Identifying regulators of associative learning using a protein-labelling approach in *C. elegans*

Aelon Rahmani[1], Anna McMillen[1], Ericka Allen[1], Radwan Ansaar[1], Anne Poljak[2], Yee Lian Chew[1,*]

1 – Flinders Health and Medical Research Institute, College of Medicine and Public Health, Flinders University, Bedford Park, South Australia 5042, Australia

2 – Bioanalytical Mass Spectrometry Facility, Mark Wainwright Analytical Centre, University of New South Wales, Sydney, New South Wales 2052, Australia

* - to whom correspondence should be addressed: yeelian.chew@flinders.edu.au

Abstract

The ability to learn and form memories is critical for animals to make choices that promote their survival. The biological processes underlying learning and memory are mediated by a variety of genes in the nervous system, acting at specific times during memory encoding, consolidation, and retrieval. Many studies have utilised candidate gene approaches or random mutagenesis screens in model animals to explore the key molecular drivers for learning and memory. We propose a complementary approach to identify this network of learning regulators: the proximity-labelling tool TurboID, which promiscuously biotinylates neighbouring proteins, to snapshot the proteomic profile of neurons during learning. To do this, we expressed the TurboID enzyme in the entire nervous system of *C. elegans* and exposed animals to biotin only during the training step of a gustatory associative learning paradigm. Our approach revealed hundreds of proteins

specific to 'trained' worms, including components of molecular pathways previously implicated in learning and memory formation in multiple species. We validated several novel regulators of learning involved in neurotransmission, including cholinergic receptors (ACC-1, ACC-3, GAR-1, LGC-46) and the putative glutaminase GLNA-3. These previously uncharacterised learning regulators show a clear impact on appetitive gustatory memory, but do not appear to have a generalised role in learning. In summary, we have shown that our approach to use proximity labelling to profile the brain of a small animal during training is a feasible and effective method to advance our knowledge on the biology of learning.

Introduction

All animals with a brain have the capacity to change their behaviour in response to changes in the environment. This capacity – to learn and remember – is essential for survival. There are numerous structural and molecular changes in the brain that modulate learning and memory in specific brain regions, occurring in a time and context-dependent manner (examples in (Huckleberry et al., 2016, Lin et al., 2010, Peixoto et al., 2015, Watteyne et al., 2020); reviewed in (Bailey et al., 2015)). Research using model organisms has been essential towards understanding the key regulatory mechanisms underlying learning, many of which involve neurotransmitter signalling, neuromodulator signalling, signal transduction pathways, and cytoskeletal dynamics (Rahmani and Chew, 2021, Peng et al., 2011, Lamprecht, 2014). Importantly, many of these mechanisms appear to be conserved across diverse species (Bailey et al., 2015, Matsumoto et al., 2018, Rahmani and Chew, 2021).

Multiple studies have demonstrated that changes in the neuronal proteome are required for learning and memory formation (Inberg et al., 2013, Barzilai et al., 1989, Rosenberg et al., 2014). New protein synthesis appears to be critical in several contexts, as the addition of a protein synthesis inhibitor (e.g. cycloheximide) has been shown to abolish long-term memory (Chen et al., 2012, Pedreira et al., 1995, Hernandez and Abel, 2008). Moreover, protein degradation together with new protein synthesis has been strongly implicated in synaptic plasticity and memory formation (Lee et al., 2008, Fazeli et al., 1993, Park and Kaang, 2019). There is also evidence that local translation in neurons, specifically the synthesis of specific proteins in dendritic regions (thereby altering local proteome composition), plays a key role in learning (Bradshaw et al., 2003, Das et al., 2023, Smith et al., 2005, Sutton and Schuman, 2006). Additionally, several key regulatory proteins have been shown to be required at specific timepoints, such as during the training/learning step, to trigger memory formation (Stefanoska et al., 2023, Watteyne et al., 2020). Taken together, these findings suggest that the spatiotemporal regulation of protein composition within neurons is critical for learning and memory formation.

The molecular requirements for learning have primarily been identified by combining genetic approaches with behavioural paradigms to test learnt associations, typically through assaying candidate genetic mutants or performing a forward genetics screen. These strategies have been extremely insightful; however, they have some limitations. The first being that candidate genetic screens are time-consuming and labour intensive, and require subjective selection of which candidate genes to test (for example (Hukema, 2006, Stein and Murphy, 2014)). The second is that large-scale screens tend to only reveal genes that have the strongest phenotypes that can be identified relatively quickly, so genes that have more subtle phenotypes (Hiroki and Iino, 2022, Lindsay et al., 2022), or act in redundant pathways (Feng et al., 2010, Gyurkó et al., 2015,

Shahmoradi, 2015), may not be identified using these approaches despite their contributions to learning (Hiroki and Iino, 2022, Feng et al., 2010).

To overcome these limitations, and to gain a holistic view of the molecular pathways that contribute to learning, we used an objective proteomics approach to snapshot the protein-level changes that occur specifically during learning. To do this, we expressed the proximity-labelling tool TurboID in the entire *C. elegans* nervous system, and used this to identify the proteins present in neurons during the training step of an associative learning paradigm we call 'salt associative learning'. TurboID is an enzyme based on the BirA* biotin ligase, engineered to provide greater catalytic efficiency (Branon et al., 2018) compared with the original BirA* enzyme used in BioID experiments (Roux et al., 2012). TurboID catalyses a reaction where biotin is covalently added onto lysine residues – as this process requires biotin, its timing can be controlled by depleting tissues of biotin, then adding it exogenously only at specific time points. Additionally, spatial control can be provided by regulating the site of TurboID expression using cell-specific transgenes. TurboID has been used in multiple studies for identification of protein-protein interactions, usually by tagging a 'bait' protein N- or C-terminally with the TurboID enzyme, allowing for rapid biotinylation of bait interactors. Through this approach, TurboID has been used for protein-tagging experiments in *C. elegans* (Artan et al., 2022, Sanchez et al., 2021, Holzer et al., 2021, Hiroki et al., 2022). For example, this approach identified cytoskeletal proteins in *C. elegans* proximal to the microtubule-binding protein PTRN-1 (Sanchez et al., 2021), and detected interactors for ELKS-1, which localises other proteins to the presynaptic active zone in the nervous system (Artan et al., 2021).

In our study, rather than focussing on specific protein-protein interactions, we expressed TurboID that was not tagged with any bait protein in the entire nervous system of *C. elegans*, to identify

as many proteins as possible within the cytoplasm. Using this approach, we identified hundreds of proteins specific to 'trained' worms, which we refer to here as the *learning proteome*, including those in molecular pathways previously shown to contribute to learning and memory formation in worms and other organisms. In addition, we validated several novel regulators of learning, including cholinergic receptors (ACC-1, ACC-3, GAR-1, and LGC-46) and the putative glutaminase GLNA-3. These previously uncharacterised learning regulators show a clear impact on salt associative learning, but do not appear to have a generalised role in learning. In summary, we have demonstrated that our approach to use proximity labelling to snapshot the brain of a small animal during training is a feasible and effective method to further our understanding of the biology of learning.

Results

TurboID expression in the nervous system of C. elegans *successfully labels proteins during learning*

To model learning in *C. elegans*, we used a simple yet robust associative learning paradigm called salt associative learning. Briefly, this assay involves training worms to associate the absence of salt (NaCl) with the presence of food. *C. elegans* normally have a mild attraction towards higher salt concentrations (100-200 mM) (Hukema et al., 2006). After training/conditioning, *C. elegans* change their behaviour to become more attracted to lower salt concentrations (group referred to 'trained'). This is because the presence of abundant food (unconditioned stimulus) is a strong innate attractive cue, and pairing this with 'no salt' (the conditioned stimulus) leads to the animals showing the same behaviour towards the conditioned stimulus as they do to the unconditioned stimulus, i.e., attraction towards no salt, reflected as a preference for lower salt concentrations (Hiroki et al., 2022, Nagashima et al., 2019).

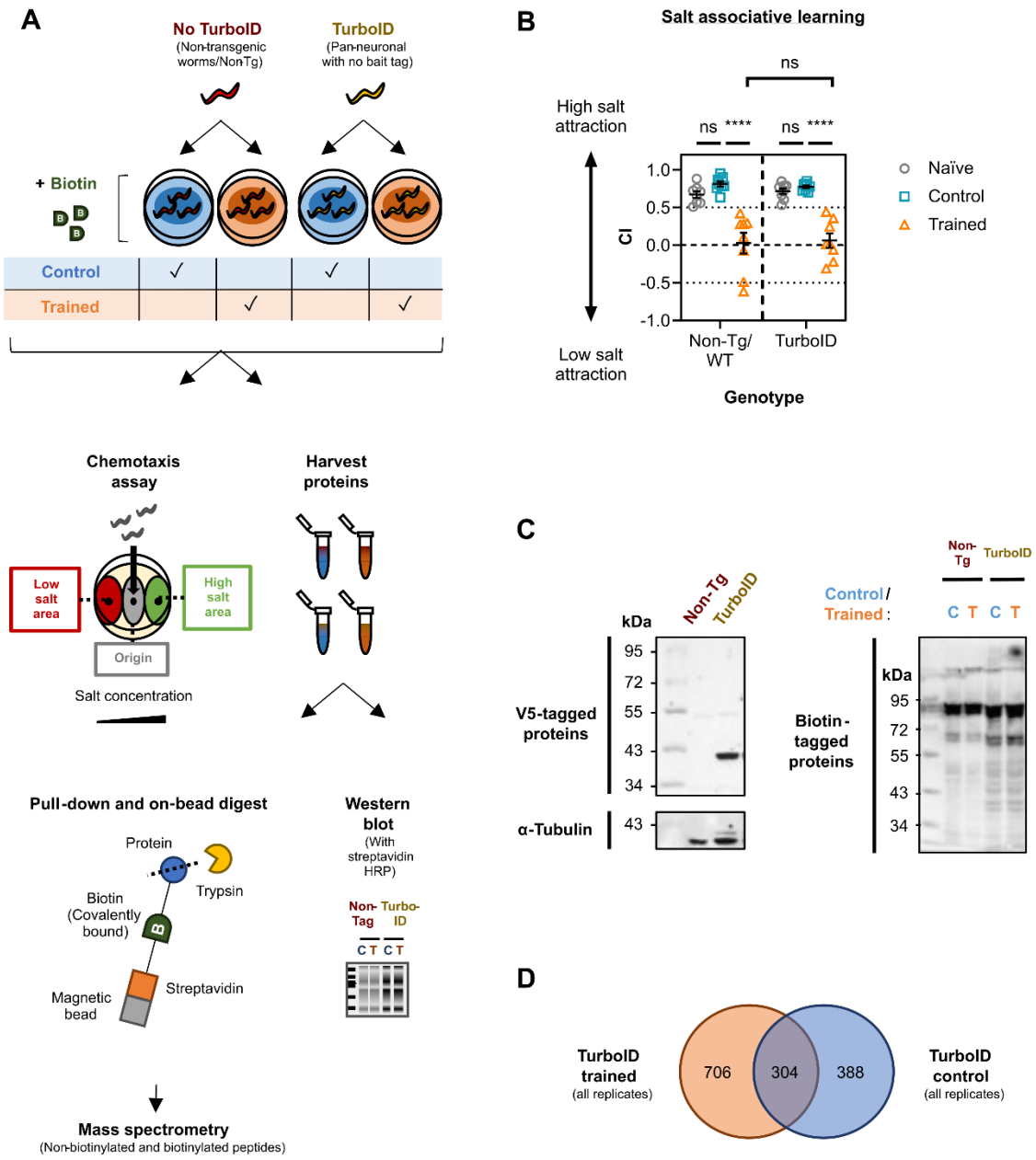


Figure 1: Summary of TurboID approach for protein labelling in all *C. elegans* neurons during learning. (A) Workflow for mass spectrometry-based analysis. Biotin-depleted animals without (non-transgenic/Non-Tg, red) or with TurboID (transgenic, yellow) were exposed to 1 mM of exogenous biotin during conditioning by pairing food with 'no salt' (orange - trained) or 'high salt' (blue - control). >3000 worms were used per group /biological replicate (n = 5) – a small proportion of each group was tested in a chemotaxis assay to assess learning capacity, while the rest was subjected to sample preparation steps for mass spectrometry (see **Experimental Procedures**). Some harvested protein was probed via Western Blot for the presence of biotinylated proteins or V5-tagged TurboID (see panel C for representative images).

