Transcriptomic profiling of sex-specific olfactory neurons reveals subset-specific receptor expression in *Caenorhabditis elegans*

Author Names and Affiliations:

Douglas K. Reilly[1,7,†], Erich M. Schwarz[2,†], Caroline S. Muirhead[1], Annalise N. Robidoux[1,3], Anusha Narayan[4,5], Meenakshi K. Doma[4,6], Paul W. Sternberg[4], Jagan Srinivasan[1,8,*]

[1]Department of Biology and Biotechnology, Worcester Polytechnic Institute, Worcester, MA 01605, USA

[2]Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14853, USA

[3]Department of Chemistry and Biochemistry, Worcester Polytechnic Institute, Worcester, MA 01605, USA

[4]Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125, USA

[5]Present address: Merck and Co, 2000 Galloping Hill Road, Kenilworth, NJ 07033, USA

[6]Present address: Individualized Cell and Gene Therapies, South San Francisco, Genentech, CA 94080, USA

[7]Present address: Biology Department, Tufts University, Medford, MA 02155, USA

[8]Bioinformatics and Computational Biology Program, Worcester Polytechnic Institute, Worcester, MA 01605, USA

*Corresponding author: Jagan Srinivasan, *jsrinivasan@wpi.edu*

[†]These authors contributed equally to this work

Key words: *C. elegans*, Pheromone, Sex-specific, G Protein-Coupled Receptors, Transcriptomes, Single-Cell

Summary

The nematode *Caenorhabditis elegans* utilizes chemosensation to navigate an ever-changing environment for its survival. A class of secreted small-molecule pheromones, termed ascarosides, play an important role in olfactory perception by affecting biological functions ranging from development to behavior. The ascaroside ascr#8 mediates sex-specific behaviors, driving avoidance in hermaphrodites and attraction in males. Males sense ascr#8 via the ciliated male-specific cephalic sensory (CEM) neurons, which exhibit radial symmetry along dorsal-ventral and left-right axes. Calcium imaging studies suggest a complex neural coding mechanism that translates stochastic physiological responses in these neurons to reliable behavioral outputs. To test the hypothesis that neurophysiological complexity arises from differential expression of genes, we performed cell-specific transcriptomic profiling; this revealed between 18 and 62 genes with at least two-fold higher expression in a specific CEM neuron type versus both other CEM neurons and adult males. These included two G protein coupled receptor (GPCR) genes, *srw-97* and *dmsr-12*, that were specifically expressed in non-overlapping subsets of CEM neurons and whose expression was confirmed by GFP reporter analysis. Single CRISPR-Cas9 knockouts of either *srw-97* or *dmsr-12* resulted in partial defects, while a double knockout of both *srw-97* and *dmsr-12* completely abolished the attractive response to ascr#8. Together, our results suggest that the evolutionarily distinct GPCRs SRW-97 and DMSR-12 act non-redundantly in discrete olfactory neurons to facilitate male-specific sensation of ascr#8.

43 **Introduction**

44 The ability of an organism to find a mate is critical to the survival of a species. Many
45 species utilize small molecule pheromones to signal mate location (PUNGALIYA *et al.* 2009;
46 NARAYAN *et al.* 2016), sexual maturity (APRISON AND RUVINSKY 2015; APRISON AND RUVINSKY
47 2017), and receptivity (HOUCK *et al.* 2007; JANG *et al.* 2017), which are sensed and processed via
48 the nervous system, driving behavioral and developmental responses.

49 The nematode *Caenorhabditis elegans* communicates with conspecifics almost exclusively
50 using pheromones called ascarosides (LUDEWIG *et al.* 2019; MCGRATH AND RUVINSKY 2019), a
51 structurally conserved, modular class of small molecule pheromones (VON REUSS AND SCHROEDER
52 2015; ZHANG *et al.* 2017) composed of a core ascarylose sugar and a fatty-acid derived side chain
53 (BUTCHER *et al.* 2007; VON REUSS *et al.* 2012; LUDEWIG *et al.* 2019). Ascarosides signal a host of
54 environmental and developmental information, including the sexual maturity and location of
55 potential mates (NARAYAN *et al.* 2016; APRISON AND RUVINSKY 2017), and this information is
56 often conserved across nematodes (CHOE *et al.* 2012; RAGSDALE *et al.* 2013; DONG *et al.* 2016;
57 DONG *et al.* 2018; REILLY *et al.* 2019).

58 While most male-specific neurons are located in the tail, the major contributor to male-
59 specific chemosensory-driven behaviors are the CEM neurons, located in the amphid. Previous
60 studies have shown these four radially symmetric neurons to be involved in sensing ascaroside #3
61 and #8 (ascr#3 and ascr#8), two pheromones that serve to attract males to mates (PUNGALIYA *et*
62 *al.* 2009; NARAYAN *et al.* 2016; REILLY *et al.* 2017). Ascr#3 also plays a role in dauer formation,
63 functioning alongside ascr#2 and ascr#4 as the "dauer pheromone" (BUTCHER *et al.* 2008). Ascr#8
64 is unique in that it contains a *p*-aminobenzoic acid moiety on its terminus (PUNGALIYA *et al.* 2009;
65 ARTYUKHIN *et al.* 2018). In animals lacking CEM the attractive response to ascr#8 is abolished,

66 while further ablation of the ASK is required for complete loss of the ascr#3 behavioral response
67 (NARAYAN *et al.* 2016). Ablating three of the four CEM neurons, while leaving one CEM intact,
68 revealed that the four CEM neurons cooperate to drive a tuned attractive response to an
69 intermediate (1 μM) concentration of both ascr#3 and ascr#8 (NARAYAN *et al.* 2016).

70 Calcium imaging and electrophysiological experiments demonstrated that CEM neurons
71 show variable responses to stimuli not only between different animals, but within a single animal
72 (NARAYAN *et al.* 2016; REILLY *et al.* 2017). While neurons such as AWC exhibit variable calcium
73 responses between different animals, such responses are consistent between the left and right AWC
74 neurons of a single animal (COCHELLA *et al.* 2014). In contrast, the responses of individual CEM
75 neurons are stochastic within single animals, and yet four CEM neurons within one animal
76 consistently generate proper behavioral responses (NARAYAN *et al.* 2016). To understand this
77 pattern of stochastic neuronal activity yielding consistent behavioral outputs, it is imperative to
78 uncover genes encoding components of the CEM response.

79 A recent study in *C. elegans* hermaphrodites generated transcriptomic landscapes of 118
80 neuronal classes from 302 neurons in order to link functional and anatomical properties of
81 individual neurons with their molecular identities (TAYLOR *et al.* 2021). Discrete neuronal classes
82 were successfully identified via their combinations of expressed neuropeptides and neuropeptide
83 receptor genes. However, a similar feat has yet to be performed on male *C. elegans*, and the
84 transcriptomic profiles of individual male neurons remain enigmatic.

85 In a more focused transcriptomic approach, gene expression profiles of extracellular
86 vesicle-releasing neurons (EVNs) were identified (WANG *et al.* 2015; KALETSKY *et al.* 2016).
87 While CEM neurons are a subset of EVN neurons, and CEM transcriptomic data are embedded in

88 these data, there were thousands of EVNs in each replicate, diluting CEM-specific expression
89 values.

90 To begin understanding how CEM neurons achieve stochastic yet reliable physiological
91 responses to ascr#8, we performed single-cell RNA-seq (SCHWARZ *et al.* 2012) on CEM neurons,
92 and uncovered a small number of highly enriched genes encoding G protein-coupled receptors
93 (GPCRs). Given that all ascaroside receptors identified to date have been GPCRs (KIM *et al.* 2009;
94 MCGRATH *et al.* 2011; PARK *et al.* 2012; GREENE *et al.* 2016a; GREENE *et al.* 2016b; CHUTE *et al.*
95 2019), we tested whether any of these enriched genes contribute to male *C. elegans* sensation of
96 ascr#8. We identified two distantly related GPCR genes that contribute to the ascr#8 behavioral
97 response, *srw-97* and *dmsr-12*. Loss of each receptor resulted in partially deficient behavioral
98 responses to ascr#8. Following generation of a double *srw-97;dmsr-12* mutant, we found that loss
99 of both receptors results in a complete loss of behavioral response to ascr#8. Phylogenetic analysis
100 further indicates that both receptors are homologous to closely related receptors present throughout
101 the *Caenorhabditis* genus, and robust ascr#8 responses may be a trait recently evolved in
102 *C. elegans*.

103

104 **Results**

105 *The transcriptomic landscape of CEM neurons is variable*

106 Individual CEM neurons were isolated from *C. elegans* expressing an integrated GFP
107 labeling male-specific EVNs (p*pkd-2*::GFP [**Figure 2A**]), as previously described (GOODMAN *et*
108 *al.* 1998; NARAYAN *et al.* 2011; NARAYAN *et al.* 2016). Cells were separated by anatomical identity
109 (i.e., CEM dorsal left [DL], dorsal right [DR], ventral left [VL], and ventral right [VR]); single-
110 cell CEM cDNA libraries were constructed and used for RNA-seq (SCHWARZ *et al.* 2012).

111 Enriched genes in each CEM neuron were identified by comparing RNA-seq profiles of
112 distinct CEM types to one another and to whole adult *C. elegans* males (**Table S1, S5**). We
113 observed 267 genes with at least two-fold higher expression in CEM neurons than in adult males.
114 Of these, we observed 164 genes that were specific to individual CEM types, defining specificity
115 as being at least two-fold higher expression in one type than in any other type: 62 genes specific
116 to CEM DL, 40 to CEM DR, 18 to CEM VL, and 44 to CEM VR. Uniquely mapped reads ranged
117 from 5.2 to 9.2 million per CEM type, matching the trend for alignment rates of each CEM neuron,
118 which ranged from 18.0% to 47.2%, (**Table S2**) with an average of 1,426 genes showing robust
119 RNA expression in each neuron (**Table S3**).

120 Five GPCR-encoding genes expressed in CEMs (*seb-3, srr-7, srw-97, dmsr-12, srd-32*)
121 were selected for study; of these, *srw-97* and *dmsr-12* showed specific expression in CEM VR and
122 CEM DL. Four of these GPCR genes were uncharacterized; one (*seb-3*) has previously been shown
123 to play roles in locomotion, stress response, and ethanol tolerance (JEE *et al.* 2013). *dmsr-12* is
124 related to *daf-37* (ROBERTSON AND THOMAS 2006), a previously identified ascaroside receptor
125 gene (PARK *et al.* 2012), although it is more closely related to the neuropeptide receptor gene *dmsr-*
126 *1*, and more distantly to *srw-97* (ROBERTSON AND THOMAS 2006). *srd-32* belongs to a divergent
127 branch of the SRD phylogeny (ROBERTSON AND THOMAS 2006). *srr-7* belongs to the one of the
128 smallest families of *C. elegans* chemoreceptor genes (ROBERTSON AND THOMAS 2006).

