J. Mol. Endocrinol.* 24, 215–224. doi: 10.1677/jme.0.0240215

Santos, C. R., Ingleton, P. M., Cavaco, J. E., Kelly, P. A., Edery, M., and Power, D. M. (2001). Cloning, characterization, and tissue distribution of prolactin receptor in the sea bream (*Sparus aurata*). *Gen. Comp. Endocrinol.* 121, 32–47. doi: 10.1006/gcen.2000.7553

Sapsford, T. J., Kokay, I. C., Ostberg, L., Bridges, R. S., and Grattan, D. R. (2012). Differential sensitivity of specific neuronal populations of the rat hypothalamus to prolactin action. *J. Comp. Neurol.* 520, 1062–1077. doi: 10.1002/cne.22775

Sartsoongnoen, N., Kosonsiriluk, S., Prakobsaeng, N., Songserm, T., Rozenboim, I., Halawani, M. E., et al. (2008). The dopaminergic system in the brain of the native thai chicken, gallus domesticus: localization and differential expression across the reproductive cycle. *Gen. Comp. Endocrinol.* 159, 107–115. doi: 10.1016/j.ygcen.2008.08.002

Sauve, D., and Woodside, B. (1996). The effect of central administration of prolactin on food intake in virgin female rats is dose-dependent, occurs in the absence of ovarian hormones and the latency to onset varies with feeding regimen. *Brain Res.* 729, 75–81. doi: 10.1016/s0006-8993(96)00227-2

Scott, N., Prigge, M., Yizhar, O., and Kimchi, T. (2015). A sexually dimorphic hypothalamic circuit controls maternal care and oxytocin secretion. *Nature* 525, 519–522. doi: 10.1038/nature15378

Seale, A. P., Watanabe, S., and Grau, E. G. (2012). Osmoreception: perspectives on signal transduction and environmental modulation. *Gen. Comp. Endocrinol.* 176, 354–360. doi: 10.1016/j.ygcen.2011.10.005

Shu, Y., Lou, Q., Dai, Z., Dai, X., He, J., Hu, W., et al. (2016). The basal function of teleost prolactin as a key regulator on ion uptake identified with zebrafish knockout models. *Sci. Rep.* 6:18597.

Sinpru, P., Sartsoongnoen, N., Rozenboim, I., Porter, T. E., El Halawani, M. E., and Chaiseha, Y. (2018). The effects of replacing eggs with chicks on mesotocin, dopamine, and prolactin in the native thai hen. *Gen. Comp. Endocrinol.* 263, 32–42. doi: 10.1016/j.ygcen.2018.04.013

Sjoeholm, A., Bridges, R. S., Grattan, D. R., and Anderson, G. M. (2011). Region-, neuron-, and signaling pathway-specific increases in prolactin responsiveness in reproductively experienced female rats. *Endocrinology* 152, 1979–1988. doi: 10.1210/en.2010-1220

Smiley, K. O. (2019). Prolactin and avian parental care: new insights and unanswered questions. *Horm. Behav.* 111, 114–130. doi: 10.1016/j.yhbeh.2019.02.012

Smiley, K. O., and Adkins-Regan, E. (2016). Relationship between prolactin, reproductive experience, and parental care in a biparental songbird, the zebra finch (*Taeniopygia guttata*). *Gen. Comp. Endocrinol.* 232, 17–24. doi: 10.1016/j.ygcen.2015.11.012

Soares, M. J., Konno, T., and Alam, S. M. (2007). The prolactin family: effectors of pregnancy-dependent adaptations. *Trends Endocrinol. Metab.* 18, 114–121. doi: 10.1016/j.tem.2007.02.005

Specker, J. L., Kishida, M., Huang, L., King, D. S., Nagahama, Y., Ueda, H., et al. (1993). Immunocytochemical and immunogold localization of two prolactin isoforms in the same pituitary cells and in the same granules in the tilapia (*Oreochromis mossambicus*). *Gen. Comp. Endocrinol.* 89, 28–38. doi: 10.1006/gcen.1993.1006

Suarez-Bregua, P., Cal, L., Canestro, C., and Rotllant, J. (2017). PTH reloaded: a new evolutionary perspective. *Front. Physiol.* 8:776. doi: 10.3389/fphys.2017.00776

Suzuki, K., Koizumi, N., Hirose, H., Hokao, R., Takemura, N., and Motoyoshi, S. (2000). Changes in plasma arginine vasopressin concentration during lactation in rats. *Comp Med* 50, 277–280.

Szawka, R. E., Ribeiro, A. B., Leite, C. M., Helena, C. V., Franci, C. R., Anderson, G. M., et al. (2011). Kisspeptin regulates prolactin release through hypothalamic dopaminergic neurons. *Endocrinology* 151, 3247–3257. doi: 10.1210/en.2009-1414

Tacon, P., Baroiller, J. F., Le Bail, P. Y., Prunet, P., and Jalabert, B. (2000). Effect of egg deprivation on sex steroids, gonadotropin, prolactin, and growth hormone profiles during the reproductive cycle of the mouthbrooding cichlid fish oreochromis niloticus. *Gen. Comp. Endocrinol.* 117, 54–65. doi: 10.1006/gcen.1999.7388

Takei, Y., Hiroi, J., Takahashi, H., and Sakamoto, T. (2014). Diverse mechanisms for body fluid regulation in teleost fishes. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 307, R778–R792.

Tindal, J. S., and Knaggs, G. S. (1969). An ascending pathway for release of prolactin in the brain of the rabbit. *J. Endocrinol.* 45, 111–120. doi: 10.1677/joe.0.0450111

Tindal, J. S., and Knaggs, G. S. (1972). Pathways in the forebrain of the rabbit concerned with the release of prolactin. *J. Endocrinol.* 52, 253–262. doi: 10.1677/joe.0.0520253

Tindal, J. S., and Knaggs, G. S. (1977). Pathways in the forebrain of the rat concerned with the release of prolactin. *Brain Res.* 119, 211–221. doi: 10.1016/0006-8993(77)90101-9

Tong, Z., Pitts, G. R., You, S., Foster, D. N., and El Halawani, M. E. (1998). Vasoactive intestinal peptide stimulates turkey prolactin gene expression by increasing transcription rate and enhancing mrna stability. *J. Mol. Endocrinol.* 21, 259–266. doi: 10.1677/jme.0.0210259

Usdin, T. B., Dobolyi, A., Ueda, H., and Palkovits, M. (2003). Emerging functions for tuberoinfundibular peptide of 39 residues. *Trends Endocrinol. Metab.* 14, 14–19. doi: 10.1016/s1043-2760(02)00002-4

Van Goor, F., Zivadinovic, D., Martinez-Fuentes, A. J., and Stojilkovic, S. S. (2001). Dependence of pituitary hormone secretion on the pattern of spontaneous

voltage-gated calcium influx. Cell type-specific action potential secretion coupling. *J. Biol. Chem.* 276, 33840–33846. doi: 10.1074/jbc.m105386200

Voogt, J. L., Lee, Y., Yang, S., and Arbogast, L. (2001). Regulation of prolactin secretion during pregnancy and lactation. *Prog. Brain Res.* 133, 173–185. doi: 10.1016/s0079-6123(01)33013-3

Walker, C. D., Toufexis, D. J., and Burlet, A. (2001). Hypothalamic and limbic expression of crf and vasopressin during lactation: implications for the control of acth secretion and stress hyporesponsiveness. *Prog. Brain Res.* 133, 99–110. doi: 10.1016/s0079-6123(01)33008-x

Wan, X. P., Xie, P., Bu, Z., Zou, X. T., and Gong, D. Q. (2019). Prolactin induces lipid synthesis of organ-cultured pigeon crops. *Poult. Sci.* 98, 1842–1853. doi: 10.3382/ps/pey540

Watanabe, H., Davis, J. B., Smart, D., Jerman, J. C., Smith, G. D., Hayes, P., et al. (2002). Activation of trpv4 channels (hvrl-2/mtrp12) by phorbol derivatives. *J. Biol. Chem.* 277, 13569–13577. doi: 10.1074/jbc.m200062200

Watanabe, S., Hirano, T., Grau, E. G., and Kaneko, T. (2009). Osmosensitivity of prolactin cells is enhanced by the water channel aquaporin-3 in a euryhaline mozambique tilapia (*Oreochromis mossambicus*). *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 296, R446–R453.

Watanabe, S., Kaneko, T., and Aida, K. (2005). Aquaporin-3 expressed in the basolateral membrane of gill chloride cells in mozambique tilapia *Oreochromis mossambicus* adapted to freshwater and seawater. *J. Exp. Biol.* 208, 2673–2682. doi: 10.1242/jeb.01684

Weber, G. M., Seale, A. P., Richman, I. N., Stetson, M. H., Hirano, T., and Grau, E. G. (2004). Hormone release is tied to changes in cell size in the osmoreceptive prolactin cell of a euryhaline teleost fish, the tilapia, *Oreochromis mossambicus*. *Gen. Comp. Endocrinol.* 138, 8–13. doi: 10.1016/j.ygcen.2004.04.006

Whittington, C. M., and Wilson, A. B. (2013). The role of prolactin in fish reproduction. *Gen. Comp. Endocrinol.* 191, 123–136. doi: 10.1016/j.ygcen.2013.05.027

Witcher, J. A., and Freeman, M. E. (1985). The proestrous surge of prolactin enhances sexual receptivity in the rat. *Biol. Reprod.* 32, 834–839. doi: 10.1095/biolreprod32.4.834

Wu, Z., Autry, A. E., Bergan, J. F., Watabe-Uchida, M., and Dulac, C. G. (2014). Galanin neurons in the medial preoptic area govern parental behaviour. *Nature* 509, 325–330. doi: 10.1038/nature13307

Xu, H., Shen, X., Zhou, M., Fang, M., Zeng, H., Nie, Q., et al. (2010). The genetic effects of the dopamine d1 receptor gene on chicken egg production and broodiness traits. *BMC Genet.* 11:17. doi: 10.1186/1471-2156-11-17

Yamaguchi, Y., Moriyama, S., Lerner, D. T., Grau, E. G., and Seale, A. P. (2016). Autocrine positive feedback regulation of prolactin release from tilapia prolactin cells and its modulation by extracellular osmolality. *Endocrinology* 157, 3505–3516. doi: 10.1210/en.2015-1969

Yip, S. H., Romano, N., Gustafson, P., Hodson, D. J., Williams, E. J., Kokay, I. C., et al. (2019). Elevated prolactin during pregnancy drives a phenotypic switch in mouse hypothalamic dopaminergic neurons. *Cell Rep.* 26, 1787–1799.

Youngren, O. M., el Halawani, M. E., Phillips, R. E., and Silsby, J. L. (1989). Effects of preoptic and hypothalamic lesions in female turkeys during a photoinduced reproductive cycle. *Biol. Reprod.* 41, 610–617. doi: 10.1095/biolreprod41.4.610

Youngren, O. M., el Halawani, M. E., Silsby, J. L., and Phillips, R. E. (1991). Intracranial prolactin perfusion induces incubation behavior in turkey hens. *Biol. Reprod.* 44, 425–431. doi: 10.1095/biolreprod44.3.425

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neuroscience

REVIEW
published: 10 July 2020
doi: 10.3389/fnins.2020.00691

# Evolution of Orexin Neuropeptide System: Structure and Function

*Shingo Soya and Takeshi Sakurai**

*Faculty of Medicine/International Institute for Integrative Sleep Medicine (WPI-IIIS), University of Tsukuba, Tsukuba, Japan*

Orexins are hypothalamic neuropeptides that were initially identified in the rat brain as endogenous ligands for an (previously) orphan G-protein-coupled receptor (GPCR). They are multitasking peptides involved in many physiological functions, including regulation of feeding behavior, wakefulness and autonomic/neuroendocrine functions, and sleep/wakefulness states in mammals. There are two isopeptides of orexin, orexin A and orexin B, which are produced from a common precursor peptide, prepro-orexin. Structures of orexins, as well as orexin genes, are highly conserved throughout mammalian species, suggesting strong evolutionary pressure that maintains the structures. Their lengths and structure suggested that orexin B is the ancestral form of the orexin neuropeptide. In mammals, orexins bind to two subtypes of GPCRs, i.e., orexin 1 receptor (OX1R) and orexin 2 receptor (OX2R). Phylogenetically, the orexin system is present exclusively in vertebrates. In genomes of species outside mammals, there is only one orexin receptor, which is similar to OX2R, suggesting that OX2R is the prototype receptor for orexins. OX1R is likely to have evolved during early mammalian evolution. Orexin-producing neurons (orexin neurons) are mainly located in the lateral hypothalamic area (LHA) in mammals and are also found in hypothalamic regions in many other vertebrates. Orexins are likely to be closely related to the regulation of active, motivated behavior in many species. The orexin system seems to have evolved as a system that supports active and purposeful behavior which is closely related with wakefulness.

**Keywords: neuropeptide, orexin, OX1R, OX2R, hypothalamus, vertebrate**

OPEN ACCESS

***Edited by:***
*Jackson Cioni Bittencourt, University of São Paulo, Brazil*

***Reviewed by:***
*Thomas Kilduff, SRI International, United States*
*Tomomi Tsunematsu, Tohoku University, Japan*

***Correspondence:**
*Takeshi Sakurai*
*sakurai.takeshi.gf@u.tsukuba.ac.jp; takeshi.sakurai@gmail.com*

***Specialty section:***
*This article was submitted to Neuroendocrine Science, a section of the journal Frontiers in Neuroscience*

***Received:*** *03 April 2020*
***Accepted:*** *08 June 2020*
***Published:*** *10 July 2020*

***Citation:***
*Soya S and Sakurai T (2020) Evolution of Orexin Neuropeptide System: Structure and Function. Front. Neurosci. 14:691. doi: 10.3389/fnins.2020.00691*

## INTRODUCTION

Orexins were initially recognized as regulators of feeding behavior. Subsequently, the finding that orexin deficiency causes narcolepsy in several mammalian species revealed that orexins play a critical role in regulation of sleep/wakefulness states, especially in maintenance of wakefulness in mammals. Orexins were also shown to be involved in the regulation of a wide range of physiological

**Abbreviations:** A, anterior; BNST, bed nucleus of the stria terminalis; Chr, chromosome number; D, dorsal; DMH, dorsomedial nuclei of hypothalamus; DR, dorsal raphe; GPCR, G-protein-coupled receptor; Hd, dorsal zone of periventricular hypothalamus; Hv, ventral zone of periventricular hypothalamus; ICV, intracerebroventricular; LC, locus coeruleus; LHA, lateral hypothalamic area; NAc, nucleus accumbens; NLT, nucleus lateralis tuberis; NPPv, nucleus posterioris periventricularis; NPY, neuropeptide Y; OC, optic chiasm; OX1R, orexin 1 receptor; OX2R, orexin 2 receptor; P, posterior; PaF, parafornical nucleus; PeF, perifornical hypothalamus; PHN, periventricular hypothalamic nucleus; POA, preoptic area; PVN, paraventricular nucleus; SCN, suprachiasmatic nucleus; TMN, tuberomammillary nucleus; V, ventral; 3V, third ventricle; VM, ventromedial thalamic nucleus.

functions, suggesting that orexins are multitasking peptides. Any purposeful behavior requires certain internal body states, including appropriate tuning of the autonomic nervous system and endocrine function. Maintenance of wakefulness and vigilance is also important for pursuing behaviors, because appropriate arousal levels are especially necessary for executing any purposeful behavior that requires high motivation. The systems involved in these functions are closely related and are interconnected with the orexin system (Sakurai, 2014). Orexin-producing neurons (orexin neurons), which locate in the LHA, receive input by forebrain structures including the extended amygdala and nucleus accumbens (NAc)—which are implicated in the processing of emotion and motivation—and send output to brain stem regions, which are implicated in the regulation of wakefulness. Orexin neurons play an important role as a link between emotional states and wakefulness states.

In non-mammalian species, sleep/wakefulness states are generally solely defined by behavioral criteria, and wakefulness is defined as a state with active behavior. Orexins are likely to play an important role in regulation of active behavior also in non-mammalian species, and these factors are also recognized as regulators of wakefulness. This review focuses on how the structures of orexins and their receptors, neuronal circuits, and their functions have evolved in the animal kingdom.

## SUMMARY OF THE MAMMALIAN OREXIN SYSTEM

The hypothalamus plays a central role in the integrated control of feeding and energy homeostasis. We identified two novel neuropeptides, both derived from the same precursor by proteolytic processing, that bind and activate two closely related previously orphan GPCRs. These peptides, termed orexins A and B, had no significant structural similarities to known families of regulatory peptides (Sakurai et al., 1998). *Prepro-orexin* mRNA and immunoreactive orexin are specifically localized in neurons within and around the lateral and posterior hypothalamus in the adult rat brain. When administered centrally to rats, these peptides increased food consumption. *Prepro-orexin* mRNA level is upregulated by fasting, suggesting a physiological role of these peptides as mediators in the central feedback mechanism that regulates feeding behavior (Sakurai et al., 1998). Molecular cloning studies showed that orexins A and B are derived from a common precursor peptide, prepro-orexin. An mRNA encoding the same precursor peptide was independently identified by De Lecea et al. (1998) as a hypothalamus-specific transcript. The authors predicted that the transcript encoded a polypeptide precursor that is cleaved to form two neuropeptides, termed hypocretin-1 and hypocretin-2 (corresponding to orexins A and B, respectively).

Our structural analysis of the purified peptides revealed that orexin A is a 33-amino-acid peptide with an N-terminal pyroglutamyl residue, two intrachain disulfide bonds, and C-terminal amidation. Strikingly, this structure is completely conserved among all mammalian species so far identified (human, gorilla, rat, mouse, cow, pig, sheep, dog, seal, and dolphin). Mammalian orexin B is a 28-amino-acid, C-terminally amidated linear peptide, which also has a highly conserved structure among mammalian species. The C-terminal half of orexin B is very similar to that of orexin A, whereas the N-terminal half is more variable (Sakurai et al., 1998) (**Figure 2**). The unusually conserved structures of orexins suggest strong evolutionary pressure that preserves the structure, which is likely to be related with the function of these peptides.

The best understood role of orexins in mammals is regulation of sleep and wakefulness states, especially in the maintenance of long, consolidated wakefulness. This is highlighted by findings that orexin deficiency caused narcolepsy in several mammalian species including mice, rats, dogs, and humans (Chemelli et al., 1999; Lin et al., 1999; Peyron et al., 2000; Thannickal et al., 2000; Sakurai, 2007). Sleep and wakefulness are regulated to occur at appropriate times, in accordance with the internal and external environments. Avoiding danger and finding food, which are life-essential activities that are regulated by emotion, reward, and energy balance, require vigilance and therefore, by definition, wakefulness. The orexin system is involved in these functions (Sakurai et al., 1998; Yamanaka et al., 2003). Other than that, orexin has been implicated in a variety of functions including regulation of food intake, emotion, the reward system, and the autonomic nervous system. These functions of orexins are mediated by two GPCRs, OX1R and OX2R. OX1R has a greater affinity for orexin A over orexin B, whereas OX2R binds both ligands with similar affinities. Orexin receptors exhibit a markedly different distribution. They are abundantly expressed by monoaminergic neurons in the brain stem (Mieda et al., 2011). Orexin neurons, which have been assumed to number around 3,000 in the rat brain and around 70,000 in the human brain, are localized exclusively in the hypothalamus, including the LHA, perifornical area, and posterior hypothalamus. These neurons send widespread projections to the brain, with particularly dense projections to monoaminergic and cholinergic nuclei in the brain stem, where OX1R and OX2R are differentially expressed.

The functions of orexins and the architecture of orexin neurons are also highly conserved among mammalian species. On the other hand, orexin-like genes are not found in invertebrates, suggesting that the orexin system originated in early vertebrates.

## EVOLUTION OF OREXIN GENES AND PEPTIDES

Thanks to genome research and previous molecular cloning studies, the amino acid sequences of orexins in several mammalian species (human, gorilla, golden monkey, baboon, gibbon, mouse, rat, pig, dog, camel, alpaca, seal, and dolphin) (Sakurai et al., 1998; Dyer et al., 1999; Peyron et al., 2000; Elbers et al., 2019), as well as reptiles (cobra and turtle) (Vonk et al., 2013), amphibians (*Xenopus laevis*) (Shibahara et al., 1999), birds (chicken, turkey, and finch) (Ohkubo et al., 2002), and fish (goldfish, zebrafish, cod, stickleback, medaka, pufferfish) (Kaslin and Panula, 2001), are currently available.

The genes encoding *prepro-orexin* show highly conserved loci throughout vertebrate evolution, including the two-exon structure, with a larger exon 2, which includes sequences encoding orexins A and B. Exon 1 generally contains 5'-UTR and part of the signal sequence. In the *prepro-orexin* sequence, orexin A sequences are directly preceded by signal peptides. Both mature peptides are followed by a putative consensus motif for C-terminal amidation (G-R/K-R/K) (**Figures 1**, **2**). The C-terminal regions of *prepro-orexin* sequences following the orexin sequence are variable among species, suggesting that no other functional peptides are encoded in the region. Rat orexin A has been purified, and its structure analyzed by peptide sequencing and mass spec analyses. It has a 33-amino-acid peptide sequence with two intrachain disulfide bridges formed by four cysteine residues (C6–C12 and C7–C14), an N-terminal glutamate residue, and C-terminal amidation. The primary sequence of mammalian orexin A was shown to be modified to have an N-terminal pyroglutamyl residue and C-terminal amidation. The structure of orexin A is completely conserved among all mammalian species thus far identified. Mammalian orexin B is a 28-amino-acid linear peptide not having disulfide bridges and has minor amino acid differences among mammalian species. In particular, the second amino acid residue is P or S depending on the species (**Figure 2**).

Non-mammalian orexins are also very similar to mammalian orexin A but show more variations than those in mammals. Generally, there is no N-terminal pyroglutamylation in the orexin A structure due to lack of an N-terminal glutamate residue.

Bird (chicken, turkey, and finch) orexin A sequences contain 34 amino-acid residues and have two intra-disulfide bridges (C7–C13 and C8–C15), while orexin B sequences are 28-amino-acid linear peptides.

In reptiles, turtle (*Terrapene carolina triunguis*) orexin A is 34 amino acids long and is predicted to have a similar structure to that of mammalian orexin A, including two intra-disulfide bridges (C7–C13 and C8–C15), while orexin B is 28 amino acids long, which is the same as mammalian orexin B. C-terminal residues (L and M for orexins A and B, respectively) are likely to be amidated as mammalian orexin B, being predicted from glycine residues preceding dibasic pairs of amino acids. Snake (cobra) orexin A is 32 amino acids long and has two intra-disulfide bridges (C5–C11 and C6–C13).

In amphibians, *Xenopus laevis* orexin A is a 31-amino-acid-residue peptide and has six-amino-acid substitutions when compared with human orexin A (Shibahara et al., 1999). *Xenopus* orexin A does not have an N-terminal pyroglutamate residue either. The relative positions of the four cysteine residues (positions 4, 5, 10, and 12) are well conserved, and it is predicted to form two intrachain disulfide bonds (C4–C10 and C5–C12).

Teleost orexin A sequences are generally much longer than mammalian orexin A. For example, *Fugu* orexin A has 43 amino acid residues. Goldfish and zebrafish orexin A have 47 amino acid residues, and cod orexin A has 50 amino acid residues. These longer sequences are due to the existence of an additional sequence between residues 24 and 25 (Kaslin and Panula, 2001; Xu and Volkoff, 2007). The inserted sequences are non-detrimental to orexin activity (Kaslin and Panula, 2001). Because teleost orexin A does not have C12, it does not have a disulfide bond between C6 and C12 as found in mammalian orexin A, although it is likely to form another disulfide bridge with a cysteine positioned at 21. Teleost orexin B consists of 28 amino acid residues, which is the same as mammals' and other species' orexin B (Kaslin and Panula, 2001), with the exception of cod orexin B (29 amino acids) (Xu and Volkoff, 2007).

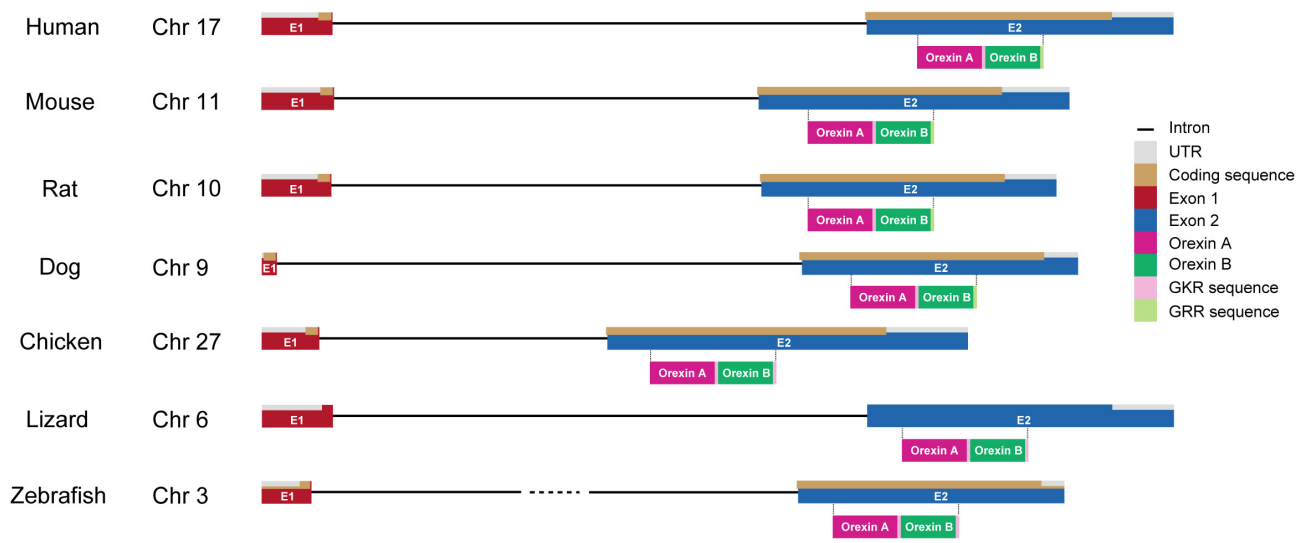


**FIGURE 1 |** Gene structure of orexin in vertebrates. Chromosome numbers (Chr) in which the orexin gene is located are shown in each vertebrate. The drawing shows the length of gene structures of orexin including intron (black line), UTR (gray), coding sequences (brown), exon 1 (E1; red), and exon 2 (E2; blue) in different species (human, mouse, rat, dog, chicken, lizard, and zebrafish). Amino acid sequences of orexin A (magenta) and orexin B (dark green) with GKR (pink) and/or GRR (light green) amidation are also shown under their coding sequences.

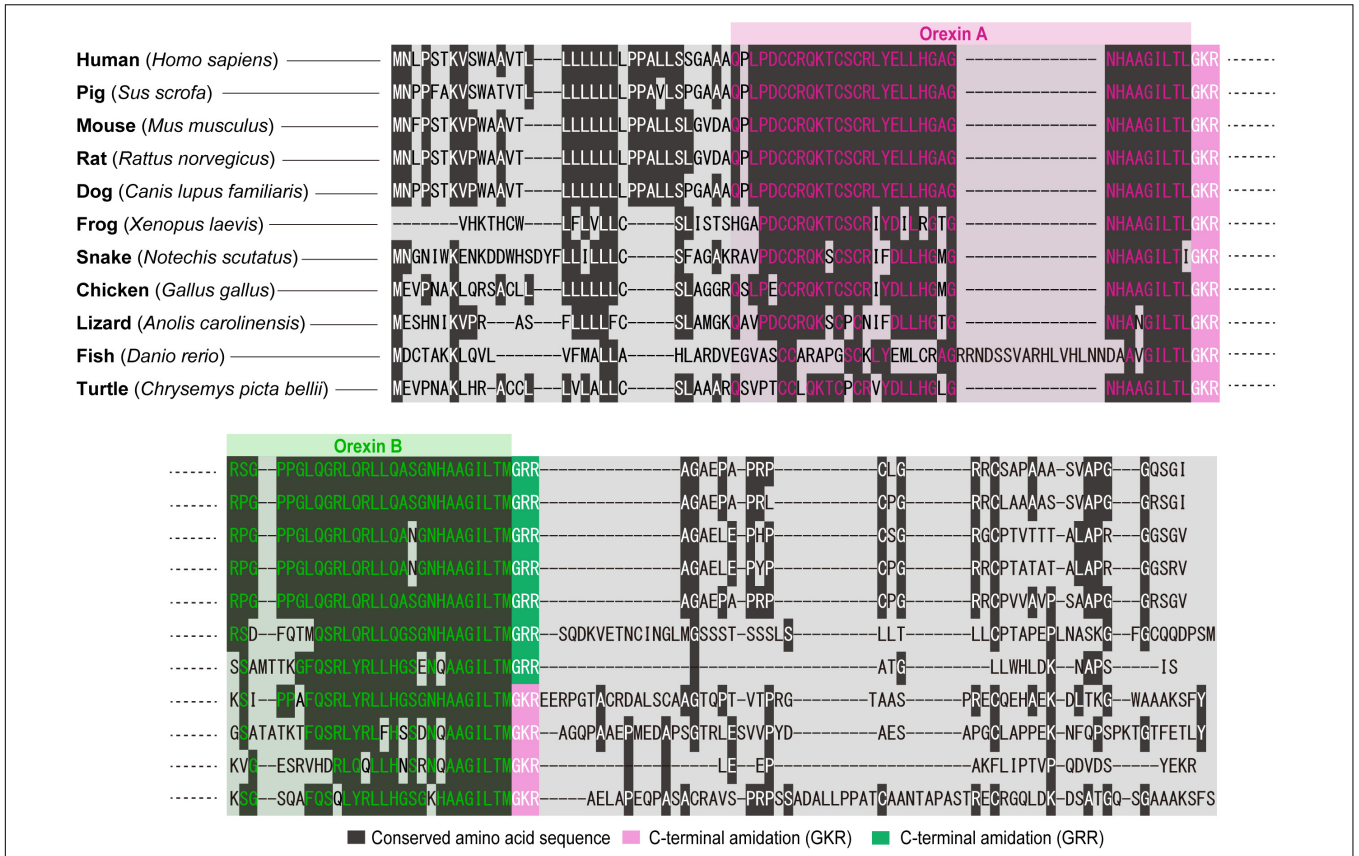


**FIGURE 2 |** Overview of amino acid sequences of orexins in vertebrates. The cartoon shows the sequences of amino acids coding orexins A and B in different species (human: *Homo sapiens*; pig: *Sus scrofa*; mouse: *Mus musculus*; rat: *Rattus norvegicus*; dog: *Canis lupus familiaris*; frog: *Xenopus laevis*; snake: *Notechis scutatus*; chicken: *Gallus gallus*; lizard: *Anolis carolinensis*; fish: *Danio rerio*; and turtle: *Chrysemys picta bellii*). Conserved amino acid sequence (dark gray), C-terminal amidation (GKR, pink), and C-terminal amidation (GRR, green) are highlighted. Amino acid sequences are aligned by ClustalW algorithm using MEGA software (Stecher et al., 2020).

Overall, structures of orexins are exceptionally well conserved in the animal kingdom from fish to mammalian species. Teleost orexin A and human orexin A still have over 52% amino acid identity. The lengths and structures of orexin B are well conserved among species as compared with orexin A, suggesting that orexin B might be a prototype of orexin peptides.

