

METHODS AND RESOURCES

An atlas of *Caenorhabditis elegans* chemoreceptor expression

Berta Vidal*, Ulkar Aghayeva, Haosheng Sun, Chen Wang, Lori Glenwinkel, Emily A. Bayer, Oliver Hobert*

Department of Biological Sciences, Howard Hughes Medical Institute, Columbia University, New York, New York, United States of America

* bvi2001@columbia.edu (BV); or38@columbia.edu (OH)



Abstract

One goal of modern day neuroscience is the establishment of molecular maps that assign unique features to individual neuron types. Such maps provide important starting points for neuron classification, for functional analysis, and for developmental studies aimed at defining the molecular mechanisms of neuron identity acquisition and neuron identity diversification. In this resource paper, we describe a nervous system-wide map of the potential expression sites of 244 members of the largest gene family in the *C. elegans* genome, rhodopsin-like (class A) G-protein-coupled receptor (GPCR) chemoreceptors, using classic *gfp* reporter gene technology. We cover representatives of all sequence families of chemoreceptor GPCRs, some of which were previously entirely uncharacterized. Most reporters are expressed in a very restricted number of cells, often just in single cells. We assign GPCR reporter expression to all but two of the 37 sensory neuron classes of the sex-shared, core nervous system. Some sensory neurons express a very small number of receptors, while others, particularly nociceptive neurons, coexpress several dozen GPCR reporter genes. GPCR reporters are also expressed in a wide range of inter- and motoneurons, as well as non-neuronal cells, suggesting that GPCRs may constitute receptors not just for environmental signals, but also for internal cues. We observe only one notable, frequent association of coexpression patterns, namely in one nociceptive amphid (ASH) and two nociceptive phasmid sensory neurons (PHA, PHB). We identified GPCRs with sexually dimorphic expression and several GPCR reporters that are expressed in a left/right asymmetric manner. We identified a substantial degree of GPCR expression plasticity; particularly in the context of the environmentally-induced dauer diapause stage when one third of all tested GPCRs alter the cellular specificity of their expression within and outside the nervous system. Intriguingly, in a number of cases, the dauer-specific alterations of GPCR reporter expression in specific neuron classes are maintained during postdauer life and in some case new patterns are induced post-dauer, demonstrating that GPCR gene expression may serve as traits of life history. Taken together, our resource provides an entry point for functional studies and also offers a host of molecular markers for studying molecular patterning and plasticity of the nervous system.

OPEN ACCESS

Citation: Vidal B, Aghayeva U, Sun H, Wang C, Glenwinkel L, Bayer EA, et al. (2018) An atlas of *Caenorhabditis elegans* chemoreceptor expression. PLoS Biol 16(1): e2004218. <https://doi.org/10.1371/journal.pbio.2004218>

Academic Editor: Leslie Vosshall, Rockefeller University, United States of America

Received: September 12, 2017

Accepted: November 22, 2017

Published: January 2, 2018

Copyright: © 2018 Vidal et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All data is contained within the paper. Transgenic strains are available via the *C.elegans* strain repository at the *Caenorhabditis* Genetics Center (CGC).

Funding: NIH (grant number 1F31NS096863) to EAB. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. NIH (grant number R21NS096343, R37NS039996) to OH. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. HHMI to OH. The funder had no role in

study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Abbreviations: Amso, amphid socket cells; AU, approximately unbiased *p*-value; bHLH, basic helix-loop-helix; BP, bootstrap probability; csGPCR, chemosensory-type GPCR; GFP, green fluorescent protein; GPCR, G-protein-coupled receptor; HSN, hermaphrodite-specific neuron; L1, first larval stage; L3, third larval stage; L4, fourth larval stage; NaN₃, sodium azide; PHsh, phasmid sheath; PHso, phasmid socket; sr, serpentine receptor; UV1, uterine/vulval cell 1.

Author summary

Maps of gene expression patterns in nervous systems provide an important resource for neuron classification, for functional analysis, and for developmental studies that ask how different neurons acquire their unique identities. By analyzing transgenic GFP reporter strains, we describe here the expression pattern of 244 putative chemosensory receptor-encoding genes, which constitute the largest gene family in *Caenorhabditis elegans*. As expected, chemoreceptor expression is enriched in chemosensory neurons. Putative chemoreceptors are also expressed in a wide range of interneurons and motor neurons, as well as non-neuronal cells, suggesting that these receptors may sense internal cues in addition to environmental signals. Each chemoreceptor is expressed sparsely, often in just one neuron type, but each neuron type can express many chemoreceptors. Chemoreceptor expression is remarkably plastic, particularly in the context of the environmentally induced dauer diapause stage. Taken together, this molecular atlas of chemosensory receptors provides an entry point for functional studies and offers a host of markers for studying neuronal patterning and plasticity.

Introduction

Molecular markers selectively expressed in individual neuron types represent invaluable tools to understand how cellular diversity in a nervous system is genetically encoded. Molecular markers that are constitutively and invariably expressed throughout the life of a specific neuron type provide static views of neuronal identity and hence provide entry points to study how invariable identity features are acquired during neuronal differentiation [1]. In contrast, some molecular features of a neuron display a remarkable plasticity in that their expression may be regulated by neuronal activity or in response to specific environmental cues. Such genes serve as markers to understand the nature of the gene regulatory programs that govern such dynamic features of a neuron. We reasoned that a significant expansion of the expression analysis of chemosensory G-protein-coupled receptors (GPCRs), initiated more than 20 years ago [2] using *gfp*-based reporter gene technology [3], may yield a significantly expanded resource of molecular markers that may label various aspects of neuronal identity and neuronal plasticity in the *C. elegans* nervous system.

Animal genomes encode five major classes of GPCRs, of which the rhodopsin class (or “class A”) is the largest class [4,5] (Table 1). Rhodopsin class GPCRs can be subdivided into phylogenetically deeply conserved neurotransmitter receptors (neuropeptides, acetylcholine, biogenic amines) as well as non-conserved, chemosensory-type GPCRs (csGPCRs) (Table 1). The csGPCRs have independently expanded in distinct animal phyla where they serve to respond to diverse, physiologically relevant external and, supposedly, internal cues [4,6,7]. The genome of the nematode *C. elegans* encodes an exceptionally large battery of csGPCRs composed of 1,341 protein-coding genes (Table 2) [2,7,8], a remarkable number given the small size of its nervous system (302 neurons constituting 118 anatomically defined neuron classes) [9]. These csGPCRs have been subdivided by sequence into families and super-families, as summarized in Table 2 [2,7].

Wormbase contains expression data for 131 csGPCRs, however for only 76 of them the expression site has been defined with single cell resolution (S1 Table). The majority of these 76 reporters revealed expression in chemosensory neurons [2]. Functional studies have linked a small subset of these receptors to the sensation of specific environmental or pheromonal cues

Table 1. The five classes of GPCRs in animal genomes and their representation in *C. elegans*. Modified from [10].

| Class ¹ | Subclass ¹ | Gene number in <i>C. elegans</i> |
|------------------------------|----------------------------|----------------------------------|
| Rhodopsin (Class A) | chemosensory | 1,341 ^a |
| | peptidergic | 153 ^b |
| | aminergic | 16 |
| | muscarinic (acetylcholine) | 3 |
| Secretin (Class B) | | 3 |
| Glutamate receptor (Class C) | | 7 |
| Adhesion | | 5 |
| Frizzled/Tas2 | | 4 |

Abbreviations: GPCR, G-protein-coupled receptor

¹ Classification after [5].

^a Will likely also contain peptide receptors (see text).

^b Defined by sequence homology to known neuropeptide receptors [10].

<https://doi.org/10.1371/journal.pbio.2004218.t001>

[12–21], but in the absence of concerted de-orphanization efforts like those seen in other organisms [22,23], the number of receptors with assigned ligands is still remarkably low.

Intriguingly, a subset of the previously characterized csGPCR genes were also expressed in non-sensory neurons [2,24–28], suggesting that these csGPCRs may also function as receptors of internal ligands of unknown identity. Providing some hints to the identity of these ligands, one csGPCR subclass, encoded by the *srw* genes, displays sequence similarities to peptide receptors [11,29]. The expression of csGPCRs in interneurons also prompted efforts to identify the function of some of these genes. Even though its ligand remains unknown, AIY-expressed *sra-11* was found to be involved in the associative learning paradigm, olfactory imprinting [30], while *sra-13* acts in the vulva to control vulval development, which is affected by food signals [26].

In spite of the relative paucity of known ligands, the previously published expression patterns of csGPCRs provided molecular indicators for a number of intriguing and generally very poorly understood nervous system features: (1) the expression pattern of the GPCR gene *str-2* revealed a left/right asymmetry in the two AWC olfactory neurons [31]; this lateralization phenomenon was later found to be required for olfactory discrimination [32] and spurred a host of studies aimed at revealing how this left/right asymmetry is developmentally programmed [33]. (2) The expression of several csGPCRs revealed a remarkable plasticity in response to changes in the environment. For example, expression of *srd-1*, *str-2* and *str-3* changes in ASI neurons in response to dauer pheromones [34], and expression of *srh-34* and *srh-234* in ADL is different in fed versus starved animals [35]. Using these dynamic reporter gene patterns, mechanisms controlling csGPCR plasticity have been elucidated [35,36]. (3) The csGPCR genes *srd-1*, *srj-54*, and *odr-10* have been found to be expressed in a sexually dimorphic manner in sex-shared sensory neurons, suggesting that sexual identity impinges on sensory perception [2,37,38].

In this resource paper, we examined the expression of 244 reporter transgenes that monitor expression of previously uncharacterized csGPCR genes. Our explicit goal in this analysis was to (1) generate more neuronal identity markers, (2) test the hypothesis that many more sensory neurons may be lateralized, (3) identify more markers of neuronal plasticity, (4) identify more markers of sexual dimorphism, and (5) examine the extent of expression in non-sensory and non-neuronal cells (suggesting roles as receivers of internal signals). Based on the molecular classification of csGPCRs into defined families, we were also interested in determining whether the expression of specific subfamilies—particularly those whose expression has not

Table 2. Overview of GPCR reporters and expression.

| Classification ^a | | Gene counts | | Reporters | Overview of expression | | |
|-----------------------------|-------------|------------------------|------------------------|---|------------------------|----------------------|-----------------|
| Super-family | Family | Old count ^a | New count ^b | Total # examined reporters ^c | Neurons only | Neurons + non-neuron | Non-neuron only |
| Str | <i>srh</i> | 218 | 223 | 43 (14) | 24 | 16 | 3 |
| | <i>str</i> | 197 * | 196 * | 42 (16) | 21 | 16 | 5 |
| | <i>sri</i> | 61 | 60 | 21 (7) | 11 | 8 | 2 |
| | <i>srd</i> | 64 | 67 | 13 (6) | 10 | 2 | 1 |
| | <i>srj</i> | 39 | 39 | 14 (1) | 7 | 6 | 1 |
| | <i>srm</i> | 5 | 6 | 6 (-) | 3 | 3 | - |
| | <i>sm</i> | 1 | 1 | 1 (-) | 1 | - | - |
| | all Str | 585 | 591 | 140 (44) | 77 | 51 | 12 |
| Sra | <i>sre</i> | 51 | 53 | 31 (20) | 13 | 13 | 5 |
| | <i>sra</i> | 32 | 35 | 22 (11) | 15 | 6 | 1 |
| | <i>srab</i> | 22 | 23 | 18 (6) | 10 | 7 | - |
| | <i>srb</i> | 14 | 16 | 10 (4) | 4 | 4 | 2 |
| | all Sra | 119 | 127 | 81 (41) | 42 | 30 | 8 |
| Srg | <i>srx</i> | 98 | 105 | 20 (6) | 12 | 7 | 1 |
| | <i>srt</i> | 61 | 67 | 16 (6) | 13 | 2 | 1 |
| | <i>srg</i> | 59 | 61 | 23 (9) | 15 | 7 | 1 |
| | <i>srw</i> | 41 | 40 | 12 (5) | 6 | 6 | - |
| | <i>srv</i> | 30 | 30 | 12 (1) | 10 | 2 | - |
| | <i>srx</i> | 17 | 17 | 8 (4) | 6 | 1 | 1 |
| | all Srg | 306 | 320 | 91 (31) | 62 | 25 | 4 |
| Solo | <i>srw</i> | 99 | 115 | 11 (7) | 8 | 1 | 2 |
| Solo | <i>srz</i> | 71 | 68 | 23 (1) | 15 | 5 | 3 |
| Solo | <i>srbc</i> | 73 | 72 | 5 (2) | 4 | 1 | - |
| Solo | <i>srsx</i> | 37 | 37 | 14 (4) | 11 | 2 | 1 |
| Solo | <i>srr</i> | 10 | 9 | 9 (-) | 4 | 5 | 1 |
| Solo | <i>sro</i> | 1 | 1 | 1 (1) | 1 | - | - |
| Totals: | | 1,277 | 1,341 | 375 (131) | 224 | 120 | 31 |

Abbreviations: GPCR, G-protein-coupled receptor, sr, serpentine receptor. “Sr” is then followed by alphabetic letter codes for each.

Only sensory-type GPCRs are shown, other GPCR systems (hormone, neurotransmitter systems) are not. See text.

Numbers in parenthesis indicate previously described reporters extracted from Wormbase.

^a Based on Thomas and Robertson [7,11], with the exception of *sro-1* which was published elsewhere [2]. Pseudogenes are excluded.

^b New counts extracted from WS246 (some previous pseudogenes have become real genes and vice versa).

^c Summarized in S1 Table

* Includes *odr-10*.

<https://doi.org/10.1371/journal.pbio.2004218.t002>

previously been examined—may reveal specific common themes (i.e., patterns of coexpression or expression in specific cells) that may provide a hint to their function. We synthesize our findings with those of previous expression pattern analyses to carve out a number of general features of csGPCR expression patterns.

Materials and methods

Mutant strains

Strains were maintained by standard methods [39]. Mutant alleles used in this study were: *pha-1(e2123)* [40], *him-5(e1490)* [41], *unc-43(n1186lf)* [42], *unc-43(n498gf)* [43], and *nsy-5(ky634)* [44].

Reporter and transgenic strain generation

GFP reporters were generated using a PCR fusion approach [45] and injected without being subcloned. Genomic fragments were fused to the GFP coding sequence, which was followed by the *unc-54* 3' untranslated region. A list of primers for all constructs can be found in the Supplementary Methods. Amplicons were injected at 50 ng/μl with the *pha-1* rescuing plasmid (pBX) as a co-injection marker (50 ng/μl). Reporters were injected into a *pha-1(e2123)* or *pha-1(e2123);him-5(e1490)* mutant background strain [40], resulting in transgenic arrays with little mosaicism. For each construct, two independent lines were scored. Reporter strains provided by the Vancouver Consortium were generated as described [46]. Further details and primer sequences used by the Vancouver Consortium can be found at <http://www.gfpworm.org>. A list of all reporter strains generated by us or provided by the Vancouver Consortium can be found in the Supplementary Methods.

Microscopy

Worms were anesthetized using 100 mM sodium azide (NaN₃) and mounted on 5% agarose on glass slides. Images were acquired using an automated fluorescence microscope (Zeiss, AXIO Imager Z.2). Acquisition of several z-stack images (each approximately 1 μm thick) was performed with the ZEN 2 pro software. Representative images are shown following max-projection of Z-stacks using the maximum intensity projection type. Image reconstruction was performed using Fiji software [47].

Neuron identification

Neurons were identified either by labeling subsets of sensory neurons with DiD (Thermo Fisher Scientific) or by crossing reporter transgenes with landmark reporter strains in which known neuron classes are labeled with a red fluorescent reporter. For dye filling, worms were washed with M9, incubated with DiD (1:500) in M9 for 1 hour at room temperature, washed 3 times with M9, and plated on agar plates coated with food for 1–3 hours before imaging. Red fluorescent reporter strains used for cell identification are: *otIs263[ceh-36p::TagRFP, rol-6(su1006)]*, *vyIs51[*str-2p::2xnl::TagRFP; ofm-1p::DsRed*] [48]*, *otIs518[*eat-4^{Fosmid}::sl2::mCherry::h2b*] [49]*, *otIs544[*cho-1^{Fosmid}::sl2::mCherry::h2b*] [50]*, *otIs564[*unc-47^{Fosmid}::sl2::mCherry::h2b*] [51]*, *otIs612[*flp-18p::NLG-1::GFP11, gpa-6p::NLG-1::GFP1-10, flp-18p::mCherry, nlp-1p::mCherry*]*, *hdIs30[*glr-1p::DsRed*]*, *otIs521[*eat-4prom8::tagRFP; ttx-3::gfp*]*.

Hierarchical clustering of neurons by GPCR reporter expression

Clustering was performed on binary expression data from 272 neuron-expressed GPCR reporters for which we had cell ID information. Expression data was from our own analysis and available data from wormbase.org [52]. Only positive neuronal cell ID information per GPCR reporter was included in the binary expression matrix with no distinction between the absence of expression and unknown expression per neuron. Data were clustered using the R *pvclust* package (<https://cran.r-project.org/web/packages/pvclust/pvclust.pdf>) [53] using the euclidean distance metric with average linkage, bootstrap 1000, and relative sample size ranging from a proportion of 0.5 to 1.4 of the original sample size. The relative proportion was incremented by 0.1 for each bootstrap resampling. Bootstrap probability (BP) value and approximately unbiased *p*-values (AUs) are derived from the multiscale-multistep bootstrap resampling. AU support values >95 indicate well-supported clusters and should be considered when evaluating dendrogram cluster relationships. Alternative distance and linkage methods

showed clustering of the PHA, PHB, and ASH neurons in all cases (42 out of 84 cases had strong support with AU/BP >95).