129 *seb-3* and *srw-97* exhibited similar enrichment profiles across the CEM neurons, with 3.0-
130 and 3.3-fold higher expression in CEM VR than in any other CEM neuron (**Table 1; Figure 1D;**
131 **Table S5**). *dmsr-12* showed 25-fold higher expression in CEM DL versus any other CEM neuron
132 (**Table 1; Figure 1A; Table S5**), while *srd-32* was only 1.8-fold higher in CEM DL (**Table 1;**
133 **Figure 1A; Table S5**). *srr-7* showed 4.2-fold higher expression in CEM VL versus any other CEM

134 neuron (**Table 1; Figure 1C; Table S5**), which correlated with previous transcriptomic analyses
135 that found *srr-7* to be enriched in *C. elegans* EVNs (WANG *et al.* 2015).
136
137 *CEM-specific receptor expression patterns*
138 To confirm our single-cell RNA-seq results, we generated transgenic GFP fusions for the
139 five receptor genes (BOULIN *et al.* 2006). Roughly 3 kb of promoter region upstream of the start
140 codon was included in these constructs, along with the majority of the coding sequence (**Table**
141 **S6**); this would have automatically included any large 5'-ward introns that might contain cis-
142 regulatory elements of these genes (FUXMAN BASS *et al.* 2014). The GFP coding sequence was
143 cloned from the Fire Kit vector pPD95.75 (BOULIN *et al.* 2006). *pha-1; lite-1; him-5* animals were
144 injected with reporter constructs and the co-injection marker pBX (*pha-1*(+)). We isolated GFP^+
145 strains and imaged GFP^+ males for expression at 63x magnification (**Figure 2; Fig. S1**). An
146 integrated p*pkd-2*::GFP line was used as a CEM-specific control (**Figure 2A**).
147 The previously characterized GPCR gene, *seb-3*, displayed a non-CEM-specific expression
148 pattern matching that previously described (**Figure 2B**). The other four receptor genes showed
149 transgenic expression patterns similar to their RNA-seq data (**Figure 2C-E**). SRR-7 was observed
150 in CEM VR, as well as another neuron that is likely to be the mechanoreceptor CEP VR (SAWIN
151 *et al.* 2000) (**Figure 2F**). Both *seb-3* and *srr-7* were excluded from further analyses, as we aimed
152 to identify CEM-specific regulators of the ascr#8 response.
153 Except for SRD-32, all GFP-tagged receptors showed subcellular localization patterns that
154 included the sensory cilia (**Figure 2, white bars**). Our failure to see this for SRD-32 may be an
155 artifact of our p*srd-32*::GFP construct design, because only 52% of the *srd-32* coding sequence
156 was included in our transgene (**Table S6**). We also observed non-GPCR genes to be specifically

157 expressed in single CEM neurons, such as *trf-1*, which encodes a TNF receptor homolog (TENOR
158 AND ABALLAY 2008) expressed in CEM VR (**Table 1**); *trf-1* has an EVN-specific promoter (WANG
159 *et al.* 2015), providing an alternative to the *pkd-2* and *klp-6* promoters typically used to drive
160 transgenes in EVNs (**Fig. S1**) (PEDEN AND BARR 2005; BAE *et al.* 2006).

161

162 *RNAi-mediated knockdown of CEM Receptors*

163 To test whether these CEM-enriched genes encode receptors that are required in ascr#8
164 sensation, we fed dsRNA to a strain that is hypersensitive to neuronal RNA interference,
165 *nre-1; lin-15B* (SCHMITZ *et al.* 2007; POOLE *et al.* 2011).

166 To confirm that our system for male neuronal RNAi could affect behavioral phenotypes,
167 we first fed animals *osm-3* and *osm-9* dsRNA clones from the Ahringer Library (FRASER *et al.*
168 2000; KAMATH *et al.* 2003). Using a Spot Retention Assay (NARAYAN *et al.* 2016), we assayed
169 animals for their behavioral dwell time in ascr#8 (**Figure S2A**). Animals fed *osm-3* dsRNA
170 showed significantly defective responses to ascr#8, as observed for loss-of-function alleles
171 (**Figure S2B**). In contrast, animals fed *osm-9* showed only a slight decrease in their times spent
172 within ascr#8 (**Figure S2A**). Because we could successfully abolish ascr#8 attraction through male
173 neuronal RNAi of *osm-3*, we were confident that RNAi would be effective at functionally verifying
174 CEM-enriched genes encoding components of the ascr#8 response.

175 We fed animals dsRNA clones targeting three CEM-enriched receptor genes: *srd-32*, *dmsr-*
176 *12*, and *srw-97*. A *srd-32* clone was unavailable in the Ahringer library (FRASER *et al.* 2000;
177 KAMATH *et al.* 2003), but a present in the Vidal library (RUAL *et al.* 2004). Each library uses the
178 same backbone vector, allowing the same control to be used for both.

179 RNAi of *srd-32* caused no defect in ascr#8 responses (**Figure S2C**). Both *dmsr-12* or *srw-*
180 97 dsRNA caused partial defects of ascr#8 dwell time (**Figure S2C**). Neither knockdown
181 statistically lowered the dwell time in ascr#8 (**Figure S2D**), though they did abolish the statistically
182 significant increase in time spent in ascr#8 over vehicle controls (**Figure S2C**).
183
184 *CRISPR-generated null mutants of candidate receptors*
185 For phenotypic analysis of null mutants, we used our Single Worm Attraction Assay
186 (SWAA), as previously described (REILLY *et al.* 2021). Wild-type (*him-5*) animals were strongly
187 attracted to ascr#8 in the SWAA, replicating previous observations (**Figure 3A, B; Fig. S3**)
188 (REILLY *et al.* 2021).
189 *srw-97(knu456)* males displayed partially defective ascr#8 attraction: their ascr#8 dwell
190 time was no different than that of vehicle, nor was it different that *him-5* animals (**Figure 3A, B;**
191 **Fig. S3).** We were able to restore normal ascr#8 attraction in *srw-97(knu456)* males through
192 transgenesis with a translational fusion construct (p*srw-97*::*srw-97*::GFP) (**Fig. S4**). Expression of
193 the rescue transgene matched that of our initial fusion, with GFP visible in both ventral CEM
194 neurons (**Figure 3C; Fig. S5**).
195 Unlike *srw-97(knu456)*, *dmsr-12(tm8706)* males exhibited completely defective responses
196 to ascr#8 (**Figure 3A, B; Fig. S3**). Similarly, the expression profile of the p*dmsr-12*::*dmsr-*
197 *12*::GFP rescue construct (**Fig. S4**) matched its earlier reporter expression (**Figure 3D; Fig. S5**),
198 and completely restored attraction to ascr#8 (**Figure 3A, B; Fig. S3**). Quantification of GFP
199 expression confirmed SRW-97 to be enriched within the CEM_VL neuron, and to a lesser extent
200 CEM_VR and CEM_DL neurons, while DMSR-12 was observed within CEM_DL and CEM_DR
201 neurons (**Table S7**).

202 Given that *dmsr-12* is expressed in dorsal CEM neurons, and *srw-97* is expressed in ventral
203 CEM neurons (**Figure 3C, D**), we speculated that single mutants of either receptor partially retain
204 the ability to sense and respond to ascr#8, with the remaining CEM-specific receptor gene
205 conferring some residual response to ascr#8. We thus generated a *dmsr-12; srw-97* double mutant
206 strain and assayed animals for their ability to respond to ascr#8. These double mutants were not
207 attracted to ascr#8 compared to the vehicle control (**Figure 3A; Fig. S3**), though they were not
208 significantly defective in their attraction compared to wild-type animals either (**Figure 3B**).
209 However, this is likely a result of lower vehicle dwell times in the double mutants. This defect was
210 specific to ascr#8; double mutants showed no defect in their responses to ascr#3 (**Figure 3E, F;**
211 **Fig. S6**).
212 Together, these data suggest that at least two GPCR receptors are expressed in discrete
213 pairs of CEM neurons and act in parallel during sensation of ascr#8.
214
215 *Phylogenetic analyses of ascr#8 receptors reveal likely gene duplication events*
216 The ability to attract mates through pheromones is often essential for a species' survival.
217 We have recently observed that different *Caenorhabditis* species show variable levels of attraction
218 to ascr#8 (REILLY *et al.* 2019). To better understand variability in levels of attraction, we analyzed
219 the evolution of both *srw-97* and *dmsr-12*.
220 We identified orthologs of *srw-97* from *C. elegans* (*Cel-srw-97*) in other *Caenorhabditis*
221 proteomes via OrthoFinder (**Figure 4A**). A closely related *C. elegans* paralog, *Cel-srw-98*,
222 underwent a species-specific expansion within *C. inopinata*. The coding DNA sequences (CDS
223 DNAs) of both *Cel-srw-97* and *Cel-srw-98* are 66.2% identical to one another, while their amino
224 acid sequences are only 57.1% identical (MADEIRA *et al.* 2019). *Cel-srw-98* shows enriched

225 expression in CEM VR that is similar to (though weaker than) that of *Cel-srw-97* (**Table 1; Table**
226 **S5**).

227 *C. tropicalis*, like *C. elegans* and *C. briggsae*, is a hermaphroditic *Caenorhabditis* species
228 (NOBLE *et al.* 2021) which encodes a reduced number of ascr#8 receptor gene paralogs, with no
229 *Cel-dmsr-12* paralogs, and only one *Cel-srw-97* paralog (**Figure 4**). We previously observed that
230 *C. tropicalis* fails to be attracted to ascr#8, and it is possible that this loss of receptor genes is one
231 reason for that failure (REILLY *et al.* 2019).

232 In contrast, the other candidate ascr#8 receptor, *dmsr-12*, does not appear in our
233 phylogenetic analysis to have undergone any species-specific expansion (**Figure 4B**). Rather, there
234 are generally fewer paralogs per non-*C. elegans* species.

235

236 **Discussion**

237 Pheromones are important for mating in many animals. In the nematode *C. elegans*, males
238 are attracted to hermaphrodites as possible mates by the small molecule pheromone ascaroside #8
239 (PUNGALIYA *et al.* 2009; NARAYAN *et al.* 2016). While the ascaroside class of small molecule
240 pheromones utilized by nematodes is rapidly being elucidated (with over 230 known ascaroside
241 structures so far; *https://smid-db.org*) (ARTYUKHIN *et al.* 2018), the neuronal receptor proteins that
242 mediate pheromone signals remain largely unknown. For only a select few ascarosides have
243 sensory components been identified at the cellular (GREENE *et al.* 2016a; GREENE *et al.* 2016b;
244 CHUTE *et al.* 2019), receptor (KIM *et al.* 2009; MCGRATH *et al.* 2011; PARK *et al.* 2012), or signal
245 transduction levels (ZWAAL *et al.* 1997).

246 Here, we identify two novel G protein-coupled receptors as active, required components in
247 sensory and behavioral responses to the mating pheromone ascr#8. Transcripts for the two GPCRs,

248 *dmsr-12* and *srw-97*, are enriched in single CEM neurons (**Table 1**), and they express in non-
249 overlapping subsets of the male-specific chemosensory neurons (**Figure 1, 2**). There are, however,
250 other receptors present in these same neurons that contribute to the navigation of a vast and ever-
251 changing array of environmental cues, such as the widely expressed ethanol sensor *seb-3* (**Figure**
252 **2B**) (JEE *et al.* 2013).

253 One mechanism for sensory flexibility in CEM neurons may be heterodimerization of
254 receptor proteins. Previous work identified two receptors for ascr#2 in the ASK neuron: DAF-37
255 and DAF-38 (PARK *et al.* 2012). While both are required for proper ascr#2-induced dauer
256 formation, only DAF-38 is involved in the sensation of other ascarosides (PARK *et al.* 2012). Such
257 heterodimers may also exist for pheromone receptors in CEM neurons.

