Experimental and Neuroinformatic Definition of Neural Circuits in *Caenorhabditis elegans*

Thesis by

Sharan Prakash

In Partial Fulfillment of the Requirements for

the degree of

Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2024

(Defended October 2nd, 2023)

Sharan Prakash
ORCID: 0000-0003-1777-0987

ACKNOWLEDGEMENTS

I would like to acknowledge the following people, whose support and friendship I've relied on and enjoyed over the course of my PhD. First, my advisor Paul Sternberg, for his enthusiasm, trust, and resourcefulness through some very challenging times. My thesis work wouldn't have been possible without the breadth of his scientific expertise, his understanding of my strengths and inclinations as a scientist, and his great patience for everything else. Second, my collaborators Kimberly van Auken and David Hill, for their generosity in training and working with me on the first project in this thesis. Their expertise is the cornerstone of the project, and allowed me to work in an area of biology that was completely new to me. Thirdly, my committee members Betty Hong, Ellen Rothenberg, and Bruce Hay, for their patience, encouragement, and advice throughout my time here. I also want to thank Marianne Bronner, Elizabeth Hong and Justin Bois for their mentorship in my early years at Caltech. I would like to thank my scientific mentors back in Australia who trained and inspired me as an undergraduate, and gave me the opportunity to be here, including Coral Warr, John Bowman, Travis Johnson, and Martin Burd. Fourth, our collaborators at Northeastern University, Vivek Venkatachalam and Maedeh Seyedolmohasedin, whose beautiful experimental work is featured in this thesis. Fifth, the members of the Sternberg lab, for their collaborative spirit, scientific inspiration, and thoughtful feedback throughout my PhD. Special mentions go to lab technicians Shahla Gharib and Stephanie Nava for their help with strain construction while I recovered from injury, to Mark Zhang for his assistance with the pheromone extraction protocol, and to Raymond Lee and Nick Markarian for helpful discussions and advice on GO-CAM. To the friends I've made here, thank you for your company, support, and generosity, and for the fond memories I'll take from this time. There are too many to name, so omissions are inevitable, but in no particular order they include Angel Merchan, Carmen Alonso, Lucy Chong, Phillip Walk, Marica Raguso, Markus Dick, Connor McMahan, Elijah Kane, Sarah Habib, Jagannadh Kumar, Chirag Mital, Zoya Valari, Lavanya Nott, Taylor Stevens, Tom O'Connell, Rory Williams, Zoila Jurado, Tom Naragon, Aditi Narayanan, Lev

Tsypin, Scott Saunders, Alicia Rogers, Lucas Meirelles, Porfirio Quintero, Sarah Cohen, Marcus Lee, Hyeongchan Jo – this list could go on! Thank you to my friends back home in Melbourne, including Rhys Murrian, Leroy D'Souza, Gilbert Oppy, Jarrod Benjamin, Liam Briginshaw, Penny Galbraith and many others, who have kept in touch over the years, supported my far away adventures, and made time and effort for me during my visits back. Finally, my family in Australia, India, and the US. This includes my parents in Australia, Ravi and Radhika, and my brother Varun. They endured distance from me over many years, including the pandemic, and gave me their unwavering support through the long emotional journey of graduate school. Special thanks are due to all my cousins, particularly Vikram Chandra, who has been part of my scientific journey for many years now. Thank you to my extended family, including Satish and Suma, who gifted me *The Selfish Gene* when I was a teenager, to Shashi and Kamala for their support during my leave of absence, and to Sunder and Anu, who spared no effort to give me a home away from home – I'm very lucky to have you all.

ABSTRACT

The free-living nematode *Caenorhabditis elegans* is an established model organism for research in molecular genetics, cell and developmental biology, evolution, and neuroscience. This thesis describes two research projects in *C. elegans* neuroscience. The first project concerns the challenge of synthesizing the accumulating neurobiological literature in *C. elegans*. I describe how an established framework for semantic modelling of cellular pathways can be adapted for semantic modelling of neural circuits and functional annotation of the nervous system, and its potential applications for systems neuroscience. A second portion describes a series of experiments investigating a decision-making process in *C. elegans* larval development that is under neuronal control. *C. elegans* larvae have the ability to decide among alternative developmental trajectories based on environmental conditions, that are detected via its nervous system. In this work, I describe the contribution of several neurons to this decision-making process, and our discoveries about the response properties of two neurons to ethologically relevant chemical stimuli.

PUBLISHED CONTENT AND CONTRIBUTIONS

This thesis includes two projects, both of which involved collaboration. The first project is summarized in a paper entitled 'Semantic Representation of Neural Circuit Knowledge' and has been published in the journal *Brain Informatics* (doi.org/10.1186/s40708-023-00208-5), authored by Sharan Prakash, Kimberly Van Auken, David P. Hill and Paul W. Sternberg. P.W.S & S.J.P. conceived and designed the study, S.J.P, K.V.A & D.P.H. conducted the study. S.J.P, K.V.A, D.P.H & P.W.S wrote and revised the manuscript. The contents of Chapter 2 largely overlap with this publication. The second project involved a collaboration with Vivek Venkatachalam and Maedeh Seyedolmohasedin at Northeastern University, and Mark G. Zhang at Caltech, to perform microfluidics experiments. S.J.P. conceived the experiments, M.S. performed the experiments in the lab of V.V., and M.G.Z provided a reagent. All other experimental work was performed by S.J.P.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

All page numbers below refer to a separate accompanying document containing all illustrations (including Supplementary Figures), available via Caltech DATA at doi:10.22002/vdn6w-65c35.

LIST OF TABLES

The following page numbers refer to this document. In addition to the following tables found in this document, Supplementary Table 1 and Supplementary Table 2 are available as separate accompanying files through Caltech DATA via doi:10.22002/vdn6w-65c35.

Chapter 1

INTRODUCTION

C. elegans as a Model System for Molecular and Systems Neuroscience

The establishment of *Caenorhabditis elegans* as a model organism was pioneered by Sydney Brenner & colleagues in the early 1960's. The short life cycle (3-5 days), bacterial diet, capacity for hermaphrodite self-fertilization, and resistance to freezing and desiccation make it an ideal organism for laboratory culture and forward genetic screens (Brenner 1974), while its optical transparency and size (1mm as adults) are ideal for observation under light and electron microscopes. The pattern of cell divisions leading to adult development has been completely described, and found to be essentially invariant among individuals, leading to an adult hermaphrodite with 959 cells, including 302 neurons (Sulston and Horvitz 1977). With this foundation, genetic analysis revealed important mechanisms of multicellular development, including lineage specification and programmed cell death (Sternberg and Horvitz 1986; Ellis and Horvitz 1986), and the description of the first non-coding RNA (Lee et al. 1993). *C. elegans* was the first animal to have the wiring diagram of its entire nervous system, or 'connectome' described (White et al. 1986). This allowed the functions of individual neurons to be investigated using laser ablation (Chalfie et al. 1985; Bargmann and Horvitz 1991; Bargmann and Avery 1995; Liu and Sternberg 1995) or via mutant strains deficient in the development of specific neurons (Trent et al. 1983). The availability of high resolution anatomical information is a major strength of the *C. elegans* model. The *C. elegans* genome was also the first animal genome to be completely sequenced (The *C. elegans* Sequencing Consortium* 1998), allowing the use of targeted reverse genetics methods such as RNAi (Fire et al. 1998). Gain-of-function and gene expression studies are possible through genetic transformation by gonadal micro-injection (Fire 1986).

More recent technology for genetic perturbation include precise genome editing via CRISPR (Friedland et al. 2013; Dickinson and Goldstein 2016; Wang et al. 2018) and a bipartite expression system, cGAL, which

allows cell-specific expression of a wide range of genetic tools for cellular manipulation (Wang et al. 2017). Stable transgenic lines for many combinations of cell and construct can be generated without the need for repeated transgenic strain construction. This includes constructs for activating or inhibiting individual neurons (Husson et al. 2013; Pokala et al. 2014). The small size and simple culture conditions of *C. elegans* make high-throughput assays of behavior relatively simple to measure (Swierczek et al. 2011). These genetic tools make it possible to perturb neurons for which specific inactivating mutations are unavailable. Genetically-encoded fluorescent sensors of calcium influx allow researchers to image the activity of individual neurons during exposure to stimulants while animals are constrained in microfluidic devices (Chung et al. 2013), making it possible to observe their dynamic responses to ethologically relevant stimuli. Using these tools, researchers are obtaining insights into the molecular and neural circuit mechanisms of behaviors such as egg-laying (Collins et al. 2016), mating (Susoy et al. 2021), and decision-making during chemotaxis (Chalasani et al. 2007; Faumont et al. 2012). More recently, these measurements have been performed in freely moving animals (Venkatachalam et al. 2016), and at the level of the entire brain during a complex behavioral sequence (Susoy et al. 2021). These measurements are mapped onto anatomy with the aid of NeuroPAL, a genetically modified strain carrying 41 reporters, each of which expresses a subset of four different fluorophores in a different, overlapping anatomical subset (Yemini et al. 2019). In combination with optogenetic perturbation, brain-wide imaging is allowing functional connectivity maps of the brain to be generated (Randi and Leifer 2020). These technologies have placed *C. elegans* at the forefront of model organism neuroscience research.

The *C. elegans* Nervous System

The nervous system of the adult *C. elegans* hermaphrodite was first described by White et al. in 1986, and reconstructed using hand drawings. More recently, researchers have used software to generate reconstructions of both sexes (Cook et al. 2019). The hermaphrodite connectome features 302 neurons, divisible by position and morphology into 118 discrete classes. Physical connectivity between neurons is mediated by approximately 5000 chemical synapses and 600 gap junctions (White et al. 1986; Cook et al. 2019). Neuronal

cell bodies are concentrated in either the head or the tail; sensory neurons (i.e., those with nerve endings exposed to the environment) are present at each end. In the head, 14 classes of sensory neurons are organized within organs known as sensilla (White et al. 1986). Two of these, known as amphids, are bilaterally symmetrical and house several ciliated sensory neurons, including ASI, ASJ, ASK, ADL, ASH, ASG and ADF. Each of these neurons may express several genes encoding chemoreceptors that bind chemicals in the environment; for instance, ASK expresses at least nine different chemoreceptor genes (Bargmann 2006). The genome is predicted to encode over 1000 G-protein-coupled receptors (GPCRs) (Bargmann 1998). A recent study examined the expression pattern of 375 reporters for putative chemosensory GPCRs, and found 84% of them to be expressed in chemosensory neurons (Vidal et al. 2018). These neurons and their associated gene families are critical for behaviors that depend on chemical cues, such as navigation towards food and mates (Zhang et al. 2014), and avoidance of noxious stimuli (Shao et al. 2019). Downstream of the sensory layer are interneurons, whose function is to integrate inputs from multiple sensory neurons, and select among appropriate motor programs (Ghosh et al. 2017).

The nervous system as a whole is highly interconnected. It is possible to draw a connection between almost any pair of neurons within three synapses (Bargmann 2012), and some (termed 'rich club' neurons) form highly interconnected hubs (Towlson et al. 2013). The relationship between connectome structure and function is an active area of *C. elegans* research (Yan et al. 2017; Towlson et al. 2018). Connections between neurons take multiple forms. Chemical synapses are traversed by small molecule neurotransmitters via vesicle exocytosis, typically binding to ligand-gated ion channels on post-synaptic neurons (Richmond 2006). Gap junctions, composed of complexes from a diverse array of related proteins, also allow rapid electrical communication between physically adjacent neurons (Hall 2017). These connection types are discernible from anatomical data. In addition, slower, long-distance connections, typically mediated by secreted neuropeptides and G-protein-coupled receptors, can modulate the excitability and dynamic properties of their target neurons, potentially broadcasting information about internal state to a large number of neurons to influence behavior (Bargmann 2012). The elucidation of the extra-synaptic connectome of the worm, and the

description of its role in nervous system control is another important goal in *C. elegans* systems neuroscience research (Bentley et al. 2016).

Semantic Modelling & Scientific Curation in Biology

Modern biological and biomedical research is a large enterprise. Its largest funding source, the NIH, supports over 2500 labs in the United States and around the world (Owens 2014). A recent report from International Association of Scientific, Technical and Medical Publishers (Johnson et al. 2018) documents the increasingly difficult challenge of keeping up with published literature, even for experts in a particular field. Therefore, there is increasing interest in developing resources to help scholars interact more efficiently and effectively with scholarly output (Landhuis 2016; Nicholson and Greene 2020), with the aid of modern computing and internet resources. The synthesis of scientific information from across the literature is of special importance in biological & biomedical research, in which most publications are focused on a small fraction of the workings of a cell or organism, while the ultimate goal of the research is to understand organisms as entire systems. Therefore, it is crucial that the vast quantity of experimental information in the literature can be integrated and made available to computational analysis. To achieve this, machine-readable representations of biological knowledge are required. Examples of this include the Reactome (Jassal et al. 2020; Good et al. 2021) and KEGG (Kanehisa and Goto 2000) knowledgebases, which contain computable representations of biological pathways.

Recently, the Gene Ontology Consortium (GOC) has approached this problem through the generation of knowledge graphs. For our purposes, a scientific knowledge graph may be described as a graphical representation of the logical relationships between the processes and entities that result from a field of inquiry, producing a model of the process under study. This is sometimes referred to as semantic modelling (Ehrlinger and Wöß 2016). In the framework described below, a knowledge graph approach to biology has grown from genome annotation. A great deal of biological knowledge generation concerns the biochemical functions of

individual genes, and their roles in biological processes. This has resulted in the creation of large databases hosting knowledge about gene function, updated and maintained by professional scientific curators (The Alliance of Genome Resources Consortium 2020). In these databases, genes are associated with objects (formally 'classes') in the Gene Ontology (or GO), a library of documented biological processes, molecular functions and anatomical entities. The GOC has developed a semantic modelling framework to create structured representations of cellular processes by linking GO Molecular Functions to their enabling genes, and to the relevant GO Biological Processes (Fig. 1). This framework is known as Gene Ontology Causal Activity Modelling (GO-CAM) (Thomas et al. 2019). In GO-CAM, nodes can be genes, GO Molecular Functions, GO Biological Process, or GO Anatomy classes, or even classes in other ontologies, such as the Chemical Entities of Biological Interest ontology (ChEBI). The edges describe the logical relationships between them, with rules that govern how they may interact with particular nodes, themselves drawn from a Relations Ontology (RO). This is in contrast to other network representations of biology, in which nodes are typically biological entities, and the edges describe chemical processes that govern their interactions. In GO-CAM, all processes are represented in nodes, where every node corresponds to a class in the GO. These models can be thought of as compositions of assertions in the form of semantic triples (subject-predicate-object)[1]. Each assertion is associated with evidence in the form of quoted author statements, and association with the relevant experimental assay via the Evidence and Conclusion Ontology (ECO) (Giglio et al. 2019). In general, the knowledge graph approach creates a degree of abstraction that allows more flexible and expressive representations of biological knowledge than traditional pathway diagrams.

It is possible to imagine a web-based resource containing knowledge graphs of the entire curated literature in a biological field, that could be subjected to natural language queries (Affolter et al. 2019). The advantage of this mode of interacting with the literature is that information from an arbitrary number of papers can be

[1] For instance, the assertion "[G-protein coupled receptor activity] *has input* [2-heptanone]" is a semantic triple, where *has input* is the predicate.

searched at once, without the need for manual compilation by researchers. Because the contents are curated, they are subject to a process similar to peer review, helping to ensure their quality. Researchers and curators can generate GO-CAM models for cell biology online via the Noctua Curation Platform, which provides a user interface for linking GO classes via relations, and provides feedback about their validity. Currently, there are over 40,000 GO-CAMs in Noctua, at various stages of completion. We envision a future scenario in which the generation of knowledge graphs via GO-CAM (or similar methodology) will become part of the publication process. This will result in up-to-date knowledge integration in line with the rate of primary paper publication, rather than relying on the labor intensive writing of periodic review articles alone.

Semantic Modelling of Neural Circuits & Behavior in *C. elegans*

In this thesis, I describe a project in which we attempted to model author statements from the *C. elegans* neurobiology literature using GO-CAM. We imagined several advantages of modelling neurobiological knowledge with GO-CAM, including the ability of the GO-CAM curation software to interface with existing biological, anatomical, chemical and relations ontologies, which already contain a very large number of neurobiological classes. We explored whether the capacity of GO-CAM to model causal networks of biological processes in an anatomical context could be used to generate knowledge graphs of neural circuit function & its relationship to behavior (*C. elegans* Neural-Circuit Causal Activity Modelling, or *Ce*N-CAM) (Fig. 1C).

In addition to the immediate application in curation, we also envision its application as a basis for functional annotation of the *C. elegans* brain, which could become a useful resource for systems neuroscience in this model. *C. elegans* research has benefited for several decades from electron microscopy data describing the wiring diagram of the hermaphrodite nervous system (White et al. 1986). This ‘connectome’ describes the direct connectivity among 302 neurons via synapses and gap junctions - a useful, but partial description of the full neural network. Additional connectomes of the adult male (Cook et al. 2019) and different stages of hermaphrodite (Witvliet et al. 2020) have yielded a more complete view of the neuroanatomy. The structure

of the *C. elegans* connectome has motivated important questions about nervous system function at a systems level. For instance, the high degree of interconnectedness in the worm nervous system raises the question of how discrete behaviors can be selected and executed in a mutually exclusive manner, and how related behaviors are coordinated (Ji et al. 2021). An important related endeavor is understanding how target selectivity or breadth is achieved using extra-synaptically secreted molecules, e.g., via brain-wide patterns of co-transmission or mutual exclusivity among secreted molecules and receptors. Researchers have also sought to use methods from network analysis to derive insights about brain function from patterns of nervous system organization (Reigl et al. 2004; Jarrell et al. 2012). Answering these and related questions requires a detailed and accurate map of the causal interactions among neurons on the scale of the entire brain, including extra-synaptic connections that cannot be captured by anatomy alone (Schafer 2005). In addition, the conservation of molecules and pathways relevant to nervous system function between *C. elegans* and mammals (Bargmann 1998) underlines the relevance of *C. elegans* research to understanding neurological disease. Therefore, the ability to functionally annotate the connectome in molecular detail in a manner amenable to computation, would be of considerable value.

To test whether the GO-CAM framework could support knowledge graphs of neural circuits, we collected a dataset of around 300 author statements from the literature describing experimentally derived information about the egg- laying and carbon dioxide-sensing circuits. For each author statement, we attempted to generate a simple assertion (i.e., a semantic triple or subject-predicate-object) that accurately modelled the author statement using relations from the Relations Ontology (Smith et al. 2005) and classes from biological ontologies including the GO (Ashburner et al. 2000; The Gene Ontology Consortium 2021) and the *C. elegans* Cell and Anatomy Ontology (Lee and Sternberg 2003). We defined an author statement as a piece of quoted literature describing each of i) an experiment or hypotheses ii) an experimental observation or result and iii) a biological interpretation.

For the egg-laying circuit, these statements largely involve interactions among interneurons, motor neurons, and the egg-laying apparatus, while the CO_2 avoidance circuit is focused on sensory neurons and their interaction with the environment. Our goal was not to generate complete knowledge graphs of the literature on each circuit, or to comprehensively curate particular papers. Rather, we set out to collect a diverse and increasingly challenging set of author statements from a substantial fraction of each literature, and asked whether they could be modelled in a biologically satisfying and semantically rigorous manner.

Summary of Findings (I)

We found that the GO-CAM framework was well-suited to modelling the causal relationships between inputs, neural circuits and behavior. The data model was able to accommodate all statements in our collection, with the addition of new GO terms for several GO Molecular Functions and GO Biological Processes. We were able to categorize these statements into three broad classes (1) *Linking Neurons, Cellular and Molecular Processes, and Behaviors*, (2) *Inputs to Neural Activity & Behavior*, and (3) *Neuron-to-Neuron Functional Connectivity*. These three classes are the primary data types that could be incorporated into a pipeline for functional annotation of the connectome.

A primary bottleneck in scientific curation is that the volume of literature grows at a faster pace than curation can be performed. Therefore, methods to automate curation and knowledge graph generation through text-mining are of increasing interest (Nicholson and Greene 2020). An important issue in this endeavor is ensuring that models generated computationally (e.g., using natural language processing) are biologically rigorous. To help address this challenge, we created generic models for different categories of experimental result, that could accommodate statements from both the egg-laying and CO_2 sensing literature. These models could serve as flexible templates either for automated or manual curation, and provide useful constraints on machine-generated graphs.

We were also able to model more challenging neurobiological phenomena, namely examples of multisensory integration and sensory adaptation. These examples demonstrated how simple computations involving AND logic and negative feedback could be modelled through GO-CAM. These results may be useful in guiding semantic modelling of similar computations in cell biological contexts.

The Dauer Decision

Under favorable environmental conditions, larval development in *C. elegans* proceeds through for stages (L1 through L4) before reaching reproductive maturity, over a period of roughly 3 days, depending on cultivation temperature. Under conditions of food deprivation, high temperature, or crowding, *C. elegans* larvae can adopt an alternative developmental trajectory at the end of the first larval stage (Cassada and Russell 1975; Klass and Hirsh 1976). Post-L1 larvae may enter an alternative L2 stage known as L2d, from which they can either return to normal development via L3, or continue to an alternative L3 stage known as 'dauer' (Fig. 2A). The dauer stage is a reversible developmental pause, allowing adaptation to stressful conditions, and is accompanied by several physiological, morphological and behavioral changes (Fielenbach and Antebi 2008).

Dauer larvae may survive without food for several months, and suffer no change in lifespan upon returning to normal development, suggesting that the dauer represents a phase where normal ageing processes are greatly attenuated (Klass and Hirsh 1976). Morphologically, dauer larvae are distinguished by radial constriction, a thickened cuticle, and closure of the mouth cavity (Cassada and Russell 1975). Germline development is paused, and the nervous system undergoes remodeling, the functional significance of which is poorly understood (Narbonne and Roy 2006; Schroeder et al. 2013). Perhaps the most striking dauer-specific behavior is nictation, in which dozens of larvae aggregate and assemble vertically into a tower-like structure, which moves back and forth in a waving motion (Lee et al. 2012). This behavior is interpreted as an attempt at hitch-hiking on a larger insect, in order to facilitate dispersal to a move favourable environment.

