






Toxin-like neuropeptides in the sea anemone *Nematostella* unravel recruitment from the nervous system to venom

Maria Y. Sachkova^{a,b,1}, Morani Landau^{a,2}, Joachim M. Surm^{a,2}, Jason Macrander^{c,d} , Shir A. Singer^a , Adam M. Reitzel^c, and Yehu Moran^{a,1} 

^aDepartment of Ecology, Evolution, and Behavior, Alexander Silberman Institute of Life Sciences, Faculty of Science, Hebrew University of Jerusalem, 9190401 Jerusalem, Israel; ^bSars International Centre for Marine Molecular Biology, University of Bergen, 5007 Bergen, Norway; ^cDepartment of Biological Sciences, University of North Carolina at Charlotte, Charlotte, NC 28223; and ^dBiology Department, Florida Southern College, Lakeland, FL 33801

Edited by Baldomero M. Olivera, University of Utah, Salt Lake City, UT, and approved September 14, 2020 (received for review May 31, 2020)

The sea anemone *Nematostella vectensis* (Anthozoa, Cnidaria) is a powerful model for characterizing the evolution of genes functioning in venom and nervous systems. Although venom has evolved independently numerous times in animals, the evolutionary origin of many toxins remains unknown. In this work, we pinpoint an ancestral gene giving rise to a new toxin and functionally characterize both genes in the same species. Thus, we report a case of protein recruitment from the cnidarian nervous to venom system. The ShK-like1 peptide has a ShKT cysteine motif, is lethal for fish larvae and packaged into nematocysts, the cnidarian venom-producing stinging capsules. Thus, ShK-like1 is a toxic venom component. Its paralog, ShK-like2, is a neuropeptide localized to neurons and is involved in development. Both peptides exhibit similarities in their functional activities: They provoke contraction in *Nematostella* polyps and are toxic to fish. Because ShK-like2 but not ShK-like1 is conserved throughout sea anemone phylogeny, we conclude that the two paralogs originated due to a *Nematostella*-specific duplication of a ShK-like2 ancestor, a neuropeptide-encoding gene, followed by diversification and partial functional specialization. ShK-like2 is represented by two gene isoforms controlled by alternative promoters conferring regulatory flexibility throughout development. Additionally, we characterized the expression patterns of four other peptides with structural similarities to studied venom components and revealed their unexpected neuronal localization. Thus, we employed genomics, transcriptomics, and functional approaches to reveal one venom component, five neuropeptides with two different cysteine motifs, and an evolutionary pathway from nervous to venom system in Cnidaria.

venom evolution | toxin recruitment | neuropeptide | nematocyst | neuron

The starlet sea anemone *Nematostella vectensis* (hereafter, *Nematostella*) has developed into a model system for characterizing the evolution of genes and their functions. It has a noncentralized nervous system, which comprises morphologically distinguishable sensory, ganglion, and nematocyte neuronal types that develop from the same neural progenitor cells (1). From a transcriptomic perspective, however, several dozen neuronal subtypes were identified (2). Unlike in bilaterian animals where neurons differentiate from and are located in the ectoderm, the nervous system of *Nematostella* spans both the ectoderm and endoderm (3, 4). Ganglion cells are found in both germ layers while sensory cells and nematocytes are present only in the ectoderm. Nematocytes are highly specialized cells bearing a nematocyst, an organelle loaded with proteinaceous venom that can be discharged upon mechanical stimulation (5–8). Thus, nematocytes function as microsyringes injecting venom into prey and predators.

Sea anemone venom is a complex mixture of toxins. Along with nematocytes, ectodermal gland cells contribute to venom production (9, 10). A significant portion of the venom consists of disulfide-rich peptides of ~4-kDa molecular mass (11). Many of these peptides possess neurotoxic activity and act through binding

to a target receptor in the nervous system of the prey or predator interfering with transmission of electric impulses. For example, Nv1 toxin from *Nematostella* inhibits inactivation of arthropod sodium channels (12), while ShK toxin from *Stichodactyla helianthus* is a potassium channel blocker (13). *Nematostella*'s nematocytes produce multiple toxins with a 6-cysteine pattern of the ShK toxin (7, 9). The ShKT superfamily is ubiquitous across sea anemones (14); however, its evolutionary origin remains unknown. Comparative phylogenetic analyses of various venom components suggest several have originated through a process of gene duplication and recruitment from other physiological systems, such as the digestive, immune, and nervous systems (15). Additionally, new toxins may originate due to divergence from other toxins following the recruitment to venom (16) or de novo (17).

While the general mechanism of toxin recruitment from other cells or tissues has been suggested (15), limited studies to date provide evidence for the function of both the ancestral physiological protein and recruited venom protein. Instead, homology is typically inferred based exclusively on sequence similarity. Furthermore, these descriptions are generally restricted to analyses at the protein superfamily level, which greatly limits our ability to reconstruct the exact evolutionary pathway for venom

Significance

Although venom has evolved independently numerous times in animals, the origin of many proteins composing venom remains poorly characterized. We report a case of protein recruitment from the nervous to venom system in the sea anemone *Nematostella*. We discovered a toxin packaged in the stinging cells and with lethal activity against fish. Surprisingly, it is similar to a peptide from the nervous system (neuropeptide) regulating *Nematostella* development. Comparisons with other sea anemones showed that these two functionally different peptides originated due to a duplication of a neuropeptide gene followed by recruitment of one of the copies into venom. Additionally, we characterized four other toxin-like neuropeptides, suggesting a recurring pattern of venom recruitment from neuropeptides in sea anemones.

Author contributions: M.Y.S., A.M.R., and Y.M. designed research; M.Y.S., M.L., J.M.S., J.M., and S.A.S. performed research; M.Y.S., J.M.S., and J.M. analyzed data; and M.Y.S. and Y.M. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Published under the PNAS license.

¹To whom correspondence may be addressed. Email: mariasachkova@gmail.com or yehu.moran@mail.huji.ac.il.

²M.L. and J.M.S. contributed equally to this work.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2011120117/-DCSupplemental>.

First published October 15, 2020.

genes due to their high sequence divergence. This is further confounded with some of the shared cysteine frameworks that are potentially recruited to venom and other physiological systems independently or even possibly originating de novo (18). Thus, the origin for many animal toxins (for example, arthropod inhibitory cysteine knot toxins, sea anemone sodium and potassium neurotoxins, conotoxins with multiple cysteine frameworks, and so forth), particularly for understudied groups like cnidarians, remains relatively unknown (11, 15).

Cnidarian neurons produce multiple neuropeptides (4, 19, 20) that act as neurotransmitters and neuromodulators. Thirty putative neuropeptide sequences belonging to families widely represented in metazoans (such as RFamide, vasopressin, galanin, tachykinin peptides) as well as cnidarian-specific families (for example, Antho-RIamide, RNamide, RPamide peptides) have been identified in the *Nematostella* genome (21). All of the reported mature sequences are relatively short (3 to 15 residues) with a C-terminal amidation signal (Gly residue) and absence of Cys residues (apart from two RPamide peptides). Recently, an approach based on tandem mass spectrometry and a novel bioinformatics pipeline identified 20 new neuropeptide sequences in *Nematostella* (22). Some of the new peptides were classified as RFamide and RPamide peptides, which were reported earlier, while others did not share similarity to any known cnidarian neuropeptides. While disulfide-rich neuropeptides have not been reported in cnidarians, preproinsulin-like peptide genes have been characterized in *Nematostella* (23). These insulin-like peptides are disulfide-rich and expressed in exocrine gland cells, with a single member also expressed in small cells adjacent to neurons. Peptides from the insulin superfamily have also been reported in the neuroendocrine system of other animals (24–26).

Cnidarian neuropeptides affect diverse physiological processes, such as muscle contraction, larval migration, metamorphosis, and neural differentiation (19, 20). *Nematostella* GLWamide neuropeptide regulates timing of transition from planula to polyp (27). Most of the studied metazoan neuropeptides act through binding to metabotropic receptors, such as peptide gated G protein-coupled receptors (28); however, RFamides from another cnidarian (*Hydra*) directly activate ion channels from the degenerin/epithelial Na⁺ channel family (4).

Venom components and neuropeptides are secreted molecules. Typically, secreted peptides are synthesized as precursor proteins equipped with a signal peptide leading to cotranslational translocation of the precursor into the endoplasmic reticulum (29). In cnidarians, the precursor also typically includes a propeptide that is cleaved off by maturation enzymes upon processing (30). Mature peptides are transported to their secretion sites, which can be a neuronal synapse in case of neuropeptides or a nematocyst in case of venom components.

Here we report two classes of neuropeptides with structural similarity to disulfide-rich venom components. We have characterized the paralogs expressed in venom and nervous systems at the genomic, transcriptomic, and functional levels. Our work reveals a first case of protein recruitment from the nervous to venom system in Cnidaria.

Results

New Toxin-like Peptides. The *Nematostella* transcriptome remains incompletely annotated; therefore, we expanded our search to identify all potential toxins from this species. We used toxin sequences deposited in UniProt (ToxProt) as a query to identify toxin-like sequences in *Nematostella* using the previously published Transcriptome Shotgun Assembly (TSA) dataset from mesenteries, nematosomes, and tentacles (PRJEB13676). This search identified Class8-like peptide, a homolog of Avd8e toxin (e-value 7×10^{-18}), and Aeq5-like1, a homolog of Acrorhagin 5a from *Actinia equina* (e-value 0.005) (31). Using Transdecoder (32), we translated the TSA sequences and obtained 1,126,514

protein sequences, with 653 sequences with the predictive ShK-like cysteine arrangement, 74 of which were predicted to have a signal region. After screening for multiple ShK-like domains, excessive cysteine residues, and sorting through ToxProt BLAST outputs, we identified transcripts encoding six uncharacterized proteins with a cysteine arrangement typical of the potassium channel blocker from *S. helianthus* known as ShK (13) (ShK-like1, ShK-like2a, ShK-like2b, ShK-like3, ShK-like4, and Class8-like) (Fig. 1A and Table 1). Additionally, we examined the *Nematostella* single-cell RNA-sequencing (RNA-seq) data (2) where one gene, *Nve8041*, was expressed in neurons and coded for a protein with the same Cys pattern as Aeq5a and Aeq5-like1; this we designated as Aeq5-like2 (Fig. 1B and Table 1).

