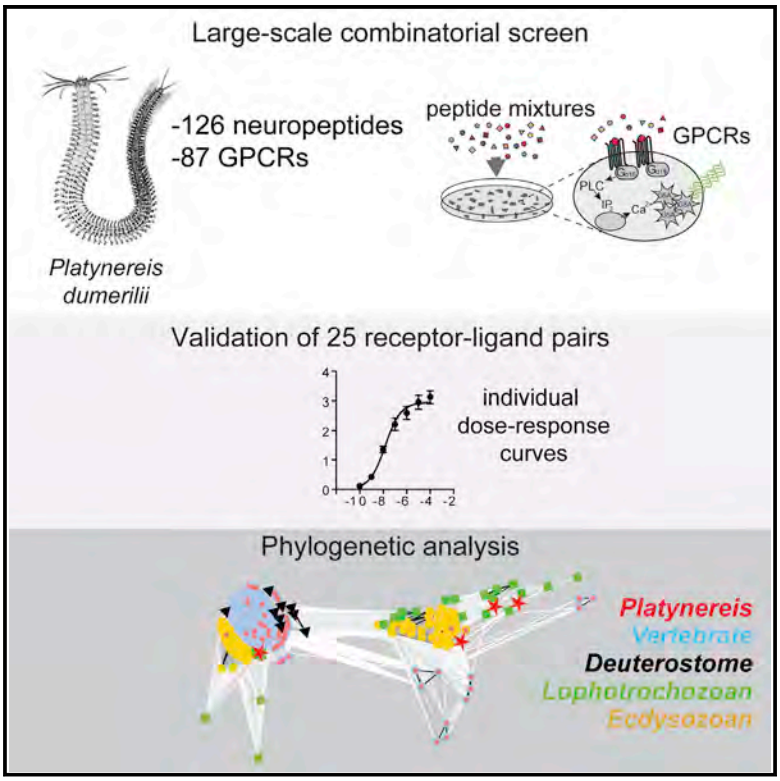
Resource

Cell Reports

Large-Scale Combinatorial Deorphanization of *Platynereis* Neuropeptide GPCRs

Graphical Abstract



Authors

Philipp Bauknecht, Gáspár Jékely

Correspondence

gaspar.jekely@tuebingen.mpg.de

In Brief

Bauknecht and Jékely report the deorphanization of 19 neuropeptide G-protein-coupled receptors from the marine annelid *Platynereis*. Among them are members of previously uncharacterized families. This work provides information about the evolution of peptidergic systems and neuropeptide signaling in bilaterians.

Highlights

- 19 GPCRs from *Platynereis* were deorphanized
- Ligands for previously uncharacterized GPCR families were found
- A *Platynereis* ortholog of thyrotropin-releasing hormone was identified
- Conserved D-peptides activate an ancient bilaterian family of achatin receptors

Accession Numbers

KP293941–KP294026
KP420212–KP420214

Bauknecht & Jékely, 2015, Cell Reports *12*, 1–10
July 28, 2015
http://dx.doi.org/10.1016/j.celrep.2015.06.052

CellPress

Please cite this article in press as: Bauknecht and Jékely, Large-Scale Combinatorial Deorphanization of *Platynereis* Neuropeptide GPCRs, Cell Reports (2015), http://dx.doi.org/10.1016/j.celrep.2015.06.052

Cell Reports
Resource

Large-Scale Combinatorial Deorphanization of *Platynereis* Neuropeptide GPCRs

Philipp Bauknecht[1] and Gáspár Jékely[1,*]
[1]Max Planck Institute for Developmental Biology, Spemannstrasse 35, 72076 Tübingen, Germany
*Correspondence: gaspar.jekely@tuebingen.mpg.de
http://dx.doi.org/10.1016/j.celrep.2015.06.052

SUMMARY

Neuropeptides, representing the largest class of neuromodulators, commonly signal by G-protein-coupled receptors (GPCRs). While the neuropeptide repertoire of several metazoans has been characterized, many GPCRs are orphans. Here, we develop a strategy to identify GPCR-peptide pairs using combinatorial screening with complex peptide mixtures. We screened 126 neuropeptides against 87 GPCRs of the annelid *Platynereis* and identified ligands for 19 receptors. We assigned many GPCRs to known families and identified conserved families of achatin, FMRFamide, RGWamide, FLamide, and elevenin receptors. We also identified a ligand for the *Platynereis* ortholog of vertebrate thyrotropin-releasing hormone (TRH) receptors, revealing the ancient origin of TRH-receptor signaling. We predicted ligands for several metazoan GPCRs and tested predicted achatin receptors. These receptors were specifically activated by an achatin D-peptide, revealing a conserved mode of activation. Our work establishes an important resource and provides information about the complexity of peptidergic signaling in the urbilaterian.

INTRODUCTION

Neuropeptides represent the largest and most diverse class of neuron-secreted signaling molecules. These peptides can have neuromodulatory, neurotransmitter, or hormonal functions and can affect development, physiology, and the activity in neural circuits. The majority of neuropeptides signal by G-protein-coupled receptors (GPCRs), with some exceptions (Chang et al., 2009; Leung et al., 1987; Lowe et al., 1989; Rechler and Nissley, 1985). While the neuropeptide repertoire of an animal can be determined using a combination of sequencing and mass-spectrometry approaches (Collins et al., 2010; Conzelmann et al., 2013a; Dircksen et al., 2011; Hauser et al., 2010; Li et al., 2008; Xie et al., 2010), the determination of neuropeptide receptors is more difficult and is usually carried out using in vitro experiments with individual peptide-receptor pairs.

Several years of effort have led to the identification of ~35 neuropeptide GPCRs in *Drosophila melanogaster* (Caers et al., 2012; Hewes and Taghert, 2001), 23 in *Caenorhabditis elegans* (Frooninckx et al., 2012), 50 in human and mouse, and only a few in non-model organisms (Bigot et al., 2014; Conzelmann et al., 2013b; Cox et al., 1997; Kim et al., 2010; Tensen et al., 1998a, 1998b).

Among the lophotrochozoans, an animal superphylum that with ecdysozoans and deuterostomes forms the Bilateria (Telford and Copley, 2011), peptidergic neuromodulation has been extensively studied in several species (Cropper et al., 1987; Kamatani et al., 1989; Rajpara et al., 1992; Hoek et al., 2005; Willows et al., 1997). However, the lack of information regarding neuropeptide receptors hinders the identification of the downstream signaling mechanisms underlying neuromodulation.

The annelid *Platynereis* has emerged in recent years as a powerful lophotrochozoan laboratory animal for the study of development, neuronal circuits, and zooplankton behavior (Jékely et al., 2008; Randel et al., 2014; Tosches et al., 2014; Zantke et al., 2014). Its larval stages represent accessible models for studying the role of neuropeptides in behavior, development, and physiology at the whole-organism level (Conzelmann et al., 2011, 2013b; Williams et al., 2015). *Platynereis* has an ancestral neuropeptide repertoire, including 30 ancestral bilaterian pro-neuropeptide families (Conzelmann et al., 2013a); however, only one neuropeptide receptor has been identified so far (Conzelmann et al., 2013b).

Here, building on established transcriptomic and peptidomic resources (Conzelmann et al., 2013a), we present a large-scale deorphanization screen of *Platynereis* neuropeptide GPCRs. We identified the peptide ligand of 19 *Platynereis* receptors. We also perform a phylogenetic analysis of *Platynereis* and other metazoan neuropeptide GPCRs to gain insights into the evolution of peptidergic signaling in bilaterians.

RESULTS

Combinatorial Screening for *Platynereis* GPCR-Neuropeptide Ligand Pairs

To facilitate the rapid identification of neuropeptide GPCRs, we developed a combinatorial cell-culture-based screening strategy (Figure 1).

We reasoned that complex peptide mixtures could be used to identify receptors activated by specific peptides present in these mixtures. Mixtures of subsets of peptides (submixtures) would

Please cite this article in press as: Bauknecht and Jékely, Large-Scale Combinatorial Deorphanization of *Platynereis* Neuropeptide GPCRs, Cell Reports (2015), http://dx.doi.org/10.1016/j.celrep.2015.06.052

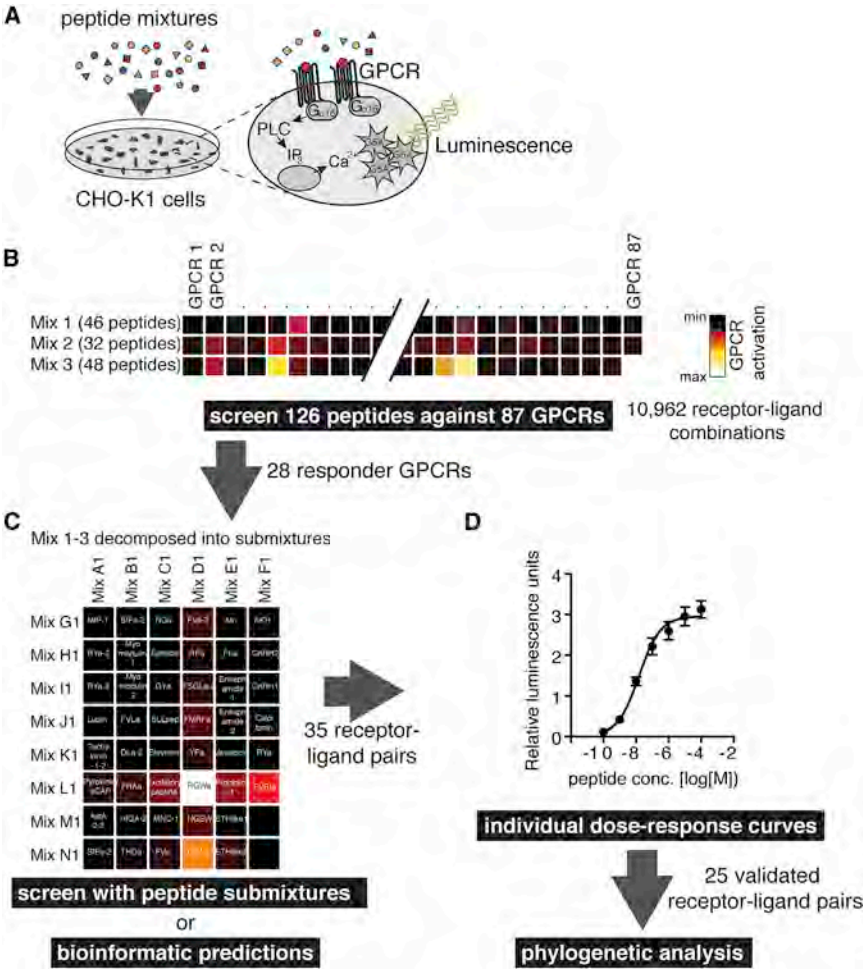


Figure 1. Schematic of the Combinatorial Screen for *Platynereis* Neuropeptide GPCRs

(A) We used a cell-culture assay with CHO-K1 cells stably transfected with a calcium-sensitive bioluminescent GFP-aequorin fusion protein (G5A).

(B) In the primary screen, 87 *Platynereis* GPCRs were tested against three complex neuropeptide mixtures. This screen identified 28 responder GPCRs.

(C) A secondary screen based on peptide submixtures and bioinformatic predictions identified 35 receptor-ligand pairs. See also Figure S1.