The entry decision is governed by the antagonistic influences of food, which promotes reproduction, versus high temperature and crowding, both of which promote dauer (Golden and Riddle 1984; Ailion and Thomas 2000). In 1982, it was shown that the crowding signal is mediated by a secreted pheromone, that acts in a dose-dependent manner, and that a food signal could antagonize its effects to promote dauer exit, also in a dose-dependent manner (Golden and Riddle 1982). Dauer entry was later shown to be governed by this same food-to-pheromone ratio (Golden and Riddle 1984). In this simple assay, pheromone extracted from a liquid culture of worms is incorporated into the agar medium on which larvae develop from eggs, along with defined quantities of bacterial food. Each larva is simultaneously secreting and detecting the chemical constituents of pheromone, a counting strategy that is directly analogous to bacterial quorum-sensing. In effect, larvae are making a calculation about the future availability of food relative to the requirements of the local population. These collective behaviors are best explained as a mechanism by which very local populations can maximize the fitness of closely related individuals within range of the pheromone signal, rather than as a mechanism for individuals to avoid competition for food, since there is otherwise little fitness advantage for an individual indicate its presence via pheromone (Viney and Franks 2004; Viney and Harvey 2017). Indeed, studies show that microhabitats of wild *C. elegans* are founded by <10 self-fertilizing individuals, with a high degree of relatedness within local populations (Richaud et al. 2018), while potentially manipulative signaling between strains has also been suggested to take place in other nematode species (Bose et al. 2014). Thus, the dauer decision-making process is likely to be a kin-selected collective behavior.

An important and under-studied question in this field comes from the observation that food and pheromone distributions are likely to be non-homogenous in natural environments, meaning that signals of varying duration and magnitude are likely to be experienced transiently as larvae navigate their environment. Over the course of 12 hours during L1, the food-pheromone ratio will therefore fluctuate over short and long timescales. A critical challenge for larvae is to arrive at an accurate average over time for this ratio. The mechanisms that could contribute to this computation are one focus of this thesis project.

Chemical Composition of Dauer Pheromone

The dauer pheromone is composed of a cocktail of chemicals derived from the 5-carbon sugar ascarylose, with great structural and functional variety coming from differences in side chain length and composition (Butcher 2017) (Fig. 2B). These molecules are critical not only for dauer entry and exit, but are also involved in other behaviors such as mating (Srinivasan et al. 2008), developmental timing (Ludewig et al. 2017) and aggregation (Srinivasan et al. 2012). These molecules are known to be of worm origin due to the existence of the mutant strain *daf-22*, from which liquid culture extracts fail to induce dauer formation (Golden and Riddle 1985).

The first ascaroside to be identified (ascr #1, previously known as C3) was isolated via activity-guided fractionation of crude pheromone extract, using a dauer assay as the activity readout (Jeong et al. 2005). However, synthetic ascr#1 could only induce dauer at toxic concentrations, suggesting that the combined or synergistic effects of different dauer pheromone components was required for their effects on dauer (Gallo and Riddle 2009). Since then, additional ascarosides, named ascr#2, ascr#3 and ascr#5, have been shown to have strong effects on dauer formation (Butcher et al. 2007, 2008), with ascr #5 showing synergistic activity with ascr#2 and ascr#3. In order to obtain a more complete view of dauer pheromone composition, NMR spectroscopy was used to compare the *daf-22* strain to wild-type, revealing additional dauer inducing molecules ascr#8 and icas#9 (Pungaliya et al. 2009). Metabolomic investigations into dauer pheromone composition and component activity are ongoing (Ludewig and Schroeder 2013; Butcher 2017).

Genetic & Molecular Basis of the Dauer Decision

The dauer entry assay provides a straightforward method to screen for mutants defective in dauer formation, where the ratio of dauers to adults is altered relative to wild type across several replicates. These screens have produced a detailed model of the molecular pathway leading to dauer formation or commitment to adult development, respectively (Fielenbach and Antebi 2008), from sensory neurons through to downstream

tissues (Fig. 3). This model can be summarized as follows. In food sensing sensory neurons ASI & ASJ, activity-dependent transcription of insulin-like and TFG-beta peptides occurs in response to food sensation, via the *tax-2/tax-4* cGMP gated ion channels, downstream of the *daf-11* guanylate cyclase (Ren et al. 1996; Murakami et al. 2001; Li et al. 2003). These peptides generally promote reproductive development via secretion from these neurons into the body cavity, and eventually interact with the insulin receptor *daf-2* in multiple tissues (Kimura et al. 1997; Apfeld and Kenyon 1998). The most important dauer-inhibiting peptides are the *daf-28, ins-6, ins-4* and *daf-7* gene products. To date, the most potent dauer promoting peptide is *ins-1*, which is also expressed in ASJ (Schackwitz et al. 1996; Pierce et al. 2001). Agonist signaling through the *daf-2* receptor promotes the phosphorylation of the transcription factor DAF-16, which removes it from the nucleus (Ogg et al. 1997; Lee et al. 2001; Lin et al. 2001), where it normally represses a genetic program required for biosynthesis of the steroid hormone, Dafachronic Acid (DA), a ligand of the DAF-12 nuclear hormone receptor, whose signaling activity promotes reproductive development (Antebi et al. 1998; Gerisch et al. 2001; Jia et al. 2002; Gerisch and Antebi 2004; Beckstead and Thummel 2006; Motola et al. 2006; Jeong et al. 2010).

The production of DA is positively regulated by the DAF-12 nuclear hormone receptor, which is itself bound and activated by DA. This positive feedback loop ensures eventual independence from upstream signals of environmental conditions, thereby committing larvae to adult development (Schaedel et al. 2012). In the absence of food-derived signals, a different feedback loop is initiated. In the absence of the DAF-7 ligand, a regulatory pathway involving the SMAD transcription factor DAF-8, which normally inhibits dauer formation by inhibiting a second SMAD transcription factor DAF-3, is inactivated (Park et al. 2010). In turn, up-regulation of SMAD-3 enables it to repress DAF-8, which in turn promotes its own expression. Thus, two feedback loops ensure eventual commitment to reproduction or dauer development, respectively.

In addition, several pheromone receptors, expressed as pairs in individual sensory neurons, have been discovered. These include DAF-37/DAF-38 (ASI neuron), SRBC-64/66 (ASK) SRG-36/SRG-37 (ASI) (Kim

et al. 2009; McGrath et al. 2011; Park et al. 2012). It is expected that many more pheromone receptors remain to be discovered.

Neuronal Basis of the Dauer Decision

In comparison to the molecular genetic pathways governing the decision, the role of the nervous system in the dauer decision has been relatively understudied, partly attributable to the fact that genetic tool for perturbing neurons at high throughput have been developed relatively recently (Husson et al. 2013; Pokala et al. 2014; Wang et al. 2017). Nevertheless, laser ablation studies provided strong evidence of the involvement of some sensory neurons in the decision (Bargmann and Horvitz 1991; Schackwitz et al. 1996). Though a large number of neurons were ablated in total, these experiments mostly ablated combinations of neurons with relatively low throughput, making it difficult to distinguish their individual effects. Clear evidence for a prominent role for ASI and ASJ was provided, along with suggestions that other sensory neurons (ASG, ADF) may be involved. The prominent role of ASJ and ASJ is consistent with previously discussed expression data for reproduction-promoting peptides.

Recently, the Sternberg lab has made use of chemical genetic tools to perturb individual neurons, and measure their effects on dauer formation. In these experiments, a histamine-gated ion channel from the *Drosophila* genome is expressed in a *C. elegans* neuron of interest via cell-specific promoter (Pokala et al. 2014). Since histamine is not synthesized by these worms, it can be added to growth media on a subset of plates to selectively inhibit neurons in these animals. These studies have confirmed the role of ASK and ADL in dauer entry, and also implicated interneurons AIA and AIB (Chai et al. 2022). The activity of these neurons were also investigated via calcium imaging experiments. Here, GCamp (an engineered GFP variant whose fluorescence depends on binding to calcium ions) allowed optical recording of neural activity following the presentation of chemical stimuli, to animals constrained within a microfluidic device (Chronis et al. 2007). This demonstrated that ASK and ADL are each depolarized by crude pheromone extract, while spontaneous AIA activity appears to be suppressed upon pheromone presentation. AIA neuron was found to secrete the

neuropeptide FLP-2 to promote growth via the broadly expressed NPR-30 receptor, and its effects are suppressed by the glutamate receptor MGL-1. The discovery of interneuron involvement in the decision is consistent with various hypotheses under which the decision is governed by a distributed neural circuit that collects integrates various inputs in multiple cells. A neural circuit may exist that implements an algorithm to interpret complex input patterns, such as signal persistence or trends in input frequency. Some interneurons may act as modulators of the response properties of sensory neurons, thereby adjusting their outputs over long timescales.

In this thesis, I performed similar experiments to further elucidate the composition of the neural circuit controlling the dauer decision, including sensory and interneurons. With collaborators, I further investigated the dynamic responses of two neurons to food and pheromone. The findings are summarized in the next section.

Summary of Findings (II)

We performed a screen of several neurons for their role in dauer entry. Among sensory neurons, we confirmed the role of ASI in inhibiting dauer entry. We found no evidence for any effect in ADF, ASG, or ASH. However, we found evidence that BAG promotes dauer entry, and preliminary evidence that URX inhibits it. We also found evidence that a group of neurons involved in mechanosensation (ALM/PLM/AVM/PVM) promote dauer entry. We screened two interneurons, AVK and AIM, and found no evidence that either is involved in regulating dauer entry. These findings suggest that multiple sensory modalities, including gas-sensing and mechanosensation, contribute to dauer. With collaborators, we performed calcium imaging experiments in response to chemical inputs delivered with microfluidic chips. We found that ASI is depolarized by OP50 supernatant, that this depolarized state persists after stimulus removal, and is antagonized by subsequent delivery of crude pheromone extract. We found that AIA neuron, previously shown to inhibit dauer entry, responds to OP50 supernatant and exhibits rapid adaptation, and that the mixture

of food and pheromone stimulus does not reduce the magnitude of the response to food stimulus alone. These findings point to complex properties of individual neurons involved in the entry decision that merit further investigation.

References

Affolter K, Stockinger K, Bernstein A (2019) A comparative survey of recent natural language interfaces for databases. The VLDB Journal 28:793–819. https://doi.org/10.1007/s00778-019-00567-8

Ailion M, Thomas JH (2000) Dauer formation induced by high temperatures in Caenorhabditis elegans. Genetics 156:1047–1067. https://doi.org/10.1093/genetics/156.3.1047

Antebi A, Culotti JG, Hedgecock EM (1998) daf-12 regulates developmental age and the dauer alternative in Caenorhabditis elegans. Development 125:1191–1205

Apfeld J, Kenyon C (1998) Cell Nonautonomy of C. elegans daf-2 Function in the Regulation of Diapause and Life Span. Cell 95:199–210. https://doi.org/10.1016/S0092-8674(00)81751-1

Ashburner M, Ball CA, Blake JA, et al (2000) Gene Ontology: tool for the unification of biology. Nature Genetics 25:25–29. https://doi.org/10.1038/75556

Bargmann CI (1998) Neurobiology of the Caenorhabditis elegans genome. Science (New York, NY) 282:2028–2033. https://doi.org/10.1126/science.282.5396.2028

Bargmann CI (2006) Chemosensation in C. elegans. WormBook : the online review of C elegans biology 1–29. https://doi.org/10.1895/wormbook.1.123.1

Bargmann CI (2012) Beyond the connectome: How neuromodulators shape neural circuits. BioEssays 34:458–465. https://doi.org/10.1002/bies.201100185

Bargmann CI, Avery L (1995) Laser killing of cells in Caenorhabditis elegans. Methods in cell biology 48:225–250. https://doi.org/10.1016/s0091-679x(08)61390-4

Bargmann CI, Horvitz HR (1991) Control of larval development by chemosensory neurons in Caenorhabditis elegans. Science 251:1243–1246. https://doi.org/10.1126/science.2006412

Beckstead RB, Thummel CS (2006) Indicted: worms caught using steroids. Cell 124:1137–1140. https://doi.org/10.1016/j.cell.2006.03.001

Bentley B, Branicky R, Barnes CL, et al (2016) The Multilayer Connectome of Caenorhabditis elegans. PLoS computational biology 12:e1005283. https://doi.org/10.1371/journal.pcbi.1005283

Bose N, Meyer JM, Yim JJ, et al (2014) Natural variation in dauer pheromone production and sensing supports intraspecific competition in nematodes. Current Biology 24:1536–1541. https://doi.org/10.1016/j.cub.2014.05.045

Brenner S (1974) The genetics of Caenorhabditis elegans. Genetics 77:71–94. https://doi.org/10.1093/genetics/77.1.71

Butcher RA (2017) Small-molecule pheromones and hormones controlling nematode development. Nature Chemical Biology 13:577–586. https://doi.org/10.1038/nchembio.2356

Butcher RA, Fujita M, Schroeder FC, Clardy J (2007) Small-molecule pheromones that control dauer development in Caenorhabditis elegans. Nature Chemical Biology 3:420–422. https://doi.org/10.1038/nchembio.2007.3

Butcher RA, Ragains JR, Kim E, Clardy J (2008) A potent dauer pheromone component in Caenorhabditis elegans that acts synergistically with other components. Proceedings of the National Academy of Sciences of the United States of America 105:14288–14292. https://doi.org/10.1073/pnas.0806676105

Cassada RC, Russell RL (1975) The dauerlarva, a post-embryonic developmental variant of the nematode Caenorhabditis elegans. Developmental Biology 46:326–342. https://doi.org/10.1016/0012-1606(75)90109-8

Chai CM, Torkashvand M, Seyedolmohadesin M, et al (2022) Interneuron control of C. elegans developmental decision-making. Current Biology 32:2316-2324.e4. https://doi.org/10.1016/j.cub.2022.03.077

Chalasani SH, Chronis N, Tsunozaki M, et al (2007) Dissecting a circuit for olfactory behavior in Caenorhabditis elegans. Nature 450:63–70. https://doi.org/10.1038/nature06292

Chalfie M, Sulston JE, White JG, et al (1985) The neural circuit for touch sensitivity in Caenorhabditis elegans. Journal of Neuroscience 5:956–964. https://doi.org/10.1523/jneurosci.05-04-00956.1985

Chronis N, Zimmer M, Bargmann CI (2007) Microfluidics for in vivo imaging of neuronal and behavioral activity in Caenorhabditis elegans. Nature Methods 4:727–731. https://doi.org/10.1038/nmeth1075

Chung SH, Sun L, Gabel CV (2013) In vivo neuronal calcium imaging in C. elegans. Journal of visualized experiments : JoVE. https://doi.org/10.3791/50357

Collins KM, Bode A, Fernandez RW, et al (2016) Activity of the C. elegans egg-laying behavior circuit is controlled by competing activation and feedback inhibition. eLife 5:1–24. https://doi.org/10.7554/eLife.21126

Cook SJ, Jarrell TA, Brittin CA, et al (2019) Whole-animal connectomes of both Caenorhabditis elegans sexes. Nature 571:63–71. https://doi.org/10.1038/s41586-019-1352-7

Dickinson DJ, Goldstein B (2016) CRISPR-Based Methods for Caenorhabditis elegans Genome Engineering. Genetics 202:885–901. https://doi.org/10.1534/genetics.115.182162

Ehrlinger L, Wöß W (2016) Towards a Definition of Knowledge Graphs. In: SEMANTiCS

Ellis HM, Horvitz HR (1986) Genetic control of programmed cell death in the nematode C. elegans. Cell 44:817–829. https://doi.org/10.1016/0092-8674(86)90004-8

Faumont S, Lindsay TH, Lockery SR (2012) Neuronal microcircuits for decision making in C. elegans. Curr Opin Neurobiol 22:580–591. https://doi.org/10.1016/j.conb.2012.05.005

Fielenbach N, Antebi A (2008) C. elegans dauer formation and the molecular basis of plasticity. Genes and Development 22:2149–2165. https://doi.org/10.1101/gad.1701508

Fire A (1986) Integrative transformation of Caenorhabditis elegans. EMBO J 5:2673–2680. https://doi.org/10.1002/j.1460-2075.1986.tb04550.x

Fire A, Xu S, Montgomery MK, et al (1998) Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature 391:806–811. https://doi.org/10.1038/35888

Friedland AE, Tzur YB, Esvelt KM, et al (2013) Heritable genome editing in C. elegans via a CRISPR-Cas9 system. Nature Methods 10:741–743. https://doi.org/10.1038/nmeth.2532

Gallo M, Riddle DL (2009) Effects of a Caenorhabditis elegans dauer pheromone ascaroside on physiology and signal transduction pathways. Journal of Chemical Ecology 35:272–279. https://doi.org/10.1007/s10886-009-9599-3

Gerisch B, Antebi A (2004) Hormonal signals produced by DAF-9/cytochrome P450 regulate C. elegans dauer diapause in response to environmental cues. Development 131:1765–1776. https://doi.org/10.1242/dev.01068

Gerisch B, Weitzel C, Kober-Eisermann C, et al (2001) A Hormonal Signaling Pathway Influencing C. elegans Metabolism, Reproductive Development, and Life Span. Developmental Cell 1:841–851. https://doi.org/10.1016/S1534-5807(01)00085-5

Ghosh DD, Nitabach MN, Zhang Y, Harris G (2017) Multisensory integration in C. elegans. Current Opinion in Neurobiology 43:110–118. https://doi.org/10.1016/j.conb.2017.01.005

Giglio M, Tauber R, Nadendla S, et al (2019) ECO, the Evidence & Conclusion Ontology: community standard for evidence information. Nucleic Acids Research 47:D1186–D1194. https://doi.org/10.1093/nar/gky1036

Golden JW, Riddle DL (1984) A pheromone-induced developmental switch in Caenorhabditis elegans: Temperature-sensitive mutants reveal a wild-type temperature-dependent process. Proceedings of the National Academy of Sciences of the United States of America 81:819–823. https://doi.org/10.1073/pnas.81.3.819

Golden JW, Riddle DL (1985) A gene affecting production of the Caenorhabditis elegans dauer-inducing pheromone. Mol Gen Genet 198:534–536. https://doi.org/10.1007/BF00332953

Golden JW, Riddle DL (1982) Pheromone Influences Larval Development in the Nematode C. elegans. Science 218:578–581

Good BM, Van Auken K, Hill DP, et al (2021) Reactome and the Gene Ontology: digital convergence of data resources. Bioinformatics 37:3343–3348. https://doi.org/10.1093/bioinformatics/btab325

Hall DH (2017) Gap junctions in C. elegans: Their roles in behavior and development. Dev Neurobiol 77:587–596. https://doi.org/10.1002/dneu.22408

Husson SJ, Gottschalk A, Leifer AM (2013) Optogenetic manipulation of neural activity in C. elegans: from synapse to circuits and behavior. Biology of the cell 105:235–250. https://doi.org/10.1111/boc.201200069

Jarrell TA, Wang Y, Bloniarz AE, et al (2012) The connectome of a decision-making neural network. Science (New York, NY) 337:437–444. https://doi.org/10.1126/science.1221762

Jassal B, Matthews L, Viteri G, et al (2020) The reactome pathway knowledgebase. Nucleic Acids Res 48:D498–D503. https://doi.org/10.1093/nar/gkz1031

Jeong MH, Kawasaki I, Shim YH (2010) A circulatory transcriptional regulation among daf-9, daf-12, and daf-16 mediates larval development upon cholesterol starvation in Caenorhabditis elegans. Developmental Dynamics 239:1931–1940. https://doi.org/10.1002/dvdy.22322

Jeong PY, Jung M, Yim YH, et al (2005) Chemical structure and biological activity of the Caenorhabditis elegans dauer-inducing pheromone. Nature 433:541–545. https://doi.org/10.1038/nature03201

Ji N, Madan GK, Fabre GI, et al (2021) A neural circuit for flexible control of persistent behavioral states. eLife 10:e62889. https://doi.org/10.7554/eLife.62889

Jia K, Albert PS, Riddle DL (2002) DAF-9, a cytochrome P450 regulating C. elegans larval development and adult longevity. Development 129:221–231. https://doi.org/10.1242/dev.129.1.221

Johnson R, Watkinson A, Mabe M (2018) The STM Report: An overview of scientific and scholarly publishing

Kanehisa M, Goto S (2000) KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Research 28:27–30. https://doi.org/10.1093/nar/28.1.27

Kim K, Sato K, Shibuya M, et al (2009) Two chemoreceptors mediate developmental effects of dauer pheromone in C. elegans. Science 326:994–998. https://doi.org/10.1126/science.1176331

Kimura KD, Tissenbaum HA, Liu Y, Ruvkun G (1997) daf-2, an Insulin Receptor-Like Gene That Regulates Longevity and Diapause in Caenorhabditis elegans. Science 277:942–946. https://doi.org/10.1126/science.277.5328.942

Klass M, Hirsh D (1976) Non-ageing developmental variant of Caenorhabditis elegans. Nature 260:523–525. https://doi.org/10.1038/260523a0

Landhuis E (2016) Scientific literature: Information overload. Nature 535:457–458. https://doi.org/10.1038/nj7612-457a

Lee H, Choi M, Lee D, et al (2012) Nictation, a dispersal behavior of the nematode Caenorhabditis elegans, is regulated by IL2 neurons. Nature Neuroscience 15:107–112. https://doi.org/10.1038/nn.2975

Lee RC, Feinbaum RL, Ambros V (1993) The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell 75:843–854. https://doi.org/10.1016/0092-8674(93)90529-y

Lee RYN, Hench J, Ruvkun G (2001) Regulation of C. elegans DAF-16 and its human ortholog FKHRL1 by the daf-2 insulin-like signaling pathway. Current Biology 11:1950–1957. https://doi.org/10.1016/S0960-9822(01)00595-4

Lee RYN, Sternberg PW (2003) Building a cell and anatomy ontology of Caenorhabditis elegans. Comparative and functional genomics 4:121–126. https://doi.org/10.1002/cfg.248

