Discovery of novel representatives of bilaterian neuropeptide families and reconstruction of neuropeptide precursor evolution in ophiuroid echinoderms

Meet Zandawala¹, Ismail Moghul¹, Luis Alfonso Yañez Guerra¹, Jerôme Delroisse¹, Nikara Abylkassimova¹, Andrew F. Hugall², Timothy D. O’Hara² and Maurice R. Elphick¹

¹School of Biological and Chemical Sciences, Queen Mary University of London, Mile End Road, London E1 4NS, UK
²Museums Victoria, GPO Box 666, Melbourne, Victoria 3001, Australia

Neuropeptides are a diverse class of intercellular signalling molecules that mediate neuronal regulation of many physiological and behavioural processes. Recent advances in genome/transcriptome sequencing are enabling identification of neuropeptide precursor proteins in species from a growing variety of animal taxa, providing new insights into the evolution of neuropeptide signalling. Here, detailed analysis of transcriptome sequence data from three brittle star species, Ophionotus victoriae, Amphiura filiformis and Ophiopsila aranea, has enabled the first comprehensive identification of neuropeptide precursors in the class Ophiuroidea of the phylum Echinodermata. Representatives of over 30 bilaterian neuropeptide precursor families were identified, some of which occur as paralogues. Furthermore, homologues of endothelin/CCHamide, eclosion hormone, neuropeptide-F/Y and nucleobinin/nesfatin were discovered here in a deuterostome/echinoderm for the first time. The majority of ophiuroid neuropeptide precursors contain a single copy of a neuropeptide, but several precursors comprise multiple copies of identical or non-identical, but structurally related, neuropeptides. Here, we performed an unprecedented investigation of the evolution of neuropeptide copy number over a period of approximately 270 Myr by analysing sequence data from over 50 ophiuroid species, with reference to a robust phylogeny. Our analysis indicates that the composition of neuropeptide ‘cocktails’ is functionally important, but with plasticity over long evolutionary time scales.

1. Introduction

The nervous systems of animals use a wide variety of chemicals for neuronal communication. These include amino acids (e.g. glutamate), biogenic amines (e.g. serotonin) and neuropeptides (e.g. vasopressin) among others. Neuropeptides are, by far, the most diverse and they control many physiological/behavioural processes, including feeding, reproduction and locomotion [1–3]. Recent advances in genome/transcriptome sequencing are enabling the identification of neuropeptide precursor proteins in species from a growing variety of animal taxa, providing new insights into the evolution of neuropeptide signalling [4–8]. The echinoderms are notable in this regard because as deuterostomian invertebrates they occupy an ‘intermediate’ phylogenetic position with respect to the vertebrates and intensely studied protostomian invertebrates such as insects (e.g. Drosophila melanogaster) and nematodes (e.g. Caenorhabditis elegans). Accordingly, characterization of neuropeptide signalling...
systems in echinoderms has recently provided key ‘missing links’ for the determination of neuropeptide relationships and reconstruction of neuropeptide evolution [8–10].

The phylum Echinodermata comprises five extant classes: Echinoidea (sea urchins and sand dollars), Holothuroidea (sea cucumbers), Asteroidea (starfish), Ophiuroidea (brittle stars and basket stars) and Crinoidea (sea lilies and feather stars). Recent molecular phylogenetic studies support the hypothesis that Echinoidea and Holothuroidea are sister groups (Echinozoa), and Asteroidea and Ophiuroidea are also sister groups (Asterozoa), with the Crinoidea basal to the Echinozoa + Asterozoa clade (Eleutherozoa) [11,12]. Echinoderms are marine organisms that have several unique features including pentaradial symmetry as adults, a remarkable ability to autotomize and regenerate body parts, and neurally controlled mutable collagenous tissue [13,14]. Previous transcriptomic analyses have identified neuropeptide precursor complements in Strongylocentrotus purpuratus (purple sea urchin), Apostichopus japonicus (Japanese sea cucumber) and Asterias rubens (common European starfish) [8,15,16]. Furthermore, the identification of neuropeptides in these species has facilitated investigation of the evolution and physiological roles of various neuropeptide signalling systems [8–10,17–21].

The recent progress in transcriptomic/genomic characterization of echinoderm neuropeptide systems has hitherto not been extended to ophiuroids or crinoids. The Ophiuroidea constitutes the largest class among extant echinoderms [22], with a long evolutionary history that extends back to the early Ordovician (around 480 Ma) [23], while the extent radiation dates from the mid-Permian (approx. 270 Ma) [12]. Available molecular data for ophiuroids have increased significantly in recent years with the emergence of numerous transcriptomic studies [20,24–29]. Here, we use transcriptome sequence data from three brittle star species, Ophiuratus victoriæ, Amphipura filiformis and Ophiopsila aranea, to perform the first comprehensive identification of neuropeptide precursors in ophiuroids. We identify representatives of over 30 neuropeptide families including homologues of endothelin/CCIHamide, eclosion hormone (EH), neuropeptide-F/Y (NPY/NPF) and nucleobin (NUCB)/nesfatin, which are the first to be discovered in a deuterostome/echinoderm.

Transcriptomes have also been employed to investigate the phylogenetic relationships of the ophiuroids, utilizing data from 52 species [12]. In the most comprehensive molecular analysis of ophiuroid phylogeny to date, previous morphology-based classification schemes [30] were rejected in favour of a new phylogeny comprising three primary ophiuroid clades and six orders [12,31,32]. This landmark study and the associated large dataset have provided a unique opportunity to investigate the conservation and diversification of neuropeptide precursor structure over a period of approximately 270 Myr of ophiuroid evolution. Our analysis reveals that the majority of ophiuroid neuropeptide precursors contain a single copy of a neuropeptide, but several precursors comprise multiple copies of identical or non-identical, but structurally related, neuropeptides. Interestingly, the number of neuropeptide copies in the majority of precursors is constant across all the ophiuroid species examined, but examples of clade-specific losses/gains of neuropeptides are also observed. This remarkable conservation in neuropeptide copy number across approximately 270 Myr of ophiuroid evolution indicates that the composition of neuropeptide ‘cocktails’ is functionally important, but with plasticity over long evolutionary time scales.

2. Results and discussion

Here, we have used transcriptome sequence data for the first comprehensive identification of neuropeptide precursors in ophiuroids (figure 1). Representatives of over 30 bilaterian neuropeptide precursor families were identified. Identification of ophiuroid representatives of these neuropeptide precursor types has, in some cases, provided new insights into neuropeptide precursor structure and evolution, as discussed in more detail below. First, however, we will highlight representatives of bilaterian neuropeptide precursor families that have been identified here for the first time in an echinoderm species.

2.1. Discovery of the first echinoderm representatives of bilaterian neuropeptide families

Comprehensive analysis of transcriptome sequence data from three ophiuroid species, O. victoriæ, A. filiformis and O. aranea, has enabled the discovery of the first echinoderm representatives of four bilaterian neuropeptide families. Specifically, we have discovered the first deuterostomian homologues of EH (figure 2), the first ambulacrarian homologue of CCIHamide/endothelin-type peptides (figure 3a), and the first echinoderm homologues of NPY/NPF (figure 3b) and NUCB/nesfatin (electronic supplementary material, figure S1), as discussed in detail below.