Upstream intergenic distances and intron length calculations

GPCR upstream intergenic regions and intron lengths were extracted from *C. elegans* exon coordinates, version WS220 using a python script. Non-coding RNA exons were excluded from the intergenic distance calculations so that intergenic distances represent the nucleotide sequence distance between coding genes. The average intron length per gene was calculated by summing the intron sequence lengths for each gene and dividing by the total number of introns. Average intron lengths for genes with multiple isoforms were calculated for each isoform and then averaged, resulting in 1 average intron length per gene.

Generation of dauers and analysis of changes in expression

To analyze GPCR reporter gene expression in dauers, mixed populations of respective strains were allowed to exhaust food for 5–7 days at 20°C. Dauers were isolated from starved plates by treatment with 1% SDS for 30 minutes and imaged within 1–2 hours of isolation. The cellular identity of expression changes in dauers was confirmed with red landmark strains, as mentioned above.

Results

Selection of csGPCRs for expression analysis and method of analysis

We chose to examine csGPCR expression patterns using *gfp*-based reporter gene technology, the standard tool of gene expression analysis in *C. elegans* [3,54]. The obvious shortcoming of this technology is that reporter genes may not capture the full *cis*-regulatory content of the respective GPCR-encoding locus, but as we will describe in more detail below, most GPCR-encoding loci are compact with small intergenic regions and introns. We emphasize that our approach is not necessarily aimed at identifying the complete set of cells expressing a GPCR, but, following ample precedent, is rather aimed at identifying novel and informative patterns of expression, as incomplete as these patterns may be.

We utilized two sources of csGPCR reporters. A consortium at the University of British Columbia (Vancouver) has generated a valuable, large panel of reporters for 1886 genes in the *C. elegans* genome [46]. However, the site of expression of these reporters has not been determined with single cell resolution in the nervous system. We obtained 100 reporters from this collection that targeted GPCR loci, and for every reporter that produced a stable pattern of expression we undertook a detailed analysis of their sites of expression in the nervous system.

In addition to these 100 reporter genes, we generated 144 of our own reporter genes. We adhered to the following principles in the choice of genes and design of reporters: first, we aimed to cover all 23 classes of chemoreceptor genes defined by Thomas and Robertson [7] (Table 2). Using phylogenetic trees assembled by Thomas and Robertson, we sampled each gene family evenly, generally avoiding the examination of close sequence paralogues, which we anticipated to reveal similar expression patterns.

Our own reporters mostly contain all 5' intergenic regions fused to *gfp* and contain, at most, 4 kb of sequence. The rationale behind this choice lies in the overall organization of GPCR loci (summarized in S1 Fig). Eighty-nine percent of the approximately 1,300 csGPCR loci contain 5' intergenic regions of less than 4 kb. We chose all of our samples from this pool, and, hence, all the reporters generated by us capture the full intergenic region. The reporters from the Vancouver consortium contain about 3 kb of 5' intergenic region, at most [46]. Furthermore,

csGPCR loci tend to have small introns (average size 432 base pairs; almost half of them <200 base pairs; [S1 Fig](#)), indicating that relatively little *cis*-regulatory information resides in these introns, which provided the basis for our focus on intergenic regions. For some genes with very short upstream intergenic regions (less than 500 bp) we included the first intron (if this was 300 base pairs or larger) in order to increase the regulatory space contained in the reporters. The coordinates for all reporter constructs can be found in the Supplementary Material.

Sites of expression within the nervous system were determined mainly for those reporters with the most robust expression patterns and was based on stereotyped cell position, cellular and process morphology, and co-labeling with either DiD (which labels a subset of sensory neurons) or by crossing with landmark strains in which specific neuron classes are labeled with a red fluorescent protein (see [Material and methods](#)). All cell identification was initially done in young adult hermaphrodite animals. As we will describe in detail later, a number of these reporter strains were also subjected to analysis at different stages, under different conditions, and in the two different sexes.

GPCRs are expressed in restricted patterns within and outside the nervous system

In our ensuing description of expression patterns of reporter genes, we summarize the expression observed with the previously described reporters, as well as the additional reporters analyzed by us. All of our expression analysis is summarized in a tabular form in [S1 Table](#). Three overall features of the 375 csGPCR reporters are immediately apparent ([Fig 1](#)): first, 92% of analyzed reporters are expressed in the nervous system; second, expression is not restricted to the nervous system: 33% of the reporters are expressed both within and outside the nervous system and 8% are expressed exclusively in non-neuronal cells; and third, the vast majority of csGPCR reporters are expressed in very restricted numbers of cells ([Fig 1A and 1B](#)). Of the neuronally-expressed reporters, 24% are expressed in single neuron pairs, 27% in 2 neuron pairs, 26% in 3–4 neuron pairs, 19% in 5–10 neuron pairs, and the remaining 4% in more than 10 neuron pairs.

Expression outside the nervous system will be described in a later section. Within the nervous system, expression is most prominent in sensory neurons ([Fig 1C](#)). 84% of the reporters are expressed in amphid sensory neurons (which are made up of 12 pairs of neurons), 20% in phasmid sensory neurons (made up of 2 pairs of neurons, PHA and PHB), and 17% in other sensory neurons. We find that every sensory neuron, except for URY and ADE neurons, expresses at least 1 GPCR ([Fig 1D](#); [S2 Table](#)). The number of GPCRs expressed in a given neuron class shows a striking range. The ASI neuron class expresses an impressive 99 GPCR reporters. After ASI, the nociceptive neurons ADL and ASH together with the phasmid neurons PHA and PHB are the sensory neuron classes with higher numbers of GPCRs, expressing 72, 51, 51, and 49 reporters, respectively. Outside the amphid and phasmid neurons, the number of reporters expressed in sensory neurons dramatically drops, with all other sensory neurons expressing less than 10 GPCRs, in some cases only a single GPCR ([Fig 1](#); [S2 Table](#)). Of course, it needs to be kept in mind that we only consider expression of a fraction of the csGPCR loci, and hence each of these total numbers is expected to increase by several fold once all csGPCR expression patterns are identified.

Twenty-four percent of the GPCR reporters for which we have information about neuron numbers are exclusively expressed in a single neuron class, and in all these cases the neuron class is a sensory neuron class ([Fig 2](#); [S3 Table](#)). In total, however, only 9 sensory neuron classes express single-neuron-specific GPCRs. The most striking one of them is the ADL nociceptive neuron class, which expresses 23 single-neuron-specific GPCR reporters (and an additional

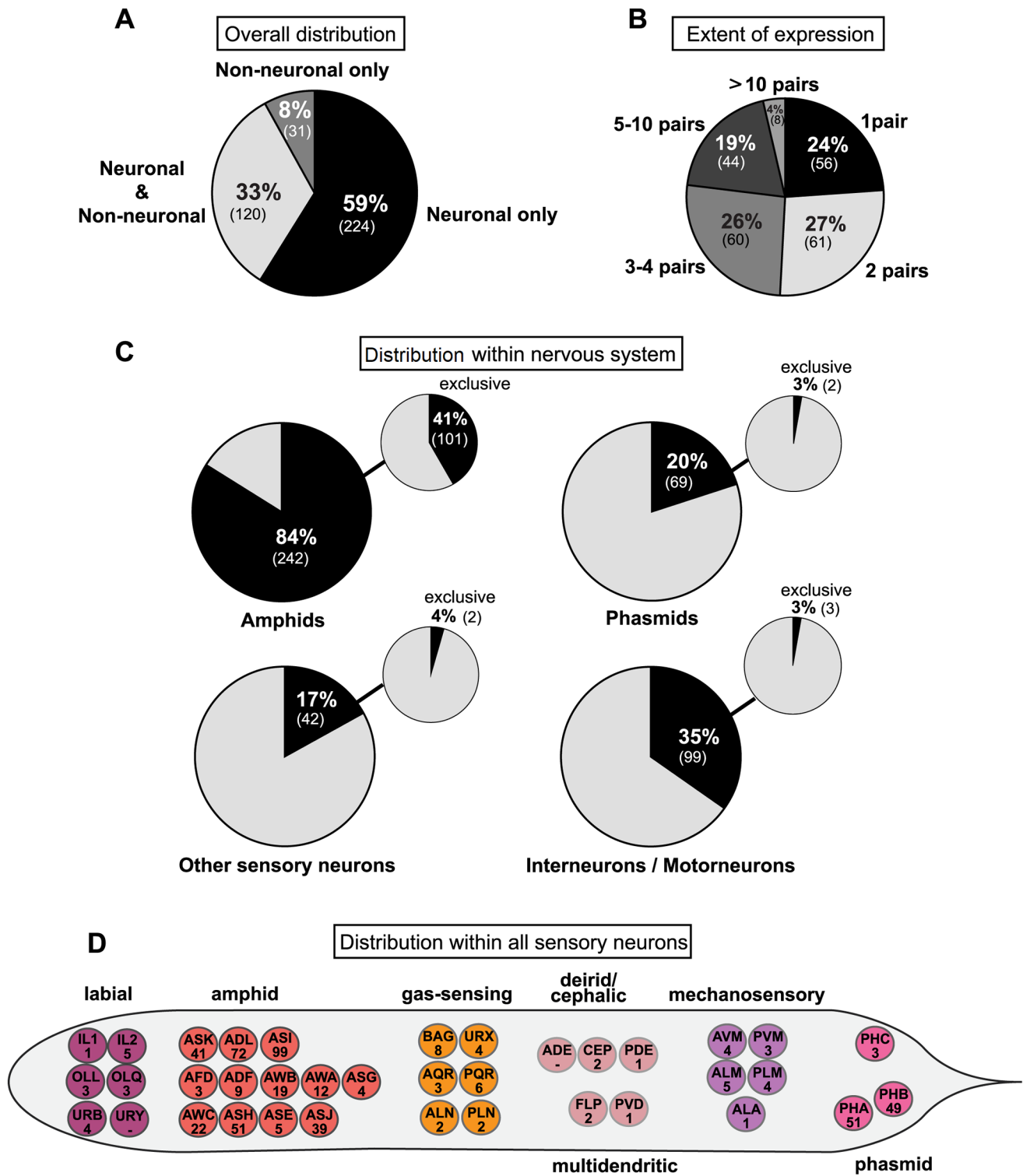


Fig 1. Summary of csGPCR reporter expression patterns. (A) Overall tissue distribution of reporter expression patterns in hermaphrodites. Pie chart showing percentage of GPCR reporters expressed exclusively in neurons, in neurons and other cells types, and exclusively in non-neuronal tissues. Numbers in parentheses represent the absolute number of reporters in each category. (B) Extent of reporter expression within the nervous system. Pie chart showing percentage of neuronal reporters expressed in 1 neuron pair, 2 pairs, 3–4 pairs, 5–10 pairs, or more than 10 pairs. Numbers in parentheses represent the absolute number of reporters in each category. (C) Distribution of reporter gene expression within the nervous system. Pie charts showing percentage of GPCR reporters expressed in amphid neurons, phasmid neurons,

other sensory neurons, and inter- or motoneurons. Small pie charts on the upper right represent the percentage of reporters exclusively expressed in amphids, phasmids, other sensory neurons, and inter- or motoneurons. Numbers in parentheses show the absolute number of reporters in each category. **(D)** Distribution within all sensory neurons of the hermaphrodite. Worm schematics showing the absolute number of GPCR reporters found to be expressed in each sensory neuron class. PHC is a phasmid neuron by name only. See [S2 Table](#) for a list of GPCR gene names expressed in the sensory neurons shown here. GPCR, G-protein-coupled receptor.

<https://doi.org/10.1371/journal.pbio.2004218.g001>

49 GPCR reporters expressed in additional neurons). The ADL-expressed, single-neuron-specific GPCRs do not fall into a specific GPCR subfamily but rather cover 7 distinct families. A small subset of the single neuron type-specific GPCRs are expressed outside the nervous system as well (genes with asterisk in [Fig 2A](#)). This may indicate that these receptors do not detect external cues, but rather sense internal signals.

Notably, expression of the csGPCR reporter collection is clearly not restricted to sensory neurons. A striking 35% of the csGPCR reporters are expressed in inter- and motoneurons ([Figs 1 and 3](#); [Table 3](#); [S1 Table](#)). There is no unifying feature of the inter- or motoneurons that express GPCR reporters. They range from ventral cord motor neurons to head interneurons, and to command interneurons in the ventral cord. One interneuron, PVT, displays a very large number of expressed csGPCR reporters (57 different reporters); however, PVT expression is generally observed in an unusually large amount of reporter genes and may, like posterior gut expression, be a reporter gene artifact that relies on cryptic regulatory elements in the reporter gene construct.

Ninety-seven percent of inter- and/or motoneuron-expressed csGPCR reporters are also expressed in sensory neurons so only 3% of them show expression exclusively in inter- or motoneurons. In light of the inter-/motoneuron expression of csGPCR reporters, we can hypothesize that csGPCR reporters that are expressed in sensory neurons may actually not function as receptors for external sensory cues, but may rather function as they likely do in inter-/motoneurons, i.e., as receptors of internal signals.

We asked whether csGPCR expression profiles cluster by neuron class. To this end, we undertook unsupervised hierarchical clustering of expression profiles. The BP value for most associations was very weak with two exceptions: csGPCR reporters are often coexpressed in the two tail phasmid neuron classes PHA and PHB (AU/BP > 95), and expression in either or both of the phasmid neurons is associated with the expression in the head neuron ASH (AU/BP > 95) ([Fig 4](#)). The ASH, PHA and PHB neuron classes are not closely related by lineage but all of these three neuron classes are nociceptive neurons that respond to similar cues and integrate sensory inputs from the head and tail [55,56] and that directly innervate command interneurons involved in reversal behavior [9]. While csGPCRs expressed in these neurons are likely involved in sensing nociceptive cues, it is notable that these coexpressed csGPCRs came from widely distinct csGPCR families ([Fig 4](#)).

Left/right asymmetric expression of csGPCR reporters

One major motivation for undertaking the csGPCR reporter analysis was to identify more lateralized neuron pairs in the nervous system. In vertebrates, there is a striking dearth of molecular correlates for widespread functional lateralization of the brain. In *C. elegans*, the chance discovery of left/right asymmetric sensory receptor expression has opened up new vistas on lateralization of the *C. elegans* nervous system [58]. Specifically, the lateralized expression of several csGPCRs in the AWC olfactory neuron pair [31] and guanylyl cyclase receptors in the gustatory ASE neuron pair [59] revealed a common theme of lateralization, providing means of sensory discrimination [32,60,61]. Since lateralization provides an elegant, straight-forward means for sensory discrimination, we speculated that such lateralization may be widespread in

A

| | |
|------------|--|
| ASK | <i>sra-7*</i> , <i>sra-9</i> , <i>srb-5</i> , <i>srbc-64</i> , <i>srbc-66</i> , <i>srd-36</i> , <i>srg-2</i> , <i>srg-8</i> , <i>srw-108*</i> , |
| ADL | <i>sre-21*</i> , <i>srh-25</i> , <i>srh-132*</i> , <i>srh-193*</i> , <i>srh-199</i> , <i>srh-279</i> , <i>srh-281</i> , <i>sri-9</i> , <i>sri-28</i> , <i>sri-45*¹</i> , <i>sri-50¹</i> , <i>sri-51</i> , <i>sru-4</i> , <i>srv-3</i> , <i>srv-4</i> , <i>srw-138</i> , <i>srz-4</i> , <i>srz-6</i> , <i>srz-28</i> , <i>srz-61*¹</i> , <i>srz-67</i> , <i>srz-99*</i> , <i>srz-103</i> |
| ASI | <i>sra-17</i> , <i>srd-1</i> , <i>srg-47</i> , <i>srj-23*</i> , <i>str-3</i> , <i>str-231</i> |
| AWA | <i>odr-10</i> |
| AWB | <i>srab-1</i> , <i>str-1</i> , <i>str-44</i> , <i>str-163*</i> , <i>str-220</i> |
| ASG | <i>srw-119</i> |
| ADF | <i>srx-12</i> |
| ASH | <i>srv-11</i> |
| PHB | <i>srab-20</i> , <i>srx-110</i> |

B

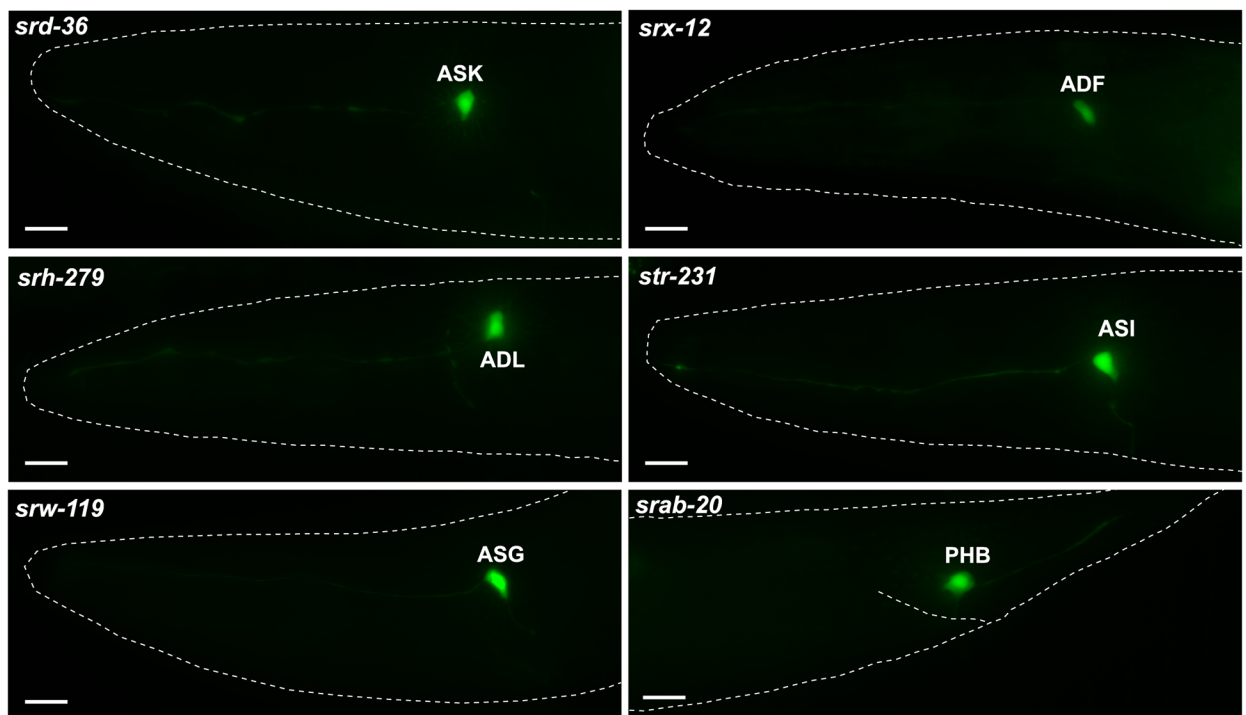


Fig 2. csGPCR reporters expressed in single sensory neuron classes. (A) Table showing all GPCR reporters expressed in a single neuron class. Genes in bold are newly identified in this paper. Genes not in bold were previously described (data extracted from www.wormbase.org). * Reporter is also expressed in some non-neuronal tissue (for details, see S1 Table). ¹ N. Masoudi, S. Finkelstein, and O. Hobert, in preparation. (B) Representative examples of reporters expressed in a single neuron class identified in this study. Young adult hermaphrodites are shown. Scale bars, 10 μm. GPCR, G-protein-coupled receptor.

