Searching for neuronal left/right asymmetry:
Genome wide analysis of nematode receptor-type guanylyl cyclases

Christopher O. Ortiz 1, John F. Etchberger 1, Shoshana L. Posy 1,2, Christian Frøkjær-Jensen 3, Shawn Lockery 3, Barry Honig 1,2 and Oliver Hobert 1,3

1 Howard Hughes Medical Institute
Department of Biochemistry and Molecular Biophysics
Columbia University Medical Center
701 W.168th Street
New York, NY 10032

2 Howard Hughes Medical Institute
Department of Biochemistry and Molecular Biophysics
Center for Computational Biology and Bioinformatics
Columbia University Medical Center
1130 St. Nicholas Avenue, Room 815
New York, NY 10032

3 Institute of Neuroscience
University of Oregon
Eugene, OR 97403

3 Corresponding author at or38@columbia.edu

9 Figures, 4 Tables, 2 Supplementary Figures

Running title: Guanylyl cyclases in nematodes
Key Words: guanylyl cyclase, C. elegans, sensory neurons, left/right asymmetry
ABSTRACT

Functional left/right asymmetry ("laterality") is a fundamental feature of many nervous systems, but only very few molecular correlates to functional laterality are known. At least two classes of chemosensory neurons in the nematode *Caenorhabditis elegans* are functionally lateralized. The gustatory neurons ASE Left (ASEL) and ASE Right (ASER) are two bilaterally symmetric neurons which sense distinct chemosensory cues and express a distinct set of four known chemoreceptors of the guanylyl cyclase (*gcy*) gene family. To examine the extent of lateralization of *gcy* gene expression patterns in the ASE neurons, we have undertaken a genome wide analysis of all *gcy* genes. We report the existence of a total of 27 receptor-type *gcy* genes and of seven soluble *gcy* genes in *C. elegans*. We describe the expression pattern of all previously uncharacterized receptor-type *gcy* genes and find them to be highly biased but not exclusively restricted to the nervous system. We find that more than 41% (11/27) of all *gcy* genes are expressed in the ASE gustatory neurons and that one third of all *gcy* genes (9/27) are expressed in a lateral, left/right asymmetric manner in the ASE neurons. The expression of all laterally expressed *gcy* genes is under the control of a gene regulatory network composed of several transcription factors and miRNAs. The complement of *gcy* genes in the related nematode *C. briggsae* differs from *C. elegans* as evidenced by differences in chromosomal localization, number of *gcy* genes and expression patterns. Differences in *gcy* expression patterns in the ASE neurons of *C. briggsae* arise from a difference in *cis*-regulatory elements and *trans*-acting factors that control ASE laterality. In sum, our results indicate the existence of a surprising multitude of putative chemoreceptors in the gustatory ASE neurons and suggest the existence of a substantial degree of laterality in gustatory signaling mechanisms in nematodes.
INTRODUCTION

The diversification of neuronal fate and function across the left/right axis of nervous systems is poorly understood but represents a fundamental problem in the neurosciences. This problem is well illustrated by a cursory comparison of structure and function of nervous systems. While the organization of nervous systems is largely bilaterally symmetric on a morphological level, brain functions are often highly lateralized (HUGDAHL and DAVIDSON 2003). Functional lateralization is presumably brought about by the diversification of neuronal function on a subanatomical level, such as differential gene expression in bilaterally symmetric structures. Indeed, quantitative comparison of transcript levels have recently revealed left/right asymmetries in gene expression profiles in the human brain (SUN et al. 2005).

The nematode C. elegans provides a simple model organism to study the lateralization of nervous system function (HOBERT et al. 2002). Such lateralization can be observed in the chemosensory system of the nematode. The best studied chemosensory neurons are a group of 12 classes of neurons called the amphid sensory neurons (Fig.1A). Each class consists of one pair of two bilaterally symmetric and morphologically indistinguishable neurons, most of which are chemosensory neurons (Fig.1B). At least two classes of these chemosensory neurons, the AWC odor sensory neuron class and the ASE gustatory neuron class are functionally lateralized, allowing the animal to sense and discriminate different sensory cues with the left and right neuron (PIERCE-SHIMOMURA et al. 2001; WES and BARGMANN 2001)(Fig.1B). Functional lateralization of AWCL/R and ASEL/R correlates with the left/right asymmetric expression of putative chemoreceptors (JOHNSTON et al. 2005; TROEMEL et al. 1999; YU et al. 1997) but the extent of lateralization of chemoreceptor gene expression in these neurons is still unclear. In the AWCL/R sensory neurons, only one left/right asymmetrically expressed chemoreceptor has been reported, a 7-transmembrane receptor (TROEMEL et al. 1999)(Fig.1B). In the ASEL/R sensory neurons, a total of four asymmetrically expressed, putative chemoreceptors were known prior to this study, which all belong to the family of receptor guanylyl cyclases, encoded by the gcy genes (JOHNSTON et al. 2005; YU et al. 1997)(Fig.1B). gcy-5 and gcy-22 are expressed in the right ASE neuron (ASER), whereas gcy-6 and gcy-7 are expressed exclusively in the
left ASE neuron (ASEL).

To further analyze the extent of lateralization of the ASE gustatory neurons, we identified the complete set of *gcy* genes in the *C. elegans* genome and undertook a genome wide analysis of their expression patterns. Previous counts of *C. elegans* receptor-type *gcy* genes were preliminary given the incomplete nature of the *C. elegans* genome sequencing project, but estimated to be in the higher twenties (BIRNBY et al. 2000; YU et al. 1997). Expression patterns had been determined for eight receptor-type *gcy* genes (BIRNBY et al. 2000; L’ETOILE and BARGMANN 2000; YU et al. 1997). We now report the final count of receptor-type guanylyl cyclases in the complete *C. elegans* genome to be 27. We present a comparative sequence analysis of all *gcy* genes and describe the expression patterns of all previously uncharacterized receptor-type *gcy* genes using *gfp* reporter gene fusions. We analyze the mechanisms of regulation of *gcy* gene expression in the context of the ASE gustatory neurons, we investigate the consequence of removal of one ASE-expressed *gcy* gene on ASE neuron function and we examine the evolutionary divergence of *gcy* gene structure and expression.
MATERIALS AND METHODS

Strains and transgenes

Wild-type and mutant strains: N2 wild-type Bristol isolate; C. briggsae AF16 wild-type strain; OH4349 lsy-6(ot71)dpy-11(e224); OH110 lim-6(nr2073) (HOBERT et al. 1999). RB1000 gcy-5(ok921), 4x outcrossed; RB1010 gcy-5(ok930), 4x outcrossed; OH2957 gcy-5(tm897), not outcrossed.

Transgenes: otIs3: Is[gcy-7prom::gfp; lin-15(+)] (CHANG et al. 2003); this transgene is expressed in ASEL and the excretory canal cell. otIs151: Is[ceh-36prom::dsRed2; rol-6(d)] (JOHNSTON and HOBERT 2003); this transgene is expressed in ASEL/R and AWCL/R. otIs133: Is[txx-3promB::rfp; pNC4.2(unc-4(+))] (WENICK and HOBERT 2004); this transgene is expressed in AIYL/R. oyls17: Is[gcy-8prom::gfp; lin-15(+)], expressed in AFDL/R, and oyls51: Is[srh-142prom::rfp; lin-15(+)], expressed in ADFL/R, both gifts from Piali Sengupta.

Sequence analysis

To identify GCY sequences, the sets of predicted proteins for C. elegans and C. briggsae were obtained from the Sanger Institute (http://www.sanger.ac.uk/Projects/C_elegans/WORMBASE/current/wormpep_download.shtml and ftp://ftp.sanger.ac.uk/pub/wormbase/cbriggsae/cb25.agp8/). Representative GCYs were used as PSI-BLAST queries to search the two proteomes. The HMMER 2.3.2 package (EDDY 1998) was used to construct a hidden Markov model from an alignment of GCY catalytic domains and search for additional GCYs.

The intracellular regions of the transmembrane GCYs and the complete soluble GCY sequences were aligned with T-coffee version 2.03 (NOTREDAME et al. 2000). Maximum parsimony phylogenetic trees were found via heuristic search with PAUP* version 4.0 beta 10 (SWOFFORD 2003). The trees were generated in 100 repeated searches with random addition of taxa to obtain the starting tree. Robustness of the tree partitions was evaluated by constructing a bootstrap consensus tree with 1000 replicates. The trees were visualized with TreeView version 1.6.6 (PAGE 1996).
Nomenclature of gcy genes and gcy gene predictions

Most but not all gcy names were previously assigned (http://www.wormbase.org). We named two previously unnamed gcy genes gcy-28 (T01A4.1) and gcy-29 (C04H5.3). Both code for receptor-type proteins. A few gcy genes have been double-named in the past. The most current names are (old names in parenthesis): gcy-18 (= gcy-26), gcy-20 (= gcy-16), gcy-17 (= gcy-24), odr-1 (= gcy-10), gcy-28 (gcy-38 in WS149), gcy-29 (gcy-39 in WS149).

Since we detected several cases where individual parts of one C. elegans gene are homologous to separate, adjacent predicted C. briggsae genes (gene prediction in Wormbase WS149), we suspected that C. briggsae genes may have been incorrectly predicted. We therefore ran the FGENESH program at www.softberry.com (SALAMOV and SOLOVYEV 2000) on chromosomal regions that contained the following predicted C. briggsae genes (from WS149): CBG07423(CBP15915) + CBG07424(CBP15916) + CBG07425(CBP15917), CBG20867(CBP04902) + CBG20868(CBP04903) and CBG19454(CBP11205) + CBG19453(CBP11204). In each of these three cases, we found that FGENESH predicted only one gene, whose product was homologous over its entire length with putative C. elegans orthologs. However, in two cases, the revised gene prediction overlooked exons predicted in the original prediction which contained homology to GCY proteins. We therefore assembled alternate gene prediction by hand, based on the similarity to known GCY proteins. We named these revised predictions CBP15915*, CBP04902* and CBP11205* (see Supp.Fig.1 and 2). Based on primary sequence homology, we also suspect that one of the two C. briggsae orthologs of C. elegans gcy-35 has been incorrectly predicted as two separate genes CBG20390 (CBP04780) and CBG20392 (CBP04781). Similarly, the CBG10472 (CBP08561) and CBG10474 (CBP08562) likely constitute one gene. However, neither of these two suspicions could be corroborated by FGENESH.