All of the identified proteins, with the exception of Aeq5-like2, possess structural features typical of sea anemone toxin precursor proteins: A signal peptide (identified by SignalP) (35), a propeptide with a dibasic processing motif (apart from Aeq5-like1), followed by a Cys-rich mature peptide (36). ShK-like2a and ShK-like2b differ in signal peptides and a portion of the propeptides, while their mature peptides are identical (and is therefore referred to as ShK-like2). ShK-like1, ShK-like2a/b, and Shk-like3 are likely closely related homologs as their mature peptides share around 50% identical residues (Fig. 1C). Signal peptides and propeptides are more diverged as their overall identity level is 36 to 40%. Aeq5-like2 has slightly different structure; instead of a signal peptide it has a transmembrane domain on its N terminus, indicating that the disulfide-rich domain is anchored on the extracellular side of the cell membrane (InterProScan prediction) (33). Furthermore, the validity of these identified transcripts as protein-coding genes was confirmed at the protein level with tandem mass spectrometry (LC-MS/MS) data (9, 37) detecting tryptic peptides of mature ShK-like1, ShK-like2, and Aeq5-like1 at multiple life-stages in *Nematostella* (SI Appendix, Figs. S1B and S2 and Table S1). For Shk-like4, a tryptic peptide mapping to the propeptide region was identified as the mature peptide lacks suitable trypsin cleavage sites.

ShK-like1 biosynthesis dynamics differed from the other proteins: it was the most abundant in larvae and significantly decreased later in the life cycle, according to transcriptomic (38) and proteomic (9) data (SI Appendix, Fig. S1). In contrast, all of the other genes apart from *ShK-like3* were expressed early in the life cycle and maintained expression through the adult stage. For *ShK-like3*, however, no reads mapped to the transcript in the NvERTx database (38) (ID: NvERTx.4.119933).

ShK-like1 Is a Venom Component. We used whole-mount in situ hybridization (ISH) in *Nematostella* planulae and primary polyps to determine the expression patterns of the discovered toxin-like genes; the experiment was performed twice with at least 50 animals per sample. Surprisingly, only *ShK-like1* was expressed in nematocytes of planulae and primary polyps (Fig. 2A). At the transcriptomic level [NvERTx database (38)], *ShK-like1* expression in planulae was comparable to the previously characterized Nep3 toxin (9). In the primary polyps, *ShK-like1* was localized predominantly in the body column nematocytes and not in the tentacles.

Notably, in the single-cell RNA-seq data generated by others (2), the NVE19767 gene model that covers part of the *ShK-like1* locus was specifically up-regulated in the C33 and C34 metacells, annotated as larval nematocytes. This fact provides independent support for the nematocyte expression of *ShK-like1*.

To confirm that ShK-like1 protein is packaged into nematocysts and thus is a venom component, we analyzed soluble protein content of nematocysts by LC-MS/MS. Because abundance of the ShK-like1 protein is higher in larvae than in adults by more than one order-of-magnitude (SI Appendix, Fig. S1B), we used capsules from 4 d postfertilization (dpf) planulae; 138 proteins with at least 2 unique tryptic peptides were identified (Dataset S1; (37)). ShK-like1 (Maxquant score 59.7, four unique

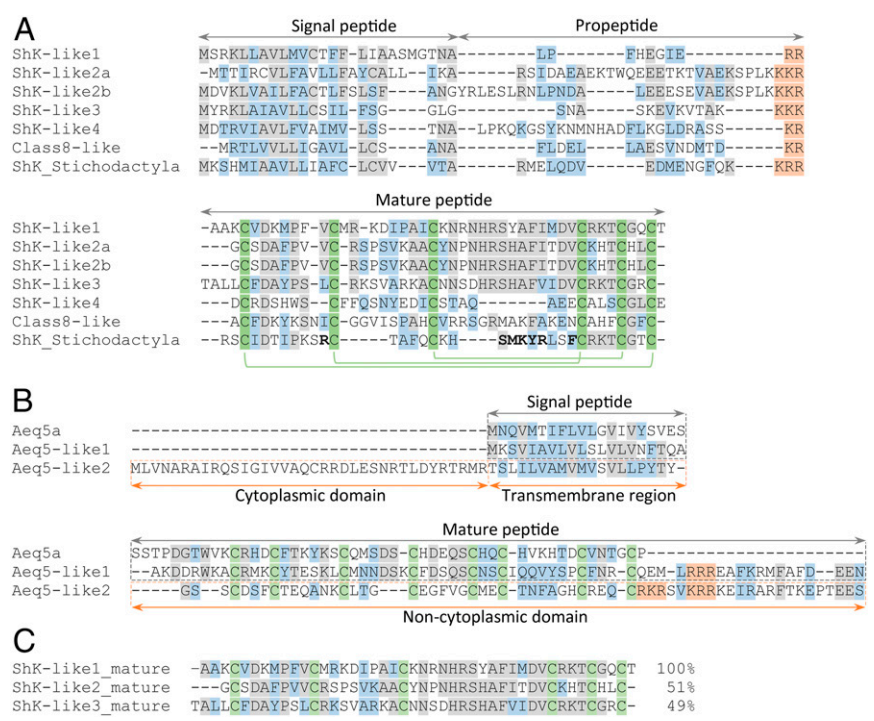


Fig. 1. Alignment of the toxin-like precursor sequences from *Nematostella* and characterized toxins ShKT from *S. helianthus* (A) and Aeq5a from *A. equina* (B). The sequences were split into signal peptide, propeptide, and mature peptide fragments and corresponding fragments were aligned separately using Muscle algorithm implemented in MEGA X and manually adjusted. Processing sites are in orange, cysteine residues making part of the cysteine motifs of the mature peptides are in green. Identical residues are in gray, conserved substitutions are in blue. (A) Sequences with ShKT cysteine motif, disulphide bonds of ShK toxin are shown as green brackets. The ShK residues important for blocking potassium channels are in bold (33). (B) Sequences with the same motif as in Aeq5a. (C) Alignment of the ShK-like1, ShK-like2, and ShK-like3 mature peptides.

peptides) (*SI Appendix, Fig. S3*), along with other previously characterized nematocyst toxins (Nep3, Nep3-like, Nep4, Nep6, Nep8-like, Nep16) (7, 9) and nematocyst structural proteins, were among the 30 most abundant proteins (Fig. 2B and *SI Appendix, Table S2*).

The mature ShK-like1 peptide was produced as a recombinant peptide (*SI Appendix, Fig. S4*). Electrospray ionization (ESI)-MS showed that molecular mass of the peptide (5105.42) corresponded to the calculated molecular weight of the isoform with all of the six cysteine residues oxidized to form three disulfide bonds (5105.14) (*SI Appendix, Fig. S4B*). To study the toxic effects of ShK-like1, the peptide was tested on zebrafish larvae: Five larvae were incubated with 0.5 mg/mL ShK-like1 with four replicates. After 1-h incubation, all of the larvae (20 in total) exhibited contraction paralysis and within 2 h they died, while all of the larvae in BSA control survived (Fig. 2C). These results are evidence of neurotoxic activity of ShK-like1. Alignment to ShK toxin from *S. helianthus* (Fig. 1A) shows that key residues required for blocking vertebrate potassium channels are missing in

ShK-like1, suggesting that it probably acts through a different mechanism (39). Thus, ShK-like1 is a nematocyst toxin highly expressed at the planula, for which the neurotoxic target is yet to be determined.

Expression of Toxin-like Genes in Neurons. In *Nematostella*, ganglion and sensory neurons can be recognized by their morphological features: ganglion neurons have two to four neural processes, while sensory cells have elongated shape, possess processes at the base, and may have a cilium at the apex (40). ISH showed that *ShK-like2a/b*, *ShK-like4*, *Class8-like*, *Aeq5-like1*, and *Aeq5-like2* were localized in various neurons in both planula and primary polyp stages but not in nematocytes or other venom-related cells, despite their structural similarity to toxins (Figs. 3–7). All of the experiments were repeated at least two times with at least 50 animals per sample.

Aeq5-like1 expression was first detected early in development in ectodermal cells of the gastrula, then increased in early planula and by the late planula had spread in expression to

Table 1. New toxin-like peptides

Name	TSA ID/Nve gene model	Homolog in ToxProt	Activity of the ToxProt homolog	Percent identity, % (ClustalO)	Cys motif
ShK-like1	HADP01047794.1/NA	<i>A. viridis</i> Avd9a	Toxin by similarity	30	
ShK-like2a	HADN01006514.1/NA	<i>A. viridis</i> Avd9a	Toxin by similarity	24	
ShK-like2b	HADO01001456.1/Nve11386	<i>A. viridis</i> Avd9a	Toxin by similarity	25	
ShK-like3	HADN01040061.1/NA	<i>A. viridis</i> Avd9c	Toxin by similarity	34	
ShK-like4	HADO01004651.1/Nve12327	<i>A. viridis</i> Avd9b	Toxin by similarity	32	
Class8-like	HADO01000486.1/Nve7851	<i>A. viridis</i> Avd8e	Toxin by similarity	51	ShKT
Aeq5-like1	HADO01002561.1/Nve8064	<i>A. equina</i> Aeq5a	Extracted from acrorhagi, toxic for crabs	39	
Aeq5-like2	Nve8041	—	—	—	Aeq5a

Gene models are according to D. Fredman et al. (34).

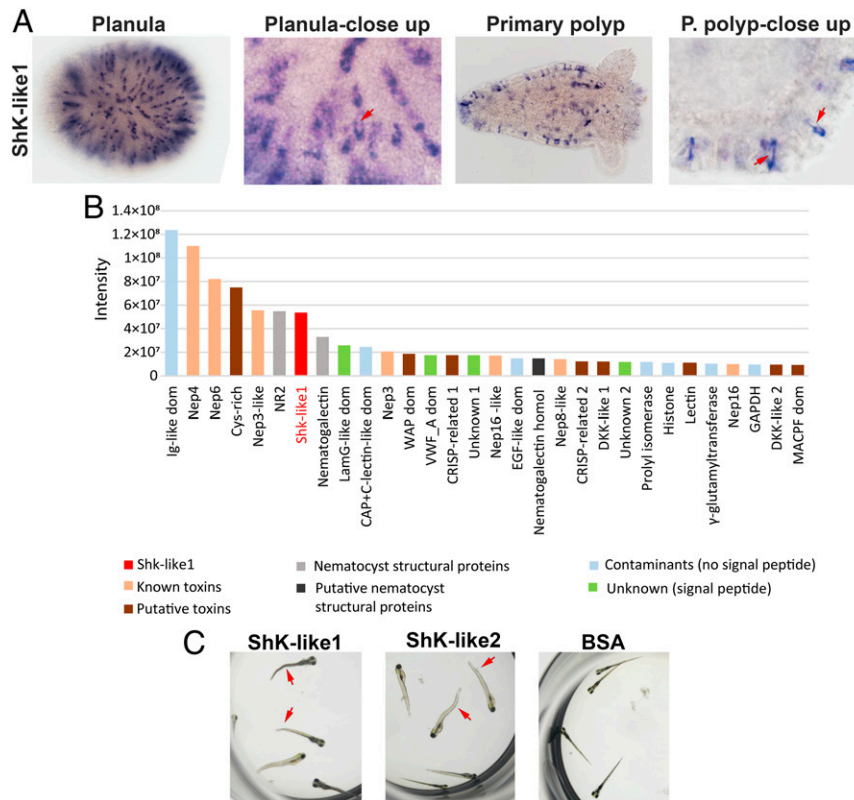


Fig. 2. ShK-like1 is a venom component. (A) *ShK-like1* gene is expressed in nematocytes in planula and primary polyps; note that the staining is distributed to the periphery of the cells and the unstained center of the cells corresponds to capsules (arrows). Magnifications: 200× (whole animals), 600× (close up). (B) The 30 most abundant proteins extracted from planula nematocysts identified by MS/MS. (C) Toxicity of recombinant ShK-like1 and ShK-like2 was tested on zebrafish larvae in parallel to BSA control. Arrows point to kinked tails evidencing for the contraction paralysis effect of the peptides.

endodermal cells (Fig. 3-I). Early ectodermal cells may represent neuronal precursors for sensory cells. In primary polyps, *Aeq5-like1* was expressed in ectodermal sensory cells and multiple endodermal ganglion cells that send their processes in different directions all around the body. *ShK-like4* was expressed in ectodermal sensory cells in planulae and oral part of polyps (Fig. 3-II). Some of the sensory cells appeared to have cilia (Fig. 3-II B and F) and processes at their base (Fig. 3-II G and H). *Class8-like* was expressed in multipolar ganglion cells with long branching processes localized in mesoglea (Fig. 4-I). *Aeq5-like2* was localized to endodermal ganglion neurons, apparently bipolar and following mesentery folds (Fig. 4-II). To the best of our knowledge these patterns are quite unique as in most cases, unlike protein-based staining, RNA-based staining of *Nematostella* neurons does not reveal neurite like structures (40, 42).