(D) Individual peptide-receptor dose-response measurements validated 25 receptor-ligand pairs. A phylogenetic analysis provided information about GPCR-ligand coevolution across bilaterians. See also Figure S1 and Table S1.

then allow the unambiguous identification of a single active peptide.

We screened 87 orphan *Platynereis* GPCRs (Table S1) against 126 *Platynereis* neuropeptides, pooled into three mixtures based on peptide pI and solubility (Table S2). We used a Ca^{2+}-mobilization assay with CHO-K1 cells stably transfected with a calcium-sensitive bioluminescent GFP-aequorin fusion protein (Tunaru et al., 2005). This screen identified 28 GPCRs that responded reproducibly to at least one of the peptide mixtures (Figure 2; Table S3).

We focused on these receptors and tested them with peptide mixtures derived by decomposing the original active mixture into the rows and columns of a matrix (Tables S2 and S3). If two mixtures representing a row and a column in the matrix activate the receptor, the peptide at the intersect likely represents a specific ligand. We defined an activation value for each peptide in the matrix as the square root of the product of the measured values of two intersecting mixtures (Figure 1; Figure S1). Using this approach, we found candidate receptors for the neuropeptides FLa (also activated by prokineticin short peptide 1 [SP-1]), allatotropin, neuropeptide KY (NKY), luqin, elevenin (two receptors), RGWa, FMRFa, excitatory peptide, neuropeptide-Y-4 (NPY-4), achatin, EFLGamide (EFLGa), and diuretic hormone 31 (DH31) (Figure S1; Table S3).

For six receptors, we were able to predict the ligands based on orthology relationships (see below). These included candidate receptors for vasotocin, allatostatin-A, allatotropin, myomodulin, allatostatin-C, and DH31 (Table S3). The GPCR-ligand pairs identified either by combinatorial screening or bioinformatic prediction were further tested in individual receptor ligand assays (Table S3).

Validation of 25 GPCR-Ligand Pairs

To confirm that the identified peptides are indeed specific ligands to the respective GPCRs, we recorded dose-response curves for each of the identified receptor-ligand pairs. After excluding a few receptors with inconsistent activation in replicate experiments, we identified 25 GPCR-ligand pairs with half maximal effective concentration (EC_{50}) values in the nanomolar or low-micromolar range (Table 1; Figure 3).

We recorded two sets of dose-response curves with different normalizations, using responses of either the *Platynereis* MIP receptor or an endogenously expressed histamine receptor as reference (Figure 3; Figure S2). For three receptors, we identified two or three ligands derived from distinct precursors (FLamide, NKY, and NPY-4 receptors), and for four peptides (allatotropin, FMRFamide, elevenin, and DH31), we identified two receptors. Overall, we established neuropeptide ligands for 19 *Platynereis* GPCRs (Table 1).

To partially characterize the signaling mechanism of these receptors, we tested them in the same assay, but without co-transfecting the Gα-16 construct. If receptor activation leads to Ca^{2+} increase without the promiscuous G protein, this indicates that the GPCR couples to the endogenous Gq-α to activate the phospholipase C (PLC)/inositol trisphosphate (IP_3)/Ca^{2+} release pathway. Ligand stimulation of ten receptors (allatotropin-1 and -2, FLamide, FMRFamide, luqin, elevenin-1 and -2, DH31-2, and vasotocin receptors) produced Ca^{2+} signals without Gα-16,

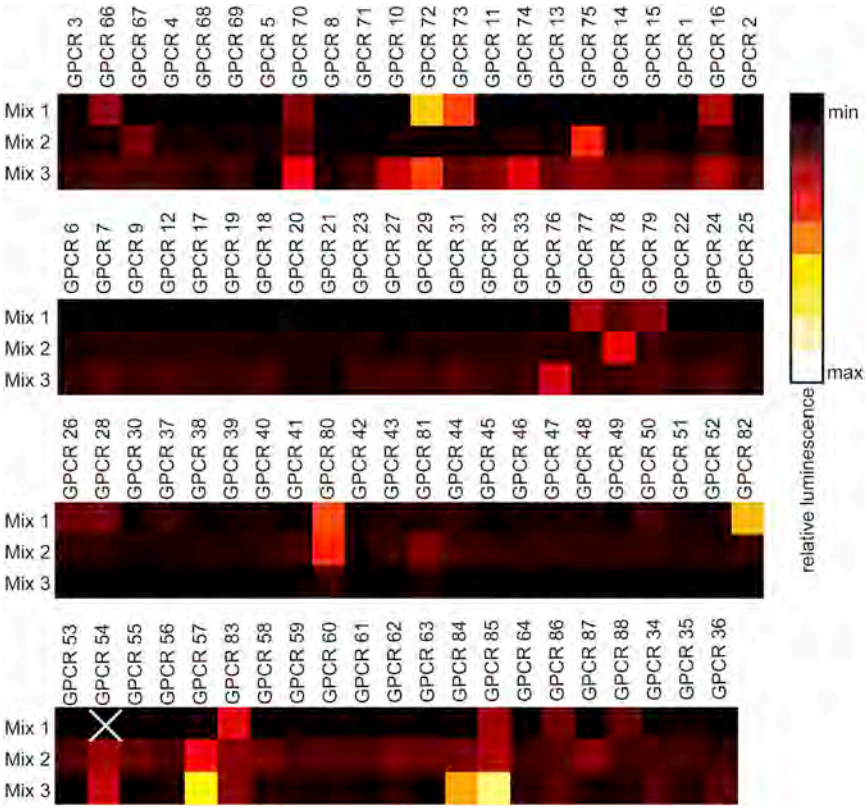


Figure 2. Primary Screen of 87 *Platynereis* GPCRs against Three Complex Peptide Mixtures

28 responder GPCRs that showed consistent activation in three replicate experiments were studied further. Mean relative luminescence values from three replicates are shown with color-coding. GPCR54 was not tested with Mix1. See also Table S3.

suggesting that these receptors couple to Gq-α (Table 1; Table S3). Given the limitations of the GFP-aequorin assay, we did not test Gi-α and Go-α signaling.

Phylogenetic Analysis of *Platynereis* Neuropeptide GPCRs

Recent bioinformatic analyses found strong support for the long-term stability of GPCR-neuropeptide ligand pairs across animal phyla (Janssen et al., 2010; Jékely, 2013; Mirabeau and Joly, 2013; Park et al., 2002). However, given the limited knowledge of GPCR ligands among the lophotrochozoans, it has not been possible to rigorously assess receptor-ligand coevolution across all three superphyla of Bilateria. Our deorphanized GPCR resource provides a large-scale dataset to test the generality of inter-phyletic receptor-ligand coevolution.

To identify orthologs of the 87 *Platynereis* GPCRs used in the screen, we performed similarity-based clustering, a method previously shown to be an efficient means of recovering orthologous groups of GPCRs (Jékely, 2013), with results similar to tree-based molecular phylogenetic analyses (Mirabeau and Joly, 2013). First, we seeded BLASTP searches with the *Platynereis* GPCR sequences in metazoan genomes. We also collected further representative neuropeptide GPCRs, including an annotated list of GPCRs whose peptide ligands have been experimentally characterized. We then separately clustered the 68 orphan *Platynereis* receptors and the 19 deorphanized *Platynereis* receptors with their respective BLASTP hits and further GPCR representatives (Figure 4; Figure S3). Clustering analysis of the 68 orphan receptors identified orthologs for 20 of them with a known peptide ligand in another species (Figure S3; Table S1). For 14 of these GPCRs, the corresponding *Platynereis* peptides were present in our mixtures. There can be several reasons why we did not see activation for these receptor-peptide pairs: (1) GPCR expressions may have failed, (2) the ligands may not have been dissolved or were unstable, (3) the ligand changed during evolution, or (4) some clusters may contain closely related paralogous receptors with different ligands.

Clustering analysis of the 19 deorphanized *Platynereis* GPCRs allowed us to assign all of them to well-resolved sequence groups (Figure 4; Figure S3). Many of these sequence clusters represented established orthology groups of metazoan GPCRs (Jékely, 2013; Mirabeau and Joly, 2013) containing already-characterized GPCRs. In most cases, the *Platynereis* receptor was activated by a peptide that was a previously recognized ortholog of known peptide ligands in that cluster (allatostatin-A, allatostatin-C, vasopressin/oxytocin, allatotropin/orexin, luqin/insect-RYamide, and DH31/calcitonin). These *Platynereis* receptors, together with previously identified mollusk receptors (Bigot et al., 2014; Tensen et al., 1998a, 1998b), represent deorphanized lophotrochozoan members of their respective families. These examples provide further evidence for receptor-ligand conservation during evolution.

Furthermore, our GPCR analysis could confirm the orthology relationships of some annelid neuropeptide families that had previously been proposed based on peptide-sequence similarity alone. Annelid and mollusk myomodulins (Cropper et al., 1987; Veenstra, 2011) were suggested to be orthologs of arthropod myosuppressins (Holman et al., 1986) based on limited peptide similarity (Jékely, 2013). The orthology of the *Platynereis* myomodulin receptor to the arthropod myosuppressin receptor confirms this (Figure 4).

Similarly, annelid excitatory peptides (Oumi et al., 1995) were suggested to belong to the bilaterian CCHamide/neuromedin-B family (Jékely, 2013; Roller et al., 2008). The *Platynereis* excitatory peptide receptor clusters with CCHamide/neuromedin-B receptors, confirming this (Figure 4).

The NKY receptors we identified are related to bilaterian NPY/NPF receptors (Bigot et al., 2014; Mertens et al., 2002) and their paralogs, the short neuropeptide F (sNPF) receptors from

Please cite this article in press as: Bauknecht and Jékely, Large-Scale Combinatorial Deorphanization of *Platynereis* Neuropeptide GPCRs, Cell Reports (2015), http://dx.doi.org/10.1016/j.celrep.2015.06.052