Generation of gfp reporter gene fusions, transgenic animals and identification of reporter gene expressing cells.

Most reporter genes were created by PCR fusion (HOBERT 2002) and some were generated by subcloning PCR amplicons into pPD95.75 (see Table 1). Primer
sequences and resulting transgenic arrays are shown in Table 1. DNA was injected at
\( \sim 10-50 \, \text{ng/\mu l} \) using either oun-122::gfp or rol-6 as injection marker. Cell identifications
were done based on overall cell position and morphology and was significantly aided by
the uses of the following four co-labeling procedures. (1) Most gfp reporters were
injected into animals carrying the otIs151 transgene in which ASEL/R and AWCL/R are
labeled with dsRed2. Since the otIs151 transgene already contains the rol-6(d) injection
marker we used uncb-122::gfp (LORIA et al. 2004) as injection marker for most injections.
This marker is expressed in coelomocytes but also yields occasional and mosaic gfp
expression within the pharynx, often in the I5 neuron. With the exception of the broadly
pharyngeal expression of one gcY gene, we therefore ignored any cell-type specific
pharyngeal expression of gcY reporter genes. Injections into wild-type C. elegans or C.
briggsae were done using rol-6(d) as injection marker. (2) Some gcY::gfp transgenic
animals were crossed with animals carrying the otIs133 transgene in which the AIY
interneurons are marked with dsRed2, thereby facilitating cell identification either by
red/green overlap (gcY-1 prom::gfp) or by the determination of relative cell position (gcY-
18 prom::gfp). (3) In many cases, a subset of the amphid neuron classes of gcY::gfp
transgenic animals was filled with DiI. DiI fills the ASK, ADL, ASI, AWB, ASH and ASJ in
the head in the PHA and PHB neuron classes in the tail (HEDGE-COCK et al. 1985),
thereby facilitating cell identification either by an overlap of red and green fluorescence
or by the determination of relative cell position. DiI was dissolved in DMF, diluted to 10
\( \mu \text{g/ml} \) in M9 or in ddH20 plus 50 mM calcium-acetate to additionally fill the IL2 neurons.
Worms were soaked in the DiI solution for at least 1 hour. (4) Expression of gcY genes
in the AFD thermosensory neurons was assessed using the oYIs17 transgene in which the AFD sensory neurons are marked with GFP. Extrachromosomal arrays carrying
gcY-18 prom::gfp and gcY-23 prom::gfp were crossed with oYIs17-containing animals and
the number of GFP expressing amphid sensory neurons was counted. To exclude
expression in the closely adjacent ADF neuron class, gfp transgenes were also crossed
with an ADF-expressed dsRed2 reporter construct, oYIs52, kindly provided by P.
Sengupta.

Chemotaxis assays
Radial population chemotaxis assays were done as previously described (Chang et al. 2004). Assay plates were 10 cm tissue culture dishes containing 20 g/L agar, 5 mM potassium phosphate (pH = 6.0), 1 mM CaCl₂, 1 mM MgSO₄. To set up the chemical gradients on the assay plates a 10 µL drop of attractant was placed 15 mm from the edge of the plate at the “attractive spot”. A 10 µL drop of ddH₂O was placed diametrically opposite and was considered the “negative control spot”. The attractant was allowed to diffuse for 14-16 hours at room temperature. To increase the steepness of the chemical gradient, 4 to 4.5 hours prior to chemotaxis assay 4 µL attractant was added to the “attractive spot” and 4 µL ddH₂O added to the “negative control spot”. The attractants NaCl and NH₄Cl (Sigma, MO, USA) were dissolved in ddH₂O to a concentration of 2.5 M and were adjusted to pH 6.0 with either NH₄OH or acetic acid. Worms were washed three times in sterile water to remove food and salts. Worms were then placed at the center of the plate and allowed to chemotax for an hour. Worms reaching either the attractant peak or the negative control spot (sterile water) were immobilized with sodium azide. Results were quantified by counting worms that were located at (A) the attractant, (B) the center of the plate or (C) the negative control. Since animals carrying two of the three gcy-5 knockout alleles (ok921 and ok930) did not disperse well from the center of the plate, we calculated a modified chemotaxis index, defined as C.I. = A/(A+C). This index therefore disregards worms that do not reach either the attractant or the negative control spot. It is doubtful that the dispersion defects of ok921 and ok930 alleles are significant since the putative null allele tm897 does not show these defects.
RESULTS

Identification of the complete set of GCY proteins in *C. elegans*

To identify the complete set of guanylyl cyclases in *C. elegans* we employed the Position-Specific Iterated Basic Local Alignment Search Tool (PSI-BLAST) to search the latest release of the complete *C. elegans* genome databases, using a set of known GCY proteins as queries. We identified a total of 34 *gcy* genes, several more than reported in previous searches of incomplete genome sequence databases (Bargmann 1998; Birnby et al. 2000; Yu et al. 1997). The identified *gcy* genes fall into two distinct families, receptor type guanylyl cyclases (27 genes) and soluble cyclases (7 genes). Both families contain a guanylyl cyclase catalytic domain (Pfam domain PF00211; Fig.2; Supp.Fig.1). In several GCY proteins catalytically important residues are not conserved in the cyclase domain (Supp.Fig.1) and it has been speculated that in these cases, heterodimerization with a catalytically active GCY protein ensures activity of the dimer (Morton 2004). Apart from the presence of the cyclase domain, receptor-type and soluble GCY proteins differ significantly (Fig.2). Soluble GCY proteins contain one other characteristic domain, a HNOB domain (for “heme NO binding” domain Pfam domain PF07700). In contrast, receptor-type guanylyl cyclases lack this HNOB domain but always contain an additional protein kinase domain (PF00069), which is likely to be inactive since it lacks a critical catalytic aspartate residue that is present in the catalytic core of all active protein kinases (“HRD motif”; Supp.Fig.2). In addition to this protein kinase-like domain, all but one of the *C. elegans* receptor-type GCY proteins contain a single transmembrane domain and a signal sequence. Moreover, the majority of receptor-type GCY proteins contain a conserved extracellular domain of unknown function that is also present in many mammalian GCY proteins and conserved in amino acid receptors in bacteria (“RFLBP” = Receptor family ligand binding region; Pfam PF01094; Fig.2). The presence of this domain strongly suggests that the *C. elegans* GCY proteins are indeed ligand-binding receptor proteins.

The overall domain topology of *C. elegans* GCY proteins is similar to that of mammalian GCY proteins. Multiple transmembrane-containing GCY proteins which can be found in unicellular eukaryotes (Wedel and Garbers 2001) are not present in *C.
elegans.

GCY-27 is an unusual receptor-type GCY protein (Fig.2). While containing all the intracellular signature motifs of receptor-type GCYs, the predicted GCY-27 protein is several hundred amino acids shorter than all other predicted receptor-type GCY proteins and does not contain a predicted signal sequence (SS), transmembrane domain (TM) or other extracellular motifs (Fig.2). In the absence of complete cDNA/EST sequences, we cannot rule out a gene prediction error, but we consider the failure to predict a complete extracellular domain to be unlikely for two reasons: (a) The upstream gene adjacent to gcy-27 is relatively close by (Fig.4), leaving little room for such a prediction oversight, particularly since in other gcy genes the extracellular domains are large and composed of many exons; (b) a C. briggsae ortholog (described in more detail in a later section) also lacks the extracellular domain. While receptor-type GCY proteins that lack a TM and SS domain have been described before (MORTON 2004), they do not appear to contain the protein kinase domain that is present in GCY-27 and all other receptor-type GCY proteins. It is interesting to note that GCY-27 is most closely related to the intracellular domain of ODR-1 (Fig.3A), a transmembrane GCY protein whose extracellular domain (which contains no canonical RFLBP domain) has previously been shown to be dispensable for its function in chemosensory signal transduction (L'ETOILE and BARGMANN 2000).

We analyzed the relatedness of soluble and receptor-type GCY proteins in more detail by generating a maximum parsimony bootstrap tree using the intracellular domain of the receptor-type GCYs and the complete sequences of the soluble GCYs. The soluble GCYs branch separately from the transmembrane-type proteins and are clustered into two subgroups. Many receptor-type GCY proteins also fall into small and well-defined subgroups (Fig.3A).

Notably, in a substantial number of cases, the degree of sequence relation correlates with proximity in the genome sequence (Fig.3B). For example, all six members of one subgroup ("gcy-5 subgroup") are located within a ~7 Mb interval on chromosome II, five of these (gcy-1 through gcy-5) map to an ~800 Kb interval and three of them are directly adjacent genes, separated by only a few hundred base pairs (Fig.3B, Fig.4). This suggests that these genes arose by relatively recent gene
duplications. Likewise, all seven members of what we term the “gcy-7 subgroup” of gcy genes (Fig.3A) reside on a single chromosome and five of them in a <8 Mb interval (Fig.3B, Fig.4). As we will demonstrate below, the degree of sequence relation and chromosomal location also correlates with similarities in gene expression patterns.

**Expression patterns of receptor-type gcy genes.**

While the expression of all seven soluble gcy genes has already been described (Cheung et al. 2004; Gray et al. 2004; Yu et al. 1997), the expression patterns of only eight of the 27 receptor-type gcy genes was previously reported, six in the context of a preliminary and incomplete genome analysis (gcy-5, gcy-6, gcy-7, gcy-8, gcy-10/odr-1, gcy-12 and gcy-22) and two in the course of a functional analysis (odr-1 and daf-11) (Birnby et al. 2000; Johnston et al. 2005; L'Etoile and Bargmann 2000; Yu et al. 1997) (summarized in Table 2). We generated gfp reporter fusions to the putative cis-regulatory regions of the remaining 19 gcy genes (schematically shown in Fig.4). In all except one case, sequences to the next upstream genes were included in the reporter gene fusions (Fig.4). While such upstream sequences are most likely to harbor gene regulatory elements, it needs to be kept in mind that additional regulatory elements may be located elsewhere and that therefore the expression patterns of the gfp reporter genes can only provide a first approximation of endogenous gene expression profiles. Since gcy genes are likely to primarily function as chemoreceptors in the mature nervous system, we restricted our gene expression analysis of transgenic animals that harbor the respective reporter genes to larval and adult stages. To allow us to reliably detect expression in the ASEL/R gustatory neurons, the neurons in which our laboratory is most interested in, we used a transgenic reporter array in the background in which the ASEL/R neurons are labeled with dsRed2 (otIs151; see Material and Methods), thereby enabling us to score for an overlap between green and red fluorescent signals. In most cases, multiple lines were analyzed per construct (see Table 1) and we did not observe any notable differences between individual lines. Reporter gene expression was observed for each of the gcy genes examined. Expression patterns are shown in Fig.5 and summarized in Table 2, which also lists previously described expression patterns. The expression patterns of all receptor-type gcy genes can be summarized as
follows.