## EVOLUTION OF OREXIN RECEPTORS

In mammals, there are two orexin receptor subtypes, OX1R and OX2R, both of which are members of the class B GPCRs. Orexin A shows similar affinity to both OX1R and OX2R, while orexin B shows higher affinity to OX2R over OX1R. The human OX1R and OX2R genes are located on chromosomes 1 and 6, respectively. Human OX1R and OX2R share 63.5% amino acid identity. They have also similarity to several other peptide receptors. For example, human neuropeptide FF receptor 1 shows 25.1 and 31.2% amino acid identity to OX1R and OX2R, respectively (Sakurai et al., 1998).

OX1R is exclusively found in mammalian species and is thought to have evolved from ancestral OX2R, presumably through gene duplication events during the evolution of early mammals. OX2R is present in all vertebrate genomes, suggesting that OX2R is the ancestral form of orexin receptors. The chromosomal localization of these receptors also suggests that OX1R is a product of a relatively recent gene duplication event from OX2R. The flanking genes of OX2R (FAM83B and GFRAL) are also well conserved in all known mammalian species. While TINAGL1 and PEF1 genes are in close proximity to the mammalian OX1R gene, they are not found in paralogous regions in non-mammalian genomes.

Because OX1R was emerged later than OX2R phylogenetically, it seems to play more complex physiological roles. We previously found that OX1R-deficient mice show anxiety-like behavior (Abbas et al., 2015). We also showed that OX1R in noradrenaline neurons in the locus coeruleus (LC) plays a role in the expression and/or consolidation of cued fear memory by exciting these neurons that send innervations to the lateral amygdala (Soya et al., 2013). Furthermore, this pathway was also involved in generalization of fear memory (Soya et al., 2017; Soya and Sakurai, 2018). OX1R was also shown to be involved in an increase of response to conditioned cues to activate motivational responses in rats (Sharf et al., 2010; Bentzley and Aston-jones,

2015) and in reward-based feeding (Kakizaki et al., 2019). These observations suggest that OX1R plays a role in emotive and motivational functions in mammals.

# EVOLUTION OF OREXIN NEURONAL SYSTEM

Orexin neurons are localized in the LHA and adjacent regions, including the dorsomedial and posterior hypothalamus, in all mammalian species (Peyron et al., 1998; Date et al., 1999; Nambu et al., 1999). These neurons send widespread axonal projections to all over the brain except the cerebellum, with especially abundant projections to monoaminergic nuclei in the brain stem. In mammals, orexin neurons receive and integrate internal and external information and regulate the autonomic and neuroendocrine systems to stabilize arousal and behavior accordingly. The hypothalamus is the main region in which orexinergic neurons are localized in many species among vertebrates. In the chicken, orexin neurons are also exclusively found in the hypothalamus, including the paraventricular hypothalamic nucleus (PVN) and LHA (Miranda et al., 2013). In reptiles, orexin neurons are also found in the hypothalamus. In the *Pseudemys scripta elegans* (turtle) and *Anolis carolinensis* (lizard), these neurons are localized in the PVN, while in the *Gekko gecko* (lizard), these cells are found in the dorsomedial nuclei (DMH) (Farrell et al., 2003; Domínguez et al., 2010).

In amphibians, orexins are found in the hypothalamus but are widely distributed in several regions outside the hypothalamus. These cells are localized especially in the suprachiasmatic nucleus (SCN) and to a lesser extent in the preoptic area (POA) and tuberal region in anurans, urodeles, and gymnophionans (**Figure 3**). Orexin-immunoreactive fibers innervate the whole-brain region, especially the POA (Shibahara et al., 1999; Galas et al., 2001; Singletary et al., 2005; Suzuki et al., 2008; López et al., 2009).

In fish, distribution of orexin neurons is more variable among species. For example, orexin neurons are localized in the POA and SCN in the lungfish and zebrafish. These neurons are found in the nucleus posterioris periventricularis (NPPv) in the medaka and in the NPPv and nucleus lateralis tuberis (NLT) in the goldfish. In zebrafish, orexin neurons are also localized in the dorsal part of the hypothalamus (**Figure 3**). Orexin fibers were shown to innervate the monoaminergic nuclei, including the dorsal raphe (DR), LC, mesopontine-like area, dopaminergic clusters, and histaminergic neurons in the tuberomammillary nucleus (TMN), showing resemblance to the mammalian orexin system (Kaslin and Panula, 2001; Huesa et al., 2005; Nakamachi et al., 2006; Amiya et al., 2007; Kojima et al., 2009). These findings suggest that orexin neurons are basically found in the hypothalamus and send rich projections to monoaminergic neurons. This basic structure is conserved in vertebrate evolution.

# EVOLUTION OF OREXIN FUNCTIONS

In mammals, orexin neurons receive and integrate internal and external information and regulate the autonomic and neuroendocrine systems during performance of various purposeful activities that require arousal. In this section, we discuss the evolution of the orexin system, especially focusing on two main functions of orexin, i.e., regulation of feeding behavior and wakefulness.

## Food Intake and Body Weight Regulation

Orexins were initially reported as factors that regulate feeding behavior, mainly because orexin neurons are distributed within the LHA (and adjacent regions), which is known to be implicated in the regulation of feeding behavior (Sakurai et al., 1998). An orexigenic effect of intracerebroventricular (ICV) administration of orexins A and B in rats was first reported in 1998, and this effect has been subsequently confirmed in many species, including mammals and other species, including fish (Nakamachi et al., 2006). Importantly, orexin signaling increased not only food intake but also energy expenditure, and an increase in the overall orexin tone generally results in decreased body weight (Funato et al., 2009). Likewise, narcoleptic mice, which lack orexin signaling, show mild obesity, especially when fed a high-fat diet.

The orexin system may contribute to the regulation of energy homeostasis by integrating information regarding metabolic state and regulating sleep/wakefulness states in order to support feeding behavior (Sakurai et al., 1998; Haynes et al., 2000; Yamada et al., 2000; Yamanaka et al., 2003; Funato et al., 2009; Sakurai, 2014). Indeed, mice lacking orexin neurons do not show an increase in wakefulness or locomotor activity in response to starvation, unlike wild-type mice (Yamanaka et al., 2003). Moreover, *prepro-orexin* mRNA is upregulated in fasted animals, and several studies report that the firing rate of orexin neurons is influenced by glucose, triglycerides, and amino acids (Chang et al., 2004; Burdakov et al., 2005; Karnani et al., 2011). Furthermore, orexin neurons are directly inhibited by leptin and excited by ghrelin and are innervated by neurons in the arcuate nucleus, which is the primary sensor for plasma leptin level (Elias et al., 1998; Yamanaka et al., 2003). Together, these observations suggest that orexin neurons sense the animal's metabolic and nutritional status and integrate this information in order to evoke arousal necessary to promote food-seeking behavior in response to a negative energy balance. The precise mechanisms by which orexins regulate feeding behavior are detailed in our previous review papers (Sakurai, 2007, 2014).

Other than in mammalian species, the roles of orexins in the regulation of food intake are not very clear. In birds, orexin neurons and fibers are present in the PVN and LHA. This distribution is similar to that of mammalian orexin neurons. However, mammalian orexins did not increase food intake in birds (chicken and pigeon) (Furuse et al., 1999; da Silva et al., 2008; Katayama et al., 2010). However, in these studies, orexins were administered during the light period, which is the active period for birds, when orexin neuronal activity might be highest in the day. This may explain why additional orexin activity did not increase food intake. In fact, when administered in the dark period, orexins do not increase food intake even in rodents, which are nocturnal animals. Also, studies using avian orexin peptides, which are structurally different from mammalian orexins, are necessary to confirm whether orexins play roles in regulation of feeding behavior.

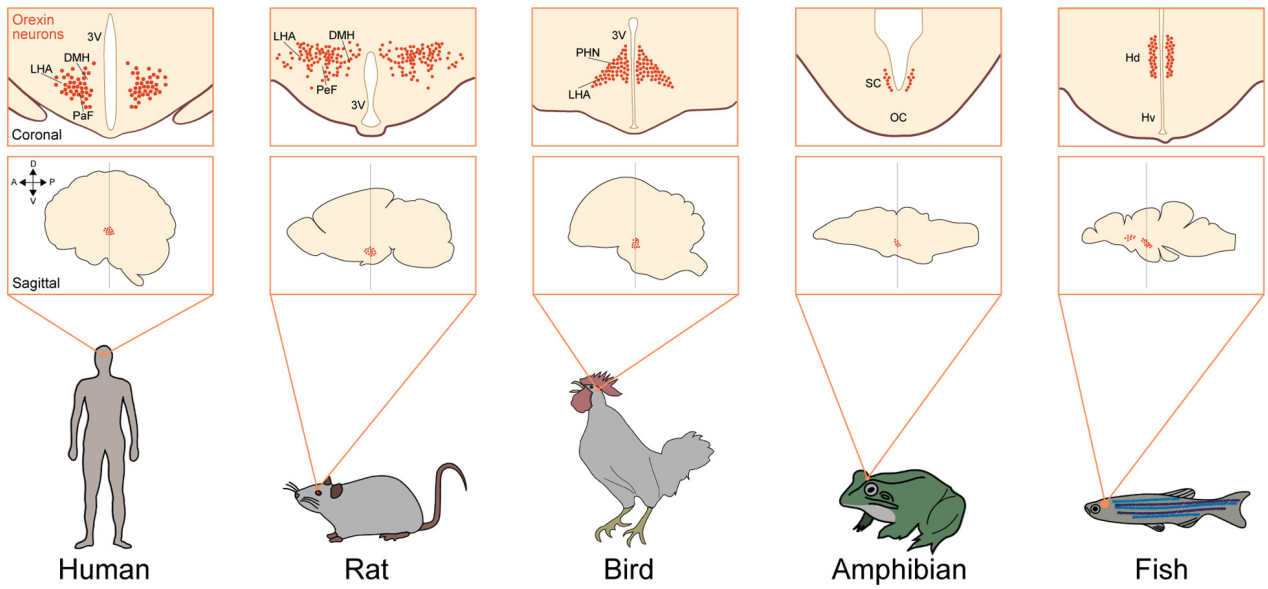


**FIGURE 3 |** Neuronal system of orexin is highly conserved in vertebrates. Upper squares show a coronal view of the brain focusing on orexin neurons (red dots) in different species [human (Peyron et al., 2000), rat (Nambu et al., 1999), bird (Ohkubo et al., 2002), amphibian (López et al., 2009), and fish (Appelbaum et al., 2009)]. Lower squares show a sagittal view of the brain focusing on the populations of orexin neurons in different species. Arrows show the anatomical orientation (A, anterior; P, posterior; D, dorsal; and V, ventral). DMH, dorsomedial nuclei of hypothalamus; LHA, lateral hypothalamic area; PaF, parafornical nucleus; 3V, third ventricle; PeF, perifornical hypothalamus; PHN, periventricular hypothalamic nucleus; VM, ventromedial thalamic nucleus; SC, suprachiasmatic nucleus; OC, optic chiasm; Hd, dorsal zone of periventricular hypothalamus; Hv, ventral zone of periventricular hypothalamus.

There are a substantial number of reports about the involvement of orexins in the regulation of feeding behavior in fish. ICV injection of human orexins increased food intake in goldfish (Volkoff et al., 1999; Nakamachi et al., 2006), and fasting increased *prepro-orexin* mRNA levels in zebrafish, as in mammals (Novak et al., 2005). Like mammals, a reciprocal relationship between orexins and ghrelin was reported in fish. Ghrelin increased the expression of *prepro-orexin* mRNA in the goldfish diencephalon when administered ICV and vice versa (Miura et al., 2007). Both neuropeptide Y (NPY)- and ghrelin-induced food intake were completely inhibited by application of an orexin receptor antagonist (Miura et al., 2007). The relationship between orexin and NPY was also shown by the colocalization of these peptides in the NPPv (Miura et al., 2007).

In other species, the roles of orexins in the regulation of feeding behavior have not been clear so far.

## Sleep/Wakefulness State Regulation

The involvement of orexins in the regulation of sleep/wakefulness states in mammals has been extensively discussed in detail in many review articles (Sakurai, 2007, 2014). The finding that orexin deficiency caused narcolepsy in humans and other mammalian species, like mice, rats, and dogs, clearly indicates that orexin plays an important role in the maintenance of long, consolidated wakefulness in mammals.

Other than in mammals, the roles of orexins in the regulation of sleep/wakefulness states are not very clear, but human orexin A induced dose-dependent arousal- and alertness-promoting behavioral effects in birds (chicken and pigeon) when administered ICV, along with a decrease in duration of sleep-like postures (da Silva et al., 2008; Katayama et al., 2010).

In zebrafish, sleep is usually defined solely by behavioral criteria based on periods of quiescence associated with a specific posture (Hendricks et al., 2000; Tononi, 2000; Zhdanova et al., 2001; Raizen et al., 2008). Genetic ablation of orexin neurons demonstrated an increase in sleep time and sleep/wakefulness transition in the daytime, with no effect on basal locomotor activity in zebrafish (Elbaz et al., 2012). Consistently, global overexpression of the orexin gene by an inducible heat-shock promoter showed an increase in wakefulness, defined by active behavior (Prober et al., 2006). A recent study showed that orexin-induced arousal is regulated via noradrenaline signaling in zebrafish (Singh et al., 2015). On the contrary, zebrafish orexins were reported to be involved in melatonin production in the pineal gland during the dark time, to regulate sleep consolidation (Appelbaum et al., 2009). Evolutional sleep loss was reported in the Mexican cavefish, *Astyanax mexicanus*, depending on their ecological conditions. The populations living in caves are blind, and their sleep time is shorter than that of other eyed populations living in surface rivers (Duboué et al., 2011). A recent study reported that the mechanism of this difference could stem from genetic and neuronal changes of orexins in the hypothalamus (Jaggard et al., 2018).

## Other Functions

Arousal responses are tightly associated with the physiological responses elicited by salient emotional stimuli. Several studies have suggested the involvement of orexins in regulating

emotional behavior. A possible role of orexins in panic disorders has been reported in human and animal studies (Johnson et al., 2010). The LHA is known as the "defense area," and orexins have functions to increase cardiovascular activity and stress response (Wilson et al., 2001; Sakamoto et al., 2004; Winsky-Sommerer et al., 2004; Zhang et al., 2009; Sakurai, 2014). Orexin neurons receive dense innervations from limbic structures like the BNST and the amygdala (González et al., 2016; Saito et al., 2018), suggesting that orexins regulate autonomic/neuroendocrine functions in response to emotional stimuli in mammals.

Other than in mammals, the roles of orexins in the regulation of emotional behavior, autonomic function, and neuroendocrine functions have not been clear, but psychomotor activity in goldfish was affected by an ICV injection of orexin A, suggesting an anxiogenic function of orexins, and this effect was abolished by injection of an OX1R antagonist (SB334867) (Nakamachi et al., 2014).

## CONCLUSION

Orexins play a highly important role in the regulation of sleep/wakefulness states in mammals. They are thought to be especially important for consolidation of wakefulness. Orexin deficiency results in narcolepsy, which is characterized by the inability to maintain long consolidated wakefulness, which is necessary to support any purposeful behaviors. Phylogenetically, orexins first appeared in vertebrates. They seem to be involved in the maintenance of wakefulness to pursue active motivated behavior in both mammals and other lower species. Even in mice, orexin neurons are relatively quiescent during quiet wakefulness, while they are active during active wakefulness, which accompanies purposeful behavior. This suggests that orexins are closely related to functions that support active behavior and consistently play a role as behavioral modulators among a wide range of species.

## AUTHOR CONTRIBUTIONS

SS and TS contributed to writing and making the figures in this article and approved its submission for publication. Both authors contributed to the article and approved the submitted version.

## FUNDING

This study was supported by a JSPS KAKENHI Grant-in-Aid for Scientific Research (B) (JP 15H03122 and 18H02595) (TS), a KAKENHI Grant-in-Aid for Exploratory Research (JP 15K12768) (TS), and a KAKENHI Grant-in-Aid for Scientific Research on Innovative Areas, "Willdynamics" (16H06401) (TS).

## REFERENCES

Abbas, M. G., Shoji, H., Soya, S., Hondo, M., Miyakawa, T., and Sakurai, T. (2015). Comprehensive behavioral analysis of Ox1r-/- mice showed implication of orexin receptor-1 in mood, anxiety and social behavior. *Front. Behav. Neurosci.* 9:324. doi: 10.3389/fnbeh.2015.00324

Amiya, N., Amano, M., Oka, Y., Iigo, M., Takahashi, A., and Yamamori, K. (2007). Immunohistochemical localization of orexin/hypocretin-like immunoreactive peptides and melanin-concentrating hormone in the brain and pituitary of medaka. *Neurosci. Lett.* 427, 16–21. doi: 10.1016/j.neulet.2007.07.043

Appelbaum, L., Wang, G. X., Maro, G. S., Mori, R., Tovin, A., Marin, W., et al. (2009). Sleep – wake regulation and hypocretin – melatonin interaction in zebrafish. *Proc. Natl. Acad. Sci. U.S.A.* 10, 1–6. doi: 10.1073/pnas.906637106

Bentzley, B. S., and Aston-jones, G. (2015). Orexin-1 receptor signaling increases motivation for cocaine-associated cues. *Eur. J. Neurosci.* 41, 1149–1156. doi: 10.1111/ejn.12866

Burdakov, D., Gerasimenko, O., and Verkhratsky, A. (2005). Physiological changes in glucose differentially modulate the excitability of hypothalamic melanin-concentrating hormone and orexin neurons in situ. *J. Neurosci.* 25, 2429–2433. doi: 10.1523/JNEUROSCI.4925-04.2005

Chang, G. Q., Karatayev, O., Davydova, Z., and Leibowitz, S. F. (2004). Circulating triglycerides impact on orexigenic peptides and neuronal activity in hypothalamus. *Endocrinology* 145, 3904–3912. doi: 10.1210/en.2003-1582

Chemelli, R. M., Willie, J. T., Sinton, C. M., Elmquist, J. K., Scammell, T., Lee, C., et al. (1999). Narcolepsy in orexin knockout mice: molecular genetics of sleep regulation. *Cell* 98, 437–451. doi: 10.1016/S0092-8674(00)81973-X

da Silva, E. S., dos Santos, T. V., Hoeller, A. A., dos Santos, T. S., Pereira, G. V., Meneghelli, C., et al. (2008). Behavioral and metabolic effects of central injections of orexins/hypocretins in pigeons (*Columba livia*). *Regul. Pept.* 147, 9–18. doi: 10.1016/j.regpep.2007.12.003

Date, Y., Ueta, Y., Yamashita, H., Yamaguchi, H., Matsukura, S., Kangawa, K., et al. (1999). Orexins, orexigenic hypothalamic peptides, interact with autonomic, neuroendocrine and neuroregulatory systems. *Proc. Natl. Acad. Sci. U.S.A.* 96, 748–753. doi: 10.1073/pnas.96.2.748

De Lecea, L., Kilduff, T. S., Peyron, C., Gao, X. B., Foye, P. E., Danielson, P. E., et al. (1998). The hypocretins: hypothalamus-specific peptides with neuroexcitatory activity. *Proc. Natl. Acad. Sci. U.S.A.* 95, 322–327. doi: 10.1073/pnas.95.1.322

Domínguez, L., Morona, R., Joven, A., González, A., and López, J. M. (2010). Immunohistochemical localization of orexins (hypocretins) in the brain of reptiles and its relation to monoaminergic systems. *J. Chem. Neuroanat.* 39, 20–34. doi: 10.1016/j.jchemneu.2009.07.007

Duboué, E. R., Keene, A. C., and Borowsky, R. L. (2011). Evolutionary convergence on sleep loss in cavefish populations. *Curr. Biol.* 21, 671–676. doi: 10.1016/j.cub.2011.03.020

Dyer, C. J., Touchette, K. J., Carroll, J. A., Allee, G. L., and Matteri, R. L. (1999). Cloning of porcine prepro-orexin cDNA and effects of an intramuscular injection of synthetic porcine orexin-B on feed intake in young pigs. *Domest. Anim. Endocrinol.* 16, 145–148. doi: 10.1016/S0739-7240(99)00011-19

Elbaz, I., Yelin-Bekerman, L., Nicenboim, J., Vatine, G., and Appelbaum, L. (2012). Genetic ablation of hypocretin neurons alters behavioral state transitions in zebrafish. *J. Neurosci.* 32, 12961–12972. doi: 10.1523/JNEUROSCI.1284-12.2012

Elbers, J. P., Rogers, M. F., Perelman, P. L., Proskuryakova, A. A., Serdyukova, N. A., Johnson, W. E., et al. (2019). Improving Illumina assemblies with Hi - C and long reads: an example with the North African dromedary. *Mol. Ecol. Resour.* 19, 1015–1026. doi: 10.1111/1755-0998.13020

Elias, C. F., Saper, C. B., Maratos-Flier, E., Tritos, N. A., Lee, C., Kelly, J., et al. (1998). Chemically defined projections linking the mediobasal hypothalamus and the lateral hypothalamic area. *J. Comp. Neurol.* 402, 442–459. doi: 10.1002/(sici)1096-9861(19981228)402:4<442::aid-cne2>3.0.co;2-r

Farrell, W. J., Delville, Y., and Wilczynski, W. (2003). Immunocytochemical localization of orexin in the brain of the green anole lizard (Anolis carolinensis). *Soc. Neur. Abstr.* 33:4.

Funato, H., Tsai, A. L., Willie, J. T., Kisanuki, Y., Williams, S. C., Sakurai, T., et al. (2009). Enhanced orexin receptor-2 signaling prevents diet-induced obesity

and improves leptin sensitivity. *Cell Metab.* 9, 64–76. doi: 10.1016/j.cmet.2008.10.010

Furuse, M., Ando, R., Bungo, T., Ao, R., Shimojo, M., and Masuda, Y. (1999). Intracerebroventricular injection of orexins does not stimulate food intake in neonatal chicks. *Br. Poult. Sci.* 40, 698–700. doi: 10.1080/00071669987115

Galas, L., Vaudry, H., Braun, B., Van Den Pol, A. N., De Lecea, L., Sutcliffe, J. G., et al. (2001). Immunohistochemical localization and biochemical characterization of hypocretin/orexin-related peptides in the central nervous system of the frog Rana ridibunda. *J. Comp. Neurol.* 429, 242–252. doi: 10.1002/1096-9861(20000108)429:2<242::aid-cne5<3.0.co;2-z

González, J. A., Iordanidou, P., Strom, M., Adamantidis, A., and Burdakov, D. (2016). Awake dynamics and brain-wide direct inputs of hypothalamic MCH and orexin networks. *Nat. Commun.* 7:11395. doi: 10.1038/ncomms11395

Haynes, A. C., Jackson, B., Chapman, H., Tadayyon, M., Johns, A., Porter, R. A., et al. (2000). A selective orexin-1 receptor antagonist reduces food consumption in male and female rats. *Regul. Pept.* 96, 45–51. doi: 10.1016/s0167-0115(00)00199-3

Hendricks, J. C., Sehgal, A., and Pack, A. I. (2000). The need for a simple animal model to understand sleep. *Prog. Neurobiol.* 61, 339–351. doi: 10.1016/S0301-0082(99)00048-49

Huesa, G., van den Pol, A. N., and Finger, T. E. (2005). Differential distribution of hypocretin (orexin) and melanin-concentrating hormone in the goldfish brain. *J. Comp. Neurol.* 488, 476–491. doi: 10.1002/cne.20610

Jaggard, J. B., Stahl, B. A., Lloyd, E., Prober, D. A., Duboue, E. R., and Keene, A. C. (2018). Hypocretin underlies the evolution of sleep loss in the Mexican cavefish. *eLife* 7, 1–22. doi: 10.7554/eLife.32637

Johnson, P. L., Truitt, W., Fitz, S. D., Minick, P. E., Dietrich, A., Sanghani, S., et al. (2010). A key role for orexin in panic anxiety. *Nat Med* 16, 111–115. doi: 10.1038/nm.2075

Kakizaki, M., Kakizaki, M., Tsuneoka, Y., Takase, K., Kim, S. J., Choi, J., et al. (2019). Differential roles of each orexin receptor signaling in obesity differential roles of each orexin receptor signaling in obesity. *Iscience* 20, 1–13. doi: 10.1016/j.isci.2019.09.003

Karnani, M. M., Apergis-Schoute, J., Adamantidis, A., Jensen, L. T., de Lecea, L., Fugger, L., et al. (2011). Activation of central orexin/hypocretin neurons by dietary amino acids. *Neuron* 72, 616–629. doi: 10.1016/j.neuron.2011.08.027

Kaslin, J., and Panula, P. (2001). Comparative anatomy of the histaminergic and other aminergic systems in zebrafish (*Danio rerio*). *J. Comp. Neurol.* 440, 342–377. doi: 10.1002/cne.1390

Katayama, S., Hamasu, K., Shigemi, K., Cline, M. A., and Furuse, M. (2010). Intracerebroventricular injection of orexin-A, but not orexin-B, induces arousal of layer-type neonatal chicks. *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* 157, 132–135. doi: 10.1016/j.cbpa.2010.05.018

Kojima, K., Kamijo, M., Kageyama, H., Uchiyama, M., Shioda, S., and Matsuda, K. (2009). Neuronal relationship between orexin-A- and neuropeptide Y-induced orexigenic actions in goldfish. *Neuropeptides* 43, 63–71. doi: 10.1016/j.npep.2009.01.004

Lin, L., Faraco, J., Li, R., Kadotani, H., Rogers, W., Lin, X., et al. (1999). The Sleep Disorder Canine Narcolepsy Is Caused by a Mutation in the Hypocretin (Orexin) Receptor 2 Gene. *Cell* 98, 365–376. doi: 10.1016/S0092-8674(00)81965-81960

López, J. M., Domínguez, L., Moreno, N., and González, A. (2009). Comparative immunohistochemical analysis of the distribution of orexins (hypocretins) in the brain of amphibians. *Peptides* 30, 873–887. doi: 10.1016/j.peptides.2009.01.013

Mieda, M., Hasegawa, E., Kisanuki, Y. Y., Sinton, C. M., Yanagisawa, M., and Sakurai, T. (2011). Differential roles of orexin receptor-1 and -2 in the regulation of non-REM and REM sleep. *J. Neurosci.* 31, 6518–6526. doi: 10.1523/jneurosci.6506-10.2011

Miranda, B., Esposito, V., De Girolamo, P., Sharp, P. J., Wilson, P. W., and Dunn, I. C. (2013). Orexin in the chicken hypothalamus: immunocytochemical localisation and comparison of mRNA concentrations during the day and night, and after chronic food restriction. *Brain Res.* 1513, 34–40. doi: 10.1016/j.brainres.2013.03.036

Miura, T., Maruyama, K., Shimakura, S. I., Kaiya, H., Uchiyama, M., Kangawa, K., et al. (2007). Regulation of food intake in the goldfish by interaction between ghrelin and orexin. *Peptides* 28, 1207–1213. doi: 10.1016/j.peptides.2007.03.023

Nakamachi, T., Matsuda, K., Maruyama, K., Miura, T., Uchiyama, M., Funahashi, H., et al. (2006). Regulation by Orexin of Feeding Behaviour and Locomotor Activity in the Goldfish. *J. Neuroendocrinol.* 18, 290–297. doi: 10.1111/j.1365-2826.2006.01415.x

Nakamachi, T., Shibata, H., Sakashita, A., Iinuma, N., Wada, K., Konno, N., et al. (2014). Orexin A enhances locomotor activity and induces anxiogenic-like action in the goldfish, Carassius auratus. *Horm. Behav.* 66, 317–323. doi: 10.1016/j.yhbeh.2014.06.004

Nambu, T., Sakurai, T., Mizukami, K., Hosoya, Y., Yanagisawa, M., and Goto, K. (1999). Distribution of orexin neurons in the adult rat brain. *Brain Res.* 827, 243–260. doi: 10.1016/s0006-8993(99)01336-1330

Novak, C. M., Jiang, X., Wang, C., Teske, J. A., Kotz, C. M., and Levine, J. A. (2005). Caloric restriction and physical activity in zebrafish (*Danio rerio*). *Neurosci. Lett.* 383, 99–104. doi: 10.1016/j.neulet.2005.03.048

Ohkubo, T., Boswell, T., and Lumineau, S. (2002). Molecular cloning of chicken prepro-orexin cDNA and preferential expression in the chicken hypothalamus. *Biochim. Biophys. Acta Gene Struct. Expr.* 1577, 476–480. doi: 10.1016/S0167-4781(02)00483-489

Peyron, C., Faraco, J., Rogers, W., Ripley, B., Overeem, S., Charnay, Y., et al. (2000). A mutation in a case of early onset narcolepsy and a generalized absence of hypocretin peptides in human narcoleptic brains. *Nat. Med.* 6, 991–997. doi: 10.1038/79690

Peyron, C., Tighe, D. K., van den Pol, A. N., de Lecea, L., Heller, H. C., Sutcliffe, J. G., et al. (1998). Neurons containing hypocretin (orexin) project to multiple neuronal systems. *J. Neurosci.* 18, 9996–10015. doi: 10.1523/JNEUROSCI.18-23-09996

Prober, D. A., Rihel, J., Onah, A. A., Sung, R. J., and Schier, A. F. (2006). Hypocretin/orexin overexpression induces an insomnia-like phenotype in zebrafish. *J. Neurosci.* 26, 13400–13410. doi: 10.1523/JNEUROSCI.4332-06.2006

Raizen, D. M., Zimmerman, J. E., Maycock, M. H., Ta, U. D., You, Y. J., Sundaram, M. V., et al. (2008). Lethargus is a Caenorhabditis elegans sleep-like state. *Nature* 451, 569–572. doi: 10.1038/nature06535

Saito, Y., Maejima, T., Nishitani, M., Hasegawa, E., Yanagawa, Y., Mieda, M., et al. (2018). Monoamines inhibit GABAergic neurons in ventrolateral preoptic area that make direct synaptic connections to hypothalamic arousal neurons. *J. Neurosci.* 38, 6366–6378. doi: 10.1523/JNEUROSCI.2835-17.2018

Sakamoto, F., Yamada, S., and Ueta, Y. (2004). Centrally administered orexin-A activates corticotropin-releasing factor-containing neurons in the hypothalamic paraventricular nucleus and central amygdaloid nucleus of rats1: possible involvement of central orexins on stress-activated central CRF neurons. *Regul. Pept.* 118, 183–191. doi: 10.1016/j.regpep.2003.12.014

Sakurai, T. (2007). The neural circuit of orexin (hypocretin): maintaining sleep and wakefulness. *Nat. Rev. Neurosci.* 8, 171–181. doi: 10.1038/nrn2092

Sakurai, T. (2014). The role of orexin in motivated behaviours. *Nat. Rev. Neurosci.* 15, 719–731. doi: 10.1038/nrn3837

Sakurai, T., Amemiya, A., Ishii, M., Matsuzaki, I., Chemelli, R. M., Tanaka, H., et al. (1998). Orexins and orexin receptors: a family of hypothalamic neuropeptides and G protein-coupled receptors that regulate feeding behavior. *Cell* 92, 573–585. doi: 10.1016/s0092-8674(00)80949-80946

Sharf, R., Sarhan, M., Brayton, C. E., Guarnieri, D. J., Taylor, J. R., and Dileone, R. J. (2010). Orexin signaling via the Orexin 1 receptor mediates operant responding for food reinforcement. *Biol. Psychiatry* 67, 753–760. doi: 10.1016/j.biopsych.2009.12.035

Shibahara, M., Sakurai, T., Nambu, T., Takenouchi, T., Iwaasa, H., Egashira, S. I., et al. (1999). Structure, tissue distribution, and pharmacological characterization of Xenopus orexins. *Peptides* 20, 1169–1176. doi: 10.1016/S0196-9781(99)00120-125