Table 1. Summary of Validated Receptor-Ligand Pairs

Receptor Name	Ligand Name	Sequence of Tested Ligand	EC_{50} Normalized to Histamine Receptor Response	EC_{50} Normalized to MIP Receptor Response	Activation without Gα-16
Allatotropin receptor1	allatotropin	GFRTGAYDRFSHGF-NH2	116 nM	840 nM	yes
Allatotropin receptor2	allatotropin	GFRTGAYDRFSHGF-NH2	7.8 nM	6.0 nM	yes
EFLGa receptor1	EFLGa	FSEFLG-NH2	350 nM	5.4 nM	no
FLamide receptor1	FLamide	AKYFL-NH2	8.9 nM	2.0 nM	yes
FLamide receptor1	prokineticin-short peptide1	GRSRPLFV-NH2	47 nM	390 nM	yes
FMRFamide receptor1	FMRFamide	FMRF-NH2	78 nM	1.5 nM	yes
NKY receptor1	NKY-1	KAFWQPMMGGPLPVETRLASFGS RIEPDRTEPGSGPNGIKAMRY-NH2	120 nM	120 nM	no
NKY receptor1	NKY-2	NNGIWIWMPAQGYVSVPHQQEGG AADEGKPGKIMRY-NH2	410 nM	390 nM	no
NKY receptor1	FMRFamide	FMRF-NH2	1.4 μM	840 nM	no
NPY-4 receptor1	NPY-4	DPSFISSGPPVRPSSFKSPEELMEY LQKVRAYYNVMSRPRF-NH2	350 nM	110 nM	no
NPY-4 receptor1	NPY-3	pGluNMEGPPPRPAIFRTPQELRDY LSDLNEYFMIVGRPRF-NH2	630 nM	1.0 μM	no
NPY-4 receptor1	NPY-1	KVLEEMPTLQQIPLKPVRPNRFRNK DELHSYLQSLRDYYSVIGRPRF-NH2	420 nM	3.7 μM	no
Luqin receptor1	luqin	WRPQGRF-NH2	5.2 nM	0.86 nM	yes
RGWamide receptor1	RGWamide	RGW-NH2	2.9 nM	10 nM	no
Excitatory peptide receptor1	excitatory peptide	KCSGQWAIHACAGGN-NH2	7.9 nM	15 nM	no
Allatostatin-A receptor1	allatostatin-A-2-2	NDALKFSGL-NH2	12 μM	15 μM	no
Elevenin receptor1	elevenin (L11)	PDCTRFVFHPSCRGVAA	62 nM	120 nM	yes
Elevenin receptor2	elevenin (L11)	PDCTRFVFHPSCRGVAA	1.3 nM	2.3 nM	yes
Achatin receptor1	D-Achatin	G{dF}GD	120 nM	150 nM	no
Achatin receptor1	L-Achatin	GFGD	not available	11 μM	no
Myomodulin receptor1	myomodulin-2	AMGMLRM-NH2	26 nM	9.6 nM	no
Myomodulin receptor1	myomodulin-1	AMSMLRM-NH2	10 nM	~10 nM	no
DH31 receptor1	DH31	RIDAGYGSRYAAGASVGSKLRALK QAADWNGP-NH2	180 nM	87 nM	no
DH31 receptor2	DH31	RIDAGYGSRYAAGASVGSKLRALK QAADWNGP-NH2	34 nM	15 nM	yes
Vasotocin receptor1	vasotocin	CFVRNCPPG-NH2	1.1 μM	920 nM	yes
Allatostatin-C receptor1	allatostatin-C	pGluPVQCLVNIVSCW-NH2	1.0 μM	1.2 μM	no
Aplysia achatin receptor	*Aplysia* D-achatin	G{d-F}FD	14 nM	62 nM	no
Aplysia achatin receptor	*Aplysia* L-achatin	GFFD	not available	190 μM	no
Branchiostoma Achatin receptor	*Branchiostoma* D-Achatin	G{d-F}GN	0.87 nM	2.4 nM	yes
Branchiostoma achatin receptor	*Branchiostoma* L-Achatin	GFGN	not available	>1 M	not available
Saccoglossus achatin receptor	*Saccoglossus* D-Achatin	G{d-F}GN	16 nM	27 nM	no
Saccoglossus achatin receptor	*Saccoglossus* L-Achatin	GFGN	15 μM	13 μM	no

Name and sequence of the tested ligands are shown. pGlu indicates N-terminal pyroglutamylation. -NH2 indicates C-terminal amidation. Cys residues that form disulfide bonds are underlined. EC_{50} values of dose-response curves are shown for the two different normalizations. All receptors were also tested without cotransfecting the promiscuous Gα-16. Activation in the absence of Gα-16 in the GFP-aequorin assay indicates that the receptor couples to the Gq-α protein endogenously present in the CHO cells.

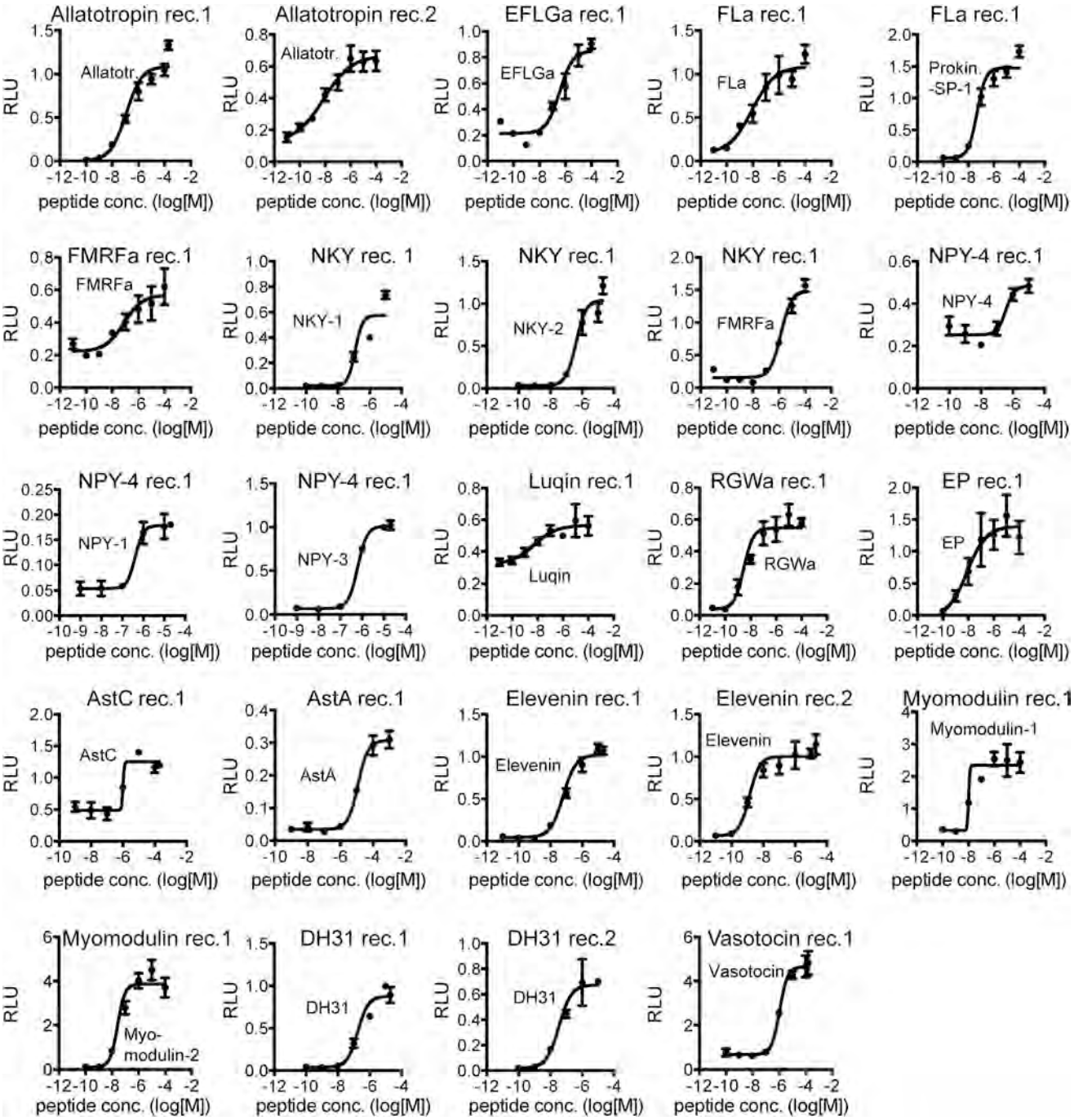


Figure 3. Dose-Response Curves of *Platynereis* Deorphanized GPCRs Treated with Varying Concentrations of Peptides

Data, representing luminescence units relative to the control response (1 mM histamine), are shown as mean ± SEM (n = 3). Dose-response curves fitted to the data are shown. Ligand names are shown beside the curves. EC_{50} values are listed in Table 1. RLU, relative luminescence unit; AstC, allatostatin-C; AstA, allatostatin-A; EP, excitatory peptide. See also Figure S2.

mollusks and insects (Bigot et al., 2014). The NKY peptides of annelids and mollusks show similarity to NPY/NPF peptides, including the RF/Yamide motif, a proline-rich stretch, and an acidic stretch (Conzelmann et al., 2013a) (Figure S4A). These results establish NKY and NPY/NPF peptides as paralogs.

Ligand Discoveries for Uncharacterized GPCR Families

We also identified six GPCR clusters where the only member with a known ligand was one of the deorphanized *Platynereis* GPCRs. These included clusters of sequences orthologous to *Platynereis* FMRFamide, L11/elevenin, achatin, RGWamide,

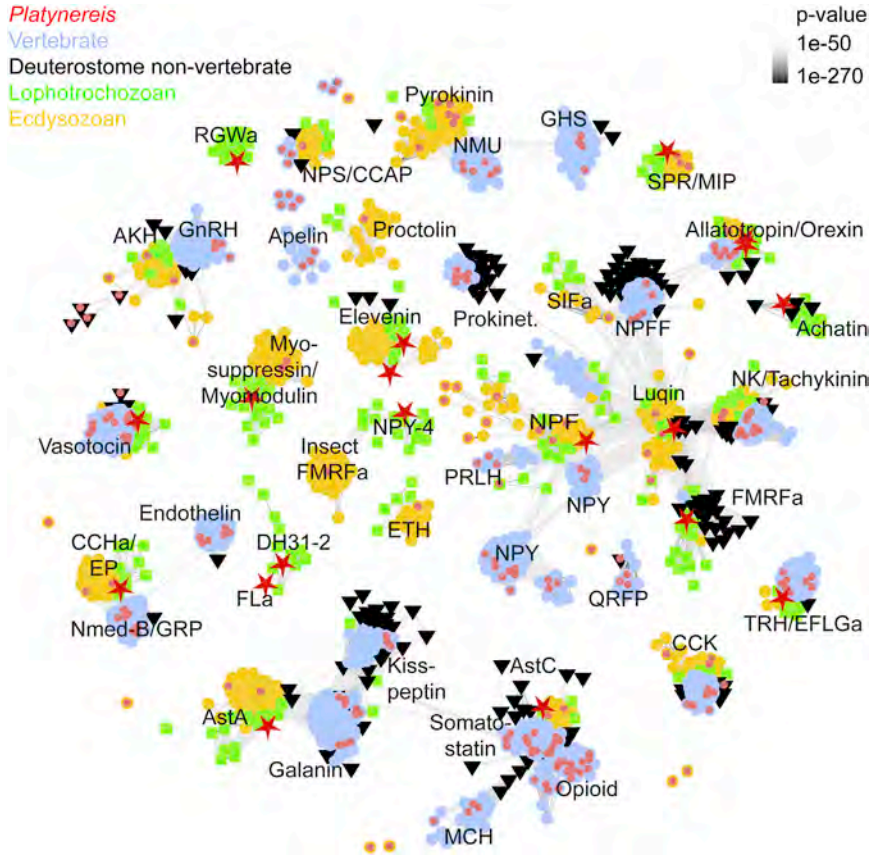


Figure 4. Sequence-Similarity-Based Clustering of Neuropeptide Class A GPCRs

Nodes represent sequences, and edges represent BLASTP connections. Edges are colored according to BLASTP p values. Nodes are colored based on taxonomy. Only deorphanized *Platynereis* GPCRs were included and are indicated as red stars. The *Platynereis* MIP receptor was described previously (Conzelmann et al., 2013b). Deorphanized receptors from other species are marked with a small red dot. Clusters are named according to the name of the deorphanized family members. DH31-receptor1 and other class B GPCRs were clustered separately and are shown in Figure S3A. AKH, adipokinetic hormone; AstC, allatostatin-C; AstA, allatostatin-A; CCK, cholecystokinin; EP, excitatory peptide; ETH, ecdysis triggering hormone; GnRH, gonadotropin-releasing hormone; MCH, melanin-concentrating hormone; MIP, myoinhibitory peptide; Nmed-B, neuromedin-B; NMU, neuromedin-U; NPF/Y, neuropeptide F/Y; NPFF, neuropeptide FF; PRLH, prolactin releasing hormone; QRFP, pyroglutamylated RFamide peptide; SK, substance-K; SPR, sex peptide receptor; TRH, thyrotropin releasing hormone. The Clans file is available at https://github.com/JekelyLab/GPCR_Clans_Maps. See also Figures S3, S4, and S6.