2.1.1. Eclosion hormone

EH was first isolated and sequenced in the insects Manduca sexta (tobacco hornworm) and Bombyx mori (silk moth), and was shown to alter the timing of adult emergence [34,35]. EH is one of the main peptide/protein hormones involved in control of ecysis (i.e. shedding of the cuticle) behaviour in arthropods [36,37]. It binds to and activates a receptor guanylyl cyclase that is expressed in epitracheal Inka cells and causes the secondary release of ecysis-triggering hormone (ETH) that is also expressed in Inka cells [38,39]. In Drosophila, EH is important for ecysis, but whether this hormone is essential for ecysis is not yet clear [40,41]. EH null mutant flies show defects in ecysis and are unable to reach adulthood, yet some flies in which EH-producing neurons have been genetically ablated (a more extreme manipulation) are able to survive till adulthood. Arthropod EHs have six conserved cysteine residues that form three disulfide bridges [38]. EHs have not been discovered previously outside of arthropods. Interestingly, four EH-like precursors were identified in A. filiformis and O. aranea, and two in O. victoriæ (electronic supplementary material, figures S2–S4; GenBank: MF155236; MF155237). The ophiuroid EH-like precursors are orthologous to neuropeptide precursors previously identified in the sea urchin S. purpuratus (Spnp11 and Spnp15, which we now rename as Spur EH1 and Spur EH2, respectively) [16] and the starfish A. rubens (Arnp11, Arnp15 and Arnp15b renamed as Arub EH1, Arub EH2a and Arub EH2b, respectively) [8]. The positions of cysteine residues are conserved across all echinoderm and insect EHs, but aside from this there is little sequence conservation (figure 2n). The echinoderm EH-like precursor sequences were also analysed using a sequence similarity-based clustering approach based on BLASTp e-values using the CLANS software [42]. The analysis shows
that echinoderm EH-like precursors (i) cluster in two compact subgroups (echinoderm EH-like precursor 1 and EH-like precursor 2), and (ii) have strong positive BLAST results with arthropod EHs and, to a lesser extent, with arthropod ion transport peptide (ITP) and vertebrate atrial natriuretic peptide (ANP) (figure 2b). ITP precursors also possess six cysteine residues; however, the position of these residues is not conserved with cysteine residues found in echinoderm EH-like precursors (not shown).

To obtain further evidence for the presence of an EH-like signalling system in echinoderms, we performed a phylogenetic analysis of EH-type receptors. Insect EHs mediate their effects by binding to membrane guanylyl cyclase receptors [39]. EH receptors are closely related to vertebrate ANP receptors and various orphan receptors [33]. Specific BLAST searches enabled the identification of transcripts in O. victoriae, A. filiformis and O. aranea that encode proteins similar to arthropod EH receptors. Maximum-likelihood and Bayesian phylogenetic analyses confirmed that these sequences group with the receptor cluster containing EH receptors (figure 2c).

The discovery of the first deuterostomian EHs suggests an ancient bilaterian origin of EHs and indicates that these hormones may have other functions in invertebrates aside from their role in ecdysis.

2.1.2. CCHamide

CCHamides are neuropeptides that were discovered relatively recently in the silkworm B. mori [43]. Later, it was found that insects have two CCHamide genes, CCHamide-1 and CCHamide-2, each encoding a single copy of the mature peptide [44]. These peptides are referred to as CCHamides because they contain two cysteine residues and a characteristic histidine-amide C-terminal motif. There are two CCHamide receptors in insects: CCHamide-1 specifically activates one receptor and CCHamide-2 specifically activates the second receptor [44,45]. CCHamide-1 has a physiological role in starvation-induced olfactory modifications [46], whereas CCHamide-2 regulates feeding, growth and developmental timing in flies [45,47]. Recent studies examining the evolution of neuropeptides in the Bilateria have shown that protostomian CCHamides are related to elevenin (another protostomian neuropeptide originally discovered from the mollusc Aplysia californica L11 neuron), lophotrochozoan GGNG peptides, endothelins and gastrin-releasing peptides (GRPs) [6,7,48,49]. The latter two are neuropeptide types that have not been found outside chordates. Furthermore, the degree of sequence/structural conservation varies across these different peptide families. Hence, CCHamides are amidated and have
Figure 2. (Caption opposite.)
Figure 2. (Opposite) Edcsion hormone (EH)-type peptides and receptors in echinoderms. (a) Partial multiple sequence alignment of EH-type precursor sequences, excluding the N-terminal signal peptide. (b) Cluster analysis of arthropod EH precursors, echinoderm EH-like precursors, arthropod ion transport peptides (ITPs) and vertebrate atrial natriuretic peptides (ANPs) shows that echinoderm EH-like precursors are more closely related to arthropod EH than ITP. (c) Maximum-likelihood and Bayesian phylogenetic analyses of membrane guanylate cyclase receptors show that EH-like receptors are found in echinoderms, but are absent in vertebrates as seen for the EH-like precursors. GCG1, 2, 3 and 4 are orphan guanylate cyclase receptors found in arthropods [33]. Echinoderm EH-like receptors are clustered with arthropod EH receptors, neuropetide-like peptide 1-VQQ receptors (NPLP1-VQQ) and OGC1 receptors. The inset shows the alternate topology obtained following Bayesian analysis. Species names: Ophionotus victoriae (Ovic), Asterias rubens (Arub), Strongylocentrotus purpuratus (Spur), Drosophila melanogaster (Dmel), Bombyx mori (Bmor) and Pediculus humanus corporis (Pcor).

(a) CCHamide

(b) Neuropeptide-F/Y

2.1.4. NUCB

Nucleobindins (NUCB1 and NUCB2) are multidomain Ca\(^{2+}\) and DNA-binding proteins. NUCB1 was first discovered in 1992 and thought to play a role in apoptosis and autoimmunity [57]. Interestingly, the NUCB1 precursor has both a signal peptide and a leucine zipper structure, suggesting that it can bind DNA and act as an endocrine factor [58]. NUCB2 is a homologue of NUCB1 and was named based on high sequence similarity between the two precursors [59]. In 2006, an 82 amino acid peptide located in the N-terminal region of NUCB2 was reported. This peptide, Nesfatin-1 (Nucleobindin-1), was discovered as a satiety-inducing factor in the rat hypothalamus [60]. Its role in inhibiting food intake in vertebrates is now well established [59,61]. Moreover, this pleiotropic peptide also modulates other processes, including glucose and lipid metabolism, and cardiovascular and reproductive functions. Recently, nesfatin-1-like peptide derived from NUCB1 was shown to be anorexigenic in goldfish [62]. Surprisingly, the presence of NUCBcs in invertebrates other than Drosophila has not been reported until now [63]. Here, we show that NUCB-type precursors are present in echinoderms (electronic supplementary material, figure S1a). Phylogenetic analysis of NUCB precursors reveals that a single copy of the NUCB precursor is found in invertebrate species and gene duplication in the vertebrate lineage gave rise to NUCB1 and NUCB2 (electronic supplementary material, figure S1b). In chordates, the NUCB precursors are predicted to generate three peptides (Nesfatin-1, 2 and 3); however, no biological role has been attributed specifically to nesfatin-2 and nesfatin-3. Interestingly, the prohormone conversion sites expected to