For *ShK-like2* ISH, we first synthesized an RNA probe that was relatively long (ShK-like2–long, 603 nt) and specific to the full-length *ShK-like2b*, which also shared significant portion of the sequence (66%) with the *ShK-like2a* mRNA (Fig. 5-I and *SI Appendix, Table S2*). The resulting pattern allowed us to conclude that *ShK-like2* was predominantly expressed in ectodermal sensory neurons in planulae and endodermal ganglion cells in primary polyps. To determine whether *ShK-like2a* and *ShK-like2b* exhibit different expression patterns, we produced probes that were specific for each isoform. Due to relatively short coding sequences, the *ShK-like2a* probe (ShK-like2a–short, 372 nt) included 84% of specific sequence while the *ShK-like2b* probe (ShK-like2b–short, 332 nt) possessed 62% of specific sequence with the rest of the sequence shared between the two isoforms (*SI Appendix, Table S2*). ISH with the specific probes showed that *ShK-like2b* was first expressed in the early planula in ectodermal cells and this pattern

persisted through the late planula stage, while *ShK-like2a* was first expressed mostly in the endoderm (with rare ectodermal cells) in the late planula. In primary polyps, expression patterns were indistinguishable: both isoforms were highly expressed in endodermal ganglion cells in the body column and tentacles with rare sensory ectodermal cells (Fig. 5-II and 5-III). This pattern was confirmed by significant overlap between stained cells in double fluorescent ISH (dFISH), especially in the tentacle and aboral endoderm (*SI Appendix, Fig. S5*). ISH with a probe for *ShK-like3* transcript did not show any specific staining (*SI Appendix, Fig. S6*). Neuronal expression patterns are summarized in *SI Appendix, Table S3*.

To determine whether the neuropeptides are colocalized with the well-studied neuronal marker *ELAV* (3), we employed double ISH (dISH) in primary polyps. *ShK-like2* expression (the long probe) showed partial overlap with *ELAV* mostly in tentacle endoderm. *ShK-like4* expression partially colocalized with *ELAV* in the ectodermal sensory cells. *Aeq5-like1* and *Class8-like* expression did not show any overlap (*SI Appendix, Figs. S7 and S8*).

Single-cell RNA-seq data (2) confirmed that *ShK-like2b*, *ShK-like4*, *Aeq5-like1*, *Class8-like*, and *Aeq5-like2* are expressed in neurons in both larvae and adult polyps (*SI Appendix, Fig. S9*). In larvae, *ShK-like2b* and *Class8-like* expression were additionally up-regulated in undifferentiated metacells that may correspond to neural precursors. In adult polyps, these data also supported partial overlap between *ELAV* and the neuropeptide expression that we observed in dISH on primary polyps: *ShK-like4* in the C32 neuronal metacell and *ShK-like2b* in C34. It also suggested an overlap of *Aeq5-like1* and *ELAV* in C30 neurons that we were not able to detect in our dISH experiments. Thus, our data revealed secretory peptides with cysteine motifs of ShK and Aeq5a toxins to be produced by neurons.

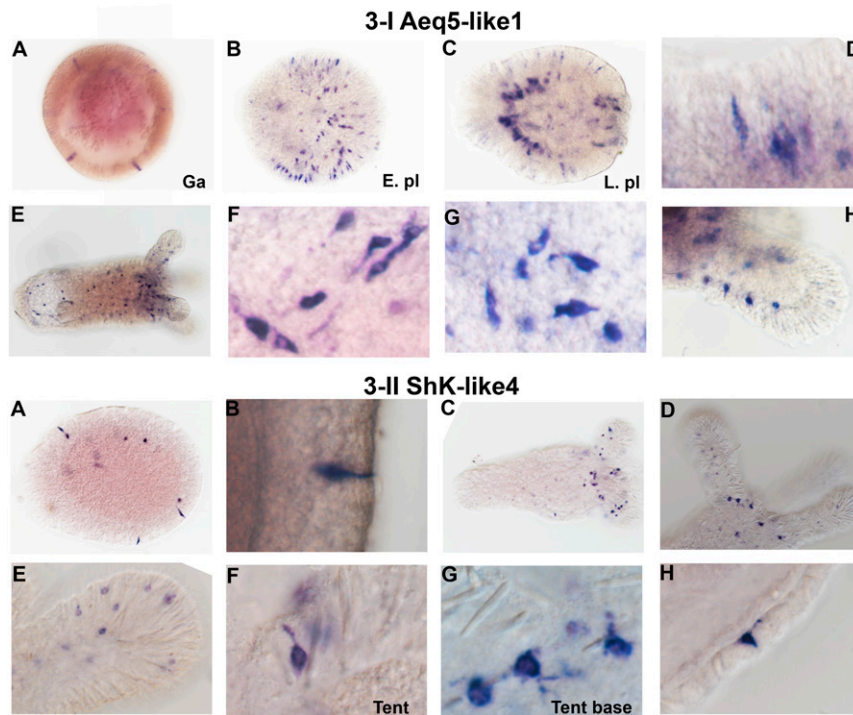


Fig. 3. Expression pattern of *Aeq5-like1* and *ShK-like4* genes by ISH. (3-I) *Aeq5-like1* expression starts in the ectodermal cells of gastrula (A) and early planula (E. pl) (B), and by the late planula (L. pl) (C) it is also noticeable in the endoderm. (D) Morphology of the planula ectodermal cells is consistent with sensory neurons. In the primary polyp, *Aeq5-like1* is expressed in both ectoderm and endoderm (E) in ganglion (F and G) and sensory neurons (H). (3-II) *ShK-like4* gene is expressed in ectodermal sensory cells in planula (A and B) and primary polyp (C–H). Tent (tentacles). Magnifications: 200 \times (3-I A–C, E; 3-II A, C), 600 \times (3-I D, F–H; 3-II, B, D–H).

ShK-like2 Gene Structure and Evolution. Alignment of the *ShK-like2a* and *ShK-like2b* transcripts to the genome (43) showed that both isoforms were encoded by the same gene (scaffold_260:199,854–210,905) with alternative transcription start sites (TSS) (Fig. 6A). The *ShK-like2* gene contains four coding exons, where exon 1 and exon 2 are alternating between the isoforms. Transcription starting from putative promoter A leads to *ShK-like2a* encoded by exons 1, 3, and 4, while alternative transcription activated by putative promoter B results in *ShK-like2b* encoded by exons 2, 3, and 4. To confirm that *ShK-like2b* expression was driven by the alternative promoter B localized upstream of exon 2, we generated a transgenic construct bearing the reporter gene mOrange2 (44) under the control of *ShK-like2b* regulatory elements. Injection of the transgenic construct into *Nematostella* zygotes resulted in reporter gene expression in ectodermal cells of the planulae (Fig. 6B–E).

Reverse transcription quantitative PCR (RT-qPCR) analysis showed that *ShK-like2* was initially expressed either in early planula (*ShK-like2b* isoform) or late planula (*ShK-like2a* isoform) and increased to the adult stage, with the exception of *ShK-like2a* in adult females that appeared to decrease its expression level (Fig. 6F). Both isoforms were expressed in the physa, mesentery, and tentacles of adult females; *ShK-like2b* had higher expression level in the mesentery (Fig. 6G).

To confirm the differences between the expression profiles of *ShK-like2a* and *ShK-like2b* isoforms observed in the qPCR and ISH staining (Fig. 5-II and 5-III), we employed DEXseq. (45). Raw reads, generated from *Nematostella* across a developmental time course (1 to 8 dpf and juveniles) (38), were mapped back to a de novo assembled transcriptome to detect differential exon usage. DEXseq analysis confirmed that exon 1 (*ShK-like2a*) and exon 2 (*ShK-like2b*) had significant differential exon usage (false-discovery rate $P < 0.001$) across development. Specifically, *ShK-like2b* was preferentially expressed from gastrula stage (1 dpf)

to late planulae stage (5 dpf), with exon 2 usage having a fold-change ≥ 2 compared to exon 1. Following metamorphosis (6 dpf), no significant difference in exon usage was observed (Fig. 6H). Thus, DEXseq analysis detected that patterns of *ShK-like2* exon expression were consistent with our qPCR results.

Interestingly, structure of the genomic locus of the *ShK-like2* ortholog is conserved in the sea anemones *Exaiptasia pallida* and *A. equina*, which are distantly related to *Nematostella* (41, 46). In these species, regulation of the *ShK-like2* gene by alternative TSSs also results in a protein with altering signal peptides (47). *ShK-like2* orthologs retrieved by BLAST search in other available sea anemone transcriptomes [*Calliactis polypus*, *Nemanthus annamensis*, *Aulactinia veratra*, *Actinia tenebrosa*, *A. equina* (48), *Anthopleura bud-demeieri* (49), *Anthopleura elegantissima*, *Edwardsiella lineata* (50), *S. helianthus* (51)] also possessed alternative signal peptides and identical mature parts (Fig. 7A and B), suggesting the regulatory mechanism is widely conserved in sea anemones. Thus, it is likely that a *ShK-like2* ortholog with this unique structure and regulation was present in the last common ancestor of all sea anemones.