(1) Broad versus cell-type specific: With the exception of the very broadly but not ubiquitously expressed \( gcy-28 \) gene, all receptor-type \( gcy \) genes are expressed in a tissue- and cell-type restricted manner (Table 2; Fig.5).

(2) Cell types: The expression of receptor-type \( gcy \) genes is strongly biased toward the nervous system. With the exception of two non-neuronally expressed \( gcy \) genes (\( gcy-9 \) and \( gcy-11 \)), all \( gcy \) genes are expressed in a restricted subset of neurons. 21/27 genes are exclusively expressed in the nervous system and 4/27 genes are expressed in restricted sets of both neuronal and non-neuronal cells (Table 2; Fig.5).

(3) Expression within the nervous system: Within the nervous system, most but not all \( C. \) \textit{elegans} \( gcy \) genes are expressed in sensory neurons. All except two pairs (ASH and ADF) of the 12 amphid sensory neuron classes shown in Fig.1A express at least one \( gcy \) gene. Unlike what the previous analyses of eight \( gcy \) genes seemed to indicate (BIRNBY et al. 2000; JOHNSTON et al. 2005; L'ETOILE and BARGMANN 2000; YU et al. 1997), neuronal \( gcy \) gene expression is, however, not restricted to sensory neurons. Non-sensory neurons that express \( gcy \) genes include the AIY, AIM, AVK, RIA and PVT interneuron classes and the RIM neuron, which is both a motorneuron and interneuron. In addition, as mentioned above, many neurons in the nervous system, including, for example, ventral cord motor neurons, express the widely-expressed \( gcy-28 \) gene. As summarized in Table 3, 25/27 of the \( C. \) \textit{elegans} receptor-type \( gcy \) genes are expressed in the nervous system, 25/27 are expressed in various types of sensory neurons plus other neurons, 15/27 are expressed exclusively in sensory neurons and 9/27 are restricted to single neuron classes. The expression of mammalian \( gcy \) genes, of which there are only 7, show roughly comparable patterns; some are expressed in non-neuronal cells and those that are expressed in neurons are strongly biased to expression in sensory structures (WEDEL and GARBERS 2001).

(4) Co-expression: A notable general feature of \( C. \) \textit{elegans} \( gcy \) gene expression profiles, both receptor- and non-receptor-type, is that a small number of neuron classes co-express a substantial number of \( gcy \) genes (Fig.5; summarized in Table 4). The most striking examples are the ASE gustatory neuron class, which expresses a total of 11
gcy genes, more than one third of all receptor-type gcy genes. In addition, six gcy genes are co-expressed in the ASI chemosensory neurons, five gcy genes are co-expressed in AWC olfactory neurons, four are co-expressed in the AFD thermosensory neurons and two are co-expressed in the ASG and PHA phasmid sensory neurons, respectively (Table 4). Besides the overlap in individual neurons, there are also a few examples of gcy genes that show similar combinations of cellular expression profiles. gcy-7 and gcy-20 are co-expressed in ASEL and in the excretory canal cell. Moreover, as previously reported, daf-19 and odr-1 are expressed in precisely the same subset of amphid sensory neurons and six soluble gcy genes are co-expressed in the AQR/PQR and URX neurons (BIRNBY et al. 2000; CHEUNG et al. 2004; GRAY et al. 2004; L’ETOILE and BARGMANN 2000; YU et al. 1997). A single receptor-type gcy gene, gcy-25, complements the expression of the six soluble gcy genes in the AQR/PQR and URX neurons (Fig.5O; Table 4).

(5) Left/right asymmetric expression: Two previous studies identified a total of four left/right asymmetrically expressed gcy genes in the ASE gustatory neuron class (gcy-5, gcy-6, gcy-7 and gcy-22; Table 2)(JOHNSON et al. 2005; YU et al. 1997). We have identified five more examples of gcy genes that are expressed in a left/right asymmetric manner in the ASE neuron class. gcy-1, gcy-3 and gcy-4 are expressed predominantly or exclusively in ASER, while gcy-14 and gcy-20 are expressed predominantly or exclusively in ASEL (Fig.5; data is quantified in Fig.6). Together with the previously reported expression patterns, a total of four gcy genes are lateralized to ASEL and five gcy genes are lateralized to ASER (summarized in Fig.1B). One third (9/27) of all gcy genes are therefore laterally expressed in the ASE neurons. There are subtle differences in the degree of laterality of ASE-expressed gcy genes. Five out of the nine L/R asymmetric gcy genes are exclusively expressed in ASER (gcy-1, gcy-5, gcy-22) or ASEL (gcy-6, gcy-7) (JOHNSON et al. 2005)(Fig.6A). The other four gcy genes are only biased to ASER (gcy-3 and gcy-4) or ASEL (gcy-14 and gcy-20)(Fig.6B); for example, while expression of gcy-4 is almost always stronger in ASER than in ASEL, there is often faint, but visible expression in ASEL (Fig.6B). Since C. elegans transgenes harbor multiple copies of reporter gene constructs, the relevance of such relatively subtle quantitative details is difficult to assess. However, we note that each
individual array of each reporter constructs behaves in a similar way (see Table 1 for number of arrays scored), arguing that these observations are not due to transgene variance.

*gcy* genes that are asymmetrically expressed in either ASER or ASEL are not asymmetrically expressed in other sensory neurons. For example, the *gcy-1* and *gcy-3* genes, two ASER-expressed genes, are expressed bilaterally in other pairs of neurons. In addition to the asymmetrically expressed *gcy* genes, we found two bilaterally, albeit weakly expressed *gcy* genes in ASEL/R, *gcy-19* and *gcy-29*, increasing the fraction of ASE-expressed *gcy* genes to 41% (11/27) of all *gcy* genes. The possible lack of regulatory elements in *gfp* reporter genes may lead to the oversight of perhaps even more ASE-expressed *gcy* genes.

We have not observed any other obvious left/right asymmetric *gcy* gene patterns in bilaterally symmetric neurons, including the AWCL/R neurons (which we could easily identify with the dsRed2-expressing *otIs151* transgene). The AWCL/R neurons are the only other known neuron pair displaying functional laterality (WES and BARGMANN 2001) and although they express five *gcy* genes (Table 4), none of them is obviously lateralized.

(6) *Similarity of chromosomal position, primary sequence and gene expression patterns.* Strikingly, the co-expression of *gcy* genes in ASEL or ASER correlates extensively with the primary sequence similarity and chromosomal location of the *gcy* genes. The ASER-expressed *gcy-1, gcy-3, gcy-4* and *gcy-5* genes fall into one sequence subgroup (Fig.3A), localize within ~800 Kb and two of them are directly adjacent to one another (Fig.3B and Fig.4). Notably, their co-expression is not simply due to a joint *cis*-regulatory element; as our reporter constructs clearly demonstrate (Fig.4), each gene contains separable *cis*-regulatory elements.

Similarly, all four ASEL-expressed *gcy* genes (*gcy-6, gcy-7, gcy-14, gcy-20*) fall into one sequence subgroup (Fig.3A) and localize at the center of chromosome V (Fig.3B). These observations argue that gene duplication events that lead to the generation of these paralogues also duplicated their regulatory regions. The only exception to this pattern is the *gcy-22* gene, which is predominantly expressed in ASER, but in terms of primary sequence identity clusters more closely with the ASEL-
expressed genes than with the ASER expressed genes (Fig.3).

A similar correlation of sequence relation and gene expression can also be observed in non-ASE expressed gcy genes. The gcy-8, gcy-18 and gcy-23 genes are related by sequence (Fig.3A), localize within a 15.6 Mb interval (Fig.3B), and are all co-expressed in the AFD sensory neurons (Fig.5). Additionally, the AFD-expressed gcy-29 gene is also closely related to gcy-8, gcy-18 and gcy-23 by primary sequence, but it localizes to a different chromosome. Lastly, the sequence-related gcy-15 and gcy-21 genes (Fig.3A) are in close chromosomal proximity (~530 Kb; Fig.3B) and are co-expressed in the ASG chemosensory neurons (Fig.5).

Taken together, all these similarities indicate that related gcy genes arose by local gene duplication events that not only duplicated the protein-coding region but also their cis-regulatory information.

Regulation of left/right asymmetric expression of gcy genes

How is the laterality of the newly characterized ASEL/R-expressed gcy genes controlled? We have previously identified a complex network of transcription factors and miRNAs that control ASE laterality after animals have passed through an initial hybrid precursor state (Chang et al. 2004; Chang et al. 2003; Johnston and Hobert 2003; Johnston et al. 2005; Johnston and Hobert 2005). These regulatory factors fall into two broad categories; first, factors that control the activity of a bistable feedback loop that determines whether an ASE neuron adopts the ASEL or ASER state and, second, factors that act outside the regulatory loop to determine specific subsets of terminal differentiation features. To test whether the newly identified, asymmetrically expressed gcy genes are subject to regulation by this network or controlled by a different set of regulatory factors, we analyzed gfp reporter expression profiles in two null mutant backgrounds that are representative for each category, Isy-6 and lim-6. Animals that lack the Isy-6 miRNA display a complete switch from the ASEL fate to the ASER fate (Johnston and Hobert 2003) and animals that lack the lim-6 LIM homeobox gene fail to activate a defined subset of ASEL features and fail to repress a subset of ASER features (Hobert et al. 1999; Johnston et al. 2005). We find that all newly identified gcy genes are components of the ASEL and ASER state that is controlled by the Isy-6-
dependent regulatory feedback loop. The ASER-specific \textit{gcy} genes are ectopically activated in ASEL upon loss of the ASEL-inducer \textit{lsy-6} and the ASEL-specific \textit{gcy} genes are lost in the ASEL neuron of \textit{lsy-6} null mutant animals. We note that in the case of the ASEL-biased \textit{gcy-14} gene, which is normally strongly expressed in ASEL and weakly expressed in ASER, the conversion of the ASEL to the ASER fate does not entail a loss of \textit{gcy-14} expression, but rather the bilateralization of weak expression (“L=R”; Fig.6B).