demonstrated for the first time the occurrence of NPY homologues in invertebrates [55]. Here, we have identified the first echinoderm representatives of the NPY/NPF family in brittle stars and starfish (figure 3b; electronic supplementary material, figure S12). The brittle star precursors contain a peptide with a C-terminal RFamide, in common with NPY in vertebrates and an ortholog in the starfish Patiria miniata. By contrast, an orthologue in the starfish A. rubens has a C-terminal RFamide, a feature that it shares with NPY/NPF-type peptides in the hemichordate Saccoglossus kowalevskii and in protostomes. Thus, our findings have revealed that NPY/NPF-type peptides with a C-terminal Yamide motif are not restricted to vertebrates, as has been previously shown in some insects [52]. Echinoderm NPY/NPF-type precursors are located immediately after the signal peptide in the precursor proteins, as is the case in other bilaterian species. Surprisingly, we did not find NPY/NPF-type precursors in the sea urchin S. purpuratus or the sea cucumber A. japonicus. However, we suspect that this may reflect sequence divergence rather than gene loss because a gene encoding an NPY/NPF-type receptor can be found in the S. purpuratus genome [56].

2.1.3. Neuropeptide-Y/Neuropeptide-F

NPY was first isolated and sequenced from the porcine hypothalamus in 1982 [30,51]. Although the NPY/NPF family of peptides are pleiotropic in nature [52], they are mainly known for their roles in regulation of feeding and stress [3,53,54]. The discovery of NPY in the tapeworm Monieza expansa in 1991 suggested that this peptide might be widespread throughout the animal kingdom. Indeed, recent studies have revealed that NPY and NPF are present in a variety of invertebrate species, including nematodes, crustaceans, molluscs, and echinoderms [55]. These findings have expanded our understanding of the evolutionary history of these peptides and provided new insights into their functional diversity.

Asterias rubens 1-VQQ receptors (NPLP1-VQQ) and OGC1 receptors. The inset shows the alternate topology obtained following Bayesian analysis. Species names: Ophionotus victoriae (Ovic), Asterias rubens (Arub), Strongylocentrotus purpuratus (Spur), Drosophila melanogaster (Dmel), Bombyx mori (Bmor) and Pediculus humanus corporis (Pcor).
generate Nesfatin-1, 2 and 3 are conserved between echinoderm and chordate NUCBs. Moreover, the O. victoriae precursor (GenBank: MF155235) has an additional predicted cleavage site within the Nesfatin-1-containing region, which is not present in other species (except for D. melanogaster). However, it remains to be determined whether or not this cleavage site in the O. victoriae precursor is functional.

2.2. First comprehensive identification of neuropeptide precursors in ophiuroids

We have identified neuropeptide precursors belonging to 32 families, which represent the first comprehensive analysis of neuropeptide precursors in ophiuroids (figure 4; electronic supplementary material, figures S2–S4). Several of these neuropeptide families have been identified previously in echinoderms and include homologues of AN peptides, bursicon (α and β) (GenBank: MF155260; MF155227), calcitonin (GenBank: MF155228), cholecystokinin (CCK) (GenBank: MF155231; MF155232) [15], corazonin (GenBank: MF155233) [10], corticotropin-releasing hormone (CRH) (GenBank: MF155234, MF155235, MF155261, MF155262), glycoprotein hormones (α2 and β5) (GenBank: MF155238; MF155239; MF155240) [64], gonadotropin-releasing hormone (GnRH) (GenBank: MF155263) [10], insulin-like peptide (GenBank: MF155264) (64), kisspeptin (KP) (GenBank: MF155241) [8], luqin (GenBank: MF155242) [7], melanin-concentrating hormone (MCH) (GenBank: MF155243) [8], NG peptides (neuropeptide-S) (GenBank: MF155244) [9,65], orexin (GenBank: MF155245; MF155246) [6,8], pedal peptides (GenBank: MF155247; MF155266; MF155267) [16], pigment-dispersing factor (PDF) (GenBank: MF155248) [8], relaxin-like peptide (GenBank: MF155249) [66], SALMFamides (L-type and F-type) (GenBank: MF155250; MF155268) [19,20,67], somatostatin (GenBank: MF155252; MF155253) [8], tachykinin (GenBank: MF155254) [8], thyrotropin-releasing hormone (TRH) (GenBank: MF155255; MF155256) [16] and vasopressin/oxytocin (GenBank: MF155257) [64,65] (figures 5–7; electronic supplementary material, figures S5–S10). With the exception of MCH (which may be unique to deuterostomes) [6,8], AN peptides and SALMFamides (which thus far have only been identified in echinoderms), the origins of all of the neuropeptide precursors identified here in ophiuroids predate the divergence of protostomes and deuterostomes [6,7]. Interestingly, insects have a single SK precursor, which comprises two neuropeptides found in insects [6,7].

2.3. Neuropeptide precursors that occur in multiple forms in Ophionotus victoriae

2.3.1. Thyrotropin-releasing hormone-type precursors

TRH (also known as thyrotropin-releasing factor or thyroliberin) was first isolated and sequenced in the 1960s [68–70]. In mammals, TRH is produced in the hypothalamus, and stimulates the release of thyroid-stimulating hormone (TSH) and prolactin from the anterior pituitary [71,72]. The recent discovery of a TRH receptor in the annelid Platynereis dumerilli indicates that the evolutionary origin of this neuropeptide signalling system predates the divergence of protostomes and deuterostomes [73].

The human TRH precursor contains six copies of the tripeptide pQHPamide [74]. Precursor proteins comprising multiple copies of TRH-like peptides have been identified previously in the sea urchin S. purpuratus, the sea cucumber A. japonicus and the starfish A. rubens [8,15,16], with a single TRH-type precursor found in each of these species. Interestingly, here we identified two TRH-type precursors (OvTRHP1 and OvTRHP2) in O. victoriae (electronic supplementary material, figure S2; figure 6a). OvTRHP1 comprises 21 copies of putative TRH-like tetrapeptides with the motif pQXXAmide (where X is variable). OvTRHP2, on the other hand, comprises two copies of the putative tetrapeptide pQGPRamide and two longer peptides that also have a C-terminal GPRamide motif but lack the N-terminal pyroglutamate.