Gene models for *ShK-like1* and *ShK-like3* had not been established previously; therefore, we aligned corresponding transcripts to the genome scaffolds. *ShK-like1* transcript aligned to a region of scaffold 53, which was interrupted by a gap corresponding to a part of the mature peptide. We PCR-amplified and sequenced that fragment of the genomic DNA to annotate the remaining part of this locus (SI Appendix, Table S3). Structure of the *ShK-like1* (scaffold_53:491,201–496,216) and *ShK-like3* (scaffold_7:1,285,234–1,288,665) genes was quite different: the *ShK-like1* gene had two coding exons while *ShK-like3* gene contained three coding exons (Fig. 7C). The *ShK-like3* gene structure was similar to *ShK-like2* gene fragment encoding *ShK-like2b*. Orthologs of *ShK-like1* and *ShK-like3* were not found in other sea anemones, not even in other members of the Edwardsiidae family. Thus, *ShK-like1* and *ShK-like3*

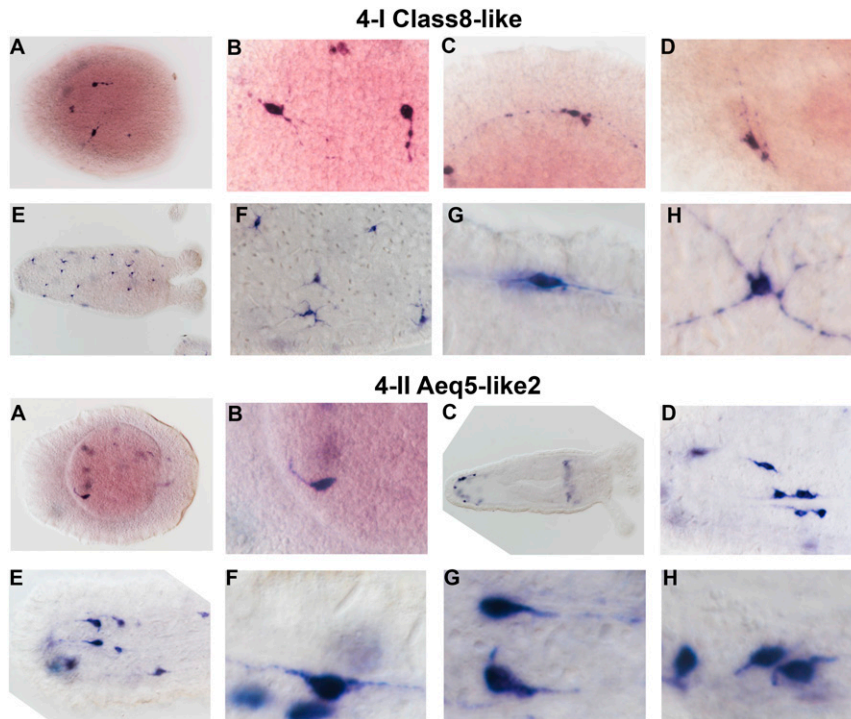


Fig. 4. Expression pattern of *Class8-like* and *Aeq5-like2* genes by ISH. (4-I) *Class8-like* gene is expressed in the ganglion neurons residing in the mesoglea of planula (A–D) and primary polyp (E–H). (4-II) *Aeq5-like2* is expressed in the endodermal ganglion neurons in planula (A and B) and primary polyp (C–H). Magnifications: 200× (4-I, A, E; 4-II A, C), 600× (4-I, B–D, F–H; 4-II, B, D–H).

probably resulted from *Nematostella*-specific gene duplications followed by sequence divergence. *ShK-like3* is expressed at a very low level and ISH did not result in any staining therefore it is probably a pseudogene.

Functional Characterization of ShK-like2. Because ShK-like2 shares high sequence similarity with the ShK-like1 toxin, we were interested in characterizing its biological activity in comparison to ShK-like1. ShK-like2 was produced as a recombinant peptide (SI Appendix, Fig. S10). ESI-MS showed that the molecular mass of the peptide (4470.3) corresponded to the calculated molecular weight of the isoform with all of the six cysteine residues oxidized to form three disulfide bonds (4470.14) (SI Appendix, Fig. S10B). To determine whether ShK-like2 provokes any noticeable physiological reaction in primary polyps, we incubated 11-dpf polyps with recombinant ShK-like2 in parallel to ShK-like1 and NvePtx1 (9) toxins (0.5 mg/mL) in addition to the BSA control (5 mg/mL). After 2 h, length-to-width ratio of the body column was significantly lower in the ShK-like2 sample compared to NvePtx1 and BSA ($P = 2.4 \times 10^3$ and $P = 3.4 \times 10^8$, respectively; Student's *t*-test) due to body contraction; there was no difference in comparison to the ShK-like1 sample ($P = 0.3$, Student's *t*-test) (Fig. 8 A and B).

To test the toxic activity of ShK-like2, we incubated five zebrafish larvae in 0.5-mg/mL ShK-like2 solution (repeated in triplicate) in parallel to ShK-like1 and BSA control (Fig. 2C). We observed paralysis and death in 4 of a total of 15 tested larvae within 2 h of incubation. Similarly to ShK-like1, ShK-like2 lacks residues important for blocking vertebrate potassium channels (Fig. 14).

ShK-like2 was first expressed at the early planula stage and therefore we aimed to test whether it played a role in development using gene knockdown with Morpholino oligonucleotides (MO). We aimed to disrupt splicing for either exon 3 or exon 4, which are shared among the isoforms (Fig. 8A), using specific

MO sequences. Animals injected by either MO targeting exon 3 (E3I3) or 4 (I3E4) had strong developmental defects (Fig. 8 C and D), resulting in a significant proportion of larvae being unable to settle compared to a morpholino control ($P = 0$, χ^2 test and $P < 0.00001$, Fisher exact test for E3I3 and I3E4, respectively). In comparison to the control MO, injection of E3I3 and I3E4 MOs on average reduced planulae settlement by $39.4 \pm 8.8\%$ (four replicates) (SI Appendix, Table S7) and $83.4 \pm 9.4\%$ (three replicates) (SI Appendix, Table S8), respectively. The difference in the effects may be explained by the difference in efficiency of each MO. PCR analysis showed that I3E4 led to nearly complete elimination of the processed mRNA, while E3I3 MO injection resulted in partial reduction of the processed mRNA and correctly processed mRNA could be detected as well (SI Appendix, Fig. S11).

Discussion

Annotation of Toxins in Cnidarians. The research reported here led to the discovery of two neuropeptide types; however, it was initially aiming to find new toxins in *Nematostella* transcriptome using UniProt sequences annotated as toxins (including “toxin by similarity”) and a search for secreted peptides with ShKT motif. Unexpectedly, all but one of the retrieved candidates were found to be expressed in ganglion and sensory neurons, not in nematocytes or ectodermal gland cells. This demonstrates that molecules with structural similarities to venom components have wide expression in different cell types in cnidarians, particularly neurons. Thus, many sequences annotated as venom components based on the sequence similarity or toxic activity without dissecting their tissue localization may have different functions. This mischaracterization is a challenge to overcome as the cnidarian venom system is composed of sparse, dispersed cells rather than an anatomically defined venom gland, leading to technical difficulties in studying venom-specific transcriptomes. Techniques enabling studies of cell-specific transcriptomes have

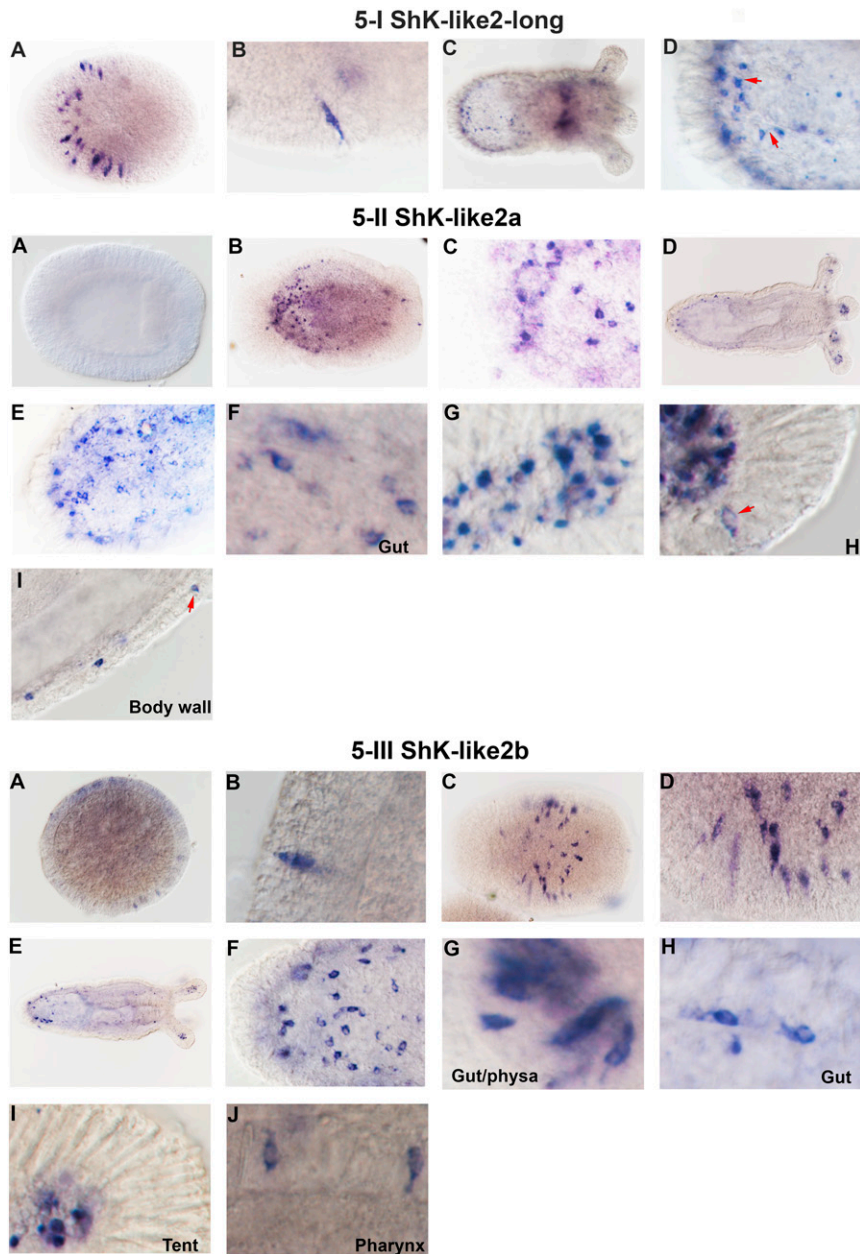


Fig. 5. *ShK-like2* expression pattern by ISH. (5-I) ISH with the *ShK-like2*-long probe reveals expression in planula's ectodermal cells (A) with a sensory neuron morphology (B) and in primary polyp's endoderm (C) in ganglion neurons with characteristic protrusions (D, arrows). (5-II) ISH with *ShK-like2a*-short probe reveals that *ShK-like2a* has no expression in early planula (A), expression starts mostly in the endoderm of late planula (B) in ganglion neurons (C). In the primary polyp (D), it is expressed mostly in the endodermal ganglion neurons in physa (E), gut (F), and tentacles (G and H). The expression is also evident in the smaller number of ectodermal sensory neurons in tentacles (H, arrow) and body wall (I, arrow). (5-III) ISH with *ShK-like2b*-short probe shows that *ShK-like2b* starts expressing in the ectoderm of early planula (A) in elongated cells (B) and continues in the ectodermal sensory neurons in the late planula (C and D). In the primary polyp (E), it is expressed in multiple endodermal ganglion neurons (F) in physa (G), gut (H), and tentacles (I). *ShK-like2b* is also found in a small number of ectodermal neurons in pharynx (J). Magnifications: 200 \times (5-I, A, C; 5-II, A, B, D; 5-III, A, C, E), 600 \times (5-I, B, D; 5-II, C, E-I; 5-III, B, D, F-J).

only recently become available in cnidarians and will make this task more manageable (2, 52).