258 Another mechanism for sensory flexibility in CEM neurons is to have diverse receptor
259 proteins with distinct specificities co-expressed within a single neuron. Multiple ascarosides are
260 sensed by the male-specific CEM neurons, including ascr#8 and ascr#3 (NARAYAN *et al.* 2016).
261 The two receptors identified here may function as ascr#8-specific receptors, as *dmsr-12;srw-97*
262 double mutant animals do not exhibit defective responses to ascr#3 (**Figure 3E, F**). The related
263 receptors *dmsr-10*, *dmsr-13*, *dmsr-16*, and *srw-98* are enriched in ventral CEM neurons, and may
264 heterodimerize with *srw-97*, to result in fine-tuned sensation of ascr#8. Similarly, there may be
265 other receptors may be enriched in the dorsal CEM neurons to heterodimerize with *dmsr-12*. The
266 role of DAF-37 as an ascr#2-specific receptor further supports our hypothesis, as *dmsr-12* is related
267 to DAF-37 (ROBERTSON AND THOMAS 2006). The ventral CEM receptor SRW-97 falls within the
268 same large family of GPCRs as well (KRISHNAN *et al.* 2014), while the promiscuous DAF-38 does
269 not (ROBERTSON AND THOMAS 2006).

270 The initial goal of our single-cell RNA-seq analysis of male-specific chemosensory CEM
271 neurons was to identify GPCRs encoding pheromone receptors. However, further analysis of other
272 genes with CEM-enriched expression should also uncover specific and novel promoter profiles,
273 which will enable optogenetic manipulation and calcium imaging of defined individual CEM
274 neurons (NARAYAN *et al.* 2016). The development of advanced transgenic reagents will also permit
275 chemical biology of CEM neurons. We have recently developed an active ascr#8 bioaffinity probe
276 to employ in the targeted elucidation of ascr#8 receptors (ZHANG *et al.* 2019). Use of this probe
277 will lead to further elucidation of the identity of acsr#8 receptors, by confirming either SRW-97
278 or DMSR-12 as receptors or identifying their heterodimeric partners. The combination of these
279 technologies should clarify how heterogeneous CEM neurons achieve homogeneous sensory
280 responses.

281 **Methods**

282 ***Single cell isolation and RT-PCR***

283 Individual ciliated male-specific cephalic sensory (CEM) neurons were isolated from adult
284 males of the *C. elegans* strain CU607 (smIs23 [p*pkd-2*::GFP+ pBX]; *him-5*(e1490)), as previously
285 described (NARAYAN *et al.* 2016). This strain expresses an integrated GFP transgene that labels
286 male-specific extracellular vesicle-releasing neurons (EVNs, p*pkd-2*::GFP [Figure 2A]). Cells
287 were separated by anatomical identity (i.e., CEM dorsal left [CEM_DL], CEM dorsal right
288 [CEM_DR], CEM ventral left [CEM_VL], and CEM ventral right [CEM_VR]). Microdissection
289 and single-cell RT-PCR of individual CEM_DL, CEM_DR, CEM_VL, and CEM_DR neurons
290 was performed essentially as described (SCHWARZ *et al.* 2012). For all four neuronal types, single-
291 end 50-nucleotide (50-nt) to 100-nt RNA-seq was performed on an Illumina HiSeq 2000. To
292 identify so-called housekeeping genes and genes primarily active outside the nervous system, we
293 compared our results from CEM_DL, CEM_DR, CEM_VL, and CEM_DR to equivalent
294 reanalyses of published single-end 38-nt RNA-seq data from mixed-stage whole *C. elegans*
295 hermaphrodite larvae (SCHWARZ *et al.* 2012), and of published single-end 36-nt RNA-seq data
296 from adult whole *C. elegans* males and females (THOMAS *et al.* 2012).

297 ***Transcriptional analysis***

298 Reads from RT-PCR-amplified CEM neurons and larvae were quality-filtered as follows:
299 neuronal reads that failed Chastity filtering were discarded (Chastity filtering had not been
300 available for the larval reads); raw 38-nt larval reads were trimmed 1 nt to 37 nt; all reads were
301 trimmed to remove any indeterminate ("N") residues or residues with a quality score of less than
302 3; and larval reads that had been trimmed below 37 nt were deleted, as were neuronal reads that
303 had been trimmed below 50 nt. For uniformity of analysis, all RNA-seq reads of 51-100 nt in

304 length were trimmed to 50 nt before further RNA-seq analysis. This left a total ranging from
305 55,902,297 (CEM_VL) to 73,412,777 (CEM_VR) filtered reads for analysis of each neuronal type,
306 versus 23,369,056 filtered reads for whole larvae (**Table S2**). Previously published RNA-seq reads
307 (THOMAS *et al.* 2012) from three biological replicates of adult males (accession numbers
308 SRR580386, SRR580387, and SRR580388) and adult females (accession numbers SRR580383,
309 SRR580384, and SRR580385) were downloaded from the European Nucleotide Archive
310 (*ftp://ftp.sra.ebi.ac.uk*) and used without further trimming or filtering.

311 We used RSEM version 1.2.17 (LI AND DEWEY 2011) with bowtie2 version 2.2.3
312 (LANGMEAD AND SALZBERG 2012) and SAMTools version 1.0 (LI *et al.* 2009) to map filtered reads
313 to a *C. elegans* gene index and generate read counts and gene expression levels in transcripts per
314 million (TPM). To create the *C. elegans* gene index, we ran RSEM's *rsem-prepare-reference* with
315 the arguments '"*--bowtie2 --transcript-to-gene-map*' upon a collection of coding DNA sequences
316 (CDSes) from both protein-coding and non-protein-coding *C. elegans* genes in WormBase release
317 WS245 (HOWE *et al.* 2016). The CDS sequences were obtained from *ftp://ftp.sanger.ac.uk/pub2/*
318 *wormbase/releases/WS245/species/c_elegans/PRJNA13758/c_elegans.*
319 *PRJNA13758.WS245.mRNA_transcripts.fa.gz* and *ftp://ftp.sanger.ac.uk/pub2/wormbase/*
320 *releases/WS245/species/c_elegans/PRJNA13758/c_elegans.PRJNA13758.WS245.ncRNA_*
321 *transcripts.fa.gz*. For each RNA-seq data set of interest, we computed mapped reads and
322 expression levels per gene by running RSEM''s *rsem-calculate-expression* with the arguments
323 '"*--bowtie2 -p 8 --no-bam-output --calc-pme --calc-ci --ci-credibility-level*
324 *0.99 --fragment-length-mean 200 --fragment-length-sd 20 --estimate-rspd --ci-memory 30000*'.
325 These arguments, in particular '*--estimate-rspd*', were aimed at dealing with single-end data from
326 3'-biased RT-PCR reactions; the argument '*--phred64-quals*' was used for the larval reads, while

327 '*--phred33-quals*' was used for all other reads (neuronal, adult males, and adult females). We
328 computed posterior mean estimates (PMEs) both for read counts and for gene expression levels,
329 and rounded PMEs of read counts down to the nearest lesser integer. We also computed 99%
330 credibility intervals (Cis) for expression data, so that we could use the minimum value in the 99%
331 CI for TPM as a robust minimum estimate of a gene's expression (minTPM).

332 We observed the following overall alignment rates of reads to the WS245 *C. elegans* gene
333 index: 47.16% for the CEM_DL read set, 27.37% for the CEM_DR read set, 40.20% for the
334 CEM_VL read set, 17.99% for the CEM_VR read set, and 76.41% for the larval read set (**Table**
335 **S2**). A similar discrepancy between lower alignment rates for hand-dissected linker cell RNA-seq
336 reads versus higher alignment rates for whole larval RNA-seq reads was previously observed, and
337 found to be due to a much higher rate of human contaminant RNA sequences in the hand-dissected
338 linker cells (SCHWARZ *et al.* 2012). For previously published adult male and female reads, we
339 observed mapping rates of 88.20% to 96.11%; this may reflect greater difficulty of getting
340 mappable reads from our RT-PCR-amplified cDNAs than from unamplified cDNA syntheses.

341 We defined detectable expression for a gene in a given RNA-seq data set by that gene
342 having an expression level of 0.1 TPM; we defined robust expression by that gene having a
343 minimum estimated expression level (termed minTPM) of at least 0.1 TPM in a credibility interval
344 of 99% (i.e., ≥ 0.1 minTPM). The numbers of genes being scored as expressed in a given neuronal
345 type above background levels, for various data sets, are given in **Table S3**. Other results from
346 RSEM analysis are given in **Table S5**.

347 We annotated *C. elegans* genes and the encoded gene products in several ways (**Table S5**).
348 For the products of protein-coding genes, we predicted classical signal and transmembrane
349 sequences with Phobius 1.01 (KÄLL *et al.* 2004), regions of low sequence complexity with pseg

350 (SEG for proteins, from *ftp://ftp.ncbi.nlm.nih.gov/pub/seg/pseg*) (WOOTTON 1994), and coiled-coil
351 domains with ncoils (from *http://www.russell.embl-heidelberg.de/coils/coils.tar.gz*) (LUPAS
352 1996). Protein domains from Pfam 31.0 (FINN *et al.* 2016) were detected with HMMER 3.1b2/
353 *hmmsearch* (EDDY 2009), using the argument '*--cut_ga*' to invoke model-specific reporting
354 thresholds. Memberships of genes in orthology groups from eggNOG 3.0 (POWELL *et al.* 2012)
355 were extracted from WormBase WS245 with the TableMaker function of ACEDB 4.3.39. Genes
356 with likely housekeeping status (based on ubiquitous expression in both larvae and linker cells)
357 were as identified in our previous work (SCHWARZ *et al.* 2012). Genes were predicted to encode
358 GPCRs on the basis of their encoding a product containing one or more of the following Pfam-A
359 protein domains: 7tm_1 [PF00001.16], 7tm_2 [PF00002.19], 7tm_3 [PF00003.17], 7tm_7
360 [PF08395.7], 7TM_GPCR_Srab [PF10292.4], 7TM_GPCR_Sra [PF02117.11],
361 7TM_GPCR_Srbc [PF10316.4], 7TM_GPCR_Srb [PF02175.11], 7TM_GPCR_Srd [PF10317.4],
362 7TM_GPCR_Srh [PF10318.4], 7TM_GPCR_Sri [PF10327.4], 7TM_GPCR_Srj [PF10319.4],
363 7TM_GPCR_Srsx [PF10320.4], 7TM_GPCR_Srt [PF10321.4], 7TM_GPCR_Sru [PF10322.4],
364 7TM_GPCR_Srv [PF10323.4], 7TM_GPCR_Srw [PF10324.4], 7TM_GPCR_Srx [PF10328.4],
365 7TM_GPCR_Srz [PF10325.4], 7TM_GPCR_Str [PF10326.4], ABA_GPCR [PF12430.3], Sre
366 [PF03125.13], and Srg [PF02118.16]. By this criterion, we identified 1,615 genes encoding
367 GPCRs in the WS245 version of the *C. elegans* genome; this resembles a previous estimate of
368 ~1,470 *C. elegans* genes encoding chemoreceptors and other GPCRs, identified through extensive
369 computational and manual analysis (HOBERT 2013). The memberships of genes in orthology
370 groups from eggNOG 3.0 (POWELL *et al.* 2012) were extracted directly from WormBase WS245
371 with the TableMaker function of ACEDB 4.3.39. Genes with likely housekeeping status (based on
372 ubiquitous expression in both larvae and linker cells) were as identified in our previous work