<https://doi.org/10.1371/journal.pbio.2004218.g002>

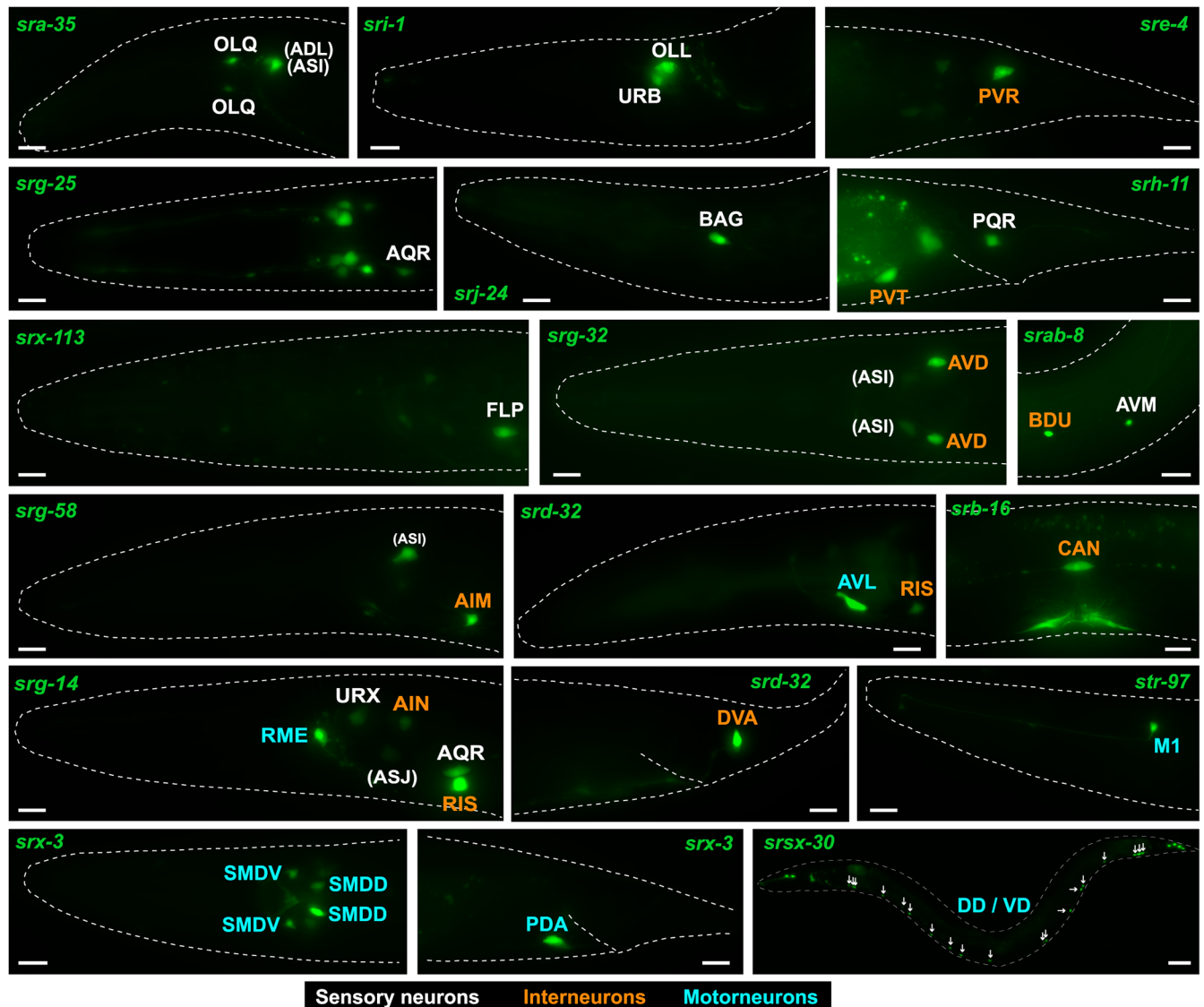


Fig 3. csGPCR reporters expressed in non-amphid/non-phasmid sensory neurons, interneurons, and motorneurons. Examples of GPCR reporters expressed in sensory neurons that are not amphids or phasmids (white font), interneurons (orange font), or motorneurons (blue font). Most examples represented here are from neuron classes that were not previously shown to express any sensory GPCR. Amphid neurons are shown in parentheses. Young adult hermaphrodites are shown. All scale bars, 10 μ m, except *srsx-30*, which is 30 μ m. See Table 3 for a complete summary of GPCR reporters expressed in inter- and motorneurons. GPCR, G-protein-coupled receptor.

<https://doi.org/10.1371/journal.pbio.2004218.g003>

the nervous system and therefore took particular care in examining whether csGPCR reporters that we analyzed are expressed in a left/right asymmetric manner.

We indeed identified 8 csGPCR reporters with left/right asymmetric gene expression in an otherwise bilaterally symmetric neuron pair. However, this laterality was only observed in the context of the AWC sensory neuron pair, which was previously known to express several csGPCRs in a left/right asymmetric manner [31,62]. Using previously described sets of mutants, we found that the asymmetry of these GPCR reporters is controlled by the same calcium-dependent signaling pathway [33] that controls all other previously known asymmetric GPCRs in the AWC neurons (Fig 5). Of course, our limited analysis does not exclude the

Table 3. Non-sensory neurons expressing csGPCR reporter.

| | | CLASS | REPORTER GENES | |
|--------------------|--------------|---------------------|--|---|
| INTERNEURONS | Head | ADA | <i>srab-12, sri-1</i> | |
| | | AIA | <i>sra-11, srab-4, (srh-269)</i> | |
| | | AIB | <i>srh-11</i> | |
| | | AIM | <i>srg-32, srg-58, srxa-14</i> | |
| | | AIN | <i>srg-14, srh-277</i> | |
| | | AIY | <i>sra-11, srab-3, sri-1, sri-12, sri-36, srx-14, str-102</i> | |
| | | AUA | <i>sre-4</i> | |
| | | AVB | <i>sra-11</i> | |
| | | AVD | <i>srg-32</i> | |
| | | AVE | <i>srab-24</i> | |
| | | AVG | <i>srn-1, (srsx-12)</i> | |
| | | RIF | <i>sra-11, srab-3, (srh-266)</i> | |
| | | RIG | <i>sre-4</i> | |
| | | RIH | <i>(srn-5), (srn-6)</i> | |
| | | RIS | <i>srd-32, srg-14</i> | |
| | SAA | <i>srx-3</i> | | |
| | Midbody | BDU | <i>srab-8, srab-12, sre-4, sri-1, sri-18, srv-27</i> | |
| | | CAN | <i>srb-16, srd-32, srv-1</i> | |
| | | SDQ | <i>srab-12</i> | |
| | Tail | DVA | <i>srd-32, srx-113</i> | |
| | | DVC | <i>srab-4</i> | |
| | | LUA | <i>srab-12</i> | |
| | | PVP | <i>srab-12</i> | |
| | | PVQ | <i>sra-6, sre-4, srg-32, srh-277, sri-1, (sru-17), srv-32, str-84</i> | |
| | | PVR | <i>sre-4</i> | |
| | | PVT | <i>sra-11, sra-28, srab-4, srb-7, srb-16, srbc-52, srd-32, sre-11, sre-22, sre-30, sre-52, srg-4, srg-14, srg-31, srg-39, srh-4, srh-5, srh-11, srh-62, srh-71, srh-210, srh-241, srh-266, sri-12, sri-36, sri-39, sri-62, srj-5, srj-20, srj-27, srj-38, srr-2, srr-7, srr-8, srsx-12, srsx-38, sru-8, sru-48, srx-10, srx-17, srxa-7, srz-13, srz-27, srz-54, srz-102, srz-104, str-31, str-52, str-123, str-143, str-178, str-217, str-233, str-236, str-247, str-249, str-250</i> | |
| | MOTORNEURONS | Head | AVL | <i>srd-32</i> |
| | | | RID | <i>sra-14, srab-3, srab-4, sre-4</i> |
| | | | RMD | <i>(sri-21)</i> |
| | | | RMDD | <i>srr-3</i> |
| | | | RMDV | <i>srr-3</i> |
| | | | RME | <i>srab-4, srg-14</i> |
| | | | RMG | <i>srab-12</i> |
| SMD | | | <i>srx-3</i> | |
| SIA | | | <i>sro-1</i> | |
| Midbody | | HSN | <i>sra-35, srab-8, srj-13</i> | |
| Ventral nerve cord | | DA | <i>sra-36</i> [DA8, DA9], <i>srb-16</i> [DA9], <i>srd-4</i> [DA9] | |
| | | DB | <i>srx-3</i> | |
| | | DD | <i>srsx-30</i> | |
| | | VA | <i>srab-4, sra-36</i> [VA11] | |
| | | VB | <i>srab-4, srx-3</i> | |
| | | VC | <i>sra-11, srb-16</i> | |
| | | VD | <i>srsx-30</i> | |
| Tail | | PDA | <i>srx-3</i> | |

Abbreviations: GPCR, G-protein-coupled receptor; HSN, hermaphrodite-specific neuron.

Bolded gene: Newly identified in this paper. Cell identifications were confirmed with neuron-specific landmark reporters (see [Material and methods](#)). Non-bolded gene: previously identified. (Gene in parentheses): ID based on position and morphology, not confirmed with neuron-specific landmark reporter.

<https://doi.org/10.1371/journal.pbio.2004218.t003>

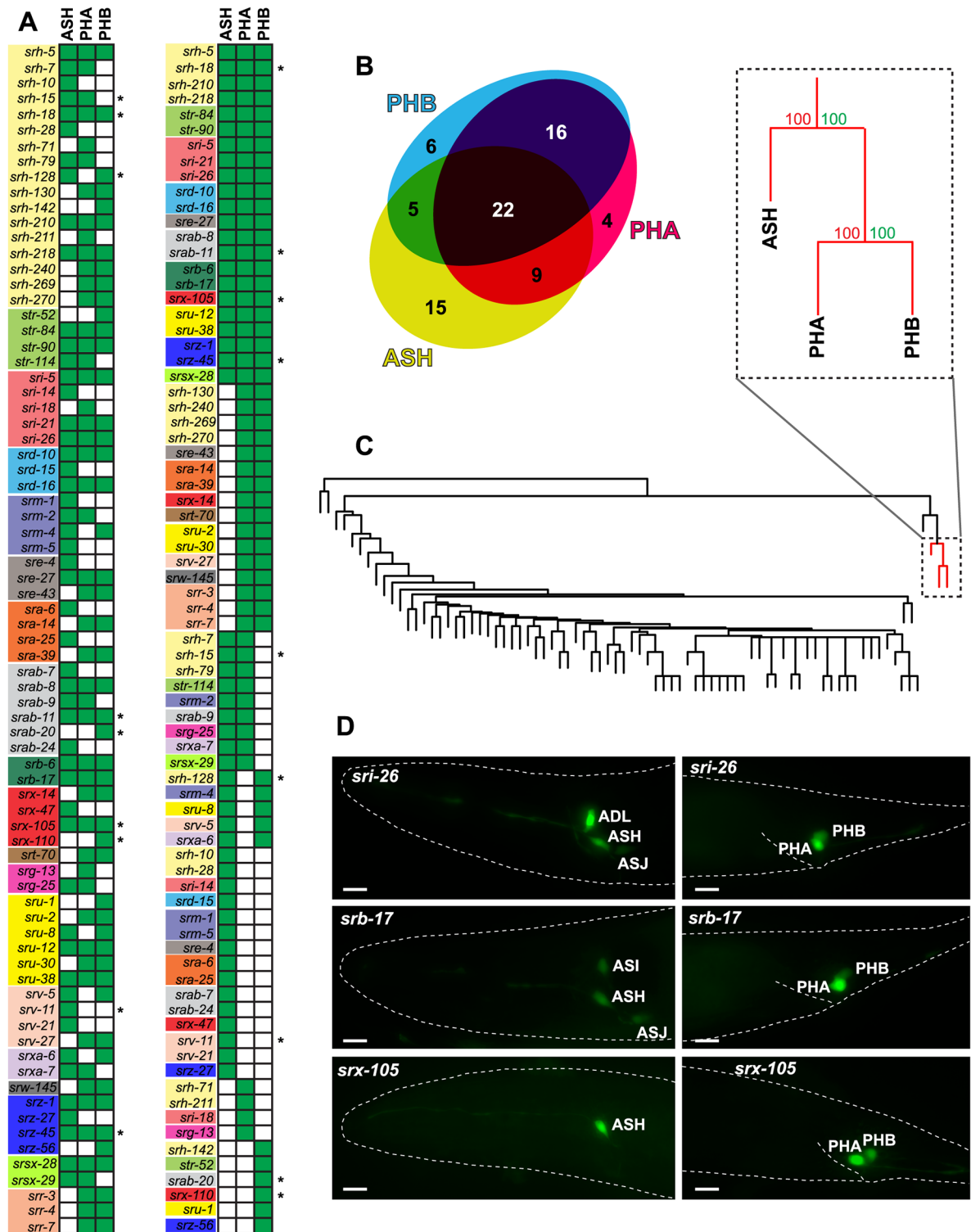


Fig 4. The only coexpression association of csGPCR reporters is in nociceptive neurons. (A,B) Graphical representation of ASH, PHA, and PHB coexpression. Green-filled square indicates expression. An asterisk denotes that the gene is exclusively expressed in the indicated neurons. Venn diagram was created with eulerAPE [57]. (C) Hierarchical clustering of neurons by GPCR reporter expression. Red lines show the well-supported ASH, PHA and PHB cluster (AU > 95%). BP values (in green) are listed in percentages. (D) Examples of reporter gene expression profiles in ASH/PHA/PHB. Young adult hermaphrodites are shown. Scale bars, 10 μm. AU, approximately unbiased *p*-value; BP, bootstrap probability value; GPCR, G-protein-coupled receptor.