FLamide, DH31-rec2 and NPY-4 receptors (Figure 4). These clusters represent conserved neuropeptide GPCR families with an identified ligand in *Platynereis*.

Three of the identified families (FMRFamide, elevenin, and achatin) have both protostome and deuterostome orthologs, representing ancient bilaterian orthology groups (Figure 4). In contrast, the GPCR clusters containing the *Platynereis* RGWamide and FLamide receptors are restricted to lophotrochozoans (Figure 4).

The FMRFamide receptor cluster we identified is related to luqin receptors and contains several mollusk and annelid sequences, as well as sequences from the non-vertebrate deuterostomes *Branchiostoma* and *Saccoglossus* (Figure 4). Interestingly, we also identified a GPCR belonging to this group from the cartilaginous fish, the elephant shark *Callorhinchus milii*. The elephant shark genome represents the slowest evolving vertebrate genome thus far identified (Venkatesh et al., 2014). We also identified FMRFamide-receptor orthologs from a hemipteran insect, but no other arthropods (Table S4). The FMRFamide-receptor family thus represents a conserved bilaterian family that has been lost in most vertebrates and arthropods but is retained in lophotrochozoans and non-vertebrate deuterostomes. An FMRFamide receptor has also been identified in *Drosophila* (Cazzamali and Grimmelikhuijzen, 2002), but this sequence belongs to an arthropod-specific group and is not closely related to the FMRFamide receptors we describe here (Figure 4). In mollusks, FaNaCs, members of the DEG/ENaC family, have been identified as FMRFamide receptors (Lingueglia et al., 1995), but no FMRFamide GPCR has yet been found.

The receptors for *Platynereis* L11/elevenin peptide also belonged to a conserved bilaterian family with members in nematodes, insects, annelids, mollusks, *Branchiostoma*, and *Saccoglossus* (Figure 4; Table S4). Elevenin orthologs are known from annelids (Veenstra, 2011), mollusks (Veenstra, 2010), and nematodes and arthropods (Jékely, 2013), but we predict that they are also present in some non-vertebrate deuterostomes.

The receptor for *Platynereis* achatin clustered with orthologs from mollusks, annelids and non-vertebrate deuterostomes (Figure 4). Achatin was described from mollusks (Veenstra, 2010), annelids (Veenstra, 2011), and *Branchiostoma* and *Saccoglossus* (Jékely, 2013), but has been lost from vertebrates and most ecdysozoans (it is present in the chelicerate *Stegodyphus mimosarum*) (Figure S4B). The identification of a receptor family showing the same phyletic distribution establishes the achatin receptor-ligand pair as an ancient bilaterian system.

RGWamide and FLamide peptides are known from annelids and mollusks (Conzelmann et al., 2013a), and we identified distinct receptor clusters containing the *Platynereis* receptors and other annelid and mollusk sequences (Figure 4; Table S4). We also identified a second DH31 peptide receptor, closely related to FLamide receptor. We could not identify members of these receptor families outside the lophotrochozoans.

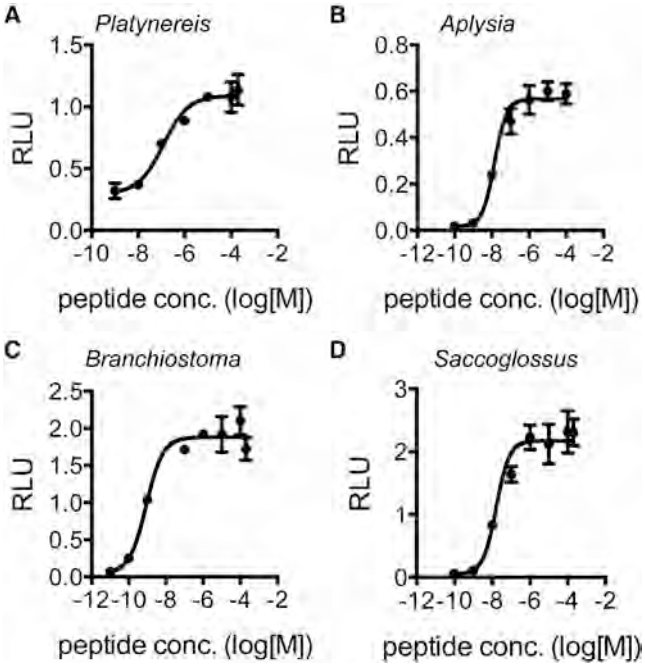


Figure 5. Dose-Response Curves of Bilaterian Achatin Receptors Treated with Varying Concentrations of D-achatin Peptides

The *Platynereis* (A), *Aplysia* (B), *Branchiostoma* (C), and *Saccoglossus* (D) receptors were tested with the species-specific achatin peptide containing a D-Phe. Data, representing luminescence units relative to the control response (1 mM histamine), are shown as mean ± SEM (n = 3). Dose-response curves fitted to the data are shown. EC_{50} values are listed in Table 1. Responses to L-achatin peptides and a separate set of measurements normalized to MIP receptor are shown in Figure S5.

A *Platynereis* Ortholog of Thyrotropin-Releasing Hormone

The identification of the ligand for the annelid ortholog of vertebrate thyrotropin-releasing hormone receptors sheds light on the evolution of this family. Thyrotropin-releasing hormones (TRHs) have so far only been identified in deuterostomes. However, the presence of GPCRs in some protostomes showing orthology to deuterostome TRH receptors suggested that TRH orthologs are present in some protostomes (Jékely, 2013; Mirabeau and Joly, 2013). We identified EFLGa as the ligand of the *Platynereis* TRH receptor ortholog (Figure 4; Table 1). EFLGa has already been described in *Platynereis*, other annelids, and mollusks (Conzelmann et al., 2013a), but its identity as a potential TRH ortholog was not recognized, since the sequence of the mature peptide (FSEFLGamide) is not similar to vertebrate TRH (pQHPamide, with pQ indicating pyroglutamate). Intriguingly, however, *Platynereis* EFLGa shows some similarity to the TRH ortholog of the sea urchin *Strongylocentrotus purpuratus* (Rowe and Elphick, 2012) (Figure S4C). Uniquely among the deuterostomes, the sea urchin peptide (Q[W/Y]PGamide) is a Gamide. This sea urchin sequence shows intermediate characteristics and bridges the gap between the protostome and deuterostome families, further suggesting the orthology of *Platynereis* EFLGa and deuterostome TRH.

An Ancient Bilaterian Family of Achatin Receptors Activated by a D-peptide

Our sequence analyses provide additional support for the widespread conservation of neuropeptide-GPCR signaling pairs. This allows us to predict ligands for several lophotrochozoan GPCRs, including receptors from *Capitella teleta*, *Aplysia californica*, and *Crassostrea gigas*, and deuterostome receptors from *Branchiostoma* and *Saccoglossus* (Table S4).

To test our predictions, we focused in more detail on the achatin family. Achatin receptors represent one of the ancient bilaterian families we identified (Figure 4), allowing us to test the feasibility of ligand predictions across Bilateria. We performed activation assays with putative achatin receptors from the sea slug *A. californica*, and the deuterostomes *S. kowalewskii* and *B. floridae* (Figure 5; Figure S5).

Achatins are 4-amino-acid peptides that share the G[FYM][GAF][DNG] motif (Figure S4B). Achatin was identified in the giant snail Achatina fulica (Kamatani et al., 1989) and was shown to contain a D-amino acid (Gly-D-Phe-Ala-Asp). Achatin is a potent neuroexcitatory peptide, and this activity is specific to the D-form. We therefore also tested species-specific achatin ligands synthesized with a D-Phe. We found that D-achatins activated all receptors with EC_{50} values in the nanomolar range, but L-achatins were poor agonists (in the high micromolar, millimolar range) (Figure 5; Figure S5; Table1).

These results show that the D-form of achatin has been conserved throughout evolution as a ligand for the bilaterian orthology group of achatin GPCRs. This indicates that our receptor-ligand predictions (Table S4) are reliable and can be used to predict receptor-ligand pairs across Bilateria.

DISCUSSION

Here, we described a large-scale screen for neuropeptide GPCRs in *Platynereis*. Our combinatorial strategy allowed us to quickly screen 10,962 receptor-ligand combinations without the need to assay all combinations individually. We could identify specific receptor-ligand pairs and study them in individual assays. However, measurements with peptide mixtures also revealed the high specificity of the interactions. We screened each receptor against 126 neuropeptides, but we found strong activation by only one or two related peptides. This strategy is generally applicable for GPCR ligand screens and could speed up ligand discovery. Here, we reported 19 deorphanized receptors and 25 validated receptor-ligand pairs from *Platynereis*. Based on these results, we now provide an updated overview (Jékely, 2013) of the phyletic distribution of peptides and peptide receptors in metazoans (Figure S6).

Importantly, many of the receptors we found represent GPCR families for which the ligand could not have been predicted based on available data. As more ligand-receptor pairs are discovered, however, ligand predictions will become increasingly straightforward.

Our results illuminate large, poorly studied areas of the GPCR sequence space within the lophotrochozoans, where only few receptors have been characterized biochemically.

The *Platynereis* GPCR-ligand pairs and our bioinformatic analyses provide further evidence for the long-term coevolution of

neuropeptides and their receptors across bilaterians. One exception may be the FMRFamide receptors of arthropods (Cazzamali and Grimmelikhuijzen, 2002). These GPCRs evolved in stem arthropods but respond to FMRFamides derived from an older FMRFamide precursor. The presence of the ancestral bilaterian FMRFamide receptor in hemipterans indicates that the two receptors coexisted for some time, but the ancestral receptor was lost from most arthropod genomes.

Receptor-ligand conservation allowed us to predict the peptide ligand for many yet uncharacterized receptors from lophotrochozoans and non-vertebrate deuterostomes. Using achatin as an example, we demonstrated that ligand predictions work across phyla.

Our results also provide information about the complexity of neuroendocrine signaling in the urbilaterian. We present receptor or ligand evidence for urbilaterian peptidergic systems, including TRH, elevenin, FMRFamide, and achatin signaling. The presence of TRH orthologs in annelids and mollusks is particularly interesting and begs the question whether TRH signaling regulates thyroid hormone synthesis in these animals, similar to its function in some vertebrates (Laudet, 2011). Thyroid hormones have been described from *Aplysia* (Heyland et al., 2006) and may have ancestrally regulated postembryonic developmental transitions in bilaterians (Laudet, 2011). This possibility is supported by the parallel loss of TRH and thyroid hormone receptors from the ecdysozoans that use ecdysone to orchestrate life-cycle transitions (Laudet, 2011).

The deorphanized GPCR dataset we describe here represents a valuable resource for the study of neuropeptide signaling in invertebrates, including annelids and mollusks. Similar combinatorial screening strategies could also be used for other species and receptor classes and could speed up GPCR ligand discovery.