373 (SCHWARZ *et al.* 2012). Gene Ontology (GO) annotations for *C. elegans* genes were extracted from
374 WormBase-computed annotations in *ftp://ftp.wormbase.org/pub/wormbase/releases/WS245/*
375 *ONTOLOGY/gene_association.WS245.wb.c_elegans*; human-readable text descriptions for GO
376 term IDs were extracted from *term.txt* in the Gene Ontology archive *http://archive.geneontology.*
377 *org/full/2014-07-01/go_201407-termdb-tables.tar.gz*.
378
379 ***GFP reporter construction***
380 Reporter fusion constructs were generated using previously described techniques (BOULIN
381 *et al.* 2006). Approximately 2-3 kb of upstream promoter region of each gene was included in
382 construct generation, as well as a portion of the coding sequence (**Table S6**). This was then fused
383 to GFP (from the Fire Vector Kit plasmid, pPD95.75), via PCR fusion (BOULIN *et al.* 2006).
384 Primers were designed using Primer 3 and ordered from IDT (Integrated DNA Technologies).
385 Primer sequences available in **Table S9**. Successful fusion was confirmed via gel electrophoresis
386 prior to injection.
387 Reporter fusion constructs were injected into the gonads of *pha-1(e2123ts); lite-1(ce314);*
388 *him-5(e1490)* animals, along with a co-injection marker of pBX(*pha-1*(+)). In this manner, positive
389 array animals will propagate normally at 20 °C. Strains were confirmed via GFP expression, with
390 multiple array lines being generated per injection (**Table S8**). Injections were performed by InVivo
391 Biosystems, with strain isolation being performed in house.
392
393 ***Imaging***
394 Animals were imaged for GFP expression using previously described techniques. GFP+
395 young adult male animals were mounted on a 1% agarose pad and immobilized with sodium azide.

396 Animals were then imaged on a spinning disk confocal microscope at 63x magnification. Z-stack
397 imaging was performed, generating 3D reconstructions of the heads of the imaged animals.
398 Central/optimal z-plane images were used to generate the images used to verify expression (**Figure**
399 **2; Fig. S1**).
400
401 ***RNAi Feeding***
402 VH624 (*rhIs13* [*unc-119*::GFP + *dpy-20*(+)]; *nre-1(hd20)*; *lin-15B(hd126)*) animals
403 (SCHMITZ *et al.* 2007; POOLE *et al.* 2011) were crossed with *him-5(e1490)* animals to integrate
404 male production into a strain hypersensitive to neuronal RNAi-knockdown, generating JSR44
405 (*nre-1(hd20)*; *lin-15B(hd126); him-5(e1490)*). During the cross, insertion of the *him-5(e1490)*
406 allele displaced the integrated array *rhIs13*, suggesting location of the array on Chromosome V.
407 Presence of *lin-15B(hd126)* in JSR44 was confirmed via sequencing. The non-annotated
408 *nre-1(hd20)* is linked with *lin-15B*, being retained alongside *lin-15B* (SCHMITZ *et al.* 2007; POOLE
409 *et al.* 2011).
410 RNAi clones were grown overnight in cultures of LB containing 50 µg/mL ampicillin.
411 Cultures were then diluted to an OD_{600} of 1.0 before plating on NGM agar plates containing 50
412 µg/mL ampicillin and 1 mM IPTG (Isopropyl-β-D-thiogalactoside) to select for RNAi clones and
413 induce expression. Lawns were allowed to grow at room temperature for 8-16 hours, before JSR44
414 young adult hermaphrodites were placed on the plates and left to propagate at 16°C. Young adult
415 males of the F1 progeny were then selected for behavioral testing suing the *Spot Retention Assay*.
416 Empty vector controls (VC-1 clone) were run alongside every targeted knockdown experiment.
417

418 ***Spot Retention Assay***

419 Following previously described methods, young adult males were isolated from
420 hermaphrodites 5-16 hours prior to testing (PUNGALIYA *et al.* 2009; NARAYAN *et al.* 2016). In
421 short, at the time of the assay, 0.6 µL of either vehicle control (-) or ascaroside #8 (+) was added
422 to the NGM plates covered in a thin lawn of OP50 *E. coli*. Ten males were then divided between
423 two pre-marked spots on the agar, equidistant from the cues. The plate was then recorded for 20
424 minutes. The time spent of each visit in either vehicle or ascaroside #8 (if greater than 10 seconds)
425 was scored and averaged. Plates in which the average was greater than two standard deviations
426 removed from the population average were removed from the final analysis as outliers. To compare
427 between strains or conditions, the vehicle was subtracted from the ascaroside dwell time for each
428 plate. The average of these differences was then compared statistically.

429

430 ***Single Worm Behavioral Assay***

431 Following previously described methods (PUNGALIYA *et al.* 2009; NARAYAN *et al.* 2016;
432 REILLY *et al.* 2021), animals were isolated and prepared in an identical manner to the Spot
433 Retention Assay. The two outside rings of wells in a 48-well tissue culture plate were filled with
434 200 µL of NGM agar, which was then seeded with 65 µL of OP50 *E. coli*. The plates and lawns
435 were then dried at 37 °C for 4 hours. Alternating wells were then prepared as spatial controls ("S",
436 nothing done), vehicle controls ("V", 0.85 µL of dH_2O was placed in the center of the well), or
437 ascaroside well ("A", 0.85 µL of ascaroside was placed in the center of the well). This was
438 performed over four quadrants. Animals were scored for their visits and duration to the center of
439 the well and/or the cue.

440 The average duration of each worm's visits was calculated, and these values were again
441 averaged together to generate a Mean Dwell Time in seconds for each plate. When comparing
442 across strains or conditions, the Spatial controls were then compared for statistical difference. If
443 none was observed, the Log(fold-change) A/V was then calculated by taking the log of the
444 ascaroside mean dwell time divided by the vehicle mean dwell time for each plate. The number of
445 times each worm visited the center was averaged to generate the Visit Counts.

446 The Percent Attraction values were calculated by first determining the "attractive" cut off
447 as two standard deviations above the vehicle average. Any visit longer than this was deemed
448 "attractive" and scored as a "1"; non-attractive visits were scored as "0". The percent attraction
449 was then calculated for each worm was the percent of visits scored a "1". The average was then
450 calculated across the plate to determine Percent Attraction.

451

452 ***CRISPR Design and Strain Generation***

453 A novel null mutation was generated for *srw-97* by InVivo Biosystems. The *srw-*
454 *97(knu456)* allele was generated in a *him-5(e1490)* strain using two sgRNAs
455 (TTTAGTAGAGCAGAAATTAA and TACAGCTTTAACTTTCAAC) to generate a 1,620-nt
456 deletion which removed the start codon and left only the terminal exon intact. The *knu456*
457 knockout allele was generated by donor homology using the pNU1361odn oligo:
458 (TTTTCTTGTATTTCCAAAAATTGTAAAAACCTTTATGAAAGTTAAAGCTGTAAGGAT
459 TTTCAGACATTTA). Following generation of a homozygous deletion by InVivo Biosystems,
460 the line was then backcrossed twice into *him-5(e1490)*.

461 The *dmsr-12(tm8706)* allele, provided by the National BioResource Group (NRBP)
462 contains a 118-nucleotide deletion that was generated by Dr. Mitani of the NRBP. The allele was

463 crossed into a *him-5(e1490)* background prior to testing. The deletion spans intron 2 and exon 3
464 of the coding sequence. Whether this results in a correctly spliced gene remains unknown, although
465 the expected remaining coding sequence remains in frame.
466
467 ***Phylogenetic Analyses***
468 For phylogenetic analysis of selected CEM genes, we downloaded proteomes for
469 *C. elegans* and related *Caenorhabditis* nematodes from WormBase (release WS275), the Blaxter
470 *Caenorhabditis* database (release 1), or our unpublished work, as listed in **Table S4**. From each
471 proteome, we extracted the longest predicted isoform for each gene with *get_largest_isoforms.pl*
472 (*https://github.com/SchwarzEM/ems_perl/blob/master/fasta/get_largest_isoforms.pl*). We
473 observed that the predicted isoform for *dmsr-12* in the WormBase WS275 release of *C. elegans*'
474 proteome was shorter than past versions of *dmsr-12*, and that the WS275 isoform omitted exons
475 that our transgenic expression data (based on older gene models for *dmsr-12*) indicated were likely
476 to be real. We therefore manually replaced the WS275 version of *dmsr-12* with an older version
477 (extracted from the *C. elegans* proteome in the WS250 release of WormBase). We then computed
478 orthology groups for *C. elegans* and its related species with OrthoFinder version 2.3.11 (EMMS
479 AND KELLY 2015; EMMS AND KELLY 2019), using the arguments '*-a 1 -S diamond -og*'. We
480 identified which orthology groups contained the *C. elegans* genes *srr-7*, *srw-97*, and *dmsr-12*, and
481 extracted their sequences from a concatenation of all 11 proteomes via *extract_fasta_subset.pl*
482 (*https://github.com/SchwarzEM/ems_perl/blob/master/fasta/extract_fasta_subset.pl*). For each
483 orthogroup's member sequences, we aligned the sequences with MAFFT version 7.455 (KATOH
484 AND STANDLEY 2013) and filtered the alignments twice with trimAl version 1.4.rev15 (CAPELLA-
485 GUTIÉRREZ *et al.* 2009), using first the argument '*-automated1*' and then the arguments

486 '*-resoverlap 0.50 -seqoverlap 50*'. From the filtered alignments, we computed maximum-
487 likelihood protein phylogenies with IQ-TREE version 2.0-rc1 (NGUYEN *et al.* 2015;
488 KALYAANAMOORTHY *et al.* 2017), using the arguments '*-m MFP -b 100 --tbe*'. In particular, we
489 used transfer bootstrap expectation ('*--tbe*') which provides more reliable confidence values than
490 classic bootstrapping (LEMOINE *et al.* 2018). We visualized the resulting phylogenies with FigTree
491 version 1.4.4 (*http://tree.bio.ed.ac.uk/software/figtree*).
492
493 ***Statistical Analyses***
494 Prior to any statistical analyses, outliers were identified and removed. Outliers were
495 defined as any data points greater than two standard deviations removed from the average. All data
496 were then tested for normality using a Shapiro-Wilk Normality Test. This test was chosen over the
497 more conventional D'Agostino-Pearson Normality Test as many data sets were below 10 in
498 number (due to the statistical power offered by the Single Worm Behavioral Assay (REILLY *et al.*
499 2021).
500 The Spot Retention Assay data was analyzed using paired *t*-tests or Wilcoxon Matched-
501 Pairs Signed Rank tests to compare vehicle control and ascaroside dwell times within strains
502 following tests for normality (**Figure S2**). When comparing the values of multiple conditions or
503 strains, the data was first normalized to account for vehicle dwell time variation between plates
504 using a base-2 exponentiation, as described previously, to transform all data points into non-zero
505 values. This allows for the calculation of the fold-change as the log(base2) of the ascaroside dwell
506 time divided by the vehicle dwell time. These normalized values were then compared using a
507 Mann-Whitney test or a One-Way ANOVA followed by a Dunnett's multiple comparisons test
508 (**Figure S2**).