(B) The graph shows chemotaxis assay data for Non-Tg/wild type (WT) and transgenic *C. elegans* following salt associative learning. Each data point represents a 'chemotaxis index' (CI) value for one biological replicate (n = 8). Each biological replicate includes three technical replicates (26-260 worms /technical replicate). Statistical analysis: Two-way ANOVA & Tukey's multiple comparisons test (**** ≤ 0.0001; ns = non-significant). Error bars = mean ± SEM. **(C)** Western blots to visualise V5-tagged TurboID and biotinylated proteins. The left side shows V5-tagged TurboID visualised using 18 µg total protein from naïve worms per lane (39 kDa). Non-Tg protein lysates acted as a negative control. α tubulin was probed as a loading control. The right side shows biotinylated proteins visualised from 25 µg total protein per lane from control (C) or trained (T) worms with streptavidin-horseradish peroxidase (HRP). **(D)** Venn diagram comparing all proteins assigned an identity by MASCOT from peptides detected by mass spectrometry from transgenic worms. Values represent the number of proteins listed as detected in 'TurboID, control' (blue) and 'TurboID, trained' (orange). These lists were generated by first subtracting proteins identified in corresponding Non-Tg lists, and then comparing both control and trained TurboID lists. The overlap represents proteins unique to 'TurboID, trained' worms in ≥1 replicate/s that were also detected in 'TurboID, control' worms in ≥1 other replicate/s.

To identify the learning proteome, we adapted this learning paradigm to incorporate TurboID-catalysed biotinylation of proteins specifically during the learning/conditioning step. We did this by (1) performing the salt associative learning assay on transgenic animals expressing TurboID in the entire nervous system (P*rab-3*::TurboID) and (2) adding biotin only when the worms are being trained (i.e., exposed to both food and 'no salt' in the 'trained' group, or to food and 'high salt' concentrations in the 'control' group). As an additional control, we performed the same assay on non-transgenic animals (**Figures 1A-1B**). We then isolated proteins from >3000 whole worms per group for both 'control' and 'trained' groups, most of which were subjected to a sample preparation pipeline for mass spectrometry, and some of which were probed via western blotting to confirm the presence of biotinylated proteins (**Figure 1A**).

The learning capacity of each experimental group was tested using a chemotaxis assay. The chemotaxis index (CI) was used as a readout of learning performance: a positive CI reflects high salt preference, a CI close to 0 represents a more neutral response, and a negative CI represents

low salt preference (**Figure S1**). We confirmed after each learning assay that naïve/untrained worms had a strongly positive CI (~0.7-0.9), whereas trained animals showed a lower CI (~0.3). We also performed a learning control (indicated as 'control') in which the presence of food (the US) is paired with high salt concentrations – worms in this group are attracted to high salt and showed a strongly positive CI (~0.7-0.9) (**Figure 1B**), displaying a similar behaviour to naïve worms. This behavioural change seen in trained animals, versus the naïve and control groups, represented successful learning as seen in (Hiroki et al., 2022, Nagashima et al., 2019). This was observed in both non-Tg and transgenic animals, confirming that introducing the transgene did not perturb learning (**Figure 1B**).

We also confirmed by western blotting that V5-tagged TurboID was expressed in transgenic worms at the expected size (39 kDa) and that biotinylated proteins could be observed in TurboID-expressing control and trained groups (**Figure 1C**). As in other *C. elegans* studies utilising TurboID, we saw background biotinylation in non-Tg controls; however, this is visually lower compared with groups from TurboID transgenic worms (**Figure 1C**) (Artan et al., 2021, Sanchez et al., 2021). Quantification of the signal within entire lanes showed a 1.1-fold increase in the 'TurboID, control' lane compared with the 'non-Tg, control' lane, and a 1.9-fold increase in the 'TurboID, trained' lane compared with the 'non-Tg, trained' lane. For all replicates, we determined that biotinylated proteins could be observed from total TurboID-positive worm lysate by western blotting before proceeding with downstream proteomic experiments.

Our sample preparation methodology for mass spectrometry is based on similar protocols used in *C. elegans* and other systems (Artan et al., 2022, Sanchez et al., 2021, Prikas et al., 2020). We performed five biological replicates, in line with other *C. elegans* studies (Artan et al., 2022, Holzer et al., 2021). To examine the learning proteome, we first subtracted proteins from

'TurboID, trained' groups also present in 'Non-Tg, trained' samples to generate a protein list specific to 'TurboID, trained' animals. We next subtracted from 'TurboID, control' lists any proteins that appeared in 'Non-Tg, control' samples to generate a revised 'TurboID, control' protein list for each replicate. We then compared these revised protein lists for 'trained' and 'control' worms from all biological replicates to examine both unique and shared proteins from these groups. We found 304 proteins that were shared between 'trained' and 'control' TurboID groups, 706 proteins unique to the 'TurboID, trained' group, and 388 proteins unique to the 'TurboID, control' group (**Figure 1D**). We refer to the *learning proteome* as proteins unique to samples for 'TurboID, trained' worms. When generating the learning proteome, we categorised proteins as 'assigned hits' based on the criteria that at least one unique peptide was identified by the MASCOT search engine for the protein identity from at least one biological replicate. We also examined peptide sequences in our peak lists that were considered 'unassigned' by MASCOT, as these sequences were not detected as unique for any protein by the software, but specific protein identities could be found by performing a Basic Local Alignment Search Tool (BLAST) query (https://blast.ncbi.nlm.nih.gov/Blast.cgi; see **Experimental Procedures** for details). The Venn diagram in **Figure 1D** shows assigned hits only. Learning proteome lists for both assigned and unassigned hits are in **Tables S2-S3**.

Examination of the learning proteome reveals known regulators of learning and memory

Our initial analysis of the learning proteome sought to validate our TurboID-based approach by searching for components of pathways previously implicated in learning. We first performed a gene ontology (GO) term analysis of 'cellular component' to obtain a broad overview of the subcellular localisation of proteins identified in trained animals (**Figure 2A**). To do this, we

generated protein-protein interaction (PPI) networks of confident protein hits within the learning proteome for subcellular components of interest (**Figure 2B-G**), using data from STRING and curated with the Cytoscape ClueGO tool (**Tables S4-S5**). We found that a majority of proteins were categorised as 'cytoplasmic' (28.1%) as expected from our approach, which utilised the TurboID enzyme not tagged to any bait protein; this means that we would anticipate the enzyme to be present relatively evenly across the cell body and to catalyse biotinylation of proteins in this space. We saw that an unexpectedly high proportion of proteins were nuclear (18.1%), despite the presence of a nuclear export signal in our TurboID transgene, which should prevent TurboID from entering the nucleus and biotinylating nuclear proteins – this could be due to some 'leaky' entry of the enzyme or biotin reactive species into the nucleus, or that some proteins are categorised solely as 'nuclear' in the ClueGO database when they are also present in other cellular components. Importantly, we found that a proportion of proteins are categorised as present in neuronal compartments – the pre-synapse (0.5%), cilia/dendrites (2.7%) and in the axon/(synaptic) vesicles (4.0%) – as expected from transgenic expression of TurboID in the nervous system.

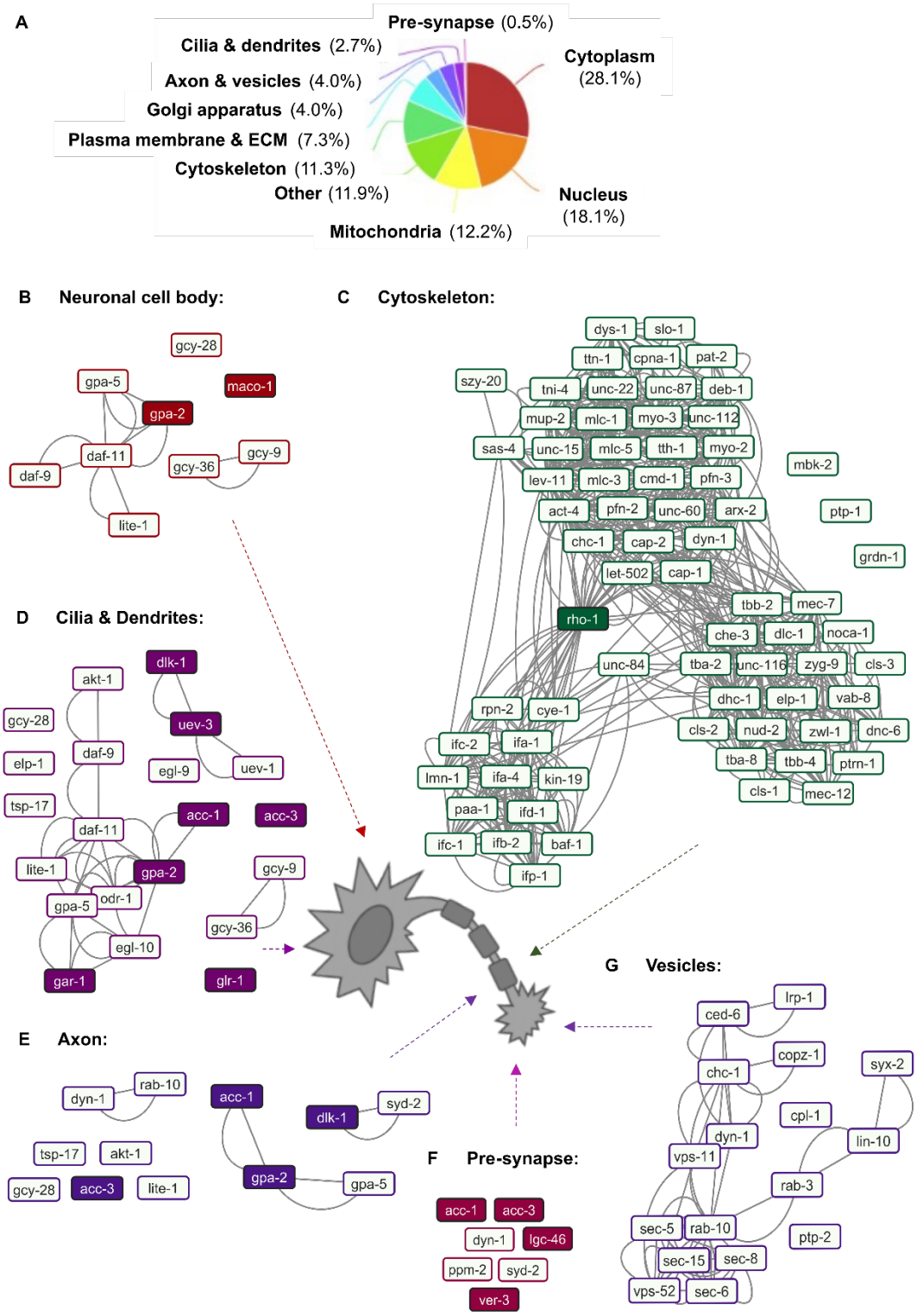


Figure 2: Analysis of subcellular localisations of confident protein hits in the learning proteome. Analyses based on gene ontology (GO) terms were performed with ClueGO (version 2.5.10) using Cytoscape (version 3.10.0). **(A)** Pie chart summarising the percentage of all

proteins in 'TurboID, trained' protein lists assigned GO terms within different 'cellular components'. Abbreviations: ECM = extracellular matrix. **(B-G)** Protein-protein interaction (PPI) networks of proteins in the learning proteome assigned to cellular components most relevant to the nervous system: **(B)** Neuronal cell body/cytoplasm, **(C)** Cytoskeleton, **(D)** Cilia and dendrites, **(E)** Axon, **(F)** Pre-synapse, and **(G)** Vesicles. PPI interactions were exported from STRING (version 12.0) and visualised with Cytoscape. Arrows from networks in **Figures 2B-2G** annotate the relevant region on a neuron schematic. Highlighted proteins were further tested in this study. Known information on subcellular localisation and predictions based on homology were also used for several nodes (Gaudet et al., 2011, Takayanagi-Kiya et al., 2016, Lans et al., 2004, Arellano-Carbajal et al., 2011, Popovici et al., 2002). The consolidated data from GO term analyses can accessed on **Table S4**.

Learning and memory formation in organisms with brains of varying sizes has been shown to involve key regulatory pathways including signalling via neurotransmitters/neuromodulators, G protein-coupled receptors (GPCR), the mitogen-activated protein kinase (MAPK) pathway and the insulin/insulin growth factor-like pathway (Matsumoto et al., 2018, Myhrer, 2003, Rahmani and Chew, 2021). We next categorised proteins within the learning proteome (**Figure 2**) based on their known roles within these signalling pathways, summarised in **Figure 3**, **Figure S3**, and **Table S6**. This included (1) several GPCR components, including the $G_{i/o}$ protein subunit GOA-1 and G_{α} protein subunit GPA-2, (2) regulators of insulin signalling including the DAF-2 insulin receptor, phosphoinositide 3-kinase AGE-1, and serine/threonine protein kinases AKT-1 and SGK-1, which were previously reported to modulate salt-based learning in the worm (Tomioka et al., 2006, Sakai et al., 2017), (3) MAPK signalling components including NSY-1/MAPKKK, MEK-2/MAPKK, and MPK-1/MAPK/ERK, (4) cAMP/PKA (protein kinase A) signalling regulators such as the regulatory PKA subunit KIN-2, and (5) multiple components that modulate synpatic vesicle release, including N-ethyl-maleimide sensitive fusion protein NSF1/NSF-1 and syntaxin/SYX-2. In addition, we identified several proteins relevant to glutamate, acetylcholine, and GABAergic signalling. Several components involved in protein synthesis and degradation

were also detected in our learning proteome, in line with studies that suggest changes in total protein composition following memory formation (Inberg et al., 2013, Barzilai et al., 1989). These data are summarised in **Figure 3**, **Figure S3**, and **Table S6**. We focussed on proteins functioning within these pathways of interest in our subsequent investigations (highlighted nodes in **Figure 2** and **Figure S3**).