2.3.2. Cholecystokinin-type precursors

A CCK-type peptide (formerly pancreozymin) was first sequenced in the 1960s [75]. CCK-type peptides play numerous roles in feeding and digestion-related physiology. CCK mediates satiety, and stimulates the release of digestive enzymes and gall bladder contractions [76–78]. CCK-type peptides are involved in mechanisms of learning and memory, and analgesia [79]. A neuropeptide precursor comprising two CCK-like peptides was recently identified in the starfish A. rubens [8]. Here, we have identified two CCK-type precursors in O. victoriae (OvCCKP1 and OvCCKP2) and orthologues of both of these precursors were also identified in the sea urchin S. purpuratus (electronic supplementary material, figure S2) [16]. The CCK-type precursor 1 comprises three CCK-like peptides in both O. victoriae and S. purpuratus, and this precursor is similar to the A. rubens CCK-type precursor, which comprises two CCK-like peptides. By contrast, the CCK-type precursor 2 comprises a single CCK-like peptide in both O. victoriae and S. purpuratus. Interestingly, the sequence of the S. purpuratus CCK-type precursor 2 was reported previously as part of a genome-wide search for neuropeptides [80], but the authors of this study did not identify it as a CCK-type precursor. However, based on the presence of a conserved tyrosine residue and a C-terminal F-amide motif in the predicted neuropeptide derived from this protein, it is evident that it belongs to the family of CCK-type precursors (figure 6b). A search of a preliminary genome assembly of the starfish P. miniata (http://www.echinobase.org) [81] did not reveal a gene encoding a CCK-type precursor 2. Therefore, it appears that this neuropeptide precursor type may have been lost in the Asteroidea; nevertheless, further analysis of a wider range of starfish species will be required to draw definitive conclusions. With a broader evolutionary perspective, CCK-type peptides in deuterostomes are orthologues of sulfakinin (SK)-type neuropeptides found in insects [6,7]. Interestingly, insects have a single SK precursor, which comprises two neuropeptides, SK-1 and SK-2 [82], and this may reflect the ancestral condition in the common ancestor of protostomes and deuterostomes. Thus, the occurrence of two CCK-type peptides on a single precursor in A. rubens and insects may be an ancestral characteristic, and the occurrence of two CCK-type precursors that comprise one and three CCK-type peptides appears to be a derived characteristic.
<table>
<thead>
<tr>
<th>Receptor Type</th>
<th>Neuropeptide Family</th>
<th>O. victoriae</th>
<th>A. filiformis</th>
<th>O. aranea</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CCHamide-like 1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>CCHamide-like 2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>corazonin</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>gonadotropin-releasing hormone</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>luquin</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>NG peptide/neuropeptide-S</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>orexin 1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>orexin 2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>tachykinin</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>10</td>
<td>thyrotrpin-releasing hormone 1</td>
<td>21</td>
<td>14*</td>
<td>17</td>
</tr>
<tr>
<td>11</td>
<td>vasopressin / oxytocin</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>kispeptin</td>
<td>2</td>
<td>1*</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>melanin-concentrating hormone</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>somatostatin 1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>15</td>
<td>bursicon alpha</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>16</td>
<td>bursicon beta</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>17</td>
<td>glycoprotein hormone alpha 2.1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>18</td>
<td>glycoprotein hormone alpha 2.2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>19</td>
<td>relaxin-like peptide</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>20</td>
<td>calcitonin</td>
<td>2</td>
<td>1/2</td>
<td>1/2</td>
</tr>
<tr>
<td>21</td>
<td>corticotropin-releasing hormone 1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>22</td>
<td>pigment-dispersing factor</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>23</td>
<td>AN peptide</td>
<td>1</td>
<td>5*</td>
<td>7</td>
</tr>
<tr>
<td>24</td>
<td>eclosion hormone 1.1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>25</td>
<td>insulin-like peptide</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>26</td>
<td>nucleobindin/nefastin</td>
<td>b</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>27</td>
<td>pedal peptide 1</td>
<td>6</td>
<td>4*</td>
<td>1*</td>
</tr>
<tr>
<td>28</td>
<td>SLMFMamide (L-type)</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>29</td>
<td>SLMFMamide (F-type)</td>
<td>12</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>30</td>
<td>neuropeptide precursor 18</td>
<td>4</td>
<td>2*</td>
<td>4</td>
</tr>
<tr>
<td>31</td>
<td>neuropeptide precursor 26</td>
<td>7</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>32</td>
<td>neuropeptide precursor 27</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

Figure 4. Summary of neuropeptide precursors identified in *Ophionotus victoriae*, *Amphiura filiformis* and *Ophiopsila aranea*. Neuropeptide precursors are classified based on the type of G-protein coupled receptor (GPCR) their constituent peptides are predicted to activate (see Mirabeau and Joly [6]). Some peptides bind to receptors other than GPCRs and these are grouped with peptides where the receptor is unknown. Ophiuroids have neuropeptide precursors from up to 32 families. The number of putative mature peptides derived from each precursor has been indicated along with the presence of amidation and pyroglutamation.
2.3.3 Somatostatin-type precursors

Somatostatin was first isolated and sequenced from sheep hypothalamus in 1973 [83]. This peptide inhibits the release of pituitary hormones such as growth hormone, prolactin and TSH [84]. Moreover, it also inhibits the release of gastrointestinal (CCK and gastrin among others) and pancreatic (insulin and glucagon) hormones [85–87]. Aside from its effects on release of hormones, somatostatin also has central actions that influence motor activity [85]. Here, we have identified two somatostatin-type precursors (OvSSP-1 and OvSSP-2) in *O. victoriae* (electronic supplementary material, figure S2; figure 6c). Homologues of both of these precursors are present in the sea urchin *S. purpuratus* (electronic supplementary material, figure S2; figure 6c), one of which was previously referred to as Spnp16 [16]. By comparison, only a single somatostatin-type precursor has been found in the starfish *A. rubens*, which is an orthologue of OvSSP-1 [8]. All somatostatin-type precursors comprise a single copy of the bioactive neuropeptide, which is located in the C-terminal region of the precursor [88,89]. Interestingly, the type-1 somatostatins in echinoderms have a phenylalanine residue located in the middle part of the peptide, and this conserved feature is found in human somatostatin. Conversely, type-2 somatostatins in echinoderms lack the phenylalanine residue, but have a neighbouring tryptophan-lysine (WK) motif that is also conserved in human and *Branchiostoma floridae* somatostatins (figure 6c). The deuterostomian somatostatins are orthologous to the allatostatin-C neuropeptide family in arthropods [7]. This family of peptides comprises three precursor types: allatostatin-C, allatostatin-CC and the recently discovered allatostatin-CCC [89,90]. Both allatostatin-C and allatostatin-CC are non-amidated, like somatostatins; however, allatostatin-CCC has a C-terminal amide. Hence, non-amidated peptides may be representative of the ancestral condition in the common ancestor of protostomes and deuterostomes. 