Singh, C., Oikonomou, G., Prober, D. A., Adamantidis, A., Zhang, F., Aravanis, A., et al. (2015). Norepinephrine is required to promote wakefulness and for hypocretin-induced arousal in zebrafish. *eLife* 4:e07000. doi: 10.7554/eLife.07000

Singletary, K. G., Delville, Y., Farrell, W. J., and Wilczynski, W. (2005). Distribution of orexin/hypocretin immunoreactivity in the nervous system of the green Treefrog. *Hyla cinerea*. *Brain Res.* 1041, 231–236. doi: 10.1016/j.brainres.2005.01.095

Soya, S., and Sakurai, T. (2018). Orexin as a modulator of fear-related behavior: hypothalamic control of noradrenaline circuit. *Brain Res.* 1731:146037. doi: 10.1016/j.brainres.2018.11.032

Soya, S., Shoji, H., Hasegawa, E., Hondo, M., Miyakawa, T., Yanagisawa, M., et al. (2013). Orexin receptor-1 in the locus coeruleus plays an important role in cue-dependent fear memory consolidation. *J. Neurosci.* 33, 14549–14557. doi: 10.1523/jneurosci.1130-13.2013

Soya, S., Takahashi, T. M., Mchugh, T. J., Maejima, T., Herlitze, S., Abe, M., et al. (2017). Orexin modulates behavioral fear expression through the locus coeruleus. *Nat. Commun.* 8, 1–14. doi: 10.1038/s41467-017-01782-z

Stecher, G., Tamura, K., and Kumar, S. (2020). Molecular Evolutionary Genetics Analysis (MEGA) for macOS. *Mol. Biol. Evol.* 37, 1237–1239. doi: 10.1093/molbev/msz312

Suzuki, H., Kubo, Y., and Yamamoto, T. (2008). Orexin-A immunoreactive cells and fibers in the central nervous system of the axolotl brain and their association with tyrosine hydroxylase and serotonin immunoreactive somata. *J. Chem. Neuroanat.* 35, 295–305. doi: 10.1016/j.jchemneu.2008.02.002

Thannickal, T. C., Moore, R. Y., Nienhuis, R., Ramanathan, L., Gulyani, S., Aldrich, M., et al. (2000). Reduced number of hypocretin neurons in human narcolepsy. *Neuron* 27, 469–474. doi: 10.1016/s0896-6273(00)00058-1

Tononi, G. (2000). Correlates of sleep and waking in Drosophila melanogaster. *Science* 287, 1834–1837. doi: 10.1126/science.287.5459.1834

Volkoff, H., Bjorklund, J. M., and Peter, R. E. (1999). Stimulation of feeding behavior and food consumption in the goldfish, *Carassius auratus*, by orexin-A and orexin-B. *Brain Res.* 846, 204–209. doi: 10.1016/s0006-8993(99)02052-2051

Vonk, F. J., Casewell, N. R., Henkel, C. V., Heimberg, A. M., Jansen, H. J., Woods, A. E., et al. (2013). The king cobra genome reveals dynamic gene evolution and adaptation in the snake venom system. *Proc. Natl. Acad. Sci. U.S.A.* 110, 20651–20656. doi: 10.1073/pnas.1314702110

Wilson, S., Baker, J., Jessop, D. S., and Harbuz, M. S. (2001). Central Orexin-A Activates Hypothalamic-pituitary-adrenal axis and stimulates hypothalamic corticotropin releasing factor and arginine vasopressin neurones in conscious rats. *J. Neuroendocrinol.* 13, 421–424. doi: 10.1046/j.1365-2826.2001.00655.x

Winsky-Sommerer, R., Yamanaka, A., Diano, S., Borok, E., Roberts, A. J., Sakurai, T., et al. (2004). Interaction between the corticotropin-releasing factor system and hypocretins (orexins): a novel circuit mediating stress response. *J. Neurosci.* 24, 11439–11448. doi: 10.1523/JNEUROSCI.3459-04.2004

Xu, M., and Volkoff, H. (2007). Molecular characterization of prepro-orexin in Atlantic cod (*Gadus morhua*): cloning, localization, developmental profile and role in food intake regulation. *Mol. Cell. Endocrinol.* 271, 28–37. doi: 10.1016/j.mce.2007.03.003

Yamada, H., Okumura, T., Motomura, W., Kobayashi, Y., and Kohgo, Y. (2000). Inhibition of food intake by central injection of anti-orexin antibody in fasted rats. *Biochem. Biophys. Res. Commun.* 531, 527–531. doi: 10.1006/bbrc.1999.1998

Yamanaka, A., Beuckmann, C. T., Willie, J. T., Hara, J., Tsujino, N., Mieda, M., et al. (2003). Hypothalamic Orexin neurons regulate arousal according to energy balance in mice. *Neuron* 38, 701–713. doi: 10.1016/S0896-6273(03)00331-333

Zhang, W., Zhang, N., Sakurai, T., and Kuwaki, T. (2009). Orexin neurons in the hypothalamus mediate cardiorespiratory responses induced by disinhibition of the amygdala and bed nucleus of the stria terminalis. *Brain Res.* 1262, 25–37. doi: 10.1016/j.brainres.2009.01.022

Zhdanova, I. V., Wang, S. Y., Leclair, O. U., and Danilova, N. P. (2001). Melatonin promotes sleep-like state in zebrafish. *Brain Res.* 903, 263–268. doi: 10.1016/S0006-8993(01)02444-2441

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 30 July 2020
doi: 10.3389/fnins.2020.00672

# Corticotropin-Releasing Hormone (CRH) Gene Family Duplications in Lampreys Correlate With Two Early Vertebrate Genome Doublings

*João C. R. Cardoso[1*†], Christina A. Bergqvist[2†] and Dan Larhammar[2]*

[1] Comparative Endocrinology and Integrative Biology, Centre of Marine Sciences, Universidade do Algarve, Faro, Portugal,
[2] Department of Neuroscience, Science for Life Laboratory, Uppsala University, Uppsala, Sweden

The ancestor of gnathostomes (jawed vertebrates) is generally considered to have undergone two rounds of whole genome duplication (WGD). The timing of these WGD events relative to the divergence of the closest relatives of the gnathostomes, the cyclostomes, has remained contentious. Lampreys and hagfishes are extant cyclostomes whose gene families can shed light on the relationship between the WGDs and the cyclostome-gnathostome divergence. Previously, we have characterized in detail the evolution of the gnathostome corticotropin-releasing hormone (CRH) family and found that its five members arose from two ancestral genes that existed before the WGDs. The two WGDs resulted, after secondary losses, in one triplet consisting of CRH1, CRH2, and UCN1, and one pair consisting of UCN2 and UCN3. All five genes exist in representatives for cartilaginous fishes, ray-finned fishes, and lobe-finned fishes. Differential losses have occurred in some lineages. We present here analyses of CRH-family members in lamprey and hagfish by comparing sequences and gene synteny with gnathostomes. We found five CRH-family genes in each of two lamprey species (*Petromyzon marinus* and *Lethenteron camtschaticum*) and two genes in a hagfish (*Eptatretus burgeri*). Synteny analyses show that all five lamprey CRH-family genes have similar chromosomal neighbors as the gnathostome genes. The most parsimonious explanation is that the lamprey CRH-family genes are orthologs of the five gnathostome genes and thus arose in the same chromosome duplications. This suggests that lampreys and gnathostomes share the same two WGD events and that these took place before the lamprey-gnathostome divergence.

**Keywords: gene duplication, tetraploidization, lamprey, paralogon, CRH**

OPEN ACCESS

***Edited by:***
*Gustavo M. Somoza, CONICET Instituto Tecnológico de Chascomús (INTECH), Argentina*

***Reviewed by:***
*Hervé Tostivint, Muséum National d'Histoire Naturelle, France*
*Bruno Querat, Université Paris Diderot, France*

***Correspondence:**
*João C. R. Cardoso*
*jccardo@ualg.pt*

*† These authors have contributed equally to this work*

***Specialty section:***
*This article was submitted to Neuroendocrine Science, a section of the journal Frontiers in Neuroscience*

***Received:*** *03 April 2020*
***Accepted:*** *02 June 2020*
***Published:*** *30 July 2020*

***Citation:***
*Cardoso JCR, Bergqvist CA and Larhammar D (2020) Corticotropin-Releasing Hormone (CRH) Gene Family Duplications in Lampreys Correlate With Two Early Vertebrate Genome Doublings. Front. Neurosci. 14:672. doi: 10.3389/fnins.2020.00672*

## INTRODUCTION

The corticotropin-releasing hormone (CRH) family consists in vertebrates of five structurally related neuropeptides that are involved in the regulation of physiological response to stress, emotional behavior, and anxiety (Vale et al., 1981; Dunn and Berridge, 1990; Koob and Heinrichs, 1999; Lovejoy and Balment, 1999; Gysling et al., 2004; Fox and Lowry, 2013). Two are named CRH (CRH1 and 2) and three are named urocortin (UCN1, 2, and 3). They have evolved through distinct

pressures during the vertebrate radiation, as reflected in their differences in evolutionary rates of amino acid change (Hwang et al., 2013; Grone and Maruska, 2015b; Cardoso et al., 2016; Endsin et al., 2017). CRH (now named CRH1) was the first family member to be discovered. It was isolated from sheep hypothalamus and consists of 41 amino acids in mammals (Vale et al., 1981). Homologs of mammalian CRH1 were subsequently found in numerous other tetrapods. Duplicate CRH1 genes now named crh1a and crh1b have been described in teleosts (Hwang et al., 2013; Lovejoy and de Lannoy, 2013; Grone and Maruska, 2015a,b; Cardoso et al., 2016) and were found to have arisen as a result of the teleost-specific genome duplication (Jaillon et al., 2004). The UCN1 peptide was the second family member to be discovered in mammals and was found to be the ortholog of the previously reported bony fish urotensin and of the amphibian sauvagine. Two additional urocortins were discovered *in silico* in mammals and named UCN2 (Reyes et al., 2001) and UCN3 (Lewis et al., 2001), both of which are 38 amino acids long in mammals. They were soon found in other classes of vertebrates, including ray-finned fishes. CRH2 is the most recently discovered member and was initially identified in cartilaginous fish and was suggested to be specific to these species (Nock et al., 2011), but subsequent reports demonstrated its presence in other vertebrate classes with the exception of placental mammals and teleosts (Grone and Maruska, 2015a; Cardoso et al., 2016).

The CRH family is one of the oldest metazoan peptide families, with homologs described in several invertebrate genomes. The closest relatives of vertebrates, the invertebrate deuterostomes such as the tunicates (*Ciona intestinalis* and *Ciona savignyi*), the cephalochordates (amphioxus *Branchiostoma floridae*), and ambulacrarians (the echinoderm *Strongylocentrotus purpuratus* and the hemichordate *Saccoglossus kowalevskii*), all have a single CRH-like gene (Kawada et al., 2010; Mirabeau and Joly, 2013). Protostomes, such as arthropods, have a related peptide named diuretic hormone 44 (DH44) (Audsley et al., 1995; Cabrero et al., 2002; Lovejoy and de Lannoy, 2013).

Diverging scenarios have been proposed to explain the origin and evolution of the CRH family in relation to the emergence of the vertebrates (Hwang et al., 2013; Cardoso et al., 2016; Endsin et al., 2017). Lovejoy and coworkers used sequence analyses to arrive at a scheme with five independent gene duplications followed by one loss (Endsin et al., 2017). However, their study did not consider adjacent genes to check for duplication of large chromosomal blocks. Already before their report, we had concluded that the five members of the gene family were established early in vertebrate evolution prior to the radiation of the gnathostomes, as based on phylogenetic sequence analyses and comparisons of gene synteny and duplicated chromosomes (Cardoso et al., 2016). The comparisons of neighboring genes showed that the two CRH subfamilies are located in different paralogons, i.e., in different sets of related chromosomal regions, with the CRH1/CRH2/UCN1 subfamily members located in a paralogon also harboring opioid peptide genes and the paralogon with the UCN2/UCN3 subfamily located in the paralogon that contains the visual opsin genes (Cardoso et al., 2016). Subsequently, the two pre-gnathostome whole genome duplications (WGD, see below) (Nakatani et al., 2007; Putnam et al., 2008) resulted in chromosome duplications that turned the first gene into three copies on separate chromosomes and the second gene into two copies on separate chromosomes. All five ancestral genes have been retained in slowly evolving lineages represented by the coelacanth (*Latimeria chalumnae*, a lobe-finned fish that diverged basal to the tetrapods), the spotted gar (*Lepisosteus oculatus*, a basal ray-finned fish that radiated prior to the teleost expansion), and the elephant shark (*Callorhinchus milii*, belonging to the holocephalans among cartilaginous fishes). Gene losses have occurred in some lineages (Cardoso et al., 2016).

The evolutionary origin of the gnathostomes is considered to have been preceded by two WGD events (Nakatani et al., 2007; Putnam et al., 2008), often referred to as 1R and 2R for the first and second round of genome doubling. However, the exact timing of these events in relation to the preceding divergence of vertebrates into the gnathostome and cyclostome lineages has been difficult to resolve. Investigation of their genomes can offer important insights into the origin and evolution of genes and gene families as well as the genomic events that have shaped vertebrate genomes. The cyclostomes, or living agnathans, consist of two major extant lineages, namely the lampreys and the hagfishes. To date, four sequenced agnathan genomes are available. Two genome assemblies are from the sea lamprey (*Petromyzon marinus)*, one of which is a somatic genome from adult liver and the other a recently assembled germline genome, which is essential because somatic lamprey cells delete much of the genome in adult tissues (Smith et al., 2013, 2018). One assembly is from the Arctic lamprey (*Lethenteron camtschaticum*, formerly known as *Lethenteron japonicum*) and was obtained from mature testis (Mehta et al., 2013). Finally, a fragmentary genome has been assembled for the inshore hagfish (*Eptatretus burgeri*)[1]. Nonetheless, analyses of agnathan gene families and genome segments have been inconclusive regarding the temporal relationship between the two WGD events and the cyclostome-gnathostome divergence, which is why different scenarios have been proposed. Analysis of the somatic sea lamprey genome suggested that the most recent WGD (2R) is likely to have occurred before the divergence of the ancestral lamprey and gnathostome lineages (Smith et al., 2013). Other investigators suggested that lampreys may have experienced distinct polyploidization events from the gnathostomes and also may have had an additional independent WGD (Mehta et al., 2013). More recently, analyses of the sea lamprey germline genome supported two possible scenarios: (1) a single shared WGD or (2) two WGD followed by extensive gene losses from the resulting daughter chromosomes, especially in the lamprey (Smith and Keinath, 2015; Smith et al., 2018). One other study concluded that cyclostomes and gnathostomes have gone through the same two WGD events before they diverged from each other (Sacerdot et al., 2018). Others have proposed that only the first WGD was shared and was followed by independent duplication and loss events in the two lineages, a WGD in gnathostomes and unclear types of duplication in lampreys (Simakov et al., 2020).

[1] www.ensembl.org

Homologs of the gnathostome CRH family members have been reported for lampreys (Roberts et al., 2014; Cardoso et al., 2016; Endsin et al., 2017). The identification of lamprey peptides representing both of the two CRH/UCN subfamilies confirmed that these arose before the divergence of the cyclostome and gnathostome lineages (Cardoso et al., 2016). However, each of the lamprey CRH/UCN-sequences did not cluster clearly with each of the five gnathostome CRH-family sequences, thus it was not possible to assign orthology based upon sequence analysis. Also, as no information on synteny was available at the time, it was not possible to use this criterion to ascertain orthology between the lamprey and gnathostome members (Cardoso et al., 2016). Thus, it could not be inferred that cyclostomes and gnathostomes share the same two WGD events.

In this study, we investigated the early vertebrate evolution of the CRH family members and the implications for understanding the timing of the WGD events in relation to the agnathan-gnathostome divergence. We used a double comparative approach combining sequence analyses of available lamprey CRH-family genes and peptides with investigation of gene synteny for 37 neighboring gene families and their sequence-based phylogenies (and two hagfish CRH-family genes). Our data show that lampreys and gnathostomes have the same number of CRH family members in both of the peptide subfamilies. Furthermore, the lamprey genes are located in gene neighborhoods that resemble those that we have previously reported for gnathostomes, although some rearrangements have taken place. The most parsimonious explanation for these similarities is that lampreys and gnathostomes share five CRH orthologs that arose by chromosome duplications of two ancestral peptide genes. This would suggest that lampreys share the same genome doubling events as gnathostomes, albeit clouded by chromosomal recombination and changes in gene order along the chromosomes.

# MATERIALS AND METHODS

## Identification of the Lampreys and Hagfish CRH-Family Genes

The mature predicted CRH-family members from our previous study (Cardoso et al., 2016), two from the sea lamprey (*Petromyzon marinus*) and four from the Arctic lamprey (*Lethenteron camtschaticum)*, were used to identify the missing genes and the scaffolds for all of the peptide genes in the sea lamprey and Arctic lamprey genome assemblies (available from NCBI database). The predicted mature peptides from lamprey were used to search for homologs in the inshore hagfish (*Eptatretus burgeri*) genome available from ENSEMBL. The identity of the CRH members that were retrieved was confirmed by submitting to the InterProScan tool[2] or by sequence homology.

[2]https://www.ebi.ac.uk/interpro/search/sequence-search

## Sequence Comparisons and Phylogeny

The complete deduced precursor sequences for both lamprey and hagfish CRH-family members were retrieved. Mature peptides were predicted by comparing with the gnathostome peptides and by localization in the sequence of putative proteolytic dibasic cleavage sites. Amino acid sequence identities were calculated using the Clustal Omega (Sievers et al., 2011), available from EMBL-EBI[3].

Phylogenetic trees of the lamprey and hagfish CRH-family members with the other vertebrate homologs were constructed using both the complete peptide precursor sequences and the mature peptides. Sequences were aligned using the MUSCLE algorithm in the AliView platform 1.18 (Larsson, 2014) and trees were built according to the maximum likelihood (ML) and Bayesian inference (BI) methods. The alignment of the complete peptide precursors was manually edited to remove sequence gaps and poorly aligned regions. ML trees were calculated using the PhyML 3.0 algorithm ATGC bioinformatics platform with the SMS automatic model selection (Lefort et al., 2017) according to the AIC (Akaike Information Criterion). ML trees were constructed according to the LG substitution model (Le and Gascuel, 2008) and reliability of internal branching was accessed using 100 bootstrap replicates. The BI trees were constructed in the CIPRES Science Gateway (Miller et al., 2010) with MrBayes (Ronquist et al., 2012) run on XSEDE using the LG substitution model (Aamodel = LG) and 1,000,000 generation sampling and probability values to support tree branching. The tunicate (*Ciona intestinalis* and *Ciona savignyi*) CRH-like orthologs were used (Mirabeau and Joly, 2013). ML and BI trees were displayed with FigTree 1.4.2 and edited in Inkscape[4].

## Gene Synteny Comparisons

The neighbors of the CRH family genes in lamprey and hagfish were identified and used to find orthologous genome regions in the spotted gar, chicken, and human. The gene environment of the sea lamprey scaffolds containing the CRH-family members (**Supplementary Table S1**) was annotated using a combination of the AUGUSTUS web interface (Stanke et al., 2004), by enquiring the species genome assembly at SIMRBASE database[5] and the somatic genome assembly available from ENSEMBL[6]. We have annotated in detail 3 Mb of the sea lamprey scaffolds (1.5 Mb in each direction from the lamprey CRH-family gene loci, **Supplementary Table S1**). AUGUSTUS predicted complete and partial genes on both strands using the Arctic lamprey (*Lethenteron camtschaticum*) and human (*Homo* sapiens) as reference species. The gene environment of the Arctic lamprey homologous genome regions were predicted using a local installation of AUGUSTUS 2.5.5 (Stanke et al., 2004, 2008) with the settings set for sea lamprey to predict genes *de novo*. Gene identity was confirmed using Swissprot through BLAST2GO (Conesa et al., 2005) comparing to human, chicken, and spotted gar non-redundant protein (nr) databases. Searches

[3]https://www.ebi.ac.uk

[4]https://inkscape.org

[5]https://genomes.stowers.org/

[6]http://www.ensembl.org/

for neighbors was complemented by procuring the species genome assemblies available from NCBI[7]. The neighboring genes of the hagfish CRH-like genome fragments were annotated using the BioMart tool available from ENSEMBL and compared with the spotted gar, chicken, and human, and common genes that were found were subsequently searched in the sea lamprey and Arctic lamprey genomes. The neighboring genes that we had previously identified (Cardoso et al., 2016) within the gnathostome CRH paralogons were also searched in lamprey and hagfish genomes.

To better comprehend the evolution of the lamprey and hagfish CRH members, phylogenetic analysis of their neighboring genes families was performed to investigate whether they had undergone similar evolutionary events. Orthologs of lamprey neighboring genes were retrieved from human, chicken, coelacanth, spotted gar, zebrafish, and elephant shark genomes available from ENSEMBL or NCBI. The invertebrate orthologs were retrieved from either two tunicates (*Ciona intestinalis* and/or *Ciona savignyi*), a cephalochordate (*Branchiostoma floridae*), or from the nematode (*Caenorhabditis elegans*) and fruit-fly (*Drosophila melanogaster*), and these were used to root the trees. Sequence alignments were performed using the AliView interface with MUSCLE, trees were carried out using the ML implemented in PhyML with automatic selection model, and sequence branching support was given by the Approximate Likelihood-Ratio Test (aLTR). The resulting trees were displayed in FigTree.

To deduce the putative ancestral pre-vertebrate CRH genomic region, we have used all the conserved cyclostome and gnathostome CRH-family neighboring genes to search for homologous genomic regions in invertebrate chordates where a CRH-like peptide gene has been described: two tunicates (*C. intestinalis* and *C. savignyi*) and two cephalochordates (*B. floridae* and *B. lanceolatum*).

# RESULTS

## The Agnathan CRH-Family Members

Blast searches with the known CRH-family members identified five CRH-family sequences in both the sea lamprey and the Arctic lamprey genomes. These correspond to the five members described in our previous report (Cardoso et al., 2016), although we could not identify the complete set in both species at that time. No additional CRH-like sequences were identified. Thus, lampreys have the same number of CRH-family genes as the gnathostome ancestor and some extant gnathostomes. Analysis of the sea lamprey germline genome assembly revealed that the five CRH-family genes map to five different genome regions: scaffold_00040 (GL480439 in ENSEMBL), scaffold_82 (GL476347 in ENSEMBL), scaffold_00003, scaffold_00017, and scaffold_00057. The three latter genome scaffolds are absent from the sea lamprey somatic genome assembly (available from ENSEMBL). Similarly, the five Arctic lamprey CRH-family genes map to five distinct genome regions (KE993827, KE994103, KE993984, KE993813, KE993959). Searches in the hagfish genome assembly identified three putative CRH members that map to separate scaffolds (FYBX02010500.1, FYBX02010617.1, and FYBX02009844.1). However, analysis of the deduced CRH peptides encoded in scaffolds FYBX02010500.1 and FYBX02010617.1 revealed that they are 100% identical, thus we only considered FYBX02010500.1 for analysis. The lamprey CRH genes were designated CRH/UCN- a, b, and c and UCN-a and b (**Table 1**) according to phylogenetic clustering based on our previous and current analyses.

**TABLE 1 |** Nomenclature adopted for the lamprey CRH-family genes.

| | *P. marinus* | *L. camtschaticum* |
|---|---|---|
| CRH/UCN_a | scaf_00017 | KE993984 |
| CRH/UCN_b | scaf_00040 | KE994103 |
| CRH/UCN_c | scaf_00003 | KE993827 |
| UCN_a | scaf_00057 | KE993813 |
| UCN_b | scaf_00082 | KE993959 |

The deduced amino acid sequences of the sea lamprey and Arctic lamprey genes were also aligned and compared with the three recently described sea lamprey cDNA sequences (Endsin et al., 2017) to confirm the genomic predictions and to correct for imprecisions in the automatic annotation. Orthologs between the two lamprey species had high sequence identity in their complete peptide precursors (>89% amino acid identity) or the corresponding deduced mature peptides (>98%), confirming the close evolutionary relationship of the two species (**Table 2**).

The mature peptide sequences deduced in the lampreys were compared with the gnathostome CRH family peptides. However, orthologies based on sequence identity alone were unclear, presumably due to lineage-specific evolutionary pressures such as lamprey GC-rich DNA sequences leading to amino acids with GC-rich codons. Lamprey CRH/UCN-a deduced mature peptide sequence shares 85% aa identity with the human and spotted gar CRH1 peptides (**Table 3**). The lamprey CRH/UCN-b and CRH/UCN-c peptides also displayed highest sequence identity to the gnathostome CRH1 peptide (66 and 63% for human CRH1, respectively). Lamprey UCN-a shares the highest sequence identity with human UCN3 (61% aa), and with fish it has slightly higher identity to spotted gar UCN2 (68% aa) than UCN3 (65% aa). Lamprey UCN-b has highest identity to gnathostome UCN2 (55% aa for human and spotted gar).

The deduced hagfish CRH/UCN-peptide encoded in scaffold FYBX02010500.1 has 59% identity to lamprey CRH/UCN-a and the hagfish UCN2/3 peptide in scaffold FYBX02009844.1 has highest identity (77% aa) to lamprey UCN-a. The absence of additional hagfish CRH-like genes is probably due to incomplete genome assembly.

## Phylogenetic Analysis

Phylogenetic analysis with the BI method of the agnathan and gnathostome deduced mature peptides (**Figure 1** and **Supplementary Figure S1**) or the complete peptide precursors (**Figure 2** and **Supplementary Figure S2**) produced tree topologies that are in agreement with our

[7]https://www.ncbi.nlm.nih.gov/assembly/

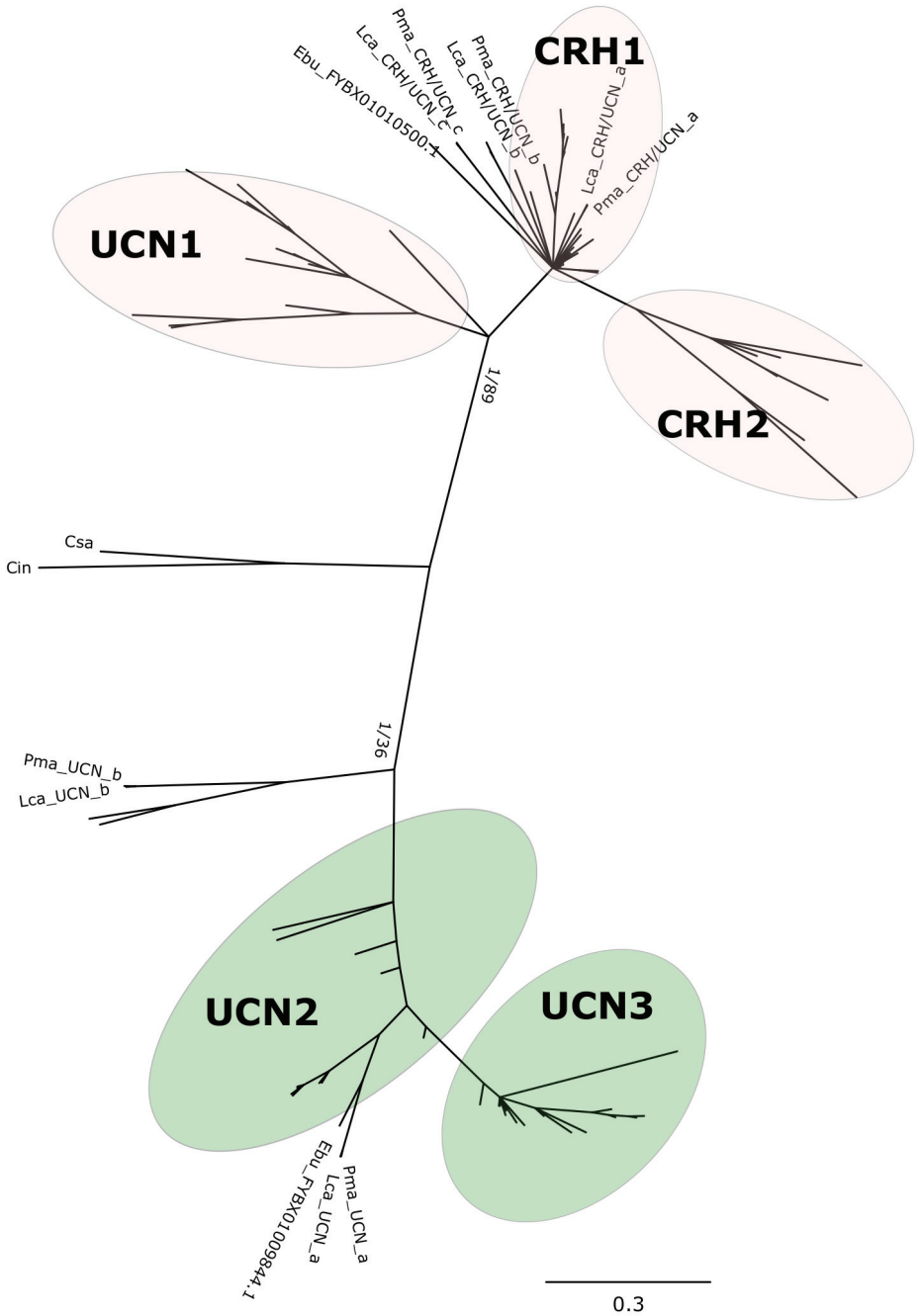


**FIGURE 1 |** Edited radial phylogenetic tree of the chordate deduced mature peptides CRH-family members. Tree was constructed with the Bayesian inference (BI) and the complete tree is available in **Supplementary Figure S1**. The maximum likelihood (ML) tree is available in **Supplementary Figure S3** and branch support values (BI posterior probability and ML bootstrap values) are shown only for the two major peptide subfamily clades. Accession numbers of the sequences used are available in **Supplementary Table S2**.

previous study (Cardoso et al., 2016). Phylogenetic trees with the ML method generated essentially identical topologies (**Supplementary Figures S3**, **S4**). All trees positioned the agnathan genes within either of the two CRH/UCN subfamilies (see **Figures 1**, **2**). Three of the lamprey CRH-members grouped in the gnathostome CRH1/CRH2/UCN1 subfamily and the two others within the gnathostome UCN2/UCN3 subfamily (**Figure 1** and **Supplementary Figures S1**, **S3**; **Figure 2** and

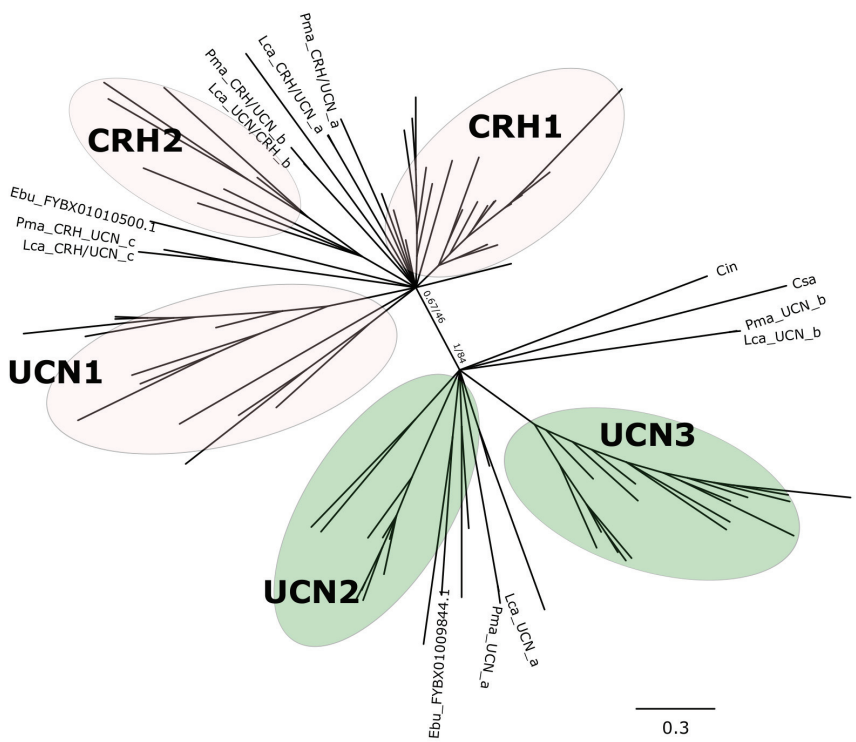


**FIGURE 2 |** Edited radial phylogenetic tree of the chordate deduced full-length CRH-family precursors. Tree was constructed with the Bayesian inference (BI) and the complete tree is available in **Supplementary Figure S2**. The maximum likelihood (ML) tree is available in **Supplementary Figure S4** and branch support values (BI posterior probability and ML bootstrap values) are shown only for the two major peptide subfamily clades. Accession numbers of the sequences used are available in **Supplementary Table S2**.