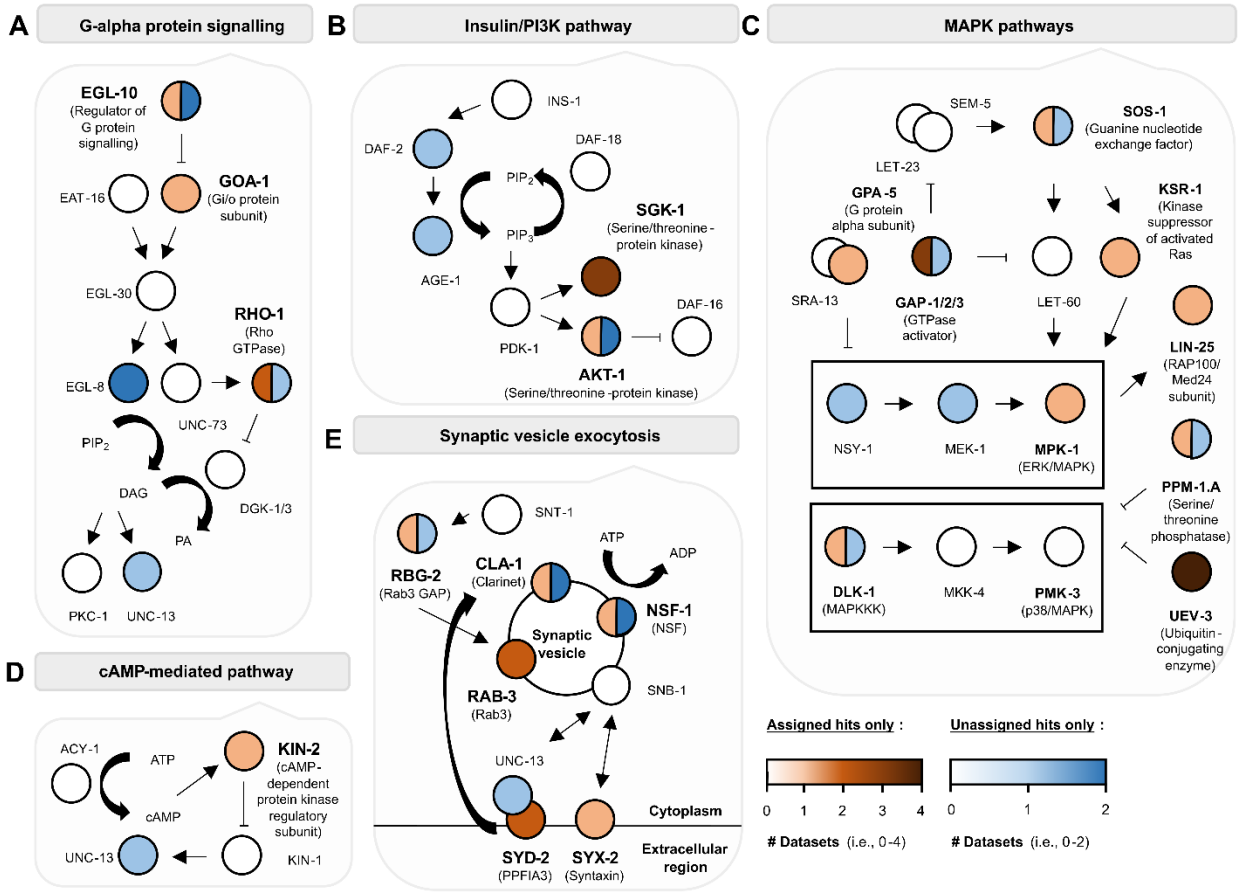


Figure 3: Molecular pathways previously implicated in associative learning are detected in our learning proteome. Proteins are represented as nodes in this schematic. Proteins detected from 'TurboID, trained' worm lysates by mass spectrometry are in bold with circles coloured as orange ('assigned hits' assigned protein identities by MASCOT) and/or blue ('unassigned hits' given protein identities by bulk BLAST searching, but not MASCOT). Darker colours mean the protein has been detected in more than one biological replicate (see legend).

Validating the requirement of learning proteome components in salt associative learning through single gene studies

Our initial analysis of learning proteome data indicates that there are multiple hits present in biological pathways important for neuron function, and that are potentially relevant to learning and memory formation. To test this directly, we performed salt associative learning experiments on selected learning proteome hits (**Figure 4**, **Figure S4).** We used the following general rules to interpret our data: if the average chemotaxis indices (CIs) for 'trained' worms were *higher* in a particular strain compared with wild-type, this strain was considered learning-defective, as this reflects a reduced magnitude of the expected behaviour change (a preference for low salt demonstrated by CIs closer to 0 or negative CIs). If the average CI for 'trained' worms are *lower* in a strain compared with wild-type, then this strain was considered to display 'better' learning, as the lower CI reflects an increased magnitude of the expected behaviour change. In general, we observed no significant difference in CIs between naïve groups for all genotypes, reflecting no gross locomotor or chemotaxis defects in the strains tested (**Figure 4**, **Figure S4**).

We first tested genetic mutants for hits involved in neurotransmission, a key function for all neurons and a requirement for learning (reviewed in Myhrer (2003) and Rahmani and Chew (2021)). We found several acetylcholine receptors that show differences in learning compared with wild-type controls (i.e., Non-Tg worms in **Figure 1**). Those tested include *acc-1*, *acc-3*, and *lgc-46* in the subfamily of anionic "ACC" acetylcholine-gated ligand-gated ion channels (Morud et al., 2021), whereas *gar-1* is a G protein-linked acetylcholine receptor (Park et al., 2000). Interestingly, *gar-1* and *lgc-46* mutants demonstrate a reduced capacity for learning compared with wild-type, whereas *acc-1* and *acc-3* mutants appear to display better learning (**Figure 4A**). Although all receptors are activated by acetylcholine (Park et al., 2000, Putrenko et al., 2005), or

have been reported to possess the same protein domain/s as known acetylcholine receptors (Takayanagi-Kiya et al., 2016), they vary substantially in expression pattern in the nervous system (Taylor et al., 2021). We postulate that acetylcholine signalling in specific neurons may therefore contribute to learning in different directions, an interesting avenue for future research. In addition, the putative glutaminase *glna-3* showed a potential role in salt associative learning: when tested, a deletion mutant for *glna-3* showed reduced learning compared to wild-type animals. We also tested other genes involved in neurotransmission, such as *gbb-2* (for GABAergic signalling) and *glr-1* (for glutamatergic signalling), but although mutants for these genes demonstrate a trend towards improved learning compared with wild-type, this did not reach statistical significance for either gene (p value between trained groups = 0.4468 (*gbb-2*) and 0.7440 (*glr-1*)) (**Figure S4G-H**). Our data therefore indicate that specific components of neurotransmitter signalling are required for salt associative learning.

We tested several components of GPCR signalling that were identified as learning proteome hits, including *gap-1*, *gap-2*, *gpa-2* and *rho-1* (**Figure S4A-D**), but found that worms with single mutations for these genes, when compared to wild-type controls, did not show statistically significant differences in learning capacity. We postulate that although GPCR signalling is broadly important for learning (reviewed in Jong et al. (2018) and Rahmani and Chew (2021)), there may be high levels of redundancy built within this pathway such that single pathway components can be compensated for by other functionally similar genes.

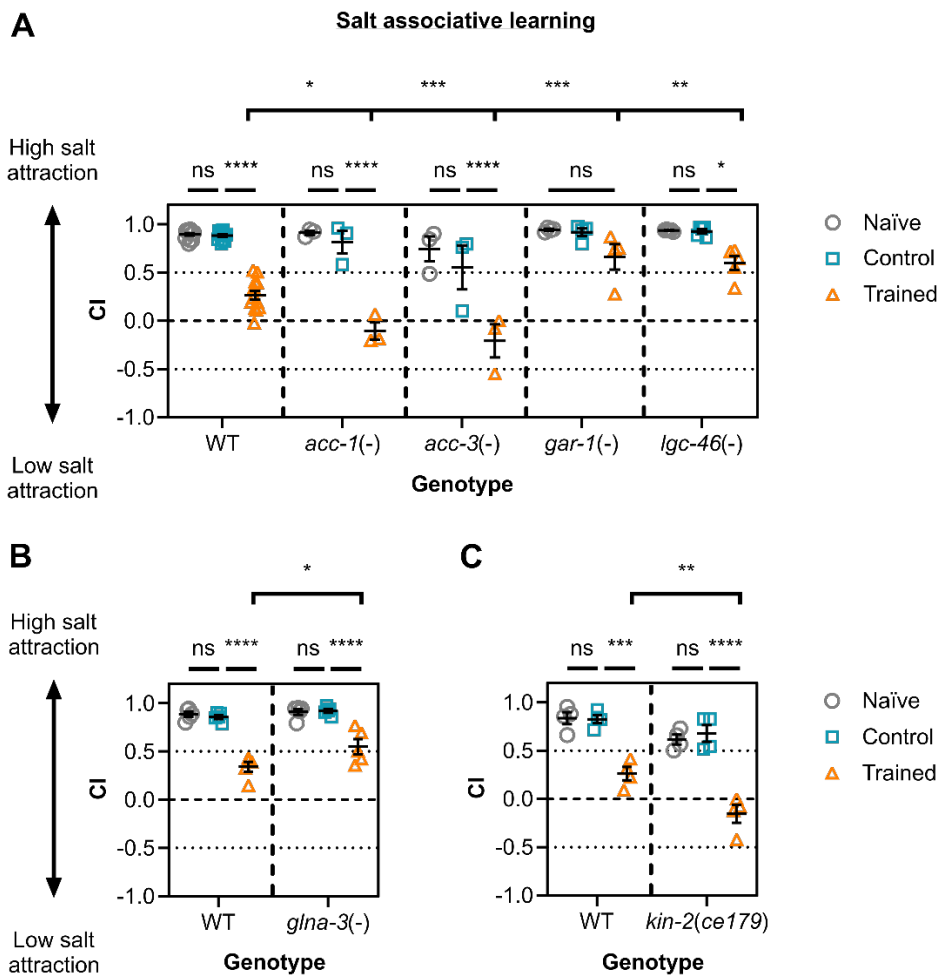


Figure 4: Validation of learning proteome hits identified through mass spectrometry experiments. Single gene mutants (mutation type in brackets) for **(A)** acetylcholine receptors *acc-1* (insertion), *acc-3* (deletion), *gar-1* (deletion), *lgc-46* (deletion), **(B)** glutaminase *glna-3* (deletion), and **(C)** PKA regulatory subunit *kin-2* (substitution) were tested using a salt associative learning assay. Detailed strain information is provided in **Table S1**. Chemotaxis indices (CI) were calculated for each biological replicate (three technical replicates per biological replicate, with 23-487 worms per technical replicate). n = 3-13 biological replicates. Statistical analysis: Two-way ANOVA & Tukey's multiple comparisons test (**** ≤ 0.0001; *** ≤ 0.001; ** ≤ 0.01; * ≤ 0.05; ns = non-significant). Error bars = mean ± SEM.

Next, we found that the regulatory subunit of PKA, *kin-2*, is involved in salt associative learning. Adenylyl cyclase is a key signalling effector for $G\alpha_s$ and $G\alpha_i$ proteins and regulates levels of the secondary messenger cyclic AMP (cAMP) within the cell. cAMP binding to PKA regulates its

activity, and therefore its downstream effects (Sassone-Corsi, 2012). We tested worms with a mutation in *kin-2* in which a conserved residue in the inhibitory domain (which normally functions to keep PKA turned off in the absence of cAMP) is mutated, resulting in increased PKA activity (Schade et al., 2005). *kin-2* has previously been shown to be required for intermediate-term memory in *C. elegans* (Stein and Murphy, 2014), with cAMP/PKA signalling previously shown to be involved in memory in multiple systems (Kandel, 2012). We found that these *kin-2* mutant animals showed an increased capacity for learning, compared with wild-type (**Figure 4C**), indicating that increased PKA activity drives salt associative learning.