<https://doi.org/10.1371/journal.pbio.2004218.g004>

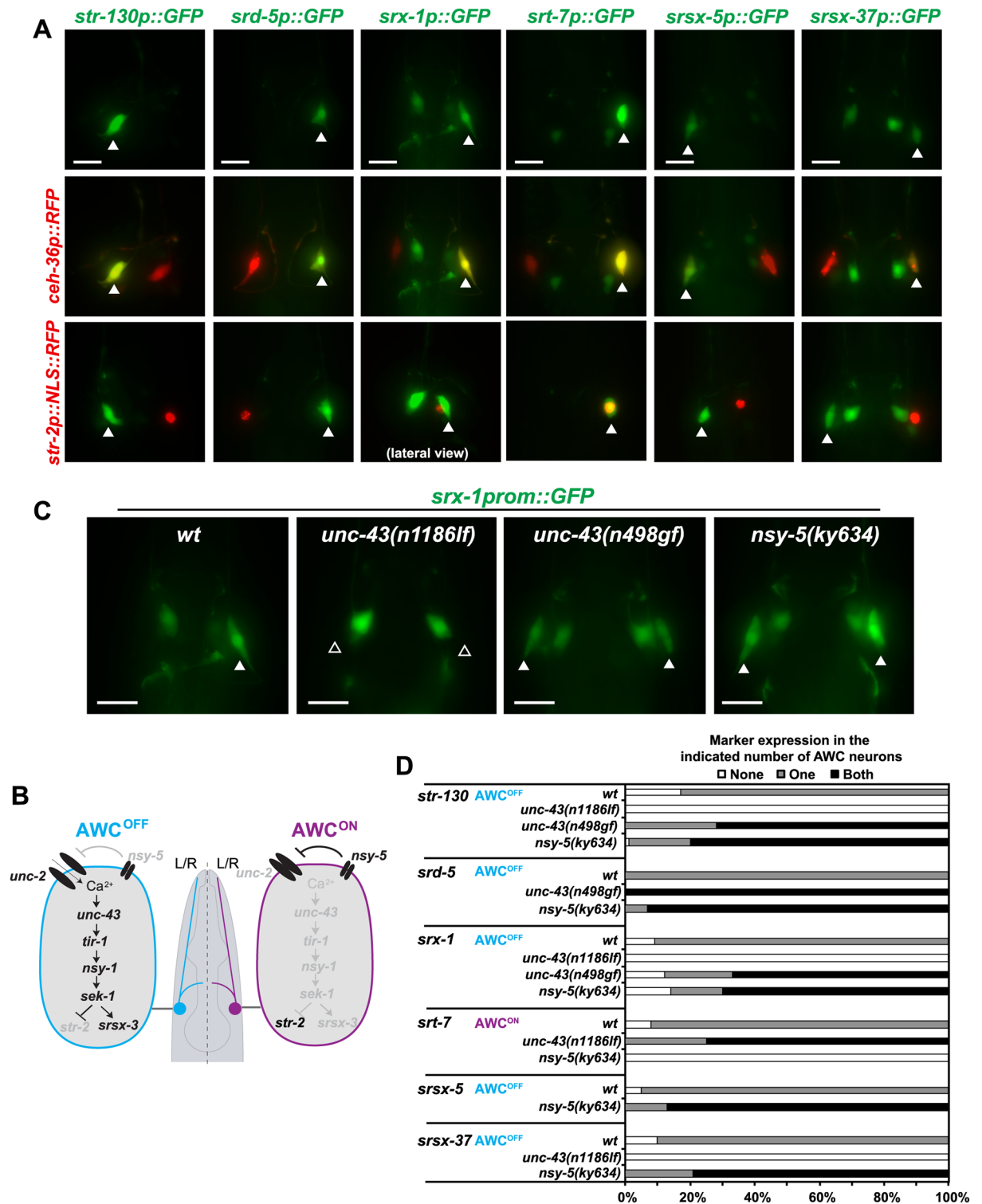


Fig 5. Lateralized csGPCR reporter expression in the AWC neuron pair. (A) Asymmetrically expressed GPCRs, indicated with arrowheads (top row), were all expressed in AWC as assessed by colocalization with the *ceh-36p::RFP* reporter (middle row). *str-130*, *srd-5*, *srx-1*, *srsx-5*, and *srsx-37* reporters were expressed in *AWC^{OFF}* while *srt-7* was expressed in *AWC^{ON}* as assessed with the *str-2p::NLS::RFP* reporter, which is an *AWC^{ON}* marker (bottom row). All pictures are dorso-ventral views unless otherwise indicated. *srt-13* and *srr-9* reporters were also found to be asymmetrically expressed in AWC; however, since these reporters were dim and not very robust, no further analysis was done. Young adult hermaphrodites are shown. Scale bars, 10 μ m. **(B)** AWC

asymmetry. Previously known components of genetic pathways that control AWC asymmetries [33]. Not all genes known to be involved are shown. Black and grey gene names indicate whether a gene is more active or more expressed (black) in one neuron compared with the other neuron. Scheme adapted from [63]. **(C)** Expression of the newly found AWC asymmetric GPCRs is regulated by previously described mechanisms. Representative pictures showing *srx-1* reporter expression (AWC^{OFF}) in different mutants of the previously described AWC asymmetry pathway. As expected, in *unc-43(n1186lf)* mutants, *srx-1* reporter is expressed in none of the AWC neurons (2 AWC^{ON} phenotype) while in *unc-43(n498gf)* and *nsy-5(ky634)* mutants *srx-1* is expressed in both AWC neurons (2 AWC^{OFF} phenotype). Scale bars, 10 μ m. **(D)** Expression quantification of AWC asymmetric GPCR reporters in *unc-43(n1186lf)*, *unc-43(n498gf)*, and *nsy-5(ky634)* mutants. Animals were scored as young adults and show the expected 2 AWC^{ON} or 2 AWC^{OFF} phenotype. Between 20 and 40 hermaphrodites were scored per genotype. GPCR, G-protein-coupled receptor.

<https://doi.org/10.1371/journal.pbio.2004218.g005>

existence of left/right asymmetrically expressed GPCR genes in other neuron classes, but it may not be as widespread as we initially hypothesized.

Sexually dimorphic expression of csGPCR reporters

Apart from brain lateralization, another domain of nervous system research displays a striking dearth of molecular markers. While the existence of sex-specific neurons is widely appreciated in the nervous system of most animals, including *C. elegans* [64], it is much less clear to what extent neurons that are shared by the two sexes of a given species display molecular differences. Recent anatomical work in *C. elegans* revealed intriguing synaptic wiring differences between sex-shared neurons in the two sexes [65], but even in *C. elegans* there is a dearth of sexually dimorphic molecular markers of sex-shared neurons. Given the distinct priorities that males and hermaphrodites display toward food and mate searching [66], and given that a number of sex-shared sensory neurons are known to respond to different cues in a sex-specific manner [49,67], we hypothesized that we may discover a multitude of sex-specifically expressed GPCRs. We indeed identified several GPCRs that are expressed in hermaphrodite-specific neurons (HSNs, VC motor neurons) or in several male-specific neurons (Fig 6); however, we did not detect differences in GPCR expression in sex-shared neurons. We emphasize here, however, that we did not systematically analyze all 244 reporters that we analyzed in the hermaphrodite for differences in expression in the male, but rather focused on those GPCRs that show expression in 1–3 pairs of neurons in the hermaphrodites.

csGPCR reporter expression outside the nervous system

Moving outside the nervous system, we found expression of individual GPCRs in essentially all tissue types (Fig 7 shows examples; summarized in Table 4). As we already mentioned above, the non-neuronal expression is often quite specific and there are only a few GPCRs that are expressed broadly in many different cell types (e.g., *srbc-58*, *srr-4*). Specific sites of non-neuronal expression include subsets of muscle cells, hypodermal cells, specialized epithelial cells, cells of the somatic gonad (distal tip cells), individual cells of the excretory system, glial cells, and others (Fig 7, Table 4). There are no obvious, specific associations of non-neuronal expression with expression in a specific set of neuron types. Also, non-neuronally expressed GPCR receptors are not biased toward a single subfamily. GPCRs expressed in non-neuronal tissues that are exposed to the environment, e.g., epidermis, could be involved in sensing external cues, but other non-neuronal cells will rather respond to internal signals. As a cautionary note, we can not presently exclude that non-neuronal expression may be the result of lack of repressor elements in the reporter constructs, but we note that in *C. elegans* there is presently little evidence for non-neuronal repressor mechanisms restricting gene expression to the nervous system (e.g., [68]).

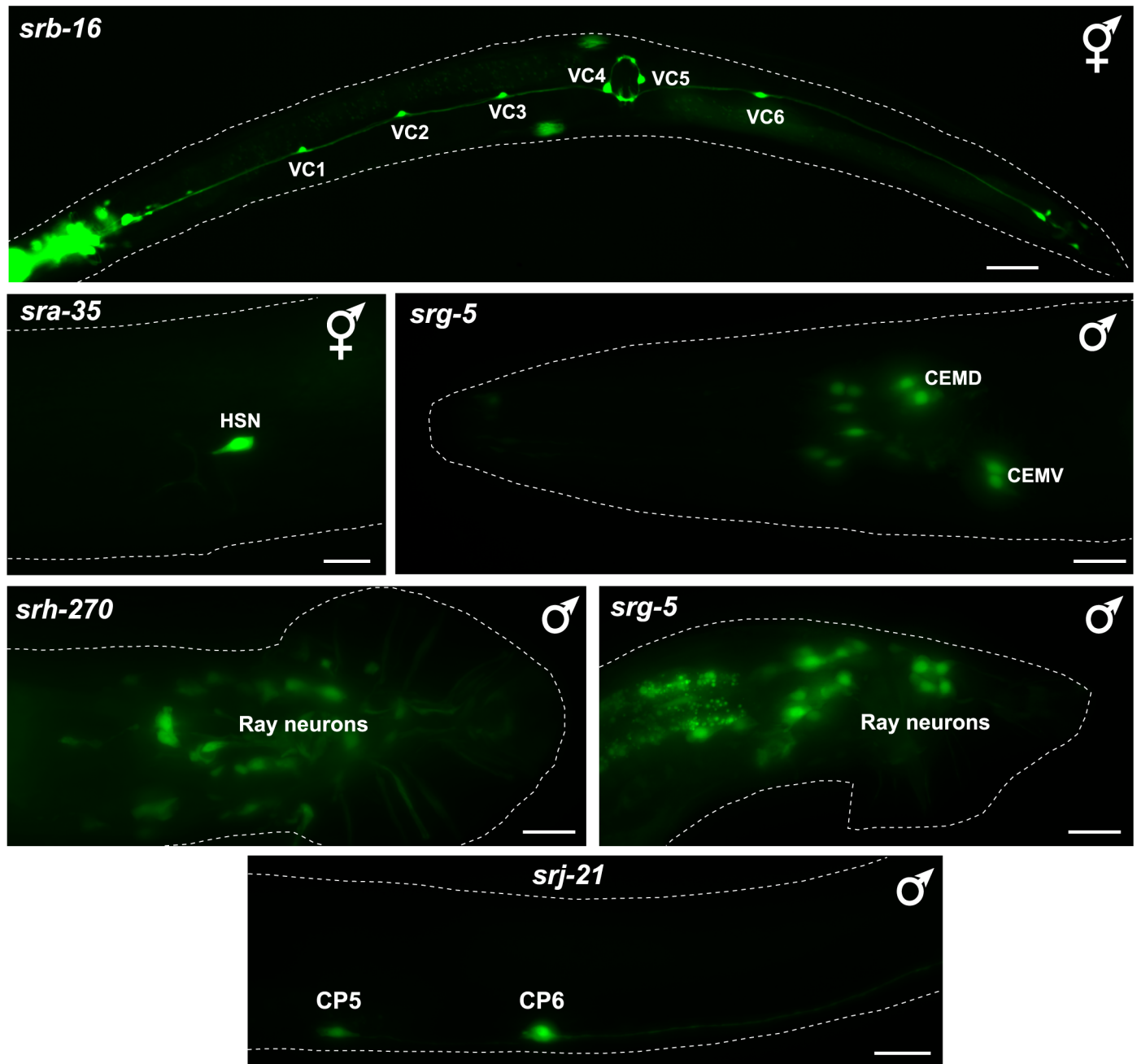


Fig 6. Expression of sex-specifically expressed csGPCR reporters. Examples of GPCR reporters expressed in hermaphrodite-specific (VCs, HSN) or male-specific neurons (CEMs, CP5, CP6, Rays). Young adult animals are shown. All scale bars, 10 μ m, except *srb-16*, which is 30 μ m. GPCR, G-protein-coupled receptor; HSN, hermaphrodite-specific neuron.

<https://doi.org/10.1371/journal.pbio.2004218.g006>

Reporter gene analysis of entire csGPCR gene families

Do any of the patterns described above cluster with sequence similarity (i.e., family membership) of the receptors? As described above, specific features of csGPCR expression patterns do not correlate with family membership, but we wanted to pursue this issue further via a more comprehensive analysis of entire chemoreceptor gene families. As defined by sequence analysis

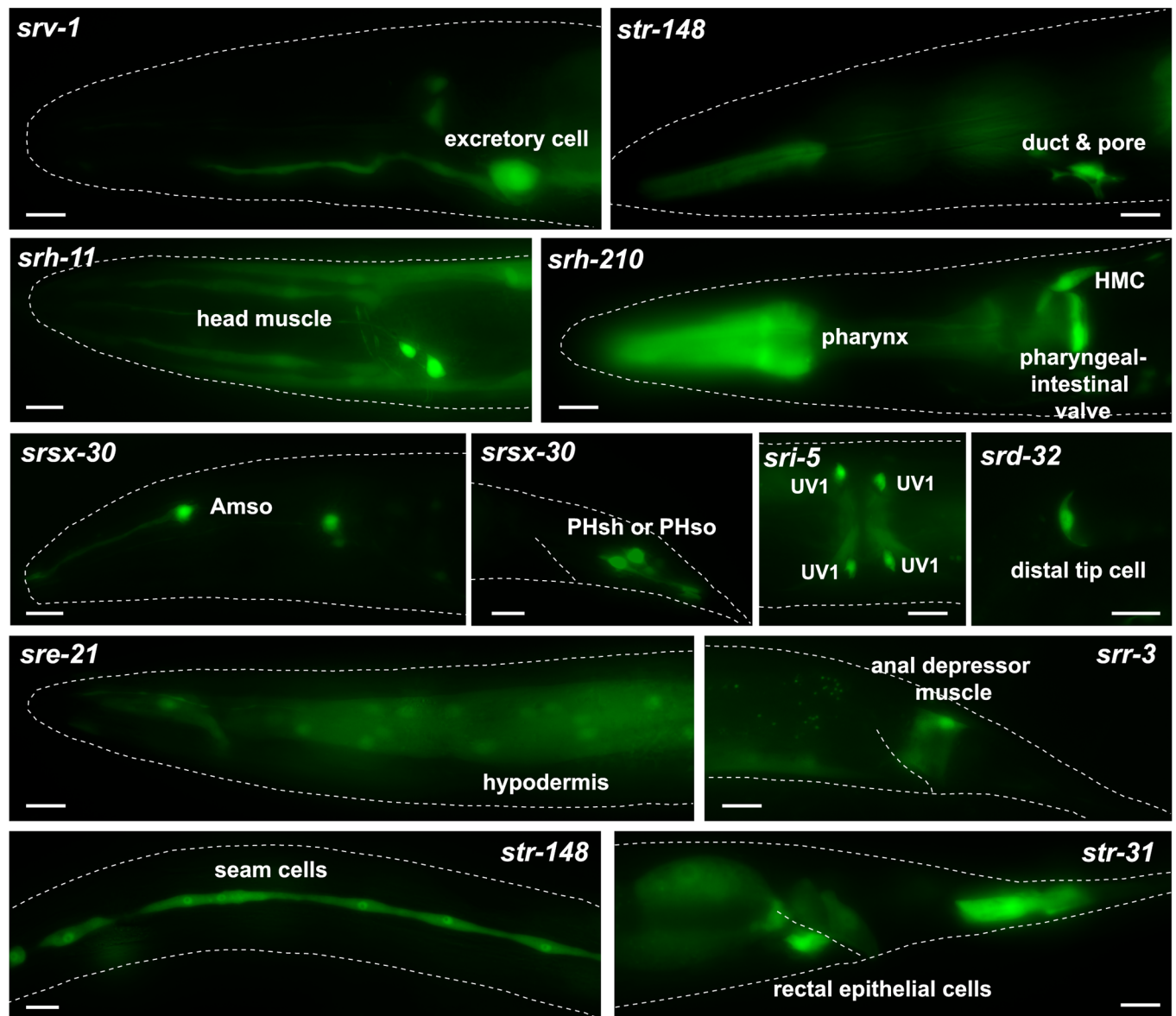


Fig 7. Expression of non-neuronal csGPCR reporters. Examples of GPCR reporters expressed in different types of non-neuronal tissue in young adult hermaphrodites. Scale bars, 10 μ m. See Table 4 for a complete summary of GPCR reporters expressed in non-neuronal tissues. Amso, amphid socket cells; GPCR, G-protein-coupled receptor; HMC, head mesodermal cell; PHsh, phasmid sheath; PHso, phasmid socket; UV1, uterine/vulval cell 1.

<https://doi.org/10.1371/journal.pbio.2004218.g007>

[7], chemoreceptor gene families have very different sizes, ranging from a single gene per family (*srn* family) to 223 genes per family (*srh* family) (Table 2). We analyzed reporter gene expression patterns of all members of two small families to examine whether there are common themes in their expression patterns, their genomic location, and *cis*-regulatory control regions. We also analyzed the expression of the one family, the *srn* family, which only has a single member and is highly conserved in other *Caenorhabditis* species, to assess whether it may show an unusual expression pattern. However, we find the *srn-1* reporter gene to be mainly expressed in amphid sensory neurons, like many other csGPCRs (Fig 8).