**TABLE 2 |** Percent amino acid sequence identity of the lamprey CRH-family members.

| | Lca CRH/UCN-a | Pma CRH/UCN-b | Lca CRH/UCN-b | Pma CRH/UCN-c | Lca CRH/UCN-c | Pma UCN-a | Lca UCN_a | Pma UCN_b | Lca UCN_b |
|---|---|---|---|---|---|---|---|---|---|
| Pma_CRH/UCN-a | **93 (100)** | 33 (61) | 31 (59) | 39 (58) | 36 (58) | 19 (27) | 18 (27) | 20 (24) | 21 (24) |
| Lca_CRH/UCN-a | 100 | 33 (61) | 32 (59) | 38 (58) | 37 (58) | 19 (27) | 20 (27) | 20 (24) | 21 (24) |
| Pma_CRH/UCN-b | | 100 | **92 (98)** | 30 (55) | 27 (55) | 22 (25) | 21 (25) | 17 (30) | 19 (30) |
| Lca_CRH/UCN-b | | | 100 | 29 (58) | 28 (58) | 19 (25) | 19 (25) | 14 (30) | 16 (30) |
| Pma_CRH/UCN-c | | | | 100 | **84 (100)** | 26 (28) | 27 (28) | 19 (38) | 20 (38) |
| Lca_CRH/UCN-c | | | | | 100 | 25 (28) | 27 (28) | 19 (38) | 19 (38) |
| Pma_UCN-a | | | | | | 100 | **98 (100)** | 33 (50) | 32 (50) |
| Lca_UCN-a | | | | | | | 100 | 33 (50) | 33 (50) |
| Pma_UCN-b | | | | | | | | 100 | **91 (98)** |

*Complete deduced precursors and deduced mature peptides (within brackets). The percent sequence identity between the lamprey orthologs is in bold.*

**TABLE 3 |** Percent amino acid identity of the deduced sea lamprey CRH-family peptides with human, spotted gar, and hagfish.

| Pma | Hsa CRH1 | Loc CRH1 | Loc CRH2 | Hsa UCN1 | Loc UCN1 | Hsa UCN2 | Loc UCN2 | Hsa UCN3 | Loc UCN3 | Ebu CRH/UCN | Ebu UCN2/3 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| CRH/UCN-a | **85** | **85** | 61 | 45 | 49 | 37 | 29 | 32 | 28 | **59** | 27 |
| CRH/UCN-b | **66** | **68** | 49 | 43 | 51 | 32 | 25 | 29 | 25 | 51 | 23 |
| CRH/UCN-c | **63** | **63** | 55 | 45 | 43 | 32 | 25 | 29 | 25 | 55 | 23 |
| UCN-a | 29 | 27 | 33 | 23 | 20 | 45 | **68** | **61** | 65 | 24 | **77** |
| UCN-b | 29 | 29 | 23 | 38 | 27 | **55** | **55** | 50 | 48 | 29 | 44 |

*The highest sequence identities are highlighted in bold.*

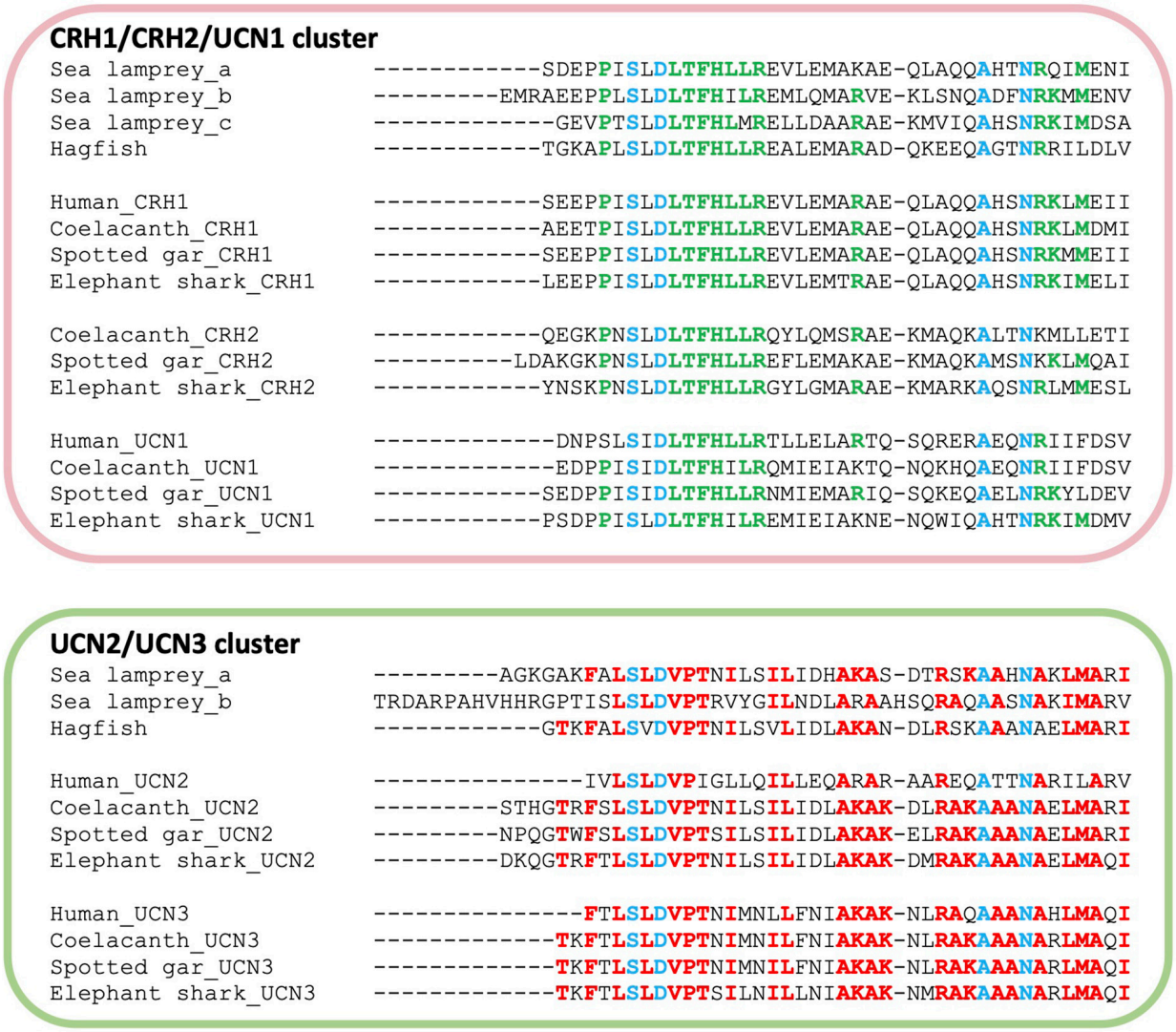


**FIGURE 3 |** Mature peptide sequence alignment of the lamprey, hagfish, and gnathostome CRH-family members. Peptides were predicted by comparing with the gnathostome peptides and by localization in the sequence of putative proteolytic dibasic cleavage sites. Amino acids conserved in all peptides are shown in blue. Positions shown in green or red are amino acids conserved within the CRH1/CRH2/UCN1 and UCN2/UCN3 subfamilies, respectively.

**Supplementary Figures S2**, **S4**). The two hagfish sequences were separated into each of the two subfamilies.

The five peptides from the two lampreys form five closely clustered pairs reflecting orthology between the species. One hagfish sequence clusters closest to the lamprey UCN_a sequences. The agnathan sequences do not cluster with clear bootstrap support with each of the five gnathostome peptide clades. The agnathan mature peptides and precursor sequences generally tended to radiate earlier and at the base of the discrete gnathostome peptide clades.

## Features of CRH-Family Peptides

The deduced mature peptide sequences of the cyclostome CRH-members were aligned and compared with the gnathostome orthologs according to clustering obtained from the phylogenetic tree (**Figure 3**). Amino acid residues and motifs that are common between the lamprey, hagfish, elephant shark, spotted gar, coelacanth, and human CRH-family members were found (**Figure 3** in blue), including those that are characteristics of the CRH1/CRH2/UCN1 (**Figure 3** in green) and UCN2/UCN3 (**Figure 3** in red) clades.

The S-L-D motif that is conserved at the N-terminus of the gnathostome CRH-members (except in the UCN clade) is also conserved in all agnathan peptides, including the asparagine (R) residue that is localized near the C-terminus and the alanine (A) residue located three positions before. Motifs characteristic of the CRH1/CRH2/UCN1 subfamily (**Figure 3**, in green) and UCN2/UCN3 subfamily (**Figure 3**, in red) are also maintained in the lampreys and hagfish. This includes the gnathostome LTFH(L/I)LR localized in the mid-region of CRH1, CRH2, and

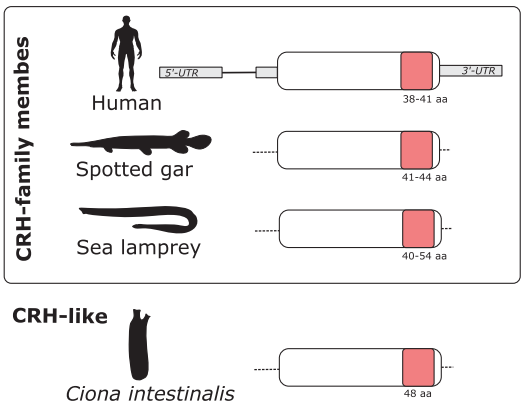


**FIGURE 4 |** Schematic representation of the gene structure of CRH-family members. In both gnathostomes and lamprey the CRH-family members are encoded by a single exon. In human the CRH-family gene structure is composed by 2 exons: exon 1 contains the 5′ UTR and exon 2 encoded the CRH-family member peptide precursors (represented in white) that contains the mature peptide (represented in pink) and the 3′UTR (represented in gray). In the spotted gar and lamprey, the complete gene structures of the CRH-family members have not yet been elucidated. The predicted sizes (aa) of the mature CRH-family peptides are indicated. Figure is not drawn to scale.

UCN1, and other amino acid residues (**Figure 3**, in green). Within the UCN2/UCN3 peptide alignment, the cyclostome peptides share the VPT motif as well as additional residues with the gnathostome UCN2 (**Figure 3**, in red).

## Gene Structure

The gnathostome CRH-family genes are composed of two exons, the second of which encodes the entire peptide precursor (Shibahara et al., 1983; Thompson et al., 1987). Characterization of the lamprey (**Figure 4**) and hagfish genes revealed that they too encode the entire prepro-peptide on a single exon. This is the situation also for the Ciona CRH-like gene (**Figure 4**).

## Neighboring Gene Family Analysis

The gnathostome CRH family genes map to two distinct paralogons: The CRH1/CRH2/UCN1 genes are located on separate chromosomes in the paralogon that also contains the genes for the opioid receptors (Dreborg et al., 2008) and the opioid peptides (Sundström et al., 2010), and the UCN2/UCN3 genes are found in a paralogon that contains the visual opsin genes (Lagman et al., 2013). Mapping of the lamprey CRH-family chromosomal regions shows that they possess similar gene repertoires to those in human, chicken, and spotted gar (**Figures 5**, **6**). This strongly suggests that agnathan and gnathostome CRH family genes most likely shared the same ancestral gene neighborhood, and that many of these genes have remained neighbors by contingency. Furthermore, the comparisons show that many of the neighboring genes belong to families consisting of quartets, triplets, or pairs, with members nearby the different CRH-family members. This is consistent with duplication of large chromosomal blocks or regions containing a large number of genes. Also, many neighboring genes that remain singletons (i.e., no duplicates have survived the chromosome duplications) support a common ancestry for these chromosomal regions in living agnathans and gnathostomes.

The conservation of synteny does not seem to extend to tunicates; despite the existence of a CRH-like peptide in Ciona, no clear homologous genomic region composed by similar genes to those of the vertebrate CRH family members was found. Likewise, no such conserved gene synteny region could be identified in the cephalochordate genome.

## The CRH1/CRH2/UCN1 (Opioid) Paralogon

The three sea lamprey CRH-family members were found in scaffolds 3, 17, and 40, respectively. A total of 21 neighboring gene families that are syntenic with the gnathostome CRH1/CRH2/UCN1 paralogon were identified and characterized (**Figure 5**). Six of these gene families also have members in a fourth sea lamprey chromosomal region contained in scaffold 1, and a few additional genes are contributed by scaffolds 2 and 11. Six of the neighboring gene families in spotted gar are quartets, namely DLGAP, EPB41, L3MBTL, SRC-B, STMN, and TRIM. Four of the neighboring families in spotted gar are triplets, i.e., SLC30A, MYB, NCOA, and XKR (**Figure 5**). In the sea lamprey, one family is a quartet (TRIM) and six families consist of triplets. This is slightly fewer than the spotted gar, so either lineage-specific gene losses occurred in the sea lamprey (or cyclostomes) lineage, or the sea lamprey genome assembly is not quite complete. Taken together, the sea lamprey CRH1/CRH2/UCN1 subfamily genes and neighboring gene families comprise a paralogon with fourfold symmetry consistent with four related chromosomal regions, i.e., two doubling events, as in gnathostomes. Two gene families in the sea lamprey, RSPO and VSNL1, even consist of five members (**Figure 7**), too many to be perfectly consistent with a chromosome quadruplication scenario, but the extra fifth gene could have arisen in an independent duplication event in the same time window as the genome doublings. A number of singletons that flank the CRH-family genes in lamprey have homologs that also occur as singletons in the gnathostome genomes.

The comparison with the Arctic lamprey (**Figure 7**) confirms much of the sea lamprey gene repertoire and organization, though it lacks some of the genes identified in sea lamprey. Only a few genes that are present in the Arctic lamprey are missing in the sea lamprey, suggesting that the latter is more complete. Orthologs in the two lamprey species are highly identical in sequence.

The hagfish genome has only a single member of the CRH1/CRH2/UCN1 subfamily (**Figure 7**). This scaffold shares three neighboring gene family members with the gnathostome species shown in **Figure 5**, namely SLC30A, TRIM, and MPV17, and shares 12 neighbors with the scaffolds in the two lamprey species. However, it is difficult to say exactly to which of the lamprey regions it is orthologous as the hagfish scaffold's gene repertoire appears to be a hybrid between sea lamprey scaffolds 2 and 17, suggesting that gene shuffling occurred after separation of the two agnathan lineages. Homologs for other lamprey CRH-like

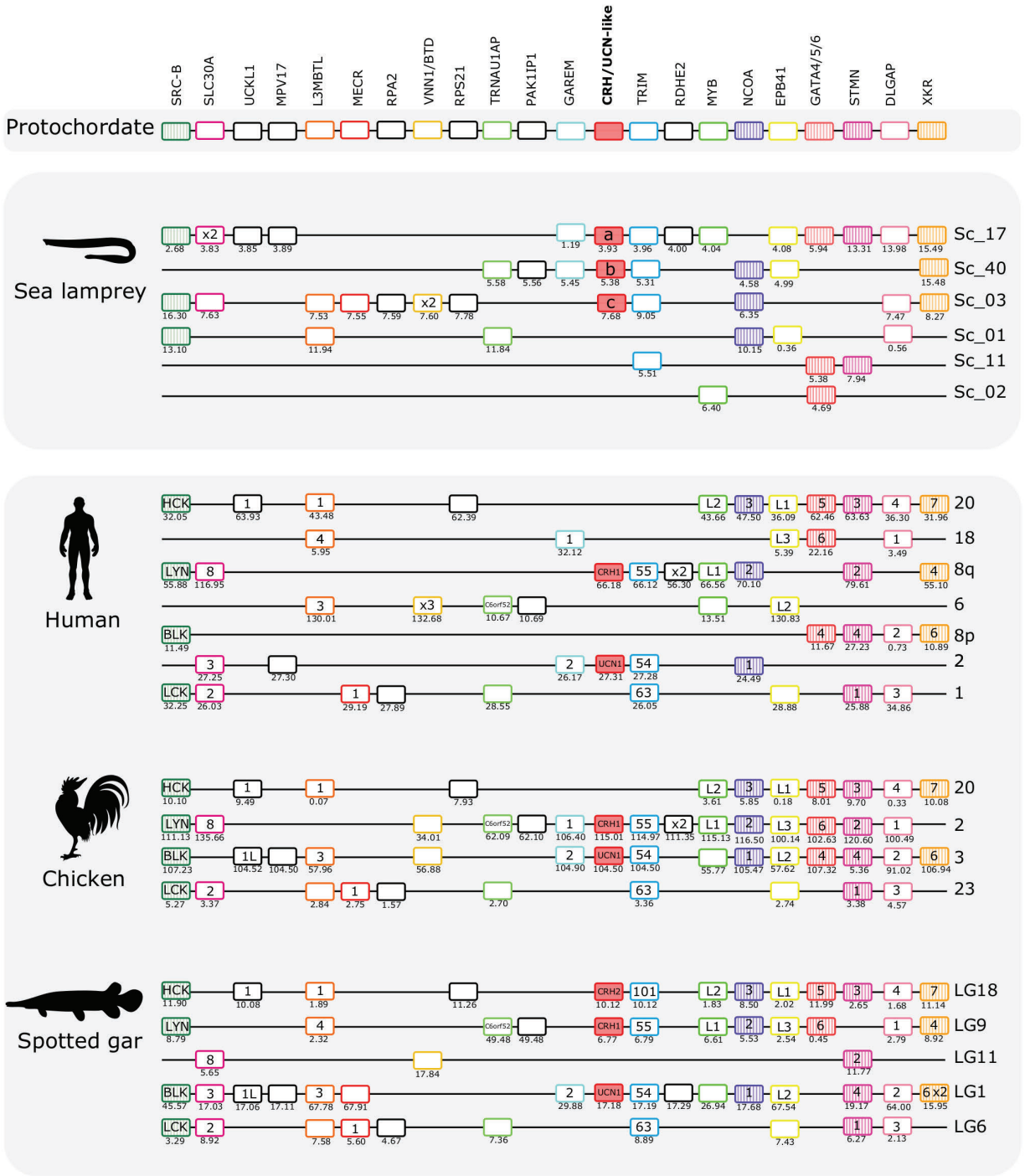


**FIGURE 5 |** Gene synteny analysis of the opioid (CRH1/CRH2/UCN1) paralogon in lamprey and gnathostomes. Chromosomal locations of the CRH-family members and of 21 neighboring genes families in lamprey, human, chicken, and spotted gar are shown. Chromosomes or scaffold numbers as well as gene positions (Mb, below each gene) are given. Genes are represented by boxes and CRH/UCN members are represented by full-red boxes and neighboring gene families are represented by different colors. The gene family symbol is shown, and the designation of the different members is provided inside the corresponding gene. Gene duplicates of the same family that map to the same genome region are represented (x2 and x3). The genes in the lamprey scaffolds are displayed according to their order in the genome region analyzed but in other species the genes were reshuffled to highlight the similarities between species. The neighboring gene families that we have previously described in the gnathostome CRH1/CRH2/UCN1 paralogon are also represented and are shown as striped colored boxes (Cardoso et al., 2016). The accession numbers and phylogenetic trees of the gene families represented are available in **Supplementary Table S3** and **Supplementary Figure S5**.

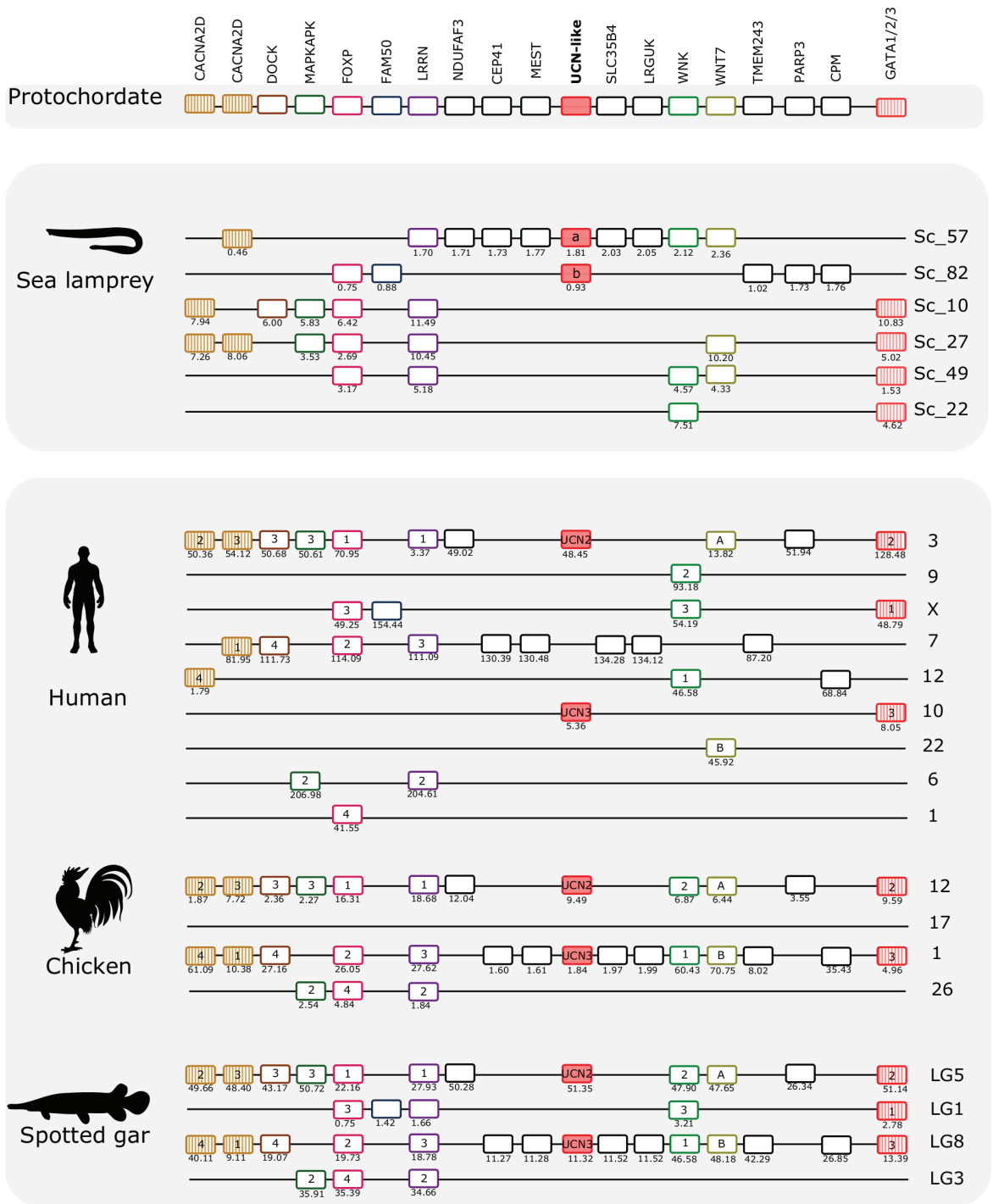


**FIGURE 6 |** Gene synteny analysis of the opsin (UCN2/UCN3) paralogon in lamprey and gnathostomes. Chromosomal locations of the CRH-family members and of 17 neighboring gene families in the sea lamprey, human, chicken, and spotted gar are shown. The chromosome or scaffold numbers as well as gene positions (Mb, below each gene) are given. Genes are represented by boxes and UCN members are represented by full-red boxes and neighboring gene families are represented by different colors. The gene family symbol is shown, and the designation of the different members is provided inside the corresponding gene. The genes in the lamprey scaffolds are represented according to their order in the genome fragment analyzed but in other species were reshuffled to highlight the similarities between species. The genes that we have previously identified in UCN2/UCN3 paralogon are represented by stripped colored boxes (Cardoso et al., 2016). The accession numbers and phylogenetic trees of the gene families represented are available in **Supplementary Table S4** and **Supplementary Figure S6**.

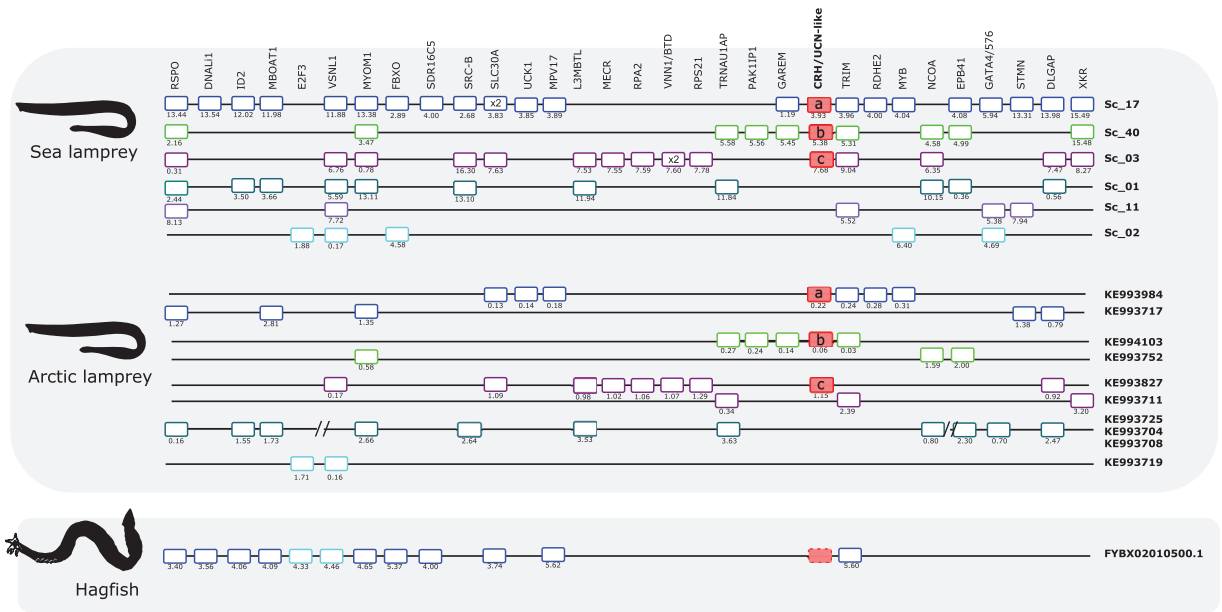


**FIGURE 7 |** Gene synteny of the opioid (CRH/UCN) paralogon in the cyclostomes. The sea lamprey, Arctic lamprey, and hagfish homolog genome regions were compared. To highlight for potential gene shuffling/ crossing-over events between the two lampreys and the two lampreys and hagfish, the sea lamprey genes that map in the same scaffold were represented by the same color. The exceptions are the CRH/UCN members which are represented by red-full boxes. Members of the same family in the sea lamprey genome or homologs in the other two cyclostome species are aligned. Gene chromosomal positions (Mb, below each gene) are shown. The sea lamprey genes are represented according to their position in the scaffolds and homologs in the Arctic lamprey and hagfish were positioned in relation to the sea lamprey genes. When possible different fragments of Arctic lamprey genome that possess a similar gene environment to a unique sea lamprey scaffold region were merged and this is marked by "//." The hagfish CRH-like gene is dashed because this gene is not predicted in the species genome assembly. Tandem duplicate genes are represented by x2. Homologs for other lamprey CRH/UCN-like neighboring genes were found in the hagfish genome assembly but are not shown as they are located on multiple small scaffolds.

neighboring genes were also found in hagfish genome, but they are located on multiple small scaffolds and do not provide synteny evidence.

Phylogenetic analyses were carried out for all of the 21 neighboring gene families and revealed a similar topology as for the CRH-family genes, i.e., the duplications seem to have taken place in the time range of early vertebrate evolution (**Supplementary Figure S5**). Occasionally, some family members display somewhat deviating species divergences compared to the established species phylogeny, especially sequences from the slowly evolving lineages represented by coelacanth, spotted gar, and elephant shark. For instance, the coelacanth sequence may cluster with the actinopterygian representatives rather than the sarcopterygian species, as in the LYN tree in **Supplementary Figure S5**. In the tree for the SLC30A family, the member SLC30A8 has coelacanth and elephant shark as closest relatives. Such slight variations in clustering are not unusual in evolutionary analyses of vertebrates due to variable evolutionary rates for the lineages. For some gene families, some members or species display a more dramatic difference in evolutionary rate, such as the human SLC30A3 gene. Nevertheless, each gene family member usually displays high statistical support for the clade comprised by the orthologs from the species included in this analysis.

The lamprey sequences vary more in their positions in the phylogenetic trees. Sometimes they branch off basally to the gnathostome members. Also, the different gene family members from lampreys may group together as if they had been duplicated in this lineage, although the synteny and paralogon analyses support simultaneous duplication along with the gene neighbors in large chromosomal blocks or entire chromosomes.

## The UCN2/UCN3 (Opsin) Paralogon

The two sea lamprey UCN2/UCN3-like genes map to scaffolds 57 and 82, respectively. Genes that belong to this paralogon were also found in four other scaffolds (**Figure 6**). A total of 17 neighboring gene families were identified and characterized that are in synteny with this paralogon in gnathostomes. Detailed analyses of these genomic regions identified two gene families that are quartets in spotted gar, and two families that are triplets. Several pairs were identified. Interestingly, both of the gene families that are quartets in spotted gar are so also in the sea lamprey genome, namely the FOXP family and the LRRN family. One of the spotted gar triplets, WNK, is likewise a triplet in sea lamprey. The other spotted gar triplet, GATA1/2/3, has a more complex situation in the sea lamprey with as many as six members, one of which might represent an ancestral fourth member in a 2R quartet (these four are shown in **Figure 6**), whereas two have arisen by duplication in the lamprey lineage (see tree in Supplementary File 5).