EXPERIMENTAL PROCEDURES

Gene Identification and Receptor Cloning

Platynereis genes were identified from a *Platynereis* mixed-stages transcriptome assembly (Conzelmann et al., 2013a). GPCRs were cloned from cDNA or expressed sequence tag clones into pcDNA3.1(+) (Thermo Fisher Scientific). Forward primers consisted of a spacer (5′-ACAATA-3′) followed by a BamHI or EcoRI restriction site, the Kozak consensus sequence (5′-CGCCACC-3′), the start codon (5′-ATG-3′) and a sequence corresponding to the target sequence. Reverse primers consisted of a spacer (5′-ACAATA-3′), a NotI restriction site, a STOP codon, and reverse complementary sequence to the target sequence. Primers were designed to end with a C or G with 72°C melting temperature. PCR was performed using Phusion polymerase (New England Biolabs GmbH).

Open reading frames coding for achatin GPCRs from *Aplysia* (XP_005106606.1), *Branchiostoma* (XM_002600016.1), and *Saccoglossus* (XM_006815704.1) were generated by gene synthesis (GenScript). The sequence of the *Branchiostoma* receptor was complemented based on information from Metazome v3.0.

Cell Culture and Receptor Deorphanization

CHO-K1 cells were kept in Ham's F12 Nut Mix medium (Thermo Fisher Scientific) with 10% fetal bovine serum and PenStrep. We used a stable cell line expressing a luminescent reporter apoaequorin-GFP fusion protein (G5A) that has been shown to emit more light than apoaequorin alone (Baubet et al., 2000). Cells were seeded in 96-well plates (Thermo Fisher Scientific) at ~10,000 cells/well. After 1 day, cells were transfected with plasmids encoding a GPCR and the promiscuous Gα-16 protein (60 ng each) using 1.5 μl of the transfection reagent TurboFect (Thermo Fisher Scientific). To measure dose-response curves, cells were also cotransfected with the G5A construct to increase the expression of the reporter. After 2 days of expression, the medium was removed and replaced with Hank's balanced salt solution (HBSS) supplemented with 1.8 mM Ca^{2+}, 10 mM glucose, and 1 mM coelenterazine h (Promega). After incubation at 37°C for 2 hr, cells were tested by adding synthetic peptides (GenScript) in HBSS supplemented with 1.8 mM Ca^{2+} and 10 mM glucose. Luminescence was recorded for 45–60 s in a plate reader (BioTek Synergy Mx or Synergy H4, BioTek). Data during the screen were normalized using the response of *Platynereis* MIP receptor to 10 nM MIP-7 (Conzelmann et al., 2013b). The final dose-response curves were normalized using the response of the same well to 1 mM histamine that was recorded following the peptide treatment. To record the second set of dose-response curves, data were normalized using the response of the *Platynereis* MIP receptor to 10 nM MIP-7. The MIP control values were recorded from three separate wells on each plate.

Deorphanization Strategy

All *Platynereis* GPCRs were first tested with three peptide mixtures containing up to 48 synthetic peptides (Table S2) at 1 μM each. Measurements were done in triplicate. Those GPCRs that showed a response compared to the negative control (empty pcDNA3.1) were tested further. We tried to predict the specific ligand from the active mixture based on receptor clustering using CLANS2 (Frickey and Lupas, 2004) and tested individual peptides where deorphanized orthologs were identified. Alternatively, GPCRs were tested with submixtures of synthetic peptides arranged in three matrices, corresponding to the decomposition of mixtures 1–3 (Table S2). The combination of mixtures that elicited a response pointed to the active ligand. These measurements were done in single wells or in duplicate. Using this information, individual ligands were tested. After identification of a bona fide ligand, dose-response curves were recorded using concentrations between 0.01 nM and 200 μM. Data for dose-response curves were recorded in triplicate for each concentration. Dose-response curves were fitted with a four-parameter curve using Prism 6 (GraphPad).

Bioinformatics

For clustering, a previous collection of GPCRs (Jékely, 2013) was complemented with deorphanized *Platynereis* sequences and deorphanized GPCR sequences from human, mouse, and rat retrieved from the IUPHAR database (Pawson et al., 2014). Deorphanized GPCRs from *D. melanogaster* and other insects (Caers et al., 2012), *C. elegans* (Frooninckx et al., 2012), and other organisms (Bigot et al., 2014; Conzelmann et al., 2013b; Cox et al., 1997; Kim et al., 2010; Tensen et al., 1998a, 1998b) were also included. Furthermore, the sequences of all *Platynereis* GPCRs tested in the screen were used to initiate BLAST searches at NCBI with an e-value cutoff of 1e-50, and all hits were downloaded and added to the collection. Deorphanized sequences were tagged "deorphanized." All sequences were complemented with taxonomic information based on the NCBI taxonomy identifier (taxid) using a bioperl script (https://github.com/JekelyLab/GPCR_Clans_Maps), or taxonomy information was added manually. Redundant sequences were removed from the collection using CD-HIT (Li and Godzik, 2006). Clustering analysis was done using CLANS2 (Frickey and Lupas, 2004) with a BLOSUM62 matrix and a p value cutoff of 1.e-50. Deorphanized and orphan *Platynereis* receptors were clustered separately with their respective orthologs. Clusters that contained no *Platynereis* sequences were removed from the map (including relaxin, melanocortin, bradykinin, urotensins, and neurotensin receptors).

ACCESSION NUMBERS

The accession numbers of all *Platynereis* GPCRs tested here are GenBank: KP293941–KP294026 and KP420212–KP420214.

SUPPLEMENTAL INFORMATION

Supplemental Information includes six figures and four tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2015.06.052.

AUTHOR CONTRIBUTIONS

G.J. conceived the experiments; G.J. and P.B. designed methodology; P.B. performed all investigations and conducted the formal analysis; G.J. wrote the first draft of the manuscript; G.J. and P.B. reviewed and edited the manuscript; and G.J. supervised the study and acquired funding.

ACKNOWLEDGMENTS

We thank Sorin Tunaru and Stefan Offermanns for providing the stable CHO cell line and for training P.B. to perform GPCR activation assays. We also thank Elizabeth Williams for comments on the manuscript. The research leading to these results received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ European Research Council Grant Agreement 260821. P.B. is supported by the International Max Planck Research School (IMPRS) "From Molecules to Organisms."

Received: February 10, 2015
Revised: May 27, 2015
Accepted: June 12, 2015
Published: July 16, 2015

REFERENCES

Baubet, V., Le Mouellic, H., Campbell, A.K., Lucas-Meunier, E., Fossier, P., and Brûlet, P. (2000). Chimeric green fluorescent protein-aequorin as bioluminescent Ca2+ reporters at the single-cell level. Proc. Natl. Acad. Sci. USA *97*, 7260–7265.

Bigot, L., Beets, I., Dubos, M.-P., Boudry, P., Schoofs, L., and Favrel, P. (2014). Functional characterization of a short neuropeptide F-related receptor in a lophotrochozoan, the mollusk Crassostrea gigas. J. Exp. Biol. *217*, 2974–2982.

Caers, J., Verlinden, H., Zels, S., Vandersmissen, H.P., Vuerinckx, K., and Schoofs, L. (2012). More than two decades of research on insect neuropeptide GPCRs: an overview. Front. Endocrinol. (Lausanne) *3*, 151.

Cazzamali, G., and Grimmelikhuijzen, C.J.P. (2002). Molecular cloning and functional expression of the first insect FMRFamide receptor. Proc. Natl. Acad. Sci. USA *99*, 12073–12078.

Chang, J.C., Yang, R.B., Adams, M.E., and Lu, K.H. (2009). Receptor guanylyl cyclases in Inka cells targeted by eclosion hormone. Proc. Natl. Acad. Sci. USA *106*, 13371–13376.

Collins, J.J., 3rd, Hou, X., Romanova, E.V., Lambrus, B.G., Miller, C.M., Saberi, A., Sweedler, J.V., and Newmark, P.A. (2010). Genome-wide analyses reveal a role for peptide hormones in planarian germline development. PLoS Biol. *8*, e1000509.

Conzelmann, M., Offenburger, S.-L., Asadulina, A., Keller, T., Münch, T.A., and Jékely, G. (2011). Neuropeptides regulate swimming depth of *Platynereis* larvae. Proc. Natl. Acad. Sci. USA *108*, E1174–E1183.

Conzelmann, M., Williams, E.A., Krug, K., Franz-Wachtel, M., Macek, B., and Jékely, G. (2013a). The neuropeptide complement of the marine annelid *Platynereis dumerilii*. BMC Genomics *14*, 906.

Conzelmann, M., Williams, E.A., Tunaru, S., Randel, N., Shahidi, R., Asadulina, A., Berger, J., Offermanns, S., and Jékely, G. (2013b). Conserved MIP receptor-ligand pair regulates *Platynereis* larval settlement. Proc. Natl. Acad. Sci. USA *110*, 8224–8229.

Cox, K.J., Tensen, C.P., Van der Schors, R.C., Li, K.W., van Heerikhuizen, H., Vreugdenhil, E., Geraerts, W.P., and Burke, J.F. (1997). Cloning, characterization, and expression of a G-protein-coupled receptor from *Lymnaea stagnalis* and identification of a leucokinin-like peptide, PSFHSWSamide, as its endogenous ligand. J. Neurosci. *17*, 1197–1205.

Cropper, E.C., Tenenbaum, R., Kolks, M.A., Kupfermann, I., and Weiss, K.R. (1987). Myomodulin: a bioactive neuropeptide present in an identified cholinergic buccal motor neuron of *Aplysia*. Proc. Natl. Acad. Sci. USA *84*, 5483–5486.

Dircksen, H., Neupert, S., Predel, R., Verleyen, P., Huybrechts, J., Strauss, J., Hauser, F., Stafflinger, E., Schneider, M., Pauwels, K., et al. (2011). Genomics, transcriptomics, and peptidomics of *Daphnia pulex* neuropeptides and protein hormones. J. Proteome Res. *10*, 4478–4504.

Frickey, T., and Lupas, A. (2004). CLANS: a Java application for visualizing protein families based on pairwise similarity. Bioinformatics *20*, 3702–3704.

Frooninckx, L., Van Rompay, L., Temmerman, L., Van Sinay, E., Beets, I., Janssen, T., Husson, S.J., and Schoofs, L. (2012). Neuropeptide GPCRs in *C. elegans*. Front. Endocrinol. (Lausanne) *3*, 167.

Hauser, F., Neupert, S., Williamson, M., Predel, R., Tanaka, Y., and Grimmelikhuijzen, C.J.P. (2010). Genomics and peptidomics of neuropeptides and protein hormones present in the parasitic wasp *Nasonia vitripennis*. J. Proteome Res. *9*, 5296–5310.

Hewes, R.S., and Taghert, P.H. (2001). Neuropeptides and neuropeptide receptors in the *Drosophila melanogaster* genome. Genome Res. *11*, 1126–1142.

Heyland, A., Price, D.A., Bodnarova-Buganova, M., and Moroz, L.L. (2006). Thyroid hormone metabolism and peroxidase function in two non-chordate animals. J. Exp. Zoolog. B Mol. Dev. Evol. *306*, 551–566.

Hoek, R.M., Li, K.W., van Minnen, J., Lodder, J.C., de Jong-Brink, M., Smit, A.B., and van Kesteren, R.E. (2005). LFRFamides: a novel family of parasitation-induced -RFamide neuropeptides that inhibit the activity of neuroendocrine cells in *Lymnaea stagnalis*. J. Neurochem. *92*, 1073–1080.