How does the \textit{lim-6} LIM homeobox gene contribute to the expression of the newly identified \textit{gcy} genes? Similar to the previously described ASEL-specific \textit{gcy-6} and \textit{gcy-7} genes, we find that the expression of the new ASEL-expressed genes \textit{gcy-14} and \textit{gcy-20} are unaffected by loss of \textit{lim-6} (Fig.6B), demonstrating that ASEL inducers such as \textit{lsy-6} independently regulate the expression of \textit{lim-6} and ASEL-specific \textit{gcy} genes (Fig.6C). Similar to the previously described ASER-specific \textit{gcy-5} and \textit{gcy-22} genes, we find that \textit{lim-6} represses the expression of the ASER-specific \textit{gcy-1}, \textit{gcy-3} and \textit{gcy-4} genes in ASEL (Fig.6A,B). In summary, all known asymmetrically expressed \textit{gcy} genes are controlled by the same categories of gene regulatory factors.

**Functional analysis of \textit{gcy-5}**

Three putative loss of function alleles have been generated by the \textit{C. elegans} knockout consortia in one of the nine asymmetrically expressed \textit{gcy} genes, \textit{gcy-5} (Fig.7A). At least one of them is a putative molecular null allele (see legend to Fig.7A). We analyzed all three mutant \textit{gcy-5} alleles in a chemotaxis assay that measures the functionality of the ASE neurons (BARGMANN and HORVITZ 1991). Our previous work has demonstrated that the ASEL neuron preferentially senses sodium, but not chloride, whereas the ASER neuron senses chloride and weakly contributes to sodium sensation (CHANG \textit{et al.} 2004; PIERCE-SHIMOMURA \textit{et al.} 2001). Neither \textit{gcy-5} mutant allele shows any significant defects in attraction to sodium or chloride (Fig.7B). We conclude that \textit{gcy-5} is not required for generic aspects of ASER development or function.

**Guanylyl cyclase genes in the nematode \textit{Caenorhabditis briggsae}**

The availability of the genome sequence of the nematode \textit{C. briggsae}, which
diverged about 100 million years ago (STEIN et al. 2003), provided us with the opportunity to examine the evolution of the gcy gene family. To identify the complete set of guanylyl cyclase genes in C. briggsae we again employed PSI-BLAST to search the latest release of the complete genome databases of C. briggsae, using a set of known GCY proteins as queries. We identified a total of 33 gcy genes in C. briggsae, one less than in C. elegans (Fig.3A). Compared to C. elegans, C. briggsae contains one additional soluble gcy gene (total = 8) and contains two receptor-type gcy genes less (total = 25). 23 nematode receptor- and soluble gcy genes show clear one-to-one ortholog matches between C. elegans and C. briggsae (Fig.3A). All the other gcy genes show species-specific gene duplication events. For example, the C. elegans gcy-6 gene has duplicated to produce two paralogs, while the directly adjacent C. elegans gcy-1, gcy-2 and gcy-3 genes are paralogs with only a single ortholog in C. briggsae (Fig.3A).

We have mentioned above the existence of a receptor-type gcy protein, GCY-27, that lacks a predicted extracellular domain. C. elegans GCY-27 is closely related to the intracellular domain of the receptor-type ODR-1; both proteins are more closely related to each other than to the two closest C. briggsae homologs (Fig.3A). One of these homologs, CBP15915* (see Material and Methods for refined gene predictions) is clearly orthologous to ODR-1 due to synteny in the chromosomal region; that is, all predicted proteins in the direct neighborhood of ODR-1 are homologous to the neighboring proteins of the putative C. briggsae ODR-1 ortholog. The protein designated CBP11205* (see Material and Methods for refined gene prediction) is the likely ortholog of GCY-27. Both proteins share the unusual feature of not containing an extracellular domain. As in the case of GCY-27, the absence of an extracellular domain in the C. briggsae ortholog is unlikely caused by a failure to predict more exons because neighboring genes are in close proximity. We conclude that the extracellular domain of an ancestral protein was lost before the C. briggsae/C. elegans split, giving rise to GCY-27 and its ortholog CBP11205*.

We sought to investigate the issue of orthology and paralogy in more detail by (a) analyzing on a sequence level the degree of synteny between gcy genes and by (b) analyzing the expression patterns of some orthologous C. briggsae genes. For the chromosomal analysis, we focused on one subgroup of left/right asymmetrically
expressed gcy subfamilies, the “gcy-5 subfamily” (Fig.3A).

We found that the chromosomal arrangement of members of the gcy-5 subfamily differ significantly between C. elegans and C. briggsae. The gcy-19 orthologs are located in distinct environments in C. elegans and C. briggsae (Fig.8A), suggesting a translocation event after the lineage split. The case of the gcy-1, gcy-2, gcy-3, gcy-4 and gcy-5 genes is more complex and provides a fascinating glimpse into evolutionary divergence. While gcy-4 and gcy-5 are direct neighbors in both C. elegans and C. briggsae, they are translocated together with their neighbors to a distinct region in C. briggsae (Fig.8B). Curiously, this distinct region contains the single C. briggsae ortholog of three C. elegans paralogs from the gcy-5 subfamily, gcy-1, gcy-2 and gcy-3 (Fig.8B). Since all paralogous members of the gcy-5 subgroup likely arose by gene duplication, the C. briggsae gene organization is probably more reflective of the ancestral situation than C. elegans. One conceivable scenario is that a common ancestor of C. elegans and C. briggsae contained, like C. briggsae, three adjacent gcy genes (which themselves arose by gene duplication). In the C. elegans lineage, this cluster split up, with gcy-4 and gcy-5 translocating to a distinct chromosomal region and the gcy-1/2/3 ortholog also translocating to a distinct location and then subsequently duplicating to generate gcy-1, gcy-2 and gcy-3 in the C. elegans lineage.

**Evolutionary divergence of left/right asymmetric gcy gene expression patterns**

Using gfp reporter genes generated from genomic C. briggsae DNA, we analyzed the expression of five gcy genes in C. briggsae. We generated gfp fusions to the single C. briggsae ortholog of the C. elegans gcy-1/gcy-2/gcy-3 genes and the C. briggsae orthologs of C. elegans gcy-4, gcy-5, gcy-7 and gcy-19. We observed no significant adult expression for C. briggsae gcy-1/2/3prom::gfp, C. briggsae gcy-5prom::gfp and C. briggsae gcy-7prom::gfp in several transgenic C. briggsae lines which is possibly due to the limited size of the reporter gene constructs (due to the size of the intergenic regions, none of the reporter constructs extended to the previous gene; data not shown). We did, however, observe interesting expression patterns for C. briggsae gcy-4prom::gfp and C. briggsae gcy-19prom::gfp and focused our analysis on these genes.

We find that C. briggsae gcy-4prom::gfp is expressed bilaterally in both ASEL and
ASER (Fig.9A). This is in striking contrast to *C. elegans gcy-4prom::gfp*, whose expression is biased to ASER in *C. elegans* (Fig.5C, Fig.6B). Such a difference could be caused by two different mechanisms. The *gcy-4* loci may contain distinct *cis*-regulatory information in *C. elegans* and *C. briggsae*, or, alternatively, the *cis*-regulatory information may be the same, but different *trans*-acting factors in *C. elegans* and *C. briggsae* interpret this information differentially. To distinguish between these possibilities, we injected the *C. briggsae gcy-4prom::gfp* construct into *C. elegans*. If the *C. elegans* and *C. briggsae gcy-4* reporter constructs contained the same *cis*-regulatory information, *C. briggsae gcy-4prom::gfp* should be expressed in the same pattern in *C. briggsae* and *C. elegans* (bilateral expression in ASEL and ASER). In contrast, if there are differences in *trans*-acting factors, then the *C. briggsae gcy-4prom::gfp* reporter should be expressed in *C. elegans* in a manner similar to that of *C.elegans gcy-4prom::gfp* (biased to ASER). We found that *C. briggsae gcy-4prom::gfp* expression becomes ASER-biased when injected into *C. elegans*, which mimics the expression of *C.elegans gcy-4prom::gfp* (Fig.9A). The difference between *gcy-4* expression in *C. elegans* and *C. briggsae* does therefore not appear to be a difference in their *cis*-regulatory architecture but a reflection of different *trans*-acting factors.

The *gcy-19* locus represents another example of evolutionary divergence of left/right asymmetric gene expression. *C. elegans gcy-19prom::gfp* is strongly expressed in the IL2 sensory neurons and weakly in both ASEL and ASER and several other head sensory neurons in *C. elegans* (Fig.5N). In striking contrast, *C. briggsae gcy-19prom::gfp* shows strong and exclusive expression in the *C. briggsae* ASER neuron, but not any other head neurons (Fig.9B). In contrast to the case of *gcy-4*, however, the difference in expression does not appear to be caused by differences in *trans*-acting factors. When injected into *C. elegans*, the *C. briggsae gcy-19prom::gfp* reporter is still expressed exclusively in ASER (Fig.9B). The *cis*-regulatory architecture of the *C. briggsae gcy-19* locus can therefore be “read out” in the same way by the *trans*-acting factors in both *C.elegans* and *C.briggsae*. Differences in the expression of *C. elegans* and *C. briggsae gcy-19* are therefore more likely caused by a difference in the *cis*-regulatory architecture of these loci. Since *C.elegans* and *C.briggsae gcy-19* reside in non-syntenic chromosomal regions (Fig.8A), it appears that the alterations in chromosomal context of
these two genes not only affected the neighboring genes of the gcy-19 loci but also their cis-regulatory architecture. As a word of caution, we note the intrinsic limitations of reporter gene constructs, which may not harbor the complete set of cis-regulatory elements thereby potentially yielding a misleading impression of gene expression patterns.