Several components of MAPK signalling have been shown to be involved in different forms of learning (reviewed in (Peng et al., 2010, Ryu and Lee, 2016)), including NSY-1/MAPKKK, MEK-2/MAPKK, and MPK-1/ERK that were identified as part of the learning proteome in our study (Ohno et al., 2014). We tested the dual leucine zipper MAPKKK-encoding gene *dlk-1,* a mutant of this gene showed no difference in learning capacity compared to wild-type (**Figure S4E**). We also tested the E2 ubiquitin-conjugating enzyme variant UEV-3, shown to be a member of the DLK-1 pathway and a potential interactor of p38/MAPK PMK-3 (Trujillo et al., 2010). Although the *uev-3* mutant strain showed a trend towards improved learning compared with wild-type, this did not reach statistical significance (p value between trained groups = 0.1806) (**Figure S4F**).

Finally, we assayed other proteins identified within our learning proteome that function within signalling pathways of interest or that were previously implicated in learning or memory paradigms in other *C. elegans* studies. *elo-6* encodes a long chain fatty acid elongase that potentially functions together with another elongase encoded by *elo-5*, although only *elo-6* is expressed in neurons (Kniazeva et al., 2004). As fatty acid composition has previously been

demonstrated to be important for learning and memory (Wallis et al., 2021, Pershina et al., 2022, Akefe et al., 2024), we tested a mutant of *elo-6* for its role in learning. Our data showed no significant learning defect or improvement in *elo-6* mutant animals (**Figure S4I**), although as mentioned, its role may be masked by functional redundancy. In addition, *maco-1* encodes a macoilin family protein that functions broadly in neurotransmission (Arellano-Carbajal et al., 2011), and it has previously been shown to regulate memory of olfactory adaptation in worms (Kitazono et al., 2017). The *maco-1* mutant tested in our study (*nj21*) did not show a locomotor defect, but also did not show any obvious learning phenotypes (**Figure S4J**). Lastly, we also tested *ver-3*, which encodes a predicted vascular endothelial growth factor (VEGF) receptor-like protein (Popovici et al., 2002). *ver-3* is expressed in the ALA neurons (Popovici et al., 2002), which function in quiescence and mechanosensation (Van Buskirk and Sternberg, 2007, Sanders et al., 2013). Moreover, VEGF/VEGFR was reported to be upregulated in rats following spatial learning (Cao et al., 2004); however, we did not observe any learning differences in *ver-3* mutants in the salt associative learning assay (**Figure S4K**).

Learning proteome hits identified in the salt associative learning paradigm do not show generalised roles in learning

Next, we explored whether the learning regulators identified from our learning proteome functioned more broadly in other types of learning. To do this, we tested worms with mutations in the same genes of interest using other learning assays, namely salt aversive and butanone appetitive learning assays.

Salt aversive learning

We first performed a salt aversive learning assay on mutant strains of interest from **Figure 4**, to test if the mutant strains generally show salt-based learning changes: in the salt associative learning assay used thus far, 'trained' worms are exposed to a pairing of no salt + food ('control' worms with high salt + food), whereas in the salt aversive learning assay, 'conditioned' worms are exposed to high salt + no food (and control 'mock-conditioned' worms with no salt + no food). In the salt aversive learning assay, 'conditioned' worms therefore learn to avoid high salt (as it associated with starvation, a strongly negative cue), whereas 'mock-conditioned' worms and naïve worms retain a preference for high salt (Nagashima et al., 2019, Hiroki et al., 2022). We found that only *gar-1* mutant worms showed a significant change in learning capacity compared with wild-type (**Figure 5A**), whereas other mutants in **Figure 4** did not show significant changes in learning capacity versus wild-type (**Figure 5B-F**). Specifically, *gar-1* mutants displayed a larger decrease in CI in trained animals compared to wild-type trained worms, demonstrating a potential learning improvement (**Figure 5A**). Interestingly, this is in the opposite direction to the *gar-1* mutant phenotype in salt associative learning, in which *gar-1* mutants were shown to display a reduced learning capacity (**Figure 4A**).

Butanone appetitive learning

We next asked if specific validated learning proteome hits were required for a different learning paradigm involving a distinct sensory modality (i.e., not gustatory/taste learning). We were most interested in testing *gar-1* mutants, as this strain was the only one that showed a defect in both salt associative learning and salt aversive learning. We selected the paradigm for butanone appetitive learning as this is a well-established olfactory learning assay used in multiple *C. elegans* studies (Kauffman et al., 2010, Stein and Murphy, 2014). Briefly, naïve (untrained) worms show a modest attraction towards butanone, but pairing butanone with abundant food

(‘conditioned’ group) increases this attraction. ‘Mock-conditioned’ worms are exposed to a pairing of ethanol (the diluent for butanone) and abundant food, and do not show this increased attraction towards butanone. Both ‘conditioned’ and ‘mock-conditioned’ worms are starved for 1 hour prior to conditioning, to increase the ‘value’ of the unconditioned stimulus (food) and their motivation to seek out environments with abundant food (Kauffman et al., 2011). Finally, ‘pre-starved’ controls, which are starved for 1 hour but are not exposed to either pairing (i.e., do not undergo training), behave similarly to naïve worms. We found that *gar-1* mutant worms do not show a significant difference in learning capacity when compared to wild-type in the butanone appetitive learning assay (**Figure 5G**). This suggests that GAR-1 is specifically involved in salt-related learning.

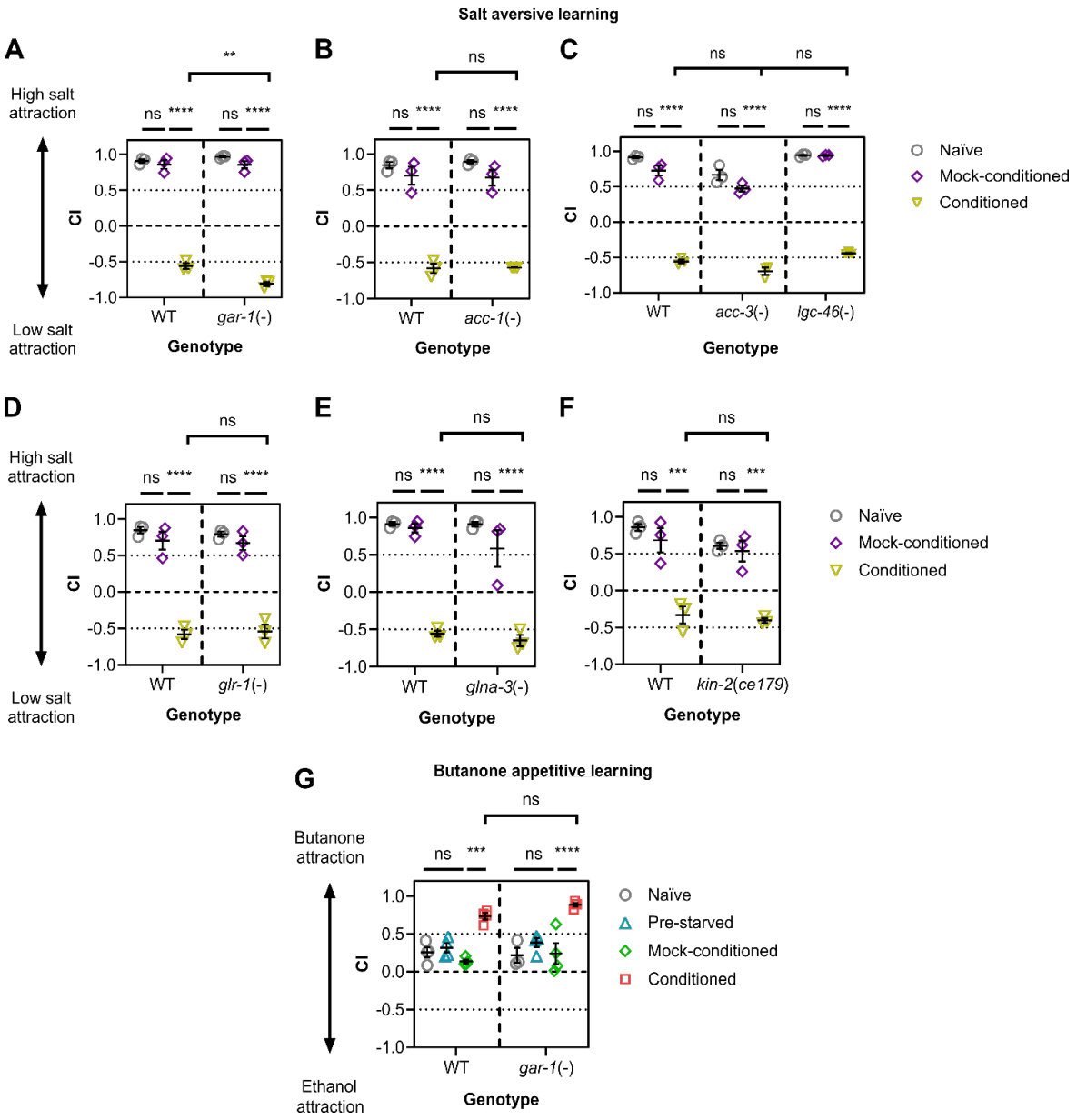


Figure 5: Learning proteome hits that impact salt associative learning generally do not affect alternate learning paradigms. (A-F) Salt aversive learning assays: Worms were naïve, mock-conditioned with no salt + no food, or conditioned with salt + no food. Three technical replicates were generated per biological replicate (26-481 worms per technical replicate). n = 3 biological replicates. **(G)** Butanone appetitive learning assays: Worms were naïve, pre-starved, mock-conditioned with ethanol + food after pre-starvation, or conditioned with butanone + food after pre-starvation. Four technical replicates were generated per biological replicate (57-243 worms per technical replicate). n = 3-4 biological replicates. Each data point represents a chemotaxis index (CI) for an individual biological replicate. Statistical analysis: Two-way ANOVA & Tukey's multiple comparisons test (**** ≤ 0.0001; *** ≤ 0.001; ** ≤ 0.01; ns = non-significant). Error bars = mean ± SEM.

Discussion

Our study demonstrates the feasibility of using the protein labelling-technique TurboID to explore *C. elegans* learning and memory. The expression of TurboID in the whole nervous system of the worm, and addition of biotin only during the training step, allowed us to label proteins present in neurons during this critical stage of memory encoding. We identified these proteins by mass spectrometry and revealed a putative 'learning proteome' from which we were able to identify multiple novel regulators of learning. We found that three acetylcholine receptors (ACC-1, ACC-3, & LGC-46) and the glutaminase GLNA-3 predicted to be involved in glutamate biosynthesis, are required for salt associative learning (**Figure 4**). We also found that a mutant of *kin-2* (regulatory subunit of PKA) that results in increased PKA activity showed improved learning, consistent with other studies demonstrating the requirement of cAMP/PKA in learning (reviewed in Kandel (2012)). Interestingly, loss-of-function mutants of the metabotropic acetylcholine receptor GAR-1 also show improved learning, suggesting that GAR-1 normally acts to inhibit salt associative learning in the worm (**Figure 4**). GAR-1 also impacts aversive gustatory learning, with *gar-1* mutant animals showing reduced learning capacity in a salt aversive learning paradigm. Notably, the other newly validated genes did not show an effect in salt aversive learning (**Figure 5**), and the *gar-1* mutation did not impact butanone appetitive learning, a form of olfactory learning (**Figure 5**). These findings indicate that the genes tested here do not play a generalised role in associative memory, and that GAR-1 may play a complex role in worms in the context of gustatory associative learning.

Learning and memory are key functions of the nervous system and are critical for survival. Forms of associative learning have been studied in invertebrate and vertebrate animals for decades, revealing many important insights on the behavioural, neuroanatomical, and molecular

requirements for learning and memory formation (reviewed in Hawkins and Byrne (2015), Jong et al. (2018), Kandel (2012), Matsumoto et al. (2018), Peng et al. (2010), Peng et al. (2011), Rahmani and Chew (2021), Ryu and Lee (2016)). However, many studies focus on single genes of interest and are therefore unable to reveal the entire network of molecular players that drive (or inhibit) learning. We used a complementary approach to take a snapshot of the proteins present in the brain during learning, using the protein-labelling tool TurboID. Our strategy identified several components of molecular pathways that are generally required for learning, including neurotransmitter signalling, MAPK signalling, insulin signalling, synaptic vesicle exocytosis, and GPCR signalling (**Figure 3**, **Figure S3**, and **Table S6**). We identified regulators of learning that may not have been obvious choices for a candidate screen, and that may have had phenotypes too subtle to be highlighted through a random mutagenesis screen. We therefore conclude that our approach is a useful and scalable method that can be used in multiple systems to delineate the molecular requirements for different forms of learning.