Holman, G.M., Cook, B.J., and Nachman, R.J. (1986). Isolation, primary structure and synthesis of leucomyosuppressin, an insect neuropeptide that inhibits spontaneous contractions of the cockroach hindgut. Comp. Biochem. Physiol. C Comp. Pharmacol. *85*, 329–333.

Janssen, T., Lindemans, M., Meelkop, E., Temmerman, L., and Schoofs, L. (2010). Coevolution of neuropeptidergic signaling systems: from worm to man. Ann. N Y Acad. Sci. *1200*, 1–14.

Jékely, G. (2013). Global view of the evolution and diversity of metazoan neuropeptide signaling. Proc. Natl. Acad. Sci. USA *110*, 8702–8707.

Jékely, G., Colombelli, J., Hausen, H., Guy, K., Stelzer, E., Nédélec, F., and Arendt, D. (2008). Mechanism of phototaxis in marine zooplankton. Nature *456*, 395–399.

Kamatani, Y., Minakata, H., Kenny, P.T., Iwashita, T., Watanabe, K., Funase, K., Sun, X.P., Yongsiri, A., Kim, K.H., Novales-Li, P., et al. (1989). Achatin-I, an endogenous neuroexcitatory tetrapeptide from *Achatina fulica Férussac* containing a D-amino acid residue. Biochem. Biophys. Res. Commun. *160*, 1015–1020.

Kim, Y.-J., Bartalska, K., Audsley, N., Yamanaka, N., Yapici, N., Lee, J.-Y., Kim, Y.-C., Markovic, M., Isaac, E., Tanaka, Y., and Dickson, B.J. (2010). MIPs are ancestral ligands for the sex peptide receptor. Proc. Natl. Acad. Sci. USA *107*, 6520–6525.

Laudet, V. (2011). The origins and evolution of vertebrate metamorphosis. Curr. Biol. *21*, R726–R737.

Leung, D.W., Spencer, S.A., Cachianes, G., Hammonds, R.G., Collins, C., Henzel, W.J., Barnard, R., Waters, M.J., and Wood, W.I. (1987). Growth hormone receptor and serum binding protein: purification, cloning and expression. Nature *330*, 537–543.

Li, W., and Godzik, A. (2006). Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics *22*, 1658–1659.

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

Lingueglia, E., Champigny, G., Lazdunski, M., and Barbry, P. (1995). Cloning of the amiloride-sensitive FMRFamide peptide-gated sodium channel. Nature *378*, 730–733.

Lowe, D.G., Chang, M.S., Hellmiss, R., Chen, E., Singh, S., Garbers, D.L., and Goeddel, D.V. (1989). Human atrial natriuretic peptide receptor defines a new paradigm for second messenger signal transduction. EMBO J. *8*, 1377–1384.

Please cite this article in press as: Bauknecht and Jékely, Large-Scale Combinatorial Deorphanization of *Platynereis* Neuropeptide GPCRs, Cell Reports (2015), http://dx.doi.org/10.1016/j.celrep.2015.06.052

Mertens, I., Meeusen, T., Huybrechts, R., De Loof, A., and Schoofs, L. (2002). Characterization of the short neuropeptide F receptor from *Drosophila melanogaster*. Biochem. Biophys. Res. Commun. *297*, 1140–1148.

Mirabeau, O., and Joly, J.-S. (2013). Molecular evolution of peptidergic signaling systems in bilaterians. Proc. Natl. Acad. Sci. USA *110*, E2028–E2037.

Oumi, T., Ukena, K., Matsushima, O., Ikeda, T., Fujita, T., Minakata, H., and Nomoto, K. (1995). The GGNG peptides: novel myoactive peptides isolated from the gut and the whole body of the earthworms. Biochem. Biophys. Res. Commun. *216*, 1072–1078.

Park, Y., Kim, Y.-J., and Adams, M.E. (2002). Identification of G protein-coupled receptors for Drosophila PRXamide peptides, CCAP, corazonin, and AKH supports a theory of ligand-receptor coevolution. Proc. Natl. Acad. Sci. USA *99*, 11423–11428.

Pawson, A.J., Sharman, J.L., Benson, H.E., Faccenda, E., Alexander, S.P.H., Buneman, O.P., Davenport, A.P., McGrath, J.C., Peters, J.A., Southan, C., et al.; NC-IUPHAR (2014). The IUPHAR/BPS Guide to PHARMACOLOGY: an expert-driven knowledgebase of drug targets and their ligands. Nucleic Acids Res. *42*, D1098–D1106.

Rajpara, S.M., Garcia, P.D., Roberts, R., Eliassen, J.C., Owens, D.F., Maltby, D., Myers, R.M., and Mayeri, E. (1992). Identification and molecular cloning of a neuropeptide Y homolog that produces prolonged inhibition in *Aplysia* neurons. Neuron *9*, 505–513.

Randel, N., Asadulina, A., Bezares-Calderón, L.A., Verasztó, C., Williams, E.A., Conzelmann, M., Shahidi, R., and Jékely, G. (2014). Neuronal connectome of a sensory-motor circuit for visual navigation. eLife *3*, e02730.

Rechler, M.M., and Nissley, S.P. (1985). The nature and regulation of the receptors for insulin-like growth factors. Annu. Rev. Physiol. *47*, 425–442.

Roller, L., Yamanaka, N., Watanabe, K., Daubnerová, I., Zitňan, D., Kataoka, H., and Tanaka, Y. (2008). The unique evolution of neuropeptide genes in the silkworm *Bombyx mori*. Insect Biochem. Mol. Biol. *38*, 1147–1157.

Rowe, M.L., and Elphick, M.R. (2012). The neuropeptide transcriptome of a model echinoderm, the sea urchin *Strongylocentrotus purpuratus*. Gen. Comp. Endocrinol. *179*, 331–344.

Telford, M.J., and Copley, R.R. (2011). Improving animal phylogenies with genomic data. Trends Genet. *27*, 186–195.

Tensen, C.P., Cox, K.J., Burke, J.F., Leurs, R., van der Schors, R.C., Geraerts, W.P., Vreugdenhil, E., and Heerikhuizen, H. (1998a). Molecular cloning and characterization of an invertebrate homologue of a neuropeptide Y receptor. Eur. J. Neurosci. *10*, 3409–3416.

Tensen, C.P., Cox, K.J., Smit, A.B., van der Schors, R.C., Meyerhof, W., Richter, D., Planta, R.J., Hermann, P.M., van Minnen, J., Geraerts, W.P., et al. (1998b). The lymnaea cardioexcitatory peptide (LyCEP) receptor: a G-protein-coupled receptor for a novel member of the RFamide neuropeptide family. J. Neurosci. *18*, 9812–9821.

Tosches, M.A., Bucher, D., Vopalensky, P., and Arendt, D. (2014). Melatonin signaling controls circadian swimming behavior in marine zooplankton. Cell *159*, 46–57.

Tunaru, S., Lättig, J., Kero, J., Krause, G., and Offermanns, S. (2005). Characterization of determinants of ligand binding to the nicotinic acid receptor GPR109A (HM74A/PUMA-G). Mol. Pharmacol. *68*, 1271–1280.

Veenstra, J.A. (2010). Neurohormones and neuropeptides encoded by the genome of *Lottia gigantea*, with reference to other mollusks and insects. Gen. Comp. Endocrinol. *167*, 86–103.

Veenstra, J.A. (2011). Neuropeptide evolution: neurohormones and neuropeptides predicted from the genomes of *Capitella teleta* and *Helobdella robusta*. Gen. Comp. Endocrinol. *171*, 160–175.

Venkatesh, B., Lee, A.P., Ravi, V., Maurya, A.K., Lian, M.M., Swann, J.B., Ohta, Y., Flajnik, M.F., Sutoh, Y., Kasahara, M., et al. (2014). Elephant shark genome provides unique insights into gnathostome evolution. Nature *505*, 174–179.

Williams, E.A., Conzelmann, M., and Jékely, G. (2015). Myoinhibitory peptide regulates feeding in the marine annelid Platynereis. Front. Zool. *12*, 1.

Willows, A.O., Pavlova, G.A., and Phillips, N.E. (1997). Modulation of ciliary beat frequency by neuropeptides from identified molluscan neurons. J. Exp. Biol. *200*, 1433–1439.

Xie, F., London, S.E., Southey, B.R., Annangudi, S.P., Amare, A., Rodriguez-Zas, S.L., Clayton, D.F., and Sweedler, J.V. (2010). The zebra finch neuropeptidome: prediction, detection and expression. BMC Biol. *8*, 28.

Zantke, J., Bannister, S., Rajan, V.B.V., Raible, F., and Tessmar-Raible, K. (2014). Genetic and genomic tools for the marine annelid *Platynereis dumerilii*. Genetics *197*, 19–31.

Cell Reports

Supplemental Information

Large-Scale Combinatorial Deorphanization of *Platynereis* Neuropeptide GPCRs

Philipp Bauknecht and Gáspár Jékely

SUPPLEMENTAL INFORMATION
Supplemental Figures

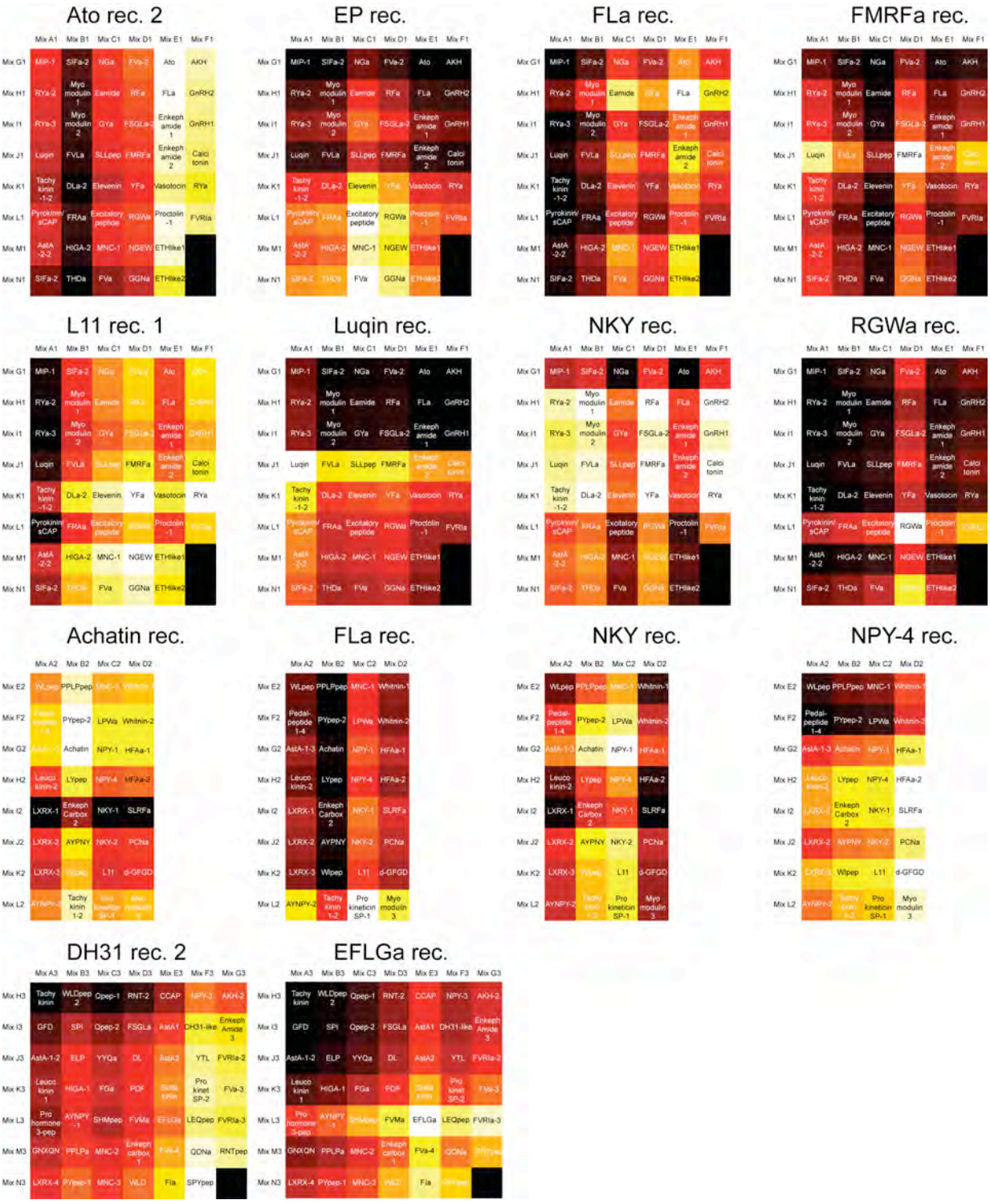


Figure S1. Combinatorial screening of responder GPCRs with peptide submixtures, Related to Figure 1.