The two small families for which we generated and analyzed reporter genes for all family members are the previously uncharacterized *srn* (six members) and *srr* (nine members). Five

Table 4. Non-neuronal sites of csGPCR reporter expression.

| TISSUE / CELL | REPORTER GENES |
|-------------------------------|--|
| Coelomocytes | <i>srh-193</i> , <i>srh-269</i> , <i>srj-4</i> , <i>str-250</i> |
| Excretory system ¹ | <i>srab-14</i> , <i>srm-3</i> , <i>srr-4</i> , <i>srr-6</i> , <i>srr-8</i> , <i>srv-1</i> , <i>str-143</i> , <i>str-148</i> |
| Glia | <i>srab-8</i> , <i>srh-270</i> , <i>srr-1</i> , <i>srsx-30</i> , <i>sru-2</i> , <i>sru-19</i> , <i>srw-29</i> , <i>srw-145</i> , <i>str-47</i> |
| Gonad | <i>srbc-58</i> , <i>srd-32</i> , <i>sre-24</i> , <i>srh-87</i> |
| Gut ² | <i>srb-17</i> , <i>srh-211</i> , <i>srm-3</i> |
| Head mesodermal cell | <i>srb-16</i> , <i>srd-32</i> , <i>srh-132</i> , <i>srh-210</i> , <i>srh-269</i> , <i>srr-3</i> , <i>srx-1</i> |
| Hypodermis | <i>sra-13</i> , <i>sra-39</i> , <i>srab-6</i> , <i>srab-13</i> , <i>srab-21</i> , <i>srbc-58</i> , <i>srd-39</i> , <i>sre-7</i> , <i>sre-21</i> , <i>sre-22</i> , <i>sre-29</i> , <i>sre-53</i> , <i>srh-76</i> , <i>srr-4</i> , <i>sru-31</i> , <i>srw-108</i> , <i>srw-118</i> , <i>srz-13</i> , <i>srz-94</i> , <i>srz-99</i> , <i>str-31</i> , <i>str-168</i> , <i>str-250</i> |
| Muscle | <i>sra-2</i> , <i>sra-13</i> , <i>srab-7</i> , <i>srb-17</i> , <i>srbc-58</i> , <i>srd-15</i> , <i>srd-32</i> , <i>sre-22</i> , <i>sre-29</i> , <i>srg-7</i> , <i>srg-29</i> , <i>srg-31</i> , <i>srh-11</i> , <i>srh-100</i> , <i>sri-19</i> , <i>srr-3</i> , <i>srt-20</i> , <i>sru-1</i> , <i>srx-1</i> , <i>srx-41</i> , <i>srx-41</i> , <i>srx-41</i> , <i>srz-94</i> , <i>str-102</i> , <i>str-111</i> , <i>str-114</i> |
| Pharynx | <i>sra-4</i> , <i>sra-10</i> , <i>sra-38</i> , <i>srb-6</i> , <i>srb-16</i> , <i>srbc-58</i> , <i>srd-15</i> , <i>srd-32</i> , <i>srg-29</i> , <i>srg-31</i> , <i>srg-39</i> , <i>srg-62</i> , <i>srh-7</i> , <i>srh-62</i> , <i>srh-71</i> , <i>srh-92</i> , <i>srh-100</i> , <i>srh-142</i> , <i>srh-201</i> , <i>srh-210</i> , <i>srh-269</i> , <i>sri-5</i> , <i>sri-36</i> , <i>srj-4</i> , <i>srj-5</i> , <i>srj-13</i> , <i>srj-38</i> , <i>srm-1</i> , <i>srm-3</i> , <i>srr-1</i> , <i>srr-2</i> , <i>srr-3</i> , <i>srr-4</i> , <i>srr-6</i> , <i>srt-65</i> , <i>sru-1</i> , <i>sru-31</i> , <i>srv-17</i> , <i>srx-10</i> , <i>srz-54</i> , <i>str-52</i> , <i>str-85</i> , <i>str-108</i> , <i>str-121</i> , <i>str-123</i> , <i>str-143</i> , <i>str-236</i> , <i>str-247</i> , <i>str-250</i> |
| Rectal epithelium | <i>srbc-58</i> , <i>srx-4</i> , <i>str-31</i> , <i>str-233</i> , <i>str-250</i> |
| Seam cells | <i>sra-13</i> , <i>srb-17</i> , <i>srbc-58</i> , <i>srd-39</i> , <i>srh-130</i> , <i>srh-266</i> , <i>srj-20</i> , <i>srz-14</i> , <i>str-31</i> , <i>str-148</i> |
| Vulva | <i>sra-13</i> , <i>srab-7</i> , <i>srab-13</i> , <i>srb-16</i> , <i>srb-17</i> , <i>srbc-58</i> , <i>sre-56</i> , <i>srh-11</i> , <i>srh-130</i> , <i>srh-210</i> , <i>srh-270</i> , <i>sri-5</i> , <i>sri-19</i> , <i>srj-13</i> , <i>srr-4</i> , <i>srsx-12</i> , <i>srx-1</i> , <i>srx-4</i> , <i>srz-102</i> , <i>str-31</i> , <i>str-52</i> , <i>str-114</i> , <i>str-247</i> , <i>str-262</i> |

Abbreviations: GPCR, G-protein-coupled receptor.

Bolded gene: newly identified in this paper. Non-bolded gene: previously identified and retrieved from Wormbase.

See S1 Table for further details about specific sites of expression.

¹ The two *str* genes are in the excretory pore and duct cells, all others are in the excretory canal cell.

² Transcriptional *gfp* reporters often show posterior gut expression, which is considered an artifact. Only reporters showing bright expression throughout the gut are listed here. Previously described reporters with annotated gut expression in Wormbase are not included here.

<https://doi.org/10.1371/journal.pbio.2004218.t004>

out of the six *srm* family genes are syntenic to other family members (Fig 8). As these direct genomic adjacencies suggest local gene duplication, we could ask the question whether such local duplications also resulted in duplication of the 5' *cis*-regulatory control regions and to what extent such duplicated *cis*-regulatory control regions retained similar expression profiles. We find that the adjacent *srm-1* and *srm-2* genes are expressed in a small set of mostly sensory neurons; some of these neurons are the same, others are different. The same theme applies to the adjacent *srm-4*, *srm-5*, and *srm-6* genes. Their 5' upstream regions direct expression to distinct, but partially overlapping sets of neurons.

The *srr* gene family is composed of nine members. Reporter genes for all members displayed expression in diverse sets of neuron types with no common theme emerging. Outside the nervous system, it is notable that half of the family members are expressed in distinct cell types of the pharynx (Fig 8), suggesting a role for these genes in sensing food.

Temporally regulated csGPCR reporter genes

We also sought to examine dynamic aspects of csGPCR expression. We focused on dynamics that relate to developmental timing and the response to harsh environmental conditions. To

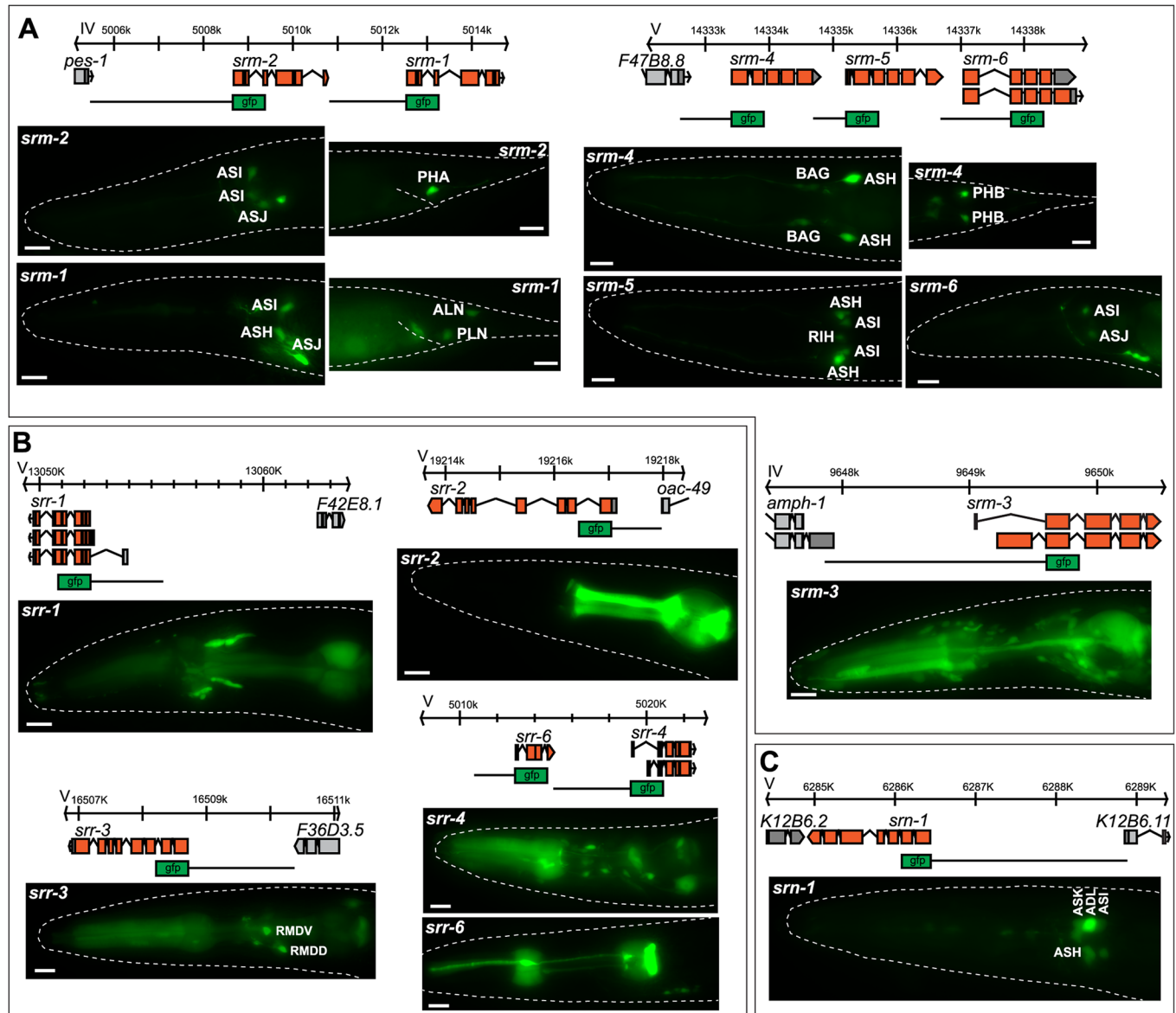


Fig 8. Reporter analysis of entire csGPCR families. Genomic loci, reporter scheme, and *gfp* expression images for the *srm* (A), *srr* (B), and *srn* (C) GPCR gene families. Only reporters expressed in the pharynx are shown for the *srr* family. Young adult hermaphrodites are shown. Scale bars, 10 μ m. GPCR, G-protein-coupled receptor.

<https://doi.org/10.1371/journal.pbio.2004218.g008>

facilitate the identification of changes in expression, we focused our analysis on GPCRs that are robustly expressed in the adult in a small number of neurons (in most cases not more than 1–3 neuron pairs in the head and/or 1–2 neuron pairs in the tail). At the first larval (L1) stage, we did not detect any differences in expression in 79 out of 82 examined reporters, compared to adults. Due to the limitations of multicopy array-based fluorescent reporters, moderate intensity changes within a cell type might be difficult to notice and could have been missed. Three reporter genes, *srh-11*, *sru-48*, and *sra-28*, show striking differences in L1 versus adult stages: all three reporter genes show expression in the ASK neuron at the L1 stage, but not at the adult stage (Fig 9). Additionally, *srh-11* is expressed brightly in the ASI neuron at the L1 stage but dimly at the adult stage (Fig 9). Furthermore, dim expression of *srh-11* and *sra-28*

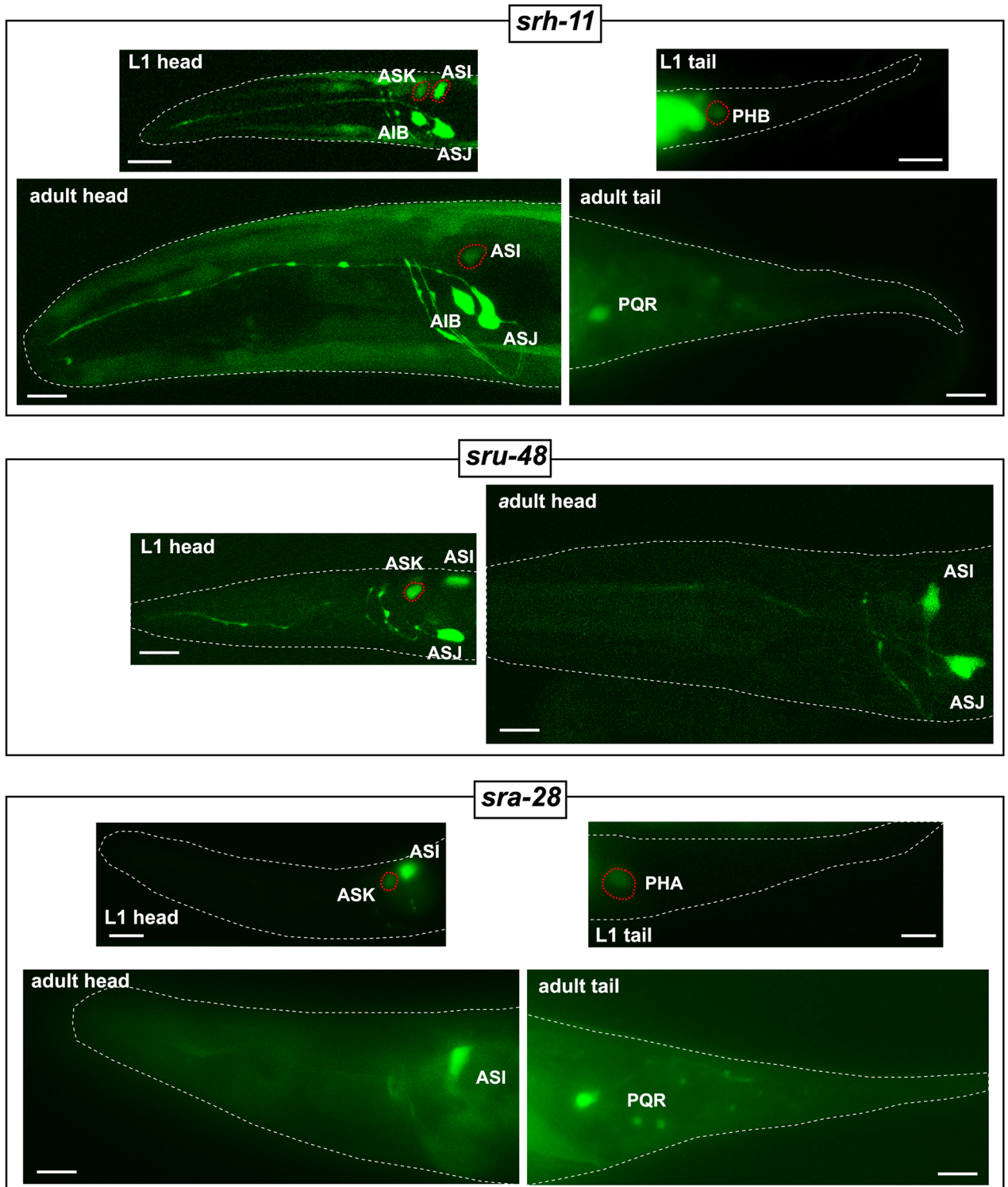


Fig 9. Temporal regulation of csGPCR reporters. GFP images showing temporal expression changes (L1 versus young adult) of *srh-11*, *sru-48* and *sra-28* reporter genes. Neurons showing temporal changes in expression are outlined with red dotted lines. Scale bars, 10 μ m. GFP, green fluorescent protein; GPCR, G-protein-coupled receptor; L1, first larval.

<https://doi.org/10.1371/journal.pbio.2004218.g009>

reporter genes in the tail phasmid PHB and PHA neurons, respectively, is only observed at the L1 stage but not at the adult stage (Fig 9).

csGPCR reporter gene expression changes in dauers

We found that a substantial number of csGPCR reporter genes were dynamically expressed when animals enter the dauer stage, an environmentally controlled diapause arrest stage that is accompanied by substantial cell, tissue, and behavioral remodeling [69,70]. Initially again focusing on reporters that are expressed in a restricted number of neurons under well-fed conditions, we found that 16 out of 46 examined reporters show a diverse set of changes in animals that were sent into the dauer stage via a standard starvation/crowding protocol (see Experimental Procedures). Many of the changes entail striking changes in the cellular specificity of GPCR reporter expression (Fig 10, Table 5). The vast majority of differences are observed in the nervous system, but some changes also occur outside the nervous system. Changes in GPCR reporter expression in the dauer stage have previously been described for two GPCR reporters [34] (summarized with our novel patterns in Table 5), but the patterns we observe here are much broader and more complex. They can be summarized as follows:

1. In most cases, there is stable and unchanged expression in several neuron classes in dauer and non-dauers, but upon dauer entry, expression is either turned on in additional neuron classes (“type I” regulation) or becomes undetectable in subsets of specific neuron classes (“type II” regulation) (Table 5; Fig 10). There are also combinations of both changes (type III regulation): in one particularly striking example, the *srh-71* reporter is expressed in some sensory neurons in both dauer and non-dauers, but undergoes a striking respecification in dauers. Reporter expression becomes undetectable in the lateral IL2, PHA, and an additional pair of tail neurons in dauer, and instead is turned on in the AIZ and ASG neurons (and increases expression levels in ASI). This hints toward the rerouting of internal sensory information.
2. In a number of cases, reporter expression is strongly down-regulated, becoming undetectable in all neurons in which the reporter is expressed (Table 5; Fig 10).
3. The changes outside the nervous system concern three tissue types: muscle, the excretory cell, and epithelial cells (Fig 10). In two cases, expression of a specific csGPCR reporter is turned on in the dauer stage, while in another case expression becomes undetectable. These findings indicate that these tissue types now became receptive to signals in a dauer-specific manner, an unanticipated finding.
4. The most recurrent set of changes in the expression of distinct reporters concern nociceptive neurons, namely the ASH, ADL, and phasmid tail neurons. Of particular note is the PHA phasmid neuron, which shows the most consistent pattern of changes: four csGPCRs are turned off or strongly down-regulated specifically in the dauer stage.
5. The most unusual novel expression pattern observed in dauer stage animals concerns the PVP tail interneuron pair. We found that in dauers, expression of the *sri-9* reporter is turned on in a left/right asymmetric manner, only in the PVPL neuron. The cellular identity of *sri-9* expression (as well as other expression changes) was corroborated by examining overlap of GPCR *gfp*-based reporters with *rfp*-based landmark strain (see Materials and methods).