Considering that there are also several gene pairs present in these chromosomal regions in the various

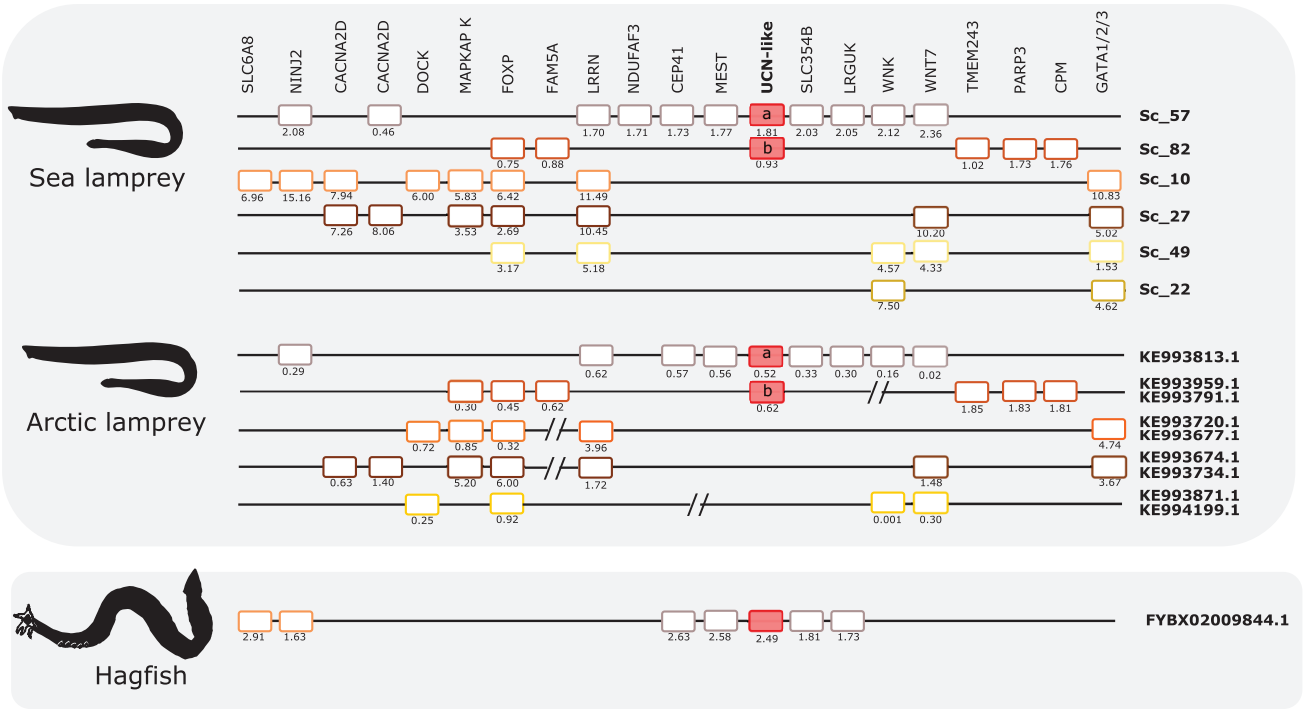


**FIGURE 8 |** Gene synteny of the opsin (UCN) paralogon in cyclostomes. The homolog genome region in sea lamprey, Arctic lamprey, and hagfish were compared. To highlight for potential gene shuffling/crossing-over events between the two lampreys and the two lampreys and hagfish, the sea lamprey genes in the same scaffold were represented by the same color. The exceptions are the UCN members which are represented by red-full boxes. Members of the same family in the sea lamprey genome or homologs in the other two cyclostome species are aligned. Gene chromosomal positions (Mb, below each gene) are shown. The sea lamprey genes are represented according to their position in the scaffolds and homologs in the Arctic lamprey and hagfish were positioned in relation to the sea lamprey genes. When possible different fragments of Arctic lamprey genome that possess a similar gene environment to a unique sea lamprey scaffold region were merged and this is marked by "//." Homologs for other lamprey UCN-like neighboring genes were found in the hagfish genome assembly but are not represented as they are located on multiple small scaffolds.

vertebrate genomes, a picture emerges of fourfold symmetry, albeit with weak representation of the fourth member in this paralogon. Overall, these observations point to a paralogon with four members, as with the CRH paralogon, thus most likely reflecting similar genomic events. Like in the other paralogon, many neighboring gene families remain as singletons both in lamprey and in gnathostomes.

Comparison of the two lamprey species reveals high similarity of not only gene repertoire, but also gene order (**Figure 8**). The hagfish scaffold with the UCN2/UCN3-like gene shares with the lampreys close synteny of four genes (CEP41, MEST, SLC35/4B, and LRGUK). Two other genes that are present in the hagfish scaffold (SLC6A8 and NINJ2) have lamprey orthologs that are syntenic with a different member of this paralogon, suggesting that gene translocation occurred after divergence of hagfishes and lampreys.

Phylogenetic analyses of all of the neighboring gene families overall display a similar topology as for the CRH/UCN genes. All of the multimember families display gene duplications in the time range of the early vertebrates WGD events (**Supplementary Figure S6**). Some of the neighboring gene families in this paralogon display similar features to those in the opioid paralogon described above, such as uneven evolutionary rates for some gene family members and variable branch positions for the cyclostome sequences.

## DISCUSSION

Our previous analyses of the CRH family allowed us to conclude that five genes arose before the gnathostome radiation as a result of the two WGD events, which triplicated one ancestral gene and duplicated the other (Cardoso et al., 2016). We also reported that a total of five CRH-family genes exist in lampreys as a group (Cardoso et al., 2016). However, in the sequence-based phylogenetic trees, the lamprey peptides did not cluster clearly with each of the five gnathostome peptides, and it was not possible to assign orthology. Thus, we could not determine whether lampreys share the same two WGD events as the gnathostomes. The recent genome assemblies of the sea lamprey *Petromyzon marinus* and the Arctic lamprey *Lethenteron camtschaticum*, and the hagfish *Eptatretus burgerii*, have allowed us to analyze scaffolds containing several neighboring genes to shed light on the agnathan-gnathostome divergence in relation to the two WGD events.

## Five CRH-Family Genes in Lampreys

Five CRH-family genes were found in both of the lamprey species, confirming our previous conclusion (Cardoso et al., 2016) and showing that both species share the same complete set. Two genes were found in the hagfish. In gnathostomes the mature peptides are encoded by a single exon, and this is the case also for the lampreys and the hagfish. Sequence-based phylogenetic analyses show that three lamprey peptides and one hagfish peptide cluster with the gnathostome CRH1/CRH2/UCN1 subfamily, and two lamprey peptides and one hagfish peptide cluster with the UCN2/UCN3 subfamily (**Figures 1**, **2**). This clearly demonstrates that the duplication that gave rise to the two ancestral CRH genes, the founders of each of the two subfamilies, had taken place well before the agnathan-gnathostome divergence. However, it was still not possible to determine orthology within each subfamily from the sequence analyses alone. Comparison of each of the five peptide genes between sea lamprey and Arctic lamprey showed high identity, consistent with recent divergence of the two species 10–30 Mya (Kuraku and Kuratani, 2006).

One of the two identified hagfish peptides clustered most closely to lamprey UCN_a (**Figures 1**, **2**), with which it displays gene environment conservation (**Figure 8**) and this is consistent with the hagfish lineage having undergone similar evolutionary events as the lamprey lineage (Fujimoto et al., 2013). Regarding orthology to the gnathostome sequences, this was as difficult to assign as for the lamprey sequences. This situation prevailed regardless whether the analysis was performed for the mature peptides (**Figure 1**) or the complete peptide precursors (**Figure 2**).

One reason that it is difficult to establish orthology between lamprey and gnathostome gene or protein family members may be that lamprey genes are under selection pressure to have high GC content, especially in coding regions (Kuraku and Kuratani, 2006; Kuraku, 2008; Smith et al., 2013; Manousaki et al., 2016). This probably explains the tendency of lamprey protein family members to cluster with one another in phylogenetic analyses rather than with their gnathostome orthologs (Qiu et al., 2011; Wotton and Shimeld, 2011; Fujimoto et al., 2013; Nah et al., 2014), as observed in the present report for both the CRH-family sequences and some of the neighboring gene families. Preference for certain synonymous codons occurs in bacteria, plants and invertebrates (Grantham et al., 1980; Qin et al., 2004), but in vertebrates it has been suggested to be gene specific and to occur at different rates across species (Palidwor et al., 2010; Romiguier et al., 2010; Aoi and Rourke, 2011). In contrast, lampreys seem to have a more general selection for GC-rich codons (Smith et al., 2013; Manousaki et al., 2016).

## Lamprey CRH-Family Genome Regions

As an additional approach to distinguish orthologs and paralogs, we and others have analyzed the repertoire of neighboring gene families in numerous studies. We therefore analyzed the CRH neighbors and their families in the scaffolds containing the lamprey and hagfish CRH-family genes. As shown in **Figures 5**, **6**, all of these agnathan scaffolds contain members of gene families that are located close to the gnathostome CRH-family genes. In total, the scaffolds covering the CRH-family gene neighborhood contain a total of 40 gene families, including the two CRH subfamilies and the two GATA subfamilies.

An important result of our analyses of the lamprey scaffolds is that some of the neighboring gene families are quartets: one in the CRH/opioid paralogon (TRIM) and two in the UCN/opsin paralogon (FOXP and LRRN). In addition, several neighboring families are triplets, six in the CRH/opioid paralogon and two in the UCN/opsin paralogon. Two of the lamprey families, RSPO and VSNL1 (**Figure 7**), actually consist of five members, and their phylogenetic trees suggest that all duplications took place in the time period of the cyclostome-gnathostome divergence (not shown). One of the UCN/opsin triplets also includes a few additional members, namely the GATA1/2/3 family, but the supernumerary members seem to be lamprey-specific duplicates (see tree in Supplementary File 5). A few such lineage-specific duplications have occurred also in the human and chicken lineages (**Figures 5**, **6**).

Counting also the CRH1/CRH2/UCN1 family, this means that in total 13 out of 37 gene families in these two sets of scaffolds contain gene families consisting of three or four (or five) members in lampreys. In addition, pairs of related genes are present in different combinations of these chromosomal regions, further corroborating relatedness. The most parsimonious explanation for this would be a quartet of related chromosomal regions. By parsimonious reasoning, the simplest explanation would be that lampreys have undergone the same two WGD events as gnathostomes, thus sharing the same 1R and 2R genome doublings. The number of gene families with quartets and triplets of genes is not quite as high as in spotted gar, suggesting that lampreys have lost some family members, as seems to be the case for human and chicken.

Assuming that lampreys have undergone the same duplication events as gnathostomes for the CRH-containing chromosome regions, it might theoretically be possible to infer which of the chromosome members in the lamprey paralogon that corresponds to which member in the gnathostomes, i.e., to assign orthology between these basal vertebrate lineages. Also, the pattern of neighboring singletons or gene losses should allow identification of orthologous chromosomal regions between lamprey and gnathostomes. However, careful scrutiny of the scaffolds shows that some lamprey regions appear to be combinations of two (or more) gnathostome regions. For instance, in the UCN/opsin paralogon, there are two UCN genes in the lamprey that might correspond to spotted gar linkage groups LG5 and LG8, both of which have rather complete representation of the neighboring gene families. One of these could correspond to lamprey scaffold 57 and the other to scaffold 82. However, when looking at the gene families that are quartets, namely FOXP and LRRN, neither of these two scaffolds can be complemented by either of the scaffolds 10, 27, or 49, because all three of these contain representatives for both of these quartet families. There are a number of such examples where combination of chromosomes in either lamprey or one of the vertebrates leads to "collision," as has also been observed for the somatostatin gene regions in *L. camtschaticum* (a.k.a. *L. japonicum*) (Tostivint et al., 2016).

One possible explanation for this is that either the lamprey or the gnathostome ancestor's chromosome underwent crossing-over, presumably rather soon after WGD when the duplicated chromosomes were still quite similar to one another also in intergenic regions. Alternately, patterns such as this might simply reflect independent paralog loss, fission, and local duplication events in lamprey vs. gnathostome genomes.

Recently, three CRH-family members were amplified from the sea lamprey brain cDNA (Endsin et al., 2017). The three peptides were suggested to be the homologs of gnathostome CRH1, UCN1, and UCN3. However, our more extensive analyses based on both phylogeny and gene synteny suggest that this may not be the case. These authors proposed that the three lamprey sequences arose by two duplication events before the cyclostome-gnathostome divergence and that two subsequent duplication events expanded the repertoire to five genes in the gnathostome lineage (Endsin et al., 2017), although our previous study had already demonstrated five CRH-family members in lampreys (Cardoso et al., 2016). Our results presented here add further support that all of the duplications of the ancestral CRH-bearing region took place before the lamprey-gnathostome divergence.

Our results clearly establish that the two CRH-family lineages separated before the lamprey-gnathostome divergence. The peptide sequences in both lamprey and hagfish show that the CRH1/CRH2/UCN1 ancestor and the UCN2/UCN3 ancestor arose before the origin of the vertebrates. These two ancestral genes, the founders of the two subfamilies, became located in separate chromosomes that subsequently formed the two paralogons in the 2R events. Interestingly, one other gene family has members in the same two paralogons: the GATA family of transcription factors. Our previous analysis (Cardoso et al., 2016) showed that GATA1/2/3 family members are located in the UCN/opsin paralogon and that GATA4/5/6 are in the CRH/opioid paralogon, as had been reported previously (Hwang et al., 2013). Both of the GATA lineages have been identified in the nematode *C. elegans*, in several arthropods including the fruit-fly *D. melanogaster,* and also in annelids, showing that the first gene duplication occurred prior to the protostome-deuterostome divergence (Gillis et al., 2007, 2008) and probably much earlier. In protostomes, homologs of vertebrate CRH have only been described in insects (DH44 peptide). We therefore wondered whether CRH and GATA were duplicated in the same event at the same time point. However, only a single gene homolog of vertebrate CRH exists in invertebrate genomes (Mirabeau and Joly, 2013), suggesting that GATA and CRH were duplicated independently and maybe ended up in the same two gene regions by chance.

### Lampreys and Gnathostomes Probably Share the Same WGD Events

Our analyses of the gnathostome CRH chromosomal regions and homologous regions in sea lamprey and Arctic lamprey suggest that lampreys share not only the first WGD event (1R) with gnathostomes but also the second WGD event. Some previous analyses of the sea lamprey genome described gene family relationships and conserved syntenic regions consistent with a shared 2R event (Caputo Barucchi et al., 2013; Decatur et al., 2013; Smith et al., 2018). However, analyses of the Hox gene clusters of the Arctic lamprey genome led to the suggestion that lamprey genomes might have undergone a third genome doubling and that 2R and perhaps even 1R could have been independent events from those in gnathostomes (Mehta et al., 2013; Zhang et al., 2017). A more recent analysis of vertebrate chromosome evolution concluded that cyclostomes share the 2R event with gnathostomes, see figure 7 in Sacerdot et al. (2018). However, other authors settled for a shared 1R event and suggested that lampreys subsequently have undergone "one or more additional duplication(s)" (Simakov et al., 2020). Our characterization of the lamprey CRH-family members and their neighboring gene environment reveals striking similarity to the chromosome regions in gnathostomes, which appears to be in line with both WGD events being shared with gnathostomes. However, it remains theoretically possible that lampreys have had independent WGD (or other duplication) events as long as the sequence-based phylogenies cannot be completely resolved for the cyclostome-gnathostome divergence in relation to the genome doublings, as has been proposed for the Hox clusters (Smith et al., 2018).

## CONCLUSION

Homologs of the five gnathostome CRH-family members and their neighboring gene families were characterized in the genomes of two lampreys and one hagfish. In lampreys, like in gnathostomes, five members of the CRH family were identified, presumably representing orthologs. However, the exact orthologous relationships could not be resolved even after synteny and paralogon analyses, presumably due to crossing-over events or other rearrangements that have changed their chromosomal positions. Although orthology between lampreys and gnathostomes cannot be determined with certainty, our detailed analyses nevertheless find strong support for lamprey chromosome quadruplication in two paralogons, and thus that lampreys (and by inference also the hagfishes if these two lineages of agnathans are monophyletic) have undergone duplications that are highly similar to those experienced by gnathostomes. Most parsimoniously, these duplication events were probably the same as in the gnathostomes, i.e., WGD events, and the lamprey-gnathostome divergence took place soon after the second WGD event.

## DATA AVAILABILITY STATEMENT

The accession numbers of all sequences used in this study were retrieved from public databases and their accession numbers are available in **Supplementary Material**.

## AUTHOR CONTRIBUTIONS

JC, CB, and DL planned the study, evaluated the results, and wrote the manuscript. JC and CB collected the data

and performed the analyses. All authors critically read and contributed to improve the manuscript.

## FUNDING

This study received Portuguese national funds from FCT - Foundation for Science and Technology through project UIDB/04326/2020. DL received a guest researcher grant for JC from Uppsala University Faculty of Medicine.

## ACKNOWLEDGMENTS

The authors acknowledge Dr. Jeramiah Smith (University of Kentucky, United States) for providing privileged access to the sea lamprey genome scaffolds. They also thank Dr. David Lagman and the two reviewers for comments on our manuscript.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins.2020.00672/full#supplementary-material

**FIGURE S1 |** Bayesian inference tree of the CRH mature peptides. Branch support values (posterior probability values) are shown and tree was rooted with the tunicate CRH-like precursors. Tree was rooted with the tunicate CRH-family members. A simplified radial tree is represented in **Figure 1**. Accession numbers of the sequences used are available in **Supplementary Table S2**.

**FIGURE S2 |** Bayesian inference tree of the CRH- peptide precursors. Branch support values (posterior probability values) are shown and tree was rooted with the tunicate CRH-like precursors. Tree was rooted with the tunicate CRH-family members. A simplified radial tree is represented in **Figure 2**. Accession numbers of the sequences used are available in **Supplementary Table S2**.

**FIGURE S3 |** Maximum likelihood tree of the CRH mature peptides. Tree was constructed using the chordate deduced full-length CRH-family precursors with the PhyML 3.0 algorithm on the ATGC bioinformatics platform using LG substitution and 100 bootstrap replicates. Tree was rooted with the tunicate CRH-family members. Accession numbers of the sequences used are available in **Supplementary Table S2**.

**FIGURE S4 |** Maximum likelihood tree of the CRH peptide precursors. Tree was constructed with the chordate deduced full-length CRH-family precursors using the PhyML 3.0 algorithm on the ATGC bioinformatics platform using LG substitution and 100 bootstrap replicates. Tree was rooted with the tunicate CRH-family members. Accession numbers of the sequences used are available in **Supplementary Table S2**.

**FIGURE S5 |** Phylogenetic analysis of the opioid (CRH1/CRH2/UCN1) paralogon neighboring gene families. Trees were constructed with the PhyML 3.0 algorithm on the ATGC bioinformatics platform with automatic selection model and sequence branching support was given by the Approximate Likelihood-Ratio Test (aLTR). All neighboring gene trees were rooted using the invertebrate family members. The accession numbers and symbols of all the sequences used are available in **Supplementary Table S3**.

**FIGURE S6 |** Phylogenetic analysis of the opsin (UCN2/UCN3) paralogon neighboring gene families. Trees were constructed with the PhyML 3.0 algorithm on the ATGC bioinformatics platform with automatic selection model and sequence branching support was given by the Approximate Likelihood-Ratio Test (aLTR). All neighboring gene trees were rooted using the invertebrate family members. The accession numbers and symbols of all the sequences used are available in **Supplementary Table S4**.

**TABLE S1 |** Size (base pairs, bp) of the sea lamprey and Arctic lamprey CRH-family scaffolds. The analyzed genome regions of the sea lamprey scaffolds were selected using as starting point the position of the lamprey CRH-family member and retrieving approximately 1.5 Kb upstream and downstream of this gene.

**TABLE S2 |** Accession numbers of all the CRH-family members used in phylogeny. The genome localization (chromosome and position) is give. ni, not identified; na, not available; un, uncharacterized.

**TABLE S3 |** Accession numbers (ENSEMBL or NCBI) and nomenclature adopted for the neighboring gene families of the CRH1/CRH2/UCN1 paralogon that were used for phylogenetic analysis (**Supplementary Figure S5**). Gene families were ordered as they appear in **Figure 5**.

**TABLE S4 |** Accession numbers (ENSEMBL or NCBI) and nomenclature adopted for the neighboring gene families of the UCN2/UCN3 paralogon that were used for phylogenetic analysis (**Supplementary Figure S6**). Gene families were ordered as they appear in **Figure 6**.

## REFERENCES

Aoi, M. C., and Rourke, B. C. (2011). Interspecific and intragenic differences in codon usage bias among vertebrate myosin heavy-chain genes. *J. Mol. Evol.* 73, 74–93. doi: 10.1007/s00239-011-9457-0

Audsley, N., Kay, I., Hayes, T. K., and Coast, G. M. (1995). Cross reactivity studies of CRF-related peptides on insect Malpighian tubules. *Comp. Biochem. Physiol. Part A Physiol.* 110, 87–93. doi: 10.1016/0300-9629(94)00132-D

Cabrero, P., Radford, J. C., Broderick, K. E., Costes, L., Veenstra, J. A., Spana, E. P., et al. (2002). The Dh gene of *Drosophila melanogaster* encodes a diuretic peptide that acts through cyclic AMP. *J. Exp. Biol.* 205, 3799–3807.

Caputo Barucchi, V., Giovannotti, M., Nisi Cerioni, P., and Splendiani, A. (2013). Genome duplication in early vertebrates: insights from agnathan cytogenetics. *Cytogenet. Genome Res.* 141, 80–89. doi: 10.1159/000354098

Cardoso, J. C. R., Bergqvist, C. A., Félix, R. C., and Larhammar, D. (2016). Corticotropin-releasing hormone family evolution: five ancestral genes remain in some lineages. *J. Mol. Endocrinol.* 57, 73–86. doi: 10.1530/JME-16-0051

Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M., and Robles, M. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21, 3674–3676. doi: 10.1093/bioinformatics/bti610

Decatur, W. A., Hall, J. A., Smith, J. J., Li, W., and Sower, S. A. (2013). Insight from the lamprey genome: glimpsing early vertebrate development via neuroendocrine-associated genes and shared synteny of gonadotropin-releasing hormone (GnRH). *Gen. Comp. Endocrinol.* 192, 237–245. doi: 10.1016/j.ygcen.2013.05.020

Dreborg, S., Sundström, G., Larsson, T. A., and Larhammar, D. (2008). Evolution of vertebrate opioid receptors. *Proc. Natl. Acad. Sci. U.S.A.* 105, 15487–15492. doi: 10.1073/pnas.0805590105

Dunn, A. J., and Berridge, C. W. (1990). Physiological and behavioral responses to corticotropin-releasing factor administration: is CRF a mediator of anxiety or stress responses? *Brain Res. Rev.* 15, 71–100. doi: 10.1016/0165-0173(90)90012-D

Endsin, M. J., Michalec, O., Manzon, L. A., Lovejoy, D. A., and Manzon, R. G. (2017). CRH peptide evolution occurred in three phases: evidence from characterizing sea lamprey CRH system members. *Gen. Comp. Endocrinol.* 240, 162–173. doi: 10.1016/j.ygcen.2016.10.009

Fox, J. H., and Lowry, C. A. (2013). Corticotropin-releasing factor-related peptides, serotonergic systems, and emotional behavior. *Front. Neurosci.* 7:169. doi: 10.3389/fnins.2013.00169

Fujimoto, S., Oisi, Y., Kuraku, S., Ota, K. G., and Kuratani, S. (2013). Non-parsimonious evolution of hagfish Dlx genes. *BMC Evol. Biol.* 13:15. doi: 10.1186/1471-2148-13-15

Gillis, W. J., Bowerman, B., and Schneider, S. Q. (2007). Ectoderm- and endomesoderm-specific GATA transcription factors in the marine annelid Platynereis dumerilli. *Evol. Dev.* 9, 39–50. doi: 10.1111/j.1525-142X.2006.00136.x

Gillis, W. Q., Bowerman, B. A., and Schneider, S. Q. (2008). The evolution of protostome GATA factors: molecular phylogenetics, synteny, and intron/exon structure reveal orthologous relationships. *BMC Evol. Biol.* 8:112. doi: 10.1186/1471-2148-8-112

Grantham, R., Gautier, C., Gouy, M., Mercier, R., and Pavé, A. (1980). Codon catalog usage and the genome hypothesis. *Nucleic Acids Res.* 8:197. doi: 10.1093/nar/8.1.197-c

Grone, B. P., and Maruska, K. P. (2015a). A second corticotropin-releasing hormone gene (CRH2) is conserved across vertebrate classes and expressed in the hindbrain of a basal Neopterygian fish, the spotted gar (Lepisosteus oculatus). *J. Comp. Neurol.* 523, 1125–1143. doi: 10.1002/cne.23729

Grone, B. P., and Maruska, K. P. (2015b). Divergent evolution of two corticotropin-releasing hormone (CRH) genes in teleost fishes. *Front. Neurosci.* 9:365. doi: 10.3389/fnins.2015.00365

Gysling, K., Forray, M. I., Haeger, P., Daza, C., and Rojas, R. (2004). Corticotropin-releasing hormone and urocortin: redundant or distinctive functions? *Brain Res. Rev.* 47, 116–125. doi: 10.1016/j.brainresrev.2004.06.001

Hwang, J. I., Moon, M. J., Park, S., Kim, D. K., Cho, E. B., Ha, N., et al. (2013). Expansion of secretin-like G protein-coupled receptors and their peptide ligands via local duplications before and after two rounds of whole-genome duplication. *Mol. Biol. Evol.* 30, 1119–1130. doi: 10.1093/molbev/mst031

Jaillon, O., Aury, J. M., Brunet, F., Petit, J. L., Stange-Thomann, N., Maucell, E., et al. (2004). Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype. *Nature* 431, 946–957. doi: 10.1038/nature03025

Kawada, T., Sekiguchi, T., Sakai, T., Aoyama, M., and Satake, H. (2010). Neuropeptides, hormone peptides, and their receptors in ciona intestinalis: an update. *Zoolog. Sci.* 27, 134–153. doi: 10.2108/zsj.27.134

Koob, G. F., and Heinrichs, S. C. (1999). A role for corticotropin releasing factor and urocortin in behavioral responses to stressors. *Brain Res.* 848, 141–152. doi: 10.1016/S0006-8993(99)01991-5

Kuraku, S. (2008). Insights into cyclostome phylogenomics: pre-2R or Post-2R. *Zool. Sci.* 25, 960–968. doi: 10.2108/zsj.25.960

Kuraku, S., and Kuratani, S. (2006). Time scale for cyclostome evolution inferred with a phylogenetic diagnosis of hagfish and lamprey cDNA sequences. *Zool. Sci.* 23, 1053–1064. doi: 10.2108/zsj.23.1053

Lagman, D., Ocampo Daza, D., Widmark, J., Abalo, X. M., Sundström, G., and Larhammar, D. (2013). The vertebrate ancestral repertoire of visual opsins, transducin alpha subunits and oxytocin/vasopressin receptors was established by duplication of their shared genomic region in the two rounds of early vertebrate genome duplications. *BMC Evol. Biol.* 13:238. doi: 10.1186/1471-2148-13-238

Larsson, A. (2014). AliView: a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics* 30, 3276–3278. doi: 10.1093/bioinformatics/btu531

Le, S. Q., and Gascuel, O. (2008). An improved general amino acid replacement matrix. *Mol. Biol. Evol.* 25, 1307–1320. doi: 10.1093/molbev/msn067

Lefort, V., Longueville, J. E., and Gascuel, O. (2017). SMS: smart Model Selection in PhyML. *Mol. Biol. Evol.* 34, 2422–2424. doi: 10.1093/molbev/msx149

Lewis, K., Li, C., Perrin, M. H., Blount, A., Kunitake, K., Donaldson, C., et al. (2001). Identification of urocortin III, an additional member of the corticotropin-releasing factor (CRF) family with high affinity for the CRF2 receptor. *Proc. Natl. Acad. Sci. U.S.A.* 98, 7570–7575. doi: 10.1073/pnas.121165198

Lovejoy, D. A., and Balment, R. J. (1999). Evolution and physiology of the corticotropin-releasing factor (CRF) family of neuropeptides in vertebrates. *Gen. Comp. Endocrinol.* 115, 1–22. doi: 10.1006/gcen.1999.7298

Lovejoy, D. A., and de Lannoy, L. (2013). Evolution and phylogeny of the corticotropin-releasing factor (CRF) family of peptides: expansion and specialization in the vertebrates. *J. Chem. Neuroanat.* 54, 50–56. doi: 10.1016/j.jchemneu.2013.09.006

Manousaki, T., Qiu, H., Noro, M., Hildebrand, F., Meyer, A., and Kuraku, S. (2016). "Molecular evolution in the lamprey genomes and its relevance to the timing of whole genome duplications," in *Jawless Fishes of the World?*, Vol. 1, eds A. Orlov and R. Beamish (Newcastle: Cambridge Scholars Publishing), 2–16.

Mehta, T. K., Raèi, È, Yamasaki, S., Lee, A. P., Lian, M. M., Tay, B. H., et al. (2013). Evidence for at least six Hox clusters in the Japanese lamprey (*Lethenteron japonicum*). *Proc. Natl. Acad. Sci. U.S.A.* 110, 16044–16049. doi: 10.1073/pnas.1315760110

Miller, M. A., Pfeiffer, W., and Schwartz, T. (2010). "Creating the CIPRES Science Gateway for inference of large phylogenetic trees," in *2010 Gateway Computing Environments Workshop, GCE 2010*, New Orleans, LA.