Li W, Kennedy SG, Ruvkun G (2003) daf-28 encodes a C. elegans insulin superfamily member that is regulated by environmental cues and acts in the DAF-2 signaling pathway. Genes and Development 17:844–858. https://doi.org/10.1101/gad.1066503

Lin K, Hsin H, Libina N, Kenyon C (2001) Regulation of the Caenorhabditis elegans longevity protein DAF-16 by insulin/IGF-1 and germline signaling. Nature Genetics 28:139–145. https://doi.org/10.1038/88850

Liu KS, Sternberg PW (1995) Sensory regulation of male mating behavior in caenorhabditis elegans. Neuron 14:79–89. https://doi.org/10.1016/0896-6273(95)90242-2

Ludewig AH, Gimond C, Judkins JC, et al (2017) Larval crowding accelerates C. elegans development and reduces lifespan. PLoS Genetics 13:1–21. https://doi.org/10.1371/journal.pgen.1006717

Ludewig AH, Schroeder FC (2013) Ascaroside signaling in C. elegans. WormBook : the online review of C elegans biology 1–22. https://doi.org/10.1895/wormbook.1.155.1

McGrath PT, Xu Y, Ailion M, et al (2011) Parallel evolution of domesticated Caenorhabditis species targets pheromone receptor genes. Nature 477:321–325. https://doi.org/10.1038/nature10378

Motola DL, Cummins CL, Rottiers V, et al (2006) Identification of Ligands for DAF-12 that Govern Dauer Formation and Reproduction in C. elegans. Cell 124:1209–1223. https://doi.org/10.1016/j.cell.2006.01.037

Murakami M, Koga M, Ohshima Y (2001) DAF-7/TGF-β expression required for the normal larval development in C. elegans is controlled by a presumed guanylyl cyclase DAF-11. Mechanisms of Development 109:27–35. https://doi.org/10.1016/S0925-4773(01)00507-X

Narbonne P, Roy R (2006) Inhibition of germline proliferation during C. elegans dauer development requires PTEN, LKB1 and AMPK signaling. Development 133:611–619. https://doi.org/10.1242/dev.02232

Nicholson DN, Greene CS (2020) Constructing knowledge graphs and their biomedical applications. Comput Struct Biotechnol J 18:1414–1428. https://doi.org/10.1016/j.csbj.2020.05.017

Ogg S, Paradis S, Gottlieb S, et al (1997) The fork head transcription factor DAF-16 transduces insulin-like metabolic and longevity signals in C. elegans. Nature 389:994–999. https://doi.org/10.1038/40194

Owens B (2014) Mapping biomedical research in the USA. The Lancet 384:11–14. https://doi.org/10.1016/S0140-6736(14)61114-4

Park D, Estevez A, Riddle DL (2010) Antagonistic Smad transcription factors control the dauer/non-dauer switch in C. elegans. Development 137:477–485. https://doi.org/10.1242/dev.043752

Park D, O'Doherty I, Somvanshi RK, et al (2012) Interaction of structure-specific and promiscuous G-protein-coupled receptors mediates small-molecule signaling in Caenorhabditis elegans. Proceedings of the National Academy of Sciences of the United States of America 109:9917–9922. https://doi.org/10.1073/pnas.1202216109

Pierce SB, Costa M, Wisotzkey R, et al (2001) Regulation of DAF-2 receptor signaling by human insulin and ins-1, a member of the unusually large and diverse C. elegans insulin gene family. Genes and Development 15:672–686. https://doi.org/10.1101/gad.867301

Pokala N, Liu Q, Gordus A, Bargmann CI (2014) Inducible and titratable silencing of Caenorhabditis elegans neurons in vivo with histamine-gated chloride channels. Proceedings of the National Academy of Sciences of the United States of America 111:2770–2775. https://doi.org/10.1073/pnas.1400615111

Pungaliya C, Srinivasan J, Fox BW, et al (2009) A shortcut to identifying small molecule signals that regulate behavior and development in Caenorhabditis elegans. Proceedings of the National Academy of Sciences of the United States of America 106:7708–7713. https://doi.org/10.1073/pnas.0811918106

Randi F, Leifer AM (2020) Measuring and modeling whole-brain neural dynamics in Caenorhabditis elegans. Curr Opin Neurobiol 65:167–175. https://doi.org/10.1016/j.conb.2020.11.001

Reigl M, Alon U, Chklovskii DB (2004) Search for computational modules in the C. elegansbrain. BMC Biology 2:25. https://doi.org/10.1186/1741-7007-2-25

Ren P, Lim CS, Johnsen R, et al (1996) Control of C. elegans larval development by neuronal expression of a TGF-β homolog. Science 274:1389–1391. https://doi.org/10.1126/science.274.5291.1389

Richaud A, Zhang G, Lee D, et al (2018) The Local Coexistence Pattern of Selfing Genotypes in Caenorhabditis elegans Natural Metapopulations. Genetics 208:807–821. https://doi.org/10.1534/genetics.117.300564

Richmond J (2006) Synaptic function. WormBook. https://doi.org/10.1895/wormbook.1.69.1

Schackwitz WS, Inoue T, Thomas JH (1996) Chemosensory neurons function in parallel to mediate a pheromone response in C. elegans. Neuron 17:719–728. https://doi.org/10.1016/S0896-6273(00)80203-2

Schaedel ON, Gerisch B, Antebi A, Sternberg PW (2012) Hormonal signal amplification mediates environmental conditions during development and controls an irreversible commitment to adulthood. PLoS Biology 10:1–18. https://doi.org/10.1371/journal.pbio.1001306

Schafer WR (2005) Deciphering the Neural and Molecular Mechanisms of C. elegans Behavior. Current Biology 15:R723–R729. https://doi.org/10.1016/j.cub.2005.08.020

Schroeder NE, Androwski RJ, Rashid A, et al (2013) Dauer-specific dendrite arborization in C. elegans is regulated by KPC-1/Furin. Curr Biol 23:1527–1535. https://doi.org/10.1016/j.cub.2013.06.058

Shao J, Zhang X, Cheng H, et al (2019) Serotonergic neuron ADF modulates avoidance behaviors by inhibiting sensory neurons in C. elegans. Pflugers Arch 471:357–363. https://doi.org/10.1007/s00424-018-2202-4

Smith B, Ceusters W, Klagges B, et al (2005) Relations in biomedical ontologies. Genome Biology 6:R46. https://doi.org/10.1186/gb-2005-6-5-r46

Srinivasan J, Kaplan F, Ajredini R, et al (2008) A blend of small molecules regulates both mating and development in Caenorhabditis elegans. Nature 454:1115–1118. https://doi.org/10.1038/nature07168

Srinivasan J, von Reuss SH, Bose N, et al (2012) A modular library of small molecule signals regulates social behaviors in Caenorhabditis elegans. PLoS Biology 10:1–14. https://doi.org/10.1371/journal.pbio.1001237

Sternberg PW, Horvitz HR (1986) Pattern formation during vulval development in C. elegans. Cell 44:761–772. https://doi.org/10.1016/0092-8674(86)90842-1

Sulston JE, Horvitz HR (1977) Post-embryonic cell lineages of the nematode, Caenorhabditis elegans. Developmental Biology 56:110–156. https://doi.org/10.1016/0012-1606(77)90158-0

Susoy V, Hung W, Witvliet D, et al (2021) Natural sensory context drives diverse brain-wide activity during C. elegans mating. Cell 184:5122-5137.e17. https://doi.org/10.1016/j.cell.2021.08.024

Swierczek NA, Giles AC, Rankin CH, Kerr RA (2011) High-throughput behavioral analysis in C. elegans. Nat Methods 8:592–598. https://doi.org/10.1038/nmeth.1625

The Alliance of Genome Resources Consortium (2020) Alliance of Genome Resources Portal: unified model organism research platform. Nucleic Acids Research 48:D650–D658. https://doi.org/10.1093/nar/gkz813

The C. elegans Sequencing Consortium* (1998) Genome Sequence of the Nematode C. elegans: A Platform for Investigating Biology. Science 282:2012–2018. https://doi.org/10.1126/science.282.5396.2012

The Gene Ontology Consortium (2021) The Gene Ontology resource: enriching a GOld mine. Nucleic Acids Research 49:D325–D334. https://doi.org/10.1093/nar/gkaa1113

Thomas PD, Hill DP, Mi H, et al (2019) Gene Ontology Causal Activity Modeling (GO-CAM) moves beyond GO annotations to structured descriptions of biological functions and systems. Nature genetics 51:1429–1433. https://doi.org/10.1038/s41588-019-0500-1

Towlson EK, Vértes PE, Ahnert SE, et al (2013) The rich club of the C. elegans neuronal connectome. J Neurosci 33:6380–6387. https://doi.org/10.1523/JNEUROSCI.3784-12.2013

Towlson EK, Vértes PE, Yan G, et al (2018) Caenorhabditis elegans and the network control framework-FAQs. Philos Trans R Soc Lond B Biol Sci 373:. https://doi.org/10.1098/rstb.2017.0372

Trent C, Tsuing N, Horvitz HR (1983) Egg-laying defective mutants of the nematode Caenorhabditis elegans. Genetics 104:619–647. https://doi.org/10.1093/genetics/104.4.619

Venkatachalam V, Ji N, Wang X, et al (2016) Pan-neuronal imaging in roaming *Caenorhabditis elegans*. Proceedings of the National Academy of Sciences 113:E1082–E1088. https://doi.org/10.1073/pnas.1507109113

Vidal B, Aghayeva U, Sun H, et al (2018) An atlas of Caenorhabditis elegans chemoreceptor expression. PLoS Biol 16:e2004218. https://doi.org/10.1371/journal.pbio.2004218

Viney M, Harvey S (2017) Reimagining pheromone signaling in the model nematode Caenorhabditis elegans. PLoS Genet 13:e1007046. https://doi.org/10.1371/journal.pgen.1007046

Viney ME, Franks NR (2004) Is dauer pheromone of Caenorhabditis elegans really a pheromone? Naturwissenschaften 91:123–124. https://doi.org/10.1007/s00114-004-0503-2

Wang H, Liu J, Gharib S, et al (2017) cGAL, a temperature-robust GAL4-UAS system for Caenorhabditis elegans. Nat Methods 14:145–148. https://doi.org/10.1038/nmeth.4109

Wang H, Park H, Liu J, Sternberg PW (2018) An efficient genome editing strategy to generate putative null mutants in Caenorhabditis elegans using CRISPR/Cas9. G3: Genes, Genomes, Genetics 8:3607–3616. https://doi.org/10.1534/g3.118.200662

White JG, Southgate E, Thomson JN, Brenner S (1986) The structure of the nervous system of the nematode Caenorhabditis elegans. Philosophical transactions of the Royal Society of London Series B, Biological sciences 314:1–340. https://doi.org/10.1098/rstb.1986.0056

Witvliet D, Mulcahy B, Mitchell JK, et al (2020) Connectomes across development reveal principles of brain maturation in C. elegans bioRxiv 2020.04.30.066209. https://doi.org/10.1101/2020.04.30.066209

Yan G, Vértes PE, Towlson EK, et al (2017) Network control principles predict neuron function in the Caenorhabditis elegans connectome. Nature 550:519–523. https://doi.org/10.1038/nature24056

Yemini E, Lin A, Nejatbakhsh A, et al (2019) NeuroPAL: A neuronal polychromatic atlas of landmarks for whole-brain imaging in C. elegans. bioRxiv. https://doi.org/10.1101/676312

Zhang C, Yan J, Chen Y, et al (2014) The olfactory signal transduction for attractive odorants in Caenorhabditis elegans. Biotechnology Advances 32:290–295. https://doi.org/10.1016/j.biotechadv.2013.10.010

CHAPTER 2

A version of this chapter has been accepted for publication at *Brain Informatics* (PMID: 37947958). I performed the vast majority of the work with co-authors Kimberly Van Auken and David Hill providing key advice on the use of ontologies, and training to understand the Gene Ontology-Causal Activity Modelling framework. Paul Sternberg and I conceived the project.

In this thesis, tables are included at the end of the chapter, while figures can be found in a separate accompanying file.

Chapter 2

SEMANTIC REPRESENTATION OF NEURAL CIRCUIT KNOWLEDGE IN *CAENORHABDITIS ELEGANS*

Abstract

In modern biology, new knowledge is generated quickly, making it challenging for researchers to efficiently acquire and synthesize new information from the large volume of primary publications. To address this problem, computational approaches that generate machine-readable representations of scientific findings in the form of knowledge graphs have been developed. These representations can integrate different types of experimental data from multiple papers and biological knowledge bases in a unifying data model, providing a complementary method to manual review for interacting with published knowledge. The Gene Ontology Consortium (GOC) has created a semantic modelling framework that extends individual functional gene annotations to structured descriptions of causal networks representing biological processes (Gene Ontology Causal Activity Modelling, or GO-CAM). In this study, we explored whether the GO-CAM framework could represent knowledge of the causal relationships between environmental inputs, neural circuits and behavior in the model nematode *C. elegans* (*C. elegans* Neural Circuit Causal Activity Modelling (*Ce*N-CAM)). We found that, given extensions to several relevant ontologies, a wide variety of author statements from the literature about the neural circuit basis of egg-laying and carbon dioxide (CO_2) avoidance behaviors could be faithfully represented with *Ce*N-CAM. Through this process, we were able to generate generic data models for several categories of experimental results. We also generated representations of multisensory integration and sensory adaptation, and we discuss how semantic modelling may be used to functionally annotate the *C. elegans* connectome. Thus, Gene Ontology-based semantic modelling has the potential to support various machine-readable representations of neurobiological knowledge.

Introduction

Caenorhabditis elegans as a Model for Systems Neuroscience

A major goal of modern neuroscience is to explain the relationship between environmental inputs and complex behaviors in terms of the properties of their underlying neural systems. *C. elegans* has been a productive model for neuroscience due to its wide range of easily measured behaviors, genetic tractability, and highly stereotyped anatomy. The function of individual *C. elegans* neurons has been studied by a variety of methods, including selective neuron ablation, either with laser microbeam irradiation (Chalfie et al. 1985; Bargmann and Horvitz 1991; Bargmann and Avery 1995; Liu and Sternberg 1995) or genetically encoded cell killing (Harbinder et al. 1997; Srinivasan et al. 2012). These physical studies, complemented by genetic screens resulting in mutant animals with distinct behavioral and neuronal phenotypes, have implicated specific neurons in behaviors (Bargmann 1993) and identified genes and neurons required for responses to environmental or pharmacological inputs (Waggoner et al. 1998). Technological advances, such as cell-specific application of optogenetic and chemical perturbations (Husson et al. 2013; Pokala et al. 2014) in combination with calcium imaging of individual neurons (Chung et al. 2013), have begun to outline the causal relationships between neurons, both locally and via long-range connections (Shen et al. 2016), while calcium imaging allows the effect of physical inputs on neural activity to be determined. Thus, causal relationships can be traced from inputs through neural circuits to behavior. In addition, traditional molecular genetic methods enable the biochemical basis of these causal relationships to be elucidated. Understanding molecular participants is particularly important for the functional description of extra-synaptic connections because they cannot be described by anatomy or gene expression alone, yet they exert powerful effects on neuronal activity (Bargmann 2012; Marder 2012; Bentley et al. 2016). In combination, the physical and molecular data allow detailed description of *C. elegans* neural circuits underlying particular behaviors.

The GO-CAM framework can be used to Represent Causal Relationships in Biology

Given the volume of biological knowledge, a method to integrate diverse types of data into causal models of biological systems, expressed in a common, machine-readable language, is highly desirable. A promising

method suitable for this application has been developed. The Gene Ontology (GO) Consortium has created a semantic modelling framework for annotating causal relationships between molecular activities in the context of functional gene annotation, known as GO-CAM (Gene Ontology Causal Activity Modelling) (Thomas et al. 2019).

Semantic models (also known as knowledge graphs) are machine-readable representations of knowledge in a given field, in which the edges of the graph describe the logical relationships between entities that comprise a field of study. In GO-CAM, curated knowledge of gene functions annotated using the Gene Ontology and other biologically relevant ontologies are used to create activity flow models of biological systems (Fig. 1) (Le Novère et al. 2009). In these graphs, the logical relationships are described via a formalism known as a semantic triple (subject-predicate-object[2]). These models can be thought of as compositions of assertions in the form of semantic triples. For instance, the assertion "[*G-protein coupled receptor activity* (GO:0004930)] *has input* [*2-heptanone* (CHEBI: 5672)]" is a semantic triple that could be included in a GO-CAM. The semantic triple format allows edges to connect many different kinds of entities, including anatomy terms and biological processes. For instance, "[*glucose-6-phosphate isomerase activity* (GO:0004374)] *part of* [*canonical glycolysis* (GO:0061621)] *occurs in* [*cytosol* (GO:0005829)]" is a pair of semantic triples that connects a GO molecular function to both a higher-level biological process and an anatomical compartment. The Gene Ontology itself follows a hierarchical structure described with semantic triples, e.g., "[*G-protein coupled receptor activity* (GO:0004930)] *is a* [*transmembrane signaling receptor* (GO:0004888)]" (here the relation '*is a*' describes a child-parent relationship). This formalism allows different kinds of entities to be connected to one another in a machine-readable format, allowing combinatorial queries and other computational analyses.

[2] Predicate and relation have the same meaning (predicate is the formal term for describing a triple).

In GO-CAM, curated knowledge of gene functions annotated using the Gene Ontology and other biologically relevant ontologies are used to create knowledge graphs of biological systems (Fig. 1). This framework extends traditional gene function annotation by capturing the causal flow of molecular activities, e.g., protein kinase activity or ion channel activity, using causal relations from the Relations Ontology (RO) and representing these interactions in the context of the relevant biological process and anatomy (Smith et al. 2005) (Box 1). These causal networks allow more in-depth computational analyses of a system than a set of stand-alone associations between genes and ontology terms, and have the potential to bridge the gap between biochemical and anatomical networks. Here, we explored whether the causal GO-CAM framework can enable the representation of the causal relationships between environmental inputs, neural circuits and behavior at varying levels of detail (Fig.1C).

CeN-CAM: GO-CAM Representation of *C. elegans* Neurobiological Knowledge

As for standard GO annotations, assertions in a GO-CAM are supported by evidence statements, ideally experimental evidence from the published literature (Ashburner et al. 2000; Giglio et al. 2019; The Gene Ontology Consortium 2021). To adapt the GO-CAM framework for modelling neurobiological statements about *C. elegans* egg-laying and carbon dioxide (CO_2)-sensing behaviors, we selected a subset of relevant papers from the *C. elegans* bibliography and identified author statements that could be used to support construction of semantically rigorous, causal models. For the egg-laying circuit, these statements largely involve interactions among interneurons, motor neurons,

and the egg-laying apparatus, e.g., vulval muscles and epithelia. The CO_2 avoidance circuit is focused on sensory neurons, their interaction with the environment, and subsequent effects on locomotory behavior.

Methods & Materials

To model neurobiological processes, we began by collecting author statements from published references. In order to ensure that our findings were broadly applicable, we collected statements from the literature on two

circuits, one centered on interneurons and motor neurons (egg-laying) and one centered on sensory neurons (CO_2 avoidance). For the egg-laying circuit we compiled 20 papers, and for CO_2 avoidance, 8 papers. We chose statements manually, according to a few criteria. To begin with, we chose statements that provided a clear interpretation and that we therefore expected to be straightforward to model with GO-CAM. Later, we selected statements describing phenomena (e.g., multi-sensory integration, neuromodulation) that were missing from the initial dataset.

We defined an author statement as text describing: i) either an experiment or hypotheses, ii) an experimental observation or result, and iii) a clear biological interpretation of the result. These typically comprised a paragraph. We then attempted to model the interpretation, along with supporting evidence using the Evidence and Conclusion Ontology (ECO) (Giglio et al. 2019) wherever possible. We avoided modelling speculative suggestions that went beyond the supporting evidence.

For each author statement, we attempted to generate one or more simple assertions (i.e., semantic triples or subject-predicate-object) that accurately modelled the author statement using classes from biological ontologies (Table 1) including the GO (Ashburner *et al.* 2000; The Gene Ontology Consortium 2021), the Chemicals of Biological Interest ontology (ChEBI) (Hastings et al. 2016), the Environmental Conditions, Treatments & Exposures Ontology (ECTO) (Chan et al. 2022), and the *C. elegans* Cell and Anatomy Ontology (WBbt) (Lee and Sternberg 2003). In a semantic triple, these classes are connected by relations from the Relations Ontology (Smith et al. 2005) (Box 1). We collected author statements and their corresponding semantic triples into a dataframe such that the triple representation can be read from left to right (unless otherwise specified). Supplementary Tables 1 and 2 provide the full list of author statements that were modelled for the egg-laying (91 unique statements comprising 128 entries from 20 papers) and CO_2-avoidance circuits (59 unique statements comprising 99 entries from 8 papers), respectively. Table 2 enumerates detailed categories of biological phenomena captured by this approach. We used this categorization process to determine whether existing ontologies contained a sufficiently rich set of classes and whether existing RO

terms were adequate to describe the relations between classes. Where applicable, we generated definitions for required novel classes and their necessary parents (Table 3). We then created illustrations of several useful examples.

In generating our empirical models, we sought as far as possible to ensure that all relations followed the conventions of the GO-CAM data model. Namely, two GO Molecular Functions can be linked by causal relations, whereas a GO Molecular Function (MF) and a GO Biological Process (BP) are linked by mereological relations (e.g., *part of*). In addition, two BPs can be linked by mereological relations when one BPs is part of another BP (i.e., a subprocess of the other). We also found it necessary, in some cases, to link distinct BPs using causal relations to accurately describe the complexity of the biology. For instance, one neuron activating another via optogenetics can be modelled by a *membrane depolarization* process causally upstream of another *membrane depolarization* process. We sought to include whichever MFs or BPs were implied by an author statement, even if the gene was missing, or the BP was not explicitly discussed, in order to denote missing information. We chose the most specific relation or GO term that we felt was justified in the circumstances. For instance, when modelling individual author statements, we used *causally upstream of*, but when modelling compilations of statements from separate papers, we were able to use the child term *positively regulates*. In generating our generic template models, we chose the highest level relations and GO terms that could reasonably represent a given statement category.