Some csGPCRs serve as molecular markers of life history

Do reporter expression changes observed in dauers recover upon re-feeding to the pattern observed in control-fed animals? Examining csGPCR reporter expression in well-fed adult

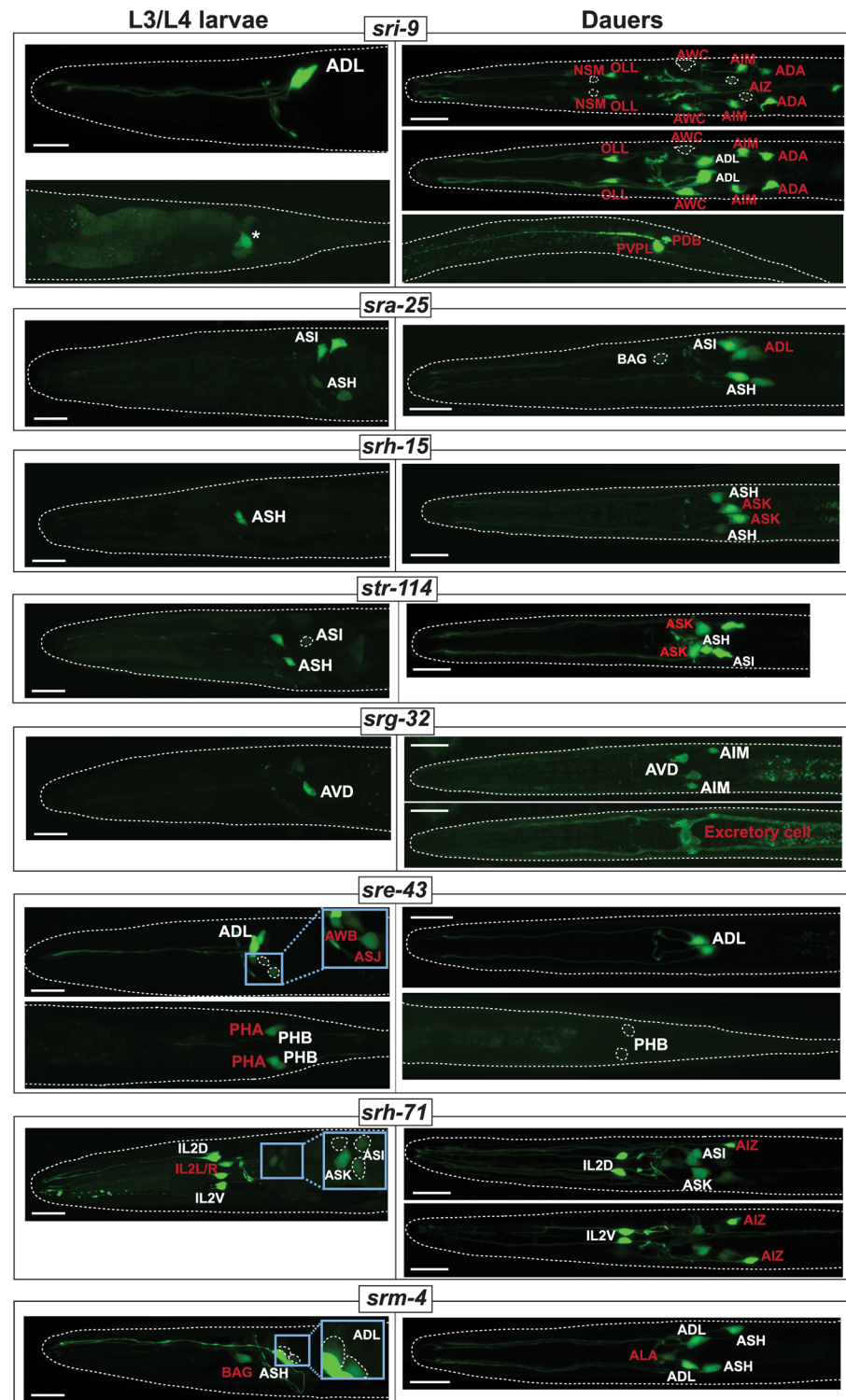


Fig 10. Examples of environment-induced changes in csGPCR expression. Examples of GPCR reporters that change expression in dauer stage animals. Designations of neuron classes that change expression are highlighted in red. Asterisk indicates posterior gut autofluorescence. Insets for *srh-71*, *sre-43*, and *srm-4* show enlarged and overexposed images of cells that are too dim to be discernable in main panels. See Table 5 for a complete summary of GPCR expression changes in dauer. Scale bars, 10 μ m. GPCR, G-protein-coupled receptor; L3, third larval stage; L4, fourth larval stage.

<https://doi.org/10.1371/journal.pbio.2004218.g010>

Table 5. Changes in csGPCR reporter expression in starvation-induced dauers, within and outside the nervous system. Reporter gene expression patterns were analyzed in starvation-induced dauers. Previously reported GPCR reporter changes are listed in the two bottom rows of the Table [34]. For the *srh-71* reporter, we also observe non-robust expression in a non-phasmid pair whose identity we have not determined.

| Type of change | Reporter Gene | Reporter expression | | | Postdauer recovery | | |
|----------------------------------|----------------|--|--|--|---|---|-------------------------------------|
| | | Constitutive expression in all stage (in dauer and non-dauer, fed L3 and post-dauer) | Cells only show expression in dauers | Constitutively expressed in fed animals only, i.e., down-regulated specifically in dauers in respective cell | Recovers | Dauer pattern is retained post-dauer | Entirely new post-dauer expression |
| dauer gains (type I) | <i>sri-9</i> | ADL | NSM (dim), OLL, AWC, AIM, AIZ, ADA, PDB, PVPL | | NSM, OLL, AWC, AIM, AIZ, ADA turn off again | PVPL remains on, PDB occasionally on | none |
| | <i>srh-15</i> | ASH, PHA | ASK | | recovers to fed condition | none | none |
| | <i>str-114</i> | ASH, ASI, PHA, head muscle | ASK, ASG | | recovers to fed condition | none | none |
| | <i>sra-25</i> | ASH, ASI, BAG (dim) | ADL, PHB | | PHB turn off again | ADL remains on, BAG becomes bright and stable | ASJ (dim) |
| | <i>str-84</i> | ASH, ASI, PHA, PHB, PVQ | Body wall muscle | | recovers to fed condition | none | none |
| | <i>srg-32</i> | AVD, AIM, PVQ | Excretory cell | | recovers to fed condition | none | none |
| dauer losses (type II) | <i>sre-43</i> | ADL, PHB (dim in dauers) | | AWB, ASJ (variable), PHA | AWB, PHA turn on again | ASJ become stable | ASH (dim) |
| | <i>srh-279</i> | ADL | | ADL down (but not off) in dauers | recovers to fed condition | none | none |
| | <i>srz-67</i> | ADL | | ADL down (but not off) in dauers | recovers to fed condition | none | none |
| | <i>srx-12</i> | | | ADF, amphid sheath glia | recovers to fed condition | none | none |
| | <i>sra-7</i> | | | ASK down (but not off) in dauers. F and U rectal epithelial cells off | recovers to fed condition | none | none |
| both gains and losses (type III) | <i>srm-4</i> | ASH, PHB, ADL (dim) | ADL(bright), ALA | BAG | recovers to fed condition | none | none |
| | <i>srx-4</i> | ASK, ASI | ADL | B and Y rectal epithelial cells | recovers to fed condition | ADL expression partially remains | none |
| | <i>srh-71</i> | ASK, ASG, ASI (dim), IL2D/V | ASI (brighter), AIZ, two ventral ganglion MN pairs | IL2L/R, PHA | PHA expression comes back to fed state | ASI, AIZ, two ventral ganglion MN pairs remain on, IL2L/R remains off | pharyngeal gland cells (ventral g1) |
| | <i>srsx-29</i> | ADF | ASH | PHA | recovers to fed condition | none | none |
| | <i>sru-12</i> | ASI, ASH, ASJ, OLL, PHB | PVQ | IL2, PHA | IL2, PHA turn on again | PVQ remains on | PLN |
| Peckol et al. 2001 | <i>srd-1</i> | | | ASI | recovers to fed condition | none | none |
| | <i>str-2</i> | ASI (dim) | ASI (brighter) | AWC | recovers to fed condition | none | none |

Abbreviations: G-protein-coupled receptor; L3, third larval stage.

<https://doi.org/10.1371/journal.pbio.2004218.t005>

animals that had passed through the dauer arrest stage during larval development, we found that the expression of 11 of the 18 reporters, which showed dauer-specific gene expression changes, recovers to that of the fed state, i.e., in these 11 cases, expression in the adult is independent on whether the animals had passed through the dauer stage or not.

For 7 csGPCR reporters we discovered intriguing, cell-type-specific alterations in animals that have passed through the dauer stage (Fig 11, Table 5). We observed three types of changes:

1. For 4 reporters (*sri-9*, *sra-25*, *srh-71*, and *sru-12*), we observed that expression, which was induced in specific neuron classes exclusively in dauers, was retained this in post-dauer animals: dauer-induced expression of the *sri-9* reporter in PVPL, of the *sra-25* reporter in ADL, of the *sru-12* reporter in PVQ as well as of the *srh-71* reporter in AIZ and two ventral ganglion head motor neuron pairs is retained in post-dauer adults. In contrast, dauer-induced loss of *srh-71* expression in the lateral IL2 pair does not recover.
2. In 4 cases (*sre-43*, *srh-71*, *sru-12*, *sra-25* reporters), we observed induction of expression in additional cell types exclusively in post-dauer animals. *sru-12* reporter expression is specifically induced in the PLN neurons of post-dauer animals, *sre-43* expression is dimly observed in the ASH neurons of post-dauer animals, *sra-25* expression is dimly observed in the ASJ neurons in post-dauer animals, and *srh-71* reporter expression was induced in a non-neuronal pair, pharyngeal gland cells, in post-dauers.
3. We found two instances in which a sporadic and weak expression observed in animals that have not passed through the dauer stage will become highly penetrant and stable if they have passed through the dauer state (*sra-25* in BAG neurons, *sre-43* in ASJ, *srh-71* in ASI).

Note that all of the reporters for which we observe changes in post-dauer recovery do recover their “fed patterns” in other neuron classes (these could be considered as internal controls that argue against the changes in expression being a reporter gene artefact). Taken together, adult animals show neuron-class specific differences in the expression of csGPCR reporters depending on whether they have passed through periods of distress. csGPCR reporters therefore serve as reporters of life history traits.

L1 starvation recapitulates some but not all csGPCR reporter changes

We tested 5 of the 16 csGPCR reporters that displayed changes in the dauer stage for whether their expression also changes in another starvation-induced arrest stage, the starvation-induced L1 arrest stage. Comparing expression in 2 day-starved L1 (egg prep into M9 medium) to fed L1, we find that two reporters (*str-114* and *sra-25*) show the same changes as observed in dauer animals (Fig 12). In contrast, two reporters (*str-84* and *srg-32*) that change their expression in dauers, do not show changes in starved L1 versus fed L1 (Fig 12). One reporter, *srh-15*, in addition to dauer-specific expression in ASK, is also expressed in ASI in starved L1. Hence, the response of csGPCR expression to arrest conditions is diverse.

Discussion

Together with previously published analyses, there are now reporter transgenes that monitor the expression of 373 of the approximately 1,300 chemosensory GPCR genes encoded in the *C. elegans* genome. One intrinsic limitation of reporter genes is that they do not necessarily capture the full complement of *cis*-regulatory control elements of a gene. However, given the compact nature of csGPCR loci, the inclusion of all 5' regions in most reporters and the small

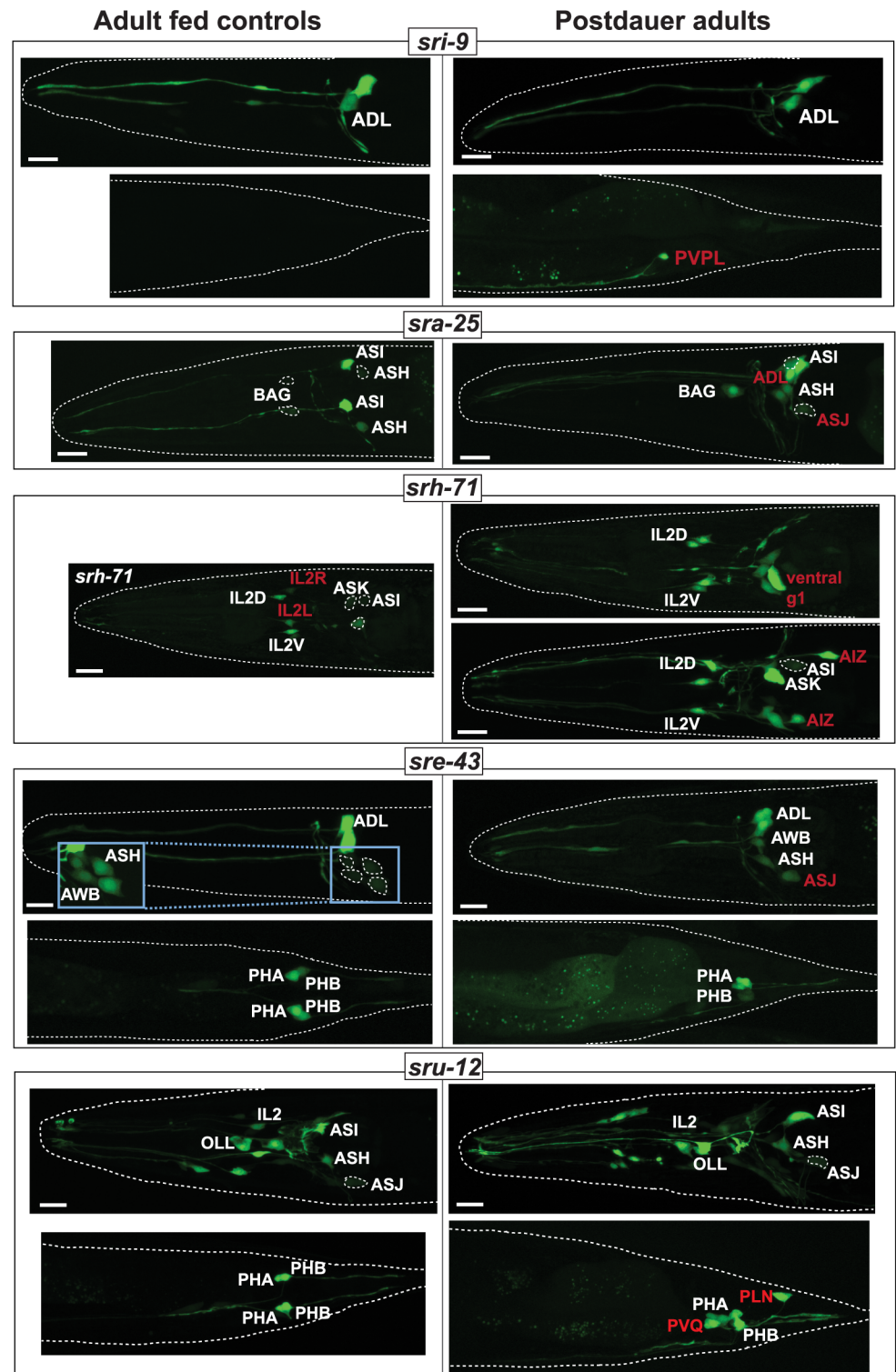


Fig 11. csGPCR expression patterns as life history traits. Comparison of GPCR expression in 1-day-old adult hermaphrodite animals that either did pass through the dauer state (right panels) or did not (age-matched fed controls; left panels). Post-dauer animals were in the dauer stage for 5–7 days. Designations of neuron classes that retain dauer-specific expression or acquire post-dauer-specific expression are highlighted in red. Inset for *sre-43* shows enlarged and overexposed images of cells that are too dim to be clearly discernable in the main panel. See Table 5 for a complete summary of GPCR expression changes in post-dauer. Scale bars, 10 μ m. GPCR, G-protein-coupled receptor.