509 The Single Worm Behavioral Assay was first analyzed within each strain by performing
510 either a Repeated-Measures ANOVA followed by a Dunnett's multiple comparisons test or
511 Friedman test followed by Dunn's correction, comparing both Spatial control and Ascaroside
512 dwell times to that of the Vehicle Control. The log(fold-change) of raw ascaroside and vehicle
513 dwell time values were calculated and compared using either a One-Way ANOVA followed by a
514 Bonferroni's multiple corrections test or Friedman Test followed by Dunn's correction (**Figure 3**).
515 The ascr#3 log(fold-change) values were compared using a Student's *t*-test (**Figure 3F**). Visit
516 Counts were compared in the same manner as the Mean Dwell Time data, while the Percent
517 Attraction data was analyzed using paired *t*-tests or Wilcoxon Matched-Pairs Signed Rank tests to
518 compare the attractive values of the vehicle and ascaroside (**Figures S3, S4, S5**).
519

520 ***Data availability***

521 RNA-seq reads have been deposited in the NCBI Sequence Read Archive (SRA) under the
522 BioProject accession number PRJNA781271.
523

524 ***Acknowledgements***

525 We thank the *Caenorhabditis* Genetics Center, which is funded by the NIH Office of
526 Research Infrastructure Programs (P40 OD01044), as well as the National BioResource Project,
527 Ding Xue (University of Colorado, Boulder), L. Rene Garcia (Texas A&M University), and
528 Douglas Portman (University of Rochester Medical Center) for providing strains. We also thank
529 Igor Antoshechkin, Caltech Genomics facility for sequencing and InVivo Biosystems for
530 generating transgenic and CRISPR knockout animals. The synthetic ascr#8 utilized in this study
531 was provided by Frank Schroeder (Cornell University). The research reported in this publication

532 was supported by NIH R01 DC016058 (J.S.), R01 GM084389 (P.W.S.), the Howard Hughes
533 Medical Institute (P.W.S.), Moore Foundation Grant No. 4551 (E.M.S.), and Cornell startup
534 funding (E.M.S.). We thank Titus Brown and the Michigan State University High-Performance
535 Computing Center (supported by U.S. Department of Agriculture grant 2010-65205-20361 and
536 NIFA–National Science Foundation (NSF) grant IOS-0923812) for computational support;
537 additional computing was enabled by start-up and research allocations from NSF XSEDE (TG-
538 MCB180039 and TG-MCB190010).

539 **Figure and Table Legends**

Gene	CEM DL	CEM DR	CEM VL	CEM VR	CEM/others	CEM/males
srw-97	0.02	0.03	0.02	**0.10**	3.33	2.50
dmsr-12	**1.26**	0.03	0.02	0.05	25.2	18.0
seb-3	0.01	0.02	0.01	**0.06**	3.00	0.02
srd-32	**0.07**	0.02	0.02	0.04	1.75	0.17
srr-7	0.03	0.03	**0.17**	0.04	4.25	0.68
pkd-2	**1.36**	0.09	0.42	1.15	1.18	0.15
trf-1	0.23	0.04	0.01	**78.02**	339.22	8.72
srw-98	0.02	0.03	0.02	**0.05**	1.67	0.20
dmsr-13	0.02	0.03	0.02	**0.05**	1.67	0.45
dmsr-10	0.02	0.04	0.03	**0.07**	1.75	1.40
dmsr-11	0.01	0.02	0.02	**0.04**	2.00	0.33
dmsr-16	0.02	0.04	0.03	**0.06**	1.50	0.67

540

541 **Table 1. Expression levels for genes of interest in CEM neurons.** For each gene, observed
542 RNA-seq expression levels in each CEM neuron type are given in transcripts per million (TPM);
543 CEM types with the highest expression are in boldface for each gene. Two measures of gene
544 enrichment and specificity are also shown: first, the ratio of a gene's expression for the highest-
545 expressing CEM neuron type divided by that gene's expression in the highest-expressing of the
546 other three CEM neuron types ('CEM/others'; 'CEM_any/CEM_non_any.max' in **Table S5**); and,
547 the ratio of that gene's highest-expressing CEM neuron type divided by that gene's expression in
548 the highest-expressing of three replicates of adult *C. elegans* males ('CEM/males';
549 'CEM_all.max_TPM/male.max_TPM' in **Table S5**). The first ratio gives the skew for a gene's
550 expression towards a single CEM type; the second ratio gives the degree to which that gene's
551 expression cannot be explained merely by its being a male-specific gene, perhaps expressed
552 throughout the male body. Genes listed here are of interest either because we tested them for
553 transgenic GFP expression or biological function, because they have previously known CEM-
554 specificity, or because they are evolutionary related to *srw-97* and *dmsr-12*.

555

556 **Figure 1. Transcriptomic landscapes of the CEM neurons. (A-D)** Expression plots of genes
557 expressed in individual CEM DL **(A)**, CEM DR **(B)**, CEM VL **(C)**, and CEM VR **(D)** neurons.
558 For each gene, its X-axis position displays the ratio of its expression in an individual CEM neuron
559 divided by its highest expression in three replicates of adult *C. elegans* males; its Y-axis position
560 displays expression in an individual CEM neuron type. Expression levels are measured in
561 transcripts per million (TPM). Genes of interest are denoted by colored symbols, defined in the
562 legend.
563
564 **Figure 2. Expression profiles of CEM-enriched GPCR genes of interest.** Transcriptional GFP
565 fusions of GPCR genes enriched in the CEM neurons. **(A)** A previously published CEM reporter,
566 p*pkd-2*::GFP. **(B)** p*seb*-3::GFP, matching previously published expression, with no discernable
567 enrichment in the CEM. **(C)** p*dmsr-12*::GFP is strongly expressed in the CEM DL neuron, as well
568 as the CEM DR soma. **(D)** p*srd-32*::GFP is weakly expressed in the somas of CEM DR and DL.
569 **(E)** p*srw-97*::GFP is expressed in both CEM VR and VL, with localization in the cilia as well as
570 the soma. **(F)** p*srr-7*::GFP is expressed in CEM VR, both the soma and cilia, as well the cilial of
571 a neighboring neuron, presumably CEP VR. Dorsal/ventral axes and anterior/posterior directions
572 shown in (B). Scale bars denote distal, cilia region.
573
574 **Figure 3. The GPCRs, SRW-97 and DMSR-12, act in non-overlapping CEM neurons to aid
575 in driving the ascr#8 response. (A)** Raw dwell times in the centers of spatial control (S), vehicle-
576 only control (V), and ascaroside (A) wells, and **(B)** Log(fold-change) of A versus V for *srw-97*
577 knockout and transgenic rescue (WorEx57), *dmsr-12* and transgenic rescue (WorEx54), as well as
578 *srw-97*;*dmsr-12* double mutant animals. **(C)** Expression of WorEx61 (p*srw-97*::*srw-97*::GFP) in

579 CEM VR and CEM VL. **(D)** Expression of WorEx60 (p*dmsr-12*::*dmsr-12*::GFP) in CEM DL and
580 CEM DR. IL2_LV and IL2_L somas denoted. Cellular identities confirmed through co-expression
581 of p*klp-6*::dtTomato (**Fig. S5**). Anterior/Posterior and Right/Left axes denoted. **(E)** Raw dwell
582 times and **(F)** Log(fold-change) of *him-5* and *dmsr-12;srw-97* animals in response to the
583 structurally related attracted pheromone, ascr#3. **(G)** Structural comparison of ascr#8 (left) and
584 ascr#3 (right). Error bars denote SEM. $n \geq 5$. **(A, E)** RM-ANOVA followed by Dunnett's post-hoc
585 test or Friedman test followed by Dunn's (within strain), **(B)** ANOVA followed by Bonferroni's
586 post-hoc test (between strain log(fold-change) values). **(F)** Student's t-test between strain log(fold-
587 change) values * $p < 0.05$, ** $p < 0.01$.
588
589 **Figure 4. Phylogenetic analysis of ascr#8-receptor candidate paralogs across the**
590 ***Caenorhabditis* genus. (A)** The phylogeny of *Cel_srw-97* reveals a *C. inopinata*-specific
591 amplification of *Cel_srw-98*. **(B)** The phylogeny of *Cel_dmsr-12* shows conserved numbers of
592 orthologs across the genus. Genes for *C. elegans* are denoted in purple, *C. inopinata* in pink,
593 *C. nigoni* in light blue, *C. briggsae* in dark blue, *C. japonica* in red, and *C. tropicalis* in dark green.
594 Distance reference bars (0.1) depict substitutions per site.

595 **References**

596 Aprison, E. Z., and I. Ruvinsky, 2015 Sex Pheromones of *C. elegans* Males Prime the Female
597 Reproductive System and Ameliorate the Effects of Heat Stress. PLoS Genet 11**:**
598 e1005729.
599 Aprison, E. Z., and I. Ruvinsky, 2017 Counteracting Ascarosides Act through Distinct Neurons
600 to Determine the Sexual Identity of *C. elegans* Pheromones. Curr Biol 27**:** 2589-
601 2599.e2583.
602 Artyukhin, A. B., Y. K. Zhang, A. E. Akagi, O. Panda, P. W. Sternberg *et al.*, 2018 Metabolomic
603 "Dark Matter" Dependent on Peroxisomal beta-Oxidation in *Caenorhabditis elegans*. J
604 Am Chem Soc 140**:** 2841-2852.
605 Bae, Y. K., H. Qin, K. M. Knobel, J. Hu, J. L. Rosenbaum *et al.*, 2006 General and cell-type
606 specific mechanisms target TRPP2/PKD-2 to cilia. Development 133**:** 3859-3870.
607 Boulin, T., J. F. Etchberger and O. Hobert, 2006 Reporter gene fusions. WormBook**:** 1-23.
608 Butcher, R. A., M. Fujita, F. C. Schroeder and J. Clardy, 2007 Small-molecule pheromones that
609 control dauer development in *Caenorhabditis elegans*. Nat Chem Biol 3**:** 420-422.
610 Butcher, R. A., J. R. Ragains, E. Kim and J. Clardy, 2008 A potent dauer pheromone component
611 in Caenorhabditis elegans that acts synergistically with other components. Proc Natl
612 Acad Sci U S A 105**:** 14288-14292.
613 Capella-Gutiérrez, S., J. M. Silla-Martínez and T. Gabaldón, 2009 trimAl: a tool for automated
614 alignment trimming in large-scale phylogenetic analyses. Bioinformatics (Oxford,
615 England) 25**:** 1972-1973.
616 Choe, A., S. H. von Reuss, D. Kogan, R. B. Gasser, E. G. Platzer *et al.*, 2012 Ascaroside
617 Signaling is Widely Conserved Among Nematodes. Current Biology 22**:** 772-780.
618 Chute, C. D., E. M. DiLoreto, Y. K. Zhang, D. K. Reilly, D. Rayes *et al.*, 2019 Co-option of
619 neurotransmitter signaling for inter-organismal communication in *C. elegans*. Nat
620 Commun 10**:** 3186.
621 Cochella, L., B. Tursun, Y. W. Hsieh, S. Galindo, R. J. Johnston *et al.*, 2014 Two distinct types
622 of neuronal asymmetries are controlled by the *Caenorhabditis elegans* zinc finger
623 transcription factor die-1. Genes Dev 28**:** 34-43.
624 Dong, C., F. Dolke and S. H. von Reuss, 2016 Selective MS screening reveals a sex pheromone
625 in Caenorhabditis briggsae and species-specificity in indole ascaroside signalling. Org
626 Biomol Chem 14**:** 7217-7225.
627 Dong, C., D. K. Reilly, C. Bergame, F. Dolke, J. Srinivasan *et al.*, 2018 Comparative Ascaroside
628 Profiling of *Caenorhabditis* Exometabolomes Reveals Species-Specific (ω) and (ω - 2)-
629 Hydroxylation Downstream of Peroxisomal beta-Oxidation. J Org Chem.
630 Eddy, S. R., 2009 A new generation of homology search tools based on probabilistic inference.
631 Genome informatics. International Conference on Genome Informatics 23**:** 205-211.
632 Emms, D. M., and S. Kelly, 2015 OrthoFinder: solving fundamental biases in whole genome
633 comparisons dramatically improves orthogroup inference accuracy. Genome Biology 16**:**
634 157.
635 Emms, D. M., and S. Kelly, 2019 OrthoFinder: phylogenetic orthology inference for comparative
636 genomics. Genome Biology 20**:** 238.
637 Finn, R. D., P. Coggill, R. Y. Eberhardt, S. R. Eddy, J. Mistry *et al.*, 2016 The Pfam protein
638 families database: towards a more sustainable future. Nucleic acids research 44**:** D279-
639 D285.