![Figure 5.](http://rsob.royalsocietypublishing.org/Content/figures/6170129a.png)

Figure 5. Multiple sequence alignments of mature peptides belonging to selected neuropeptide families. (a) Corazonin alignment; (b) gonadotropin-releasing hormone (GnRH) alignment; (c) orexin alignment; (d) luqin alignment; (e) vasopresin/oxytocin (VP/OT) alignment; (f) Ovnp18 alignment; (g) MCH alignment; (h) NG peptide alignment; (i) pigment-dispersing factor (PDF) alignment (see electronic supplementary material, figure S10 for a multiple sequence alignment of PDF-type precursors). Species names: *Ophionotus victoriae* (Ovic), *Asterias rubens* (Arub), *Strongylocentrotus purpuratus* (Spur), *Apostichopus japonicus* (Ajap), *Saccoglossus kowalevskii* (Skow), *Branchiostoma floridae* (Bflo), *Anopheles gambiae* (Agam), *Daphnia pulex* (Dpul), *Strigamia maritima* (Smar), *Lottia gigantea* (Lgig) and *Homo sapiens* (Hsap). 

2.3.3. Somatostatin-type precursors

Somatostatin was first isolated and sequenced from sheep hypothalamus in 1973 [83]. This peptide inhibits the release of pituitary hormones such as growth hormone, prolactin and TSH [84]. Moreover, it also inhibits the release of gastrointestinal (CCK and gastrin among others) and pancreatic (insulin and glucagon) hormones [85–87]. Aside from its effects on release of hormones, somatostatin also has central actions that influence motor activity [85]. Here, we have identified two somatostatin-type precursors (OvSSP-1 and OvSSP-2) in *O. victoriae* (electronic supplementary material, figure S2; figure 6c). Homologues of both of these precursors are present in the sea urchin *S. purpuratus* (electronic supplementary material, figure S2; figure 6c), one of which was previously referred to as Spnp16 [16]. By comparison, only a single somatostatin-type precursor has been found in the starfish *A. rubens*, which is an orthologue of OvSSP-1 [8]. All somatostatin-type precursors comprise a single copy of the bioactive neuropeptide, which is located in the C-terminal region of the precursor [88,89]. Interestingly, the type-1 somatostatins in echinoderms have a phenylalanine residue located in the middle part of the peptide, and this conserved feature is found in human somatostatin. Conversely, type-2 somatostatins in echinoderms lack the phenylalanine residue, but have a neighbouring tryptophan-lysine (WK) motif that is also conserved in human and *Branchiostoma floridae* somatostatins (figure 6c). The deuterostomian somatostatins are orthologous to the allatostatin-C neuropeptide family in arthropods [7]. This family of peptides comprises three precursor types: allatostatin-C, allatostatin-CC and the recently discovered allatostatin-CCC [89,90]. Both allatostatin-C and allatostatin-CC are non-amidated, like somatostatins; however, allatostatin-CCC has a C-terminal amide. Hence, non-amidated peptides may be representative of the ancestral condition in the common ancestor of protostomes and deuterostomes.
deuterostomes, with the amidated allatostatin-CCC probably having evolved only within the arthropod lineage [90]. It remains to be determined whether or not the duplication of somatostatin-type precursors in echinoderms and the duplication of allatostatin-C (to give rise to allatostatin-CC) represent independent duplications. Further insights into this issue may be obtained if the receptors for somatostatin-type peptides in echinoderms are deorphanized.

2.3.4. Corticotropin-releasing hormone-type precursors

CRH-type peptides are a family of related neuropeptides that include CRH, urocortins and urotensin-I in chordates, egg-laying hormone (ELH) in lophotrochozoans and diuretic hormone 44 (DH44) in arthropods [6,7]. Arthropods usually have a single DH44 precursor, which comprises a single copy of the mature peptide. In some insects, such as Tribolium castaneum and B. mori, alternative splicing of DH44 transcripts results in multiple mature peptide isoforms of varying lengths [43,91]. The situation in lophotrochozoans is more complex, with several species having multiple precursors and some of these precursors comprising multiple ELH mature peptides [4,92]. A single CRH-type precursor was found previously in the starfish A. rubens, whereas here we have identified four CRH-type precursors in O. victoriae (electronic supplementary material, figure S2; figure 6d). Thus, expanded families of CRH-type peptides and receptors appear to have evolved independently in multiple animal lineages, including chordates and ophiuroid echinoderms [93,94].

2.4. Diversity in neuropeptide precursor structure: new insights from ophiuroids

2.4.1. Tachykinins

The mammalian neuropeptide substance P was the first tachykinin-type peptide to be isolated and sequenced [95–97]. Subsequently, tachykinin-type peptides were discovered in other animals including tunicates [98], insects [99,100], annelids [101] and molluscs [102]. Tachykinin-type peptides regulate various physiological processes, including muscle contractility [103], nociception [104] and stress responses [105], among others [106]. Analysis of genomic/transcriptomic sequence data from the sea urchin S. purpuratus and the sea cucumber A. japonicus did not identify candidate tachykinin-type precursors [6,7,15,16]. However, recently a putative tachykinin-type precursor was discovered in the starfish A. rubens (ArTKP), indicating that this signalling system does occur in some echinoderms [8]. Here, we have identified orthologues of ArTKP in O. victoriae and other ophiuroids (figures 4 and 7a). Collectively, these findings indicate that this signalling system has been retained in the Asterozoa but lost in the Echinozoa.
Comparison of the structure of the asterozoan tachykinin-type precursors reveals that the *A. rubens* precursor (ArTKP) comprises two putative mature peptides, whereas the *O. victoriae* precursor comprises four mature peptides (figure 7b). It remains to be determined, however, which of these two conditions represent the ancestral state in the common ancestor of the Asterozoa. Further insights into this issue may be obtained if sequence data from a variety of starfish species are analysed.

### 2.4.2. Kisspeptins

KP (formerly known as metastin) is encoded by the *KiSS1* gene in humans. *KiSS1* was originally discovered as a gene that may suppress the metastatic potential of malignant melanoma cells [107]. Subsequently, it was found to play a vital role in regulating the onset of puberty. Thus, in vertebrates, KP binds to its receptor GPR54 to stimulate pituitary release of GnRH [108]. The first KP-type precursors to be identified in non-chordates were discovered recently in ambulacrarians—the echinoderms *A. rubens* and *S. purpuratus* and the hemichordate *S. kowalevskii* [8]. Accordingly, here we have identified KP-type precursors in *O. victoriae* and other ophiuroids. All of the ambulacrarian precursor proteins comprise two KP-type peptides, and the first putative neuropeptide in the echinoderm precursors has two cysteine residues at the N-terminus, which could form an N-terminal disulfide bridge similar that of calcitonin-type peptides (see below). By contrast, the second putative neuropeptide does not contain any cysteine residues and is typically shorter than the first peptide (figure 7c, d). Interestingly, comparison of the sequences of the first (long) and second (short) KP-type peptides in echinoderms reveals that the long and short peptides share less sequence similarity with each other within a species than they do with respective peptides in other species (figure 7c). This indicates that the duplication event that gave