There are several interesting unanswered questions: Firstly, why are the proteins we identified not general learning regulators? Mutants of ACC-1, ACC-3, GAR-1, GLNA-3, LGC-46, and KIN-2 show significantly different learning capacities in the salt associative learning assay compared with wild-type (**Figure 4**); however, only *gar-1* mutants showed a significant learning difference in another salt-based paradigm, and showed no difference in an olfactory learning assay (**Figure 5**). In addition, *kin-2(ce179)* mutants show increased learning capacity in the salt associative learning assay but show intact salt aversive learning **(Figure 5F)** and butanone appetitive learning (Stein and Murphy, 2014). One possibility is that our method for selecting hits from our mass spectrometry data for downstream validation introduced an unintended bias: to identify the learning proteome, we subtracted proteins in the list for 'control' worms from the

protein list for 'trained' worms, for each biological replicate. The main difference between control and trained worms was whether they were exposed to a pairing of high salt + food (control) or low salt + food (trained). It is possible that proteins present in trained worms, but not control worms, during conditioning are those that are strongly regulated by changes in salt concentration (and therefore impact mainly salt-based learning). Indeed, our 'simple subtraction' approach may be an overly conservative method for selecting learning regulators, as it is likely that many neuronal proteins generally important for learning are present in both groups, but at higher (or lower) levels in trained worms. Another approach may be to perform a protocol for quantitative mass spectrometry, thereby allowing for differences in protein levels between trained and control groups to be measured. This may reveal additional learning regulators that have more generalised impacts on learning and memory that are not limited to a specific cue or sensory modality.

Secondly, why do the different acetylcholine receptors that we identified in our study impact learning differently? We showed that loss-of-function mutants of ACC-1 and ACC-3 show improved learning, whereas mutants of GAR-1 and LGC-46 show learning defects compared with wild-type (**Figure 4**). One study suggested that ACC-1 and ACC-3 may function together: ACC-3 homomers do not respond robustly to acetylcholine, but ACC-1 and ACC-3 can form a functional heteromer, albeit with lower sensitivity to acetylcholine that ACC-1 homomers (Putrenko et al., 2005). ACC-1 and ACC-3 functioning as heteromers may be why these proteins impact learning in the same direction. On the other hand, LGC-46 is also a member of the acetylcholine-gated chloride channel (ACC) family but impacts salt associative learning in the opposite direction to ACC-1 and ACC-3. LGC-46 (in 81 neurons) has a much broader expression pattern than ACC-1 (in 32 neurons) or ACC-3 (in 14 neurons), and is expressed in many more

interneurons and motor neurons (Taylor et al., 2021). It is therefore possible that some of the LGC-46-expressing neurons function to regulate learning in a different manner to ACC-1- or ACC-3-expressing cells. Lastly, GAR-1 showed a very interesting phenotype in salt-based learning: *gar-1* deletion mutants show a reduced capacity of salt associative learning but have the opposite phenotype in salt aversive learning. The main difference between these assays is that salt associative learning uses a positive association (salt + food/satiety) whereas salt aversive learning uses a negative association (salt + no food/starvation). One possibility is that GAR-1 can modulate salt-based learning differently based on the presence or absence of food. Single cell expression data indicates that GAR-1 is expressed in the salt-sensing neurons ASEL and ASER (Taylor et al., 2021), suggesting that it is likely to contribute to salt sensation and downstream behaviours. How this is linked to learning and the impact of satiety factors such as the presence/absence of food, remains to be determined.

Our approach demonstrates a potentially powerful method to uncover regulatory networks for a variety of behaviours; however, one factor we aim to improve in future is the amount of background 'noise' observed in the learning proteome. While the five biological replicates of learning proteome data did reveal genes and molecular pathways implicated in learning and memory, there was also potential background. For example, when assessing the overlap between all proteins (i.e., both assigned and unassigned proteins (**Figure S2B**)), a large proportion of genes identified were categorised as ribosomal (4-12%), mitochondrial (8-12%), or involved in reproduction (16- 19%) (STRING database, accessed Feb 2024). Mass spectrometers detect peptides by abundance (Bakalarski et al., 2008), so many of these highly abundant proteins may have preferentially been detected over less abundant neuronal proteins. There are several potential ways to reduce background and improve the signal-to-noise ratio: (1) using an

integrated transgenic line that expresses TurboID with 100% transmission, as the line used in our study had a 70-80% transmission rate, (2) using a cell-sorting strategy (e.g., by flow cytometry) to isolate a tissue of interest (Beets et al., 2020), this is particularly useful if the tissue of interest is only a low proportion of the worm biomass (e.g., neurons are 1% (Froehlich et al. (2021)), as proteins from larger tissues, such the gut and germline, may interfere with detection, and (3) removing highly abundant background proteins during mass spectrometry sample preparation, such as mitochondrial carboxylases that are endogenously biotinylated, as in Artan et al. (2022).

Here, we have described an approach to identify the network of molecular processes that drive learning and memory formation, using the compact *C. elegans* nervous system. Our data reveal proteins from known biological pathways relevant to associative memory, and through which we have identified novel regulators of gustatory associative learning. Future studies using this approach to identify learning regulators in other contexts will advance our understanding of the complex spatiotemporal regulation of learning and memory. This may help to elucidate the principles through which different memory types arise from the combination of specific neuronal signals, individual brain regions/cells, and different sensory modalities, relevant to brains of many sizes.

Experimental Procedures

C. elegans strain maintenance

Young adult (day 1) hermaphrodite *C. elegans* were grown using standard conditions on nematode growth medium (NGM) agar in petri dishes at 22 °C for all experiments (Brenner, 1974). This was done for at least two generations for salt associative learning assays involving TurboID, as well as assays involving salt aversive learning involving *kin-2(ce179)* mutants, and all butanone appetitive learning assays (assays defined below). Animals were otherwise cultured at 22 °C for one generation, and at 15 °C prior to this. For TurboID-based labelling experiments, worms were cultured for at least two generations with the biotin-auxotrophic strain as their food source (Branon et al., 2018). For all other experiments, animals were fed *Escherichia coli* (*E. coli*) strain OP50. *C. elegans* lines used in this paper are listed in **Table S1**.

Biotin treatment

The biotin treatment strategy was adapted from Artan et al. (2021). Briefly, a solution of 100 mM biotin, 250 mM KOH, 5 mM K_3PO_4 (pH 6.0), 1 mM $CaCl_2$, and 1 mM $MgSO_4$ was diluted 1:100 in *E. coli* MG1655 *bioB::kan* washed with modified Luria Broth (LB; 25 mM NaCl, 5 mM K_3PO_4 (pH 6.0), 1 mM $CaCl_2$, 1 mM $MgSO_4$, 1.0% (w/v) Bacto™ Tryptone, 0.5% (w/v) yeast extract, 0.05 mg/mL Kanamycin). Biotin-depleted worms were fed *E. coli* MG1655 on NGM agar during a 6 h conditioning period (see the '**Salt associative learning**' section). The bacterial pH was increased by a negligible amount (i.e., 0.1) from the addition of KOH, thus it is expected it will not significantly impact worm physiology (Khanna et al., 1997, Cong et al., 2020). At least 3,000 worms per group were utilised for downstream proteomic experiments.

Protein extraction and quantification

The following protocol was adapted from Liang et al. (2014) and Artan et al. (2021): After biotin treatment (end of learning assay), worms were washed twice using 'washing buffer' (50 mM NaCl, 5 mM K_3PO_4 (pH 6.0), 1 mM $CaCl_2$, 1 mM $MgSO_4$) and then stored as pellets in a -80 °C freezer. Each pellet was suspended in 200 µL Radioimmunoprecipitation assay (RIPA) buffer containing 2 M urea (prepared as in Sanchez and Feldman (2021)), as well as 150 mM NaCl, 50 mM Tris-Cl (pH 8.0), 5 mM EDTA, 10 mM NaF, 2 mM Na_3VO_4, 1 mM NaPP, 1% (v/v) Nonidet-P40, 1% (w/v) SDS, 0.5% (w/v) sodium deoxycholate, 0.1% (w/v) β-glycerophosphate, and 1× cOmplete Mini Protease Inhibitor (Merck). Worm pellets were sonicated (10× 4 sec total sonication time, 2 sec 'on' pulse, 3 sec 'off' pulse', and 25% amplitude) with the Q125 Sonicator (Q Sonica) in a temperature-controlled room (~2-8 ºC) and allowed to rest on ice for >20 sec between each sonication. All samples were vortexed for ~5 sec at room temperature, and then centrifuged (14,000 rcf, 4 °C, 10 min) to separate carcasses/debris from supernatant containing proteins. The supernatant was then used for protein quantification by BCA assay (ThermoFisher Scientific, #23225).

SDS-PAGE, protein transfer, and western blotting

20-40 µL protein samples, containing 1× sample buffer (0.25 M Tris (pH 6.8), 10% (v/v) β-mercaptoethanol, 10% (v/v) glycerol, 10% (w/v) SDS, 0.25% (w/v) bromophenol blue), were boiled at 95 °C for 3 min. Proteins were electrophoresed through 8% (v/v) polyacrylamide gels for 40-60 min at 80 V, before the voltage was increased to 100 V for an additional 40-60 min, and then increased to 120 V until completion of electrophoresis (Jeong et al., 2018, Stefanoska et al., 2022). The standard protocol was utilised for semi-dry protein transfer (constant 25 V, 40-45 min) onto nitrocellulose membranes and western blotting, using the following antibodies/probes (dilution in brackets, all made in 5% (w/v) bovine serum albumin): rabbit anti-α tubulin (1:1000)

(Abcam, #ab4074), mouse anti-V5 (1:1000) (Cell Signalling Technology, #80076), goat anti-rabbit HRP (1:20,000) (Cell Signalling Technology, #7074P2), goat anti-mouse HRP (1:20,000) (ThermoFisher Scientific, #G-21040), and streptavidin-HRP (1:5000) (Cell Signalling Technology, #3999) (Jeong et al., 2018, Stefanoska et al., 2022). Probed proteins were visualised by chemiluminescence using Clarity Western ECL Substrate (Bio-Rad, #1705060) according to manufacturer's instructions.

Mass spectrometry

1. Sample preparation

Our protocol was adapted from Artan et al. (2022), Sanchez et al. (2021), and Prikas et al. (2020): ~1 mg of protein was desalted using 7 kDa molecular weight cut-off desalting spin column (ThermoFisher Scientific, #89883) by buffer exchange with RIPA buffer containing a lower SDS content (i.e., 0.1% (w/v) SDS) and no urea (Artan et al., 2022). Desalted protein samples were quantified by BCA assay so that ~0.6-1.0 mg of total protein per sample could be used in subsequent pull-down experiments to enrich for biotinylated peptides using Streptavidin Magnetic Beads (NEB, #S1420S) equilibrated with TBS-T (150 mM NaCl, 10 mM Tris (pH 7.4), 0.1% (v/v) Tween-20). Total protein was gently agitated in a tube rotator in 1 mL of total volume (4.5 µL bead:8.0 µg total protein ratio) for 18 h at 4 °C. Magnetic beads were sequentially washed with the following solutions (number of washes in brackets): TBST (x3), 1 M KCl (1x), 0.1 M Na_2CO_3 (1x), and PBS (5x; ThermoFisher Scientific, #10010023) (Sanchez et al., 2021).

The beads were then incubated in a ThermoMixer (Eppendorf # 5384000063) at 800 rpm, 55 °C for 1 h in 200 µL per sample of reducing solution containing 4 M urea, 50 mM NH_4HCO_3, 5 mM

dithiothreitol, and 0.1% (w/v) Protease-Max Surfactant (Promega, #V2071) in 50 mM NH_4HCO_3. Alkylation was promoted by adding 4 µL of 0.5 M iodoacetamide to each solution and re-incubating each sample at 800 rpm, 55 °C for 20 min in the dark. Finally, samples were incubated at 800 rpm, 37 °C, for 18 h) with the addition of 162 µL of digesting solution per sample (50 mM NH_4HCO_3, 0.1% (w/v) Protease-Max Surfactant (Promega, #V2071) in 50 mM NH_4HCO_3, 0.01 µg/µL of Sequencing Grade Modified Trypsin (Promega #V5111)) to facilitate an on-bead protein digest.