640 Fraser, A. G., R. S. Kamath, P. Zipperlen, M. Martinez-Campos, M. Sohrmann *et al.*, 2000
641 Functional genomic analysis of *C. elegans* chromosome I by systematic RNA
642 interference. Nature 408**:** 325-330.
643 Fuxman Bass, J. I., A. M. Tamburino, A. Mori, N. Beittel, M. T. Weirauch *et al.*, 2014
644 Transcription factor binding to Caenorhabditis elegans first introns reveals lack of
645 redundancy with gene promoters. Nucleic Acids Res 42**:** 153-162.
646 Goodman, M. B., D. H. Hall, L. Avery and S. R. Lockery, 1998 Active currents regulate
647 sensitivity and dynamic range in *C. elegans* neurons. Neuron 20**:** 763-772.
648 Greene, J. S., M. Brown, M. Dobosiewicz, I. G. Ishida, E. Z. Macosko *et al.*, 2016b Balancing
649 selection shapes density-dependent foraging behaviour. Nature 539**:** 254-258.
650 Greene, J. S., M. Dobosiewicz, R. A. Butcher, P. T. McGrath and C. I. Bargmann, 2016a
651 Regulatory changes in two chemoreceptor genes contribute to a *Caenorhabditis elegans*
652 QTL for foraging behavior. Elife 5.
653 Hobert, O., 2013 The neuronal genome of *Caenorhabditis elegans*. WormBook**:** 1-106.
654 Houck, L., C. Palmer, R. Watts, S. Arnold, P. Feldhoff *et al.*, 2007 A new vertebrate courtship
655 pheromone, PMF, affects female receptivity in a terrestrial salamander. Animal
656 Behaviour 73**:** 315-320.
657 Howe, K. L., B. J. Bolt, S. Cain, J. Chan, W. J. Chen *et al.*, 2016 WormBase 2016: expanding to
658 enable helminth genomic research. Nucleic acids research 44**:** D774-D780.
659 Jang, Y. H., H. S. Chae and Y. J. Kim, 2017 Female-specific myoinhibitory peptide neurons
660 regulate mating receptivity in *Drosophila melanogaster*. Nat Commun 8**:** 1630.
661 Jee, C., J. Lee, J. P. Lim, D. Parry, R. O. Messing *et al.*, 2013 SEB-3, a CRF receptor-like
662 GPCR, regulates locomotor activity states, stress responses, and ethanol tolerance in *C.*
663 *elegans*. Genes, brain, and behavior 12**:** 10.1111/j.1601-1183X.2012.00829.x.
664 Kaletsky, R., V. Lakhina, R. Arey, A. Williams, J. Landis *et al.*, 2016 The C. elegans adult
665 neuronal IIS/FOXO transcriptome reveals adult phenotype regulators. Nature 529**:** 92-96.
666 Käll, L., A. Krogh and E. L. L. Sonnhammer, 2004 A combined transmembrane topology and
667 signal peptide prediction method. Journal of molecular biology 338**:** 1027-1036.
668 Kalyaanamoorthy, S., B. Q. Minh, T. K. F. Wong, A. von Haeseler and L. S. Jermiin, 2017
669 ModelFinder: fast model selection for accurate phylogenetic estimates. Nature methods
670 14**:** 587-589.
671 Kamath, R. S., A. G. Fraser, Y. Dong, G. Poulin, R. Durbin *et al.*, 2003 Systematic functional
672 analysis of the *Caenorhabditis elegans* genome using RNAi. Nature 421**:** 231-237.
673 Katoh, K., and D. M. Standley, 2013 MAFFT multiple sequence alignment software version 7:
674 improvements in performance and usability. Molecular biology and evolution 30**:** 772-
675 780.
676 Kim, K., K. Sato, M. Shibuya, D. M. Zeiger, R. A. Butcher *et al.*, 2009 Two Chemoreceptors
677 Mediate Developmental Effects of Dauer Pheromone in *C. elegans*. Science 326**:** 994-
678 998.
679 Krishnan, A., M. S. Almén, R. Fredriksson and H. B. Schiöth, 2014 Insights into the Origin of
680 Nematode Chemosensory GPCRs: Putative Orthologs of the Srw Family Are Found
681 across Several Phyla of Protostomes. PLoS ONE 9**:** e93048.
682 Langmead, B., and S. L. Salzberg, 2012 Fast gapped-read alignment with Bowtie 2. Nature
683 methods 9**:** 357-359.

684 Lemoine, F., J. B. Domelevo Entfellner, E. Wilkinson, D. Correia, M. Dávila Felipe *et al.*, 2018
685 Renewing Felsenstein's phylogenetic bootstrap in the era of big data. Nature 556**:** 452-
686 456.
687 Li, B., and C. N. Dewey, 2011 RSEM: accurate transcript quantification from RNA-Seq data
688 with or without a reference genome. BMC bioinformatics 12**:** 323-323.
689 Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan *et al.*, 2009 The Sequence
690 Alignment/Map format and SAMtools. Bioinformatics (Oxford, England) 25**:** 2078-2079.
691 Ludewig, A. H., A. B. Artyukhin, E. Z. Aprison, P. R. Rodrigues, D. C. Pulido *et al.*, 2019 An
692 excreted small molecule promotes *C. elegans* reproductive development and aging. Nat
693 Chem Biol 15**:** 838-845.
694 Lupas, A., 1996 Prediction and analysis of coiled-coil structures. Methods in enzymology 266**:**
695 513-525.
696 Madeira, F., Y. M. Park, J. Lee, N. Buso, T. Gur *et al.*, 2019 The EMBL-EBI search and
697 sequence analysis tools APIs in 2019. Nucleic acids research 47**:** W636-W641.
698 McGrath, P. T., and I. Ruvinsky, 2019 A primer on pheromone signaling in *Caenorhabditis*
699 *elegans* for systems biologists. Curr Opin Syst Biol 13**:** 23-30.
700 McGrath, P. T., Y. Xu, M. Ailion, J. L. Garrison, R. A. Butcher *et al.*, 2011 Parallel evolution of
701 domesticated *Caenorhabditis* species targets pheromone receptor genes. Nature 477**:** 321-
702 325.
703 Narayan, A., G. Laurent and P. W. Sternberg, 2011 Transfer characteristics of a thermosensory
704 synapse in *Caenorhabditis elegans*. Proceedings of the National Academy of Sciences of
705 the United States of America 108**:** 9667-9672.
706 Narayan, A., V. Venkatachalam, O. Durak, D. K. Reilly, N. Bose *et al.*, 2016 Contrasting
707 responses within a single neuron class enable sex-specific attraction in *Caenorhabditis*
708 *elegans*. Proc Natl Acad Sci U S A 113**:** E1392-1401.
709 Nguyen, L.-T., H. A. Schmidt, A. von Haeseler and B. Q. Minh, 2015 IQ-TREE: a fast and
710 effective stochastic algorithm for estimating maximum-likelihood phylogenies.
711 Molecular biology and evolution 32**:** 268-274.
712 Noble, L. M., J. Yuen, L. Stevens, N. Moya, R. Persaud *et al.*, 2021 Selfing is the safest sex for
713 Caenorhabditis tropicalis. Elife 10.
714 Park, D., I. O'Doherty, R. K. Somvanshi, A. Bethke, F. C. Schroeder *et al.*, 2012 Interaction of
715 structure-specific and promiscuous G-protein-coupled receptors mediates small-molecule
716 signaling in *Caenorhabditis elegans*. PNAS 109**:** 9917-9922.
717 Peden, E. M., and M. M. Barr, 2005 The KLP-6 kinesin is required for male mating behaviors
718 and polycystin localization in *Caenorhabditis elegans*. Curr Biol 15.
719 Poole, R. J., E. Bashllari, L. Cochella, E. B. Flowers and O. Hobert, 2011 A Genome-Wide
720 RNAi Screen for Factors Involved in Neuronal Specification in *Caenorhabditis elegans*.
721 PLoS Genet 7**:** e1002109.
722 Powell, S., D. Szklarczyk, K. Trachana, A. Roth, M. Kuhn *et al.*, 2012 eggNOG v3.0:
723 orthologous groups covering 1133 organisms at 41 different taxonomic ranges. Nucleic
724 acids research 40**:** D284-D289.
725 Pungaliya, C., J. Srinivasan, B. W. Fox, R. U. Malik, A. H. Ludewig *et al.*, 2009 A shortcut to
726 identifying small molecule signals that regulate behavior and development in
727 *Caenorhabditis elegans*. Proceedings of the National Academy of Sciences of the United
728 States of America 106**:** 7708-7713.