Recruitment to Venom. To the best of our knowledge, the recruitment of ShK-like1 is unique as a case of a toxin evolving from a neuropeptide and evaluated at the functional level of individual genes in an evolutionary context; it is also unique as a study of toxin recruitment in cnidarians. Processes surrounding venom recruitment have been widely debated, with inconclusive or conflicting findings based on a comparative tissue-specific gene-expression profiling and

evolutionary genomic approach (53–55). The high sequence similarity between mature ShK-like1 and ShK-like2 peptides (51% amino acid identity) and occurrence of ShK-like1 only in *Nematostella* provides evidence for recent gene duplication and recruitment events. In cone snails, recruitment of insulin from neuroendocrine cells to venom was suggested. While the similarity between the signaling and venom insulins is mostly limited to the conserved cysteine framework, conserved intron–exon boundaries (26) provide evidence for an ancient gene-duplication event. Thus, while in most other cases the protein recruitment pathway was identified at the level of protein

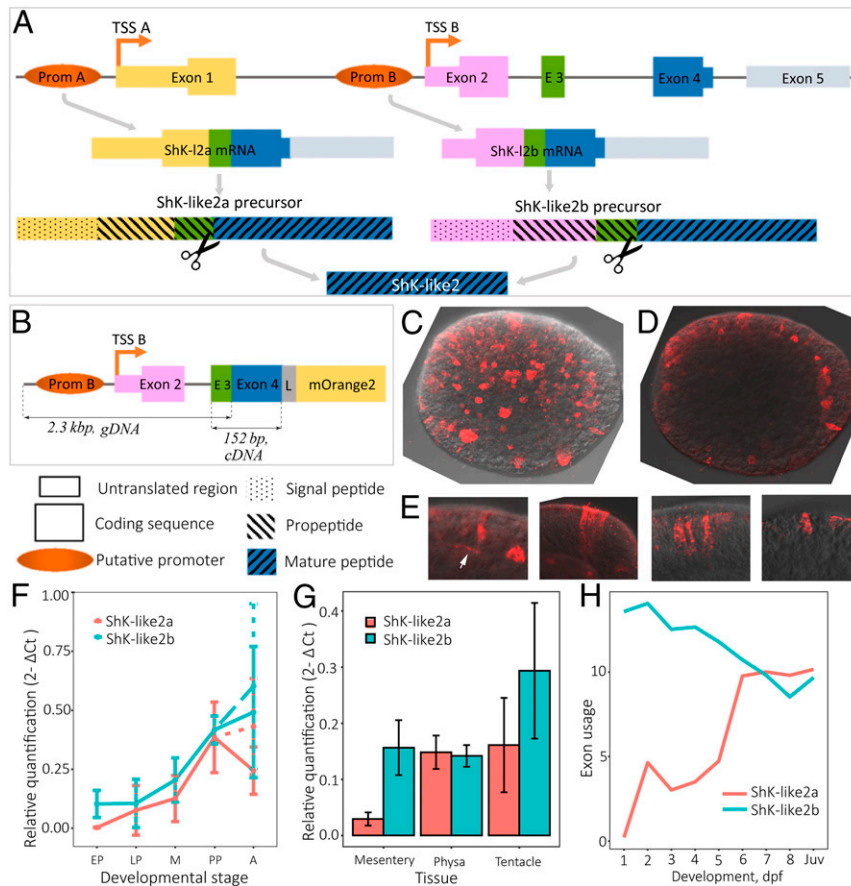


Fig. 6. Structure and regulation of *ShK-like2* gene. (A) *ShK-like2* gene structure and biosynthesis of the two isoforms of ShK-like2 precursor protein leading to the same mature peptide. (B) Scheme of the transgenic construct aimed to label cells expressing ShK-like2b isoform. (C–E) Immunostaining of transgenic planulae, confocal imaging; red, mOrange2; magnification 400 \times ; (whole animals). (C) Staining in multiple ectodermal cells (z-stack of 42 maximum-intensity projections). (D) Individual stained ectodermal cells (z-stack of 4 maximum-intensity projections). (E) Stained cells are morphologically sensory neurons with processes at the base (arrow) and uneven distribution of the staining evidencing for mOrange2 being packed in vesicles (z-stacks of 3 to 4 maximum-intensity projections). (F) *ShK-like2a* and *ShK-like2b* gene-expression dynamics in development measured by qPCR. A, adult; EP, early planula; LP, late planula; M, metamorphosis; PP, primary polyp. Adult females, solid line; adult males, dashed line. (G) *ShK-like2a* and *ShK-like2b* expression levels in adult female tissues measured by qPCR. In F and G, error bars represent SDs. (H) Differential usage of exon 1 (*ShK-like2a*) and compared to exon 2 (*ShK-like2b*) detected by DEXseq. Juv, juvenile polyp (6-wk-old).

superfamily (15), here we pinpoint an ancestral gene giving rise to a previously undescribed toxin. This unique finding allowed us to characterize both the toxin and the ancestral physiological protein in the same species.

Our data provide evidence that ShK-like1, a component of nematocyte venom, evolved via a *Nematostella*-specific duplication of a neuronally expressed gene followed by sequence divergence and recruitment of one of the copies into the venom system. The other copy likely retained its expression pattern, giving rise to ShK-like2. *ShK-like3* appears to result from a duplication of an ancestral *ShK-like2* gene as well; however, unlike *ShK-like1*, it became a pseudogene. We have previously described a similar evolutionary pattern of gene gain and loss for another family of *Nematostella* venom components (56). The fact that nematocytes represent a highly specialized neuronal cell type in *Nematostella* (1) may facilitate recruitment of proteins with typical neural functions to venom, and it is plausible that this evolutionary pathway is recurrent in sea anemones. This is further supported by our discovery of two other neuropeptides with ShK cysteine motif (Class8-like and ShK-like4) and multiple venom components with ShK domains, reported earlier (7); however, theoretically the recruitment in the opposite direction (from venom to nervous system) is also possible (57).

As neuropeptides affect the nervous system, it might seem plausible that they will be frequently recruited to venoms. Disulfide-rich venom components with high sequence similarity to hormone-like peptides from the nervous system were also reported in cone snails (58–60) and spiders (61). In other reported cases, the similarity was mostly limited to a shared cysteine framework [e.g., snake three-finger toxins (62, 63)] or three-dimensional structure (64). An example of short linear neuropeptides recruited to venom in multiple instances is invertebrate tachykinins (65). However, such events have been rarely described and it is possible that most neuropeptides make poor toxins because they act mainly via G protein-coupled receptors, which deliver signal too slowly to be an effective target for toxins (28). Additionally, linear neuropeptides might be less proteolytically stable (66) compared to disulfide-rich peptides (67, 68), limiting their chances of being efficient venom components. Yet, the small number of known linear neuropeptides recruited in venoms might be partially explained by technical difficulties of their detection by similarity search tools due to short length.

Partial Functional Specialization of ShK-like1 Compared to ShK-like2. When we compared the toxicity of ShK-like1 and ShK-like2 on fish, it was noticeable that the latter is less effective, suggesting that ShK-like1 evolved higher pharmacological efficiency against

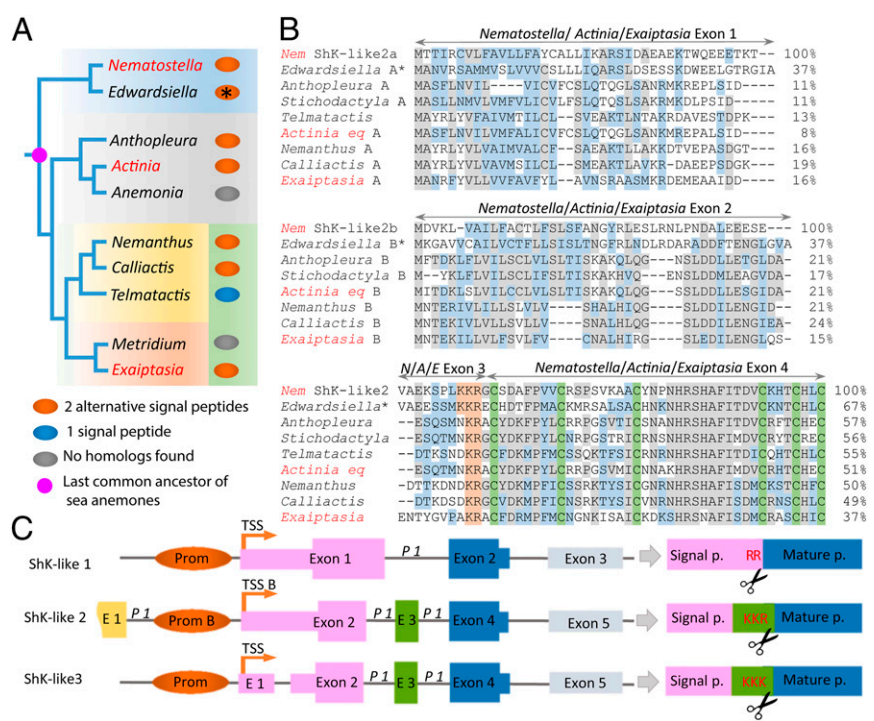


Fig. 7. Evolution of *ShK-like2* gene. (A) Distribution of the *ShK-like2* homologs across the phylogenetic tree of sea anemones (adapted from ref. 41); the species with sequences genomes are in red. The number of alternative signal peptides for each *ShK-like2* homolog identified in each genus (transcriptome or genome) is represented by ovals (orange, 2; blue, 1; gray, no homologs); the asterisk denotes that full-length sequences were obtained by joining shorter fragments manually. (B) Alignment of *ShK-like2* homologs from different sea anemones. The alignment was built using Clustal Omega online tool with minor manual adjustments. Exon-intron structure is conserved between *Nematostella*, *Actinia*, and *Exaiptasia* (in red) and shown above the alignment. (C) Alignment of exon-intron structures of *ShK-like1*, *ShK-like2*, and *ShK-like3* genes.

fish. When new peptides are recruited into the venom system as toxins, they may lose their original function. However, when we tested the ability of *ShK-like1* to induce effects in *Nematostella*, it was as effective as *ShK-like2* (Fig. 8A). Thus, it seems that *ShK-like1* retained its ancestral activity to affect the *Nematostella* nervous system, despite the increased activity on fish. This suggests that increasing potency as a toxin against one species does

not necessarily result in decreasing its efficacy on another. It is currently unknown what the pharmacological targets or receptors of *ShK-like1* and *ShK-like2* are and whether these targets are shared between fish and sea anemones. Both *ShK-like1* and *ShK-like2* lack residues that have been reported to be important for potassium channel-blocking activity of *ShK* (Fig. 1A) (39) and therefore they probably act through a different mechanism.