<https://doi.org/10.1371/journal.pbio.2004218.g011>

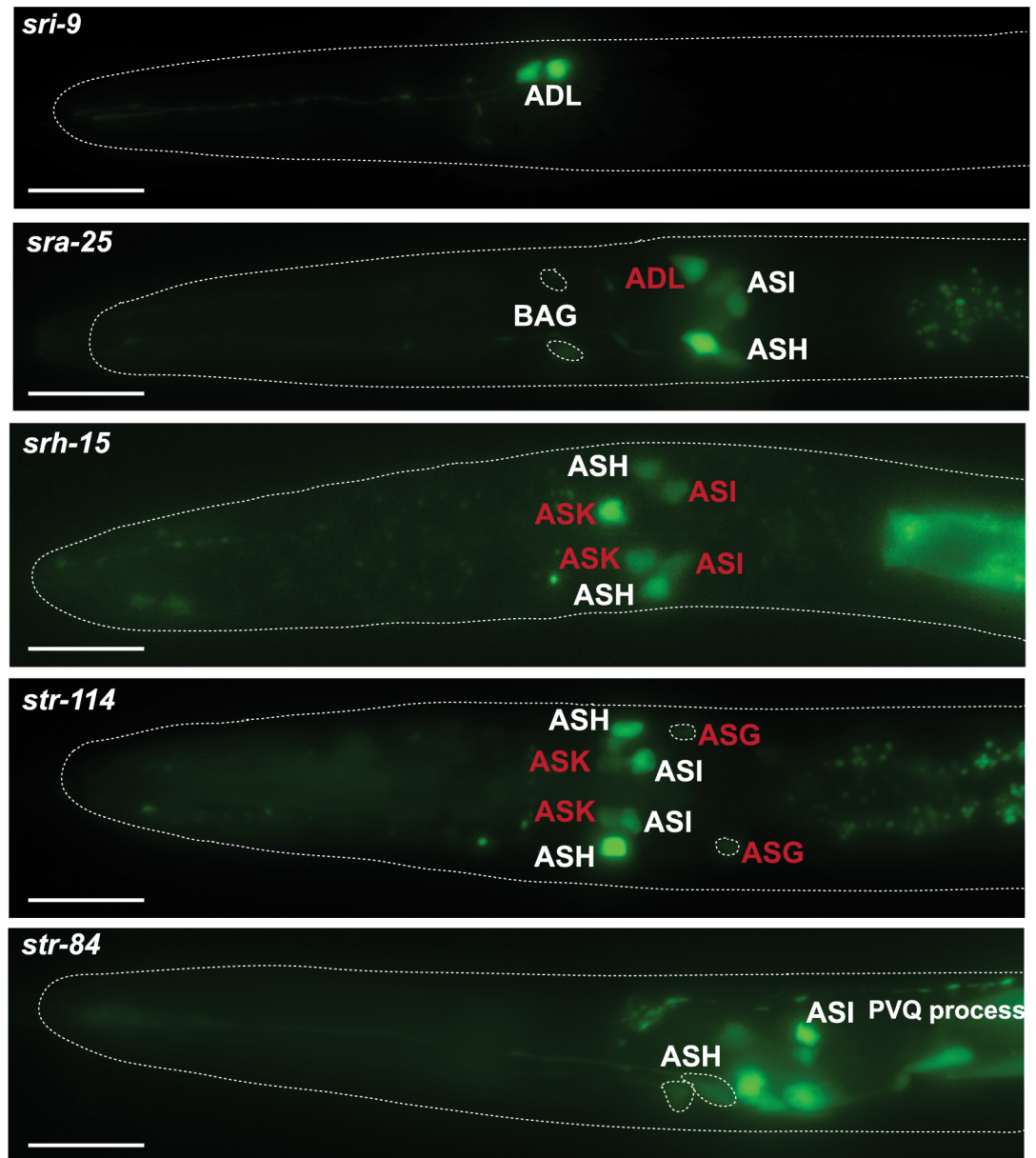


Fig 12. csGPCR reporter expression in starved L1 animals. Examples of GPCR reporter expression in starved L1 worms. Eggs isolated by bleach treatment were allowed to hatch and were kept in M9 for 48 hours. Designations of neuron classes that change expression compared to fed L1 worms are highlighted in red. Scale bars, 10 μ m. GPCR, G-protein-coupled receptor; L1, first larval stage.

<https://doi.org/10.1371/journal.pbio.2004218.g012>

size of introns, the number of inaccuracies may be quite limited. Irrespective of whether the reporters are a reflection of the complete expression of a csGPCR, they nevertheless function as highly valuable molecular markers of cellular identity and plasticity. Meaning, reporter gene analysis decodes *cis*-regulatory information and provides read-outs of regulatory states of specific cell types. The key conclusions of the expression patterns inferred from the reporter genes are as follows:

Restricted expression

Most csGPCRs show a very restricted expression in few cell types. Many are expressed in single neuron classes. Those csGPCRs that express in multiple neuron classes do not display a coherent set of coexpressing neurons, with one notable exception: the nociceptive ASH, PHA, and PHB neurons express similar (but not identical) sets of csGPCRs.

csGPCR coexpression within a neuron class

Some neurons coexpress a remarkably large number of GPCRs. The ASI neuron displays the most csGPCR genes at 99, followed by many distinct types of nociceptive neurons. While csGPCRs have been found for all but two sensory neurons (URY and ADE), there is a striking disparity in the number of csGPCRs coexpressed in sensory neuron classes. Amphid sensory neurons clearly coexpress the largest number of csGPCRs, while other sensory neurons express many fewer csGPCRs. The nociceptive ADL stands out in the list of amphid neurons, as it is the neuron expressing the most single-neuron-specific csGPCR reporters.

Expression in sensory and non-sensory neuron classes

While expression of csGPCRs clearly predominates in sensory neurons, they are also expressed in inter- and motorneurons and in a diverse set of non-neuronal cells. In most cases, each GPCR is restrictively expressed, suggesting that many different cell types in an organism show very distinct and cell-type-specific responses, possibly to internal signals. The similarity of one GPCR family, the *srw* family, to peptide receptors of other animal species provides hints to the nature of these ligands [11,29]. The expression of many members of the *srr* family in pharyngeal tissues suggests another source of ligands; perhaps these receptors respond to cues from ingested bacteria. In vertebrates, chemosensory GPCRs are now also becoming increasingly appreciated as being expressed in non-neuronal cells [6].

Polymodality of sensory neurons

csGPCRs were detected in sensory neurons that are known to express distinct types of sensory receptors and engage in non-chemosensory behavior, e.g., in gas-sensing neurons or different types of mechanosensory neurons. The expression of csGPCRs in these neuron classes may hint toward these neurons perceiving different sensory inputs, i.e., they are likely polymodal. However, as discussed above, csGPCRs expressed in these neurons may not be involved in detecting external sensory cues, but measuring internal states.

Absence of gene family themes

The absence of any overarching expression theme within gene families is striking. We did not observe that the expression of family members clusters in specific neuron classes or share any other specific expression features. Specifically: (a) left/right asymmetrically-expressed csGPCRs in the AWC neurons do not fall into the same family; (b) csGPCR reporters that are differentially regulated in larval stages or in the dauer stage do not come from a single family; (c) csGPCRs that share specific expression pattern themes (e.g., coexpression in the nociceptive ASH, PHA, and PHB neurons) do not derive from specific families; (d) non-sensory neuron-expressed or non-neuronal expressed csGPCRs do not fall into a specific family. The only glimpse of perhaps some common function is observed in the small *srr* family (nine genes), half of which appear to be expressed in non-neuronal pharyngeal tissue. An important note of caution is that these conclusions are based only on reporter genes. However, the substantial sample size on which these conclusions are based lends some credence to these conclusions.

Combinatorial complexity

csGPCRs generally act as homo- or heterodimers [71], thereby hugely increasing the amount of distinct sensory receptor complexes expressed in a cell. This combinatorial activity also makes prediction of function of a given csGPCR very difficult in that a csGPCR may have one function expressed in one cell (in combination with another csGPCR), while it may have a very different function in another cell (in combination with yet another csGPCR).

Left/right asymmetric csGPCR expression patterns

While we recovered novel csGPCR genes expressed in a left/right asymmetric manner in the AWC neuron pair, we were surprised to find no other obvious left/right asymmetries in other sensory neuron pairs. Of course, such asymmetries may still be found with currently not analyzed csGPCR genes, but the number of AWC asymmetries recovered suggests that AWC neurons may be exceptional in their extent of lateralization.

The only other asymmetry that we found revealed itself not in a sensory, but an interneuron, and only in a non-anticipated context. The *sri-9* reporter transgene becomes induced in dauer animals in PVPL, but not PVPR, and PVPL expression is retained in postdauer animals. Molecular asymmetries in PVP neurons have not previously been reported but can perhaps be inferred by the fact that PVPL and PVPR are innervated in a left/right asymmetric manner by unilateral neurons. Specifically, PVPL, but not PVPR, is innervated by the unilateral DVB neuron. Perhaps *sri-9* may play a role in this synaptic signaling context, but why this should be dauer-specific is unclear.

Plasticity of csGPCR expression

One notable feature of our analysis was the extent of plasticity that csGPCR reporters show in the context of the dauer stage. Dauer animals are thought to remodel most tissue types and significantly alter behavioral patterns. Changes in csGPCR expression, and hence changes in the external and internal signal perception, fit very well into the mold of organismal plasticity and illustrate the plasticity of many different tissue types (note, for example, the changes in csGPCR expression in muscle). We find it particularly intriguing that several csGPCRs represent markers of life history. Some of the changes in csGPCR reporter gene expression in dauers is retained in post-dauer animals and some csGPCR reporters turn on only in postdauer animals. Animal-wide expression transcriptomic analysis has previously identified large cohorts of transcripts that, like our csGPCR reporters, serve as markers of dauer life history, i.e., transcripts change in dauers and some of these transcript changes persist in post-dauer animals [72]. However, due to the whole animal nature of this analysis, this previous study lacked cellular resolution. Our findings add a novel spatial component to these previous findings, since we find the life history traits to be strikingly neuron class-specific. The expression of the TRP channel gene *osm-9* has also previously been shown to be modulated during dauer and post-dauer stages in a neuron class-specific manner; in this case, *osm-9* expression is down-regulated in the ADL (but not AWA) chemosensory neurons and the repression is retained post-dauer, using RNAi and chromatin-based mechanisms [73]. In all except one case that we report here, we observe the opposite post-dauer effect; reporters that are turned on in dauers persist in non-dauers. The mechanistic basis of this may hence be distinct from the *osm-9* case.

It is important to remember here that the life history trait observations are based on transcriptional reporter genes which, on the one hand, may not accurately reflect expression of the endogenous locus, but, on the other hand, clearly provide a definitive molecular “read out” of changes in the “regulatory state” of specific neuron types, depending on whether they have passed through the dauer stage or not. Moreover, our transcriptional reporters also argue that

the life history regulatory phenomenon must be transcriptional in nature. These csGPCR reporters will therefore provide excellent starting points to analyze the molecular mechanisms controlling this plasticity.

Future uses of the csGPCR expression map

The csGPCR reporter atlas can be put to a number of future uses. The sites of expression of specific csGPCRs point to potential functions of the csGPCRs, guiding the future analysis of csGPCR knockout strains. For example, csGPCRs expressed in the polymodal nociceptive ASH, ADL and phasmid neurons may be mediating the response to a number of distinct sensory cues processed by these neurons [56,74].

csGPCR expression patterns point to perhaps unexpected cellular sites of internal signal perception that warrant further investigation. For example, the excretory canal cell expresses at least six csGPCRs reporters (considering that we only examined reporters for approximately 20% of GPCR loci, this number may increase several fold). The relevance of this expression could be tested through the excretory cell-specific expression of dominant negative versions of G-protein downstream signaling components. Similarly, the cellular dynamics in csGPCR expression patterns point to specific cells undergoing changes that warrant future characterization. For example, the induction or suppression of csGPCR reporter expression during the dauer stage in specific sensory and interneurons that were not previously associated with dauer-specific functions may warrant a closer examination of other molecular and functional changes of these neurons during the dauer stage.

Because csGPCR reporter fusions also link precisely delineated sequences (used for reporter construction) to specific cellular sites of gene expression, patterns of coexpression of GPCRs can be used to extract *cis*-regulatory information, which in turn may point to *trans*-acting factors involved in controlling GPCR gene expression. A proof of principle for this type of analysis has already been conducted, pointing to a critical function of, for example, a basic helix-loop-helix (bHLH) factor in controlling csGPCR expression in the ADL nociceptive neuron [28], and with now substantially more expression information available can be further extended to additional cell types.

Lastly, green fluorescent protein (GFP) reporter transgenes have generally served as invaluable starting points for genetic mutant screens in which the genetic control of specific biological processes can be investigated. The csGPCR reporter collection provides a multitude of entry points. For example, the post-dauer expression of multiple reporter genes can be used to screen for mutants in which these life history traits fail to be properly expressed. GFP reporter genes have also served as invaluable cellular identity markers and here again the csGPCR reporter collection can be used to assess how the identity of specific cell types is genetically controlled.

Supporting information

S1 Fig. csGPCR gene locus analysis. (A) Histogram of upstream intergenic region distances of all *C. elegans* csGPCR genes.

The average size of the 5' intergenic region (= distance to next gene) is 1.8 kb.

Eighty-nine percent of all loci have a 5' intergenic region smaller than 4 kb.

(B) Histogram of average combined intron length (bp) per GPCR gene.

(C) The intergenic region of the majority of GPCR is substantially larger than the combined intronic region. bp, base pair; csGPCR, chemosensory-type GPCRs; GPCR, G-protein-coupled receptor.

(TIF)

S1 Table. Masterlist of all examined GPCR reporters. GPCR, G-protein-coupled receptor protein.
(XLSX)

S2 Table. List of all identified sensory neurons with GPCR expression. Gene in bold: newly identified in this paper
Gene in non-bold: previously identified
(Gene in parenthesis): ID based on position and morphology, not confirmed with neuron-specific reporter.
GPCR, G-protein-coupled receptor protein.
(XLSX)

S3 Table. Primers. Primer sequences for the reporters generated by the Vancouver consortium (BC strains) can be found at <http://www.gfpworm.org>.
(XLSX)

Acknowledgments

We thank Q. Chen for generating transgenic strains, the Vancouver gene expression consortium for generating strains, and the CGC for providing these strains to us. We thank Neda Masoudi and Sarah Finkelstein for communicating unpublished results.

Author Contributions

Conceptualization: Berta Vidal, Oliver Hobert.

Data curation: Berta Vidal.

Formal analysis: Berta Vidal.

Funding acquisition: Oliver Hobert.

Investigation: Berta Vidal, Ulkar Aghayeva, Haosheng Sun, Chen Wang, Lori Glenwinkel, Emily A. Bayer.

Methodology: Berta Vidal.

Supervision: Oliver Hobert.

Visualization: Berta Vidal.

Writing – original draft: Oliver Hobert.

Writing – review & editing: Berta Vidal, Ulkar Aghayeva, Oliver Hobert.

References

1. Hobert O, Carrera I, Stefanakis N. The molecular and gene regulatory signature of a neuron. *Trends in neurosciences*. 2010; 33(10):435–45. Epub 2010/07/29. <https://doi.org/10.1016/j.tins.2010.05.006> PMID: 20663572.
2. Troemel ER, Chou JH, Dwyer ND, Colbert HA, Bargmann CI. Divergent seven transmembrane receptors are candidate chemosensory receptors in *C. elegans*. *Cell*. 1995; 83(2):207–18. PMID: 7585938
3. Chalfie M, Tu Y, Euskirchen G, Ward WW, Prasher DC. Green fluorescent protein as a marker for gene expression. *Science*. 1994; 263(5148):802–5. PMID: 8303295.
4. Lagerstrom MC, Schioth HB. Structural diversity of G protein-coupled receptors and significance for drug discovery. *Nat Rev Drug Discov*. 2008; 7(4):339–57. Epub 2008/04/03. <https://doi.org/10.1038/nrd2518> PMID: 18382464.