The composition of the submixtures is described in Table S2. The activation value for each peptide in the matrices is defined as the square root of the product of the measured values of two intersecting mixtures.

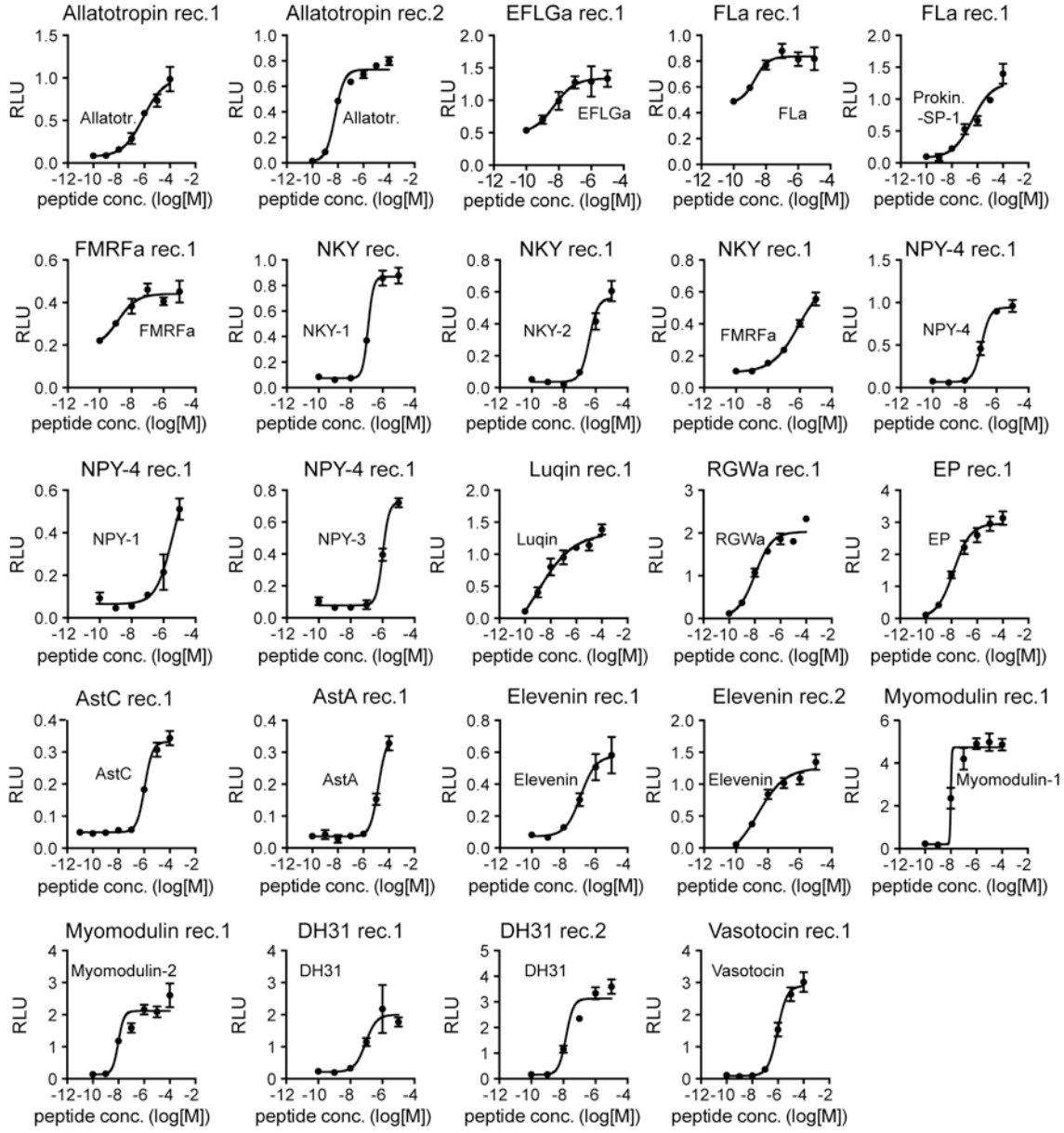


Figure S2. Dose-response curves of *Platynereis* deorphanized GPCRs treated with varying concentrations of peptides, Related to Figure 3.

Data represent luminescence units relative to the response of the *Platynereis* MIP receptor to 10 nM MIP. Data are shown as mean ± SEM (n = 3). Dose-response curves fitted to the data are shown. Ligand names are shown beside the curves. EC_{50} values are listed in Table 1. Abbreviations: RLU, relative luminescence unit; AstC, allatostatin-C; AstA, allatostatin-A; EP, excitatory peptide.

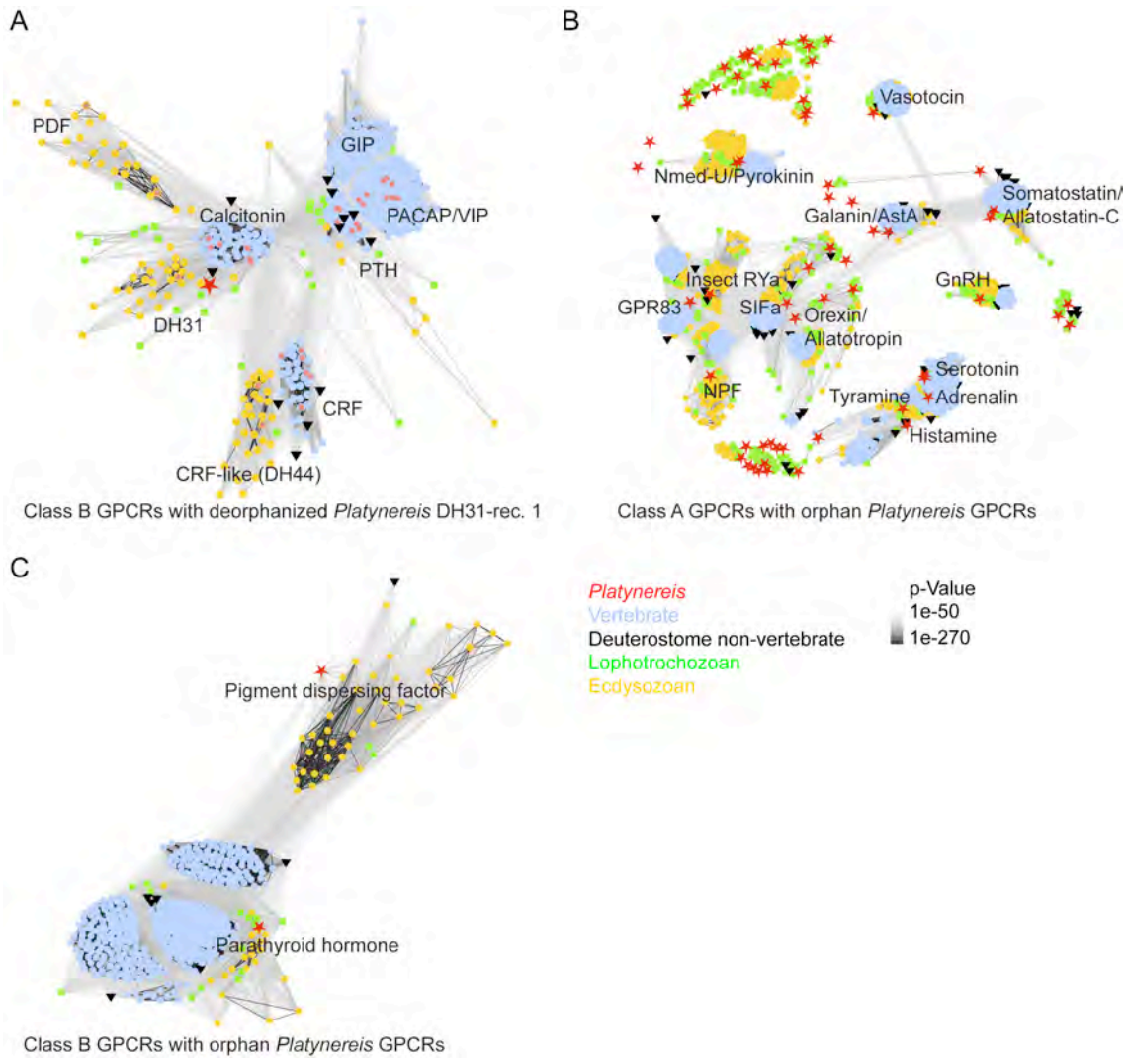


Figure S3. Sequence-similarity-based clustering of neuropeptide GPCRs, Related to Figure 4.

Nodes represent sequences, edges represent BLASTP connections. Edges are colored according to BLASTP p-values. Nodes are colored based on taxonomy. (A) Class-B neuropeptide GPCRs including the deorphanized *Platynereis* DH31 receptor, indicated with a red star. Deorphanized receptors from other species are marked with a red dot. Clusters are named according to the name of the deorphanized family members. Abbreviations: DH31, diuretic hormone 31; GIP, gastric inhibitory polypeptide; PACAP, pituitary adenylate cyclase-activating peptide; PDF, pigment dispersing factor; PTH, parathyroid hormone, VIP, vasoactive intestinal polypeptide. (B) Class-A neuropeptide GPCRs including non-deorphanized *Platynereis* GPCRs. Orphan *Platynereis* GPCRs were used to initiate BLAST searches in Uniprot. Only clusters that contain orphan *Platynereis* GPCRs and deorphanized orthologs are labeled. (C) Class-B neuropeptide GPCRs including orphan *Platynereis* GPCRs. Only clusters that contain *Platynereis* GPCRs and deorphanized orthologs are labeled.