**DISCUSSION**

Biochemical properties and functions of receptor-type guanylyl cyclases have been summarized and discussed in several reviews over the past few years (LUCAS et al. 2000; MORTON 2004; WEDEL and GARBERS 2001). Recent findings on the function of soluble gcy gene in oxygen sensation (CHEUNG et al. 2004; CHEUNG et al. 2005; GRAY et al. 2004) have also been reviewed (RANKIN 2005). We focus here on several specific outcomes of our studies.

**The function of nematode receptor-type gcy genes.** *C. elegans* and *C. briggsae* contain an unusual number of receptor-type gcy genes. Insects such as *Drosophila melanogaster* or *Anopheles gambiae* contain six receptor-type guanylyl cyclases (MORTON 2004), mammals contain seven (four orphan and three peptide-binding receptors) (LUCAS et al. 2000; WEDEL and GARBERS 2001), but *C. elegans* contains 27 and *C. briggsae* 25 (this study). The physiological function of insect gcy genes is entirely unknown, though the expression of the only two analyzed receptors in sensory neurons (among other neurons) has been noted (MORTON 2004). Vertebrate gcy genes are expressed in several different tissue types including chemosensory neurons (WEDEL and GARBERS 2001).

We propose that the significant expansion of receptor-type gcy genes in the nematode lineage is a reflection of their employment as chemoreceptors used to sense and navigate through a complex soil habitat. This hypothesis, which was also put forward by Yu et al. (1997), is mainly based on the observation that almost 90% (24/27) of gcy genes are expressed in sensory neurons (lack of sensory neuron expression of three gcy genes may merely be caused by a lack of the complete set of cis-regulatory
element in the reporter genes used). Moreover, 41% (11/27) of gcy genes are expressed in the main gustatory neuron class of *C. elegans*, ASE. This neuron class has previously been shown to be functionally lateralized in that it can sense different chemosensory cues (Pierce-Shimomura *et al.* 2001). Consistent with a role of gcy genes as chemoreceptors we find that 9/11 ASE-expressed gcy genes are expressed in a left/right asymmetric manner, thereby providing molecular correlates to functional lateralization.

Amino acids are among the several classes of chemicals that can be sensed by the ASE neurons (Bargmann and Horvitz 1991). The presence of a domain in the extracellular parts of many GCY proteins that is homologous to bacterial amino acid binding proteins makes receptor-type GCY proteins good candidates to be amino acid receptors. More sensitive assays (Faumont and Lockery 2005; Faumont *et al.* 2005; Miller *et al.* 2005; Wicks *et al.* 2000), than those previously used (Ward 1973) will be required to establish the full spectrum of amino acids and other possible sensory cues that signal through ASE. Such a systematic cataloguing of sensory cues processed by *C. elegans* needs to be followed by a systematic analysis of strains harboring deletions in gcy genes as well as misexpressing gcy genes to establish their roles as amino acid receptors. A role for GCY proteins as salt receptors is also conceivable but highly speculative at present. There is as yet no consensus about the molecular identity of salt receptors in the vertebrate gustatory system. Notably, the crystal structure of the ANP receptor, a mammalian GCY protein that like most *C. elegans* GCY proteins contains an extracellular RFLPB domain, revealed a high affinity chloride binding site (van den Akker *et al.* 2000).

Our attempt to establish a mutant phenotype for a gcy gene, the ASER-expressed gcy-5 gene, has failed so far, but we note that we have only tested one of the two known cues sensed in an ASER-specific manner, namely chloride ions. The failure to detect a mutant phenotype does, however, allow the conclusion that gcy genes such as gcy-5 are unlikely to control a fundamental, non-redundant aspect of the development or overall function of the neuron. It appears more likely that gcy-5 and, by inference, other ASE-expressed gcy genes fulfill a sensory modality specific function in ASE such as being a receptor for a specific class of gustatory cues.
Whereas two thirds of *C. elegans* gcy genes (16/24) are expressed exclusively in sensory neurons, one third are also expressed in non-sensory neurons, including inter/motorneurons and non-neuronal cells, suggesting that GCY proteins also respond to endogenously produced ligands. Since all known ligands for mammalian GCY proteins are peptidergic signaling molecules (Lucas et al. 2000) and since *C. elegans* contains scores of neuropeptide encoding genes (Li et al. 1999), we propose that non-sensory *C. elegans* GCY proteins may be receptors for peptidergic ligands.

The function of receptor-type GCY proteins may not be restricted to a receptor function. The *C. elegans* ODR-1 protein does not require its extracellular domain to fulfill its function in transducing odorsensory signals (L'ETOILE and Bargmann 2000). Moreover, one GCY protein that we describe here, the ODR-1-related GCY-27 protein, entirely lacks an extracellular domain. These proteins may heterodimerize with ligand-binding GCY receptors to constitute a receptor complex and/or they may serve as second-messenger producing signaling proteins that are embedded in signal transduction cascades triggered by other receptor systems.

**Co-expression of gcy genes.** Another notable feature of gcy gene expression patterns is their degree of co-expression. Six gcy genes are co-expressed in the ASI sensory neuron class, five are co-expressed in ASER, four in ASEL, five in the AWC olfactory neuron class, four in the AFD thermosensory neuron class and two gcy genes are each co-expressed in the ASG and PHA phasmid sensory neuron classes. In addition, several gcy genes that are expressed in multiple cell types show similar and non-intuitive combinations of expression patterns. gcy-7 and gcy-20 are co-expressed in ASEL and the excretory cell (two cells of no obvious relation) and daf-11 and odr-1 show a precisely overlapping expression pattern in five amphid sensory neurons (Birnby et al. 2000; L'ETOILE and Bargmann 2000). Co-expression of receptor-type GCY proteins raises at least two different possibilities: (1) The proteins are independent receptors for distinct sensory inputs. While an attractive possibility for the ASE-expressed gcy genes, this is unlikely to be the case for daf-11 and odr-1, which are both independently required for AWC-mediated olfaction to several distinct odorants (Birnby et al. 2000; L'ETOILE and Bargmann 2000). (2) Since GCY proteins dimerize, it is possible that the defined set of GCY proteins of one cell-type can homo- and
heterodimerize to form an even larger repertoire of dimerized receptor complexes. As previously suggested, such heterodimerization may be obligatory in those cases where a given GCY protein lacks residues that are necessary for cyclase activity (MORTON 2004). The cyclase domains of several GCY proteins, including at least GCY-11, DAF-11 and GCY-29, are predicted to be inactive based on the substitution of conserved residues required for activity (Supp.Fig.1), but all three genes are co-expressed with other GCY proteins that are predicted to be active (Table 4).

**Laterality in the nematode nervous system.** With the identification of a total of nine asymmetrically expressed gcy genes, the ASE neurons present so far the most striking example of a lateralized chemosensory neuron. As mentioned above, our previous work has provided a conceptual framework for the functional relevance of laterality in the ASE neurons (PIERCE-SHIMOMURA et al. 2001). We expect that other neurons may similarly employ the principle of lateralizing chemosensory function. However, our expression analysis has not revealed further examples of laterality in gcy gene expression profiles in other bilaterally symmetric amphid neurons, therefore leaving only the AWC odorsensory and ASE gustatory neurons as neuron pairs with lateralized functions (Fig.1B). Considering the potential usefulness of lateralizing chemoreceptor function, it is our expectation that the analysis of expression of the hundreds of chemoreceptors of the 7-transmembrane receptor family (BARGMANN 1998) will reveal more examples of laterality in the nervous system.

Our analysis of the laterality of gcy gene expression in *C. briggsae* revealed several striking examples of evolutionary plasticity in the laterality of the gustatory system of nematodes. The variation that we observe appears to be caused by a variation in both cis-acting elements and trans-acting factors, a conclusion that we can draw from our comparison of *C. briggsae* gcy promoter activity in *C. briggsae* and *C. elegans*. Variations in cis-regulatory control have been recognized to be a major feature of evolutionary processes (CARROLL et al. 2001) but our cross-species analysis also provides strong support for the more conventional view of evolution of trans-acting factors. The clearly distinct nature of cis-regulatory architecture of at least some gcy genes in *C. elegans* and *C. briggsae* is a strong reminder that the use of bioinformatic tools which use phylogenetic conservation to identify cis-regulatory elements in genomic
sequences (e.g. (BIGelow et al. 2004); for review see (BULyK 2003)) may easily lead to false-negative predictions. It is difficult at this point to speculate about differences in the trans-acting factors that control laterality in C. elegans and C. briggsae. We have demonstrated that so far all known left/right asymmetrically expressed terminal differentiation markers in the ASEL/R neurons, including all gcy genes, are under control of a previously described bistable feedback loop that is composed of several transcription factors and miRNAs (JOHNSTON et al. 2005). The activity of transcription factors that are controlled by the output of the loop, for example the lim-6 LIM homeobox gene (Fig.6B), may have functionally diverged in C. briggsae. This can perhaps be best illustrated with the gcy-4 gene, which in C. elegans is repressed by lim-6 in ASEL. In C. briggsae, gcy-4 is not repressed in ASEL, yet the C. briggsae gcy-4 promoter can be repressed in C. elegans. These observations are consistent with the presence of a lim-6 responsive repressor element in both C. elegans and C. briggsae gcy-4 promoters but with an inability of C. briggsae lim-6 to control this element in C. briggsae.

Mutant screening approaches that identify the complete set of trans-acting factors controlling L/R asymmetric gcy gene expression in C. elegans and experimental promoter dissection approaches that identify cis-regulatory elements through which these trans-acting factors act are currently ongoing in our laboratory and are likely to reveal the molecular control of laterality in the gustatory system of C. elegans and its divergence in C. briggsae.
ACKNOWLEDGEMENTS

We thank Qi Chen for expert technical assistance, the *C. elegans* gene knockout consortia at Tokyo Women’s Medical University School of Medicine (led by Shohei Mitani) and at the Oklahoma Medical Research Foundation (led by Bob Barstead) for providing mutant strains, the CGC and members of the worm community for providing strains, Piali Sengupta for help with cell identifications and members of the Hobert lab and Cori Bargmann for comments on the manuscript. This work was supported by the NIH Medical Scientist Training Program (to C.O.O.), NIH R01 NS050266-01 (O.H.) and R01 NS39996-05 (O.H.). B.H. and O.H. are Investigators of the HHMI.
FIGURE LEGENDS

Fig.1: An introduction to C. elegans sensory anatomy.