5. Fredriksson R, Lagerstrom MC, Lundin LG, Schiöth HB. The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Mol Pharmacol*. 2003; 63(6):1256–72. <https://doi.org/10.1124/mol.63.6.1256> PMID: 12761335.
6. Foster SR, Roura E, Thomas WG. Extrasensory perception: odorant and taste receptors beyond the nose and mouth. *Pharmacol Ther*. 2014; 142(1):41–61. <https://doi.org/10.1016/j.pharmthera.2013.11.004> PMID: 24280065.
7. Thomas JH, Robertson HM. The *Caenorhabditis* chemoreceptor gene families. *BMC Biol*. 2008; 6:42. <https://doi.org/10.1186/1741-7007-6-42> PMID: 18837995.
8. The *C.elegans*-Sequencing-Consortium. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science*. 1998; 282(5396):2012–8. PMID: 9851916
9. White JG, Southgate E, Thomson JN, Brenner S. The structure of the nervous system of the nematode *Caenorhabditis elegans*. *Philosophical Transactions of the Royal Society of London B Biological Sciences*. 1986; 314:1–340. PMID: 22462104
10. Hobert O. The neuronal genome of *Caenorhabditis elegans*. *WormBook*. 2013:1–106. <https://doi.org/10.1895/wormbook.1.161.1> PMID: 24081909.
11. Robertson HM, Thomas JH. The putative chemoreceptor families of *C. elegans*. *WormBook*. 2006:1–12. Epub 2007/12/01. <https://doi.org/10.1895/wormbook.1.66.1> PMID: 18050473.
12. Taniguchi G, Uozumi T, Kiriya K, Kamizaki T, Hirotsu T. Screening of odor-receptor pairs in *Caenorhabditis elegans* reveals different receptors for high and low odor concentrations. *Sci Signal*. 2014; 7(323):ra39. <https://doi.org/10.1126/scisignal.2005136> PMID: 24782565.
13. Sengupta P, Chou JH, Bargmann CI. odr-10 encodes a seven transmembrane domain olfactory receptor required for responses to the odorant diacetyl. *Cell*. 1996; 84(6):899–909. Epub 1996/03/22. PMID: 8601313.
14. Zhang Y, Chou JH, Bradley J, Bargmann CI, Zinn K. The *Caenorhabditis elegans* seven-transmembrane protein ODR-10 functions as an odorant receptor in mammalian cells. *Proc Natl Acad Sci U S A*. 1997; 94(22):12162–7. PMID: 9342380.
15. Park D, O'Doherty I, Somvanshi RK, Bethke A, Schroeder FC, Kumar U, et al. Interaction of structure-specific and promiscuous G-protein-coupled receptors mediates small-molecule signaling in *Caenorhabditis elegans*. *Proc Natl Acad Sci U S A*. 2012. Epub 2012/06/06. <https://doi.org/10.1073/pnas.1202216109> PMID: 22665789.
16. McGrath PT, Xu Y, Ailion M, Garrison JL, Butcher RA, Bargmann CI. Parallel evolution of domesticated *Caenorhabditis* species targets pheromone receptor genes. *Nature*. 2011; 477(7364):321–5. Epub 2011/08/19. <https://doi.org/10.1038/nature10378> PMID: 21849976.
17. Kim K, Sato K, Shibuya M, Zeiger DM, Butcher RA, Ragains JR, et al. Two chemoreceptors mediate developmental effects of dauer pheromone in *C. elegans*. *Science*. 2009; 326(5955):994–8. Epub 2009/10/03. <https://doi.org/10.1126/science.1176331> PMID: 19797623.
18. Greene JS, Dobosiewicz M, Butcher RA, McGrath PT, Bargmann CI. Regulatory changes in two chemoreceptor genes contribute to a *Caenorhabditis elegans* QTL for foraging behavior. *eLife*. 2016; 5. <https://doi.org/10.7554/eLife.21454> PMID: 27893361.
19. Greene JS, Brown M, Dobosiewicz M, Ishida IG, Macosko EZ, Zhang X, et al. Balancing selection shapes density-dependent foraging behaviour. *Nature*. 2016; 539(7628):254–8. <https://doi.org/10.1038/nature19848> PMID: 27799655.
20. Zhang C, Zhao N, Chen Y, Zhang D, Yan J, Zou W, et al. The Signaling Pathway of *Caenorhabditis elegans* Mediates Chemotaxis Response to the Attractant 2-Heptanone in a Trojan Horse-like Pathogenesis. *J Biol Chem*. 2016; 291(45):23618–27. <https://doi.org/10.1074/jbc.M116.741132> PMID: 27660389.
21. Aoki R, Yagami T, Sasakura H, Ogura K, Kajihara Y, Ibi M, et al. A seven-transmembrane receptor that mediates avoidance response to dihydrocaffeic acid, a water-soluble repellent in *Caenorhabditis elegans*. *J Neurosci*. 2011; 31(46):16603–10. <https://doi.org/10.1523/JNEUROSCI.4018-11.2011> PMID: 22090488.
22. Peterlin Z, Firestein S, Rogers ME. The state of the art of odorant receptor deorphanization: a report from the orphanage. *J Gen Physiol*. 2014; 143(5):527–42. <https://doi.org/10.1085/jgp.201311151> PMID: 24733839.
23. Montagne N, de Fouchier A, Newcomb RD, Jacquín-Joly E. Advances in the identification and characterization of olfactory receptors in insects. *Prog Mol Biol Transl Sci*. 2015; 130:55–80. <https://doi.org/10.1016/bs.pmbts.2014.11.003> PMID: 25623337.
24. Chen N, Pai S, Zhao Z, Mah A, Newbury R, Johnsen RC, et al. Identification of a nematode chemosensory gene family. *Proc Natl Acad Sci U S A*. 2005; 102(1):146–51. <https://doi.org/10.1073/pnas.0408307102> PMID: 15618405.

25. Wenick AS, Hobert O. Genomic cis-Regulatory Architecture and trans-Acting Regulators of a Single Interneuron-Specific Gene Battery in *C. elegans*. *Dev Cell*. 2004; 6(6):757–70. <https://doi.org/10.1016/j.devcel.2004.05.004> PMID: 15177025.
26. Battu G, Hoier EF, Hajnal A. The *C. elegans* G-protein-coupled receptor SRA-13 inhibits RAS/MAPK signalling during olfaction and vulval development. *Development*. 2003; 130(12):2567–77. PMID: 12736202.
27. Colosimo ME, Brown A, Mukhopadhyay S, Gabel C, Lanjuin AE, Samuel AD, et al. Identification of thermosensory and olfactory neuron-specific genes via expression profiling of single neuron types. *Curr Biol*. 2004; 14(24):2245–51. <https://doi.org/10.1016/j.cub.2004.12.030> PMID: 15620651.
28. McCarroll SA, Li H, Bargmann CI. Identification of transcriptional regulatory elements in chemosensory receptor genes by probabilistic segmentation. *Curr Biol*. 2005; 15(4):347–52. <https://doi.org/10.1016/j.cub.2005.02.023> PMID: 15723796.
29. Krishnan A, Almen MS, Fredriksson R, Schiöth HB. Insights into the origin of nematode chemosensory GPCRs: putative orthologs of the *Strv* family are found across several phyla of protostomes. *PLoS ONE*. 2014; 9(3):e93048. <https://doi.org/10.1371/journal.pone.0093048> PMID: 24663674.
30. Remy JJ, Hobert O. An interneuronal chemoreceptor required for olfactory imprinting in *C. elegans*. *Science*. 2005; 309(5735):787–90. <https://doi.org/10.1126/science.1114209> PMID: 16051801.
31. Troemel ER, Sagasti A, Bargmann CI. Lateral signaling mediated by axon contact and calcium entry regulates asymmetric odorant receptor expression in *C. elegans*. *Cell*. 1999; 99(4):387–98. PMID: 10571181
32. Wes PD, Bargmann CI. *C. elegans* odour discrimination requires asymmetric diversity in olfactory neurons. *Nature*. 2001; 410(6829):698–701. <https://doi.org/10.1038/35070581> PMID: 11287957.
33. Hsieh YW, Alqadah A, Chuang CF. Asymmetric neural development in the *Caenorhabditis elegans* olfactory system. *Genesis*. 2014; 52(6):544–54. <https://doi.org/10.1002/dvg.22744> PMID: 24478264.
34. Peckol EL, Troemel ER, Bargmann CI. Sensory experience and sensory activity regulate chemosensory receptor gene expression in *Caenorhabditis elegans*. *Proc Natl Acad Sci U S A*. 2001; 98(20):11032–8. <https://doi.org/10.1073/pnas.191352498> PMID: 11572964
35. Gruner M, Nelson D, Winbush A, Hintz R, Ryu L, Chung SH, et al. Feeding state, insulin and NPR-1 modulate chemoreceptor gene expression via integration of sensory and circuit inputs. *PLoS Genet*. 2014; 10(10):e1004707. <https://doi.org/10.1371/journal.pgen.1004707> PMID: 25357003.
36. Nolan KM, Sarafi-Reinach TR, Horne JG, Saffer AM, Sengupta P. The DAF-7 TGF-beta signaling pathway regulates chemosensory receptor gene expression in *C. elegans*. *Genes Dev*. 2002; 16(23):3061–73. <https://doi.org/10.1101/gad.1027702> PMID: 12464635.
37. Ryan DA, Miller RM, Lee K, Neal SJ, Fagan KA, Sengupta P, et al. Sex, age, and hunger regulate behavioral prioritization through dynamic modulation of chemoreceptor expression. *Curr Biol*. 2014; 24(21):2509–17. <https://doi.org/10.1016/j.cub.2014.09.032> PMID: 25438941.
38. Portman DS. Genetic control of sex differences in *C. elegans* neurobiology and behavior. *Adv Genet*. 2007; 59:1–37. [https://doi.org/10.1016/S0065-2660\(07\)59001-2](https://doi.org/10.1016/S0065-2660(07)59001-2) PMID: 17888793.
39. Brenner S. The genetics of *Caenorhabditis elegans*. *Genetics*. 1974; 77(1):71–94. PMID: 4366476
40. Granato M, Schnabel H, Schnabel R. pha-1, a selectable marker for gene transfer in *C. elegans*. *Nucleic Acids Res*. 1994; 22(9):1762–3. PMID: 8202383.
41. Hodgkin J, Horvitz HR, Brenner S. Nondisjunction Mutants of the Nematode *CAENORHABDITIS ELEGANS*. *Genetics*. 1979; 91(1):67–94. PMID: 17248881.
42. Reiner DJ, Newton EM, Tian H, Thomas JH. Diverse behavioural defects caused by mutations in *Caenorhabditis elegans* unc-43 CaM kinase II. *Nature*. 1999; 402(6758):199–203. <https://doi.org/10.1038/46072> PMID: 10647014.
43. Park EC, Horvitz HR. *C. elegans* unc-105 mutations affect muscle and are suppressed by other mutations that affect muscle. *Genetics*. 1986; 113(4):853–67. PMID: 3744029.
44. Chuang CF, Vanhoven MK, Fetter RD, Verselis VK, Bargmann CI. An innexin-dependent cell network establishes left-right neuronal asymmetry in *C. elegans*. *Cell*. 2007; 129(4):787–99. Epub 2007/05/22. <https://doi.org/10.1016/j.cell.2007.02.052> PMID: 17512411.
45. Hobert O. PCR fusion-based approach to create reporter gene constructs for expression analysis in transgenic *C. elegans*. *BioTechniques*. 2002; 32(4):728–30. PMID: 11962590.
46. Hunt-Newbury R, Viveiros R, Johnsen R, Mah A, Anastas D, Fang L, et al. High-throughput in vivo analysis of gene expression in *Caenorhabditis elegans*. *PLoS Biol*. 2007; 5(9):e237. Epub 2007/09/14. <https://doi.org/10.1371/journal.pbio.0050237> PMID: 17850180.

47. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. *Nat Methods*. 2012; 9(7):676–82. <https://doi.org/10.1038/nmeth.2019> PMID: 22743772.
48. Chang C, Hsieh YW, Lesch BJ, Bargmann CI, Chuang CF. Microtubule-based localization of a synaptic calcium-signaling complex is required for left-right neuronal asymmetry in *C. elegans*. *Development*. 2011; 138(16):3509–18. <https://doi.org/10.1242/dev.069740> PMID: 21771813.
49. Serrano-Saiz E, Oren-Suissa M, Bayer EA, Hobert O. Sexually Dimorphic Differentiation of a *C. elegans* Hub Neuron Is Cell Autonomously Controlled by a Conserved Transcription Factor. *Curr Biol*. 2017; 27(2):199–209. <https://doi.org/10.1016/j.cub.2016.11.045> PMID: 28065609.
50. Pereira L, Kratsios P, Serrano-Saiz E, Sheftel H, Mayo AE, Hall DH, et al. A cellular and regulatory map of the cholinergic nervous system of *C. elegans*. *eLife*. 2015; 4. <https://doi.org/10.7554/eLife.12432> PMID: 26705699.
51. Gendrel M, Atlas EG, Hobert O. A cellular and regulatory map of the GABAergic nervous system of *C. elegans*. *eLife*. 2016; 5. <https://doi.org/10.7554/eLife.17686> PMID: 27740909.
52. Howe KL, Bolt BJ, Cain S, Chan J, Chen WJ, Davis P, et al. WormBase 2016: expanding to enable helminth genomic research. *Nucleic Acids Res*. 2016; 44(D1):D774–80. <https://doi.org/10.1093/nar/gkv1217> PMID: 26578572.
53. Shimodaira H. Approximately unbiased tests of regions using multistep-multiscale bootstrap resampling. *The Annals of Statistics*. 2004; 32(6):2616–41.
54. Boulin T, Etchberger JF, Hobert O. Reporter gene fusions. *WormBook*. 2006:1–23. Epub 2007/12/01. <https://doi.org/10.1895/wormbook.1.106.1> PMID: 18050449.
55. Hilliard MA, Bargmann CI, Bazzicalupo P. *C. elegans* responds to chemical repellents by integrating sensory inputs from the head and the tail. *Curr Biol*. 2002; 12(9):730–4. PMID: 12007416.
56. Zou W, Cheng H, Li S, Yue X, Xue Y, Chen S, et al. Polymodal Responses in *C. elegans* Phasmid Neurons Rely on Multiple Intracellular and Intercellular Signaling Pathways. *Scientific reports*. 2017; 7:42295. <https://doi.org/10.1038/srep42295> PMID: 28195191.
57. Micallef L, Rodgers P. eulerAPE: drawing area-proportional 3-Venn diagrams using ellipses. *PLoS ONE*. 2014; 9(7):e101717. <https://doi.org/10.1371/journal.pone.0101717> PMID: 25032825.
58. Hobert O, Johnston RJ Jr., Chang S. Left-right asymmetry in the nervous system: the *Caenorhabditis elegans* model. *Nat Rev Neurosci*. 2002; 3(8):629–40. <https://doi.org/10.1038/nrn897> PMID: 12154364.
59. Yu S, Avery L, Baude E, Garbers DL. Guanylyl cyclase expression in specific sensory neurons: a new family of chemosensory receptors. *Proc Natl Acad Sci U S A*. 1997; 94(7):3384–7. PMID: 9096403
60. Ortiz CO, Faumont S, Takayama J, Ahmed HK, Goldsmith AD, Pocock R, et al. Lateralized gustatory behavior of *C. elegans* is controlled by specific receptor-type guanylyl cyclases. *Curr Biol*. 2009; 19(12):996–1004. Epub 2009/06/16. <https://doi.org/10.1016/j.cub.2009.05.043> PMID: 19523832.
61. Pierce-Shimomura JT, Morse TM, Lockery SR. The fundamental role of pirouettes in *Caenorhabditis elegans* chemotaxis. *J Neurosci*. 1999; 19(21):9557–69. PMID: 10531458.
62. Lesch BJ, Gehrke AR, Bulyk ML, Bargmann CI. Transcriptional regulation and stabilization of left-right neuronal identity in *C. elegans*. *Genes Dev*. 2009; 23(3):345–58. Epub 2009/02/11. <https://doi.org/10.1101/gad.1763509> PMID: 19204119.
63. Cochella L, Tursun B, Hsieh YW, Galindo S, Johnston RJ, Chuang CF, et al. Two distinct types of neuronal asymmetries are controlled by the *Caenorhabditis elegans* zinc finger transcription factor *die-1*. *Genes Dev*. 2014; 28(1):34–43. <https://doi.org/10.1101/gad.233643.113> PMID: 24361693.
64. Emmons SW. The development of sexual dimorphism: studies of the *Caenorhabditis elegans* male. *Wiley interdisciplinary reviews Developmental biology*. 2014; 3(4):239–62. <https://doi.org/10.1002/wdev.136> PMID: 25262817.
65. Jarrell TA, Wang Y, Bloniarz AE, Brittin CA, Xu M, Thomson JN, et al. The connectome of a decision-making neural network. *Science*. 2012; 337(6093):437–44. Epub 2012/07/28. <https://doi.org/10.1126/science.1221762> PMID: 22837521.
66. Lipton J, Kleemann G, Ghosh R, Lints R, Emmons SW. Mate searching in *Caenorhabditis elegans*: a genetic model for sex drive in a simple invertebrate. *J Neurosci*. 2004; 24(34):7427–34. <https://doi.org/10.1523/JNEUROSCI.1746-04.2004> PMID: 15329389.
67. Oren-Suissa M, Bayer EA, Hobert O. Sex-specific pruning of neuronal synapses in *Caenorhabditis elegans*. *Nature*. 2016; 533(7602):206–11. <https://doi.org/10.1038/nature17977> PMID: 27144354.
68. Stefanakis N, Carrera I, Hobert O. Regulatory Logic of Pan-Neuronal Gene Expression in *C. elegans*. *Neuron*. 2015; 87(4):733–50. <https://doi.org/10.1016/j.neuron.2015.07.031> PMID: 26291158.

69. Cassada RC, Russell RL. The dauerlarva, a post-embryonic developmental variant of the nematode *Caenorhabditis elegans*. *Dev Biol.* 1975; 46(2):326–42. Epub 1975/10/01. PMID: [1183723](#).
70. Androwski RJ, Flatt KM, Schroeder NE. Phenotypic plasticity and remodeling in the stress-induced *Caenorhabditis elegans* dauer. *Wiley interdisciplinary reviews Developmental biology.* 2017. <https://doi.org/10.1002/wdev.278> PMID: [28544390](#).
71. Pierce KL, Premont RT, Lefkowitz RJ. Seven-transmembrane receptors. *Nat Rev Mol Cell Biol.* 2002; 3(9):639–50. <https://doi.org/10.1038/nrm908> PMID: [12209124](#).
72. Hall SE, Beverly M, Russ C, Nusbaum C, Sengupta P. A cellular memory of developmental history generates phenotypic diversity in *C. elegans*. *Curr Biol.* 2010; 20(2):149–55. Epub 2010/01/19. <https://doi.org/10.1016/j.cub.2009.11.035> PMID: [20079644](#).
73. Sims JR, Ow MC, Nishiguchi MA, Kim K, Sengupta P, Hall SE. Developmental programming modulates olfactory behavior in *C. elegans* via endogenous RNAi pathways. *eLife.* 2016; 5. <https://doi.org/10.7554/eLife.11642> PMID: [27351255](#).
74. Hilliard MA, Apicella AJ, Kerr R, Suzuki H, Bazzicalupo P, Schafer WR. In vivo imaging of *C. elegans* ASH neurons: cellular response and adaptation to chemical repellents. *EMBO J.* 2005; 24(1):63–72. <https://doi.org/10.1038/sj.emboj.7600493> PMID: [15577941](#).