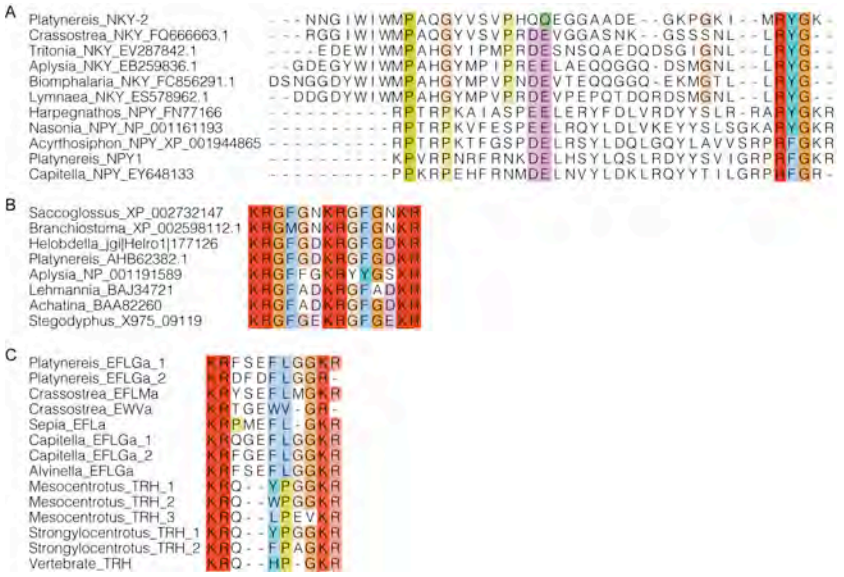


Figure S4. Multiple sequence alignment of NKY/NPY, achatin and ELFGamide/TRH neuropeptides, Related to Figure 4.

(A) Multiple sequence alignment of protostome NKY and NPY neuropeptides. The NKY peptides are related to NPY peptides, as also supported by the orthology of their receptors. (B) Multiple sequence alignment of protostome and deuterostome achatin peptides. Two consecutive peptide stretches with the dibasic cleavage sites are shown. The Phe residue was shown to be in the D-form in the giant snail *Achatina fulica* (Kamatani et al., 1989). The corresponding residue was also tested in both the L- and the D-form for the *Platynereis*, *Aplysia*, *Saccoglossus* and *Branchiostoma* peptides. (C) Multiple sequence alignment of protostome ELFGamide and deuterostome TRH peptides. The orthology of TRH and EFLGa receptors and the limited sequence similarity of the peptides suggest that TRH and EFLGa peptides are orthologous.

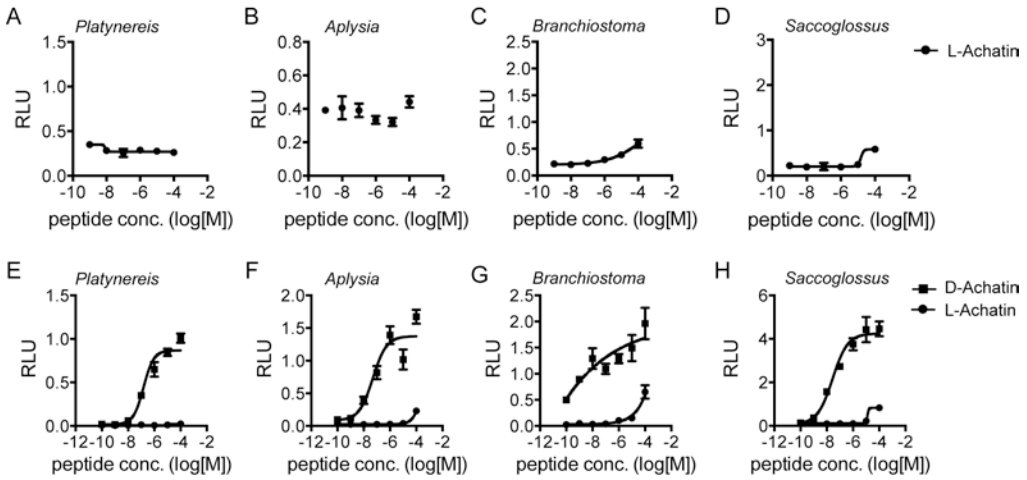


Figure S5. Dose-response curves of bilaterian achatin receptors treated with varying concentrations of achatin peptides, Related to Figure 5.

Each receptor was tested with the species-specific peptide containing a L-Phe or D-Phe. Data representing luminescence units relative to the control response are shown as mean ± SEM (n = 3). Data were normalized either to histamine receptor (A-D) or to MIP receptor responses (E-H). Dose-response curves fitted to the data are shown.

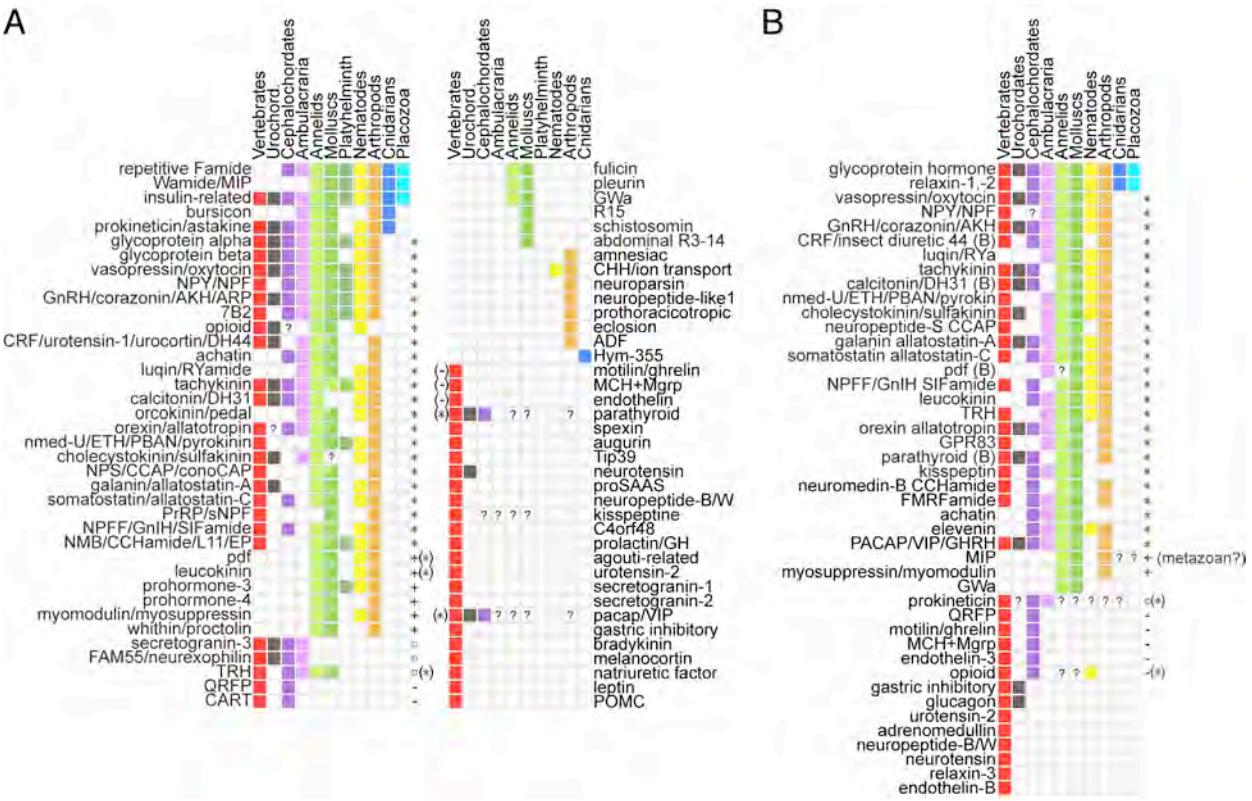


Figure S6. Phyletic distribution of metazoan pNP and neuropeptide GPCR families, Related to Figure 4.

Table updated from (Jékely, 2013). (A) Phyletic distribution of metazoan pNP families. (B) Phyletic distribution of metazoan class-A and class-B neuropeptide GPCR families. Class-B GPCRs are indicated as (B). Ancestral bilaterian (*), protostome (+), deuterostome (o) and chordate (-) families are indicated. Question marks indicate receptors or ligands that are expected to be in the indicated taxonomic group but have not yet been described. Besides the updates described in the main text, the following changes were made: achatin was found in a chelicerate, *Stegodyphus mimosarum* (GenBank KFM77812), Wamides and insulins were described in *Trichoplax* (Nikitin, 2014). Pyrokinin is present in Platyhelminthes (Collins et al., 2010). sNPF was found in mollusks (Bigot et al., 2014). DH44 has been identified in annelids and mollusks and shown to be related to mollusk egg-laying hormone (Conzelmann et al., 2013; Mirabeau and Joly, 2013). Sulfakinin was identified in annelids (Conzelmann et al., 2013). NPY, CRF/urotensin-1/urocortin/DH44, and cholecystokinin/sulfakinin orthologs were identified in Ambulacraria (Mirabeau and Joly, 2013). Parathyroid hormone and PACAP orthologs were identified in *Ciona* and *Branchiostoma* (Mirabeau and Joly, 2013). Putative neurotensin and opioid peptides were described in *Ciona* (Mirabeau and Joly, 2013). A somatostatin ortholog was identified in *Branchiostoma* (Mirabeau and Joly, 2013). Tachykinin (T07C12.15), luqin (Y75B8A.11), DH31 (ZK287.3), and SIFamide orthologs (nlp-10) were identified in *C. elegans* (Mirabeau and Joly, 2013). An opioid peptide and its receptor were described in *C. elegans* (Cheong et al., 2015). The *C. elegans* FLP-18 RFamide peptides represent the nematode orthologs of NPY/NPF since these RFamide peptides activate NPY-receptor orthologs (npr-4 and npr-5) (Cohen et al., 2009).

Supplemental Tables

Table S1. List of 87 *Platynereis* GPCRs tested in the primary screen, Related to Figure 1.
The names, GenBank identifiers and identified ligands are listed for the 87 *Platynereis* GPCRs tested. The predicted ligands for orphan *Platynereis* GPCRs are indicated.

Table S2. Name and sequence of *Platynereis* neuropeptides in Mixtures 1-3 and Submixtures 1-3, Related to Figure 3.
Mixtures 1 to 3 contain 46, 32 and 48 peptides, respectively. pGlu indicates N-terminal pyroglutamylation. -NH2 indicates C-terminal amidation. Cys residues that form disulphide bonds are underlined.

Table S3. GPCR activation data from the screen, Related to Figure 3.
Data are shown from the screen with peptide mixtures, from the combinatorial screen with peptide submixtures, the targeted screen based on ligand-receptor predictions, and the measurements without Gα16. For the combinatorial screen, raw data are shown with peptide submixtures. For single peptide-GPCR measurements the data are also shown for measurements with complex mixtures lacking the active peptide. In order to identify GPCR signaling by Gq-α, the measurements with the identified ligands were repeated with and without the promiscuous Gα16.

Table S4. Ligand predictions for GPCRs based on phylogenetic information, Related to Figure 4.
The file contains GenBank or Uniprot sequence identifiers for metazoan GPCR sequences that are orthologous to the identified neuropeptide GPCR families. The predicted ligands for the *Branchiostoma* FMRFamide receptor orthologs are also shown.

Supplemental References

Cheong, M.C., Artyukhin, A.B., You, Y.-J., and Avery, L. (2015). An opioid-like system regulating feeding behavior in *C. elegans*. Elife *4*.

Cohen, M., Reale, V., Olofsson, B., Knights, A., Evans, P., and de Bono, M. (2009). Coordinated regulation of foraging and metabolism in *C. elegans* by RFamide neuropeptide signaling. Cell Metab. *9*, 375–385.

Nikitin, M. (2014). Bioinformatic prediction of *Trichoplax adhaerens* regulatory peptides. Gen Comp Endocrinol.