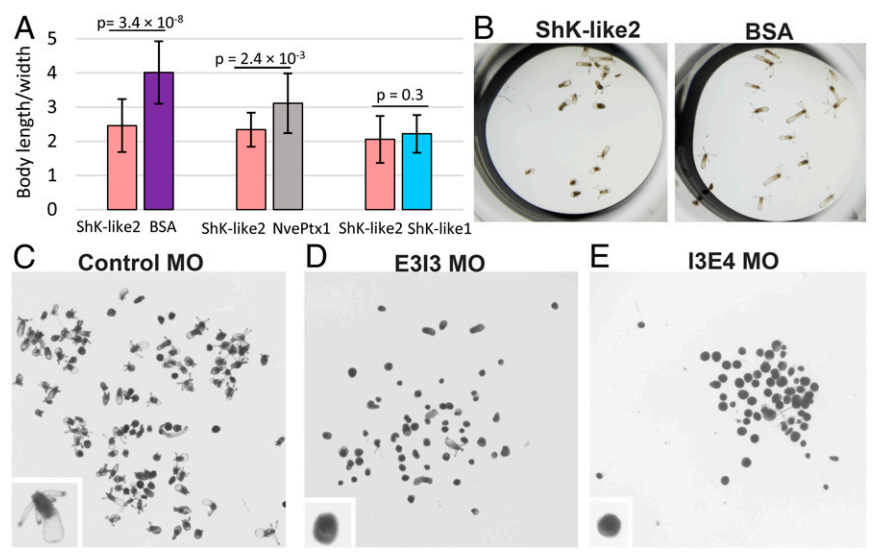


Fig. 8. Effects of *ShK-like2* on *Nematostella*. (A) Ratios of body length to width in primary polyps after 2 h of incubation with *ShK-like2* in comparison to BSA, NvePtx1, and *ShK-like1*. (B) Effect of the 2-h incubation with *ShK-like2* in comparison to BSA. (C–E) Developmental effects of MOs in 11-dpf polyps, magnification 10 \times ; close-ups to embryos with typical phenotypes are in the lower left corners. (C) Control MO, (D) E3I3 MO, (E) I3E4 MO.

Several examples of sea anemone peptides with an ShK fold but lacking potassium channel blocking activity have been reported recently (69, 70).

Cysteine-Rich Peptides in the Nervous System and Other Regulatory Systems. Secreted peptides with ShK-like cysteine motifs have been reported as venom components in sea anemones; however, to the best of our knowledge, the present work shows evidence of ShK-like peptides being components of the nervous system is unique. The cysteine arrangement of Aeq5-like1 and Aeq5-like2 had been reported in Aeq5 peptide extracted from total lysate of *A. equina*, where they were shown to be lethal for crabs and hence classified as a toxin (31). Neuronal localization of peptides of this type has not been reported before. Because multiple disulfide-rich venom components (including ShK toxin) are neurotoxins targeting neuroreceptors, the molecules may potentially act through binding to endogenous receptors in *Nematostella*.

Disulfide-rich peptides from insulin superfamily are widespread in both vertebrates and invertebrates (71). In insects and mollusks, some of these peptides are produced by neurosecretory cells and regulate a wide range of physiological functions, including metabolism and learning (24, 25, 72). However, the insulin superfamily is not restricted to the nervous system and produced by several cell types and tissues (24, 71).

Toxin-like molecules have been reported in vertebrate, mollusk, and insect nervous systems (25, 26, 63, 73–75), but in many cases their sequence similarity to toxins is very low, making convergent evolution a realistic scenario. One previously reported example is Agatoxin-like peptides in the insect neuroendocrine system, where the sequence similarity to the corresponding spider toxin is quite high (74). The most characterized group of toxin-like neuroregulators are mammalian prototoxins with significant structural similarity to snake three-fingered toxins (63). Members of these family can be secreted or glycosylphosphatidylinositol-anchored (which makes them similar to the membrane-bound Aeq5-like2). They regulate function of nicotinic acetylcholine receptors at multiple levels, including direct binding to the mature receptor, leading to modulation or inhibition of its activity as well as by affecting subunit stoichiometry during receptor assembly.

Regulation of ShK-like2 by Alternative Promoters and TSSs. *ShK-like2* is regulated by alternative promoters, resulting in different dynamics of the transcripts encoding the two precursor isoforms in development and probably in tissues. As our exon usage and ISH analyses showed, *ShK-like2b* isoform is preferentially expressed during the first 5 d of development with initial expression in the ectoderm of early planula. *ShK-like2a* was first expressed in the late planula and mostly in the endoderm. Additionally, qPCR analysis suggests that the two isoforms might be expressed at different levels in the mesentery of adult females. Thus, the *ShK-like2* gene is equipped with a sophisticated regulatory system allowing expression flexibility.

However, *ShK-like2a* and *-2b* expression patterns largely overlap in primary polyps. Eukaryotic promoters not only initiate transcription but also affect mRNA stability, translation efficiency, and subcellular localization (76, 77), particularly in neurons (78). Thus, even though *ShK-like2a* and *-2b* are coexpressed in the same cells, the regulation by different promoters may potentially result in different abundance and localization of protein precursors. Additionally, different promoters might respond differently to stimuli. Regulation by alternative promoters and TSSs is common in development (76, 79) and has been reported for multiple neuronal proteins, such as signaling molecules (80), receptors (81), and enzymes (82).

Alternative TSSs result in different prepropeptides in the ShK-like2 isoforms. Propeptides are known to affect the fate of mature peptides in multiple ways, including its stability and localization (83). Different propeptide sequences may potentially affect sorting of mature ShK-like2 and lead to different

intracellular localization of the two isoforms. For example, in the mammalian neuroendocrine system, prohormones are sorted into regulated secretory pathway through binding of their propeptide to a specific receptor in the *trans*-Golgi network (84), and therefore different propeptide sequences might affect the binding. Furthermore, in the sea hare *Aplysia californica*, neuropeptides derived from processing of the same precursor are sorted to different classes of secretory vesicles and transported to different neural processes of the same cell, depending on the sequences of intermediates during the stepwise maturation (85).

Thus, the structure of the *ShK-like2* gene underlies regulatory flexibility at the transcriptional level in development and may potentially affect regulation at posttranscriptional, translational, and posttranslational levels. The conservation of such gene structure among ShK-like2 homologs among sea anemones supports its functional importance. To the best of our knowledge, this report of gene regulation by alternative promoters and TSS in a cnidarian and, particularly in cnidarian neuropeptides, is unique.

Methods

For a full description of methods used, see *SI Appendix, Supplementary Information Text*.

Sequence Search and Analysis. The sequences were retrieved by a combination of online BLAST (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) in the TSA databases in the National Center for Biotechnology Information (NCBI), tBLASTn on Reef Genomics (reefgenomics.org), and search for the consistent amino acid arrangement of the last three cysteine residues in ShK domains (CXXCXXC) using TransDecoder (32). To resolve isoforms of the ShK-like2 ortholog in *A. equina*, a de novo transcriptome was assembled using Trinity (v2.6.6) (32) from previously published raw reads (accession no. SRR7507858) (48).

Transgenic Construct and Immunostaining. The *ShK-like2* gene fragment encoding putative promoter B, exon 2, intron 2, and exon 3, was fused to the exon 4 and cloned into an expression plasmid (9) to be located upstream of the mOrange2 sequence. The plasmid was injected into *Nematostella* zygotes along with the yeast Meganuclease to facilitate genomic integration according to an established protocol (86). Immunostaining was performed using a rabbit polyclonal antibody against mCherry (Abcam), as previously described (87). Imaging was performed with an Olympus FV-1000 inverted confocal microscope (Olympus).

ISH, dISH, and dFISH. ISH and dFISH were performed as previously described (3, 7, 88). Probe sequences are available in *SI Appendix, Table S2*.

qPCR and RNA-Seq. Total RNA from *Nematostella* larvae, adults, and female body parts was extracted and cDNA was synthesized as described earlier (9). We designed qPCR primers specific to the ShK-like2a and ShK-like2b transcripts (*SI Appendix, Table S3*). To quantify transcripts, the StepOnePlus Real-Time PCR System (ABI instrument, Thermo Fisher Scientific) was used. ANOVA was used to determine if the expression of ShK-like2a and ShK-like2b was significantly different across development and tissues. Ct values were used to calculate relative abundance of the transcripts. DEXseq analysis (45, 89) was performed using a *Nematostella* transcriptome assembled de novo with Trinity (v2.6.6) (32), previously published raw reads (PRJEB13676) (90), and the Trinity SuperTranscripts pipeline (32, 89). *Nematostella* raw reads (38) generated across development were downloaded from the NCBI Sequence Read Archive (PRJNA418421 and PRJNA419631) and mapped to the SuperTranscripts assembly using STAR (91).

Recombinant Expression, Purification, and MS. Fragments encoding ShK-like1 and ShK-like2 mature peptides were cloned into pET40 vector (MilliporeSigma), expressed in BL21(DE3) *Escherichia coli* (MilliporeSigma) and purified by FPLC, as described previously (9). ESI-MS was employed to determine the molecular weight of the purified recombinant peptides following an established protocol (92).

Toxicity Assay. ShK-like1 and ShK-like2 recombinant peptides were added to 0.5 mg/mL concentration to zebrafish media and incubated with 4-dpf zebrafish (*Danio rerio*) larvae in 500- μ L wells, five larvae per well. Next, 5 mg/mL BSA was used as a negative control. At this developmental stage

(larvae younger than 120 hpf), any ethical permits are not required for experimental use according to the European and Israeli laws.