Mirabeau, O., and Joly, J. S. (2013). Molecular evolution of peptidergic signaling systems in bilaterians. *Proc. Natl. Acad. Sci. U.S.A.* 110, E2028–E2203. doi: 10.1073/pnas.1219956110

Nah, G. S. S., Tay, B. H., Brenner, S., Osato, M., and Venkatesh, B. (2014). Characterization of the runx gene family in a jawless vertebrate, the Japanese lamprey (*Lethenteron japonicum*). *PLoS One* 9:e113445. doi: 10.1371/journal.pone.0113445

Nakatani, Y., Takeda, H., Kohara, Y., and Morishita, S. (2007). Reconstruction of the vertebrate ancestral genome reveals dynamic genome reorganization in early vertebrates. *Genome Res.* 17, 1254–1265. doi: 10.1101/gr.6316407

Nock, T. G., Chand, D., and Lovejoy, D. A. (2011). Identification of members of the gonadotropin-releasing hormone (GnRH), corticotropin-releasing factor (CRF) families in the genome of the holocephalan, *Callorhinchus milii* (elephant shark). *Gen. Comp. Endocrinol.* 171, 237–244. doi: 10.1016/j.ygcen.2011.02.001

Palidwor, G. A., Perkins, T. J., and Xia, X. (2010). A general model of Codon bias due to GC mutational bias. *PLoS One* 5:e13431. doi: 10.1371/journal.pone.0013431

Putnam, N. H., Butts, T., Ferrier, D. E. K., Furlong, R. F., Hellsten, U., Kawashima, T., et al. (2008). The amphioxus genome and the evolution of the chordate karyotype. *Nature* 453, 1064–1071. doi: 10.1038/nature06967

Qin, H., Wu, W. B., Comeron, J. M., Kreitman, M., and Li, W. H. (2004). Intragenic spatial patterns of codon usage bias in prokaryotic and eukaryotic genomes. *Genetics* 68, 2245–2260. doi: 10.1534/genetics.104.030866

Qiu, H., Hildebrand, F., Kuraku, S., and Meyer, A. (2011). Unresolved orthology and peculiar coding sequence properties of lamprey genes: the KCNA gene family as test case. *BMC Genomics* 12:325. doi: 10.1186/1471-2164-12-325

Reyes, T. M., Lewis, K., Perrin, M. H., Kunitake, K. S., Vaughan, J., Arias, C. A., et al. (2001). Urocortin II: a member of the corticotropin-releasing factor (CRF) neuropeptide family that is selectively bound by type 2 CRF receptors. *Proc. Natl. Acad. Sci. U.S.A.* 98, 2843–2848. doi: 10.1073/pnas.051626398

Roberts, B. W., Didier, W., Rai, S., Johnson, N. S., Libants, S., Yun, S. S., et al. (2014). Regulation of a putative corticosteroid, 17,21-dihydroxypregn-4-ene,3,20-one, in sea lamprey, *Petromyzon marinus*. *Gen. Comp. Endocrinol.* 196, 17–25. doi: 10.1016/j.ygcen.2013.11.008

Romiguier, J., Ranwez, V., Douzery, E. J. P., and Galtier, N. (2010). Contrasting GC-content dynamics across 33 mammalian genomes: relationship with life-history traits and chromosome sizes. *Genome Res.* 20, 1001–1009. doi: 10.1101/gr.104372.109

Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D. L., Darling, A., Höhna, S., et al. (2012). MrBayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61, 539–542. doi: 10.1093/sysbio/sys029

Sacerdot, C., Louis, A., Bon, C., Berthelot, C., and Roest Crollius, H. (2018). Chromosome evolution at the origin of the ancestral vertebrate genome. *Genome Biol.* 19:166. doi: 10.1186/s13059-018-1559-1

Shibahara, S., Morimoto, Y., Furutani, Y., Notake, M., Takahashi, H., Shimizu, S., et al. (1983). Isolation and sequence analysis of the human corticotropin-releasing factor precursor gene. *EMBO J.* 2, 775–779. doi: 10.1002/j.1460-2075.1983.tb01499.x

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using clustal omega. *Mol. Syst. Biol.* 7:539. doi: 10.1038/msb.2011.75

Simakov, O., Marlétaz, F., Yue, J.-X., O'Connell, B., Jenkins, J., Brandt, A., et al. (2020). Deeply conserved synteny resolves early events in vertebrate evolution. *Nat. Ecol. Evol.* 4, 820–830. doi: 10.1038/s41559-020-1156-z

Smith, J. J., and Keinath, M. C. (2015). The sea lamprey meiotic map improves resolution of ancient vertebrate genome duplications. *Genome Res.* 25, 1081–1090. doi: 10.1101/gr.184135.114

Smith, J. J., Kuraku, S., Holt, C., Sauka-Spengler, T., Jiang, N., Campbell, M. S., et al. (2013). Sequencing of the sea lamprey (*Petromyzon marinus*) genome provides insights into vertebrate evolution. *Nat. Genet.* 45, 415–421. doi: 10.1038/ng.2568

Smith, J. J., Timoshevskaya, N., Ye, C., Holt, C., Keinath, M. C., Parker, H. J., et al. (2018). The sea lamprey germline genome provides insights into programmed genome rearrangement and vertebrate evolution. *Nat. Genet.* 50, 270–277. doi: 10.1038/s41588-017-0036-1

Stanke, M., Diekhans, M., Baertsch, R., and Haussler, D. (2008). Using native and syntenically mapped cDNA alignments to improve de novo gene finding. *Bioinformatics* 24, 637–644. doi: 10.1093/bioinformatics/btn013

Stanke, M., Steinkamp, R., Waack, S., and Morgenstern, B. (2004). AUGUSTUS: a web server for gene finding in eukaryotes. *Nucleic Acids Res.* 32, W309–W312. doi: 10.1093/nar/gkh379

Sundström, G., Dreborg, S., and Larhammar, D. (2010). Concomitant duplications of opioid peptide and receptor genes before the origin of jawed vertebrates. *PLoS One* 5:e10512. doi: 10.1371/journal.pone.0010512

Thompson, R. C., Seasholtz, A. F., and Herbert, E. (1987). Rat corticotropin-releasing hormone gene: sequence and tissue-specific expression. *Mol. Endocrinol.* 1, 363–370. doi: 10.1210/mend-1-5-3

Tostivint, H., Dettaï, A., Quan, F. B., Ravi, V., Tay, B. H., Rodicio, M. C., et al. (2016). Identification of three somatostatin genes in lampreys. *Gen. Comp. Endocrinol.* 237, 89–97. doi: 10.1016/j.ygcen.2016.08.006

Vale, W., Spiess, J., Rivier, C., and Rivier, J. (1981). Characterization of a 41-residue ovine hypothalamic peptide that stimulates secretion of corticotropin and β-endorphin. *Science* 213, 1394–1397. doi: 10.1126/science.6267699

Wotton, K. R., and Shimeld, S. M. (2011). Analysis of lamprey clustered Fox genes: insight into Fox gene evolution and expression in vertebrates. *Gene* 489, 30–40. doi: 10.1016/j.gene.2011.08.007

Zhang, H., Ravi, V., Tay, B. H., Tohari, S., Pillai, N. E., Prasad, A., et al. (2017). Lampreys, the jawless vertebrates, contain only two ParaHox gene clusters. *Proc. Natl. Acad. Sci. U.S.A.* 114, 9146–9151. doi: 10.1073/pnas.1704457114

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Endocrinology

ORIGINAL RESEARCH
published: 27 August 2020
doi: 10.3389/fendo.2020.00529

# Corticotropin-Releasing Factor: An Ancient Peptide Family Related to the Secretin Peptide Superfamily

*Ola M. Michalec*[1], *Belinda S. W. Chang*[1], *Nathan R. Lovejoy*[2] *and David A. Lovejoy*[1]*

[1] Department of Cell and Systems Biology, University of Toronto - St. George Campus, Toronto, ON, Canada, [2] Department of Biological Sciences, University of Toronto Scarborough, Toronto, ON, Canada

Corticotropin-releasing factor (CRF) is the hypothalamic releasing peptide that regulates the hypothalamic-pituitary-adrenal/inter-renal (HPA/I) axis in vertebrates. Over the last 25 years, there has been considerable discussion on its paralogs genes, urotensin-I/urocortin-1, and urocortins-2 and-3 and their subsequent role in the vertebrate stress response. Phylogenetically, the CRF family of peptides also belong to the diverse assemblage of Secretin- and Calcitonin-based peptides as evidenced by comparative-based studies of both their ligand and G-protein-coupled receptor (GPCR) structures. Despite this, the common origin of this large assemblage of peptides has not been ascertained. An unusual peptide, teneurin-C-terminal associated peptide (TCAP), reported in 2004, comprises the distal extracellular tip of the teneurin transmembrane proteins. Further studies indicated that this teneurin region binds to the latrophilin family of GPCRs. Initially thought to be a member of the Secretin GPCR family, evidence indicates that the latrophilins are a member of the Adhesion family of GPCRs and are related to the common ancestor of both Adhesion and Secretin GPCR families. In this study, we posit that TCAP may be a distantly related ancestor of the CRF-Calcitonin-Secretin peptide family and evolved near the base of metazoan phylogeny.

**Keywords: secretin superfamily, TCAP, teneurin, adhesion GPCRs, evolution, metabolism**

OPEN ACCESS

***Edited by:***
*Vance L. Trudeau,*
*University of Ottawa, Canada*

***Reviewed by:***
*Richard Giuseppe Manzon,*
*University of Regina, Canada*
*James A. Carr,*
*Texas Tech University, United States*

***Correspondence:**
*David A. Lovejoy*
*david.lovejoy@.utoronto.ca*

***Specialty section:***
*This article was submitted to Neuroendocrine Science, a section of the journal Frontiers in Endocrinology*

***Received:*** *20 April 2020*
***Accepted:*** *29 June 2020*
***Published:*** *27 August 2020*

***Citation:***
*Michalec OM, Chang BSW, Lovejoy NR and Lovejoy DA (2020) Corticotropin-Releasing Factor: An Ancient Peptide Family Related to the Secretin Peptide Superfamily. Front. Endocrinol. 11:529. doi: 10.3389/fendo.2020.00529*

## INTRODUCTION

Corticotropin-releasing factor (CRF) is the critical hypothalamic releasing factor that regulates the hypothalamus-pituitary-adrenal/inter-renal (HPA/1) axis in vertebrates, yet after some 40 years after its discovery, numerous questions still exist regarding when, why, and how this peptide evolved. We hypothesize that due to the high level of primary structure similarity among CRF paralogs and related peptide lineages (e.g., calcitonin, secretin) there was likely an ancestor peptide common to this cluster. We further suggest that the "teneurin C-terminal associated peptides" (TCAP) represent an extant candidate lineage related to the hypothetical common ancestor.

The discovery of CRF in the early 1980s (1) occurred about the same time as the discovery of other peptides of similar structure [sauvagine (2); urotensin-I (3)]. Later, Vale and his laboratory characterized a mammalian version of sauvagine/urotensin-I in rat brain that they termed, urocortin (4). Further phylogenetic studies suggested that mammalian urocortin, amphibian sauvagine, and fish urotensin-I were orthologs of the same gene (5). In 2001, the structures of two novel related peptides were reported by the Vale laboratory who named the peptides, urocortin 2

and 3 (6, 7) and by Hsu and Hsieh (8) who termed the peptides as "stresscopin" and "stresscopin-related peptide." These novel CRF family homologs were subsequently established to be a separate paralogs lineage of CRF and the urotensin-I/sauvagine/urocortin grouping (5, 9–13). In parallel to studies of vertebrate CRF isoforms, the presence of related peptides were reported in insects and arthropods (12, 14–16). Therefore, the high degree of structural similarity among CRF-like peptides in both deuterostome (e.g., chordates) and protostomes (e.g., arthropods) indicated that an ancestral peptide with CRF family primary structure attributes was present before the bifurcation of these metazoan lineages. Importantly, this ancestral peptide appeared to exist in a physiologically mature form indicative of a distant lineage that likely radiated as other ancestral peptides with distinct but overlapping functions. The identity of these hypothetical ancestral peptides has remained elusive, however, it is plausible that these lineages led to the evolution and expansion of the secretin and calcitonin family of peptides (11, 12).

The Secretin superfamily of peptides is a diverse assemblage of peptide lineages with overlapping functions utilizing structurally related receptors. The nomenclature describing the phylogeny of the secretin grouping of peptides and receptors is confusing. In order to clarify this, we have used the term "secretin family" to denote those peptides that are thought to be part of a direct monophyletic clade (e.g., secretin, PACAP, VIP, and glucagon paralogs). For the inclusion of the wider group which include CRF and calcitonin, we have referred to this as the "Secretin superfamily." Due, in part, to the similarity and structural conservation of their cognate receptors, the Secretin family G-Protein Coupled Receptors (GPCR) was defined as a distinct clade (17). The Secretin superfamily of peptides is one of the five main families of ligands that bind to GPCRs. The GPCRs have most recently been classified into five main families using the GRAFS system; Glutamate (G), Rhodopsin (R), Adhesion (A), Frizzled/Taste2 (F), and Secretin (S) (17). Notably, both CRF and calcitonin receptors are included within the Secretion GPCR family. Among these, Adhesion and Secretin GPCRs are the most evolutionarily ancient (18). Adhesion GPCRs have a characteristically long N-terminus rich in serine and threonine residues whereas Secretin GPCRs have a characteristic hormone-binding domain (HBD) in their N-terminal region (18). Secretin-related receptors form a single monophyletic clade that derived from the Adhesion GPCRs (18, 19). Adhesion GPCR genes have been identified in choanoflagellate and sea anemone genomes but Secretin GPCR genes have not suggesting that Adhesion GPCRs are more evolutionarily ancient than Secretin GPCRs (18). Interestingly, some derived phylogenetically younger Adhesion GPCR members possess an HBD with highly conserved amino acid sequences and similar splice site motifs as Secretin GPCRs. These observations led, in part, to the hypothesis that the Secretin GPCR clade was derived from an offshoot of the Adhesion GPCR lineage. However, although the data linking the Adhesion and Secretin superfamilies were compelling, evidence of a structurally related peptide ligand linking the two receptor clades was lacking.

One such lineage of Adhesion GPCRs that does possess a HBD with similar structural motifs to Secretin GPCRs are the latrophilins (LPHN) or ADGRL (Adhesion G-protein coupled receptors, subfamily L). It was originally considered a new type of Secretin GPCR, due to its characteristic HBD, but has now been re-classified as an Adhesion GPCR (17). The first identified ligand for ADGRL was α-latrotoxin, a peptide component of black widow spider toxin venom that specifically targets vertebrates (20) and shares major sequence similarity with other Secretin superfamily ligands (21). The data suggest that these peptides have a common origin. Although, α-latrotoxin was an exogenous ligand, the high affinity binding of this soluble peptide to ADGRL indicated that this receptor had the potential to bind and be activated by an endogenous peptide similar to the structure of α-latrotoxin. The search for this theoretical ligand led to the identification of the teneurin transmembrane proteins as a likely suspect.

Several recent studies established that the distal region of the extracellular domain of the teneurin transmembrane proteins binds ADGRL with high affinity and activates the receptor. Silva et al. (22) first discovered that teneurin-2, expressed on post-synaptic dendritic branches, binds LPHN-1 expressed on pre-synaptic nerve terminals to form a trans-synaptic complex. Similar trans-cellular interactions were observed between teneurins-2 and 4 and all three LPHNs (23) and between teneurin-1 and LPHN-3 (24). A C-terminal fragment of teneurin-2, named Lasso, triggered an increase in cytosolic $Ca^{2+}$ in Nb2a cells overexpressing LPHN-1 and in pre-synaptic nerve terminals of hippocampal cells (22). This distal region of the teneurin extracellular domain contains a peptide-like sequence termed "teneurin C-terminal associated peptide" (TCAP). The TCAPs are a family of four bioactive peptides that are 40–41 amino acids in length and are located at the C-terminus of each of the teneurin transmembrane proteins (25, 26). TCAPs possess a cleavage motif at the N-terminus and an amidation motif at the C-terminus (27) and may be autolytically cleaved from teneurins upon binding with LPHN (28, 29). TCAP shares about 20% sequence similarity with CRF and calcitonin, members of the Secretin superfamily of ligands, suggesting a common evolutionary origin (27). Moreover, our laboratory has recently identified that teneurin C-terminal associated peptide (TCAP)-1 is likely an endogenous ligand that interacts with the HBD of LPHN (30).

Therefore, as TCAP binds to an Adhesion GPCR and shares sequence similarity to CRF and calcitonin, ligands that bind to Secretin GPCR receptors that are classified as being most closely related to ancestral Adhesion GPCRs, this prompted the investigation of TCAP as a progenitor of the Secretin superfamily. The hypothesis that the teneurin-TCAP system is an ancient system that arose prior to the emergence of the Metazoa as a result of a horizontal gene transfer (HGT) event from a prokaryote to a choanoflagellate ancestor has previously been raised (31–33). However, the TCAP family has not been previously examined. Thus, TCAP may be associated with an early evolving lineage of peptides that is a sister lineage to the CRF, calcitonin, and secretin families of peptides (11, 34). We therefore examined the phylogenetic relationships of these peptides using TCAP as an outgroup.

# MATERIALS AND METHODS

## Collection of Sequences

Peptide sequences of Secretin GPCR ligands, including CRF, calcitonin and secretin families, and Adhesion GPCR ligands, including TCAP 1–4, as well as reference groups including neuropeptide Y (NPY) and insulin were collected among a range of extant protostomes and deuterostomes, using the GenBank genome sequence analysis program on the NCBI website. The peptides were organized by organism, phylum, class, and order and were tabulated and their accession numbers were recorded (**Table 1**). Sequences were divided into pre-propeptides (or propeptides for TCAP) and mature peptides, after which were imported to MEGA 6.0 for analysis (38). Downloaded from http://www.megasoftware.net/.

## Sequence Alignments

Peptide sequences were aligned using the MUSCLE algorithm (39). The alignment was examined, reviewed for duplicate sequences using pairwise distances ($d = 0.0$ was identical) and excess sequence was cut at both 5′ and 3′ ends, as these fragments did not contribute to the alignment. Modifications to the alignment were made to ensure that the characteristic residue motifs were conserved. This included highly conserved cysteine (C), tryptophan (W), arginine (R), and lysine (K) residues throughout as well as motifs characteristic of each family. For the CRF family this was the 5′ leucine (L), serine (S), and the 3′ asparagine (N) motif that is conserved throughout the entire family, the "TCV" or "TCXV" motif that is conserved among the calcitonin family and the "PELAD" motif that is conserved among the TCAP family.

## Phylogenetic Analysis

Phylogenetic tree construction and statistical analyses were carried out in MEGA 6.0 (38). A multi-step approach was undertaken in order to understand the relationship of each family relative to TCAP prior to conducting a comprehensive analysis of all of the families.

### Maximum Likelihood (ML) Method

The amino acid substitution model and the rate among sites were both chosen based on the model that resulted in the greatest log likelihood, the lowest Akaike Information Criterion (AIC) and the lowest Bayesian Information Criterion (BIC), parameters calculated by MEGA 6.0. To ensure the most accurate analysis, these parameters were calculated for each constructed tree. The model that maximized the log likelihood was used for analysis. A partial deletion of sequences with too many gaps/missing data was applied with a cutoff of 95%, so sites that were not found in at least 95% of sequences were not used toward the analysis. The applied heuristic method was Nearest-Neighbor Interchange (NNI), so the initial trees were obtained using the NJ method to a matrix of pairwise distances estimated using a JTT model. Reliability of the tree was tested using 1,000 bootstrap replicates.

### Pre-propeptide and Mature Peptide Analysis

Two sets of analyses were performed. The first involved Secretin superfamily pre-propeptides, which are composed of a signal, cryptic, and mature peptide and TCAP propeptides, as TCAP does not possess a signal peptide. Given the functional importance, bioactivity, and high level of conservation throughout evolution, a second separate analysis was performed on mature peptides of both Secretin superfamily and TCAP family members.

For analysis involving Secretin superfamily pre-propeptides and TCAP family propeptides, a total of 181 amino acid sequences were used, with a total of 44 positions in the final dataset after all positions with <95% site coverage were eliminated.

### Mature Peptide Analysis

For analysis involving Secretin superfamily mature peptides and TCAP mature peptides, a multi-step analysis was undertaken in order to elucidate the relationships of each family with respect to one another and TCAP. As insulin has a tertiary structure where the peptide folds and the two mature chains are connected by sets of disulfide bonds from the cysteine residues (40), the mature peptide had to be divided into A and B chains for the purpose of this analysis. Due to the high sequence conservation of NPY that may have resulted in the odd placement of the NPY reference group in the pre-propeptide analysis and given that the NPY mature peptide is even so more highly conserved, it was not included as a reference group in the analysis of mature peptides. For insulin and the calcitonin family, analysis involved 72 amino acid sequences, with a total of 14 positions in the final dataset. For insulin, calcitonin, and TCAP, analysis involved 95 amino acid sequences, with a total of 14 positions in the data set. For insulin, calcitonin, CRF and TCAP families, analysis involved 135 amino acid sequences leaving 12 positions in the final data set. Lastly, for insulin, calcitonin, CRF, secretin, and TCAP families, analysis included 179 amino acid sequences leaving 15 positions in the final dataset.

# RESULTS

## Sequence Analysis of TCAP Paralogs and Orthologs

TCAP paralogs, those that diverged as a result of a genome duplication event, demonstrated a high degree of conservation (**Figure 1**). When TCAP 1–4 are aligned in mouse, residues Q2, L4, G7, V9, Q10, G11, Y12, G14, V17, V20, E21, Q22, Y23, E25, L26, D28, S29, N32, I33, F35, R37, Q38, and E40 are all conserved among the four paralogs (**Figure 1**). Similarly, TCAP orthologs, those that arose as a result of a species divergence, demonstrate a high degree of conservation among vertebrates (**Figure 2**). When mammalian, bird, amphibian, and fish TCAP 1-4 sequences are aligned residues L3, G7, V9, G11, Y12, G14, L18, Q22, E25, L26, D28, N32, R37 are conserved among TCAP-1 orthologs (**Figure 2A**). Among TCAP-2 orthologs, residues Q2, L3, L4, G7, G11, Y12, E13, G14, Y15, Y16, V17, L18, P19, V20, E21, Q22, Y23, P24, E25, L26, A27, D28, S29, S30, N32, I33, Q34, F35, L36, Q38, N39, E40, M41 are conserved (**Figure 2B**). Among TCAP-3 orthologs, Q2, L3, L4, S5, K8, V9, G11, Y12, D13, G14, Y15, V17, L18, S19, V20, E21, Q22, Y23, E25, L26, D28, S29, N32, F35, R37, Q38, E40, I41 are conserved (**Figure 2C**). Lastly, among TCAP-4


**FIGURE 1 |** Multiple sequence alignment of the TCAP family of peptides in mouse. The mature peptide sequences were aligned using MUSCLE (MUltiple Sequence Comparison by Log-Expectation). Dark gray boxes indicate amino acid identity and light gray boxes indicate a functional replacement.

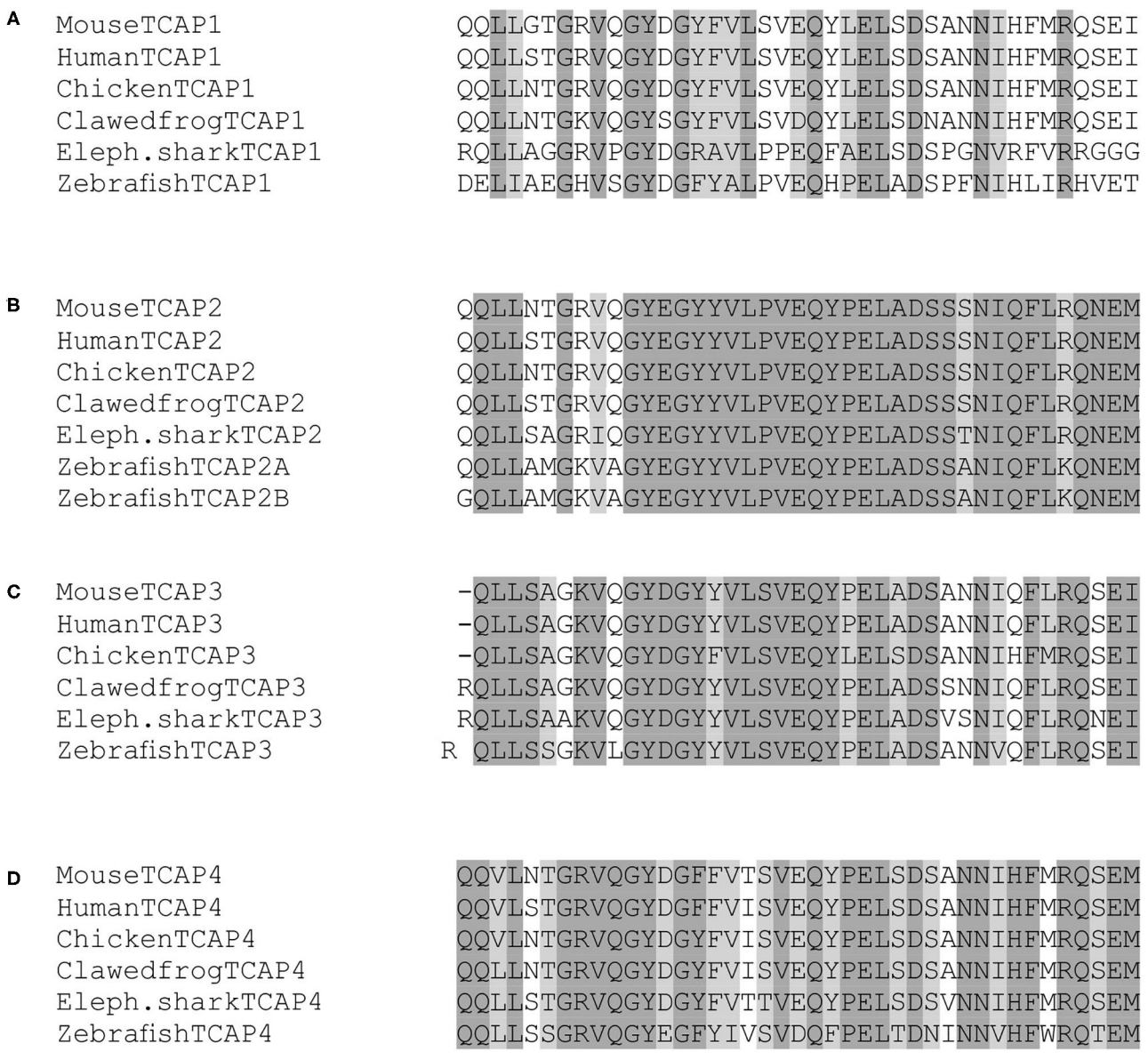


**FIGURE 2 |** Multiple sequence alignment TCAP family members among various species. **(A)** TCAP-1; **(B)** TCAP-2; **(C)** TCAP-3; **(D)** TCAP-4. The mature peptide sequences were aligned using MUSCLE (Multiple Sequence Comparison by Log-Expectation). Dark gray boxes indicate amino acid identity and light gray boxes indicate a functional replacement.

orthologs, Q1, Q2, L4, G7, R8, V9, Q10, G11, Y12, G14, F15, V20, Q22, P24, E25, L26, D28, N31, N32, H34, F35, R37, Q38, E40, M41. Overall, TCAP-2 orthologs (**Figure 2B**) are the most highly conserved and TCAP-1 orthologs (**Figure 2A**) are the least highly

conserved. Also, a characteristic "PELAD" motif at positions 24–28 from the N-terminus is conserved among the TCAP paralogs and orthologs. The high level of conservation of the "PELAD" motif suggests that it possesses a functional attribute, such as a receptor binding or activation site (27). This family of peptides contains the "PELAD" motif at residues 24–28 from the N-terminus. Therefore, both TCAP orthologs and paralogs demonstrate a high degree of conservation among vertebrates.

### Evolutionary Analysis of Pre-propeptides and Mature Peptides of Secretin Superfamily and TCAP Family Members

Phylogenetic analysis of CRF, calcitonin, and secretin pre-propeptide families and TCAP family propeptides revealed that each family formed a distinct group. TCAP, CRF, and secretin families form distinct clades and insulin forms a sister group with the calcitonin family (**Figure 3**). Also, CRF and calcitonin are closely related sister lineages and they, in turn, form a sister lineage to the secretin family. TCAP, the putative progenitor, is most distantly related to these families relative to their relationships to one another.

A separate analysis was performed with mature peptide sequences of the Secretin superfamily and TCAP mature peptides due to their high conservation and functional importance throughout evolution. Phylogenetic analysis of calcitonin mature peptides, insulin A and B mature chains and TCAP demonstrated that calcitonin and insulin families are sister lineages (**Figure 4**). Insulin A chains were more closely related to the calcitonin family than insulin B chains (**Figure 4**). Phylogenetic analysis of calcitonin, insulin A and B chains, CRF, and TCAP mature peptides confirmed that calcitonin and insulin families were sister lineages and that CRF formed a separate group to these two families (**Figure 5**). Lastly, phylogenetic analysis of calcitonin, insulin A and B chains, CRF, secretin, and TCAP mature peptides revealed that calcitonin and insulin families were sister lineages and that both CRF and secretin formed separate groups from these two families (**Figure 6**). Therefore, the multi-step mature peptide analysis confirmed that insulin and calcitonin are sister lineages, that form distinct groups from CRF and secretin families and in turn, that the TCAP family is a distinct clade from Secretin superfamily members.

## DISCUSSION

In this study, the TCAP family is presented as a putative progenitor of the Secretin superfamily of ligands for the first time. The evolutionary relationships among the receptors of these peptides are well-established (18, 19). However, the relationships among members of the Secretin superfamily of ligands as well as a progenitor for this family of peptides have not been elucidated. We considered TCAP as a putative progenitor of the Secretin superfamily for the following reasons. First, evolutionary relationships among the receptors of these ligands demonstrate that Secretin GPCRs derived from Adhesion GPCRs (19) and as TCAP-1 binds to LPHN, an Adhesion GPCR with a HBD characteristic of Secretin GPCRs (17). It is possible that a similar course of evolution occurred for the ligands. Second, the sequence similarity that TCAP shares with CRF and calcitonin (27), both Secretin superfamily members whose receptors are the most closely related to Adhesion GPCRs, suggests that these peptides may have evolved from TCAP, a candidate progenitor peptide.

The teneurin-TCAP system is well-established as being evolutionarily ancient. Evidence suggests that this system arose before the Metazoa evolved about 1 billion years ago and prior to the emergence of the Secretin superfamily that arose around the time of the protostome-deuterostome divergence, about 600 million years ago. As a result, although the TCAP sequence shows some amino acid similarity with the Secretin superfamily, there are a number of differences indicating that the two lineages are evolutionarily divergent. Indeed, we could not determine any significant binding or activation capacity of TCAP with any members of the Secretin GPCRs [(11, 34); Lovejoy, unpublished observations]. In contrast, TCAP binds to the latrophilin HBD and activates this receptor [(30); Reid et al., submitted]. As proposed by Zhang et al. (33), the teneurin-TCAP system likely evolved from a polymorphic proteinaceous toxin (PPT) gene that arose as a result of a HGT event from a prokaryote to a choanoflagellate, a primitive unicellular organism. Importantly, the teneurin gene has been identified in the choanoflagellate, *Monosiga brevicollis* (32). Choanoflagellates are thought to be a progenitor to the Metazoans (42). This supports the hypothesis that a choanoflagellate may have engulfed a prokaryote containing the PPT gene, which became integrated into its genome and lost its toxic role over time (32, 33). With respect to structural evidence, the teneurins share characteristics of PPTs: the same type II orientation, rearrangement hotspot (RHS) domains and close similarity to the C-terminal domain to the histidine-asparagine-histidine (HNH) bacterial toxin of the glycine-histidine-histidine (GHH) clade (33, 43). The GHH domain may be an ancestor of TCAP that lost its toxic role and functioned as an intracellular signaling molecule (33). Additionally, the C-terminal region of the *M. brevicollis* teneurin protein contains tyrosine-aspartate (YD) repeats characteristic of proteobacteria and most of the extracellular domain is encoded on one large 6,829 base pair exon characteristic of prokaryotic genomes and of HGT (32). Therefore, evidence suggests that the teneurin-TCAP system is ancient as it evolved as a result of a HGT event prior to the emergence of the Metazoa.