This digest was stopped using a protocol modified from Prikas et al. (2020). Beads were placed on a magnetic rack (NEB, #S1506S) so the supernatant could be transferred to new 1.5 mL tubes without the beads; peptides from digested proteins in the supernatant will henceforth be referred to as 'unbound' samples, whereas peptides attached to beads will be called 'bound' samples. For 'unbound samples', trifluoroacetic acid (TFA) was added to each sample to a final concentration of 0.1% (w/v), centrifuged (20,000 rcf, 22 °C, 20 min), and the supernatant transferred to new 1.5 mL tubes. To ensure as much protein as possible was recovered, 0.1% (w/v) of TFA was then added to the previously centrifuged 1.5 mL tubes, re-centrifuged (20,000 rcf, 22 ˚C, 10 min), and the resulting supernatant pooled with the first supernatant sample. For 'bound' samples, magnetic beads were first gently agitated on a tube rotator in 250 µL of elution buffer (EB) solution (80% (v/v) acetonitrile, 0.2% (w/v) TFA, 0.1% (v/v) formic acid), magnetised so the supernatant could be transferred, and then boiled in 200 µL of EB solution at 800 rpm, 95 ˚C for 5 min. These beads were then magnetised again so that supernatant could be transferred for TFA treatment.

Peptides for both 'unbound' and 'bound' samples were desalted using tC18 cartridges (Waters, #WAT036810), vacuum-dried at ambient temperature for ~3 h, and then resuspended in a

compatible solution (0.2% (v/v) heptafluorobutyric acid in 1% (v/v) formic acid) for liquid chromatography with tandem mass spectrometry (LC-MSMS).

2. Mass spectrometry

For technical reasons, we used two mass spectrometers as outlined in Prikas et al. (2020) – the ThermoFisher Scientific Q-Exactive Orbitrap (QE) and ThermoScientific Orbitrap Exploris (Exploris). Samples from biological replicates 1 and 2 were run on the QE, replicates 4 and 5 were run on the Exploris, and replicate 3 was run on both machines. We treated these as 6 separate experiments, although there were only 5 biological replicates, as considerably more proteins were identified using the Exploris compared with the QE – for this reason, the Exploris was used for subsequent experiments. The overlap between learning proteomes for each biological replicate (i.e., proteins unique to 'TurboID, trained') has been summarised in **Figure S2**, based on the mass spectrometer used. The resulting data was processed using the MASCOT search engine (Matrix Science) and *C. elegans* Swiss-Prot database (downloaded on 01/02/2021). An MS/MS ion search was performed with the following settings: 'semi-trypsin' enzyme, 'monoisotopic' mass values, 'unrestricted' protein mass, '± 5 ppm' peptide tolerance, '± 0.05 Da' fragment mass tolerance, and '3' maximum missed cleavages. Biotin (K), Carbamidomethyl (C), Oxidation (M), Phospho (ST), and Phospho (Y) were selected as variable modifications. Data from the MASCOT search engine will be made accessible on the Dryad platform upon publication.

3. Data analysis

We designated protein identities from all mass spectrometry experiments as either 'assigned' or 'unassigned' hits (see **Tables S2-S3** for full lists). 'Assigned proteins/hits' are defined as proteins

identified by MASCOT, based on peptide sequences detected during mass spectrometry, with at least one unique peptide detected for that protein. We did not restrict our definition of 'assigned hits' to any peptide or protein score threshold. In contrast, 'unassigned hits' were determined using peptide sequences with a peptide score ≥ 15, but that were not assigned a protein identity by MASCOT, as the peptide was not detected as unique for a specific protein by MASCOT. Our criteria for 'unassigned hits' was the protein identity required (1) at least one peptide for the protein with a peptide score ≥ 15 calculated by MASCOT, (2) an 100% identity match between peptide sequences with a peptide score ≥ 15 and the protein determined by BLAST, and (3) an e value < 0.05 for the identity match calculated by BLAST, such that a smaller number represents an increased probability that the identity is true and not given by random chance. To collect a list of unassigned hits, we used a custom Python script to perform bulk BLAST-p searches for these sequences using the 'Reference proteins (refseq_protein)' database and '*Caenorhabditis elegans* (taxid:6239)' organism. This Python code uses Anaconda 2023.03 (Anaconda Software Distribution, November, 2016), BioPython 1.78 (Cock et al., 2009), and pandas software 1.5.3 (The pandas development team, 2020). Our pipeline then assigned protein accession numbers to searched peptide sequences where the percent identity = 100% and e value < 0.05. Finally, the Batch Entrez online software available on the NCBI website was used to convert accession numbers to protein identities. This generated lists of unassigned hits for samples in each biological replicate.

Proteins that met the following criteria were selected for testing in the learning assays described below: (i) Present in ≥1 biological replicate of mass spectrometry data for only TurboID, trained worms, (ii) Expressed in neuron/s according to CenGEN (threshold = 2) (Taylor et al., 2021), excluding *elo-6* due to its specific expression in the ASER neuron (threshold = unfiltered), and

(iii) Documented previously as a non-mitochondrial protein. Proteins detected in TurboID, trained worms only were calculated as follows: (1) proteins detected from 'non-transgenic, control' and 'non-transgenic, trained' worms were subtracted from corresponding protein lists generated from TurboID worms, and then (2) proteins that overlap between lists for 'TurboID, control' and 'TurboID, trained' worms were subtracted from each other. Venn diagrams were generated using an online tool (https://bioinformatics.psb.ugent.be/webtools/Venn/).

GO term analyses were achieved using STRING (version 12.0), Cytoscape (version 3.10.0), and the Cytoscape App ClueGO (version 2.5.10). All 1010 proteins from 'TurboID, trained' lists were entered into STRING as a single list, so k-means clustering could be utilised to separate proteins into 10 clusters, to parse data into more accessible clusters with enough proteins to output enriched GO terms. The 'tabular text' protein-protein interaction information exported from STRING for each cluster was then uploaded onto Cytoscape. GO term analyses were performed with ClueGO for the following categories per cluster using default settings (network specificity in brackets): (i) Cellular component (medium), (ii) Biological process (detailed), and (iii) Molecular function (medium). The ClueGO results were exported to spreadsheets for each cluster as **Tables S4-S5**, such that each row corresponds to a GO term based on gene/s within a specific cluster, **Table S4** (cellular component) and **Table S5** (biological process/molecular function). Each list of genes in each row was consolidated with other gene lists with a matching or similar corresponding GO term, to generate the data shown in **Tables S4-S5**, **Figure 2**, and **Figure S3**. Nodes that did not show protein-protein interactions with other nodes and/or were not categorised into any GO term by ClueGO were manually categorised through a literature search and added onto these figures through Cytoscape.

Behavioural assays

We adapted previously established methods to perform three behavioural paradigms that model associative learning: (1) Salt associative learning (Leo et al., 2023, Hiroki et al., 2022, Nagashima et al., 2019), (2) Salt aversive learning (Lim et al., 2018), and (3) Butanone appetitive learning (Kauffman et al., 2010).

1. Salt associative learning

This experiment had three groups: naïve worms that did not undergo training, 'control' worms that were paired with 100 mM NaCl and food, and 'trained' worms that were paired with no NaCl and food. Worms from all groups were washed off agar plates using washing buffer as in the '**Protein extraction and quantification**' section above. The naïve group could be then immediately tested for their innate response to salt using the salt chemotaxis assay (Rahmani et al., 2024). For experimental groups, worms were washed a third time with washing buffer (containing 50 mM NaCl; for 'control') or no-salt buffer (washing buffer without NaCl; for 'trained'). These washing steps were completed within 10 min per group. Trained animals were placed on 9 cm 'conditioning plates' containing salt-deficient agar (5 mM K_3PO_4 (pH 6.0), 1 mM $CaCl_2$, 1 mM $MgSO_4$, 2.0% (w/v) agar) and their bacterial food source *E. coli* MG1655 *bioB::kan*. Control worms were placed on conditioning plates containing salt-deficient agar supplemented with 100 mM NaCl. Worms were left on these plates for 6 h at 22 °C in the dark (Nagashima et al., 2019). Following the training step, worms were washed twice with washing buffer and then transferred within 2 min to test their learning capacity using a salt chemotaxis assay (described below in **Section 3**).

The food source (*E. coli* MG1655 *bioB::kan*) was prepared 3-4 d before each experiment, which involved (1) pelleting 1.5 mL of bacteria by centrifugation at 11,000 rpm for 30 sec in a 2 mL tube, (2) discarding the supernatant, and then (3) vortexing cells in 750 µL of modified LB. 750

μL of washed bacteria was transferred onto each conditioning plate, left to dry at room temperature overnight, and then left at 22 °C for 3-4 d before use. Notably, the salt-deficient agar in conditioning plates for training contains 0.728 mM of NaCl due to the use of modified LB, meaning that it is not completely lacking NaCl but has only a very small amount.

2. Salt aversive learning

This assay involves three groups: naïve (did not undergo training), mock-conditioned worms (paired with no salt and no food), and conditioned worms (paired with 50 mM NaCl and no food). All groups were first washed as described in the **Salt associative learning** section, except the third wash was performed with no-salt buffer (for 'mock-conditioned') or washing buffer (for 'conditioned'). Naïve worms were placed on chemotaxis assay plates after two washes with washing buffer. Worms undergoing conditioning were incubated at room temperature for 3 h in 1.5 mL tubes, placed on a shaker at 175 rpm. These tubes contained no-salt buffer for 'mock-conditioned' groups or washing buffer (containing 50 mM NaCl) for 'conditioned' groups. Worms were pelleted by sedimentation for 1-2 min before use in salt chemotaxis assays as described below.

3. Salt chemotaxis assay

Salt chemotaxis assay plates contain a salt gradient prepared by placing 5 mm cubes of salt-deficient agar on top of salt-deficient agar, with one cube containing 0 mM salt on one side and another cube containing 200 mM salt on the other side (Jang et al., 2019) (see **Figure S1A** for a schematic). Worms were allowed to crawl freely on chemotaxis assay plates for 45 min in the dark at 22 °C (Nagashima et al., 2019), becoming immobilised when they encountered the paralytic agent sodium azide at the extremes of the salt gradient. These animals were then

counted within the regions outlined in **Figure S1B** to calculate the salt chemotaxis index (CI) based on the below equation:

$$\frac{\#\ worms\ on\ high\ salt - \#\ worms\ on\ low\ salt}{\#\ total\ population - \#\ worms\ on\ origin}$$

CI values range from -1.0 (strong preference for low salt concentrations) to +1.0 (strong preference for high salt concentrations). Data was excluded for salt associative learning experiments when wild-type N2 displayed (1) an average CI ≤ 0.6499 for naïve or control groups and/or (2) an average CI either < -0.0499 or >0.5499 for trained groups.

TurboID worm populations exposed to biotin during training by salt associative learning were scored for the total percentage of transgenic animals on chemotaxis assay plates (**Figure 1B**). This is because TurboID worms expressed the enzyme from an extrachromosomal array, and this scoring confirmed the presence of TurboID-positive animals in worm pellets to be used for downstream proteomics. To do this, fluorescing worms only (identified by co-injection marker P*unc-122::rfp*) were counted in all zones excluding the origin and compared to the total number of worms (fluorescing and non-fluorescing worms) in these zones. The average percentage for TurboID-positive worms was 27-57% per biological replicate.

4. Butanone appetitive learning

The following experimental groups were used for this assay: naïve (did not undergo training), pre-starved (starved for 1 h), mock-conditioned (starved for 1 h and then exposed to ethanol and food), and conditioned worms (starved for 1 h and then exposed to butanone and food). Briefly, *C. elegans* populations were washed twice in M9 buffer (86 mM NaCl, 42 mM Na_2HPO_4, 22 mM K_3PO_4 (pH 6.0), 1 mM M_gSO_4) to constitute the 'naïve' group in butanone chemotaxis

assays. Remaining groups were then subjected to 1 h 'pre-starvation' in M9 buffer on a shaker (175 rpm at room temperature). Mock-conditioned and conditioned groups were then transferred onto NGM plates with *E. coli* OP50 for 'mock-conditioning' or 'conditioning'. The mock-conditioned groups were exposed to 95% (v/v) ethanol (Chem-Supply. # EA043) streaked onto plate lids, whereas conditioned groups were exposed to 10% (v/v) butanone (Sigma-Aldrich, #360473; final concentration, diluted in 95% ethanol) for 1 h. All groups were transferred onto salt-deficient agar in 6 cm petri dishes to perform butanone chemotaxis assays as described in Kauffman et al. (2010) and Rahmani et al. (2024). Butanone chemotaxis assay plates were prepared with two equidistant points both spotted with sodium azide, as well as 95% (v/v) ethanol on one position and 10% (v/v) butanone in 95% (v/v) ethanol on the other spot. Butanone chemotaxis behaviour was quantified using the following equation:

$$\frac{\#\ worms\ on\ butanone - \#\ worms\ on\ ethanol}{\#\ total\ population - \#\ worms\ on\ origin}$$

The chemotaxis index (CI) ranges between -1.0 (complete ethanol attraction) to +1.0 (complete butanone attraction). Each assay plate constituted one technical replicate; technical replicates generated from <50 total worms or >250 worms were excluded from the data shown in **Figure 5**.