---

**Figure 7.** Comparative analysis of ophiuroid tachykinin, KP and calcitonin-type precursors and neuropeptides. (a) Alignment of tachykinin-type peptides in *O. victoriae* (Ophiuroidea) and *A. rubens* (Asteroidea). (b) Schematic diagrams of the *O. victoriae* and *A. rubens* tachykinin precursors showing the location of the signal peptide (SP) and predicted neuropeptides (labelled 1–4). (c) Alignments of the long and short forms of kisspeptin (KP)-type neuropeptides in *O. victoriae*, *A. rubens* and *S. purpuratus* (Echinoidea). (d) Schematic diagrams of the *O. victoriae* and *A. rubens* KP precursors showing the locations of the SP, short and long orthocopies and cysteine (C) residues. (e) Alignment of calcitonin-type peptides from *O. victoriae*, *A. rubens*, *S. purpuratus* and *A. japonicus* (Holothuroidea). (f) Predicted alternative splicing of the calcitonin gene in ophiuroids, with the location of the SP and neuropeptides (CT1 and CT2) labelled. Species names: Ophionotus victoriae (Ovic), Asterias rubens (Arub), Strongylocentrotus purpuratus (Spur) and Apostichopus japonicus (Ajap).
rise to the occurrence of the long and short peptides occurred before the divergence of the Asterozoa and Echinodermata. Interestingly, previous studies have revealed that there has been an expansion of KP-type receptors in ambulacraria (S. purpuratus and S. kowalevskii) and in the cephalochordate, B. floridæ, with 16 KP receptors present in the latter [6,56]. Furthermore, studies are now needed to identify the proteins that act as receptors for the KP-type peptides identified here in ophiuroids and previously in other echinoderms [8].

### 2.4.3. Calcitonin

Calcitonin was first discovered in 1962 by Copp & Cheney [109]. The sequencing of the porcine calcitonin in 1968 revealed that this polypeptide is composed of 32 amino acids [110]. In vertebrates, calcitonin is produced by the thyroid gland [111] and regulates calcium (Ca^{2+}) levels in the blood, antagonizing the effects of parathyroid hormone [112,113]. The evolutionary antiquity of calcitonin-related peptides is further evidenced by the discovery that this polypeptide is composed of 32 amino acids [110]. The sequencing of the porcine calcitonin in 1968 revealed that this polypeptide is composed of 32 amino acids [110]. In vertebrates, calcitonin is produced by the thyroid gland [111] and regulates calcium (Ca^{2+}) levels in the blood, antagonizing the effects of parathyroid hormone [112,113]. The evolutionary antiquity of calcitonin-related peptides is further evidenced by the discovery that this polypeptide is composed of 32 amino acids [110].

Interestingly, previous studies have revealed that there has been an expansion of KP-type receptors in ambulacraria (S. purpuratus and S. kowalevskii) and in the cephalochordate, B. floridæ, with 16 KP receptors present in the latter [6,56]. Furthermore, studies are now needed to identify the proteins that act as receptors for the KP-type peptides identified here in ophiuroids and previously in other echinoderms [8].

Calcitonin was first discovered in 1962 by Copp & Cheney [109]. The sequencing of the porcine calcitonin in 1968 revealed that this polypeptide is composed of 32 amino acids [110]. In vertebrates, calcitonin is produced by the thyroid gland [111] and regulates calcium (Ca^{2+}) levels in the blood, antagonizing the effects of parathyroid hormone [112,113]. The evolutionary antiquity of calcitonin-related peptides is further evidenced by the discovery that this polypeptide is composed of 32 amino acids [110].

### 2.4.4. GPA2 and GPB5

The vertebrate glycoprotein hormone family comprises luteinizing hormone follicle-stimulating hormone, chorionic gonadotropin, TSH and the recently discovered thyrostimulin (TS) [122,123]. TS is a heterodimer composed of two subunits, glycoprotein alpha 2 (GPA2) and glycoprotein beta 5 (GPB5). Orthologues of GPA2 and GPB5 have been identified and characterized in the insect D. melanogaster [124] and in other invertebrates, including echinoderms [125]. Insect GPA2 and GPB5 both contain 10 conserved cysteine residues that are important in forming a heterodimeric cysteine-knot structure. Surprisingly, A. japonicus GPA2 contains only seven cysteine residues (having lost residues 7, 8 and 9), whereas O. victoriae GPA2 contains 10 conserved cysteine residues (having lost the final two cysteine residues) (electronic supplementary material, figure S5). It is difficult to predict the structural differences that may arise in the heterodimer due to this variability in the number of cysteine residues. The possibility of GPA2 and/or GPB5 monomers or homodimers exerting their own biological functions has not been ruled out [126]. Additional investigations are needed to investigate if GPA2 and GPB5 are co-localized in echinoderms and if the monomers and dimers (both homo and hetero) exert different effects.

### 2.5. Uncharacterized neuropeptides

In addition to the neuropeptides discussed above, we have also identified three neuropeptide precursors that could not be classified into any known neuropeptide families. These include O. victoriae neuropeptide precursor (Ovnp) 18 (O. victoriae orthologue of Spnp18 in S. purpuratus) [16], Ovnp26 and Ovnp27, with the latter two identified for the first time in echinoderms. The choice of nomenclature for Ovnp26 and Ovnp27 is based on a previously used numerical nomenclature in S. purpuratus and/or A. rubens, which goes up to Arnp25 in A. rubens.

#### 2.5.1. Ovnp18

Ovnp18 comprises four copies of a predicted mature peptide with the sequence LFVWD and the C-terminal region of the precursor (partial sequence) contains at least four cysteine residues (figure 5f; GenBank MF155258). Interestingly, this precursor type only comprises a single mature peptide in A. rubens, S. purpuratus and A. japonicus, and the C-terminal region contains 9, 8 and 8 cysteine residues, respectively (data not shown) [8,15,16].
2.5.2. Ophionotus victoriae neuropeptide precursor 26

Ovnp26 was identified following an analysis of O. victoriae transcriptome sequence using NpSearch [8]. Orthologues of Ovnp26 (GenBank: MF155259) were identified in other brittle stars but not in other echinoderms (electronic supplementary material, figures S2–S4). Ovnp26 comprises seven copies of peptides with a conserved C-terminal GW motif, whereas
orthologues in *O. annae* and *A. filiformis* are predicted to generate eight copies of the mature peptide. Some of the mature peptides have a C-terminal SGW motif, which is similar to the C-terminus of predicted mature peptides derived from *O. victoriae* pedal peptide precursor 3 (electronic supplementary material, figure S7). However, the lack of sequence similarity in other parts of the peptide suggests that the C-terminal similarity may reflect convergence rather than homology.