Results

CeN-CAM: GO-CAM provides a framework to model neurobiological statements

As a first step in converting information from the scientific literature to a causal model using the GO-CAM framework, we created semantic triples to represent author statements (Supplementary Tables 1 and 2). As an example, a statement by Banerjee *et al.* describing the results of an optogenetic experiment that activates membrane depolarization in uv1 neurons shows that the uv1 cells control the duration of egg deposition during egg-laying behavior. We created a semantic triples to represent this finding: [*membrane depolarization*

(GO:0051899)] *occurs in* [*uv1* (WBbt:0006791)] *part of* [*negative regulation of egg deposition* (GO: proposed)] *part of* [*egg-laying behavior* (GO:0018991)] (Supplementary Table 1, local identifier EL12). In creating triples for 123 egg-laying and 98 CO_2 avoidance author statements, we found that the set of relations used in the GO-CAM data model were sufficient to model all author statements in our dataset. However, we required new classes in several other ontologies (the Gene Ontology (GO), the Evidence & Conclusion Ontology (ECO), and the Environmental Conditions, Treatments & Exposures Ontology (ECTO) (Chan *et al.* 2022)) to describe some statements in both datasets (25/123 statements in the egg-laying dataset, and 84/99 in CO_2 avoidance) (Table 3). These results show that author statements describing *C. elegans* neurobiology can be faithfully captured using the framework of the GO-CAM data model.

We also found it necessary to re-evaluate some existing definitions and classifications of biological processes under the GO class *behavior (GO:0007610)*. For example since the primary term *oviposition* is a subclass of *reproductive behavior (GO:0019098)* in GO and oviposition can be used to describe both the entire behavior of egg laying and to describe the actual deposition of an egg onto a substrate, we requested to switch the primary label of *oviposition (GO:0046662)* with the GO synonym *egg-laying behavior*. We also requested a refinement of the definition of *egg-laying behavior* to 'A reproductive behavior that results in the deposition of eggs (either fertilized or not) upon a surface or into a medium such as water'. In addition, we created a new term *egg deposition (GO:0160027)*, defined as 'The multicellular organismal reproductive process that results in the movement of an egg from within an organism into the external environment'. In this way, the mechanical process of *egg deposition* is clearly distinguished from *egg-laying behavior*, which includes its regulation by the nervous system. We requested new terms for the positive and negative regulation of egg deposition, defined as nervous system processes. In addition, we proposed definitions for new classes required to describe CO_2 avoidance, including *carbon dioxide avoidance behavior* and its parent *behavioral response to carbon dioxide* (Table 3).

Many statements describe findings from genetic perturbations, implicating specific pathways, whereas others, such as cell ablation, leave open a variety of genetic mechanisms by which a phenotype is manifested. Here, we describe the use of different relations and processes to refine models according to the range of conclusions available in each case.

Statement Category: Linking Neurons, Cellular and Molecular Processes, and Behaviors

Fully elucidating functional neural circuits requires an understanding of the cells (e.g., neurons and muscles) involved in the behavior, the molecular basis of the behavior (e.g., the relevant gene products and their activities), and the coordinated relationships among them to affect the behavior. As with all biological processes, however, the full understanding of a neural circuit and a behavior is produced from individual, granular observations that, together and over time, combine to complete the picture. Leading up to a complete understanding, we need to also have the ability to represent the current state of knowledge at the organismal, anatomical, cellular and molecular level. Thus, in our first category of statements, we aimed to capture atomized statements that link cells and genes to cellular and molecular level processes and those processes to a specific behavior.

A traditional experiment for linking neurons to behavior is to ablate a neuron of interest and observe behavioral effects, an experiment that gives us information at the cellular level (Chalfie *et al.* 1985). When an ablation results in a behavioral change, it is interpreted that one or more processes (either in series or in parallel) occurring in that cell has a causal effect on the behavior (Table 4A). Since cell ablation disrupts unknown cellular processes, we chose to model this result using the high level GO biological process term *cellular process (GO:0009987)*, and the *occurs in (BFO:0000066)* relation to contextualize the cellular process with respect to the ablated neuron. We then used the children of the broader causal relation *causally upstream of or within (RO:0002418)* (or preferably a *positive (RO:0004047)* or *negative (RO:0004046)* effect child term) to tie the *cellular process* to a *nervous system process (GO:0050877)* (Fig. 4A) (Table 4A). We used the *part of* relation in cases where more specific perturbations were made (e.g., neuronal activation or

inhibition, genetic knockouts and rescues), allowing an assertion about the composition of the processes involved.

For an illustrative example of this distinction, it is useful to consider experiments from our collection that generated insights by deletion and cell-specific rescue of genes involved in neurotransmitter biosynthesis. We reasoned that since the biosynthesis can proceed even while the neuron is at rest (i.e. independent of the induction of behavior), it should not be considered *part of* the asserted *nervous system process*, but *causally upstream of, positive effect (RO:0002304)* (Fig. 4B, Table 4B) to a secretion process that is *part of* the *nervous system process* (on the assumption that this secretion depends on the depolarization of the neuron). We provide a more detailed explanation of how results of this type are modelled in a later section.

A more recent experimental technology for discerning the effect of neurons on behavior is optogenetic activation. In these experiments, a specific neuron is activated by opening the light-sensitive Channelrhodopsin ion channel, transgenically expressed in specific neurons of interest (Guo et al. 2009). We modelled these results similarly to cell ablation, except that in this case, we were able to say that the *membrane depolarization* that *occurs in* a specific cell is *part of* the *nervous system process* (in this case, the *negative regulation of egg deposition)* that regulates *egg-laying behavior* (Fig. 4C, Table 4C).

Statement Category: Inputs to Neural Activity & Behavior

A second category of experiment provides insight into the molecular basis of behavior or neural activity induced by an environmental or internal stimulus. In this type of study, a behavior or neural activity that is typically induced by some environmental or experimental (i.e. pharmacological) condition is eliminated under the same conditions when a gene is inactivated. The gene activity is often tied to a cell via rescue of a behavioral mutant phenotype by cell-specific expression of the wild-type allele in the loss-of-function background.

In these cases, we can tie the rescue gene functions to cells, (e.g., in Figure 5A, [*G protein-coupled serotonin receptor signaling pathway* (GO:0098664)] *occurs in* [*VM* (WBbt:0006917)]), and to implied GO biological process terms via *part of* (e.g., in Figure 5B [*intracellular receptor signaling pathway* (GO:0030522)] *part of* [*positive regulation of negative chemotaxis* (GO:0050924)]. In contrast to the case of cell ablation, where unknown cellular processes are disrupted, these more specific biological or cellular process terms can in turn be assigned as *part of* the *nervous system process*. Additional ontology terms and relations can be used to further specify processes or functions. For example, the Chemicals of Biological Interest ontology (ChEBI) contains neurotransmitter classes (e.g., *serotonin* (CHEBI:28790)), as well as environmental chemicals (e.g., *carbon dioxide* (CHEBI:16526) which may be linked to GO receptor activities or other GO molecular functions via *has small molecule activator* (RO:0012001) (Fig. 5A. Table 4D, Fig. 5B, Table 4E).

In some cases, the response to a stimulus is measured in a neuron without knowledge of the receptor molecule. For instance, AFD neurons respond to removal of CO_2, but the experiment does not identify the receptor molecule (Bretscher et al. 2011) (Fig 5C) (Table 4J). Because the receptor molecule is unknown, a rescue experiment cannot localise the receptor activity to a cell, meaning that the response may depend on receptor activity in another neuron. This is indicated by the absence of a relationship between the receptor activity and a neuron (similarly, a gene knockout experiment that disrupts neural activity without cell-specific rescue would tie only the membrane depolarization GO term to the neuron). These examples also demonstrate the use of a *nervous system process* term as an intermediate between the *cellular process* terms and the *behavior* terms. For instance, in our model of the role of *npr-1* in the carbon dioxide sensing circuit, a CO_2 receptor activity is implied, but not tied to a gene or cell. However, the *nervous system process* term provides a natural point of integration by which the receptor activity (and by implication, the cell in which it acts) can be included as part of the same neural circuit.

Statement Category: Neuron-to-Neuron Functional Connectivity

An additional type of information necessary for fully modelling neural circuits and behaviors is the functional link between neurons. We were able to model statements describing functional connectivity between neurons. For example, an optogenetic experiment in which one neuron is depolarized by a light stimulus and electrical currents are recorded in another neuron may show how a membrane depolarization process occurring in the upstream neuron results in a subsequent membrane depolarization process in the downstream neuron. To capture this relationship, we can connect two *membrane depolarization* (GO:0051899) processes to one another with the *causally upstream of, positive effect* relation (Fig. 6A, Table 4G).

GO also contains classes sufficient to indicate that the transmission occurs through a synapse, when this is explicitly tested by authors. For instance, Kopchock et al. (2021) showed a synapse-dependent inhibitory connection between HSN and VC, using tetanus toxin to perturb synaptic transmission. This could be modelled using the GO term *chemical synaptic transmission* (GO:0007268) or one of its children, and the *causally upstream of, negative effect (RO:0002305)* relation to describe the inhibition (Fig. 6B, Table 4H). A similar representation would be appropriate for an experiment describing increase or loss of activity from a recorded neuron in mutants defective for synaptic transmission via mutation of *unc-13* (encodes Munc13), which is required for synaptic vessel exocytosis (Richmond et al. 1999). In contrast, mutation of *unc-31* (encodes CAPS), which disrupts dense-core vesicle exocytosis, is required for extra-synaptic transmission (Speese et al. 2007)[3]. GO does not have an explicit term for extra-synaptic signaling, or neuropeptide ligand activity. We include an example representation for an extra-synaptic peptidergic connection between two neurons (Supplementary Figure 2C), and provide a definition for the required new GO classes (Table 3). Finally, we include an example that illustrates how *Ce*N-CAM models can represent sub-cellular phenomena involved in neuron-to-neuron functional connectivity in molecular detail (Fig. 6C) (Table J-L). This model

[3] There is still debate in the literature as to how *unc-13* and *unc-31* may regulate distinct or common processes in synaptic and extra-synaptic transmission (for instance, see Sieburth *et al.* (2007)). In addition, tetanus toxin may disrupt dense core vesicle exocytosis as well as synaptic vesicle exocytosis, as in humans (Hoogstraaten *et al.* 2020). Our modelling here reflects the interpretations of the authors.

compiles findings from Choi et al. (2021), who use the connection between RIA and BAG neurons to investigate mechanisms by which neurotransmitters are loaded into synaptic vesicles.

Generic Data Models for Statement Categories

In modeling author statements, we found it possible to construct models with varying levels of detail, e.g., cell types, gene products, etc. For instance, Figure 6D represents a 'minimal model' of the same statement described in Figure 5B, representing the rescue of CO_2 avoidance by expression of the *npr-1* gene in URX. We sought to provide a set of standards for the ideal model of a given category of experimental finding. In our view, a satisfying model will have a structure that corresponds to the conceptual framework of the field (here, the causal flow from inputs to circuits to behavior), and will explicitly illustrate missing knowledge. By modelling the biology that results from different categories of experimental studies, we were able to produce such generic data models for every category (Supplementary Figures 1, 2). In these models, the availability of GO terms and RO relations is constrained by parentage, i.e. only the generic term in the model or one of its children should be used. Importantly, the models are intended to be flexible, i.e. editable using the Noctua GO-CAM modelling software (Thomas et al. 2019). In particular, high-level *cellular process* and *nervous system process* terms can be attached to as many GO molecular functions and genes as required to represent the biology. These generic models could accommodate results from both the egg-laying and CO_2 circuits, suggesting that they may be more broadly applicable to *C. elegans* neurobiology. These models can serve as useful starting points for researchers or biocurators to generate representations of the experimental results, with minimal prerequisite knowledge of the underlying data model.

GO-CAM can model neural circuits

Systems neuroscience seeks to understand the causal relationships between neural circuits, the behaviors they control, and the inputs that stimulate these circuits, in molecular detail. Having established that a wide variety of author statements describing neurobiological knowledge can be represented in semantic triples, and describing the required GO classes, we generated a model that captures some of the causal relationships within

a single circuit. This graph represents interactions between four of the cells that influence egg-laying behavior, from a limited subset of statements in our collection (Fig. 7).

Though this diagram does not contain all cells, or all known connections that contribute to egg-laying, it illustrates several useful features of using the GO-CAM framework to model this biology. For instance, the influence of AWC in the circuit is connected to the rescue of HSN inhibition through AWC-specific expression of *tax-4* (Fenk and de Bono 2015). This presumably involves chemical output from AWC that depends on its electrical activity; however, the author statement does not assert this specifically. Similarly, the serotonin synthesized in HSN is likely to be causally involved in the activation of VC, via activity-dependent release into the synapse connecting these two neurons, but this has not been demonstrated directly – only that exogenous serotonin can substitute for the absence of HSN, where there is evidence for *tph-1*-dependent serotonin biosynthesis (Zhang et al. 2008). Finally, we used two nodes to represent serotonin, because it allows the possibility that the HSN-VC serotonergic connection may be synaptic, while the HSN-VM connection is extra-synaptic. Thus, *Ce*N-CAM models can represent causal flow within anatomical networks in molecular detail, at the level of what is known, supported by statements in the published literature, and as a result, also indicate what knowledge is missing. In addition, we show that it is possible to use more informative relations in the context of a model that integrates various findings from the egg-laying literature, compared to those used to model individual author statements. In the case of representing author statements, our models were restricted to the use of information contained in those statements. Here, in the larger *Ce*N-CAM model, we are able to use relations that reflect an overall interpretation of the biology, such as *positively regulates (RO:0002213)* (a child of *causally upstream of, positive effect*) to describe interactions between processes in different neurons.

GO-CAM can model simple circuit phenomena

Many studies of neural circuits investigate the mechanistic basis for information processing capabilities in the brain, such as the integration of inputs from multiple sensory modalities, and changes in behavior that depend

on memory of past experience. We extended our modelling efforts to represent some of these findings, primarily from our CO_2 avoidance behavior dataset.

Context-Dependence & Multisensory Integration

An important function of nervous systems in any organism is the ability to execute behavioral responses in a context-dependent manner. This requires integrating multiple kinds of environmental information, 'computing' on that information and eliciting an appropriate response. This integration may commonly be performed either by individual neurons responsive to multiple inputs, or by small circuits of three or more neurons, e.g., single interneurons that integrate input from multiple sensory neurons (Ghosh et al. 2017). Capturing this type of integration requires relations that imply the necessity of multiple conditions toward a single response, sometimes referred to as AND logic. We found a relevant example in our CO_2 avoidance dataset. In one study, *tax-2*-dependent rescue of CO_2 avoidance was found to depend on the presence of food (Bretscher et al. 2011) (Table 5A). We considered whether any of the GO-CAM relations can be interpreted as conveying necessity, in particular the relation *part of*. When considering processes such as those represented in a model, if one process is part of another process, then the latter process necessarily has the former process as a part (or subprocess), meaning that in these contexts *part of* and *has part (BFO:0000051)* are inverse relations (Smith *et al.* 2005). Figures 8A-8C show how the necessity for AND logic might be modelled. The cell-specific rescue of CO_2 avoidance via *tax-2* expression in BAG neurons, along with the inferred CO_2 receptor activity, constitutes one 'branch' of the model. A second 'branch' represents the involvement of food, via an inferred *signal transduction (GO:000716)* process. These two branches converge on a proposed GO term s*ignal integration process* via *part of* relations, capturing their joint necessity. We chose to include a new GO biological process for their integration (rather than having them converge on positive regulation of CO_2 avoidance) in order to represent that the mechanism enabling the AND logic should be asserted. This representation leaves open many possible biological models for the mechanism by which the asserted integration might occur (for example, one in which food and CO_2 are sensed by distinct sensory neurons, and integrated in a third interneuron), while capturing AND logic.

Sensory Adaptation

Some important modelling challenges arose in our attempts to model sensory adaptation. This required representing feedback regulation, and the different qualities of input (in terms of concentration and duration) relevant at different points in the process. Waggoner et al. (2000) describe sensory adaptation to nicotine mediated by down-regulation of the UNC-29 receptor (Table 4G, 4H), in a *tpa-1*-dependent manner. Using this example, we found that feedback is best modelled by creating a duplicate temporal instances of a given input, molecular activity and associated biological process (Fig. 9). One set is used for the initiation of the gene expression changes that mediate desensitization to acetylcholine (nicotine), and another for the altered behavioral response to the same input. It was also necessary to model how this process depends on *tpa-1,* given that the statement doesn't show precisely how the *tpa-1*-enabled activity is related to the *unc-29*-enabled activity. We found that this was best achieved by invoking an instance of p*rotein kinase activity* (GO:0004672) enabled by *tpa-1* that is part of a s*ignal transduction* (GO:0007165) process that is, in turn *part of* an instance of *sensory adaptation to nicotine* (a proposed novel GO Biological Process), and *causally upstream of, positive effect* to [*negative regulation of gene expression* (GO:0010269)]. We then modelled a causal effect of the *acetylcholine receptor signaling pathway* (GO:0095500) on the *tpa-1* linked *signal transduction process* (GO:007165). This completes the causal chain required to represent the feedback inherent in sensory adaptation, while allowing that *tpa-1* may be required in parallel to *unc-29*, rather than downstream. We were able to capture the requirement for sustained input for gene regulation, and the requirement for elevated input concentration in the post-adaptation response using terms from (PATO) and the associated relation *has quality* (RO:0000086). However, as discussed in a later section, this approach has limitations, and we would prefer to develop a systematic set of ontology terms for temporal dimensions of environmental input hosted on ECTO.

Neuromodulation

An important goal of neural modelling is to capture neuromodulatory effects, which may be defined as changes in neuronal excitability or dynamics, due to changes in internal state or external context (Bargmann

2012). We found a small number of entries in our egg-laying dataset that described changes in membrane excitability (e.g., Table 5F). We chose to model these with the GO term *regulation of resting membrane potential* (GO:0060075), with the view that changing the ability of the cell to maintain its resting potential is the primary mechanism for regulating neuronal excitability. However, it may be more appropriate to use a parent GO class that can model changes in excitability, rather than implying any mechanism. For instance, one might imagine induced changes in receptor expression that could alter excitability or responsiveness, without changing the resting membrane potential (Shine et al. 2021). The term *regulation of membrane depolarization* (GO:0003245) and its children may be more appropriate when the mechanism is not known.

Extending Existing Ontology Classes for Modelling Neurobiology

We found that the use of existing ontologies provided the correct classes for building our models of neural circuitry. However, in some cases we found that additional classes would be useful for a complete and accurate description of the type of biology we are modeling. These proposed additional classes would be added to GO, the Evidence and Conclusion Ontology (ECO) and the Environmental Conditions, Treatments and Exposures Ontology (ECTO) and are listed in Table 3. Evidence supported by four types of experiments, chemical inhibition of neurons via histamine chloride (Pokala et al. 2014), inhibition of synaptic transmission (Sweeney et al. 1995), mechanical perturbation, and long-term exposure experiments require additional classes in ECO. The categories below describe the biological phenomena that require new GO Biological Process terms, GO Molecular Function terms and ECTO terms to model. Inclusion of these new classes would enrich the kinds of queries that could be supported by *Ce*N-CAM (for instance, we may want a list of all interneurons whose activity is known to be modulated by peptidergic output from ASI neuron).

Particularly useful would be the addition of the previously mentioned requirement for GO terms describing extra-synaptic neuropeptide signaling and neuropeptide activity. OBO ontologies are carefully managed, and ontology developers provide processes for the addition of new classes. For instance, we were able to add a GO term for *carbon dioxide receptor activity* (GO:0170015) via the GO GitHub repository by providing the

necessary information for its incorporation into the ontology (see Choi et al. (2021)). We discuss other proposed classes below.

Fine Temporal Dynamics of Neural Activity & Behavior

Many statements described neural activity in fine temporal detail. Experimental treatments are sometimes reported to result in changes to either magnitude, duration and/or frequency of membrane depolarization or hyperpolarization (e.g., Table 5H). In some cases, these phenotypes lead authors to the interpretation that these parameters of a neuron's behavior are under selection in wild-type organisms, and required to perform the given behavioral task (for instance, changes in the frequency of calcium transients in neurons of the egg-laying circuit are thought to reflect shifts from 'active' to 'inactive' states of the circuit, reflecting phases of the behavior (Collins et al. 2016). However, the GO class for *membrane depolarization (GO:0051899)* does not distinguish these variations, and related terms such as *positive regulation of membrane depolarization (GO:1904181)* explicitly groups these phenomena together under one term. In the future, it may be useful to have these classes separated into explicit categories for a more comprehensive and informative view of how neural activity is regulated.

Likewise, many assays of egg-laying behavior document its temporal features, dividing it into active and inactive phases, and measuring the effect of various perturbations on their duration and frequency (e.g., Table 5I). In the CO_2 avoidance literature, a small number of entries described fine details in motor output as a result of neuronal perturbations, such as changes in rates of reversal or frequency of omega turns (Bretscher *et al.* 2011) (Table 5K). We were unable to model these features due to a lack of sufficiently fine-grained GO terms in the Biological Process ontology. However, we note that WormBase has a phenotype ontology to describe behavior in many of the appropriate ways (for example *turning frequency increased (WBPhenotype:0002313)*) (Schindelman *et al.* 2011). Since these are mutant phenotypes and not Biological Processes, these ontology classes are a poor fit for *Ce*N-CAM. Conversion of these phenotype classes into meaningful GO Biological Processes would be helpful to create more fine-grained models of behavior.

Temporal Features of Environmental Input

In modelling environmental inputs, we found it necessary to model several temporal features. As discussed previously, our sensory adaptation example required representing sustained input, and the relative decrease in concentration required for the future behavior. We approximated these using PATO terms *chronic (PATO: 0001863)* and *increased amount (PATO:0000470)*, respectively. However, PATO lacked terms required to model changes in input concentration or intensity over time, as required to model the OFF response to CO_2 in ADF neurons (Fig 5B). We found that terms in the Environmental Conditions, Treatments & Exposures Ontology (ECTO) came closer to these requirements (e.g., *exposure to decreased methane (ECTO:4000005)*), but a specific exposure term for many chemicals, such as carbon dioxide, does not exist. We propose and define new classes specifying temporal properties that could be hosted in ECTO (Fig. 10) (Table 3).