5. Statistical analyses

For learning assays, an ordinary two-way ANOVA with Tukey's multiple comparisons post-test ($\alpha = 0.05$) was performed to compare differences between mean CI values for each group. Exact p-values for each statistical comparison are reported in **Table S7.**

Data Availability

The source data from this publication, including mass spectrometry data (files from MASCOT search), raw data from learning assays, videos from behavioural tracking, have been uploaded to the **Dryad database:** this will be made available publicly following publication. In the interim, please contact the corresponding author for requests for raw data. Custom Python code is available through **GitHub** (https://github.com/ChewWormLab/Chew-Worm-Lab-Post-Mass-Spectrometry-Peptide-processing). *C. elegans* strains used in this study are available upon request.

Acknowledgements

Many thanks go to our colleagues in the Worm Neuroscience lab (Flinders) for thoughtful discussions, particularly Dr Michaela Johnson for proofreading this manuscript, and to our colleagues at Flinders University – A/Prof Arne Ittner, Prof Karin Nordström, Dr Alyce Martin – for scientific feedback and advice. We are grateful to Dr Emmanuel Prikas for expert technical advice and assistance. Mass spectrometry data were obtained in the Bioanalytical Mass Spectrometry Facility of the University of New South Wales. We sincerely thank Prof John E Cronan (University of Illinois, USA) for providing *E. coli* strain MG1655. We gratefully acknowledge the *Caenorhabditis* Genetics Centre, which is supported by the National Institutes of Health (P40 OD010440), for providing some of the strains used in this study.

Funding: A.R is funded by a PhD scholarship supported through the Australian Graduate Research Training Program (Flinders University). A.M is funded by a Flinders University Research Scholarship (Flinders University). Y.L.C is funded by the National Health and Medical Research Council (NHMRC) (GNT1173448), the Australian Research Council (DP220102511),

the Flinders University Parental Leave Research Support Scheme, the Flinders University Impact Seed Funding Grant for Early Career Researchers, and the Flinders Foundation Mary Overton Senior Research Fellowship.

Author contributions: Conceptualisation/Methodology/Project administration/Supervision – AP, YLC. Data curation/Formal analysis/Investigation – AR, AM, EA, RA, AP, YLC. Funding acquisition – YLC. Visualisation/Writing – original draft/Writing – review & editing – AR, YLC.

Conflicts of Interest: The authors have no conflicts of interest to declare.

References

AKEFE, I. O., SABER, S. H., MATTHEWS, B., VENKATESH, B. G., GORMAL, R. S., BLACKMORE, D. G., ALEXANDER, S., SIERIECKI, E., GAMBIN, Y., BERTRAN-GONZALEZ, J., VITALE, N., HUMEAU, Y., GAUDIN, A., ELLIS, S. A., MICHAELS, A. A., XUE, M., CRAVATT, B., JOENSUU, M., WALLIS, T. P. & MEUNIER, F. A. 2024. The DDHD2-STXBP1 interaction mediates long-term memory via generation of saturated free fatty acids. *EMBO J*.

ANACONDA SOFTWARE DISTRIBUTION. November, 2016. *Anaconda: Computer software* [Online]. Vers. 2-2.4.0. Available: https://anaconda.com [Accessed 2023].

ARELLANO-CARBAJAL, F., BRISEÑO-ROA, L., COUTO, A., CHEUNG, B. H., LABOUESSE, M. & DE BONO, M. 2011. Macoilin, a conserved nervous system-specific ER membrane protein that regulates neuronal excitability. *PLoS Genet,* 7**,** e1001341.

ARTAN, M., BARRATT, S., FLYNN, S. M., BEGUM, F., SKEHEL, M., NICOLAS, A. & DE BONO, M. 2021. Interactome analysis of Caenorhabditis elegans synapses by TurboID-based proximity labeling. *J Biol Chem,* 297**,** 101094.

ARTAN, M., HARTL, M., CHEN, W. & DE BONO, M. 2022. Depletion of endogenously biotinylated carboxylases enhances the sensitivity of TurboID-mediated proximity labeling in *Caenorhabditis elegans*. *J Biol Chem,* 298**,** 102343.

BAILEY, C. H., KANDEL, E. R. & HARRIS, K. M. 2015. Structural Components of Synaptic Plasticity and Memory Consolidation. *Cold Spring Harb Perspect Biol,* 7**,** a021758.

BAKALARSKI, C. E., ELIAS, J. E., VILLÉN, J., HAAS, W., GERBER, S. A., EVERLEY, P. A. & GYGI, S. P. 2008. The impact of peptide abundance and dynamic range on stable-isotope-based quantitative proteomic analyses. *J Proteome Res,* 7**,** 4756-65.

BARZILAI, A., KENNEDY, T. E., SWEATT, J. D. & KANDEL, E. R. 1989. 5-HT modulates protein synthesis and the expression of specific proteins during long-term facilitation in *Aplysia* sensory neurons. *Neuron,* 2**,** 1577-1586.

BEETS, I., ZHANG, G., FENK, L. A., CHEN, C., NELSON, G. M., FÉLIX, M. A. & DE BONO, M. 2020. Natural Variation in a Dendritic Scaffold Protein Remodels Experience-Dependent Plasticity by Altering Neuropeptide Expression. *Neuron,* 105**,** 106-121.e10.

BRADSHAW, K. D., EMPTAGE, N. J. & BLISS, T. V. 2003. A role for dendritic protein synthesis in hippocampal late LTP. *Eur J Neurosci,* 18**,** 3150-2.

BRANON, T. C., BOSCH, J. A., SANCHEZ, A. D., UDESHI, N. D., SVINKINA, T., CARR, S. A., FELDMAN, J. L., PERRIMON, N. & TING, A. Y. 2018. Efficient proximity labeling in living cells and organisms with TurboID. *Nature biotechnology,* 36**,** 880-887.

BRENNER, S. 1974. The genetics of Caenorhabditis elegans. *Genetics,* 77**,** 71-94.

CAO, L., JIAO, X., ZUZGA, D. S., LIU, Y., FONG, D. M., YOUNG, D. & DURING, M. J. 2004. VEGF links hippocampal activity with neurogenesis, learning and memory. *Nature Genetics,* 36**,** 827-835.

CHEN, C.-C., WU, J.-K., LIN, H.-W., PAI, T.-P., FU, T.-F., WU, C.-L., TULLY, T. & CHIANG, A.-S. 2012. Visualizing Long-Term Memory Formation in Two Neurons of the *Drosophila* Brain. *Science,* 335**,** 678-685.

COCK, P. J. A., ANTAO, T., CHANG, J. T., CHAPMAN, B. A., COX, C. J., DALKE, A., FRIEDBERG, I., HAMELRYCK, T., KAUFF, F., WILCZYNSKI, B. & DE HOON, M. J. L. 2009. Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics,* 25**,** 1422-1423.

CONG, Y., YANG, H., ZHANG, P., XIE, Y., CAO, X. & ZHANG, L. 2020. Transcriptome Analysis of the Nematode *Caenorhabditis elegans* in Acidic Stress Environments. *Front Physiol,* 11**,** 1107.

DAS, S., LITUMA, P. J., CASTILLO, P. E. & SINGER, R. H. 2023. Maintenance of a short-lived protein required for long-term memory involves cycles of transcription and local translation. *Neuron,* 111**,** 2051-2064.e6.

FAZELI, M. S., CORBET, J., DUNN, M. J., DOLPHIN, A. C. & BLISS, T. V. 1993. Changes in protein synthesis accompanying long-term potentiation in the dentate gyrus *in vivo*. *J Neurosci,* 13**,** 1346-53.

FENG, J., ZHOU, Y., CAMPBELL, S. L., LE, T., LI, E., SWEATT, J. D., SILVA, A. J. & FAN, G. 2010. Dnmt1 and Dnmt3a maintain DNA methylation and regulate synaptic function in adult forebrain neurons. *Nat Neurosci,* 13**,** 423-30.

FROEHLICH, J. J., RAJEWSKY, N. & EWALD, C. Y. 2021. Estimation of *C. elegans* cell- and tissue volumes. *MicroPubl Biol,* 2021.

GAUDET, P., LIVSTONE, M. S., LEWIS, S. E. & THOMAS, P. D. 2011. Phylogenetic-based propagation of functional annotations within the Gene Ontology consortium. *Briefings in Bioinformatics,* 12**,** 449-462.

GOTTSCHLING, D. C., DÖRING, F. & LÜERSEN, K. 2017. Locomotion Behavior Is Affected by the Gα(S) Pathway and the Two-Pore-Domain K(+) Channel TWK-7 Interacting in GABAergic Motor Neurons in *Caenorhabditis elegans*. *Genetics,* 206**,** 283-297.

GYURKÓ, M. D., CSERMELY, P., SŐTI, C. & STETÁK, A. 2015. Distinct roles of the RasGAP family proteins in *C. elegans* associative learning and memory. *Sci Rep,* 5**,** 15084.

HAWKINS, R. D. & BYRNE, J. H. 2015. Associative learning in invertebrates. *Cold Spring Harb Perspect Biol,* 7.

HAYASHIZAKI, S., IINO, Y. & YAMAMOTO, M. 1998. Characterization of the *C. elegans gap-2* gene encoding a novel Ras-GTPase activating protein and its possible role in larval development. *Genes to Cells,* 3**,** 189-202.

HERNANDEZ, P. J. & ABEL, T. 2008. The role of protein synthesis in memory consolidation: Progress amid decades of debate. *Neurobiology of Learning and Memory,* 89**,** 293-311.

HIROKI, S. & IINO, Y. 2022. The redundancy and diversity between two novel PKC isotypes that regulate learning in *Caenorhabditis elegans*. *Proc Natl Acad Sci U S A,* 119.

HIROKI, S., YOSHITANE, H., MITSUI, H., SATO, H., UMATANI, C., KANDA, S., FUKADA, Y. & IINO, Y. 2022. Molecular encoding and synaptic decoding of context during salt chemotaxis in *C. elegans*. *Nature Communications,* 13**,** 2928.

HOLZER, E., RUMPF-KIENZL, C., FALK, S. & DAMMERMANN, A. 2021. A Modified TurboID Approach Identifies Tissue-Specific Centriolar Components In *C. elegans*. *bioRxiv***,** 2021.12.20.473533.

HUCKLEBERRY, K. A., FERGUSON, L. B. & DREW, M. R. 2016. Behavioral mechanisms of context fear generalization in mice. *Learn Mem,* 23**,** 703-709.

HUKEMA, R. 2006. *Gustatory behaviour in C. elegans.*

HUKEMA, R. K., RADEMAKERS, S., DEKKERS, M. P. J., BURGHOORN, J. & JANSEN, G. 2006. Antagonistic sensory cues generate gustatory plasticity in *Caenorhabditis elegans*. *The EMBO journal,* 25**,** 312-322.

INBERG, S., ELKOBI, A., EDRI, E. & ROSENBLUM, K. 2013. Taste Familiarity Is Inversely Correlated with Arc/Arg3.1 Hemispheric Lateralization. *The Journal of Neuroscience,* 33**,** 11734.

JANG, M. S., TOYOSHIMA, Y., TOMIOKA, M., KUNITOMO, H. & IINO, Y. 2019. Multiple sensory neurons mediate starvation-dependent aversive navigation in *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences,* 116**,** 18673-18683.

JEONG, D.-E., LEE, Y. & LEE, S.-J. V. 2018. Western Blot Analysis of *C. elegans* Proteins. *In:* HUANG, L. E. (ed.) *Hypoxia: Methods and Protocols.* New York, NY: Springer New York.

JONG, Y. I., HARMON, S. K. & O'MALLEY, K. L. 2018. Intracellular GPCRs Play Key Roles in Synaptic Plasticity. *ACS Chem Neurosci,* 9**,** 2162-2172.

KANDEL, E. R. 2012. The molecular biology of memory: cAMP, PKA, CRE, CREB-1, CREB-2, and CPEB. *Molecular Brain,* 5**,** 14.

KAUFFMAN, A., PARSONS, L., STEIN, G., WILLS, A., KALETSKY, R. & MURPHY, C. 2011. C. elegans positive butanone learning, short-term, and long-term associative memory assays. *J Vis Exp*.

KAUFFMAN, A. L., ASHRAF, J. M., CORCES-ZIMMERMAN, M. R., LANDIS, J. N. & MURPHY, C. T. 2010. Insulin signaling and dietary restriction differentially influence the decline of learning and memory with age. *PLoS Biol,* 8**,** e1000372.

KHANNA, N., CRESSMAN III, C. P., TATARA, C. P. & WILLIAMS, P. L. 1997. Tolerance of the Nematode *Caenorhabditis elegans* to pH, Salinity, and Hardness in Aquatic Media. *Archives of Environmental Contamination and Toxicology,* 32**,** 110-114.