### 2.5.3 *Ophionotus victoriae* neuropeptide precursor 27

Ovpn27 (GenBank: MF155251) was identified following a Hidden Markov Model (HMM)-based search for SIFamide-type peptides [127,128], albeit with a high E-value. This neuropeptide precursor comprises two putative amidated mature peptides that are located immediately after the signal peptide (electronic supplementary material, figures S2–S4), as seen in SIFamide precursors [129]. The first peptide of the *O. victoriae* precursor has a C-terminal IFamide motif just like in insect SIFamides (electronic supplementary material, figure S9). However, there is no sequence similarity with SIFamides in the rest of the peptide. This coupled with the fact that SIFamide-type receptors have not been identified in echinoderms [6] suggests that the sequence similarity that peptides derived from Ovpn27-type precursors share with SIFamides may reflect convergence rather than homology.

### 2.6. Neuropeptide precursors not found in brittle stars

Our analysis of ophiuroid transcriptome sequence data did not reveal orthologues of the Spnp9 precursor from *S. purpuratus* or the AmpR21, AmpR22, AmpR23 and AmpR24 precursors from *A. rubens* [8,16]. An Spnp9 orthologue is found in *A. japonicus* but not in *A. rubens* [15]; therefore, this neuropeptide precursor type may be restricted to the Echinozoa. Orthologues of AmpR21–24 have not been found in *O. victoriae, S. purpuratus* or *A. japonicus*, which suggest that these may be Asteroidea-specific precursors.

Previous studies have shown that receptors for leucokinin, ETH, QRFP, parathyroid hormone, galanin/allatostatins-A and Neuromedin-U/CAPA are present in ambulacaria [6,15]. The presence of these receptors suggests that their cognate ligands should also be present in ambulacaria. However, our search approaches failed to identify any proteins in ophiuroids that resemble precursors of these neuropeptides.

### 2.7. Evolutionary conservation and variation of neuropeptide copy number in the Ophiuroidea

Many neuropeptide precursors comprise several structurally similar but non-identical bioactive peptides (i.e. the precursor protein gives rise to a neuropeptide ‘cocktail’). This feature of neuropeptide precursors occurs throughout metazoans. But how do these ‘cocktails’ of neuropeptides evolve and what is their functional significance? Are the copies of mature peptides functionally redundant or do they have their own specific functions? These are important questions in neuroendocrinology for which answers remain elusive.

Evidence that neuropeptide copy number may be functionally important has been obtained from comparison of the sequences of neuropeptide precursors in 12 *Drosophila* species, the common ancestor of which dates back approximately 50 Myr [130]. The number of peptide copies in each neuropeptide precursor was found to be identical (except for the FMRFamide precursor) when compared between the 12 species, suggesting that stabilizing selection has acted to conserve neuropeptide ‘cocktails’ in the *Drosophila* lineage.

Here, a comparison of *O. victoriae, A. filiformis* and *O. annae* neuropeptide precursors and their putative mature peptides revealed that 14 neuropeptide precursors comprised multiple neuropeptide copies. In certain cases, the number of the mature peptides derived from a particular precursor varied across species, whereas in other cases the numbers remained constant (figure 4). Interestingly, these three species belong to two of the three major clades of brittle stars that evolved approximately 270 Ma [12]. While *O. victoriae* belongs to the order Ophiurida (in clade A), *A. filiformis* and *O. annae* belong to the order Amphilepidida (in clade C). Hence, this prompted us to examine the evolution of neuropeptides and neuropeptide copy number variation at a higher level of phylogenetic resolution. To do this, we used a unique dataset comprising 52 ophiuroid transcriptomes. These transcriptomes were recently used as part of a phylotranscriptomic approach to reconstruct the phylogeny of ophiuroids, generating a robust phylogenetic tree that comprises three major clades [12]. Hence, this dataset allowed us to explore the evolution of neuropeptide precursors in the context of an established phylogenetic framework spanning over an unprecedented time scale of approximately 270 Myr.

We selected for analysis neuropeptide precursors comprising more than one putative mature neuropeptide, which include AN peptide, calcitonin, CCK1, KP, np18, np26, np27, NG peptide, PDF, SALMFamide (L-type and F-type), tachykinin and TRH (1 and 2). Pedal peptide precursors (1, 2 and 3) were excluded from the analysis because orthology relationships between these precursors could not be established with confidence across all species (data not shown). We used *O. victoriae* representatives of these neuropeptide precursor families and the *A. filiformis* AN peptide precursor to mine 52 ophiuroid transcriptomes using BLAST. Multiple sequence alignments were generated based on the search hits (electronic supplementary material, figure S11) and the number of predicted mature peptides was compared (figure 8). Interestingly, the number of peptides within the majority of precursors remained constant across all the species examined, which share a common ancestor estimated to date from approximately 270 Ma [12].

Some studies that have investigated the physiological significance of neuropeptide ‘cocktails’ indicate that neuropeptides derived from the same precursor protein are functionally redundant. For example, this was found for myomodulin neuropeptides in the mollusc *Aplysia californica* using the accessory radula closer muscle preparation as a bioassay [131], and for FMRFamide-related neuropeptides in *D. melanogaster* when analysing effects on nerve-stimulated contraction of larval body wall muscles [132]. However, the authors of the latter study cautiously highlighted the need to ‘search for additional functions or processes in which these peptides may act differentially’. Importantly, studies employing the use of multiple bioassays have obtained data indicating that neuropeptides derived from a single precursor protein are not functionally redundant. For example, when the actions of 14 structurally related neuropeptides derived from a precursor of *Mytilus* inhibitory peptide-related peptides in *Aplysia* were tested on...
three organ preparations (oesophagus, penis retractor and body wall), it was found that the rank order of potency for the peptides differed between preparations [133]. Similarly, when assaying the effects of allatostatin neuropeptides in cockroaches, tissue-specific differences in potency were observed [134]. The conservation of peptide copy number across a time scale of approximately 270 Myr in the Ophiuroidea supports the idea that the occurrence of multiple copies of identical or structurally related neuropeptides is functionally important.

For those neuropeptide precursors that did exhibit variation in neuropeptide copy number, TRH-type precursors exhibited the highest variation, with numbers ranging from 16 to 20 copies (figure 9). F-type SALMFamide precursors also showed variation in copy numbers (figure 10), but loss of peptides was more frequent in F-type SALMFamide precursors than in TRH-type precursors. Furthermore, detailed analysis of sequence alignments for these precursors revealed that loss of neuropeptide copies is usually a consequence of non-synonymous mutations in codons for residues that form dibasic cleavage sites or for glycine residues that are substrates for the C-terminal amidation. This is not surprising because the C-terminal amide in smaller-sized peptides is usually important for receptor binding and activation. What is unclear at the moment is how the peptide copy number increases within a given precursor. Perhaps the increase in peptide copy number occurs as a result of unequal crossing-over during recombination [130].