Analysis of Nematocyst Content. Nematocysts were purified from 4-dpf planulae using a Percoll (MilliporeSigma) gradient. The purified nematocysts were disrupted by ultrasonication and the soluble proteins were reduced with DTT, alkylated with iodoacetamide, digested by trypsin, and analyzed by LC-MS/MS with a Q Exactive Plus mass spectrometer (Thermo Fisher Scientific). Protein identification was performed using MaxQuant software (93) as described earlier (9). The raw LC-MS/MS files along with MaxQuant output were submitted to ProteomeXchange Consortium via the PRIDE (94) partner repository with identifier PXD019085 (95). Annotated spectra were obtained using Proteome Discoverer v2.4.1.15 (Thermo Scientific). To rank protein abundancies, intensity value was normalized against the number of unique and razor peptides identified for each protein group. Uncharacterized proteins were annotated using InterProScan (<https://www.ebi.ac.uk/interpro>).

ShK-like1 and ShK-like2 Activity on Polyyps. Eleven-day-old primary polyyps were incubated with recombinant peptides (0.5 mg/mL). BSA (5 mg/mL) and NvePtx1 toxin (9) were used for comparison. The effects were recorded every 30 min during 2 h avoiding any physical disturbance of the animals.

Knockdown with MOs. MOs (*SI Appendix, Table S4*) were ordered from Gene Tools, LLC (Philomath). MOs were designed to complement sequences

corresponding to the boundaries of exon 3–intron 3 (E3–I3) and intron 3–exon 4 (I3–E4) of the *ShK-like2* gene to interfere with splicing. MOs were injected into *Nematostella* zygotes at the concentration of 0.9 mM in parallel to a control MO, which has no binding target in *Nematostella*. After 10 d, the proportion of polyyps that were unsettled vs. settled was recorded, and Fisher's exact test and χ^2 test used to calculate significance. Effects on splicing were assessed by PCR (*SI Appendix, Table S3*).

Data Availability. Proteomics data have been deposited in the publicly accessible ProteomeXchange Consortium, www.proteomexchange.org (accession no. PXD019085, <https://www.ebi.ac.uk/pride/archive/projects/PXD019085>) (37).

ACKNOWLEDGMENTS. We thank Prof. Norman Metanis and Reem Mousa (Institute of Chemistry, The Hebrew University) for the help with electrospray ionization mass spectrometry; and the following researchers of the core facilities of the Alexander Silberman Institute of Life Sciences, The Hebrew University: Dr. William Breuer (Interdepartmental Unit) for the help with tandem mass spectrometry, Dr. Mario Lebendiker (Protein Expression and Purification Unit) for the help with chromatography, and Dr. Naomi Melamed-Book (Advanced Imaging Unit) for her help with confocal microscopy. This research was supported by an Israel Science Foundation Grant 869/18 (to Y.M.) and United States–Israel Binational Science Foundation Grant 2014667 (NSF Award 1536530) (to Y.M. and A.M.R.).

- G. S. Richards, F. Rentsch, Transgenic analysis of a SoxB gene reveals neural progenitor cells in the cnidarian *Nematostella vectensis*. *Development* **141**, 4681–4689 (2014).
- A. Sebe-Pedros *et al.*, Cnidarian cell type diversity and regulation revealed by whole-organism single-cell RNA-seq. *Cell* **173**, 1520 (2018).
- N. Nakanishi, E. Renfer, U. Technau, F. Rentsch, Nervous systems of the sea anemone *Nematostella vectensis* are generated by ectoderm and endoderm and shaped by distinct mechanisms. *Development* **139**, 347–357 (2012).
- S. Gründer, M. Assmann, Peptide-gated ion channels and the simple nervous system of Hydra. *J. Exp. Biol.* **218**, 551–561 (2015).
- C. N. David *et al.*, Evolution of complex structures: Minicollagens shape the cnidarian nematocyst. *Trends Genet.* **24**, 431–438 (2008).
- G. Kass-Simon, A. A. Scappaticci, The behavioral and developmental physiology of nematocysts. *Can. J. Zool.* **80**, 1772–1794 (2002).
- Y. Moran *et al.*, Analysis of soluble protein contents from the nematocysts of a model sea anemone sheds light on venom evolution. *Mar. Biotechnol. (NY)* **15**, 329–339 (2013).
- T. Rachamim *et al.*, The dynamically evolving nematocyst content of an anthozoan, a scyphozoan, and a hydrozoan. *Mol. Biol. Evol.* **32**, 740–753 (2015).
- Y. Y. Columbus-Shenkar *et al.*, Dynamics of venom composition across a complex life cycle. *eLife* **7**, e35014 (2018).
- Y. Moran *et al.*, Neurotoxin localization to ectodermal gland cells uncovers an alternative mechanism of venom delivery in sea anemones. *Proc. Biol. Sci.* **279**, 1351–1358 (2012).
- B. Madio, G. F. King, E. A. B. Undheim, Sea anemone toxins: A structural overview. *Mar. Drugs* **17**, 325 (2019).
- Y. Moran *et al.*, Intron retention as a posttranscriptional regulatory mechanism of neurotoxin expression at early life stages of the starlet anemone *Nematostella vectensis*. *J. Mol. Biol.* **380**, 437–443 (2008).
- O. Castañeda *et al.*, Characterization of a potassium channel toxin from the Caribbean Sea anemone *Stichodactyla helianthus*. *Toxicon* **33**, 603–613 (1995).
- T. Shafee, M. L. Mitchell, R. S. Norton, Mapping the chemical and sequence space of the ShKT superfamily. *Toxicon* **165**, 95–102 (2019).
- B. G. Fry *et al.*, The toxicogenomic multiverse: Convergent recruitment of proteins into animal venoms. *Annu. Rev. Genomics Hum. Genet.* **10**, 483–511 (2009).
- S. S. Pineda *et al.*, Structural venomomics reveals evolution of a complex venom by duplication and diversification of an ancient peptide-encoding gene. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 11399–11408 (2020).
- S. H. Drukevitze, L. Bokelmann, E. A. B. Undheim, B. M. von Reumont, Toxins from scratch? Diverse, multimodal gene origins in the predatory robber fly *Dasypteron diadema* indicate a dynamic venom evolution in dipteran insects. *Gigascience* **8**, giz081 (2019).
- E. A. B. Undheim, M. Mobli, G. F. King, Toxin structures as evolutionary tools: Using conserved 3D folds to study the evolution of rapidly evolving peptides. *BioEssays* **38**, 539–548 (2016).
- B. Galliot *et al.*, Origins of neurogenesis, a cnidarian view. *Dev. Biol.* **332**, 2–24 (2009).
- T. Takahashi, N. Takeda, Insight into the molecular and functional diversity of cnidarian neuropeptides. *Int. J. Mol. Sci.* **16**, 2610–2625 (2015).
- M. Ancia, Chemical transmission in the sea anemone *Nematostella vectensis*: A genomic perspective. *Comp. Biochem. Physiol. Part D Genomics Proteomics* **4**, 268–289 (2009).
- E. Hayakawa *et al.*, A combined strategy of neuropeptide prediction and tandem mass spectrometry identifies evolutionarily conserved ancient neuropeptides in the sea anemone *Nematostella vectensis*. *PLoS One* **14**, e0215185 (2019).
- P. R. H. Steinmetz, A. Aman, J. E. M. Kraus, U. Technau, Gut-like ectodermal tissue in a sea anemone challenges germ layer homology. *Nat. Ecol. Evol.* **1**, 1535–1542 (2017).
- D. R. Nässel, J. Vanden Broeck, Insulin/IGF signaling in *Drosophila* and other insects: Factors that regulate production, release and post-release action of the insulin-like peptides. *Cell. Mol. Life Sci.* **73**, 271–290 (2016).
- P. D. Floyd *et al.*, Insulin prohormone processing, distribution, and relation to metabolism in *Aplysia californica*. *J. Neurosci.* **19**, 7732–7741 (1999).
- H. Safavi-Hemami *et al.*, Venom insulins of cone snails diversify rapidly and track prey taxa. *Mol. Biol. Evol.* **33**, 2924–2934 (2016).
- N. Nakanishi, M. Q. Martindale, CRISPR knockouts reveal an endogenous role for ancient neuropeptides in regulating developmental timing in a sea anemone. *eLife* **7**, e39742 (2018).
- G. Jékely *et al.*, The long and the short of it—A perspective on peptidergic regulation of circuits and behaviour. *J. Exp. Biol.* **221**, jeb166710 (2018).
- A. M. Benham, Protein secretion and the endoplasmic reticulum. *Cold Spring Harb. Perspect. Biol.* **4**, a012872 (2012).
- G. Anderluh, Z. Podlesek, P. Macek, A common motif in propeptides of cnidarian toxins and nematocyst collagens and its putative role. *Biochim. Biophys. Acta* **1476**, 372–376 (2000).
- T. Honma *et al.*, Novel peptide toxins from acrorhagi, aggressive organs of the sea anemone *Actinia equina*. *Toxicon* **46**, 768–774 (2005).
- B. J. Haas *et al.*, De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* **8**, 1494–1512 (2013).
- P. Jones *et al.*, InterProScan 5: Genome-scale protein function classification. *Bioinformatics* **30**, 1236–1240 (2014).
- D. Fredman, M. Schwaiger, F. Rentsch, U. Technau, *Nematostella vectensis* transcriptome and gene models v2.0. *figshare* (2013), <https://doi.org/10.6084/m9.figshare.807696.v1>. Accessed 1 November 2018.
- T. N. Petersen, S. Brunak, G. von Heijne, H. Nielsen, SignalP 4.0: Discriminating signal peptides from transmembrane regions. *Nat. Methods* **8**, 785–786 (2011).
- Y. Moran, D. Gordon, M. Gurevitz, Sea anemone toxins affecting voltage-gated sodium channels—Molecular and evolutionary features. *Toxicon* **54**, 1089–1101 (2009).
- M. Y. Sachkova, Y. Moran, Soluble proteome of Nematocysts from *Nematostella*. ProteomeXchange Consortium via PRIDE partner repository. <https://www.ebi.ac.uk/pride/archive/projects/PXD019085>. Deposited 11 May 2020.
- J. F. Warner *et al.*, NVERTX: A gene expression database to compare embryogenesis and regeneration in the sea anemone *Nematostella vectensis*. *Development* **145**, dev162867 (2018).
- H. Rauer, M. Pennington, M. Cahalan, K. G. Chandy, Structural conservation of the pores of calcium-activated and voltage-gated potassium channels determined by a sea anemone toxin. *J. Biol. Chem.* **274**, 21885–21892 (1999).
- H. Q. Marlow, M. Srivastava, D. Q. Matus, D. Rokhsar, M. Q. Martindale, Anatomy and development of the nervous system of *Nematostella vectensis*, an anthozoan cnidarian. *Dev. Neurobiol.* **69**, 235–254 (2009).
- E. Rodríguez *et al.*, Hidden among sea anemones: The first comprehensive phylogenetic reconstruction of the order Actiniaria (Cnidaria, Anthozoa, Hexacorallia) reveals a novel group of hexacorals. *PLoS One* **9**, e96998 (2014).
- D. Z. Faltine-Gonzalez, M. J. Layden, Characterization of nAChRs in *Nematostella vectensis* supports neuronal and non-neuronal roles in the cnidarian-bilaterian common ancestor. *Evodevo* **10**, 27 (2019).
- N. H. Putnam *et al.*, Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. *Science* **317**, 86–94 (2007).
- N. C. Shaner *et al.*, Improving the photostability of bright monomeric orange and red fluorescent proteins. *Nat. Methods* **5**, 545–551 (2008).