Discussion

Given the size, scope and rapid growth of the biological literature, new methods are required to integrate, represent and interpret accumulating knowledge at varying levels of detail. In this work, we demonstrate the applicability of the GO-CAM framework for representing neural circuits in *C. elegans.* By capturing author statements in select papers, we were able to construct simple semantic statements and then link those statements together to begin building causal models of two *C. elegans* behaviors, egg-laying and carbon dioxide avoidance. We found that the existing Relations Ontology (RO) relations used in GO-CAMs are adequate, but new classes are required in several ontologies, including the Gene Ontology (GO), the Evidence & Conclusions Ontology (ECO) and the Experimental Conditions, Treatments & Exposures Ontology (ECTO) to fully represent the statements in our collection.

In general, the GO contains a rich vocabulary for neurobiology, in part due to projects such as SynGO (Koopmans et al. 2019), which expanded GO's representation of synaptic function, and deposited corresponding annotations in the GO repository as GO-CAM models. In addition, the Reactome knowledgebase contains pathways for synaptic transmission, and these have been converted to GO-CAMs

(Good et al. 2021). To complement the synaptic transmission part of the ontology, new terms will be required to describe features of extra-synaptic (i.e. peptidergic) connectivity. We also anticipate a more widespread need to model temporal details of sensory neuron input, since chemotactic behaviors typically involve sensing of spatial gradients, experienced by sensory neurons as change over time, resulting in movement towards or away from the odor source. For instance, the sensory neuron AWA adapts to a given concentration of diacetyl, requiring increasing concentration for continued depolarization and associated positive chemotaxis (Larsch et al. 2015). Adding these temporal details to the inputs of individual neurons would allow for more expressive representations. In addition, many of the GO Biological Process terms that we propose as additions to the Gene Ontology are the result of describing the processes at the level of an organism or cell and are not derived from attempts to annotate gene function. Such temporal details are often derived from phenotypic measurements resulting from non-genetic perturbation (e.g., cell ablation, pharmacological inputs), in anticipation of the involvement of gene activities in the programmed regulation of these processes. In practice, new GO BP terms based on these observations will likely need genetic evidence before they can be included in the GO; however, we include them here as suggestions, which may guide future proposals as the need arises.

In addition, the models presented here go beyond the minimal requirements for the conversion of author statements into semantic triple format. According to our criteria, a satisfying model should reflect the conceptual framework of the field (in this case, representing causal flow from inputs through circuits to behavior). In this way, the models indicate which knowledge is missing. For instance, Figure 3B depicts the role of *npr-1* in the URX neuron in carbon dioxide avoidance behavior. By including a *nervous system process* term indicating the involvement of neural circuit, it is possible to indicate that a carbon dioxide receptor, whose encoding gene and cellular site of action require identification, are part of the circuit. The data modelling work presented here also provided us with an empirical basis for creating generic models or templates for each of the statement categories described above (Supplementary Figs. 1, 2). In constructing these generic models, we followed structures that reflected the relevant conceptual framework into which

particular classes of experimental results should fit. For instance, the full description of a peptidergic connection between neurons should involve the relevant ligand(s), receptor(s), ion channel(s) and encoding genes (Supplementary Fig. 2C). Including the overarching biological process term *neuron-to-neuron signaling by neuropeptide* allows a database to be indexed for these types of connections. In this way, scientists and biocurators can collaborate to generate models with a common understanding of their proper criteria.

We also tried to capture simple 'computations' important for nervous system function, and arrived at some modelling principles that are noteworthy. Firstly, when representing the AND logic involved in multisensory integration, it is important to use relations that convey necessity, and have separate causal flows that converge on a single biological process. We note that the proposed GO Biological Process terms (*signal integration process* and *behavior coordination process*) describe an information processing event that could in theory be carried out via any molecular mechanism that satisfies the task. Representing similar kinds of neurobiological knowledge in the GO may require further understanding of the types of molecular mechanisms that typically underlie this type of nervous system process (Ghosh *et al*. 2017). We also found that, when representing feedback (as in our sensory adaptation example), duplicating the representation of molecular functions and/or biological processes, occurring in defined temporal sequences, is important for clear and consistent modelling.

In this study, we focused on modelling interactions within neural circuits, and their relationship to broad features of behavior, rather than the detailed mechanics of motor programs that they control. In principle, it is possible to link neural activities to the mechanical outputs of neural activity, where both are considered *part of* the organismal behavior under study. In the case of egg-laying, this motor output is simple, involving only the contraction of the vulval muscles. However, CO_2 avoidance involves a complex series of locomotory processes, each of which is regulated by specific patterns of neural activity (for example, see Bretscher *et al.* 2011). As discussed above, inference of new biological process terms by conversion of the appropriate terms from the *C. elegans* Phenotype Ontology will allow modelling of these features of behavior. These motor

outputs could then be modelled as *part of* the organismal behavior *carbon dioxide avoidance behavior* (i.e. they are the targets of the *regulation of chemotaxis* term in the models diagrammed here).

One limitation not previously discussed is that GO-CAM currently has no way of incorporating negative data. In some cases, this prevented documentation of important discoveries from our literature search. For instance, (Shyn et al. 2003) found that in the absence of VC neurons and HSN neurons, spontaneous Ca^{2+} transients continued in the vulval muscles, suggesting that these neurons are not necessary for VM activity (Table 5J). These are arguably important omissions from these knowledge graphs.

With these adjustments, this work demonstrates the possibility of creating a machine-readable knowledge base for neurobiology that can return information based on queries. An important part of this resource will be to generate a representation of the *C. elegans* brain that is computable, since the current anatomy ontology does not contain synaptic or gap junction connections between neurons (Lee and Sternberg 2003). Incorporating connectome data that contains the appropriate neuron to neuron relations and property chain algebra (i.e. (Neuron A synapses to Neuron B) and (Neuron B synapses to Neuron C) implies that (Neuron A connects with Neuron C)) will allow queries that include or depend on synaptic connectivity information.

The application and widespread use of this technology depends on the amount of information incorporated into the knowledgebase, much of which at this point is directly dependent on manual input by curators. Given our definition of an author statement as a passage of text following a stereotyped form (hypothesis, observation, interpretation), it is possible to envision how author statements could be identified automatically. We envision a scenario in which machine intelligence could be applied to identify not only author statements, but identify the category of experiment they describe, and the GO terms that correspond to words within them. Using the generic data models described here as templates could help to ensure that machine-generated models are constrained by a desirable structure. With these capabilities, a large volume of the *C. elegans* neural circuit

literature could potentially be converted into *Ce*N-CAM models computationally. The author statements that we collected as part of this work will serve as training data to pursue this type of approach.

It is also important for biologists to have usable and intuitive ways of interacting with and analyzing synthesized knowledge. One way to achieve this is by representing compiled neurobiological data in an anatomical context. For instance, the Virtual Fly Brain project has used an ontology-based approach to integrate connectivity and single-cell gene expression data, which can be visualized in a 3-dimensional visualization of the brain, using a semantic integration framework (Milyaev et al. 2012; Court et al. 2023). This allows users to run queries to explore gene expression and phenotype data in an anatomic context. We are exploring the possibility of functionally annotating the *C. elegans* connectome in molecular detail using *Ce*N-CAM (Fig. 11). The relevant data are the same as those captured by the statement categories for which we have generated templates, namely causal relationships between inputs to neurons, neurons to behavior, and causal connection between neurons. In addition to populating template data models, the GO terms in the relevant author statements could be used to populate a dataframe of the kind used by visualization software such as Cytoscape (Shannon et al. 2003) (Fig. 10), ideally in an automated manner. This visualization could serve as an intuitive entry point for exploring neural circuit function on a connectome scale, where evidence behind individual elements of the graph could be accessed by linking to the corresponding *Ce*N-CAM models. An anatomical visualization that includes functional and connectivity data would allow predictions to be made about functional relationships between different circuits. For instance, CO_2 has been shown to inhibit egg-laying (Fenk and de Bono 2015) in an AWC-dependent manner. Representing neurons that respond to CO_2 along with neurons that control egg-laying in a connectome context (Fig. 11) suggests that ASH is a CO_2 responsive neuron synaptically linked to HSN. Indeed, ASH was later shown to inhibit both egg-laying and HSN activity (Wen et al. 2020). Functional connectome annotation would also enable various kinds of system-wide analysis of the *C. elegans* brain – a research avenue that has so far been pursued in the absence of functional information (Reigl et al. 2004; Alon 2007; Jarrell et al. 2012).

We also envision the ability to make useful predictions using the underlying semantic models. For instance, the graphs may include causal links between molecular functions and behaviors that result from synthesis of disparate literature, leading to new predictions about how genetic or pharmacological perturbations may affect behavior. Thus, the work described here provides semantically and biologically rigorous foundations for an integrated systems neuroscience resource combining knowledge representation, connectome annotation and associated computational analyses of *C. elegans* nervous system function.

Tables

Table 1: Biological Ontologies Used To Generate *Ce*N-CAM models

GO-CAM element	Ontology	Example
Molecular activity	GO molecular function	serotonin receptor activity (GO:0099589)
Biological process	GO biological process	Membrane depolarization (GO:0051899)
Location	GO cellular component	cytosol (GO:0005829)
Cell	WormBase Anatomy Ontology	HSN (WBbt:0006830)
Active Entity (Gene/Gene Product)	WormBase	*tph-1* (WBGene00006600)
Chemical inputs	Chemical Entities of Biological Interest (ChEBI)	dioxygen (CHEBI:15379)
Relations arrows	Relations Ontology (RO), Basic Formal Ontology (BFO)	occurs in (RO_0002479)
Evidence codes	Evidence & Conclusions Ontology (ECO)	optogenetic evidence used in manual assertion (ECO:0006033)
Environmental Conditions	Phenotype and Trait Ontology (PATO) Environmental Conditions, Treatments and Exposure Ontology (ECTO)	Increased duration (PATO_0000498)

Table 2: Categories of Neurobiological Phenomena Modelled with *Ce*N-CAM

Neuronal basis of behavior	**Receptor or G-protein activity regulates behavior or cell activity**
Neuron regulates behavior	G protein activity in specific neuron regulates behavior
Cellular process regulates behavior	GPCR regulates G -protein-activity in specific neuron to regulate behavior
Neuronal activity regulates behavior	Neuromodulation of specific neuron by G protein signaling
Neuronal activity dependent secretion from identified neuron regulates behavior	G protein activity regulates gene expression
Neuronal activity dependent neurotransmitter secretion	G protein activity regulates gene expression cell autonomously
Neuron-neuron interaction	G protein activity regulates neurotransmitter biosynthesis
Activity of Neuron A regulates activity of Neuron B	G protein activity regulates phospholipase activation
Activity of Neuron A regulates activity of Neuron B (synapse-dependent)	G protein signaling activity regulates neuronal activity cell autonomously
Mechanical stimulation of Neuron A regulates activity of Neuron B	GPCR regulates ion channel
Negative autoregulation of neuronal activity	Ion channel regulates neuronal activity via GPCR
Neural activity depends on extra-synaptic signaling	G protein activity regulates neurotransmitter biosynthesis cell autonomously
Environmental influence on behavior or cell activity	GPCR regulates G -protein-activity
Environmental input regulates behavior	G protein activity regulates neurotransmitter biosynthesis
Environmental input regulates neuronal activity	Neuromodulation of specific neuron by G protein signaling
Environmental condition regulates neuronal activity	Receptor activity regulates neuronal activity cell autonomously
Environmental condition regulates behavior	**Cellular Process**
Mechanical process regulates neural activity	Neurotransmitter biosynthesis
Environmental input regulates behavior via defined neuron	Neurotransmitter signaling pathway affects behavior
Environmental input regulates gene expression	Biochemical process regulates neural activity
Receptor-ligand Interaction	Dense core vesicle exocytosis from identified neuron regulates behavior
Receptor-ligand interaction	Gene activity regulates neural activity
Neurotransmitter regulates neuronal activity via ion channel	Gene activity in identified neuron regulates behavior
Neurotransmitter regulates behavior via specific receptor	Dense core vesicle exocytosis regulates behavior

Neurotransmitter regulates behavior via ion channel in identified neuron	Neuropeptide signaling pathway affects behavior
Neurotransmitter regulates behavior via specific receptor in identified neuron	Neuropeptide signaling pathway affects behavior via identified cell
Neurotransmitter affects identified receptor class	Regulation of gene expression in identified neuron
Ion channel activity regulates behavior or cell activity	**Neurotransmitter/neuropeptide activity regulates behavior or cell activity**
Ion channel regulates neural activity	Neurotransmitter biosynthesis from identified source neuron regulates behavior
Neuromodulation of specific neuron by ion channel activity	Neurotransmitter biosynthesis regulates behavior
Ion channel regulates membrane potential	Neuropeptide from specific neuron regulates behavior
Ion channel activity in defined neuron regulates behavior	Neurotransmitter activity depends on ion channel
Ion channel activity regulates behavior	Neurotransmitter regulates behavior
Nervous System Process	Neurotransmitter regulates neuronal activity
Adaptation to chemical stimulus	Regulation of neurotransmitter activity by upstream neuropeptide activity
Co-ordination of locomotion and neural activity to influence behavior	Regulation of secretion by upstream neuropeptide activity

Table 3: Definitions & Classification for Proposed New GO Classes

Category	**Proposed Modified Classes**	**Modification**	**Classification**
GO Biological Process	*Oviposition (GO:0018991)*	Change primary term name to synonym 'egg-laying behavior' New definition: The muscle system process resulting in the deposition of eggs (either fertilized or not) upon a surface or into a medium such as water.	is a *muscle system process (GO:0003012)*, part of 'egg-laying behavior'
GO Biological Process	*Behavior (GO:0007610)*	New classification (see right)	is a *response to stimulus (GO:0050896)*
Category	**Proposed New Class**	**Definitions**	**Classification**
GO Biological Process	Egg deposition	The multicellular organismal reproductive process that results in the movement of an egg from within an organism into the external environment.	is a *reproductive behavior (GO:00198098);* part of *oviposition/egg-laying behavior (GO:0018991)*
GO Biological Process	Positive regulation of egg deposition	Any process that positively regulates the rate, frequency or extent of egg deposition	is a *nervous system process (GO:0050877);* part of *oviposition/egg-laying behavior (GO:0018991)*
GO Biological Process	Negative regulation of egg deposition	Any process that negatively regulates the rate, frequency or extent of egg deposition	is a *nervous system process (GO:0050877)*; part of *oviposition/egg-laying behavior (GO:0018991)*
GO Biological Process	Neuron-to-neuron extra-synaptic peptide signaling	Any process by which a cellular process within one neuron influences a cellular process within another neuron via a secreted gene product, where this secretion occurs independently of synapses	is a *neuron-to-neuron chemical signaling process*
GO Biological Process	Neuron-to-neuron chemical signaling	Any process by which a cellular process within one neuron influences a cellular process within another neuron via a secreted molecule	is a *cell-cell signaling (GO:0007267)* process
GO Biological Process	Behavioral response to carbon dioxide	The behavior of an organism in response to a carbon dioxide stimulus.	is a *behavior (GO:007610)*; is a *response to carbon dioxide (GO:0010037)*
GO Biological Process	Carbon dioxide avoidance behavior	The behavioral response to carbon dioxide which results in the directed movement of a motile cell or organism towards a lower carbon dioxide concentration	is a *negative chemotaxis (GO:0050919)*; is a *behavioral response to carbon dioxide*
GO Biological Process	Negative regulation of carbon dioxide avoidance behavior	Any process that negatively influences locomotory behavior directed away from a source or gradient of carbon dioxide .	is a *nervous system process (GO:0050877);* part of *behavioral response to carbon dioxide*

GO Biological Process	Positive regulation of carbon dioxide avoidance behavior	Any process that positively influences locomotory behavior directed away from a source or gradient of carbon dioxide.	is a *nervous system process (GO:0050877);* part of *behavioral response to carbon dioxide*
GO Biological Process	Adaptation of neuron to stimulus	Any process that results in an increased threshold for induction of neural activity due to prior exposure to the same stimulus	is a *negative regulation of membrane depolarization (GO1904180)*
GO Biological Process	Sensitization of neuronal response to stimulus	Any process that results in a reduced threshold for induction of neural activity due to prior exposure to the same stimulus.	is a *positive regulation of response to stimulus (GO:0048584)*
GO Biological Process	Sensitization of behavioral response to stimulus	Any process that results in a reduced threshold for induction of behavioral response due to prior exposure to the same stimulus.	is a *positive regulation of behavior (GO:0048520)*
GO Biological Process	Sensory adaptation in behavioral response to stimulus	Any process that results in an increased threshold for induction of behavioral response due to prior exposure to the same stimulus.	is a *negative regulation of behavior (GO:0048521)*
GO Biological Process	Behavior co-ordination process	Any neural process that links the execution or cessation of one behavior to the induction or cessation of another behavior.	is a *regulation of behavior (GO:0050795)*
GO Biological Process	Signal integration Process	Any nervous system process by which different types of input to the nervous system contribute in combination to a behavioral or physiological output.	is a *nervous system process (GO:0050877)*
GO Molecular Function	CO2 receptor activity	Binding to and responding, e.g., by conformational change, to changes in the cellular level of carbon dioxide (CO2) or its dissociation products in water.	is a *signaling receptor activity (GO:0038023)*
ECO Evidence Class	Neuron Chemical Inhibition Assay Evidence used in Manual Assertion	A type of experimental phenotypic evidence that is used in a manual assertion, arising from experiments in which the output from a neuron is inhibited by a chemical	is a *experimental phenotypic evidence (ECO:0000059)*
ECO Evidence Class	Synaptic Transmission Inhibition Evidence used in Manual Assertion	A type of experimental phenotypic evidence, that is used in a manual assertion, arising from experiment in which neuron-to-neuron synaptic transmission is manipulated using inhibitors of synaptic transmission, that is used in a manual assertion.	is a *experimental phenotypic evidence (ECO:0000059)*
ECO Evidence Class	Mechanical Perturbation Evidence used in Manual Assertion	A type of experimental phenotypic evidence, that is used in a manual assertion, arising from experiment in which cellular responses are manipulated using mechanical force, , that is used in a manual assertion.	is a *experimental phenotypic evidence (ECO:0000059)*

ECO Evidence Class	Long-term Exposure or Conditioning Evidence used in Manual Assertion	A type of experimental phenotypic evidence, that is used in a manual assertion, arising from experimental treatment involving sustained exposure of an organism to one or more environmental conditions.	is a *experimental phenotypic evidence (ECO:0000059)*
ECTO Class	exposure to increasing carbon dioxide	An exposure event involving the interaction of an exposure receptor to increasing amount of carbon dioxide	is a *exposure to chemical (ECTO: 0000231)*
ECTO Class	exposure to decreasing carbon dioxide	An exposure event involving the interaction of an exposure receptor to increasing amount of carbon dioxide	is a *exposure to chemical (ECTO: 0000231)*

Table 4: Author Statements Collection A

	Author	Author Statement
A	Waggoner *et al.* 1998	"The roles of individual neurons in controlling the timing of egg-laying events can be determined with high precision by eliminating specific neurons by laser ablation and assaying the effect of the ablation on behavior. We therefore eliminated the neurons with prominent synaptic input to the egg-laying muscles to determine how their absence affected the timing of egg-laying events. We first investigated the involvement of the HSNs, a pair of serotonergic motor neurons that are required for efficient egg laying. By tracking the behavior of animals lacking both HSNs, we found that elimination of the HSNs did not qualitatively alter the pattern of egg laying: eggs were still laid in clusters, and the intervals between clusters and between egg-laying events within a cluster were still exponentially distributed. However, HSN ablation did cause a substantial lengthening of the inactive phase, which led to a slower overall rate of egg laying (Figure 2A). Since loss of the HSNs decreased the frequency of egg-laying clusters (i.e., $\lambda 2$ was decreased; Table 1) but did not slow the egg-laying rate within these clusters ($\lambda 1$ was actually increased), these results suggest that the HSNs stimulate egg laying by inducing the active state."
B	Bany *et al.* 2003	"Because the VC neurons appear to inhibit egg laying and are cholinergic, we tested whether the VCs release acetylcholine to inhibit egg laying. The VCs are the only cells of the egg-laying system that express the UNC-4 complex, CHA-1, and UNC-17 (Lickteig *et al.*, 2001); however, because unc-4, cha-1, and unc-17 are each expressed in other neurons, it was necessary to determine whether mutations in these genes cause hyperactive egg laying specifically attributable to their effects on the VC neurons. For this purpose, we expressed the unc-4, cha-1,or unc-17 cDNAs in the VC neurons and determined whether this rescued the hyperactive egg-laying defects of the corresponding mutants. To direct VC expression, we used a modified lin-11 promoter similar to that used to express GFP in Figure 3A (see Materials and Methods). Expression of the unc-4 cDNA using this promoter rescued the hyperactive egg-laying defect of unc-4 mutants, returning the percentage of early-stage eggs laid to near-wild-type levels (Fig. 4A). Furthermore, expressing the cha-1 cDNA in the VC neurons of cha-1 mutants also rescued their hyperactive egg-laying phenotype (Fig. 4B). Similar experiments with unc-17 gave analogous results (data not shown). Restoring the inhibition of egg laying by restoring the ability of the VC neurons to signal with acetylcholine provides our most compelling evidence that it is the VC neurons that inhibit egg laying."
C	Banerjee *et al.* 2017	"We next sought to determine whether uv1 activation is sufficient to inhibit egg-laying. To address this question, we expressed channelrhodopsin (ChR2) in uv1 cells using the regulatory regions of ocr-2 as above. Light stimulation immediately following the initial egg-laying event of an active phase (see Methods) significantly delays subsequent egg-laying events, and also significantly reduces the total number of egg-laying events within an active phase (Fig 2)(S1 Movie). For example, under control conditions a majority (~80%) of animals show a delay between the first and second egg-laying events within an active phase of <20 s (light stimulation, -ATR) (Fig 2B). This proportion is reduced dramatically (to around 10%) when uv1 cells are activated (light stimulation, +ATR)...Taken together, our findings provide evidence that uv1-mediated inhibition of egg-laying promotes periods of quiescence in the egg-laying program and plays a key role in setting their duration."
D	Carnell *et al.* 2005	"The expression of gfp in the vulval muscles suggests that ser-1 may be acting in vulval muscles to mediate the stimulatory effect of 5-HT on egg laying. To test this hypothesis, we expressed the ser-1 cDNA using the muscle-specific myo-3 promoter (Okkema *et al.*, 1993) to determine whether it could rescue 5-HT-induced egg laying. Consistent with this hypothesis, we found the Pmyo-3::ser-1(+) transgene partially restored 5-HT-dependent egg laying to ser-1(ok345) animals (Fig. 2A). A wild-type ser-1 transgene with the same 3.4 kB promoter that failed to express gfp in the vulval muscles also failed to rescue the egg-laying defects of the ser-1 mutant animals (Fig. 2A). These results indicate that ser-1 expression in muscle can restore egg laying. Previous studies have indicated that 5-HT acts on vulval muscle to stimulate egg laying (Trent *et al.*, 1983; Brundage *et al.*, 1996; Waggoner *et al.*, 1998; Bastiani *et al.*, 2003; Shyn *et al.*, 2003). Our results indicate that ser-1 mediates this response."
E	Carillo *et al.* 2013	"NPR-1 is not expressed in BAG neurons but is expressed in a number of other sensory neurons as well as some interneurons (Macosko *et al.*, 2009). To identify the site of action for the regulation of CO2 response by npr-1, we introduced the N2 allele of npr-1 into npr-1(lf) mutants in different subsets of neurons and assayed CO2 response. We found that expressing npr-1 in neuronal subsets that included the O2-sensing URX neurons (Cheung *et al.*, 2004; Gray *et al.*, 2004) restored CO2 response (Fig. 3A). These results suggest that NPR-1 activity in URX neurons is sufficient to enable CO2 avoidance. However, we cannot exclude the possibility that NPR-1 function in other neurons also contributes to CO2 avoidance."