A: A prominent and well-characterized subset of C. elegans sensory neurons, the amphid sensory neurons, are schematically shown. Like most other neuron classes, amphid sensory neuron classes consist of one pair of two bilaterally symmetric cells (see also panel B). Each of the 12 pairs of amphid sensory neurons extends a dendrite to the tip of the nose and an axon into the nerve ring, a nerve bundle where synaptic connections are made (WHITE et al. 1986). Delineated functions of amphid sensory neurons are indicated (BARGMANN and MORI 1997).

B: Amphid sensory neuron classes consist of two bilaterally symmetric pairs of neurons, two of which, AWCL/R and ASEL/R, are functionally lateralized. While some other amphid sensory neurons appear to contribute to gustation, ASE is the main gustatory neuron class in C. elegans, mediating responses to salts, amino acids and small metabolites (BARGMANN and HORVITZ 1991). ASE not only mediates attractive, but also repulsive responses to specific chemicals (SAMBONGI et al. 1999). The salts sodium, chloride and potassium are sensed in a left/right asymmetric manner, with ASEL sensing sodium, but not chloride and potassium and ASER sensing chloride and potassium (PIERCE-SHIMOMURA et al. 2001). It is not yet known whether other ASE-sensed chemicals may also activate ASEL and ASER differentially. ASE-expressed gcy genes are shown; those newly described in this paper are shaded; those that are asymmetric are colored. The AWCL/R neurons can discriminate benzaldehyde and butanone based on the left/right asymmetric expression of str-2 (colored), which is stochastically expressed in either AWCL or AWCR. Newly identified gcy genes in AWC are shown and shaded.

Fig.2: Domain structure of GCY proteins.

SS = signal sequence, TM = transmembrane domain, RFLPB = “Receptor family ligand binding region” (Pfam PF01094), Protein Kinase-like = ‘Protein kinase domain” (PF00069), Cyclase = “Adenylate and Guanylate cyclase catalytic domain ” (PF00211), HNOB = “heme NO binding domain” (PF07700). See Supp.Fig.1 and Supp.Fig.2 for the
primary sequence alignment of individual domains. We note that several of the receptor-type GCY proteins, such as GCY-11, lack a clear SS at the N-terminus but the presence of a clear TM and/or RFLPB domains make us suspect that the absence is due to an incorrectly predicted N-terminus of the respective genes and we therefore grouped these genes together with other clear-cut SS/TM-containing proteins. GCY-22 has an unusual and phylogenetically conserved insertion between the transmembrane and protein kinase domain (Supp.Fig.2).

**Fig.3: Sequence similarity and chromosomal localization of gcy genes.**

**A:** Phylogenetic tree based on the intracellular domain of the receptor-type guanylyl cyclases and the complete sequences of the soluble guanylyl cyclases. Numbers at the tree nodes are bootstrap values, which indicate the frequency (in %) of occurrence of a given partition in the 1000 replicate trees. *C. elegans* proteins are shaded; *C. briggsae* proteins all carry the prefix “CBP”. A select number of cells that co-express multiple gcy genes are indicated by color-coded shades, as indicated in the figure. See Material and Methods for comments on the gcy gene names. Note that the CBP15915*, CBP11205* and CBP04902* proteins used here differ from those in Wormbase WS149 based on an alternative gene prediction that we performed (see Material and Methods).

**B:** Chromosomal localization of gcy genes. See Table 2 for detailed map position. Chromosome III does not contain any gcy genes.

**Fig.4: Reporter gene constructs.**

Representation of genomic loci are adapted from www.wormbase.org. Reporter gene constructs are indicated by the 5’ upstream region used (red box), usually the intergenic region up to the next gene, and the gfp coding region (green box; not drawn to scale). See Table 1 for primer sequences and a list of transgenic arrays containing the individual reporter gene constructs.

**Fig.5: Expression patterns of gcy reporter gene fusions.**

Transgenic animals expressing gfp reporter gene fusions are shown. Images shown are representative animals from several independent lines (see Table 1 for a list of
transgenic reporter arrays used). Most transgenic animals are scored in the late larval
and adult stage and contain otls151 in the background to facilitate the identification of
the ASE neurons (see Material and Methods); blue circles indicate ASER, red circles
indicate ASEL. The quantification of the L/R asymmetric expression in ASE is shown in
Fig.6.
A: gcy-1prom::gfp. Dorsal view (left panel) and ventral view (right panel) of two different
focal plains of the head region. The inset in the right panel shows the overlap of the gfp
signal with otls133, an AIY-expressed rfp marker (lateral view).
B: gcy-3prom::gfp. Dorsal view of the head region (left panel) of an animal whose amphid
sensory neurons have been filled with Dil. The right panel shows a full length worm with
expression in the PVT interneuron.
C: gcy-4prom::gfp. Dorsal view of the head region.
D: gcy-14prom::gfp. Dorsal view of the head region.
E: gcy-20prom::gfp. Dorsal view of the head region (left panel). The right panel shows a
full length worm with expression in the excretory system. EXG = excretory gland cell;
EXC = excretory canal cell.
F: gcy-7prom::gfp. Lateral view. Expression in ASEL, but not expression in the excretory
canal cell (EXC), had been previously reported (Yu et al. 1997).
G: gcy-2prom::gfp. Lateral view of the head region. A defined subset of the amphid
neurons are filled with Dil to allow for easier assessment of cell position. The inset
shows a dorsal view to illustrate bilateral symmetry of gfp expressing cells.
H: gcy-11prom::gfp. Lateral view of the head region. The strong neuronal expression (N)
in the pharynx is likely due to the injection marker, but pharyngeal muscle expression is
due to the reporter gene.
I: gcy-13prom::gfp. Dorsal view of the head region. The inset shows a lateral view of a
Dil-filled animal.
J: gcy-15prom::gfp. Lateral view of the head region. A defined subset of the amphid
neurons are filled with Dil to allow for easier assessment of cell position. The inset
shows a ventral view to illustrate bilateral symmetry.
K: gcy-17prom::gfp. Dorsal view of the tail region. The inset shows a Dil-filled animal.
L: gcy-18prom::gfp. Lateral view of the head region. The inset shows a lateral view of gfp
expression in relation to the rfp expression from *otls133*, an AIY-specific cell marker.

**M:** *gcy-21prom::gfp*. Lateral view of the head region of a Dil-filled animal. The inset shows a ventral view to illustrate bilateral symmetry.

**N:** *gcy-19prom::gfp*. Lateral view (left panel) and dorsal view (right panel) of the head region. The inset in left panel shows the overlap of the *gfp* signal with Dil-filled IL2 neurons.

**O:** *gcy-25prom::gfp*. Left panel shows an oblique view of the head region, the right panel shows a lateral view of the tail region.

**P:** *gcy-23prom::gfp*. Ventral view of the head region.

**Q:** *gcy-27prom::gfp*. Dorsal view of the head region. The inset shows a lateral view of a Dil-filled animal.

**R:** *gcy-26prom::gfp*. VNC = ventral nerve cord; HG = head ganglia; N = non-neuronal cells. The animal is mosaic and does not show muscle expression.

**S:** *gcy-29prom::gfp*. Ventral view of the head region. The reporter is also expressed in AWCL/R and AVKL/R (not shown in this animal).

**Fig.6: Regulation of the expression of asymmetric gcy genes.**

**A, B:** Quantification of the asymmetry of ASE-expressed *gcy* genes in wild-type and mutant backgrounds. In panel A, the quantification and a representative example of *gcy-1prom::gfp* reporter gene-carrying animals are shown, in panel B only the quantification of the results obtained with the other asymmetrically expressed and previously uncharacterized *gcy* reporter gene constructs are shown. Black, white and grey circles indicate relative expression levels of *gfp* in the ASEL and ASER neurons. Data is shown for one representative array each (*otEx2419 for gcy-1prom::gfp, otEx2423 for gcy-3prom::gfp, otEx2409 for gcy-4prom::gfp, otEx2322 for gcy-14prom::gfp, otEx2327 for gcy-20prom::gfp*). These arrays were each crossed into the indicated mutant backgrounds. Comparable results were obtained with several independent arrays (not shown). Like in other figures, color coding is red to indicate ASEL expression and blue for ASER expression.

**C:** Summary of the gene regulatory interactions. Grey shading indicates genes identified in this paper. For more details on the ASEL and ASER inducers, see
Fig.7: Functional analysis of gcy-5.
A: Mutant alleles of gcy-5. Color coding for the domains encoded by the individual exons is shown in Fig.2. Reading frames are indicated to illustrate the effects of the respective deletion alleles. tm897 contains a 691 bp deletion from position 462 - 1045 of the coding sequence and replacement with 5’-GGGGTAGAAGAGGC. Within the genomic locus, the deletion starts in exon 4 and ends in exon 7. With the deletion and insertion, a frame-shift is created leading to an early stop codon. This allele is therefore a putative null allele. The effect of the other alleles is more difficult to predict since the respective deletions start in exons and end in introns. If one assumes splicing around the half-deleted exons, then the two ok alleles produce large, but in-frame deletions.
B: Chemotaxis to soluble ions of wild-type and gcy-5 mutant animals. NaCl measures the functionality of both ASEL and ASER. NH₄Cl mainly measures ASER function since NH₄⁺ is sensed in a ASE-independent manner (NH₄⁺ sensation is unaffected in che-1 mutants in which ASEL/R fail to develop (CHANG et al. 2004)). All strains were grown and assayed at room temperature (21-23°C). Population chemotaxis assays were performed in a radial gradient of the indicated salt (see Material and Methods). Each experiment was done with at least two plates in parallel and for each assay plate at least 20 worms reached either attractant or negative control spot. 3-9 independent experiments were done for each condition. Error bars indicate standard error of the mean. For statistical analysis, an one-way ANOVA was performed for each attractant, with Dunnet’s post-test comparing all three alleles to wild-type control data. None of the means were significantly different.