Moreover, with respect to the course of evolution of the receptors, evidence demonstrates that Adhesion GPCRs evolved prior to Secretin GPCRs and that Secretin GPCRs are derived from Adhesion GPCRs. Adhesion GPCR genes have been identified in the genome of amphioxus, *Branchiostoma floridae*, the choanoflagellate, *M. brevicollis,* and the sea anemone, *Nematostella vectensis* (18), meaning that these lineages were present prior to the protostome-deuterostome divergence. On the other hand, Secretin GPCRs have not been identified in these species and therefore, receptor lineages of the Secretin superfamily likely expanded and radiated around the time of the bifurcation of protostomes and deuterostomes. Also, Nordström et al. (18) demonstrated Secretin GPCRs evolved from Adhesion

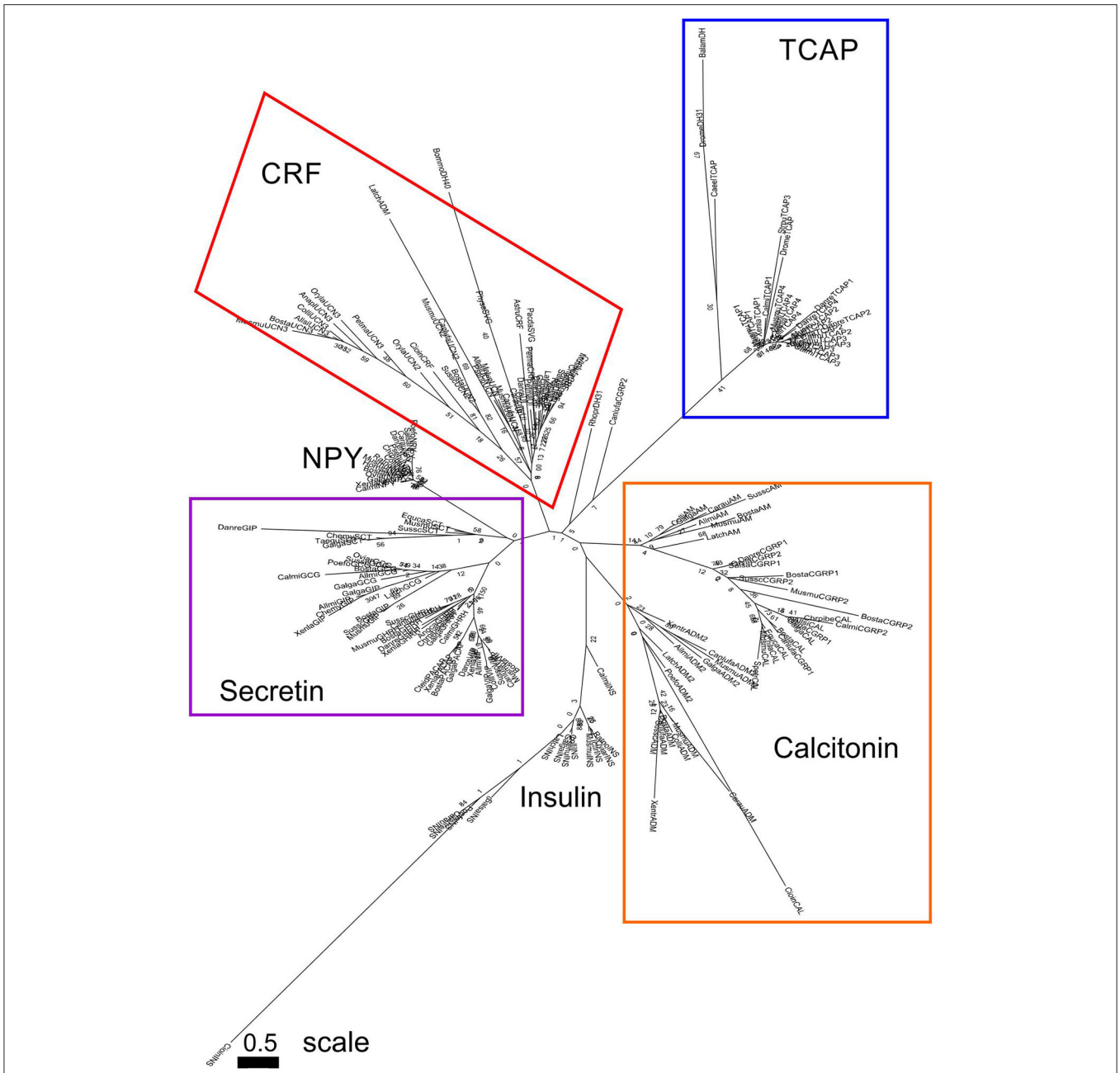


**FIGURE 3 |** Phylogenetic analysis of CRF, calcitonin, insulin, and secretin family pre-propeptides with TCAP propeptides (rooted to TCAP). Each family is highlighted with a different color: CRF (red), calcitonin (orange), secretin (purple), TCAP (blue). Analysis was conducted using the maximum likelihood method based on the JTT+G matrix-based model (lnL = −11224.5064; +G, parameter = 1.3976) (41). Initial trees for the heuristic search were obtained by applying the NJ method to a matrix of pairwise distances estimated using a JTT model. Branch lengths represent the number of substitutions per site, with the tree shown to scale. Bootstrap analysis involved 1,000 replicates. CRF family: CRF, corticotropin-releasing factor; TCN, teleocortin; UCN, urocortin; UCN2, urocortin 2; UCN3, urocortin 3; UI, urotensin; SVG, sauvagine; DH, diuretic hormone; Calcitonin family: CALC, calcitonin; CGRP1, calcitonin-gene-related peptide 1; CGRP2, calcitonin-gene-related peptide 2; AM, amylin; ADM, adrenomedullin; ADM2, adrenomedullin 2; Secretin family: SCT, secretin; GHRH, growth hormone releasing hormone; GIP, gastric inhibitory peptide; GCG, glucagon; PACAP, pituitary adenylate cyclase-activating peptide; VIP, vasoactive intestinal peptide; Reference groups: NPY, neuropeptide Y; INS, insulin; Outgroup: TCAP, teneurin C-terminal associated peptide. The scale bar indicates the level of magnification for the tree.

GCPRs using phylogenetic analysis. Therefore, evidence that the teneurin-TCAP system arose prior to the emergence of the Metazoa as well as the characterization of Adhesion GPCRs but not Secretin GPCRs prior to the protostome-deuterostome divergence suggests that the teneurin-TCAP system predates members of the Secretin superfamily. We suggest that if the

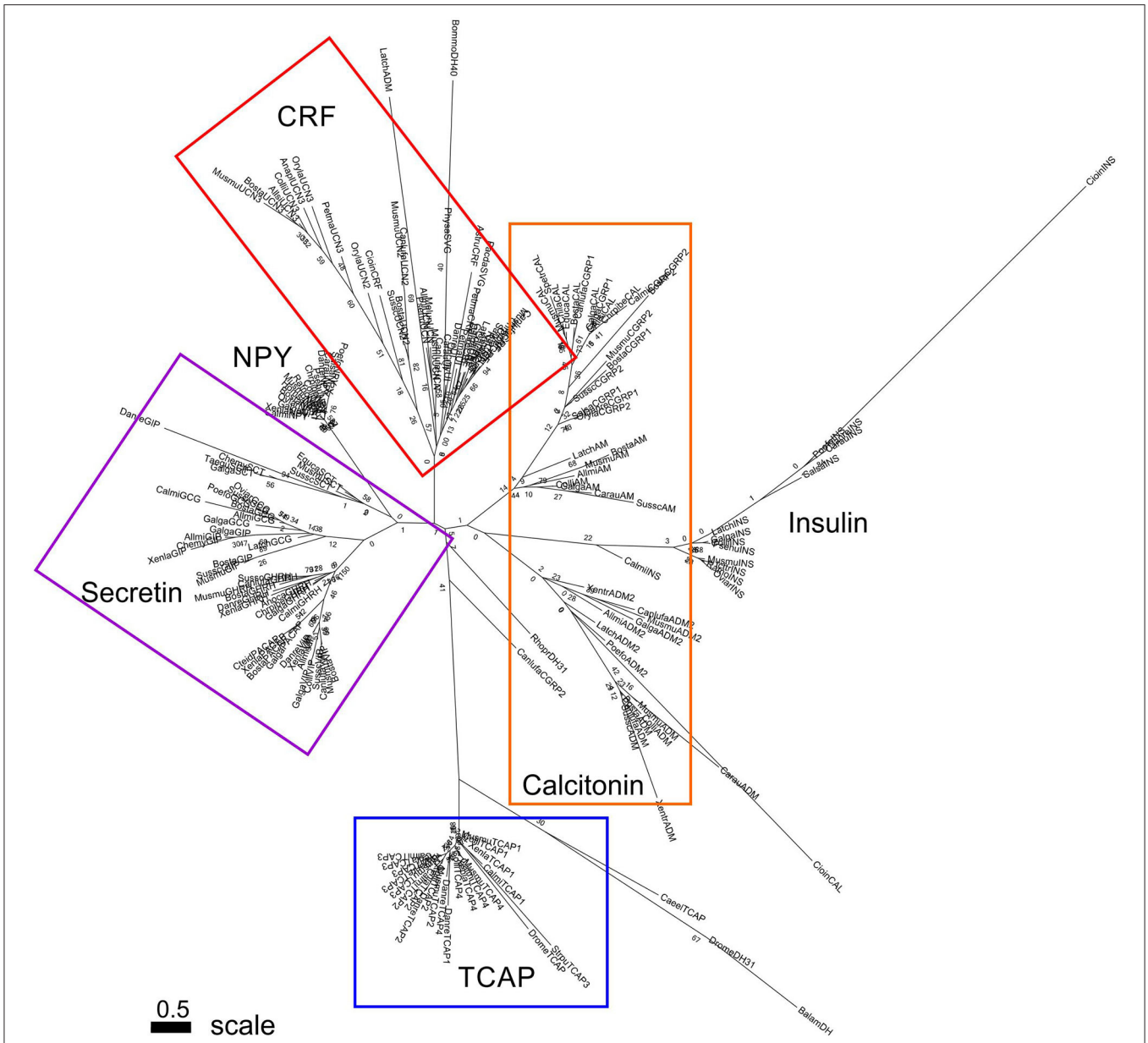


**FIGURE 4 |** Phylogenetic analysis of insulin, calcitonin and TCAP mature peptides (rooted to TCAP). Each family is highlighted with a different color: calcitonin (orange), insulin (green), TCAP (blue). Analysis was conducted using the maximum likelihood method based on the JTT matrix-based model (lnL = −919.2846; +G, parameter = 6.6766) (41) Initial trees for the heuristic search were obtained by applying the NJ method to a matrix of pairwise distances estimated using a JTT model. Branch lengths represent the number of substitutions per site, with the tree shown to scale. Bootstrap analysis involved 1,000 replicates. Calcitonin family: CALC, calcitonin; CGRP1, calcitonin-gene-related peptide 1; CGRP2, calcitonin-gene-related peptide 2; AM, amylin; ADM, adrenomedullin; ADM2, adrenomedullin 2; Insulin: INSa, insulin A chain; INSb, insulin B chain; Outgroup: TCAP, teneurin C-terminal associated peptide. The scale bar indicates the level of magnification for the tree.

ligands for these receptors underwent a similar course in evolution, the TCAP family may be a putative progenitor to the Secretin superfamily.

In light of the evidence to suggest that the teneurin-TCAP system evolved prior to the emergence of the Metazoa, the previously established relationship that Secretin GPCRs derived from Adhesion GPCRs [(Nordstom et al., 2009); (19)], the evidence that TCAP binds to LPHN, an Adhesion GPCR with a HBD characteristic of Secretin GPCRs (17) and given the sequence similarity that TCAP shares with Secretin superfamily members, CRF, and calcitonin (27), a phylogenetic investigation using TCAP as a putative progenitor of the Secretin superfamily was undertaken. A putative progenitor of the Secretin superfamily of ligands has not been previously established. Sequence analysis of TCAP family members demonstrated a highly conserved peptide and

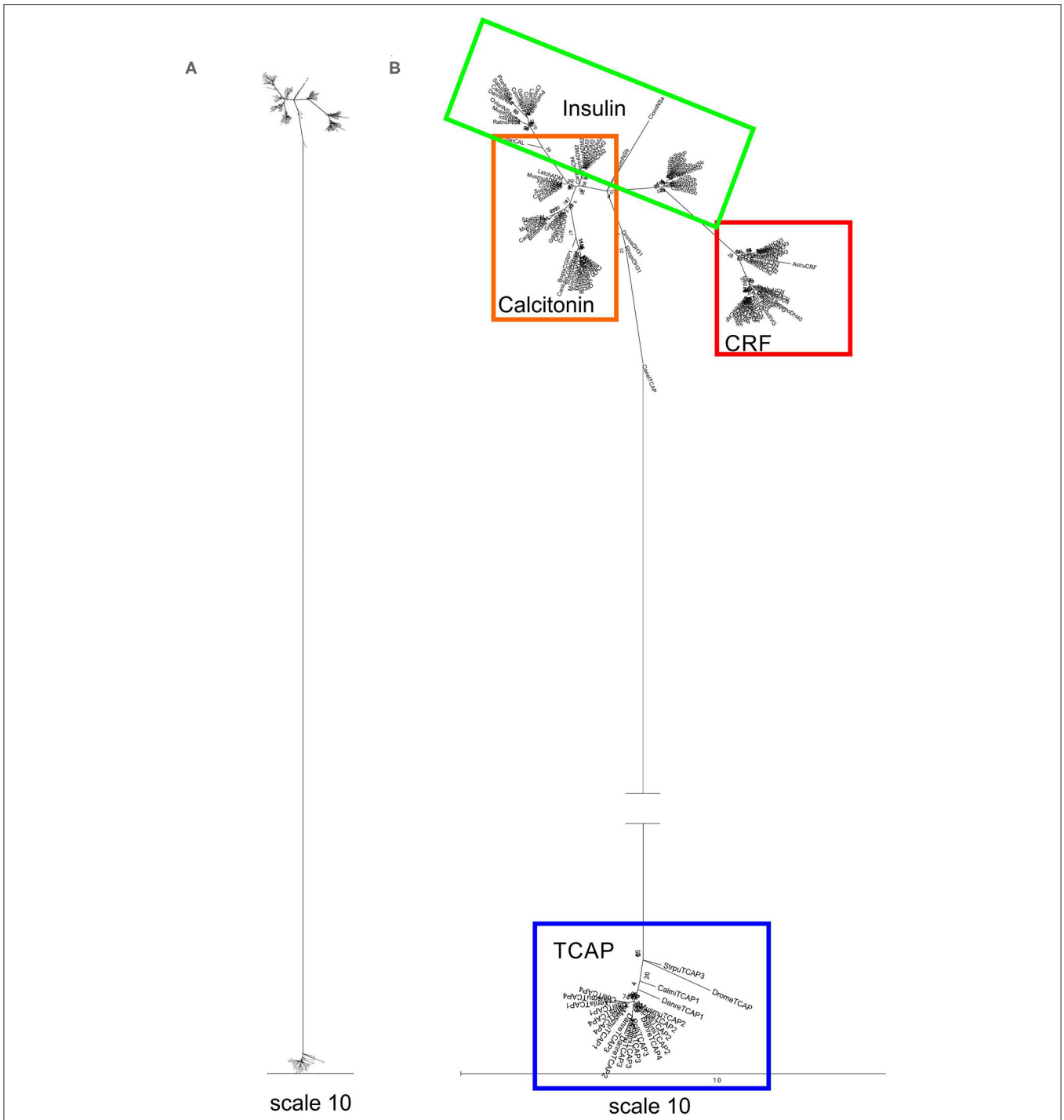


**FIGURE 5 |** Phylogenetic analysis of insulin, calcitonin, CRF and TCAP mature peptides (rooted to TCAP). The trees are represented as **(A)** original tree with the appropriate scale **(B)** magnified and rooted to TCAP. Each family is highlighted with a different color: CRF (red), calcitonin (orange), insulin (green), TCAP (blue). Analysis was conducted using the maximum likelihood method based on the Dayhoff matrix-based model (lnL = −1019.5552; +G, parameter = 6.6766) (41). Initial trees for the heuristic search were obtained by applying the NJ method to a matrix of pairwise distances estimated using a JTT model. Branch lengths represent the number of substitutions per site, with the tree shown to scale. Bootstrap analysis involved 1,000 replicates. Calcitonin family: CALC, calcitonin; CGRP1, calcitonin-gene-related peptide 1; CGRP2, calcitonin-gene-related peptide 2; AM, amylin; ADM, adrenomedullin; ADM2, adrenomedullin 2; Insulin: INSa, insulin A chain; INSb, insulin B chain; CRF family: CRF, corticotropin-releasing factor; TCN, teleocortin; UCN, urocortin; UCN2, urocortin 2; UCN3, urocortin 3; UI, urotensin; SVG, sauvagine; DH, diuretic hormone; Outgroup: TCAP, teneurin C-terminal associated peptide. The scale bar indicates the level of magnification for the tree.

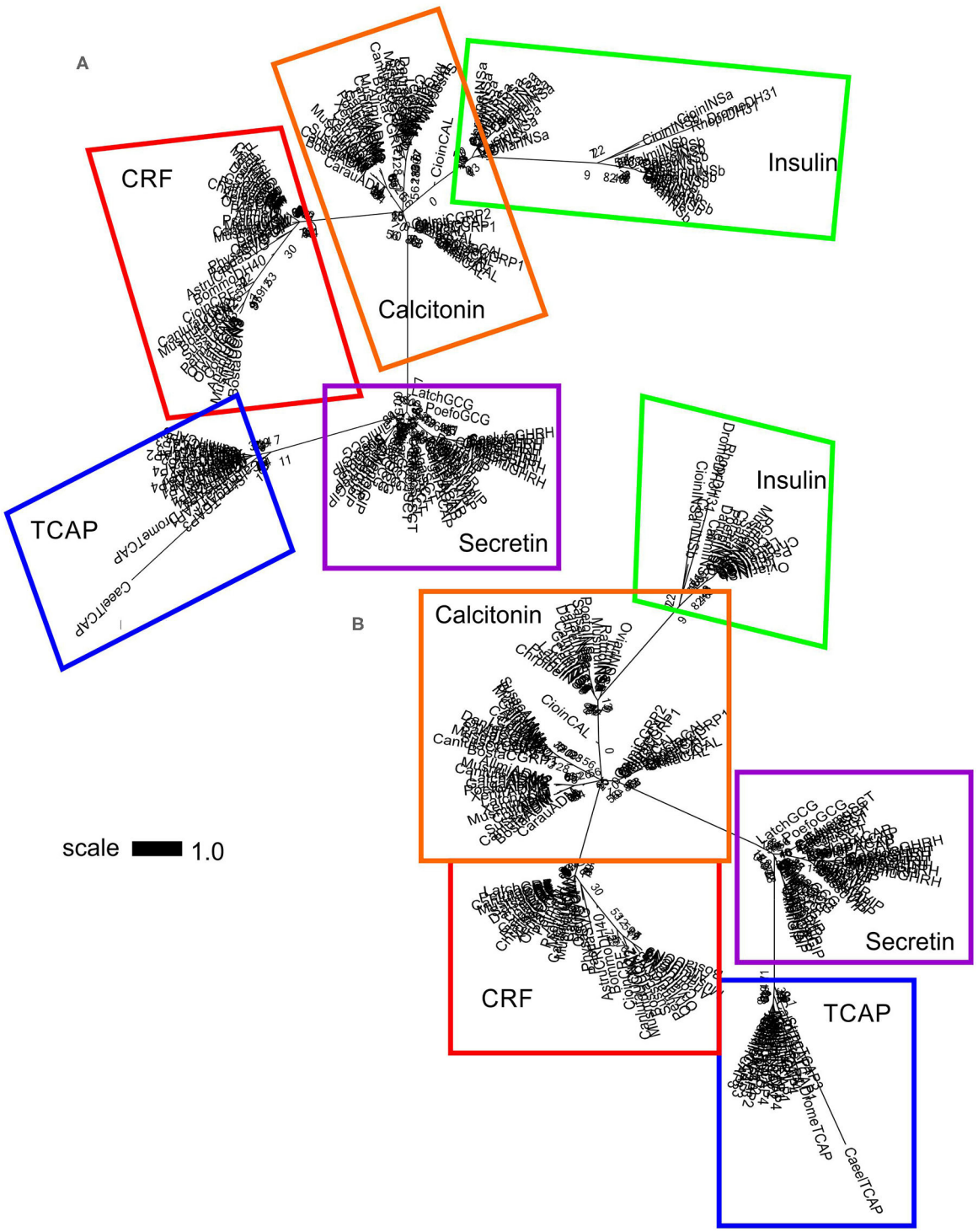


**FIGURE 6 |** Phylogenetic analysis of insulin, calcitonin, CRF, secretin, and TCAP mature peptides. The trees are represented as **(A)** unrooted and **(B)** rooted to TCAP. Each family is highlighted with a different color: CRF (red), calcitonin (orange), insulin (green), secretin (purple), and TCAP (blue). Analysis was conducted using the maximum likelihood method based on the Whelan and Goldman model (lnL = −1781.0007; +G, parameter = 23.5912) (41). Initial trees for the heuristic search were obtained by applying the NJ method to a matrix of pairwise distances estimated using a JTT model. Branch lengths represent the number of substitutions per site, with the tree shown to scale. Bootstrap analysis involved 1,000 replicates. Calcitonin family: CALC, calcitonin; CGRP1, calcitonin-gene-related peptide 1; CGRP2, calcitonin-gene-related peptide 2; AM, amylin; ADM, adrenomedullin; ADM2, adrenomedullin 2; INSa, insulin A chain; INSb, insulin B chain; CRF family: CRF, corticotropin-releasing factor; TCN, teleocortin; UCN, urocortin; UCN2, urocortin 2; UCN3, urocortin 3; UI, urotensin; SVG, sauvagine; DH, diuretic hormone; Secretin family: SCT, secretin; GHRH, growth hormone releasing hormone; GIP, gastric inhibitory peptide; GCG, glucagon; PACAP, pituitary adenylate cyclase-activating peptide; VIP, vasoactive intestinal peptide; Outgroup: TCAP, teneurin C-terminal associated peptide. The scale bar indicates the level of magnification for the tree.

**TABLE 1 |** List of Sequences used to construct phylogenetic trees.

| Hormone | Organism | Phylum | Class | Order | Accession number |
|---|---|---|---|---|---|
| **CRF FAMILY** | | | | | |
| CRF | *Xenopus laevis* (African clawed frog) | Chordata | Amphibia | Anura | NM_001172210 |
| CRF | Rana sylvatica (Wood frog) | Chordata | Amphibia | Anura | HQ630608 |
| CRF | *Gallus gallus* (Chicken) | Chordata | Aves | Galliformes | NM_001123031 |
| CRF | *Columba livia* (Rock pigeon) | Chordata | Aves | Columbiformes | XM_005506466 |
| CRF | *Danio rerio* (Zebrafish) | Chordata | Actinopterygii | Cypriniformes | JN859047 |
| CRF | *Salmo salar* (Atlantic salmon) | Chordata | Actinopterygii | Salmoniformes | NM_001141590 |
| CRF | *Mus musculus* (House mouse) | Chordata | Mammalia | Rodentia | NM_205769 XM_357335 |
| CRF | *Canis lupus familiaris* (Dog) | Chordata | Mammalia | Carnivora | NM_001014278 XM_544106 |
| CRF | *Petromyzon marinus* (Sea lamprey) | Chordata | Cephalaspidomorphi | Petromyzontiformes | Endsin et al. (35) |
| CRF | *Latimeria chalumnae* (Coelacanth) | Chordata | Sarcopterygii | Coelacanthiformes | XM_006009123 |
| CRF | *Chrysemys picta bellii* (Western painted turtle) | Chordata | Reptilia | Testudines | XM_005288085 |
| CRF | *Asterias rubens* (Common starfish) | Echinodermata | Asteroidea | Forcipulatida | Semmens et al. (36) |
| CRF | *Ciona intestinalis* (Sea tunicate) | Urochordata | Ascidiacea | Enterogona | Lovejoy and Barsyte-Lovejoy, (37) |
| TCN | *Oryzias latipes* (Japanese rice fish) | Chordata | Actinopterygii | Beloniformes | BAR90710.1 |
| SVG | *Pachymedusa dacnicolor* (Mexican tree frog) | Chordata | Amphibia | Anura | FR846380 |
| SVG | *Phyllomedusa sauvagii* | Chordata | Amphibia | Anura | AAY21509.1 |
| UCN | *Pseudopodoces humilis* (Tibetan ground-tit) | Chordata | Aves | Passeriformes | XM_005532743 |
| UCN | *Melopsittacus undulatus* (Budgerigar) | Chordata | Aves | Psittaciformes | XM_005140144 |
| UCN | *Mus musculus* (House mouse) | Chordata | Mammalia | Rodentia | NM_021290 |
| UCN | *Canis lupus familiaris* (Dog) | Chordata | Mammalia | Carnivora | XM_848667 |
| UCN | *Alligator mississippiensis* (American alligator) | Chordata | Reptilia | Crocodylia | XM_006273603 |
| UCN2 | *Sus scrofa* (Pig) | Chordata | Mammalia | Artiodactyla | XM_003132185 |
| UCN2 | *Bos taurus* (Cattle) | Chordata | Mammalia | Artiodactyla | XM_005223016 |
| UCN2 | *Mus musculus* (House mouse) | Chordata | Mammalia | Rodentia | NM_145077 XM_910603 |
| UCN2 | *Canis lupus familiaris* (Dog) | Chordata | Mammalia | Carnivora | XM_005632587 |
| UCN2 | *Oryzias latipes* (Japanese rice fish) | Chordata | Actinopterygii | Beloniformes | NP_001121991.1 |
| UCN3 | *Anas platyrhynchos* (Mallard) | Chordata | Aves | Anseriformes | XM_005014551 |
| UCN3 | *Mus musculus* (House mouse) | Chordata | Mammalia | Rodentia | NM_031250 |
| UCN3 | *Bos taurus* (Cattle) | Chordata | Mammalia | Artiodactyla | NM_001076527 |
| UCN3 | *Alligator sinensis* (Chinese alligator) | Chordata | Reptilia | Crocodylia | XM_006035307 |
| UCN3 | *Petromyzon marinus* (Sea lamprey) | Chordata | Cephalaspidomorphi | Petromyzontiformes | Endsin et al. (35) |
| UCN3 | *Oryzias latipes* (Japanese rice fish) | Chordata | Actinopterygii | Beloniformes | NP_001121992.1 |
| UI | *Oncorhynchus mykiss* (Rainbow trout) | Chordata | Actinopterygii | Salmoniformes | NM_001124343 |
| UI | *Danio rerio* (Zebrafish) | Chordata | Actinopterygii | Cypriniformes | NM_001030180 XM_687090 |
| UI | *Carassius auratus* (Goldfish) | Chordata | Actinopterygii | Cypriniformes | AF129115 |
| UI | *Petromyzon marinus* (Sea lamprey) | Chordata | Cephalaspidomorphi | Petromyzontiformes | Endsin et al. (35) |
| DH 31 | *Drosophila melanogaster* (Fruit fly) | Arthropoda | Insecta | Diptera | NM_078790 |
| DH 40 | *Bombyx mori* (Domestoc silkworm) | Arthropoda | Insecta | Lepidoptera | AB298934 |
| DH 31 | *Rhodnius prolixus* (Assassin bug) | Arthropoda | Insecta | Hemiptera | HM030716 |
| DH | *Balanus amphitrite* (Barnacle) | Arthropoda | Maxillopoda | Sessilia | JQ864196 |
| **CALCITONIN FAMILY** | | | | | |
| CALC | *Gallus gallus* (Chicken) | Chordata | Aves | Galliformes | X03012 |
| CALC | *Bos taurus* (Cattle) | Chordata | Mammalia | Artiodactyla | AB462435 |
| CALC | *Mus musculus* (House mouse) | Chordata | Mammalia | Rodentia | X97991 |
| CALC | *Alligator sinensis* (Chinese alligator) | Chordata | Reptilia | Crocodylia | XM_006018232 |
| CALC | *Chrysemys picta bellii* (Western painted turtle) | Chordata | Reptilia | Testudines | XM_005303304 |
| CALC | *Ciona intestinalis* (Sea tunicate) | Urochordata | Ascidiacea | Enterogona | AB485672 |
| CALC | *Equus caballus* (Horse) | Chordata | Mammalia | Perissodactyla | AF249307 |
| CALC | *Chinchilla lanigera* (Long-tailed chinchilla) | Chordata | Mammalia | Rodentia | XP_005380394.1 |

*(Continued)*

**TABLE 1 |** Continued

| Hormone | Organism | Phylum | Class | Order | Accession number |
|---|---|---|---|---|---|
| CALC | Spermophilus tridecemlineatus (Squirrel) | Chordata | Mammalia | Rodentia | XM_005326775 |
| CGRP1 | *Gallus gallus* (Chicken) | Chordata | Aves | Galliformes | NM_001113708 |
| CGRP1 | *Danio rerio* (Zebrafish) | Chordata | Actinopterygii | Cypriniformes | NM_001002471 |
| CGRP1 | *Salmo salar* (Atlantic salmon) | Chordata | Actinopterygii | Salmoniformes | NM_001146580 |
| CGRP1 | *Bos taurus* (Cattle) | Chordata | Mammalia | Artiodactyla | NM_001076340 |
| CGRP1 | *Canis lupus familiaris* (Dog) | Chordata | Mammalia | Carnivora | NM_001003266 |
| CGRP2 | *Callorhinchus milii* (Elelphant shark) | Chordata | Chondrichthyes | Chimaeriformes | XM_007887527 |
| CGRP2 | *Bos taurus* (Cattle) | Chordata | Mammalia | Artiodactyla | NM_001134662 |
| CGRP2 | *Canis lupus familiaris* (Dog) | Chordata | Mammalia | Carnivora | NM_001002948 |
| CGRP2 | *Mus musculus* (House mouse) | Chordata | Mammalia | Rodentia | NM_054084 |
| CGRP2 | *Sus scrofa* (Pig) | Chordata | Mammalia | Artiodactyla | NM_001102473 |
| CGRP2 | *Oryzias latipes* (Japanese rice fish) | Chordata | Actinopterygii | Beloniformes | NM_001104894 |
| Amylin | *Columba livia* (Rock pigeon) | Chordata | Aves | Columbiformes | XM_005504095 |
| Amylin | *Gallus gallus* (Chicken) | Chordata | Aves | Galliformes | NM_205397 |
| Amylin | *Carassius auratus* (Goldfish) | Chordata | Actinopterygii | Cypriniformes | EU000530 |
| Amylin | *Latimeria chalumnae* (Coelacanth) | Chordata | Sarcopterygii | Coelacanthiformes | XM_005998862 |
| Amylin | *Bos taurus* (Cattle) | Chordata | Mammalia | Artiodactyla | NM_001195038 |
| Amylin | *Mus musculus* (House mouse) | Chordata | Mammalia | Rodentia | NM_010491 |
| Amylin | *Sus scrofa* (Pig) | Chordata | Mammalia | Artiodactyla | XM_003126437 |
| Amylin | *Alligator mississippiensis* (American alligator) | Chordata | Reptilia | Crocodylia | XM_006270879 |
| ADM | *Xenopus tropicalis* (Western clawed frog) | Chordata | Amphibia | Anura | XM_002936741 |
| ADM | *Columba livia* (Rock pigeon) | Chordata | Aves | Columbiformes | XM_005499377 |
| ADM | *Carassius auratus* (Goldfish) | Chordata | Actinopterygii | Cypriniformes | EU000533 |
| ADM | *Latimeria chalumnae* (Coelacanth) | Chordata | Sarcopterygii | Coelacanthiformes | XM_005989689 |
| ADM | *Sus scrofa* (Pig) | Chordata | Mammalia | Artiodactyla | D14875 |
| ADM | *Bos taurus* (Cattle) | Chordata | Mammalia | Artiodactyla | NM_173888 |
| ADM | *Canis lupus familiaris* (Dog) | Chordata | Mammalia | Carnivora | AB191461 |
| ADM | *Mus musculus* (House mouse) | Chordata | Mammalia | Rodentia | NM_009627 |
| ADM2 | *Xenopus tropicalis* (Western clawed frog) | Chordata | Amphibia | Anura | XM_002939371 |
| ADM2 | *Gallus gallus* (Chicken) | Chordata | Aves | Galliformes | XM_004937395 |
| ADM2 | *Latimeria chalumnae* (Coelacanth) | Chordata | Sarcopterygii | Coelacanthiformes | XM_006013419 |
| ADM2 | *Poecilia formosa* (Amazon molly) | Chordata | Actinopterygii | Cyprinodontiformes | XM_007545759 |
| ADM2 | *Canis lupus familiaris* (Dog) | Chordata | Mammalia | Carnivora | XM_843399 |
| ADM2 | *Mus musculus* (House mouse) | Chordata | Mammalia | Rodentia | AB121035 |
| ADM2 | *Alligator mississippiensis* (American alligator) | Chordata | Reptilia | Crocodylia | XM_006275255 |
| **SECRETIN FAMILY** | | | | | |
| GHRH | *Xenopus laevis* (African clawed frog) | Chordata | Amphibia | Anura | NM_001096728 |
| GHRH | *Gallus gallus* (Chicken) | Chordata | Aves | Galliformes | NM_001040464 |
| GHRH | *Callorhinchus milii* (Elelphant shark) | Chordata | Chondrichthyes | Chimaeriformes | XM_007885752 |
| GHRH | *Danio rerio* (Zebrafish) | Chordata | Actinopterygii | Cypriniformes | NM_001080092 |
| GHRH | *Bos taurus* (Cattle) | Chordata | Mammalia | Artiodactyla | NM_178325 |
| GHRH | *Canis lupus familiaris* (Dog) | Chordata | Mammalia | Carnivora | NM_001290112 |
| GHRH | *Mus musculus* (House mouse) | Chordata | Mammalia | Rodentia | NM_010285 |
| GHRH | *Sus scrofa* (Pig) | Chordata | Mammalia | Artiodactyla | NM_001195118 |
| GHRH | *Anolis carolinensis* (Green anole) | Chordata | Reptilia | Squamata | XM_003225298 |
| GHRH | *Chrysemys picta bellii* (Western painted turtle) | Chordata | Reptilia | Testudines | XM_005295322 |
| GIP | *Xenopus laevis* (African clawed frog) | Chordata | Amphibia | Anura | NM_001097922 |
| GIP | *Gallus gallus* (Chicken) | Chordata | Aves | Galliformes | NM_001080104 |
| GIP | *Danio rerio* (Zebrafish) | Chordata | Actinopterygii | Cypriniformes | NM_001080059 |
| GIP | *Bos taurus* (Cattle) | Chordata | Mammalia | Artiodactyla | NM_001166605 |
| GIP | *Mus musculus* (House mouse) | Chordata | Mammalia | Rodentia | NM_008119 |