F	Bretscher *et al*. 2011	"Strikingly, AFD also responded to removal of CO2 with a fast Ca2+ spike that peaked within 10 s ("CO2-OFF" response...CO2-evoked activity in AFD could be due to synaptic input to AFD. To test this, we imaged CO2 responses in unc-13 mutants, which have severe defects in synaptic release (Richmond *et al.*, 1999). The AFD CO2 responses of unc-13 animals were indistinguishable from wild-type (Figures 2H and S1C). These data suggest that, as well as being a thermosensory neuron (Mori and Ohshima, 1995; Kimura *et al*., 2004; Clark *et al*., 2007), AFD is a CO2 sensor with both ON and OFF responses."
G	Collins *et al*. 2016	"To directly test how neurotransmitter signaling from the HSNs regulates egg-laying circuit activity, we used the egl-6 promoter to express Channelrhodopsin-2 in the HSNs (Emtage *et al.*, 2012), allowing us to drive neurotransmitter release specifically from the HSNs with blue light. ...We found that activation of HSNs resulted in circuit activity reminiscent of a spontaneous active state, including rhythmic Ca2+ activity of both VCs and vulval muscles, and egg-laying events that accompanied a subset of these Ca2+ transients... These results suggest that the high level of HSN activity after optogenetic activation induces strong coupling of VC and vulval muscle excitation."
H	Kopchock *et al*. 2021	"To determine whether VC synaptic transmission regulates egg laying via HSN, we recorded HSN Ca2+ activity in WT and transgenic animals expressing TeTx in the VCs (Fig. 6A). During the egg-laying active state, the HSNs drive egg laying during periods of increased Ca2+ transient frequency in the form of burst firing (Fig. 6B) (Collins *et al*., 2016; Ravi *et al.*, 2018a). We observed a significant increase in HSN Ca2+ transient frequency when VC synaptic transmission was blocked compared with nontransgenic control animals (Fig. 6C). WT animals spent ~11% of their time exhibiting high-frequency burst activity in the HSN neurons, whereas transgenic animals expressing TeTx in the VC neurons spent ~21% of their time exhibiting HSN burst firing activity (Fig. 6D). These results are consistent with the interpretation that VC neurotransmission is inhibitory toward the HSNs, such as proposed in previous studies (Bany *et al.*, 2003; Zhang *et al.*, 2008)."
I	Choi *et al*. 2021	"VGLUTs are members of a family of anion transporters that move diverse solutes, including inorganic phosphate, acidic sugars, negatively charged amino acids, and phosphorylated adenosine nucleotides33. As a member of the SLC17 family of transporters, VST-1 is likely an anion transporter and there are different ways an anion transporter in the synaptic vesicle membrane could limit glutamate uptake. ...we used synaptopHluorin to measure vesicular pH in wild-type and vst-1 BAG neurons. Measurements of total and surface-accessible pHluorin (Fig. 3g) allow computation of vesicular pH42...Importantly, we found that loss of VST-1 caused a measurable increase in vesicular pH (Fig. 3h), consistent with a model in which VST-1 supports anion influx into synaptic vesicles. We also measured vesicular pH in BAG neurons lacking EAT-4/VGLUT (Fig. 3h). Unlike loss of VST-1, loss of EAT-4/VGLUT did not cause a measurable change in vesicular pH. The effect of VST-1 mutation on vesicular pH provides additional evidence that VST-1 functions in the synaptic vesicle membrane. These data are also consistent with a model in which VST-1 is an anion transporter that competes with EAT-4/VGLUT for the electrochemical gradient required for glutamate uptake into synaptic vesicles. However, some SLC17 family transporters can cotransport cations such as Na+ and H+ 33, and we cannot rule out the possibility that cation efflux (rather than anion influx) contributes to the effect of VST-1 on vesicular pH."
J	Choi *et al*. 2021	"We further tested whether the effects of vst-1 mutation on RIA activation by BAGs require GLR-1 glutamate receptors, as predicted by our model. In mutants lacking GLR-1, there was no clear effect of vst-1 mutation (Fig. 6d, e), indicating that the increased activation of RIAs observed in vst-1 mutants requires signaling through GLR-1."

Table 5: Author Statements Collection B

	Author	Author Statement
A	Bretscher *et al.* 2011	"When placed in a 5%-0% CO2 gradient, *C. elegans* migrate away from high CO2 (Figures 1A and 1B) (Bretscher *et al.*, 2008). We used this assay to identify potential CO2-sensing neurons..... We next attempted to rescue the tax-2(p694) defect by expressing tax-2 cDNA from neuron-specific promoters, confirming appropriate expression by polycistronic constructs that coexpress tax-2 and gfp (Coates and de Bono, 2002). Expressing tax-2 cDNA in the AFD thermosensory neurons strongly rescued CO2 avoidance, both on and off food (Figure 1D). In contrast, restoring tax-2 to the BAG O2-sensing neurons rescued CO2 avoidance on food, as shown previously (Hallem and Sternberg, 2008), but not off food. Expressing tax-2 cDNA in the ASE taste neurons or in the AQR, PQR, and URX O2-sensing neurons also partially rescued CO2 avoidance, both on food and off food (Figure 1D). These data implicate functionally diverse sensory neurons in CO2 avoidance.""
B	Kopchock *et al.* 2021	"Optogenetic stimulation of the vulval muscles triggered an immediate rise in vulval muscle cytosolic Ca2+, tonic contraction of the vulval muscles, vulval opening, and egg release (Fig. 7B,C). Although optogenetic stimulation resulted in sustained vulval muscle Ca2+ activity and contraction, vulval opening and egg release remained rhythmic and phased with locomotion, as previously observed in WT animals (Collins and Koelle, 2013; Collins *et al.*, 2016). Simultaneous brightfield recordings showed the vulva only opened for egg release when the adjacent ventral body wall muscles were in a relaxed phase (Movie 5). We have previously shown that eggs are preferentially released when the vulva is at a particular phase of the body bend, typically as the ventral body wall muscles anterior to the vulva go into a more relaxed state (Collins and Koelle, 2013; Collins *et al.*, 2016). We now interpret this phasing of egg release with locomotion as evidence that vulval muscle Ca2+ activity drives contraction, but the vulva only opens for successful egg release when contraction is initiated during relaxation of the adjacent body wall muscles. Together, these results show that optogenetic stimulation of the vulval muscles is sufficient to induce vulval muscle Ca2+ activity for egg release in a locomotion phase-dependent manner."
C	Branicky *et al.* 2016	"Because clh-3 encodes chloride channels, we reasoned that it might affect HSN activity by affecting HSN excitability. To test this, we crossed the clh-3 mutants with an integrated transgenic line that expresses Channelrhodopsin-2 (ChR2), the blue-light- activated cation channel (Nagel *et al.*, 2005), in the HSNs (wzIs6 [pegl-6::ChR2]; Leifer *et al.*, 2011; Emtage *et al.*, 2012). In wild- type worms, egg laying is robustly stimulated by ChR2 activation (Leifer *et al.*, 2011; Fig. 7). The magnitude of the response, as indicated by both the percentage of stimulations resulting in egg- laying events and the number of eggs laid per stimulation, is dependent on both the strength and duration of the light stimu- lus (Fig. 7A). The response is also completely dependent on the addition of all-trans retinal, the cofactor for ChR2, to the plates (Fig. 7B), as well as the presence of the HSNs (Fig. 7D). We observed that the clh-3(n995gf) mutant laid significantly fewer eggs per stimulation than the wild-type and blue light stimulation elicited an egg-laying event significantly less frequently in mutant animals than in wild-type. Conversely, the clh-3(ok768 and ok763) mutants laid significantly more eggs than the wild-type and blue light stimulation elicited egg-laying events, including the laying of multiple eggs, more frequently than for the wild- type (Fig. 7C,D). Together, these data support a role for the clh- 3-encoded channels in inhibiting HSN excitability: increased channel activity inhibits HSN excitability whereas loss of the channel promotes HSN excitability."
D	Emtage *et al.* 2012	"Having established a method for exciting the HSN neurons in freely behaving animals, we next tested whether Go signaling controls the sensitivity of the HSNs to ChR2-mediated stimulation. egl-10 encodes an RGS family GTPase-activating protein (GAP) that accelerates hydrolysis of GTP by Goα and thereby antagonizes Go signaling (Koelle and Horvitz, 1996). egl-10 mutants carrying a Promegl-6::ChR2 transgene did not lay eggs in response to a photostimulus that reliably evoked egg-laying behavior when applied to wild-type transgenic animals (Fig. 6E), indicating that globally increasing Go signaling reduced the excitability of the HSN neurons. We next measured the effect of activating Go signaling downstream of the EGL-6 GPCR by testing the behavioral re- sponses of transgenic egl-6(gf) mutants to photostimulation. Like egl-10 mutants, transgenic egl-6(gf) mutants had reduced behavioral responses to photostimulation of HSN neurons (Fig. 6F). Deletion of irk-1 significantly restored the response of egl-6(gf) mutants to excitatory input (Fig. 6F). "

E	Collins *et al.* 2016	"We have previously shown that two Cl- extruding transporters, KCC-2 and ABTS-1, are expressed in the HSNs where they promote the development of inhibitory ligand-gated Cl- channel signaling (Tanis *et al.*, 2009; Bellemer *et al.*, 2011). These data suggest that tyramine signaling through LGC-55 would hyperpolarize the HSN and inhibit activity. To test this directly, we compared HSN activity in wild-type and lgc-55 mutant animals. We observed a significant increase in the frequency of Ca2+ transients in HSNs of lgc-55 mutant animals (Figure 6E and F) in both the inactive and active states of egg-laying behavior. Mean HSN inter-transient intervals in wild-type animals were 41 ± 5 s in the inactive state and 17 ± 2 s during the active state, while intervals in lgc-55 mutants were reduced to 22 ± 2 s in the inactive state and 13 ± 1 s during the active state. Thus, the absence of inhibitory feedback by tyramine signaling onto the HSNs leads to increased activity in both the active and inactive egg-laying behavior states."
F	Bretscher *et al.* 2011	"The timing of CO2-evoked Ca2+ responses in both AFD and BAG correlated with peaks in locomotory activity (Figure 6A). We investigated these correlations directly by ablating AFD and/or BAG and examining behavioral responses (Figure 6B). For statistical comparison, we chose time intervals before and after gas switches according to the occurrence of peaks in wild-type behavioral rates. In the absence of food, neither AFD nor BAG ablation abolished modulation of speed across shifts in CO2 (Figures 6B and S4). Stronger phenotypes were observed for reversal and omega rates (Figure 6B). Unexpectedly, ablation of AFD increased reversal and omega rates following a sharp CO2 rise (ttx-1, Figures Figures6B,6B, B,7B,7B, 7C, 7H, and 7I) and reduced suppression of omega turns following a CO2 fall (ttx-1, Figures Figures6B,6B, B,7K,7K, and 7L), suggesting that AFD acts to suppress reversals and omega turns at these two time points. Ablation of BAG abolished reversal and omega responses to a rise in CO2 (pBAG::egl-1, Figures Figures6B,6B, B,7B,7B, 7C, 7H, and 7I) and reduced the suppression of omega turns following a CO2 fall (pBAG::egl-1, Figures Figures6B,6B, B,7K,7K, and 7L), consistent with BAG excitation promoting reversals and omega turns. Coablation of AFD and BAG abolished the suppression of reversals and omega turns following a fall in CO2 (ttx-1; pBAG::egl-1, Figures 7F and 7L). This effect was due to reduced reversal and omega rates under prolonged high CO2 (ttx-1; pBAG::egl-1, red bars, Figures 7E and 7K). These data suggest that together BAG and AFD act to suppress reversals and omega turns when CO2 decreases."
G	Shyn *et al.* (2003)	"Behavioral data implicated serotonin, a neuromodulator released from the HSN egg-laying motorneurons, in the control of egg-laying behavior 2, 3, 4. When we treated animals with exogenous 5HT, we observed a significant increase in the frequency of Ca2+ events from a baseline of 5.63 min−1 to a rate of 35.01 min−1 ($p < 0.001$, Kolmogorov-Smirnov test). ...Thus, exogenous serotonin appeared to modulate the functional state of the vulval muscles, switching them from a pattern of sporadic Ca2+ activity to a pattern of continual Ca2+ activity. In principle, serotonin could exert its effects directly on the vulval muscles, or it could act indirectly by altering the activity of the egg-laying motorneurons. To resolve this issue, we ablated the egg-laying motorneurons and assayed the effect of serotonin on vulval muscle Ca2+ transients. We found that ablated animals exhibited a continuous train of Ca2+ transients on serotonin essentially identical to that exhibited by unablated wild-type animals (Figure 2, Table 1). Thus, the ability of serotonin to increase the frequency of Ca2+ events was not markedly affected by the absence of the egg-laying motorneurons, indicating that serotonin directly stimulates the activity of the vulval muscles."

Abbreviations

GO (Gene Ontology), GOC (Gene Ontology Consortium), GO-CAM (Gene Ontology Causal Activity Modeling), CeN-CAM (*Caenorhabditis elegans* Neural Circuit Causal Activity Modeling), ChEBI (Chemicals of Biological Interest Ontology), Relations Ontology (RO), Evidence & Conclusions Ontology (ECO), Basic Formal Ontology (BFO), Phenotype and Trait Ontology (PATO), Environmental Conditions, Treatments and Exposure Ontology (ECTO).

Data & Materials Availability

All data generated or analysed during this study are included in this published article (and its supplementary information files).

Author Contributions

P.W.S & S.J.P. conceived and designed the study, S.J.P, K.V.A & D.P.H. conducted the study. S.J.P, K.V.A, D.P.H & P.W.S wrote and revised the manuscript. All authors approved the manuscript.

Acknowledgements

We thank all members of the Sternberg lab at Caltech for their feedback during the course of the project. We also thank Raymond Lee (WormBase) and members of the Gene Ontology Consortium, as well as Susan Bello (Mouse Genome Informatics) and members of the Unified Phenotype Ontology working group for helpful discussions.

Funding

K.V.A. & D.P.H. are funded by the National Human Genome Research Institute (U24HG012212). S.J.P. is funded by NIH U24HG010859-03S2.

References

Alon U (2007) Network motifs: Theory and experimental approaches. Nature Reviews Genetics 8:450–461. https://doi.org/10.1038/nrg2102

Ashburner M, Ball CA, Blake JA, et al (2000) Gene Ontology: tool for the unification of biology. Nature Genetics 25:25–29. https://doi.org/10.1038/75556

Bargmann CI (1993) Genetic and Cellular Analysis of Behavior in C. Elegans. Annual Review of Neuroscience 16:47–71. https://doi.org/10.1146/annurev.ne.16.030193.000403

Bargmann CI (2012) Beyond the connectome: How neuromodulators shape neural circuits. BioEssays 34:458–465. https://doi.org/10.1002/bies.201100185

Bargmann CI, Avery L (1995) Laser killing of cells in Caenorhabditis elegans. Methods in cell biology 48:225–250. https://doi.org/10.1016/s0091-679x(08)61390-4

Bargmann CI, Horvitz HR (1991) Control of larval development by chemosensory neurons in Caenorhabditis elegans. Science 251:1243–1246. https://doi.org/10.1126/science.2006412

Bentley B, Branicky R, Barnes CL, et al (2016) The Multilayer Connectome of Caenorhabditis elegans. PLoS computational biology 12:e1005283. https://doi.org/10.1371/journal.pcbi.1005283

Bretscher AJ, Kodama-Namba E, Busch KE, et al (2011) Temperature, oxygen, and salt-sensing neurons in C. elegans are carbon dioxide sensors that control avoidance behavior. Neuron 69:1099–1113. https://doi.org/10.1016/j.neuron.2011.02.023

Chalfie M, Sulston JE, White JG, et al (1985) The neural circuit for touch sensitivity in Caenorhabditis elegans. Journal of Neuroscience 5:956–964. https://doi.org/10.1523/jneurosci.05-04-00956.1985

Chan LE, Thessen AE, Duncan WD, et al (2022) The Environmental Conditions, Treatments, and Exposures Ontology (ECTO): Connecting Toxicology and Exposure to Human Health and Beyond

Choi J-H, Horowitz LB, Ringstad N (2021) Opponent vesicular transporters regulate the strength of glutamatergic neurotransmission in a C. elegans sensory circuit. Nature Communications 12:6334. https://doi.org/10.1038/s41467-021-26575-3

Chung SH, Sun L, Gabel CV (2013) In vivo neuronal calcium imaging in C. elegans. Journal of visualized experiments : JoVE. https://doi.org/10.3791/50357

Collins KM, Bode A, Fernandez RW, et al (2016) Activity of the C. elegans egg-laying behavior circuit is controlled by competing activation and feedback inhibition. eLife 5:1–24. https://doi.org/10.7554/eLife.21126

Court R, Costa M, Pilgrim C, et al (2023) Virtual Fly Brain—An interactive atlas of the Drosophila nervous system. Frontiers in Physiology 14:

Fenk LA, de Bono M (2015) Environmental CO2 inhibits Caenorhabditis elegans egg-laying by modulating olfactory neurons and evokes widespread changes in neural activity. Proceedings of the National Academy of Sciences of the United States of America 112:E3525-34. https://doi.org/10.1073/pnas.1423808112

Ghosh DD, Nitabach MN, Zhang Y, Harris G (2017) Multisensory integration in C. elegans. Current Opinion in Neurobiology 43:110–118. https://doi.org/10.1016/j.conb.2017.01.005

Giglio M, Tauber R, Nadendla S, et al (2019) ECO, the Evidence & Conclusion Ontology: community standard for evidence information. Nucleic Acids Research 47:D1186–D1194. https://doi.org/10.1093/nar/gky1036

Good BM, Van Auken K, Hill DP, et al (2021) Reactome and the Gene Ontology: digital convergence of data resources. Bioinformatics 37:3343–3348. https://doi.org/10.1093/bioinformatics/btab325

Guo ZV, Hart AC, Ramanathan S (2009) Optical interrogation of neural circuits in Caenorhabditis elegans. Nature Methods 6:891–896. https://doi.org/10.1038/nmeth.1397

Harbinder S, Tavernarakis N, Herndon LA, et al (1997) Genetically targeted cell disruption in Caenorhabditis elegans. Proc Natl Acad Sci U S A 94:13128–13133. https://doi.org/10.1073/pnas.94.24.13128

Hastings J, Owen G, Dekker A, et al (2016) ChEBI in 2016: Improved services and an expanding collection of metabolites. Nucleic acids research 44:D1214-9. https://doi.org/10.1093/nar/gkv1031

Husson SJ, Gottschalk A, Leifer AM (2013) Optogenetic manipulation of neural activity in C. elegans: from synapse to circuits and behavior. Biology of the cell 105:235–250. https://doi.org/10.1111/boc.201200069

Jarrell TA, Wang Y, Bloniarz AE, et al (2012) The connectome of a decision-making neural network. Science (New York, NY) 337:437–444. https://doi.org/10.1126/science.1221762

Koopmans F, van Nierop P, Andres-Alonso M, et al (2019) SynGO: An Evidence-Based, Expert-Curated Knowledge Base for the Synapse. Neuron 103:217-234.e4. https://doi.org/10.1016/j.neuron.2019.05.002

Kopchock RJ, Ravi B, Bode A, Collins KM (2021) The Sex-Specific VC Neurons Are Mechanically Activated Motor Neurons That Facilitate Serotonin-Induced Egg Laying in C. elegans. The Journal of neuroscience : the official journal of the Society for Neuroscience 41:3635–3650. https://doi.org/10.1523/JNEUROSCI.2150-20.2021

Larsch J, Flavell SW, Liu Q, et al (2015) A Circuit for Gradient Climbing in C. elegans Chemotaxis. Cell reports 12:1748–1760. https://doi.org/10.1016/j.celrep.2015.08.032

Le Novère N, Hucka M, Mi H, et al (2009) The Systems Biology Graphical Notation. Nat Biotechnol 27:735–741. https://doi.org/10.1038/nbt.1558

Lee RYN, Sternberg PW (2003) Building a cell and anatomy ontology of Caenorhabditis elegans. Comparative and functional genomics 4:121–126. https://doi.org/10.1002/cfg.248