Fig.8: Analysis of synteny of gcy genes in C. elegans and C. briggsae.
Gene predictions were taken from WormBase release WS149 (www.wormbase.org). The red line indicates the genomic regions included in gfp reporter gene constructs. The C. elegans reporter constructs are also shown in Fig.4. The size of the C. briggsae gcy-19 construct is 2 kb, the size of the C. briggsae gcy-4 construct is 433 bp (up to preceding gene).
Fig.9: Phylogenetic conservation of gcy gene expression profiles.
C. briggsae reporter gene constructs are shown in Fig.8, C. elegans reporter gene constructs are shown in Fig.4. Species names in pictures indicate into which species the respective reporter gene was injected.

A: gcy-4 expression. C. briggsae gcy-4prom::gfp is expressed in both ASEL and ASER in C. briggsae (3 lines; data for one representative array, otEx2508, is shown), but is expressed predominantly in ASER in C. elegans (3 lines; n>40 each; data for one representative array, otEx2510, is shown). Expression of C. elegans gcy-4prom::gfp (construct in Fig.4) in C. elegans is also biased to ASER (Fig.5C, Fig.6B).

B: gcy-19 expression. C. briggsae gcy-19prom::gfp is expressed exclusively in ASER in C. briggsae (3 lines; data for one representative array, otEx2139, is shown) and in C. elegans (3 lines; data for one representative array, otEx2141, is shown).

Supp.Fig.1: Alignment of the guanylyl cyclase domain of C. elegans GCY proteins. An asparagine and arginine residue essential for catalytic activity (MORTON 2004; YAN et al. 1997) are boxed. Note that one or both of these two residues is mutated in several of the receptor and soluble-type kinases suggesting that they may be catalytically inactive.

Supp.Fig.2: Alignment of the protein kinase-like domain of C. elegans GCY proteins. A set of canonical protein kinase domains is included in this alignment of the kinase domains of the GCY proteins. Note the labeled HRD motif (boxed in figure); the D is conserved in all active protein kinases and is required for catalytic activity (TAYLOR et al. 1992). It is not conserved in GCY proteins, therefore making it unlikely that these protein kinases have catalytic activity.
LITERATURE CITED


<table>
<thead>
<tr>
<th>Promoter</th>
<th>PCR fusion primers</th>
<th>Resulting transgenic arrays</th>
</tr>
</thead>
<tbody>
<tr>
<td>gcy-1</td>
<td>gaagctggagtcaggtgt</td>
<td>otEx2419, otEx2420, otEx2421</td>
</tr>
<tr>
<td>gcy-2</td>
<td>ccattacgagagatcag</td>
<td>otEx2312, otEx2313, otEx2314</td>
</tr>
<tr>
<td>gcy-3</td>
<td>gagctctgagagatcag</td>
<td>otEx2422, otEx2423</td>
</tr>
<tr>
<td>gcy-4</td>
<td>ggaatctcagagatcag</td>
<td>otEx2407, otEx2408, otEx2409</td>
</tr>
<tr>
<td>gcy-9</td>
<td>gatatactcagagatcag</td>
<td>otEx2307, otEx2308, otEx2541, otEx2542</td>
</tr>
<tr>
<td>gcy-11</td>
<td>ctgattgagagatcag</td>
<td>otEx2304, otEx2305, otEx2306</td>
</tr>
<tr>
<td>gcy-13</td>
<td>gccattcagagagatcag</td>
<td>otEx2410, otEx2411, otEx2412</td>
</tr>
<tr>
<td>gcy-14</td>
<td>gtcagattcagagagatcag</td>
<td>otEx2322, otEx2323, otEx2324</td>
</tr>
<tr>
<td>gcy-15</td>
<td>gttatcctacgagagatcag</td>
<td>otEx2497</td>
</tr>
<tr>
<td>gcy-17</td>
<td>ggtgagagagatcag</td>
<td>otEx2413, otEx2414, otEx2415</td>
</tr>
<tr>
<td>gcy-18</td>
<td>ggtgagagagatcag</td>
<td>otEx2498</td>
</tr>
<tr>
<td>gcy-19</td>
<td>ggtgagagagatcag</td>
<td>otEx2309, otEx2310, otEx2311, otEx2535, otEx2536, otEx2537, otEx2538</td>
</tr>
<tr>
<td>gcy-20</td>
<td>ggtgagagagatcag</td>
<td>otEx2325, otEx2326, otEx2327</td>
</tr>
<tr>
<td>gcy-21</td>
<td>ggtgagagagatcag</td>
<td>otEx2416, otEx2417, otEx2418</td>
</tr>
<tr>
<td>gcy-23</td>
<td>ggtgagagagatcag</td>
<td>otEx2499, otEx2500, otEx2501</td>
</tr>
<tr>
<td>gcy-25</td>
<td>ggtgagagagatcag</td>
<td>otEx2424, otEx2425, otEx2426</td>
</tr>
<tr>
<td>gcy-27</td>
<td>ggtgagagagatcag</td>
<td>otEx2502, otEx2503, otEx2504, otEx2505, otEx2506, otEx2507, otEx2508</td>
</tr>
<tr>
<td>gcy-28</td>
<td>ggtgagagagatcag</td>
<td>otEx2490, otEx2491, otEx2492</td>
</tr>
</tbody>
</table>

**Subcloning** 5

<table>
<thead>
<tr>
<th>5' primer</th>
<th>3' primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>C.b.gcy-4</td>
<td>ttaagattAAAACGTATCCCTTTCTCCTT</td>
</tr>
<tr>
<td>C.b.gcy-19</td>
<td>ttaagattCAGCGGTTGAAAAAATAATTCG</td>
</tr>
</tbody>
</table>

1 Primer nomenclature according to Hobert, 2002. All sequences go 5' to 3'. 2 If not indicated otherwise, *C.elegans* animals containing the *otIs151* transgene were injected, using *unc-122::gfp* as injection marker. 3 Injection was done into *otIs133* animals, using *rol-6(d)* as injection marker. 4 Injection was done into N2 animals, using *rol-6(d)* as injection marker. 5 Primers used for subcloning contained restriction sites at their end and were subcloned into pPD95.75. 6 Injected into *C.briggsae*, using *rol-6(d)* as injection marker.
### Table 2: Receptor guanylyl cyclases in *C.elegans*

<table>
<thead>
<tr>
<th>Gene</th>
<th>Other names 1</th>
<th>Map position 2</th>
<th>Expression pattern in adult animals</th>
<th>Previous reports 3-7</th>
<th>This report</th>
</tr>
</thead>
<tbody>
<tr>
<td>daf-11</td>
<td>B0240.3</td>
<td>V +3.27</td>
<td>ASIL/R, ASJL/R, ASKL/R, AWBL/R, AWCL/R 3</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>gcy-1</td>
<td>AH6.1</td>
<td>II +1.49</td>
<td>no signal 4</td>
<td>ASER, ASIL/R, PVT, URXL/R, AILY/R, intestine</td>
<td></td>
</tr>
<tr>
<td>gcy-2</td>
<td>R134.2</td>
<td>II +1.48</td>
<td>-</td>
<td>AWAL/R, ASIL/R, RIAL/R, PVT,</td>
<td></td>
</tr>
<tr>
<td>gcy-3</td>
<td>R134.1</td>
<td>II +1.48</td>
<td>-</td>
<td>ASER, ASIL/R, PVT</td>
<td></td>
</tr>
<tr>
<td>gcy-4</td>
<td>ZK970.5</td>
<td>II +2.31</td>
<td>no signal 4</td>
<td>ASER-biased 8</td>
<td></td>
</tr>
<tr>
<td>gcy-5</td>
<td>ZK970.6</td>
<td>II +2.32</td>
<td>ASER 4</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>gcy-6</td>
<td>B0024.6</td>
<td>V +2.46</td>
<td>ASEL 4,7</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>gcy-7</td>
<td>F52E1.4</td>
<td>V +1.37</td>
<td>ASEL 4,7</td>
<td>also expressed in excretory canal cell (only in adults)</td>
<td></td>
</tr>
<tr>
<td>gcy-8</td>
<td>C49H3.1</td>
<td>IV +3.50</td>
<td>AFDL/R 4</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>gcy-9</td>
<td>ZK455.2</td>
<td>X +3.0</td>
<td>no signal 4</td>
<td>weak, occasional and variable expression in non-neuronal tissues</td>
<td></td>
</tr>
<tr>
<td>gcy-11</td>
<td>C30G4.3</td>
<td>X +24.07</td>
<td>-</td>
<td>pharyngeal muscle</td>
<td></td>
</tr>
<tr>
<td>gcy-12</td>
<td>F08B1.2</td>
<td>II –1.80</td>
<td>PHAL/R 4</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>gcy-13</td>
<td>F23H12.6</td>
<td>V +4.02</td>
<td>no signal 4</td>
<td>RIML/R</td>
<td></td>
</tr>
<tr>
<td>gcy-14</td>
<td>ZC412.2</td>
<td>V +6.93</td>
<td>no signal 4</td>
<td>ASEL-biased 8, AWCL/R (faint), PVT</td>
<td></td>
</tr>
<tr>
<td>gcy-15</td>
<td>ZC239.7</td>
<td>II –8.08</td>
<td>-</td>
<td>ASGL/R (faint)</td>
<td></td>
</tr>
<tr>
<td>gcy-17</td>
<td>W03F11.2, gcy-24</td>
<td>V –10.32</td>
<td>-</td>
<td>PHAL/R</td>
<td></td>
</tr>
<tr>
<td>gcy-18</td>
<td>ZK896.8, gcy-26</td>
<td>IV +6.48</td>
<td>-</td>
<td>AFDL/R, AIML/R</td>
<td></td>
</tr>
<tr>
<td>gcy-19</td>
<td>C17F4.6</td>
<td>II –7.80</td>
<td>-</td>
<td>IL2 (strong), ASEL/R (faint) 9, additional faint sensory neurons (3 pairs)</td>
<td></td>
</tr>
<tr>
<td>gcy-20</td>
<td>F21H7.9, gcy-16</td>
<td>V +9.87</td>
<td>-</td>
<td>ASEL, AWCL/R (faint) excretory gland and canal cells</td>
<td></td>
</tr>
<tr>
<td>gcy-21</td>
<td>F22E5.3</td>
<td>II –12.28</td>
<td>-</td>
<td>ASGL/R, ADLL/R (faint)</td>
<td></td>
</tr>
<tr>
<td>gcy-22</td>
<td>T03D8.5</td>
<td>V +25.25</td>
<td>ASER 6</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>gcy-23</td>
<td>T26C12.4</td>
<td>IV –5.01</td>
<td>-</td>
<td>AFDL/R</td>
<td></td>
</tr>
<tr>
<td>gcy-25</td>
<td>Y105C5B.2</td>
<td>IV +14.17</td>
<td>-</td>
<td>AQR, PQR, URXL/R</td>
<td></td>
</tr>
<tr>
<td>gcy-28</td>
<td>T01A4.1</td>
<td>I –1.17</td>
<td>-</td>
<td>Many head neurons, ventral cord and tail neurons, body wall muscle, hypodermis, somatic gonad, intestine 10</td>
<td></td>
</tr>
<tr>
<td>gcy-29</td>
<td>C04H5.3</td>
<td>II +23.04</td>
<td>-</td>
<td>ASELR, AWCL/R, AVKL/R, AFDL/R, few variable other neurons (weak)</td>
<td></td>
</tr>
<tr>
<td>odr-1</td>
<td>R01E6.1, gcy-10</td>
<td>X +12.7</td>
<td>ASIL/R, ASJL/R, ASKL/R, AWBL/R, AWCL/R 4,5</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