45. S. Anders, A. Reyes, W. Huber, Detecting differential usage of exons from RNA-seq data. *Genome Res.* **22**, 2008–2017 (2012).
46. C. S. Wilding *et al.*, The genome of the sea anemone *Actinia equina* (L.): Meiotic toolkit genes and the question of sexual reproduction. *Mar. Genomics* **53**, 100753 (2020).
47. S. Baumgarten *et al.*, The genome of *Aiptasia*, a sea anemone model for coral symbiosis. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 11893–11898 (2015).
48. F. M. Waldron, G. N. Stone, D. J. Obbard, Metagenomic sequencing suggests a diversity of RNA interference-like responses to viruses across multicellular eukaryotes. *PLoS Genet.* **14**, e1007533 (2018).
49. C. A. van der Burg, P. J. Prentis, J. M. Surm, A. Pavasovic, Insights into the innate immunome of actinarians using a comparative genomic approach. *BMC Genomics* **17**, 850 (2016).
50. D. J. Stefanik *et al.*, Production of a reference transcriptome and transcriptomic database (EdwardsiellaBase) for the lined sea anemone, *Edwardsiella lineata*, a parasitic cnidarian. *BMC Genomics* **15**, 71 (2014).
51. E. Rivera-de-Torre, A. Martinez-Del-Pozo, J. E. Garb, *Stichodactyla helianthus*' de novo transcriptome assembly: Discovery of a new actinoporin isoform. *Toxicon* **150**, 105–114 (2018).
52. K. Sunagar *et al.*, Cell type-specific expression profiling unravels the development and evolution of stinging cells in sea anemone. *BMC Biol.* **16**, 108 (2018).
53. A. D. Hargreaves, M. T. Swain, M. J. Hegarty, D. W. Logan, J. F. Mulley, Restriction and recruitment-gene duplication and the origin and evolution of snake venom toxins. *Genome Biol. Evol.* **6**, 2088–2095 (2014).
54. I. L. M. Junqueira-de-Azevedo *et al.*, Venom-related transcripts from *Bothrops jararaca* tissues provide novel molecular insights into the production and evolution of snake venom. *Mol. Biol. Evol.* **32**, 754–766 (2015).
55. J. Reyes-Velasco *et al.*, Expression of venom gene homologs in diverse python tissues suggests a new model for the evolution of snake venom. *Mol. Biol. Evol.* **32**, 173–183 (2015).
56. M. Y. Sachkova *et al.*, The birth and death of toxins with distinct functions: A case study in the sea anemone *Nematostella*. *Mol. Biol. Evol.* **36**, 2001–2012 (2019).
57. N. R. Casewell, G. A. Huttley, W. Wüster, Dynamic evolution of venom proteins in squamate reptiles. *Nat. Commun.* **3**, 1066 (2012).
58. S. D. Robinson *et al.*, Hormone-like peptides in the venoms of marine cone snails. *Gen. Comp. Endocrinol.* **244**, 11–18 (2017).
59. C. W. Gruber, J. Koehbach, M. Muttenthaler, Exploring bioactive peptides from natural sources for oxytocin and vasopressin drug discovery. *Future Med. Chem.* **4**, 1791–1798 (2012).
60. S. Kumar, M. Vijayasarathy, M. A. Venkatesha, P. Sunita, P. Balaram, Cone snail analogs of the pituitary hormones oxytocin/vasopressin and their carrier protein neurophysin. Proteomic and transcriptomic identification of conopressins and conophysins. *Bba-Proteins Proteom* **1868**, 140391 (2020).
61. C. McCowan, J. E. Garb, Recruitment and diversification of an ecdysozoan family of neuropeptide hormones for black widow spider venom expression. *Gene* **536**, 366–375 (2014).
62. J. M. Miwa *et al.*, *lynx1*, an endogenous toxin-like modulator of nicotinic acetylcholine receptors in the mammalian CNS. *Neuron* **23**, 105–114 (1999).
63. J. M. Miwa, K. R. Anderson, K. M. Hoffman, *Lynx* protoxins: Roles of endogenous mammalian neurotoxin-like proteins in modulating nicotinic acetylcholine receptor function to influence complex biological processes. *Front. Pharmacol.* **10**, 343 (2019).
64. E. A. B. Undheim *et al.*, Weaponization of a hormone: Convergent recruitment of hyperglycemic hormone into the venom of arthropod predators. *Structure* **23**, 1283–1292 (2015).
65. D. R. Nassel, M. Zandawala, T. Kawada, H. Satake, Tachykinins: Neuropeptides that are ancient, diverse, widespread and functionally pleiotropic. *Front. Neurosci. Switz* **13**, 1262 (2019).
66. J. Yi, D. Warunek, D. Craft, Degradation and stabilization of peptide hormones in human blood specimens. *PLoS One* **10**, e0134427 (2015).
67. V. Herzig, G. F. King, The cystine knot is responsible for the exceptional stability of the insecticidal spider toxin ω -Hexatoxin-Hv1a. *Toxins (Basel)* **7**, 4366–4380 (2015).
68. K. Kikuchi, M. Sugiura, T. Kimura, High proteolytic resistance of spider-derived inhibitor cystine knots. *Int. J. Pept.* **2015**, 537508 (2015).
69. B. Krishnarjuna *et al.*, Structure, folding and stability of a minimal homologue from *Anemonia sulcata* of the sea anemone potassium channel blocker ShK. *Peptides* **99**, 169–178 (2018).
70. B. Krishnarjuna *et al.*, Synthesis, folding, structure and activity of a predicted peptide from the sea anemone *Oulactis* sp. with an ShKT fold. *Toxicon* **150**, 50–59 (2018).
71. R. H. M. Ebberink, A. B. Smit, J. Vanminnen, The insulin family—Evolution of structure and function in vertebrates and invertebrates. *Biol. Bull.* **177**, 176–182 (1989).
72. S. Kojima *et al.*, Function of insulin in snail brain in associative learning. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* **201**, 969–981 (2015).
73. N. Kaplan, N. Murgio, M. Linial, Novel families of toxin-like peptides in insects and mammals: A computational approach. *J. Mol. Biol.* **369**, 553–566 (2007).
74. S. Sturm, D. Ramesh, A. Brockmann, S. Neupert, R. Predel, Agatoxin-like peptides in the neuroendocrine system of the honey bee and other insects. *J. Proteomics* **132**, 77–84 (2016).
75. E. Georgaka, V. Nastopoulos, S. Eleftheriou, I. K. Zarkadis, A toxin-like gene in rainbow trout: Cloning, expression, and gene organization. *Toxicon* **49**, 1002–1009 (2007).
76. T. A. Y. Ayoubi, W. J. M. Van De Ven, Regulation of gene expression by alternative promoters. *FASEB J.* **10**, 453–460 (1996).
77. B. Slobodin, R. Agami, Transcription initiation determines its end. *Mol. Cell* **57**, 205–206 (2015).
78. K. Oktaba *et al.*, ELAV links paused Pol II to alternative polyadenylation in the *Drosophila* nervous system. *Mol. Cell* **57**, 341–348 (2015).
79. P. Batut, A. Dobin, C. Plessy, P. Carninci, T. R. Gingeras, High-fidelity promoter profiling reveals widespread alternative promoter usage and transposon-driven developmental gene expression. *Genome Res.* **23**, 169–180 (2013).
80. D. Fleck, A. N. Garratt, C. Haass, M. Willem, BACE1 dependent neuregulin processing: Review. *Curr. Alzheimer Res.* **9**, 178–183 (2012).
81. L. N. Wei, P. Y. Law, H. H. Loh, Post-transcriptional regulation of opioid receptors in the nervous system. *Front. Biosci.* **9**, 1665–1679 (2004).
82. J. P. Boissel, P. M. Schwarz, U. Förstermann, Neuronal-type NO synthase: Transcript diversity and expressional regulation. *Nitric Oxide* **2**, 337–349 (1998).
83. U. Shinde, M. Inouye, Intramolecular chaperones: Polypeptide extensions that modulate protein folding. *Semin. Cell Dev. Biol.* **11**, 35–44 (2000).
84. D. R. Cool *et al.*, Carboxypeptidase E is a regulated secretory pathway sorting receptor: Genetic obliteration leads to endocrine disorders in *Cpe(fat)* mice. *Cell* **88**, 73–83 (1997).
85. W. S. Sossin, A. Sweet-Cordero, R. H. Scheller, Dale's hypothesis revisited: Different neuropeptides derived from a common prohormone are targeted to different processes. *Proc. Natl. Acad. Sci. U.S.A.* **87**, 4845–4848 (1990).
86. E. Renfer, U. Technau, Meganuclease-assisted generation of stable transgenics in the sea anemone *Nematostella vectensis*. *Nat. Protoc.* **12**, 1844–1854 (2017).
87. Y. Moran *et al.*, Neurotoxin localization to ectodermal gland cells uncovers an alternative mechanism of venom delivery in sea anemones. *Proc. Biol. Sci.* **279**, 1351–1358 (2012).
88. G. Genikhovich, U. Technau, In situ hybridization of starlet sea anemone (*Nematostella vectensis*) embryos, larvae, and polyps. *Cold Spring Harb. Protoc.* **2009**, pdb.prot5282 (2009).
89. N. M. Davidson, A. D. K. Hawkins, A. Oshlack, SuperTranscripts: A data driven reference for analysis and visualisation of transcriptomes. *Genome Biol.* **18**, 148 (2017).
90. L. S. Babonis, M. Q. Martindale, J. F. Ryan, Do novel genes drive morphological novelty? An investigation of the nematosomes in the sea anemone *Nematostella vectensis*. *BMC Evol. Biol.* **16**, 114 (2016).
91. A. Dobin *et al.*, STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
92. R. Mousa, S. Lansky, G. Shoham, N. Metanis, BPTI folding revisited: Switching a disulfide into methylene thioacetal reveals a previously hidden path. *Chem. Sci. (Camb.)* **9**, 4814–4820 (2018).
93. J. Cox, M. Mann, MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).
94. Y. Perez-Riverol *et al.*, The PRIDE database and related tools and resources in 2019: Improving support for quantification data. *Nucleic Acids Res.* **47**, D442–D450 (2019).
95. M. Y. Sachkova, Y. Moran, Dynamics of venom composition across a complex life cycle. ProteomeXchange Consortium via PRIDE partner repository. <https://www.ebi.ac.uk/pride/archive/projects/PXD008218>. Deposited 15 January 2018.