Liu KS, Sternberg PW (1995) Sensory regulation of male mating behavior in caenorhabditis elegans. Neuron 14:79–89. https://doi.org/10.1016/0896-6273(95)90242-2

Marder E (2012) Neuromodulation of neuronal circuits: back to the future. Neuron 76:1–11. https://doi.org/10.1016/j.neuron.2012.09.010

Milyaev N, Osumi-Sutherland D, Reeve S, et al (2012) The Virtual Fly Brain browser and query interface. Bioinformatics 28:411–415. https://doi.org/10.1093/bioinformatics/btr677

Pokala N, Liu Q, Gordus A, Bargmann CI (2014) Inducible and titratable silencing of Caenorhabditis elegans neurons in vivo with histamine-gated chloride channels. Proceedings of the National Academy of Sciences of the United States of America 111:2770–2775. https://doi.org/10.1073/pnas.1400615111

Reigl M, Alon U, Chklovskii DB (2004) Search for computational modules in the C. elegansbrain. BMC Biology 2:25. https://doi.org/10.1186/1741-7007-2-25

Richmond JE, Davis WS, Jorgensen EM (1999) UNC-13 is required for synaptic vesicle fusion in C. elegans. Nature neuroscience 2:959–964. https://doi.org/10.1038/14755

Shannon P, Markiel A, Ozier O, et al (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome research 13:2498–2504. https://doi.org/10.1101/gr.1239303

Shen Y, Wen Q, Liu H, et al (2016) An extrasynaptic GABAergic signal modulates a pattern of forward movement in Caenorhabditis elegans. eLife 5:. https://doi.org/10.7554/eLife.14197

Shine JM, Müller EJ, Munn B, et al (2021) Computational models link cellular mechanisms of neuromodulation to large-scale neural dynamics. Nature Neuroscience 24:765–776. https://doi.org/10.1038/s41593-021-00824-6

Shyn SI, Kerr R, Schafer WR (2003) Serotonin and Go Modulate Functional States of Neurons and Muscles Controlling C. elegans Egg-Laying Behavior. Current Biology 13:1910–1915. https://doi.org/10.1016/j.cub.2003.10.025

Smith B, Ceusters W, Klagges B, et al (2005) Relations in biomedical ontologies. Genome Biology 6:R46. https://doi.org/10.1186/gb-2005-6-5-r46

Speese S, Petrie M, Schuske K, et al (2007) UNC-31 (CAPS) is required for dense-core vesicle but not synaptic vesicle exocytosis in Caenorhabditis elegans. The Journal of neuroscience : the official journal of the Society for Neuroscience 27:6150–6162. https://doi.org/10.1523/JNEUROSCI.1466-07.2007

Srinivasan J, von Reuss SH, Bose N, et al (2012) A modular library of small molecule signals regulates social behaviors in Caenorhabditis elegans. PLoS Biology 10:1–14. https://doi.org/10.1371/journal.pbio.1001237

Sweeney ST, Broadie K, Keane J, et al (1995) Targeted expression of tetanus toxin light chain in Drosophila specifically eliminates synaptic transmission and causes behavioral defects. Neuron 14:341–351. https://doi.org/10.1016/0896-6273(95)90290-2

The Gene Ontology Consortium (2021) The Gene Ontology resource: enriching a GOld mine. Nucleic Acids Research 49:D325–D334. https://doi.org/10.1093/nar/gkaa1113

Thomas PD, Hill DP, Mi H, et al (2019) Gene Ontology Causal Activity Modeling (GO-CAM) moves beyond GO annotations to structured descriptions of biological functions and systems. Nature genetics 51:1429–1433. https://doi.org/10.1038/s41588-019-0500-1

Waggoner LE, Dickinson KA, Poole DS, et al (2000) Long-term nicotine adaptation in Caenorhabditis elegans involves PKC-dependent changes in nicotinic receptor abundance. The Journal of neuroscience : the official journal of the Society for Neuroscience 20:8802–8811. https://doi.org/10.1523/JNEUROSCI.20-23-08802.2000

Waggoner LE, Zhou GT, Schafer RW, Schafer WR (1998) Control of alternative behavioral states by serotonin in Caenorhabditis elegans. Neuron 21:203–214. https://doi.org/10.1016/S0896-6273(00)80527-9

Wen X, Chen Y-H, Li R, et al (2020) Signal Decoding for Glutamate Modulating Egg Laying Oppositely in Caenorhabditis elegans under Varied Environmental Conditions. iScience 23:101588. https://doi.org/10.1016/j.isci.2020.101588

Zhang M, Chung SH, Fang-Yen C, et al (2008) A Self-Regulating Feed-Forward Circuit Controlling C. elegans Egg-Laying Behavior. Current Biology 18:1445–1455. https://doi.org/10.1016/j.cub.2008.08.047

CHAPTER 3

This chapter includes neuronal silencing experiments that I performed in the Sternberg lab. Calcium imaging experiments that I conceived were performed by Maedeh Seyedolmohasedin in the Venkatachalam lab at Northeastern University. For the imaging experiments, crude pheromone prepared by Mark Zhang was used. The GAL-4-UAS lines for silencing of ASH, BAG and URX were generated by Shahla Gharib and Stephanie Nava, and the line for silencing of the touch response neurons was generated by James Lee.

Chapter 3

NEURONAL BASIS OF THE DAUER ENTRY DECISION IN *CAENORHABDITIS ELEGANS*

Abstract

The model nematode *Caenorhabditis elegans* can choose between normal adult development, or enter a long-term diapause, known as the dauer stage, under conditions of crowding or low food availability. Previous studies showed that the ratio of signals originating from bacterial food versus secreted larval pheromone determined the propensity to enter dauer, and that the decision is under neuronal control, involving both sensory and interneurons. In nature, these environmental cues are likely to fluctuate over long and short timescales, and this information must be collected and compared over a 25 hour period of development. The composition of this circuit and the neuronal mechanisms for the decision-making process are unknown. Here, we use modern genetic tools to silence several sensory and interneurons to further elucidate circuit composition. We confirmed the role of ASI in inhibiting dauer entry, and report the additional involvement of the gas-sensing BAG and URX neurons, as well as a set of mechanosensory neurons. In addition, we interrogated neuronal responses to food and pheromone inputs by measuring calcium traces from ASI and AIA. We found that ASI exhibits a striking memory of food exposure via its electrical activity, which is antagonized by pheromone. In contrast, AIA adapts quickly to food exposure, and is unaffected by pheromone. We discuss how these response properties may inform decision-making in complex natural environments.

Introduction

In many organisms, a single genome can give rise to discrete alternative developmental programs in order to adapt to changes in environmental conditions, a phenomenon known as phenotypic plasticity. In some cases, this results in two or more distinct morphologies, known as polyphenisms. Well-known examples include

caste differentiation in social insects (Wheeler 1986), and discrete variation in male horn size in *Onthophagus* beetles (Emlen 1997). These phenomena imply mechanisms by which environmental conditions are sensed and transduced into alternative developmental trajectories. The dauer diapause stage of *C. elegans* provides an opportunity to study the mechanisms of polyphenism in a genetically tractable organism. In response to crowding, food deprivation or high temperature, larvae undergo a reversible diapause (Cassada and Russell 1975). The genetic and molecular basis of the dauer entry decision has been extensively studied (Hu 2007; Fielenbach and Antebi 2008; Schaedel et al. 2012). However, the neural circuit underlying the decision and the mechanisms underlying the corresponding algorithm are poorly understood. Previous studies using laser ablation implicated sensory neurons ASI, ASG, ADF and ASJ in the dauer entry decision (Bargmann and Horvitz 1991; Schackwitz et al. 1996). However, the technical difficulty of laser ablation made it difficult to isolate the effect of individual neurons at high throughput and thus limits population based assays. Recently, cell-specific expression of the *Drosophila* histamine-gated chloride ion channel HisCl1 was used to pharmacologically silence individual neurons (Pokala et al. 2014; Chai et al. 2022b). This approach allows silencing over timescale of the dauer decision. Plates containing histamine were compared to plates without histamine for rates of dauer entry during exposure to pheromone, implicating interneurons AIA (dauer-inhibiting) and AIB (dauer-inhibiting) in the decision, as well as sensory neurons ADL and ASK (both dauer-inhibiting pheromone sensors). Here, we applied the same genetic tools to examine the additional sensory neurons and interneurons individually. We also examined the dynamic responses of ASI and AIA - both potential sites of food-pheromone signal comparison - to these inputs.

Results

Sensory Neurons and Interneurons Involved in Dauer Entry

We found that among sensory neurons previously implicated, ASI promotes dauer, but found no evidence that ASG or ADF are involved (Fig 13A-13C), in contrast to previous studies (Bargmann and Horvitz 1991). We were unable to generate a robust cell-specific promoter for ASJ. Previous genetic studies in the Sternberg laboratory implicated touch and gas sensation as having a role in dauer decision (J. Lee, P. Shih, C. Chai and

PWS, unpublished observations) and that many neuropeptides and receptors affect dauer formation (Lee et al. 2017; Chai et al. 2022a). We thus also tested neurons involved in other sensory modes, namely gas avoidance, light touch avoidance, and general nociception. Among gas-sensing neurons (Zimmer et al. 2009; Bretscher et al. 2011; Carrillo et al. 2013), we found that BAG and URX promote and inhibit dauer respectively (Fig. 13D, 13E). Finally, the ALM/PLM/AVM/PVM group of neurons (hereafter the TRNs or Touch Response Neurons (Chalfie et al. 1985)) was found to promote dauer (Fig. 13F), while ASH neuron had no effect (Fig. 13G). We also started to test interneurons with interesting connectivity. We tested two additional interneurons, AVK and AIM, and found no evidence for their influence on dauer (Fig. 13H, 13I).

Food sensing neurons ASI and AIA respond to inputs on different timescales

ASI neuron expresses several pheromone receptors (Kim et al. 2009; McGrath et al. 2011; Park et al. 2012). ASI is also depolarized by soluble chemicals present in liquid OP50 solution (Chalasani et al. 2007) and synthesizes peptides that inhibit dauer formation in an activity-dependent manner (Ren et al. 1996; Murakami et al. 2001; Li et al. 2003). Therefore, we reasoned that ASI may compare food and pheromone cues via changes in its electrical activity. To probe this, we imaged calcium dynamics in ASI using GCamp6s expressed in sensory neurons using a microfluidic chip. We found that ASI responds strongly to liquid OP50, such that the extent of depolarization progresses with stimulus duration. It also displayed a striking memory of food exposure, remaining depolarized for at least 15 seconds after the stimulus is removed, perhaps due to an OFF-response (Fig. 14). We also found the initial depolarization is antagonized by subsequent exposure to crude pheromone, consistent with patterns of pheromone receptor expression.

We also examined the responses of AIA interneuron. AIA was previously shown to inhibit dauer entry, and occupies an interesting position in the dauer circuit, since it is anatomically connected to both food-sensing neurons (ASI, AWA) (Larsch et al. 2015) and pheromone-sensing neurons (ADL). In contrast to ASI, AIA adapted quickly to OP50 exposure (Fig. 15, left panel). Previous work had suggested that pheromone inhibits spontaneous AIA activity (Chai et al. 2022b). In our assays, spontaneous activity was lacking, so this was not

observed, nor did we observe any hyperpolarization by pheromone (Fig. 15, middle panel). We further examined the possible inhibition of AIA by pheromone by mixing food and pheromone inputs. We saw no suppression of the AIA food response by pheromone (Fig 15, right panel), suggesting that the role of AIA in dauer (inhibiting dauer formation) is due to its food-sensing role only.

Discussion

In *C. elegans,* the decision to enter larval diapause is under neuronal control. We hypothesized that this requires a distributed neural circuit that can integrate inputs from multiple sensory modalities, and involves memory formation over long timescales. Here, we screened sensory and interneurons for their role in the dauer entry decision, and compare the dynamic responses of two important circuit components.

Among sensory neurons, the finding that ASI inhibits dauer formation is consistent with previous laser ablation studies, its role as a food-sensor, and with established patterns of localized gene expression. The absence of any effect of ASG and ADF suggests that their previously attributed effects are in fact due to ASI (Bargmann and Horvitz 1991). We also found evidence for a role in gas-sensing in dauer entry. URX neurons are normally depolarized by upward shifts in O_2 concentration, while BAG is depolarized by downward shifts in O_2 concentration (Zimmer et al. 2009). BAG (but not URX) is depolarized by pulses of CO_2 (Bretscher et al. 2011; Carrillo et al. 2013). In addition, BAG-expressed peptide FLP-17 is a putative ligand for FRPR-8 (expressed in URX) (Isabel Beets et al. 2022), suggesting a mechanism by which BAG activity may antagonize URX. These opposite effects on neural activity are mirrored by opposite effects on the decision, where URX appears to inhibit dauer entry while BAG activity appears to promote it. More data will be required to confirm the effect of URX on the decision, including in experimental conditions where gas conditions are tightly controlled. However, these data are also consistent with previous observations (data not shown) in which BAG-expressed receptors *gcy-9* and *gcy-31* promote dauer entry, while URX-expressed *gcy-36* inhibits dauer entry. (C. Chai and P. Sternberg, personal communication). Finally, both the *flp-17* and *frpr-18* gene products promote dauer (Lee et al. 2017). These results suggest that either O_2 itself, or the ratio of O_2

and CO_2 may influence the decision, and perhaps that BAG activity can antagonize URX. Intuitively, a higher CO_2 to O_2 ratio may reflect a crowded environment. However, an environment with abundant bacteria may also have a high CO_2 to O_2 ratio, in which case the opposite signs would be expected. Further experiments involving these neurons under different gas conditions, as well as investigations of gas ratios in natural environments, are needed to further explore the role of these neurons.

We also found that a set of neurons (ALM/PLM/AVM/PVM) involved in touch (each of which expresses the *mec-14* gene product required for touch response) also promote dauer entry. These results are consistent with several of our previous observations (data not shown) (J. Lee and P. Sternberg, personal communication). Firstly, mechanical stimulation promotes dauer. In addition, *mec-3* mutants, in which differentiation of TRNs is perturbed, are defective in dauer formation, as are loss-of-function mutants for *mec-4* and *mec-10* (members of the MEC-4/MEC-10/MEC-2/MEC-6 complex, required for mechanosensation (O'Hagan et al. 2005)). Intuitively, more crowded environments should result in more contact with other organisms, resulting in more frequent touch neuron activation. More research is needed to understand patterns of mechanical stimulation in natural environments, how these related to dynamics of these neuron's activities, and how this input is integrated with other sensory modalities. Among the interneurons we tested, AIM and AVK had no effect, despite a reported effect of AVK peptide output on transcription of dauer inhibiting peptides in ASI (Une et al. 2022). Other interneurons implicated in dauer formation are AIA and AIB (Chai et al. 2022a). How interneurons contribute to the decision-making circuit requires further research.

We also examined the dynamic responses of two key neurons of interest, ASI and AIA, to food and pheromone inputs. ASI expresses both food and pheromone receptors, implicating it as a potential site of food-pheromone comparison. Similarly, AIA is downstream of food sensing neurons ASI and AWA, as well as the pheromone sensor ADL. Our imaging experiments revealed contrasting dynamics for different neurons. The ability of crude pheromone to inhibit the food-induced depolarization of ASI reveals that the antagonistic effect of these inputs plays out in the electrical activity of a single neuron. Our finding that ASI exhibits a long-timescale

memory of food is challenging to explain. We speculate this may be part of a mechanism to create long-lasting, activity-dependent changes in circuit properties that could help bridge the short timescales of fluctuating input and the long timescale of the developmental decision. The mechanisms by which a long-lasting depolarized state can be maintained in a single neuron deserves further investigation. In contrast to ASI, AIA adapted quickly to OP50 exposure. Previous work had suggested that pheromone inhibits spontaneous AIA activity (Chai et al. 2022b), but we saw no suppression of the AIA food response by pheromone, suggesting that AIA may not be receptive to pheromone input from sensory neurons. Notably, AIA is connected synaptically and by gap junctions to the volatile sensors AWA, AWB and AWC. We speculate that AIA may contribute to the dauer decision by integrating and transducing volatile food signals received through these neurons.

Materials and Methods

Animal maintenance and strains

Animals were cultivated at 21°C on standard nematode growth media (NGM) plates seeded with *Escherichia coli* OP50 cultured in lysogeny broth (LB). All cGAL strains were generated by crossing UAS effector and GAL-4 driver strains generated by the Sternberg lab. The following strains were used in this study:

Strain Number	Description	Genotype
PS9517	AIM::His-Cl	syIs374[15xUAS::HisCL::SL2::GFP::let-858 3'UTR + unc-22p::GFP]; syIs704 [pdf-1p::NLS::cGAL(DBD)::gp41-1-N-intein::let-858 3'UTR + mbr-1p::NLS::gp41-1-C-intein::cGAL(AD)::let-858 3'UTR + unc-122p::RFP + 1kb DNA ladder (NEB)]; syIs300[15xUAS::GFP, pttx-3::RFP]
PS9845	ASH::His-Cl	syIs374[15xUAS::HisCL::SL2::GFP::let-858 3'UTR + unc-22p::GFP]; syIs337 [15xUAS::GFP]; syIs1901 [nlp-76p::NLS::GAL4(sk)::VP64::let-858 3'UTR + unc-122p::RFP]
PS9925	URX::His-Cl	syIs374[15xUAS::HisCL::SL2::GFP::let-858 3'UTR + unc-22p::GFP]; syIs300[15xUAS::GFP, pttx-3::RFP]; syIs678 [flp-

		8p::NLS::GAL4(sk)::VP64::let-858 3'UTR + unc-122p::RFP + 1kb DNA ladder (NEB)]
PS9940	BAG::His-Cl	syIs374[15xUAS::HisCL::SL2::GFP::let-858 3'UTR + unc-22p::GFP]; syIs300[15xUAS::GFP, pttx-3::RFP]; syIs682 [pgcy-17-GAL4(sk)-VP64 + unc-22::RFP]
PS7334	AIA::His-Cl	syIs371 [15xUAS::HisCL::SL2::GFP::let-858 3'UTR + unc-22p::GFP], syIs448 [gcy-28d::GAL4 - VP64 - let 858 3'UTR]; syIs300[15xUAS::GFP, pttx-3::RFP]
PS9370	ADF::His-Cl	syIs374[15xUAS::HisCL::SL2::GFP::let-858 3'UTR + unc-22p::GFP]; syIs591[psrh-142-Gal4(sk)-VP64 + unc-122p::RFP]
PS9349	ASG::His-Cl	syIs371 [15xUAS::HisCL::SL2::GFP::let-858 3'UTR + unc-22p::GFP]; syIs300[15xUAS::GFP, pttx-3::RFP]; sIs741 [pY41C4A.2-GAL4(sk)-VP64 + unc-22p::GFP]
PS9190	ASI::His-Cl	syIs371[15xUAS::HisCL::SL2::GFP::let-858 3'UTR + unc-22p::GFP]; syIs686 [pgpa-4-GAL4(sk)-VP64 + unc-22::rfp]
PS10239	AVK::His-Cl	syIs374[15xUAS::HisCL::SL2::GFP::let-858 3'UTR + unc-22p::GFP]; syIs300[15xUAS::GFP, pttx-3::RFP]; syIs589 [15xUAS::wrmScarlet::let-858 3'UTR + ttx-3p::RFP + 1kb DNA ladder (NEB)]
PS8323	TRN::His-Cl	syIs371[15xUAS::HisCL::SL2::GFP::let-858 3'UTR, marker: coel::GFP]; syIs#597[mec-17p::cGAL, ofm-1p::rfp]

Pheromone Preparation

Crude pheromone was prepared as described previously (Schroeder and Flatt 2014). Briefly, *C. elegans* were cultured in flasks over several days until starved. Worms were separated from liquid by centrifugation and filtration. Crude pheromone was isolated by heating and ethanol extraction, and resuspended in water.

Dauer assays

Dauer assays were performed as described previously (Lee et al. 2017). On day 1, histamine chloride inhibition plates were prepared as follows. For each plate, 2ml of peptone-free Nematode Growth Medium (NGM) was mixed with either 6ul or 10ul of crude pheromone extract. For experimental plates, 30-40mg histamine dihydrochloride powder was first dissolved in 15ml NGM solution at 60°C, while control plates lacked histamine. Four plates were prepared per treatment for each experiment. One control plate lacking

pheromone, but containing histamine was also prepared for each experiment from the same histamine (+) NGM dilution. Roughly 100 L4 adults were picked onto new seeded plates, and a single colony of OP50 was used to inoculate LB at 37°C overnight. On day 2, OP50 was concentrated to 8% w/v in S-basal. 2ul of this OP50 was added to each plate to allow 80 adults to lay 70-90 eggs at 25°C. As a control, the line PS8720, expressing the His-Cl effector transgene under the *myo-3* promoter was transferred to the histamine (+) control plate. Successful inhibition results in cessation of locomotion in these animals. Remaining OP50 was heat killed at 95°-100°C for 10 minutes, and 18ul added per plate. Once dried, all plates were parafilmed and moved to 25°C for 72 hours. Dauer and non-dauer worms were counted for each plate. All statistical analysis was performed with the scipy Python package.

Microfluidic device fabrication

The Venkatachalam lab designed a 2-layer microfluidic chip capable of delivering sequences of stimuli with a worm trap suitable for housing worms at L4 larval stage. The chip was designed in AutoCAD software, and sent to Artnet Pro Inc. for photomask printing. Photolithography in a clean room was performed on a silicon wafer to make the 2-layer mold from the photomask. For the first layer, which included the worm trap, SU-8 2025 was spin coated on the silicon wafer at 4000 rpm to achieve 25 μm thickness. For the second layer, the same photoresist was spin coated at 1250 rpm for a thickness of 70 μm. Polydimethylsiloxane (PDMS) was poured over the mold and cured on a 90˚C hotplate to solidify. Each PDMS chip was then punched with a 1 mm biopsy punch and was bonded to a cover slip using a handheld corona treater.