*(Continued)*

**TABLE 1 |** Continued

| Hormone | Organism | Phylum | Class | Order | Accession number |
|---|---|---|---|---|---|
| GIP | *Sus scrofa* (Pig) | Chordata | Mammalia | Artiodactyla | NM_001287408 |
| GIP | *Alligator mississippiensis* (American alligator) | Chordata | Reptilia | Crocodylia | XM_006277905 |
| GIP | *Chelonia mydas* (Green sea turtle) | Chordata | Reptilia | Testudines | XM_007061917 |
| GCG | *Gallus gallus* (Chicken) | Chordata | Aves | Galliformes | NM_205260 |
| GCG | *Callorhinchus milii* (Elelphant shark) | Chordata | Chondrichthyes | Chimaeriformes | XM_007889848 |
| GCG | *Bos taurus* (Cattle) | Chordata | Mammalia | Artiodactyla | NM_173916 |
| GCG | *Sus scrofa* (Pig) | Chordata | Mammalia | Artiodactyla | NM_214324 |
| GCG | *Ovis aries* (Sheep) | Chordata | Mammalia | Cetartiodactyla | XM_004004659 |
| GCG | *Alligator mississippiensis* (American alligator) | Chordata | Reptilia | Crocodylia | XM_006277994 |
| GCG | *Poecilia formosa* (Amazon molly) | Chordata | Actinopterygii | Cyprinodontiformes | XM_007546594 |
| GCG | *Latimeria chalumnae* (Coelacanth) | Chordata | Sarcopterygii | Coelacanthiformes | XM_006004345 |
| PACAP | *Xenopus laevis* (African clawed frog) | Chordata | Amphibia | Anura | AF187877 |
| PACAP | *Gallus gallus* (Chicken) | Chordata | Aves | Galliformes | AY956323 |
| PACAP | *Ctenopharyngodon idella* (Grass carp) | Chordata | Actinopterygii | Cypriniformes | EF592488 |
| PACAP | *Bos taurus* (Cattle) | Chordata | Mammalia | Artiodactyla | AY924308 |
| SCT | *Gallus gallus* (Chicken) | Chordata | Aves | Galliformes | NM_001024833 |
| SCT | *Taeniopygia guttata* (Zebra finch) | Chordata | Aves | Passeriformes | NM_001256233 |
| SCT | *Mus musculus* (House mouse) | Chordata | Mammalia | Rodentia | X73580 |
| SCT | *Sus scrofa* (Pig) | Chordata | Mammalia | Artiodactyla | XM_003122391 |
| SCT | *Equus caballus* (Horse) | Chordata | Mammalia | Perissodactyla | XM_003362642 |
| SCT | *Chelonia mydas* (Green sea turtle) | Chordata | Reptilia | Testudines | XM_007060911 |
| VIP | *Xenopus laevis* (African clawed frog) | Chordata | Amphibia | Anura | NM_001085714 |
| VIP | *Columba livia* (Rock pigeon) | Chordata | Aves | Columbiformes | XM_005507654 |
| VIP | *Gallus gallus* (Chicken) | Chordata | Aves | Galliformes | NM_205366 |
| VIP | *Danio rerio* (Zebrafish) | Chordata | Actinopterygii | Cypriniformes | NM_001114553 |
| VIP | *Bos taurus* (Cattle) | Chordata | Mammalia | Artiodactyla | AF503910 |
| VIP | *Canis lupus familiaris* (Dog) | Chordata | Mammalia | Carnivora | XM_005615524 |
| VIP | *Mus musculus* (House mouse) | Chordata | Mammalia | Rodentia | NM_011702 |
| VIP | *Sus scrofa* (Pig) | Chordata | Mammalia | Artiodactyla | NM_001195233 |
| VIP | *Alligator mississippiensis* (American alligator) | Chordata | Reptilia | Crocodylia | XM_006265239 |
| **OTHER** | | | | | |
| NPY | *Mus musculus* (House mouse) | Chordata | Mammalia | Rodentia | EDK98613.1 |
| NPY | *Danio rerio* (Zebrafish) | Chordata | Actinopterygii | Cypriniformes | AAI62071.1 |
| NPY | *Carassius auratus* (Goldfish) | Chordata | Actinopterygii | Cypriniformes | AAA49186.1 |
| NPY | *Rattus norvegicus* (Norway rat) | Chordata | Mammalia | Rodentia | NP_036746.1 |
| NPY | *Gallus gallus* (Chicken) | Chordata | Aves | Galliformes | NP_990804.1 |
| NPY | *Xenopus laevis* (African clawed frog) | Chordata | Amphibia | Anura | AAH80115.1 |
| NPY | *Ovis aries* (Sheep) | Chordata | Mammalia | Cetartiodactyla | NP_001009452.1 |
| NPY | *Bos taurus* (Cattle) | Chordata | Mammalia | Artiodactyla | ACH61954.1 |
| NPY | *Homo sapiens* (Human) | Chordata | Mammalia | Primate | NP_000896.1 |
| NPY | *Columba livia* (Rock pigeon) | Chordata | Aves | Columbiformes | NP_001269740.1 |
| NPY | *Pseudopodoces humilis* (Ground tit) | Chordata | Aves | Passeriformes | XP_005518939.1 |
| NPY | *Callorhinchus milii* (Elelphant shark) | Chordata | Chondrichthyes | Chimaeriformes | ACF22970.1 |
| NPY | *Chrysemys picta bellii* (Western painted turtle) | Chordata | Reptilia | Testudines | XP_005290923.1 |
| INS | *Ciona intestinalis* (Sea tunicate) | Urochordata | Ascidiacea | Enterogona | NP_001123204.1 |
| INS | *Mus musculus* (House mouse) | Chordata | Mammalia | Rodentia | ABF48502.1 |
| INS | *Rattus norvegicus* (Norway rat) | Chordata | Mammalia | Rodentia | AAA41439.1 |
| INS | *Ictidomys tridecemlineatus* (Ground squirrel) | Chordata | Mammalia | Rodentia | AAK72558.1 |
| INS | *Ovis aries* (Sheep) | Chordata | Mammalia | Cetartiodactyla | AAB60625.1 |
| INS | *Gallus gallus* (Chicken) | Chordata | Aves | Galliformes | NP_990553.1 |

*(Continued)*

**TABLE 1** | Continued

| Hormone | Organism | Phylum | Class | Order | Accession number |
|---|---|---|---|---|---|
| INS | *Columba livia* (Rock pigeon) | Chordata | Aves | Columbiformes | EMC88047.1 |
| INS | *Pseudopodoces humilis* (Ground tit) | Chordata | Aves | Passeriformes | XP_005522396.1 |
| INS | *Callorhinchus milii* (Elelphant shark) | Chordata | Chondrichthyes | Chimaeriformes | XP_007902984.1 |
| INS | *Danio rerio* (Zebrafish) | Chordata | Actinopterygii | Cypriniformes | NP_571131.1 |
| INS | *Carassius auratus* (Goldfish) | Chordata | Actinopterygii | Cypriniformes | ALO24192.1 |
| INS | *Salmo salar* (Atlantic salmon) | Chordata | Actinopterygii | Salmoniformes | ACI69187.1 |
| INS | *Latimeria chalumnae* (Coelacanth) | Chordata | Sarcopterygii | Coelacanthiformes | XP_006008147.1 |
| INS | *Poecilia formosa* (Amazon molly) | Chordata | Actinopterygii | Cyprinodontiformes | XP_016521686.1 |
| INS | *Chrysemys picta bellii* (Western painted turtle) | Chordata | Reptilia | Testudines | XP_005312438.1 |
| TCAP1 | *Xenopus laevis* (African clawed frog) | Chordata | Amphibia | Anura | XP_017951867.1 |
| TCAP1 | *Callorhinchus milii* (Elelphant shark) | Chordata | Chondrichthyes | Chimaeriformes | XP_007893009.1 |
| TCAP | *Caenorhabditis elegans* (Roundworm) | Nematoda | Secernentea | Rhabditida | NM_171175 |
| TCAP | *Drosophila melanogaster* (Fruit fly) | Arthropoda | Insecta | Diptera | NP_001097661 |
| TCAP1 | *Mus musculus* (House mouse) | Chordata | Mammalia | Rodentia | NP_035985.2 |
| TCAP1 | *Danio rerio* (Zebrafish) | Chordata | Actinopterygii | Cypriniformes | XP_691552.5 |
| TCAP1 | *Columba livia* (Rock pigeon) | Chordata | Aves | Columbiformes | EMC88689.1 |
| TCAP2 | *Xenopus laevis* (African clawed frog) | Chordata | Amphibia | Anura | XP_012815129.1 |
| TCAP2 | *Callorhinchus milii* (Elelphant shark) | Chordata | Chondrichthyes | Chimaeriformes | XP_007900206.1 |
| TCAP2 | *Mus musculus* (House mouse) | Chordata | Mammalia | Rodentia | NP_035986.3 |
| TCAP2 | *Danio rerio* (Zebrafish) | Chordata | Actinopterygii | Cypriniformes | XP_017208443.1 |
| TCAP2 | *Columba livia* (Rock pigeon) | Chordata | Aves | Columbiformes | EMC78205.1 |
| TCAP3 | *Xenopus laevis* (African clawed frog) | Chordata | Amphibia | Anura | NP_001096158 |
| TCAP3 | *Callorhinchus milii* (Elelphant shark) | Chordata | Chondrichthyes | Chimaeriformes | XP_007894102.1 |
| TCAP3 | *Strongylocentrotus purpuratus* (Sea urchin) | Echinodermata | Echinoidea | Echinoidea | XM_001180001 |
| TCAP3 | *Mus musculus* (House mouse) | Chordata | Mammalia | Rodentia | NP_035987.3 |
| TCAP3 | *Danio rerio* (Zebrafish) | Chordata | Actinopterygii | Cypriniformes | NP_571043.1 |
| TCAP3 | *Columba livia* (Rock pigeon) | Chordata | Aves | Columbiformes | XP_005505621.1 |
| TCAP4 | *Xenopus laevis* (African clawed frog) | Chordata | Amphibia | Anura | NP_001096158.1 |
| TCAP4 | *Callorhinchus milii* (Elelphant shark) | Chordata | Chondrichthyes | Chimaeriformes | XP_007900970.1 |
| TCAP4 | *Mus musculus* (House mouse) | Chordata | Mammalia | Rodentia | NP_001297689.1 |
| TCAP4 | *Danio rerio* (Zebrafish) | Chordata | Actinopterygii | Cypriniformes | NP_571044.2 |
| TCAP4 | *Columba livia* (Rock pigeon) | Chordata | Aves | Columbiformes | XP_005500626.1 |

phylogenetic analysis of the Secretin superfamily in relation to TCAP as a putative progenitor revealed relationships among Secretin superfamily members. Calcitonin and insulin families are sister lineages and they are much more closely related to one another than was previously thought. Also, calcitonin and insulin are sister lineages that form distinct lineages to CRF and secretin families. Therefore, placing TCAP as an ancestor of the Secretin superfamily allowed a novel interpretation of evolutionary relationships among Secretin superfamily members.

## Sequence Analysis of TCAP Paralogs and Orthologs

Sequence analysis of both TCAP paralogs and orthologs revealed that this family of peptides is highly conserved. The presence of a conserved "PELAD" motif among TCAP orthologs and paralogs, suggests that it may possess a functional attribute, such as a receptor-binding or activation site (27). Also, some characteristic amino acids are retained throughout orthologs and paralogs. Arginine (R) and lysine (K) residues are retained in some parts of the mature peptide and they are often characteristic of the presence of cleavage sites. Glycine (G) and proline (P) are also highly conserved and these amino acids have a tendency to be retained as their secondary structure can break the α-helical structure of peptides. A peptide system with such a large amount of conservation is indicative of great functional importance that may have been selected for. Therefore, the high sequence conservation among TCAP orthologs and paralogs suggests that this peptide system is evolutionarily ancient and may have been strongly selected for throughout evolutionary time.

## Evolutionary Analysis of Pre-propeptides and Mature Peptides of Secretin Superfamily and TCAP Family Members

Phylogenetic analysis of Secretin superfamily pre-propeptides (composed of the signal, cryptic, and mature peptide) and TCAP family pro-peptides (composed of the cryptic and mature peptide) was undertaken in order to elucidate the relationships among these peptides. Analysis revealed that calcitonin, CRF,

secretin, and TCAP families formed distinct groups. Despite being chosen to serve as a reference group because it binds to a tyrosine kinase receptor and not a GPCR, insulin formed a group with calcitonin, suggesting that they may be sister lineages (**Figure 3**). The close relationship between calcitonin and insulin has previously been explored where Wimalawansa (44) suggested that insulin and calcitonin families are closely related. This is supported by phylogenetic analysis of the pre-propeptides and suggests that insulin and calcitonin are sister lineages. When the tree was rooted to TCAP (**Figure 3**), to establish the assumption that TCAP is the ancestor, CRF, calcitonin, and secretin families formed distinct groups. This evolutionary analysis suggests that the secretin family forms a separate clade that is a sister to CRF and calcitonin families, which, in turn, are sisters to one another. This is consistent with what has been observed with respect to Secretin GPCR evolution, where CRF and calcitonin receptors share the greatest amount of sequence similarity among Secretin GPCRs (17). Therefore, it is possible that a similar evolutionary scheme occurred with respect to the ligands. Thus, analysis of Secretin superfamily pre-propeptides with TCAP propeptides suggests that insulin and calcitonin are closely related sister lineages, that calcitonin-insulin and CRF lineages are closely related and that calcitonin-insulin and CRF form a distinct sister lineage to the secretin family.

Subsequently, phylogenetic analysis was performed with the mature peptides of Secretin superfamily members and the TCAP family. The analysis of TCAP family mature peptide sequences with calcitonin and insulin mature sequences (**Figure 4**) demonstrated that insulin A chains were closely related to mature calcitonin peptides. This suggests that the insulin A mature chain is more closely related to the calcitonin family than the insulin B mature chain, which is different from what was previously suggested by Wimalawansa (44). Subsequent analyses involving CRF, calcitonin, insulin, and TCAP mature peptides (**Figure 5**) as well as secretin, CRF, calcitonin, insulin, and TCAP mature peptides (**Figure 6**) confirmed that the insulin A chain was more closely related to the calcitonin family than the insulin B chain. Taken together, insulin and calcitonin are closely related sister groups, which was also observed with the pre-propeptide analysis (**Figure 3**). Moreover, with respect to relationships among Secretin superfamily members, calcitonin-insulin, and CRF families are more closely related to one another than they are to secretin or TCAP, which is supported by the evolutionary scheme of their receptors, which also appear to be very closely related. Finally, secretin forms a sister lineage to a lineage that comprises both calcitonin-insulin and CRF families. This is consistent with what was observed for analysis of the pre-propeptides (**Figure 3**).

Considering the evidence with respect to the ancestral origin of the teneurin-TCAP system and in light of the findings presented here, it is possible to present two hypotheses for the evolutionary scheme of these peptides. The first suggests that an ancient TCAP-like peptide may have been the ancestor of the Secretin superfamily and that it evolved prior to the emergence of CRF, calcitonin, and secretin families. This is supported by the identification of TCAP in organisms prior to the protostome-deuterostome divergence, where as members of the Secretin superfamily have not been identified this early in evolution (31, 32, 34). The possibility of a second hypothesis, suggesting that the Secretin superfamily forms a parallel lineage to extant TCAP and that these two lineages evolved from a proto-CRF-calcitonin-secretin-TCAP ancestor that was related to all of these families, cannot be discounted. Due to sequence availability, phylogenetic analysis was performed using extant Secretin superfamily and TCAP sequences. As a result, both of these hypotheses are plausible. Future analysis should be undertaken in order to further investigate whether TCAP is a progenitor of the Secretin superfamily of ligands.

## CONCLUSIONS

Taken together, phylogenetic analysis of members of the Secretin superfamily using TCAP as a putative progenitor demonstrated relationships among Secretin superfamily members. First, calcitonin formed a closely related sister lineage to insulin, particularly the insulin A chain with respect to mature peptides, but this was also observed with the pre-propeptides. Also, calcitonin-insulin and CRF families are more closely related to one another than they are to secretin or TCAP, which is supported by the evolutionary scheme of their receptors. Finally, secretin forms a sister lineage to a group that comprises both calcitonin-insulin and CRF. Therefore, given evidence that the teneurin-TCAP system arose as a result of a HGT event prior to the emergence of the Metazoa, as well as the previously established structural similarity of TCAP to calcitonin and CRF, members of the Secretin superfamily, the presented phylogenetic analysis allowed for the elucidation of relationships among members of the Secretin superfamily. To conclude, this is the first time that relationships among this family of peptides were resolved and because a progenitor peptide for the Secretin superfamily has not been elucidated, we present TCAP as a candidate progenitor.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## AUTHOR CONTRIBUTIONS

OM performed all analyses and completed the first draft of the paper. BC and NL provided technical guidance on the construction of the phylogenetic tree. DL oversaw the research program and completed the final draft of the manuscript. All authors contributed to the article and approved the submitted version.

## FUNDING

This research was supported by a grant from the Natural Sciences and Engineering Council of Canada (NSERC)-Canada Graduate Scholarship-Master's (CGSM) program in addition to funding from the University of Toronto for OM.

## REFERENCES


1. Vale W, Spiess J, Rivier C, Rivier J. Characterization of a 41-residue ovine hypothalamic peptide that stimulates secretion of corticotropin and β-endorphin. *Science*. (1981) 213:1394–7. doi: 10.1126/science.6267699
2. Monticucchi PA, Neschen A, Erspamer V. Structure of sauvagine, as vasoactive peptide from the skin of a frog. *Hoppe-Seyler's Z Physiol Chem*. (1979) 360:1178.
3. Lederis K, Letter A, McMaster D, Moore G, Schlesinger D. Complete amino acid sequence of urotensin I, a hypotensive and corticotropin-releasing neuropeptide from Catostomus. *Science*. (1982) 218:162–4. doi: 10.1126/science.6981844
4. Vaughan J, Donaldson C, Bittencourt J, Perrin MH, Lewis K, Sutton S, et al. Characterization of urocortin, a novel mammalian neuropeptide related to fish urotensin-I and to CRF. *Nature*. (1995) 378:287–92. doi: 10.1038/378287a0
5. Lovejoy DA, Balment RJ. Evolution and physiology of the corticotropin-releasing factor (CRF) family of neuropeptides in vertebrates. *GenComp Endocrinol*. (1999) 155:1–22. doi: 10.1006/gcen.1999.7298
6. Lewis K, Li C, Perrin MH, Blount A, Kunitake K, Donaldson C, et al. Identification of urocortin III, an additional member of the corticotropin-releasing factor (CRF) family with high affinity for the CRF2 receptor. *Proc Natl Acad Sci USA*. (2001) 98:7570–5. doi: 10.1073/pnas.121165198
7. Reyes TM, Lewis K, Perrin MH, Kunitake KS, Vaughan J, Arias CA, et al. Urocortin II: a member of the corticotropin releasing factor (CRF) neuropeptide family that is selectively bound by type 2 receptors. *Proc Natl Acad Sci USA*. (2001) 98:2843–8. doi: 10.1073/pnas.051626398
8. Hsu SY, Hseuh AJW. Human stresscopin and stresscopin-related peptide are selective ligands for the type 2 corticotropin-releasing hormone receptor. *Nat Med*. (2001) 7:605–11. doi: 10.1038/87936
9. Cardoso JC, Bergqvist CA, Félix RC, Larhammar D. Corticotropin-releasing hormone family evolution: five ancestral genes remain in some lineages. *J Mol Endocrinol*. (2016) 57:73–86. doi: 10.1530/JME-16-0051
10. Lovejoy DA. Structural evolution of urotensin-I: retaining ancestral functions before corticotropin-releasing hormone evolution. *Gen Comp Endocrinol*. (2009) 164:15–9. doi: 10.1016/j.ygcen.2009.04.014
11. Lovejoy DA, De Lannoy L. Evolution and phylogeny of the corticotropin-releasing factor (CRF) family of peptides: expansion and specialization in the vertebrates. *J Neurochem Anat*. (2013) 54:50–6. doi: 10.1016/j.jchemneu.2013.09.006
12. Lovejoy DA, Jahan S. Phylogeny and Evolution of the corticotropin releasing factor family of peptides. *Gen Comp Endocrinol*. (2006) 146:1–8. doi: 10.1016/j.ygcen.2005.11.019
13. Lovejoy DA, Chang B, Lovejoy N, Del Castillo J. Molecular evolution and origin of the corticotrophin-releasing hormone receptors. *J Mol Endocrinol*. (2014) 52:6043–79. doi: 10.1530/JME-13-0238
14. Coast GM. Insect diuretic peptides: structures, evolution and actions. *Amer Zool*. (1998) 38:442–9. doi: 10.1093/icb/38.3.442
15. Kataoka H, Troetschler RG, Li JP, Kramer SJ, Carney RL, Schooley DA. Isolation and identification of a diuretic hormone from the tobacco hornworm, *Manduca sexta*. *Proc Natl Acad Sci USA*. (1989) 86:2976–80. doi: 10.1073/pnas.86.8.2976
16. Zandawala M. Calcitonin-like diuretic hormones in insects. *Insect Biochem Mol Biol*. (2012) 42:816–25. doi: 10.1016/j.ibmb.2012.06.006
17. Fredriksson R, Lagerström M, Lundin L, Schiöth H. The G-protein coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups and fingerprints. *Mol Pharmacol*. (2003) 63:1256–72. doi: 10.1124/mol.63.6.1256
18. Nordström K, Lagerström M, Wallér L, Fredriksson R, Schiöth H. The secretin GPCRs descended from the family of adhesion GPCRs. *Mol Biol Evol*. (2009) 26:71–84. doi: 10.1093/molbev/msn228
19. Schiöth H, Nordström K, Fredriksson R. The adhesion GPCRs: gene repertoire, phylogeny and evolution. In: Yona S, Stacey M, editors. *Adhesion GPCRs: Structure to Function*. Berlin: Landes Bioscience and Springer Science and Business Media (2010). p. 1–13. doi: 10.1007/978-1-4419-7913-1_1
20. Lelianova V, Davletov B, Sterling A, Rahman M, Grishin E, Totty N, et al. A-latrotoxin, receptor, latrophilin, is a novel member of the secretin family of G protein-coupled receptors. *J Biol Chem*. (1997) 272:21504–8. doi: 10.1074/jbc.272.34.21504
21. Holz GG, Habener JF. Black widow spider a-latrotoxin: a presynaptic neurotoxin that shares structural homology with the glucagon-like peptide-1 family of insulin secretagogic hormones. *Comp Biochem Physiol B Biochem Mol Biol*. (1989) 121:177–84. doi: 10.1016/S0305-0491(98)10088-3
22. Silva JP, Lelianova VG, Ermolyuk YS, Hitchen PG, Berninhausen O, Rahman MA, et al. Latrophilin 1 and its endogenous ligand Lasso/teneurin-2 form a high affinity transsynaptic receptor pair with signaling capabilities. *Proc Natl Acad Sci USA*. (2011) 108:12113–8. doi: 10.1073/pnas.1019434108
23. Boucard AA, Maxeiner S, Sudhof TC. Latrophilins function as heterophilic cell-adhesion molecules by binding to teneurins: regulation by alternative splicing. *J Biol Chem*. (2014) 289:387–402. doi: 10.1074/jbc.M113.504779
24. O'Sullivan ML, Martini F, von Daake S, Comoletti D, Ghosh A. LPHN3, a presynaptic adhesion-GPCR implicated in ADHD, regulates the strength of neocortical layer 2/3 synaptic input to layer 5. *Neural Dev*. (2014) 9:7. doi: 10.1186/1749-8104-9-7
25. Qian X, Barsyte-Lovejoy D, Wang L, Chewpoy B, Gautam N, Al Chawaf A, et al. Cloning and characterization of teneurin C-terminus associated peptide (TCAP)-3 from the hypothalamus of an adult rainbow trout (*Oncorhynchus mykiss*). *Gen Comp Endocrinol*. (2004) 137:205–16. doi: 10.1016/j.ygcen.2004.02.007
26. Wang L, Rotzinger S, Al Chawaf A, Elias C, Barsyte-Lovejoy D, Qian X, et al. Teneurin proteins possess a carboxy terminal sequence with neuromdulatory activity. *Mol Brain Res*. (2005) 133:253–65. doi: 10.1016/j.molbrainres.2004.10.019
27. Lovejoy DA, Al Chawaf A, Cadinouche A. Teneurin C-terminal associated peptides: An enigmatic family of neuropeptides with structural similarity to the corticotropin-releasing factor and calcitonin families of peptides. *Gen Comp Endocrinol*. (2006) 148:299–305. doi: 10.1016/j.ygcen.2006.01.012
28. Jackson VA, Meijer DH, Carrasquero M, van Bezouwen LS, Lowe ED, Kleanthous C, et al. Structures of Teneurin adhesion receptors reveal an ancient fold for cell-cell interaction. *Nat Commun*. (2018) 9:1079. doi: 10.1038/s41467-018-03460-0
29. Li J, Shalev-Benami M, Sando R, Jiang X, Kibrom A, Wang J, et al. Structural basis for teneurin function in circuit-wiring: a toxin motif at the synapse. *Cell*. (2018) 173:735–48. doi: 10.1016/j.cell.2018.03.036
30. Husić M, Barsyte-Lovejoy D, Lovejoy DA. Teneurin C-terminal associated peptide (TCAP)-1 and latrophilin interaction in HEK293 Cells: evidence for modulation of intercellular adhesion. *Front. Endocrinol.* (2019) 10:22. doi: 10.3389/fendo.2019.00022
31. Tucker RP. Horizontal gene transfer in choanoflagellates. *J Exp Zool B Mol Dev Evol*. (2013) 320:1–9. doi: 10.1002/jez.b.22480
32. Tucker R, Beckmann J, Leachman N, Schöler J, Chiquet-Ehrismann R. Phylogenetic analysis of the teneurins: conserved features and premetazoan ancestry. *Mol Biol Evol*. (2012) 29:1019–29. doi: 10.1093/molbev/msr271
33. Zhang D, de Souza R, Anantharaman V, Iyer L, Aravind L. Polymorphic toxin systems: comprehensive characterization of trafficking modes, processing, mechanisms of action, immunity and ecology using comparative genomics. *Biol Direct*. (2012) 7:18. doi: 10.1186/1745-6150-7-18
34. Chand D, de Lannoy L, Tucker RP, Lovejoy DA. Origin of chordate peptides by horizontal protozoan gene transfer in early metazoans and protists: evolution of the teneurin C-terminal associated peptides. *Gen Comp Endocrinol*. (2013) 188:144–50. doi: 10.1016/j.ygcen.2013.02.006
35. Endsin MJ, Michalec OM, Manzon LA, Lovejoy DA, Manzon RA. CRH peptide evolution occurred in three phases: evidence from characterizing sea lamprey CRH system members. *Gen Comp Endocrinol*. (2017) 240:162–73. doi: 10.1016/j.ygcen.2016.10.009
36. Semmens DC, Mirabeau O, Moghul I, Pancholi MR, Wurm Y, Elphick MR. Transcriptomic identification of starfish neuropeptide precursors yields new insights into neuropeptide evolution. *Open Biol*. (2016) 6:150224. doi: 10.1098/rsob.150224
37. Lovejoy, DA, Barsyte-Lovejoy D. Characterization of a diuretic hormone-like peptide from tunicates: insight into the origins of the vertebrate corticotropin-releasing factor (CRF) family. *Gen Comp Endocrinol*. (2010) 165:330–6. doi: 10.1016/j.ygcen.2009.07.013

38. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA 6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol*. (2013) 30:2725–9. doi: 10.1093/molbev/mst197
39. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. (2004) 32:1792–7. doi: 10.1093/nar/gkh340
40. Weiss M, Steiner DF, Philipson LH. Insulin biosynthesis, secretion, structure and structure-activity relationships. In: De Groot LJ, Chrousos G, Dungan K, editor. *Endotext*. South Dartmouth, MA: NIH/NLM/NCBI Books (2014). p. 1–6.
41. Jones DT, Taylor WR, Thornton JM. The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci*. (1992) 8:275–82.
42. Lang BF, O'Kelly C, Nerad T, Gray MW, Burger G. The closest unicellular relatives of animals. *Curr Biol*. (2002) 12:1773–8. doi: 10.1016/S0960-9822(02)01187-9
43. Minet A, Rubin B, Tucker R, Baumgartner S, Chiquet-Ehrismann R. Teneurin-1, a vertebrate homologue of the Drosophila pair-rule gene Ten-m, is a neuronal protein with a novel type of heparin-binding domain. *J Cell Sci*. (1999) 112:2019–32.
44. Wimalawansa SJ. Amylin, calcitonin gene-related peptide, calcitonin and adrenomedullin: a peptide superfamily. *Critic Rev Neurobiol*. (1997) 11:167–239. doi: 10.1615/CritRevNeurobiol.v11.i2-3.40

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

# Advantages of publishing in Frontiers

**OPEN ACCESS**
Articles are free to read for greatest visibility and readership

**FAST PUBLICATION**
Around 90 days from submission to decision

**HIGH QUALITY PEER-REVIEW**
Rigorous, collaborative, and constructive peer-review

**TRANSPARENT PEER-REVIEW**
Editors and reviewers acknowledged by name on published articles

**Frontiers**
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

**Visit us:** www.frontiersin.org
**Contact us:** frontiersin.org/about/contact

**REPRODUCIBILITY OF RESEARCH**
Support open data and methods to enhance research reproducibility

**DIGITAL PUBLISHING**
Articles designed for optimal readership across devices

**FOLLOW US**
@frontiersin

**IMPACT METRICS**
Advanced article metrics track visibility across digital media

**EXTENSIVE PROMOTION**
Marketing and promotion of impactful research

**LOOP RESEARCH NETWORK**
Our network increases your article's readership