729 Ragsdale, E. J., M. R. Muller, C. Rodelsperger and R. J. Sommer, 2013 A developmental switch
730 coupled to the evolution of plasticity acts through a sulfatase. Cell 155**:** 922-933.
731 Reilly, D. K., D. E. Lawler, D. R. Albrecht and J. Srinivasan, 2017 Using an Adapted
732 Microfluidic Olfactory Chip for the Imaging of Neuronal Activity in Response to
733 Pheromones in Male *C. Elegans* Head Neurons. Journal of Visualized Experiments**:**
734 e56026.
735 Reilly, D. K., E. J. McGlame, E. Vandewyer, A. N. Robidoux, C. S. Muirhead *et al.*, 2021
736 Distinct neuropeptide-receptor modules regulate a sex-specific behavioral response to a
737 pheromone. Communications Biology 4**:** 1018.
738 Reilly, D. K., L. J. Randle and J. Srinivasan, 2019 Evolution of hermaphroditism decreases
739 efficacy of Ascaroside#8-mediated mate attraction in *Caenorhabditis* nematodes.
740 microPublication Biology.
741 Robertson, H. M., and J. H. Thomas, 2006 The putative chemoreceptor families of *C. elegans*.
742 WormBook**:** 1-12.
743 Rual, J. F., J. Ceron, J. Koreth, T. Hao, A. S. Nicot *et al.*, 2004 Toward improving
744 *Caenorhabditis elegans* phenome mapping with an ORFeome-based RNAi library.
745 Genome Res 14**:** 2162-2168.
746 Sawin, E. R., R. Ranganathan and H. R. Horvitz, 2000 C. elegans locomotory rate is modulated
747 by the environment through a dopaminergic pathway and by experience through a
748 serotonergic pathway. Neuron 26**:** 619-631.
749 Schmitz, C., P. Kinge and H. Hutter, 2007 Axon guidance genes identified in a large-scale RNAi
750 screen using the RNAi-hypersensitive *Caenorhabditis elegans* strain *nre-1(hd20) lin-*
751 *15b(hd126)*. Proceedings of the National Academy of Sciences 104**:** 834-839.
752 Schwarz, E. M., M. Kato and P. W. Sternberg, 2012 Functional transcriptomics of a migrating
753 cell in Caenorhabditis elegans. Proceedings of the National Academy of Sciences of the
754 United States of America 109**:** 16246-16251.
755 Taylor, S. R., G. Santpere, A. Weinreb, A. Barrett, M. B. Reilly *et al.*, 2021 Molecular
756 topography of an entire nervous system. Cell 184**:** 4329-4347.e4323.
757 Tenor, J. L., and A. Aballay, 2008 A conserved Toll-like receptor is required for Caenorhabditis
758 elegans innate immunity. EMBO Rep 9**:** 103-109.
759 Thomas, C. G., R. Li, H. E. Smith, G. C. Woodruff, B. Oliver *et al.*, 2012 Simplification and
760 desexualization of gene expression in self-fertile nematodes. Curr Biol 22**:** 2167-2172.
761 von Reuss, S. H., N. Bose, J. Srinivasan, J. J. Yim, J. C. Judkins *et al.*, 2012 Comparative
762 Metabolomics Reveals Biogenesis of Ascarosides, a Modular Library of Small-Molecule
763 Signals in *C. elegans*. Journal of the American Chemical Society 134**:** 1817-1824.
764 von Reuss, S. H., and F. C. Schroeder, 2015 Combinatorial chemistry in nematodes: modular
765 assembly of primary metabolism-derived building blocks. Natural product reports 32**:**
766 994-1006.
767 Wang, J., R. Kaletsky, M. Silva, A. Williams, L. A. Haas *et al.*, 2015 Cell-Specific
768 Transcriptional Profiling of Ciliated Sensory Neurons Reveals Regulators of Behavior
769 and Extracellular Vesicle Biogenesis. Curr Biol 25**:** 3232-3238.
770 Wootton, J. C., 1994 Non-globular domains in protein sequences: automated segmentation using
771 complexity measures. Computers & chemistry 18**:** 269-285.
772 Zhang, Y. K., D. K. Reilly, J. Yu, J. Srinivasan and F. C. Schroeder, 2019 Photoaffinity probes
773 for nematode pheromone receptor identification. Journal of Organic & Biomolecular
774 Chemistry**:** 10.1039/C1039OB02099C.

775 Zhang, Y. K., M. A. Sanchez-Ayala, P. W. Sternberg, J. Srinivasan and F. C. Schroeder, 2017
776 Improved Synthesis for Modular Ascarosides Uncovers Biological Activity. Org Lett 19**:**
777 2837-2840.
778 Zwaal, R. R., J. E. Mendel, P. W. Sternberg and R. H. Plasterk, 1997 Two neuronal G proteins
779 are involved in chemosensation of the *Caenorhabditis* elegans Dauer-inducing
780 pheromone. Genetics 145**:** 715-727.

781

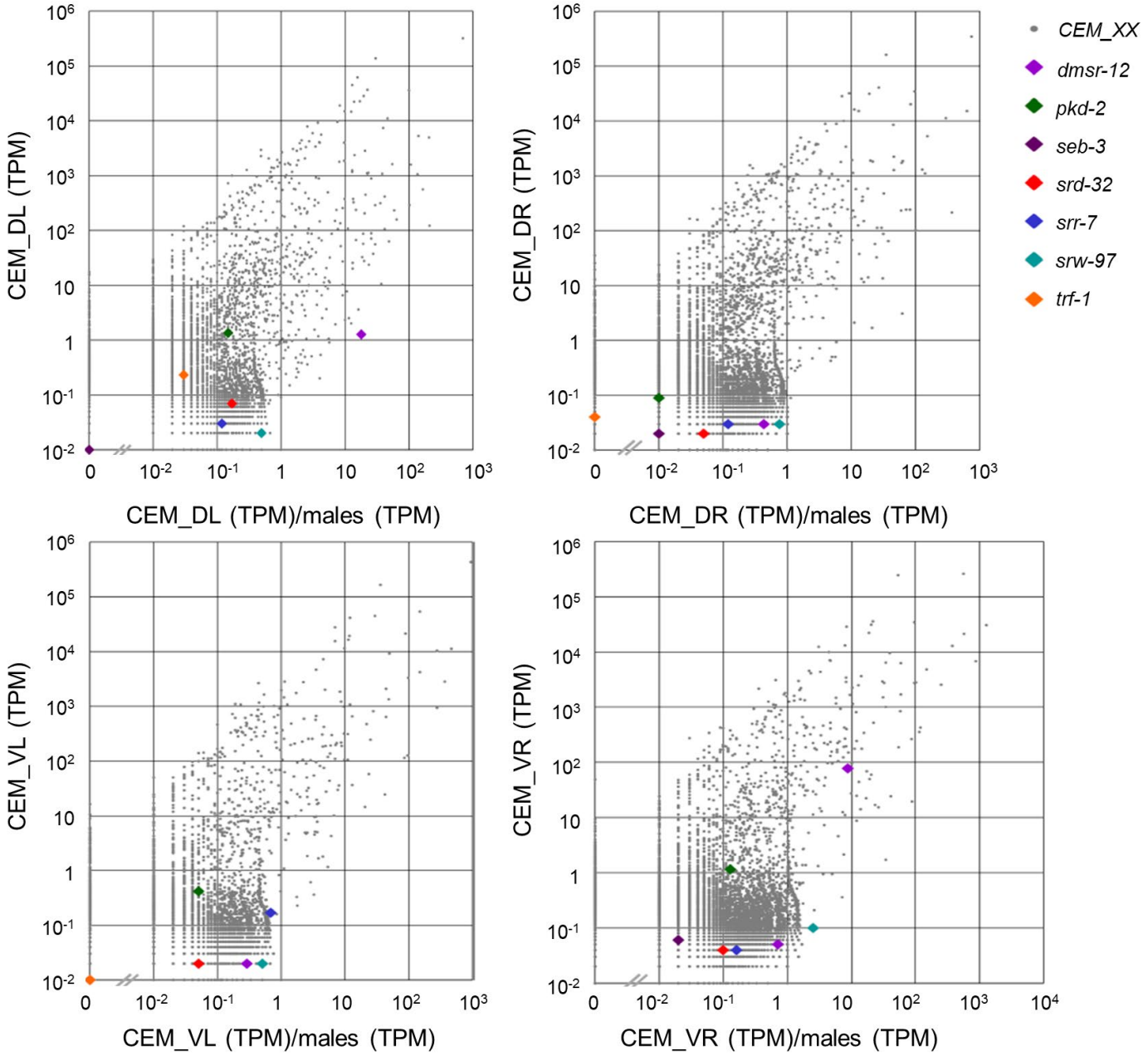
CEM_XX
dmsr-12
pkd-2
seb-3
srd-32
srr-7
srw-97
trf-1
CEM_DL (TPM)
CEM_DL (TPM)/males (TPM)
CEM_DR (TPM)
CEM_DR (TPM)/males (TPM)
CEM_VL (TPM)
CEM_VL (TPM)/males (TPM)
CEM_VR (TPM)
CEM_VR (TPM)/males (TPM)

Figure - 1

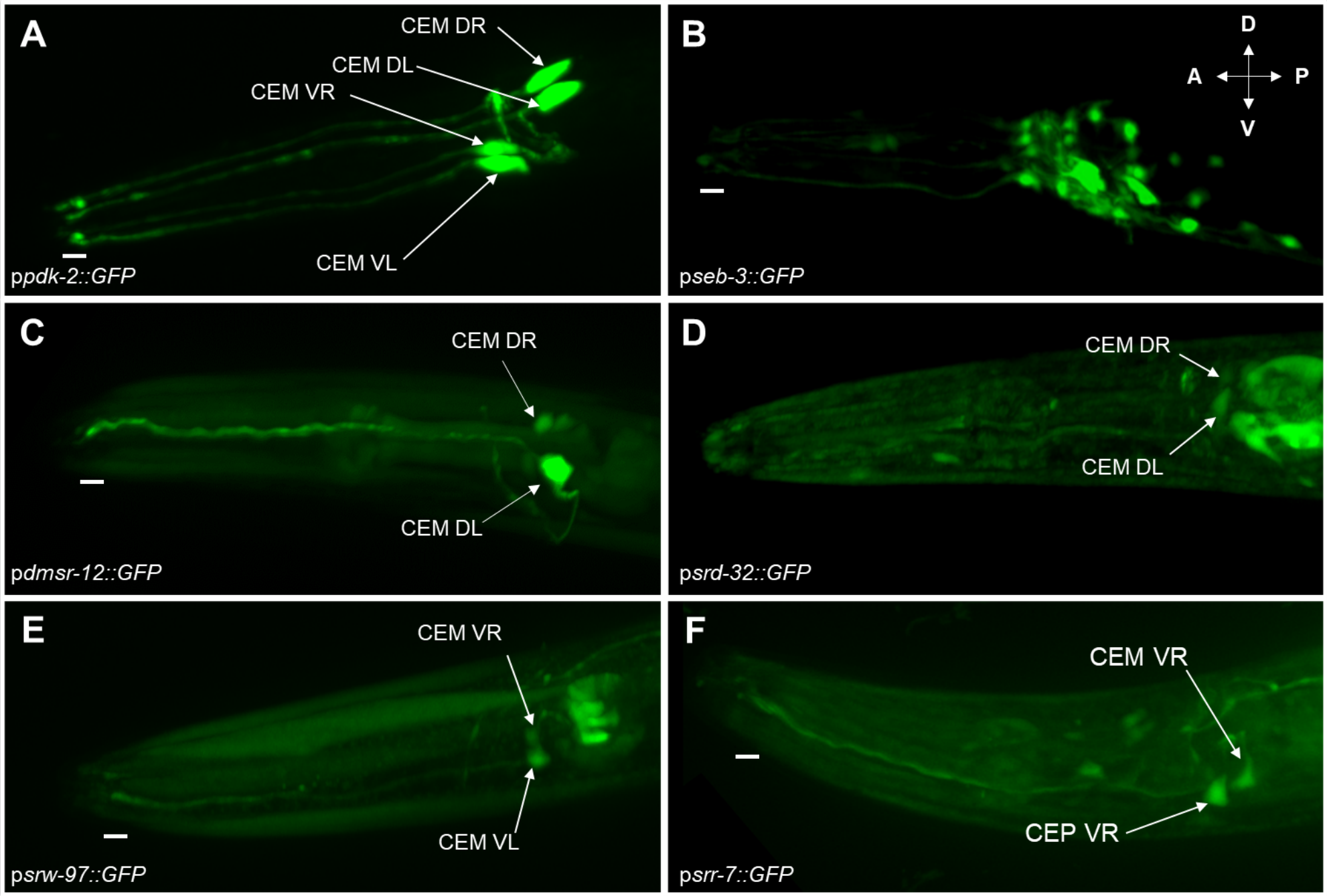
A
CEM DR
CEM DL
CEM VR
CEM VL
ppdk-2::GFP
B
D
A
P
V
pseb-3::GFP
C
CEM DR
CEM DL
pdmsr-12::GFP
D
CEM DR
CEM DL
psrd-32::GFP
E
CEM VR
CEM VL
psrw-97::GFP
F
CEM VR
CEP VR
psrr-7::GFP