The number of peptides within the F-type SALMFamide precursors appears to be clade-specific. Thus, the average/median number of F-type SALMFamides in precursors from clade A is 13, clade B is 12 and clade C is 11, with a few exceptions (figure 8). Similarly, the number of peptides within NP26-type precursors also appears to be clade-specific. Hence, the number of peptides is highly stable at seven peptides within clades A and B, but a high variation in peptide copy number is observed in clade C. When examining peptide copy number within clades, there are a few cases where the number of peptides within a given precursor for certain species appears to be an exception/outlier. For instance, 16 copies of the mature peptide in Ophioplax lamellosa TRH-1 precursor...
are distinctly different from the 19 copies found in other species within that clade (clade C). Likewise, Ophiactis savignyi only has three copies of KP-type peptides compared with four copies found in other species of that clade (figure 8).

It could be argued that misalignments during transcriptome assembly may have influenced the number of predicted peptides found in a given precursor. However, it is unlikely that misalignments have affected the predicted sequences of neuropeptide precursors comprising multiple copies of peptides that are similar but non-identical, which applies to the majority of the precursor proteins analysed here in ophiuroids. The only exception to this are the TRH-type precursors, where the encoded peptide sequences are short and often identical, even at the nucleotide level (data not shown). Another limitation of using transcriptome data is that the sequences of neuropeptide precursors may be partial or unknown for some species, and where this applies a peptide copy number is not shown in figure 8. An extreme example of this is the AN peptide precursor, where complete precursor sequences were only obtained from the three reference species and three other species. However, for the majority of precursor types, sequence data were obtained from a variety of species from each of the three clades of ophiuroids. For example, complete F-type SALMFamide precursor sequences were found in most of the investigated species (39 species + 3 reference species).

3. Conclusion

Here, we report the first detailed analysis of the neuropeptide precursor complement of ophiuroids and the most comprehensive identification of echinoderm neuropeptide precursors to date. We have identified novel representatives of several bilaterian neuropeptide families in echinoderms for the first time, which include orthologues of endothelin/CCHamide, EH, NPF/NPY and NUCB/nesfatin. Furthermore, analysis of precursor proteins comprising multiple copies of identical or related neuropeptides across approximately 270 Myr of ophiuroid evolution indicates that the precise composition of neuropeptide ‘cocktails’ is functionally important as evident from the conservation of neuropeptide copy number for multiple precursors.

4. Material and methods

4.1. Sequencing and assembly of transcriptomes

Ophiuroid transcriptomes used in this study were sequenced and assembled as reported previously [12,20,24].
4.2. Identification of neuropeptide precursors in ophiuroids

To identify neuropeptide precursors in *O. victoriae*, *A. filiformis* and *O. aranea*, sequences of neuropeptide precursors identified previously in other echinoderms (including the starfish, *A. rubens*, the sea urchin *S. purpuratus* and the sea cucumber, *A. japonicus*) were used as queries for tBLASTn analysis of a transcriptome database, using an *e*-value of 1000. Sequences identified as potential neuropeptide precursors by BLAST were translated using the ExPasy Translate tool (http://web.expasy.org/translate/) and then analysed for features of neuropeptide precursors. Specifically, sequences were evaluated based on (i) the presence of an N-terminal signal peptide (using SIGNAL P v. 4.1 with the sensitive cut-off of 0.34) and (ii) the presence of monobasic or dibasic cleavage sites flanking the putative bioactive peptide(s).

To identify novel neuropeptide precursors or highly divergent precursors with low sequence similarity to known precursors, we used two additional approaches. In the first approach, we used NiSEARCH [8], software that identifies putative neuropeptide precursors based on various characteristics (presence of signal peptide and dibasic cleavage sites among others). In the second approach, NpHMMER (http://nphmmer.sbc.soton.ac.uk/), an HMM-based software, was used to identify neuropeptides not found using the above approaches.

Neuropeptide precursors identified in *O. victoriae* (which represented a more comprehensive neuropeptide precursor repertoire compared to *A. filiformis* and *O. aranea*) were then submitted as queries for BLAST analysis of sequence data from 52 Ophiuroidea species, using an *E*-value of 1 × 10^−6*. BLAST hits were then further analysed using an automated ruby script (available at https://github.com/IsmailM/ophiuroid_neuropeptidome). Each BLAST hit was translated using the ExPasy Translate tool (http://web.expasy.org/translate/) and then analysed for features of neuropeptide precursors. Specifically, sequences were evaluated based on (i) the presence of an N-terminal signal peptide (using SIGNAL P v. 4.1 with the sensitive cut-off of 0.34) and (ii) the presence of monobasic or dibasic cleavage sites flanking the putative bioactive peptide(s).

4.3. Phylogenetic and clustering analyses of sequence data

Phylogenetic analysis of membrane guanylyl cyclase receptors and nucleobindins was performed using maximum-likelihood and Bayesian methods. Prior to these analyses, corresponding multiple alignments were trimmed using BMGE [135] with the following options: BLOSUM30, max − *h* = 1, −*b* = 1, as described previously [10,94]. The maximum-likelihood method was implemented in the PhyML program (v. 3.1/3.0 aLRT). The WAG substitution model was selected assuming an estimated proportion of invariant sites of (0.112) and four gamma-distributed rate categories to account for rate heterogeneity across sites. The gamma shape parameter was estimated directly from the data. Reliability for internal branch was assessed using the bootstrap method (500 bootstrap replicates). The Bayesian inference method was implemented in the MrBayes program (v. 3.2.3). The number of substitution types was fixed to 6. The Poisson model was used for substitution, while rate variation across sites was fixed to ‘invgamma’. Four Markov chain Monte Carlo (MCMC) chains were run for 100 000 generations, sampling every 100 generations, with the first 500 sampled trees discarded as ‘burn-in’. Finally, a 50% majority rule consensus tree was constructed.

CLANs analysis was performed on echinoderm EH-like, arthropod EH, arthropod ITP and vertebrates ANP precursors based on all-against-all sequence similarity (BLAST searches) using the BLOSUM 45 matrix (https://toolkit.tuebingen.mpg.de/clans/) [42] and the significant HSPs. Neuropeptide precursors were clustered in a three-dimensional graph represented here in two dimensions.

Data accessibility. Raw sequence data used to assemble the transcriptomes have been deposited in the NCBI Sequence Read Archive (SRA) under the accession no. SRP107914 (https://www.ncbi.nlm.nih.gov/sra/?term=SRP107914) and in the NCBI BioProject under the accession no. PRJNA311384 (https://www.ncbi.nlm.nih.gov/bioproject/311384).

Authors’ contributions. M.Z., T.D.O. and M.R.E. designed the research; I.M. lyised the data; M.Z., J.D. and M.R.E. wrote the manuscript with input from other authors. M.Z. and M.R.E. supervised the study.

Competing interests. We declare we have no competing interests.

Funding. This work was supported by Leverhulme Trust grant no. (RPG-2013-351) and a BBSRC grant no. (BB/M001644/1) awarded to M.R.E. L.A.Y.G. is supported by a PhD studentship awarded by CONACYT (studentship no. 418612). I.M. is supported by the Biotechnology and Biological Sciences Research Council (grant no. BB/M009513/1).

Acknowledgement. The authors acknowledge Zuzuia Waffa and Giulia Oluwabunmi Olayemi for their assistance with sequence alignments.

References