Calcium imaging

L4 stage animals were assayed and placed in the microfluidic device. For each experiment, *E. coli* OP50 was cultured overnight in LB, and its supernatant was collected. Crude pheromone extract was diluted to a concentration of 2.5% (v/v) in either buffer (H2O) or OP50 supernatant. Each stimulus was delivered to the animal's nose for a duration of 15 seconds, followed by either buffer (H2O) or pheromone. Fluorescence was recorded with a spinning disc confocal microscope (Dragonfly 200, Andor) and a sCMOS camera

(Photometrics Kinetix). The fluorescence was captured from GCaMP6s at a rate of 10 ms per 1.0 μm z-slice, with 25 z-slices per volume and 4 volumes per second. To extract calcium activity from the recorded data, we performed the following steps: 1. Background intensity was subtracted from each recorded volume. 2. The center of the ROI was annotated at one timepoint, and then the center was tracked throughout the entire recording using the Zephir tracking algorithm (Yu, 2022). 3. For ASI neurons, the ROI was defined as the neuronal nucleolus, and for AIA neurons, the ROI was defined as the processes located in the gap junction. 4. Average pixel intensity from each ROI was calculated. ΔF/F0 was computed, where F0 was defined as the average intensity during the 5-second window preceding stimulus delivery.

References

Bargmann CI, Horvitz HR (1991) Control of larval development by chemosensory neurons in Caenorhabditis elegans. Science 251:1243–1246. https://doi.org/10.1126/science.2006412

Bretscher AJ, Kodama-Namba E, Busch KE, et al (2011) Temperature, oxygen, and salt-sensing neurons in C. elegans are carbon dioxide sensors that control avoidance behavior. Neuron 69:1099–1113. https://doi.org/10.1016/j.neuron.2011.02.023

Carrillo MA, Guillermin ML, Rengarajan S, et al (2013) O2-sensing neurons control CO2 response in C. elegans. J Neurosci 33:9675–9683. https://doi.org/10.1523/JNEUROSCI.4541-12.2013

Cassada RC, Russell RL (1975) The dauerlarva, a post-embryonic developmental variant of the nematode Caenorhabditis elegans. Developmental Biology 46:326–342. https://doi.org/10.1016/0012-1606(75)90109-8

Chai CM, Park H, Sternberg PW (2022a) Brain-wide bidirectional neuropeptide modulation of individual neuron classes regulates a developmental decision. Curr Biol 32:3365-3373.e6. https://doi.org/10.1016/j.cub.2022.05.048

Chai CM, Torkashvand M, Seyedolmohadesin M, et al (2022b) Interneuron control of C. elegans developmental decision-making. Current Biology 32:2316-2324.e4. https://doi.org/10.1016/j.cub.2022.03.077

Chalasani SH, Chronis N, Tsunozaki M, et al (2007) Dissecting a circuit for olfactory behavior in Caenorhabditis elegans. Nature 450:63–70. https://doi.org/10.1038/nature06292

Chalfie M, Sulston JE, White JG, et al (1985) The neural circuit for touch sensitivity in Caenorhabditis elegans. Journal of Neuroscience 5:956–964. https://doi.org/10.1523/jneurosci.05-04-00956.1985

Emlen DJ (1997) Alternative reproductive tactics and male-dimorphism in the horned beetle Onthophagus acuminatus (Coleoptera: Scarabaeidae). Behavioral Ecology and Sociobiology 41:335–341. https://doi.org/10.1007/s002650050393

Fielenbach N, Antebi A (2008) C. elegans dauer formation and the molecular basis of plasticity. Genes and Development 22:2149–2165. https://doi.org/10.1101/gad.1701508

Hu PJ (2007) Dauer. WormBook : the online review of C elegans biology 1–19. https://doi.org/10.1895/wormbook.1.144.1

Isabel Beets, Sven Zels, Elke Vandewyer, et al (2022) System-wide mapping of neuropeptide-GPCR interactions in C. elegans. bioRxiv 2022.10.30.514428. https://doi.org/10.1101/2022.10.30.514428

Kim K, Sato K, Shibuya M, et al (2009) Two chemoreceptors mediate developmental effects of dauer pheromone in C. elegans. Science 326:994–998. https://doi.org/10.1126/science.1176331

Larsch J, Flavell SW, Liu Q, et al (2015) A Circuit for Gradient Climbing in C. elegans Chemotaxis. Cell reports 12:1748–1760. https://doi.org/10.1016/j.celrep.2015.08.032

Lee JS, Shih PY, Schaedel ON, et al (2017) FMRFamide-like peptides expand the behavioral repertoire of a densely connected nervous system. Proceedings of the National Academy of Sciences of the United States of America 114:E10726–E10735. https://doi.org/10.1073/pnas.1710374114

Li W, Kennedy SG, Ruvkun G (2003) daf-28 encodes a C. elegans insulin superfamily member that is regulated by environmental cues and acts in the DAF-2 signaling pathway. Genes and Development 17:844–858. https://doi.org/10.1101/gad.1066503

McGrath PT, Xu Y, Ailion M, et al (2011) Parallel evolution of domesticated Caenorhabditis species targets pheromone receptor genes. Nature 477:321–325. https://doi.org/10.1038/nature10378

Murakami M, Koga M, Ohshima Y (2001) DAF-7/TGF-β expression required for the normal larval development in C. elegans is controlled by a presumed guanylyl cyclase DAF-11. Mechanisms of Development 109:27–35. https://doi.org/10.1016/S0925-4773(01)00507-X

O'Hagan R, Chalfie M, Goodman MB (2005) The MEC-4 DEG/ENaC channel of Caenorhabditis elegans touch receptor neurons transduces mechanical signals. Nat Neurosci 8:43–50. https://doi.org/10.1038/nn1362

Park D, O'Doherty I, Somvanshi RK, et al (2012) Interaction of structure-specific and promiscuous G-protein-coupled receptors mediates small-molecule signaling in Caenorhabditis elegans. Proceedings of the National Academy of Sciences of the United States of America 109:9917–9922. https://doi.org/10.1073/pnas.1202216109

Pokala N, Liu Q, Gordus A, Bargmann CI (2014) Inducible and titratable silencing of Caenorhabditis elegans neurons in vivo with histamine-gated chloride channels. Proceedings of the National Academy of Sciences of the United States of America 111:2770–2775. https://doi.org/10.1073/pnas.1400615111

Ren P, Lim CS, Johnsen R, et al (1996) Control of C. elegans larval development by neuronal expression of a TGF-β homolog. Science 274:1389–1391. https://doi.org/10.1126/science.274.5291.1389

Schackwitz WS, Inoue T, Thomas JH (1996) Chemosensory neurons function in parallel to mediate a pheromone response in C. elegans. Neuron 17:719–728. https://doi.org/10.1016/S0896-6273(00)80203-2

Schaedel ON, Gerisch B, Antebi A, Sternberg PW (2012) Hormonal signal amplification mediates environmental conditions during development and controls an irreversible commitment to adulthood. PLoS Biology 10:1–18. https://doi.org/10.1371/journal.pbio.1001306

Schroeder NE, Flatt KM (2014) In vivo imaging of Dauer-specific neuronal remodeling in C. elegans. https://doi.org/10.3791/51834

Une R, Kageyama N, Ono M, et al (2022) The FMRFamide-like peptide FLP-1 modulates larval development by regulating the production and secretion of the insulin-like peptide DAF-28 in

Caenorhabditis elegans. Bioscience, Biotechnology, and Biochemistry zbac187. https://doi.org/10.1093/bbb/zbac187

Wheeler DE (1986) Developmental and physiological determinants of caste in social Hymenoptera: evolutionary implications. The American Naturalist 128:13–34

Zimmer M, Gray JM, Pokala N, et al (2009) Neurons detect increases and decreases in oxygen levels using distinct guanylate cyclases. Neuron 61:865–879. https://doi.org/10.1016/j.neuron.2009.02.013

Chapter 4

CONCLUDING REMARKS

Mechanisms of the Dauer Circuit

Circuit composition

Prior to the histamine-chloride neuronal silencing experiments performed by our lab, it was possible to envision various ways in which neuronal activity could regulate the dauer decision. On one view, the circuit may have involved only one or two neurons (such as ASI and ASJ) that were each receptive to both food and pheromone inputs, and secreted neuropeptides in an activity-dependent manner directly to downstream tissues, i.e. without any interneuron involvement. On another view, we might expect that other sensory modalities (e.g. gentle touch) could influence dauer, and the integration of all of these inputs may require a larger circuit involving interneurons. We might also expect that a decision-making algorithm involving a more complicated interaction between many neurons was required. Finally, it was possible that downstream tissues were sensitive to the peptidergic output of most of the nervous system, meaning that any neuron could in some way influence the decision.

The initial series of histamine-chloride silencing experiments by Chai et al. (2022b) showed clearly that additional sensory neurons and some interneurons were involved, motivating further experiments of this kind. The results described in this thesis suggest that, while the dauer circuit may extend beyond sensory neurons, not every neuron influence the decision in an obvious way. This is interesting to consider in light of other results from our lab, in which GPCR activities from across the brain were implicated in dauer entry (Chai et al. 2022a). Notably, these GPCRs were distributed such that, within neuron classes, the set of receptors do not have a consistent directional effect on the decision (e.g. the CAN neuron expressed 4 receptors inhibiting dauer and 6 receptors promoting dauer). Therefore, one way to reconcile these findings may be that while peptide activity from across the brain may have a modulatory effect, few cells can be said to have a dedicated 'sign' on the decision. It may also be worthwhile to consider these results in the context of the role of modulatory signaling the *C. elegans* nervous system more generally. In a compact nervous system with relatively few physical connections (and perhaps fewer physically separated circuits), one general strategy for dealing with changing environments over long timescales may be to shift the brain as a whole between several different possible states, each adapted to different environmental challenges. On this view, the activities of individual peptides and receptors may have less to do with individual behaviors than moving the nervous system as a whole between these states. In contrast, the overall output from individual neurons is more likely

to reflect immediate environmental inputs than the internal state of the animal, and may more accurately represent the activity of a specialized circuit. We may also expect that a physiological decision such as dauer may be sensitive to these state changes, though whether this is a design feature or simply a consequence of diffuse and non-specific signaling may be difficult to demonstrate clearly.

Overall, the histamine-chloride inhibition assay has proven to be an informative paradigm for implicating neurons in this behavior, and elucidating circuit composition. Despite the 'negative' results for a number of neurons, it will be worthwhile to continue these experiments for several additional cells of interest, notably CAN, AVF, and PVW, due to the high number of implicated GPCRs they express. In addition, it will be important to extend the screen to sensors of volatile signals such as AWA, AWB and AWC, particularly since they are each connected to AIA. Involvement of these sensory neurons would implicate volatile food signals as inputs in the decision.

Neuronal responses to inputs

The results of calcium imaging experiments presented here and in previous experiments (Chai et al. 2022b) have shown that it is possible to link environmental inputs to the activities of neurons implicated in the dauer decision, and generate insights into the decision-making process. So far, this has involved probing the response profiles of individual neurons in response to the primary chemical inputs into the decision. It will be important to continue this neuron-by-neuron approach in order to demonstrate clearly the involvement of mechanosensation and gas-sensing in the decision, and interpret their roles. For instance, to test whether mechanosensation is involved in detecting crowding, it may be useful to vary the number of *daf-22* larvae on a plate while silencing the TRNs. In addition, it will be important to perform dauer assays with varying oxygen and carbon dioxide conditions, while silencing BAG and URX neurons respectively.

Additional calcium experiments will provide more important insights. The availability of synthetic ascarosides will allow the mapping of each ascarosides to the individual (or sets) of sensory neurons that are receptive to each them. The mechanism by which ASI maintains a sustained depolarization after removal of the stimulus should also be investigated as far as possible. The first steps in this project should be a thorough characterization of its response properties to varying patterns of food input, over long time periods. For instance, how long does the depolarization last after stimulus removal? Is the duration of depolarization proportional to the duration of input? Understanding the response properties of ASI to both food and pheromone will be crucial to understanding how fluctuating and opposing inputs are processed over long periods. Describing the mechanisms may be more challenging, but a first step may be to test whether the *tax-*

2 and *tax-4* encoded ion channels, or the *daf-11* guanylate cyclase, are required for the depolarization of ASI. These proteins may bind to other proteins that mediate its response properties.

Neural circuit mechanisms

Single-neuron imaging of ASI and AIA allowed us to investigate mechanisms by which food and pheromone signals could be compared at the level of individual neurons. However, at this stage it is worth considering a shift in focus towards understanding the flow of sensory information through the implicated neurons and their neighbors. This could be achieved by pan-neuronal imaging (or through imaging of subsets of connected neurons) and would allow several lines of inquiry. First, this could provide more efficient way to obtain a picture of the circuit. Not all neurons that respond to inputs will be involved in the behavior (something that could be easily tested) but unresponsive neurons can be ruled out. Second, it will help uncover functional connectivity that is not implied by anatomy, and that could be highly informative. For instance, in one model, the balance of outputs of BAG, URX, and signals from volatile-sensing neurons AWA, AWB and AWC might be reflected in the relative activities of AIA and AIB, which would help explain both the roles of these interneurons and the mechanism of transduction for these signals.

It would be especially exciting to be able to perform these experiments on long timescales (perhaps with a series of movies taken over several hours), which may be better suited to imaging freely-moving animals, provided that inputs could be delivered appropriately. This could reveal long-lasting changes in circuit properties that could bridge the short timescales on which environmental inputs stimulate the nervous system, and the several-hour long process of physiological commitment to dauer. This might take the form of changes in synaptic weight between connected neurons, or changes in the sensitivity of sensory neurons to input. Indeed, pheromone exposure has been shown to cause changes in GPCR expression in ASI (Peckol et al. 2001). It could also take the form of feedback loops within the circuit, that could sustain neuropeptide synthesis and secretion in the absence of input, similar to the commitment mechanism for adult development involving Dafachronic Acid biosynthesis (Schaedel et al. 2012). A mechanism of this kind could be part of the explanation for the sustained activity observed in ASI. The question of how memory is encoded in circuit properties over several hours is arguably the central question in the study of the dauer decision. A reasonable expectation is that this memory will be distributed among various mechanisms involving long neuropeptide half-lives, connectivity and synaptic weight changes, and other changes in the sensitivity and output of various neurons.

These experiments will be technically challenging to execute. One difficulty in establishing a multi-neuron imaging paradigm will be in finding promoter combinations that will tag the right combinations of cells, with

sufficiently uniform calcium indicator expression to allow meaningful interpretation of their dynamics. Brain-wide imaging could be used, but suffers the drawback that the majority of signals from a moving animal will be irrelevant. However, it is remarkable and encouraging to consider that all of the required technology exists. The study of the dauer decision offers the possibility of a detailed understanding of the mechanisms of a decision-making process at the level of environmental inputs, molecules, genes and circuit properties. This represents just one of many exciting research directions in *C. elegans* neuroscience today.

Semantic Modeling of Neural Circuits

Summary

Overall, the GO-CAM approach was fairly successful in representing a wide variety of neural circuit knowledge in terms of GO-based semantic triples and WormBase Anatomy. With appropriate thought and effort, it was possible to go beyond simple text-to-triple conversion, and specify models that reflected the conceptual framework of the field and highlight missing knowledge. This will be useful for both connectome annotation, and potentially for automating curation for certain kinds of neural circuit knowledge. We also explored how simple computations, such as AND logic and negative feedback, could be represented. These motifs are also prevalent in cellular pathways, and the modeling approaches developed here may be useful for creating GO-CAM models of these types of pathways. Finally, the scope of this project did not extend to capturing quantitative information. Developing this will be important given that many quantities (such as concentrations of chemical inputs) are known, and that measurements (such as time-varying readouts of neural activity) can be made, and mathematical modeling of neural circuits are an active area of *C. elegans* research (Rakowski et al. 2013; Kuramochi and Doi 2017). One approach for this is discussed below.

Representing Quantitative Information in *Ce*N-CAM

It is most important to point out that in semantic graphs, edges can only take on verbal form, and that this is a strength of this approach, allowing queries of their contents in broad strokes. However, we can specify quantities in additional nodes given the right relations. For instance, we could say that some [carbon dioxide receptor activity] that *occurs in* some [Neuron A] *has input* [carbon dioxide], and that this [carbon dioxide] *has concentration* [0.04%] or *has minimum concentration* [0.04%], or *has concentration range* [0.04%-0.08%], provided that these relations could themselves be structured appropriately. It may be desirable to mandate inclusion of these details in any *Ce*N-CAM model, so that the condition-dependence of any experimental result is clearly expressed, and used for queries. We could also imagine representing mathematical models that describe the functional relationships between neurons (Fig. 16a), by specifying the functional connection as a process with a model as one of its properties, described in a connected node. This

need not be restricted to two neurons at a time i.e. groups of neurons that form a circuit could be assigned as *part of* a [nervous system process] or similar term.

The main challenge in implementing this would be for these models to be curated and stored in an appropriate digital repository that is interoperable with GO-CAM. Notably, databases hosting computational models for neuroscience (e.g. ModelDB (Hines et al. 2004)) and systems biology (BioModels (Malik-Sheriff et al. 2020)) exist. In BioModels, mathematical models of biological systems are stored in various machine-readable format, and given unique identifiers, and are individually curated. Their components, including genes, proteins and chemicals, are linked to the GO, UniProt and ChEBI, and other relevant ontologies. Preserving these features will be important to a GO-based *C. elegans* neuroscience database. Thus, while ModelDB is neuroscience focused, the design of BioModels may provide a better example for how to integrate these databases with GO-CAM.

Automating curation

A primary motivation for creating a knowledgebase in which curated results from many papers and data sources can be integrated is the sheer volume of publications. However, this also makes curation itself a daunting task. A method to automate the conversion of text into realistic and satisfying semantic models would be very useful.

Recently, researchers have been exploring whether pre-composed 'knowledge schema' can be used to guide artificial intelligence to automate curation. These knowledge schema are essentially the same concept as the generic 'templates' that were generated in this project, in that they specify the required nodes and relationships to constrain the structure of the data models in desirable ways. In a recent study (Caufield et al. 2023), authors explore the possibility of using a combination of text and knowledge schema as input into a Large Language Model, which is able to populate the schema using appropriate ontology terms through a recursive process. An advantage of this approach is that it requires no training on a prepared dataset. In theory, the ability to interact with the LLM through prompts could allow the design of curation algorithms that could read entire papers in pre-conceived ways, e.g. one paragraph at a time.

Ensuring the accuracy of this approach will be a challenging task, but could be made easier by our demonstration that experimental knowledge in *C. elegans* neuroscience can be grouped into relatively simple, stereotyped categories that are individually meaningful, and whose combination on a large scale will be valuable. However, these developments are unlikely to replace the need for manual curation in the near future. Going forward, it will be important to include lab scientists in curation, for instance by inviting them to

generate *Ce*N-CAM models as part of the normal publication process. The curation templates generated here should make this much more intuitive.

Visualization, Connectome Annotation and Analysis

As discussed previously, an exciting application for *Ce*N-CAM is to use it as a data modeling framework to support connectome functional annotation using a software such as Cytoscape. This would provide a way to display evidence for the experimental information and link its contents to literature (or to large datasets such as *Ce*NGEN (Hammarlund et al. 2018). Cytoscape also has the capability to perform various kinds of network analysis, which will increase in scope and utility as functional annotation proceeds. Software development expertise will be required to develop an interface that can host each of these applications, with a user-friendly interface that populates the dataframes and models from author input. Figure 16b illustrates one schema for how this might work. Any piece of information in the connectome visualization could be linked to a data model that describes the underlying experimental evidence. In theory, this type of resource could allow the synthesis of the entire *C. elegans* neurobiology literature. The causal relationships between inputs, circuits and behavior across the entire *C. elegans* brain, described in terms of cellular and molecular mechanisms, could be displayed together, and made available to computational analysis. In this way, the traditional strengths of *C. elegans* as a model organism for molecular genetics and its emerging capabilities in systems neuroscience may be brought together.

References

Caufield JH, Hegde H, Emonet V, et al (2023) Structured prompt interrogation and recursive extraction of semantics (SPIRES): A method for populating knowledge bases using zero-shot learning

Chai CM, Park H, Sternberg PW (2022a) Brain-wide bidirectional neuropeptide modulation of individual neuron classes regulates a developmental decision. Curr Biol 32:3365-3373.e6. https://doi.org/10.1016/j.cub.2022.05.048

Chai CM, Torkashvand M, Seyedolmohadesin M, et al (2022b) Interneuron control of C. elegans developmental decision-making. Current Biology 32:2316-2324.e4. https://doi.org/10.1016/j.cub.2022.03.077

Hammarlund M, Hobert O, Miller DM 3rd, Sestan N (2018) The CeNGEN Project: The Complete Gene Expression Map of an Entire Nervous System. Neuron 99:430–433. https://doi.org/10.1016/j.neuron.2018.07.042

Hines ML, Morse T, Migliore M, et al (2004) ModelDB: A Database to Support Computational Neuroscience. Journal of Computational Neuroscience 17:7–11. https://doi.org/10.1023/B:JCNS.0000023869.22017.2e

Kuramochi M, Doi M (2017) A Computational Model Based on Multi-Regional Calcium Imaging Represents the Spatio-Temporal Dynamics in a Caenorhabditis elegans Sensory Neuron. PLoS One 12:e0168415. https://doi.org/10.1371/journal.pone.0168415

Malik-Sheriff RS, Glont M, Nguyen TVN, et al (2020) BioModels-15 years of sharing computational models in life science. Nucleic Acids Res 48:D407–D415. https://doi.org/10.1093/nar/gkz1055

Peckol EL, Troemel ER, Bargmann CI (2001) Sensory experience and sensory activity regulate chemosensory receptor gene expression in Caenorhabditis elegans. Proceedings of the National Academy of Sciences of the United States of America 98:11032–11038. https://doi.org/10.1073/pnas.191352498

Rakowski F, Srinivasan J, Sternberg PW, Karbowski J (2013) Synaptic polarity of the interneuron circuit controlling C. elegans locomotion. Front Comput Neurosci 7:128. https://doi.org/10.3389/fncom.2013.00128

Schaedel ON, Gerisch B, Antebi A, Sternberg PW (2012) Hormonal signal amplification mediates environmental conditions during development and controls an irreversible commitment to adulthood. PLoS Biology 10:1–18. https://doi.org/10.1371/journal.pbio.1001306