Figure - 2

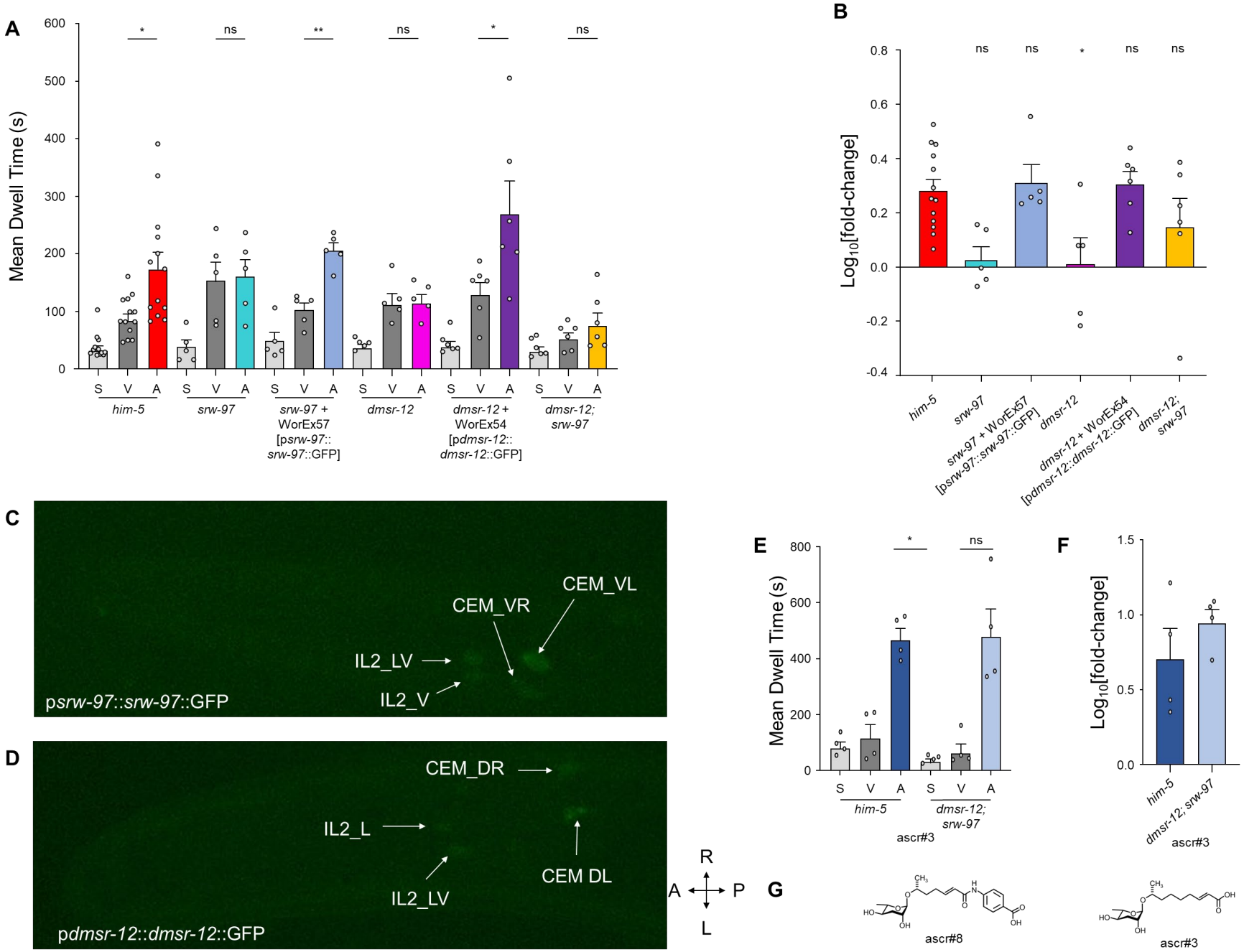
A
Mean Dwell Time (s)
600
500
400
300
200
100
0
*
ns
**
ns
*
ns
S V A S V A S V A S V A S V A S V A
him-5
srw-97
srw-97 + WorEx57 [psrw-97::srw-97::GFP]
dmsr-12
dmsr-12 + WorEx54 [pdmsr-12::dmsr-12::GFP]
dmsr-12; srw-97
B
Log10[fold-change]
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
ns
ns
*
ns
ns
him-5
srw-97
srw-97 + WorEx57 [psrw-97::srw-97::GFP]
dmsr-12
dmsr-12 + WorEx54 [pdmsr-12::dmsr-12::GFP]
dmsr-12; srw-97
C
CEM_VL
CEM_VR
IL2_LV
IL2_V
psrw-97::srw-97::GFP
D
CEM_DR
IL2_L
IL2_LV
CEM DL
pdmsr-12::dmsr-12::GFP
R
A
P
L
E
Mean Dwell Time (s)
800
600
400
200
0
*
ns
S V A S V A
him-5
dmsr-12; srw-97
ascr#3
F
Log10[fold-change]
1.5
1.0
0.5
0.0
him-5
dmsr-12; srw-97
ascr#3
G
ascr#8
ascr#3

Figure - 3

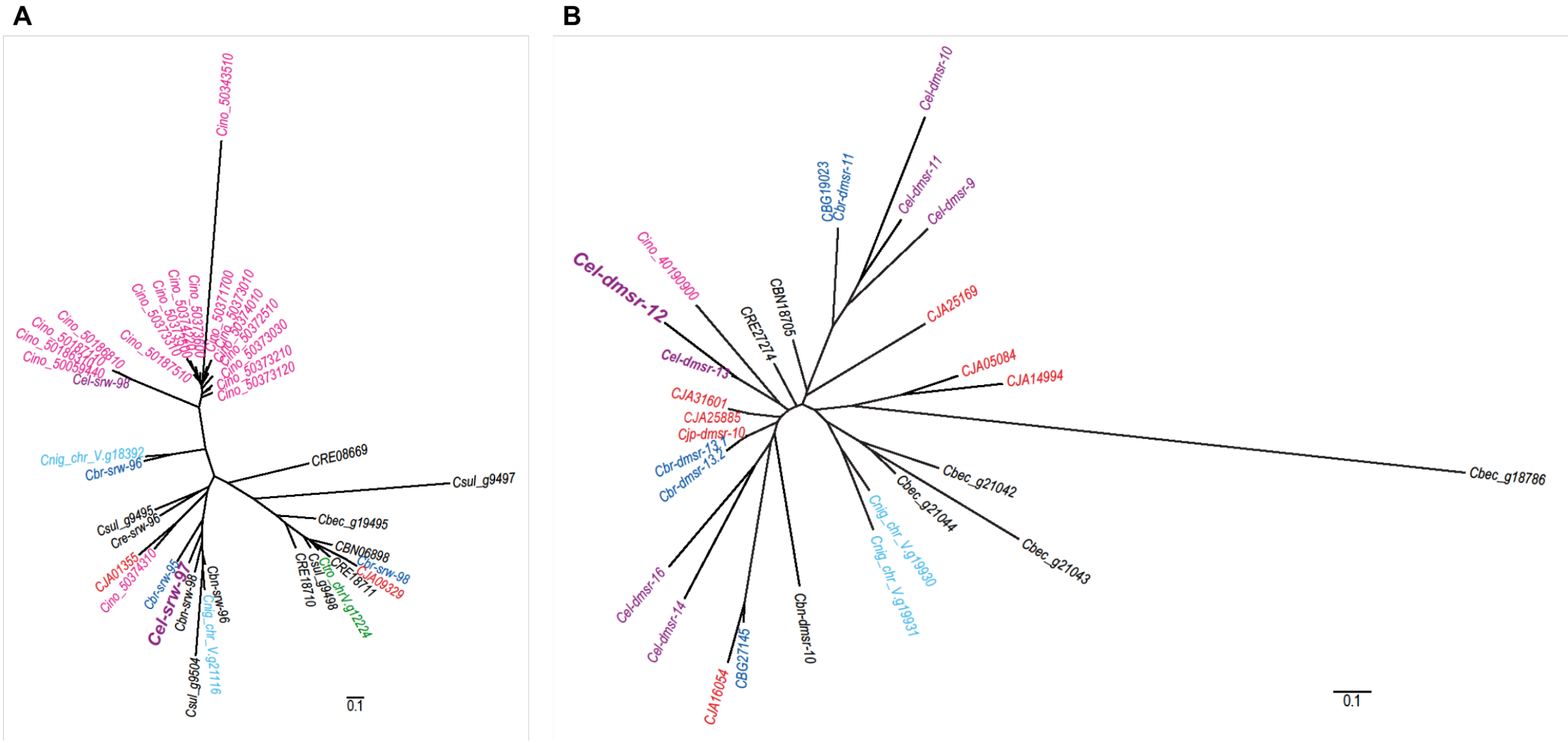
A
Cino_50343510
Cino_50371700
Cino_50373010
Cino_50374010
Cino_50372510
Cino_50373030
Cino_50373210
Cino_50373120
Cino_50187510
Cino_50186810
Cino_50059440
Cel-srw-98
Cnig_chr_V.g18392
Cbr-srw-96
CRE08669
Csul_g9497
Csul_g9495
Cre-srw-96
Cbec_g19495
CBN06898
Cbr-srw-98
CJA09329
CRE18711
Cino_chrV.g12224
Csul_g9498
CRE18710
CJA01355
Cino_50374310
Cbr-srw-95
Cel-srw-97
Cbn-srw-98
Cbn-srw-96
Cnig_chr_V.g21116
Csul_g9504
0.1
B
Cel-dmsr-10
CBG19023
Cbr-dmsr-11
Cel-dmsr-11
Cel-dmsr-9
Cino_40190900
Cel-dmsr-12
CBN18705
CRE27274
CJA25169
Cel-dmsr-13
CJA05084
CJA14994
CJA31601
CJA25885
Cjp-dmsr-10
Cbr-dmsr-13.1
Cbr-dmsr-13.2
Cbec_g18786
Cbec_g21042
Cbec_g21044
Cbec_g21043
Cnig_chr_V.g19930
Cnig_chr_V.g19931
Cel-dmsr-16
Cel-dmsr-14
Cbn-dmsr-10
CJA16054
CBG27145
0.1

Figure - 4