This Table summarizes previously characterized and newly determined receptor-gcy expression patterns. “-” indicates that expression was not analyzed.

1 See Material and Methods for comments on gene nomenclature. 2 From www.wormbase.org. 3 Birnby et al., 2000. 4 Yu et al., 1997. We suppose that we detected clear expression in several cases where no signal was observed by Yu et al. since we (a) used a gfp variant that is much brighter than the old gfp version used by Yu et al. and (b) since our reporter constructs may encompass more cis-regulatory information that those of Yu et al. 5 L’Etoile and Bargmann, 2000. 6 Johnston et al., 2005. Additional weak expression that fades in adults is observed in two additional, unidentified head neurons.
As described in Johnston et al., 2005, gcy-6 and gcy-7 are embryonically expressed in both ASEL and ASER and only become restricted to ASEL postembryonically. A similar scenario may apply for other ASEL-expressed gcy genes, but has not been explicitly examined. "ASER-biased" incorporates two categories of expression patterns in a given transgenic line: expression only in ASER in some animals and stronger expression in ASER than in ASEL in other animals. The opposite holds for "ASEL-biased" expression. Such weak and occasional expression in the other cell could be caused by array-overexpression artifacts and we therefore do not wish to emphasize that biased expression is fundamentally different from exclusive expression. Expression in ASEL/R is very dim and not completely penetrant and potential biases to ASEL or ASER are therefore difficult to determine. Expression in all tissue types is often mosaic. Expression in the ASE and AWC neuron classes could not be observed, but as with any other reporter construct described in this Table it is possible that additional regulatory elements not contained within the respective constructs may yield expression in these neurons.
Table 3: Summary of cell-type specificity of gcy reporter gene expression.

<table>
<thead>
<tr>
<th>gcy gene expression pattern</th>
<th>#</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-neuronal only</td>
<td>2/27</td>
</tr>
<tr>
<td>Nervous system</td>
<td>25/27</td>
</tr>
<tr>
<td>Nervous system only</td>
<td>21/27</td>
</tr>
<tr>
<td>Sensory neurons + other neurons</td>
<td>25/27</td>
</tr>
<tr>
<td>Sensory neurons only</td>
<td>15/27</td>
</tr>
<tr>
<td>Single neuron class-specific</td>
<td>9/27</td>
</tr>
</tbody>
</table>

Table 4: Summary of co-expressed gcy reporter genes.

<table>
<thead>
<tr>
<th>Sensory neuron class</th>
<th>Co-expressed gcy genes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>#</td>
</tr>
<tr>
<td>ASE</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>AQR, PQR, URX</td>
<td>7</td>
</tr>
<tr>
<td>ASI</td>
<td>6</td>
</tr>
<tr>
<td>AWC</td>
<td>5</td>
</tr>
<tr>
<td>AFD</td>
<td>4</td>
</tr>
<tr>
<td>ASG</td>
<td>2</td>
</tr>
<tr>
<td>PHA</td>
<td>2</td>
</tr>
</tbody>
</table>
Figure 1

AWA, AWB, AWC: odorsensory
AFD: thermosensory
ASE, ASG, ADF, ASI, ASJ, ADL, ASK: chemosensory
ASH: polymodal

ASEL/R
Salts
Amino acids
Small metabolites

AWCL/R
Odorants

ASEL
str-2
gcy-6
gcy-7
gcy-14
gcy-20
gcy-29

AWL
odr-1
daf-11
gcy-14
gcy-20
gcy-29

AWCL
odr-1
daf-11
gcy-14
gcy-20
gcy-29

Definition:

ASEL/R
ASL/R

AWCL
ASR/R

ASL
Odorants

gcy-19
gcy-29
Receptor-type GCYs

Soluble GCYs

Figure 2
Figure 3
Figure 4
Figure 5
Figure 6

A

**gcy-1**

- Wild type
- **lim-6(nr2073)**
- **lsy-6(ot71)**

**gcy-3**

- Wild type
- **lim-6(nr2073)**
- **lsy-6(ot71)**

**gcy-4**

- Wild type
- **lim-6(nr2073)**
- **lsy-6(ot71)**

B

**gcy-14**

- Wild type
- **lim-6(nr2073)**
- **lsy-6(ot71)**

**gcy-20**

- Wild type
- **lim-6(nr2073)**
- **lsy-6(ot71)**

C

ASEL-inducing genes

- (die-1, lsy-6)

ASER-inducing genes

- (mir-273, cog-1)

red - ASEL
blue - ASER

**lim-6**

- **gcy-1**
- **gcy-3**
- **gcy-4**
- **gcy-6**
- **gcy-7**
- **gcy-14**
- **gcy-20**

**gcy-22**
Figure 7

A

B

Figure 7
Orthologs at ~2220K

CBG17049 ≈ CBP04086
Orthologue on contig cb25.fpc4248

C.elegans
C.briggsae

LG II

CB17049=CBP04086

Ortholog as shown
Orthologous gcy genes as shown
Genomic region included in gfp reporter construct

Figure 8
**A**

*C. briggsae gcy-4 (CBP13906)*

![Image of C. briggsae and C. elegans.](image)

- **C. briggsae**
  - Left-biased: 61
  - Equal: 0
  - Right-biased: 4

- **C. elegans**
  - Left-biased: 50
  - Equal: 10
  - Right-biased: 4

**B**

*C. briggsae gcy-19 (CBP04086)*

![Image of C. briggsae and C. elegans.](image)

- **C. briggsae**
  - Left-biased: 22
  - Equal: 2
  - Right-biased: 8

- **C. elegans**
  - Left-biased: 32
  - Equal: 12
  - Right-biased: 6

**Figure 9**
Supp. Fig. 2

<table>
<thead>
<tr>
<th>Protein</th>
<th>Start</th>
<th>End</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDC5_YEAST/82-337</td>
<td>256</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARBK1_BOVIN/191-453</td>
<td>263</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PKD1_DICDI/36-291</td>
<td>256</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCY-11</td>
<td>345</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCY-21</td>
<td>334</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCY-15</td>
<td>334</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBP00662</td>
<td>302</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBP05559</td>
<td>279</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCY-22</td>
<td>266</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBP01427</td>
<td>266</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBP08562</td>
<td>237</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCY-6</td>
<td>278</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCY-7</td>
<td>273</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCY-17/24</td>
<td>371</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBP14936</td>
<td>268</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBP15625</td>
<td>268</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCY-14</td>
<td>268</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCY-4</td>
<td>270</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCY-3</td>
<td>270</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCY-2</td>
<td>243</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCY-1</td>
<td>270</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCY-5</td>
<td>269</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCY-19</td>
<td>269</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBP04086</td>
<td>269</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCY-9</td>
<td>302</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBP04902*</td>
<td>295</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCY-23</td>
<td>297</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCY-18/26</td>
<td>297</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBP13439</td>
<td>334</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBP00094</td>
<td>292</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCY-8</td>
<td>325</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBP01237</td>
<td>81</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCY-4</td>
<td>85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCY-3</td>
<td>85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCY-2</td>
<td>85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCY-1</td>
<td>82</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBP13905</td>
<td>84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBP04086</td>
<td>81</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCY-9</td>
<td>84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCY-29</td>
<td>83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBP20608</td>
<td>84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBP11205*</td>
<td>84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCY-27</td>
<td>84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCY-10</td>
<td>85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBP00094</td>
<td>82</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCY-8</td>
<td>110</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TTK_HUMAN/509-775 | 94 | |
CDC5_YEAST/82-337 | 94 | |
CHK1_SCHPO/10-272 | 99 | |
ARBK1_BOVIN/191-453 | 98 | |
KGP1_DROME/457-717 | 97 | |
KIN82_YEAST/324-602 | 97 | |
GCY-11 | 128 | |
CBP18607 | 128 | |
CBP04548 | 138 | |
GCY-28 | 244 | |
CBP03512 | 230 | |
GCY-12 | 93 | |
CBP00662 | 94 | |
GCY-7 | 86 | |
CBP02750 | 86 | |
GCY-17/24 | 185 | |
CBP14936 | 81 | |
CBP15625 | 81 | |
CBP01237 | 81 | |
GCY-4 | 85 | |
GCY-1 | 82 | |
CBP13905 | 84 | |
CBP04086 | 81 | |
GCY-9 | 84 | |
GCY-29 | 83 | |
CBP20608 | 84 | |
CBP11205* | 84 | |
GCY-27 | 84 | |
GCY-10 | 85 | |
CBP00094 | 82 | |