KITAZONO, T., HARA-KUGE, S., MATSUDA, O., INOUE, A., FUJIWARA, M. & ISHIHARA, T. 2017. Multiple Signaling Pathways Coordinately Regulate Forgetting of Olfactory Adaptation through Control of Sensory Responses in Caenorhabditis elegans. *J Neurosci,* 37**,** 10240-10251.

KNIAZEVA, M., CRAWFORD, Q. T., SEIBER, M., WANG, C. Y. & HAN, M. 2004. Monomethyl branched-chain fatty acids play an essential role in *Caenorhabditis elegans* development. *PLoS Biol,* 2**,** E257.

LAMPRECHT, R. 2014. The actin cytoskeleton in memory formation. *Progress in Neurobiology,* 117**,** 1-19.

LANS, H., RADEMAKERS, S. & JANSEN, G. 2004. A Network of Stimulatory and Inhibitory Gα-Subunits Regulates Olfaction in *Caenorhabditis elegans*. *Genetics,* 167**,** 1677-1687.

LEE, S. H., CHOI, J. H., LEE, N., LEE, H. R., KIM, J. I., YU, N. K., CHOI, S. L., LEE, S. H., KIM, H. & KAANG, B. K. 2008. Synaptic protein degradation underlies destabilization of retrieved fear memory. *Science,* 319**,** 1253-6.

LEO, T. H. T., GARRETT, A. L., STEVEN, J. C., JACQUELIN, H., CASSANDRA, C. P. & HANNES, E. B. 2023. Restructuring of an asymmetric neural circuit during associative learning. *bioRxiv***,** 2023.01.12.523604.

LIANG, V., ULLRICH, M., LAM, H., CHEW, Y. L., BANISTER, S., SONG, X., ZAW, T., KASSIOU, M., GOTZ, J. & NICHOLAS, H. R. 2014. Altered proteostasis in aging and heat shock response in C. elegans revealed by analysis of the global and de novo synthesized proteome. *Cell Mol Life Sci,* 71**,** 3339-61.

LIM, J. P., FEHLAUER, H., DAS, A., SARO, G., GLAUSER, D. A., BRUNET, A. & GOODMAN, M. B. 2018. Loss of CaMKI Function Disrupts Salt Aversive Learning in *C. elegans*. *The Journal of neuroscience : the official journal of the Society for Neuroscience,* 38**,** 6114-6129.

LIN, C. H. A., TOMIOKA, M., PEREIRA, S., SELLINGS, L., IINO, Y. & KOOY, D. V. D. 2010. Insulin Signaling Plays a Dual Role in *Caenorhabditis elegans* Memory Acquisition and Memory Retrieval. *The Journal of Neuroscience,* 30**,** 8001.

LINDSAY, J. H., MATHIES, L. D., DAVIES, A. G. & BETTINGER, J. C. 2022. A neuropeptide signal confers ethanol state dependency during olfactory learning in *Caenorhabditis elegans*. *Proc Natl Acad Sci U S A,* 119**,** e2210462119.

MATSUMOTO, Y., MATSUMOTO, C. S. & MIZUNAMI, M. 2018. Signaling Pathways for Long-Term Memory Formation in the Cricket. *Front Psychol,* 9**,** 1014.

MORUD, J., HARDEGE, I., LIU, H., WU, T., CHOI, M.-K., BASU, S., ZHANG, Y. & SCHAFER, W. R. 2021. Deorphanization of novel biogenic amine-gated ion channels identifies a new serotonin receptor for learning. *Current Biology,* 31**,** 4282-4292.e6.

MYHRER, T. 2003. Neurotransmitter systems involved in learning and memory in the rat: a meta-analysis based on studies of four behavioral tasks. *Brain Research Reviews,* 41**,** 268-287.

NAGASHIMA, T., IINO, Y. & TOMIOKA, M. 2019. DAF-16/FOXO promotes taste avoidance learning independently of axonal insulin-like signaling. *PLoS genetics,* 15**,** e1008297-e1008297.

OHNO, H., KATO, S., NAITO, Y., KUNITOMO, H., TOMIOKA, M. & IINO, Y. 2014. Role of synaptic phosphatidylinositol 3-kinase in a behavioral learning response in *C. elegans*. *Science,* 345**,** 313-317.

PARK, H. & KAANG, B. K. 2019. Balanced actions of protein synthesis and degradation in memory formation. *Learn Mem,* 26**,** 299-306.

PARK, Y. S., LEE, Y. S., CHO, N. J. & KAANG, B. K. 2000. Alternative splicing of *gar-1*, a *Caenorhabditis elegans* G-protein-linked acetylcholine receptor gene. *Biochem Biophys Res Commun,* 268**,** 354-8.

PEDREIRA, M. E., DIMANT, B., TOMSIC, D., QUESADA-ALLUE, L. A. & MALDONADO, H. 1995. Cycloheximide inhibits context memory and long-term habituation in the crab *Chasmagnathus*. *Pharmacology Biochemistry and Behavior,* 52**,** 385-395.

PEIXOTO, L. L., WIMMER, M. E., POPLAWSKI, S. G., TUDOR, J. C., KENWORTHY, C. A., LIU, S., MIZUNO, K., GARCIA, B. A., ZHANG, N. R., GIESE, K. P. & ABEL, T. 2015. Memory acquisition and retrieval impact different epigenetic processes that regulate gene expression. *BMC Genomics,* 16**,** S5.

PENG, S., ZHANG, Y., ZHANG, J., WANG, H. & REN, B. 2010. ERK in learning and memory: a review of recent research. *Int J Mol Sci,* 11**,** 222-232.

PENG, S., ZHANG, Y., ZHANG, J., WANG, H. & REN, B. 2011. Glutamate receptors and signal transduction in learning and memory. *Molecular Biology Reports,* 38**,** 453-460.

PERSHINA, E. V., KULAGINA, T. P., SAVINA, T. A., ARIPOVSKY, A. V., LEVIN, S. G. & ARKHIPOV, V. I. 2022. Changes in the level of fatty acids in the brain of rats during memory acquisition. *Behav Brain Res,* 417**,** 113599.

POPOVICI, C., ISNARDON, D., BIRNBAUM, D. & ROUBIN, R. 2002. *Caenorhabditis elegans* receptors related to mammalian vascular endothelial growth factor receptors are expressed in neural cells. *Neuroscience Letters,* 329**,** 116-120.

PRIKAS, E., POLJAK, A. & ITTNER, A. 2020. Mapping p38α mitogen-activated protein kinase signaling by proximity-dependent labeling. *Protein Science,* 29**,** 1196-1210.

PUTRENKO, I., ZAKIKHANI, M. & DENT, J. A. 2005. A Family of Acetylcholine-gated Chloride Channel Subunits in *Caenorhabditis elegans*. *Journal of Biological Chemistry,* 280**,** 6392-6398.

RAHMANI, A. & CHEW, Y. L. 2021. Investigating the molecular mechanisms of learning and memory using *Caenorhabditis elegans*. *Journal of Neurochemistry,* 159**,** 417-451.

RAHMANI, A., MCMILLEN, A., ALLEN, E., MINERVINI, C. & CHEW, Y. L. 2024. Behavioral Tests for Associative Learning in *Caenorhabditis elegans*. *In:* DWORKIN, S. (ed.) *Neurobiology: Methods and Protocols.* New York, NY: Springer US.

ROSENBERG, T., GAL-BEN-ARI, S., DIETERICH, D. C., KREUTZ, M. R., ZIV, N. E., GUNDELFINGER, E. D. & ROSENBLUM, K. 2014. The roles of protein expression in synaptic plasticity and memory consolidation. *Frontiers in Molecular Neuroscience,* 7.

ROUX, K. J., KIM, D. I., RAIDA, M. & BURKE, B. 2012. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J Cell Biol,* 196**,** 801-10.

RYU, H.-H. & LEE, Y.-S. 2016. Cell type-specific roles of RAS-MAPK signaling in learning and memory: Implications in neurodevelopmental disorders. *Neurobiology of Learning and Memory,* 135**,** 13-21.

SAKAI, N., OHNO, H., TOMIOKA, M. & IINO, Y. 2017. The intestinal TORC2 signaling pathway contributes to associative learning in *Caenorhabditis elegans*. *PLOS ONE,* 12**,** e0177900.

SANCHEZ, A. D., BRANON, T. C., COTE, L. E., PAPAGIANNAKIS, A., LIANG, X., PICKETT, M. A., SHEN, K., JACOBS-WAGNER, C., TING, A. Y. & FELDMAN, J. L. 2021. Proximity labeling reveals non-centrosomal microtubule-organizing center components required for microtubule growth and localization. *Current Biology,* 31**,** 3586-3600.e11.

SANCHEZ, A. D. & FELDMAN, J. L. 2021. A proximity labeling protocol to probe proximity interactions in *C. elegans*. *STAR Protoc,* 2**,** 100986.

SANDERS, J., NAGY, S., FETTERMAN, G., WRIGHT, C., TREININ, M. & BIRON, D. 2013. The *Caenorhabditis elegans* interneuron ALA is (also) a high-threshold mechanosensor. *BMC Neurosci,* 14**,** 156.

SASSONE-CORSI, P. 2012. The cyclic AMP pathway. *Cold Spring Harb Perspect Biol,* 4.

SCHADE, M. A., REYNOLDS, N. K., DOLLINS, C. M. & MILLER, K. G. 2005. Mutations that rescue the paralysis of *Caenorhabditis elegans ric-8* (synembryn) mutants activate the G alpha(s) pathway and define a third major branch of the synaptic signaling network. *Genetics,* 169**,** 631-49.

SHAHMORADI, A. 2015. *Analyses of improved long term memory in SHARP1 and SHARP2 double knockout mice.* University of Göttingen.

SMITH, W. B., STARCK, S. R., ROBERTS, R. W. & SCHUMAN, E. M. 2005. Dopaminergic stimulation of local protein synthesis enhances surface expression of GluR1 and synaptic transmission in hippocampal neurons. *Neuron,* 45**,** 765-79.

STEFANOSKA, K., GAJWANI, M., TAN, A. R. P., AHEL, H. I., ASIH, P. R., VOLKERLING, A., POLJAK, A. & ITTNER, A. 2022. Alzheimer's disease: Ablating single master site abolishes tau hyperphosphorylation. *Sci Adv,* 8**,** eabl8809.

STEFANOSKA, K., PRIKAS, E., LIN, Y., KOSONEN, R. & ITTNER, A. 2023. Remote memory engrams are controlled by encoding-specific tau phosphorylation. *bioRxiv***,** 2023.12.01.569663.

STEIN, G. M. & MURPHY, C. T. 2014. *C. elegans* positive olfactory associative memory is a molecularly conserved behavioral paradigm. *Neurobiology of Learning and Memory,* 115**,** 86-94.

SUTTON, M. A. & SCHUMAN, E. M. 2006. Dendritic Protein Synthesis, Synaptic Plasticity, and Memory. *Cell,* 127**,** 49-58.

TAKAYANAGI-KIYA, S., ZHOU, K. & JIN, Y. 2016. Release-dependent feedback inhibition by a presynaptically localized ligand-gated anion channel. *eLife,* 5**,** e21734.

TAYLOR, S. R., SANTPERE, G., WEINREB, A., BARRETT, A., REILLY, M. B., XU, C., VAROL, E., OIKONOMOU, P., GLENWINKEL, L., MCWHIRTER, R., POFF, A., BASAVARAJU, M., RAFI, I., YEMINI, E., COOK, S. J., ABRAMS, A., VIDAL, B., CROS, C., TAVAZOIE, S., SESTAN, N., HAMMARLUND, M., HOBERT, O. & MILLER, D. M. 2021. Molecular topography of an entire nervous system. *Cell,* 184**,** 4329-4347.e23.

THE PANDAS DEVELOPMENT TEAM. 2020. *pandas-dev/pandas: Pandas* [Online]. Latest ver.: Zenodo. Available: https://doi.org/10.5281/zenodo.3509134 [Accessed 2023].

TOMIOKA, M., ADACHI, T., SUZUKI, H., KUNITOMO, H., SCHAFER, W. R. & IINO, Y. 2006. The insulin/PI 3-kinase pathway regulates salt chemotaxis learning in Caenorhabditis elegans. *Neuron,* 51**,** 613-25.

TRUJILLO, G., NAKATA, K., YAN, D., MARUYAMA, I. N. & JIN, Y. 2010. A ubiquitin E2 variant protein acts in axon termination and synaptogenesis in *Caenorhabditis elegans*. *Genetics,* 186**,** 135-45.

VAN BUSKIRK, C. & STERNBERG, P. W. 2007. Epidermal growth factor signaling induces behavioral quiescence in *Caenorhabditis elegans*. *Nat Neurosci,* 10**,** 1300-7.

WALLIS, T. P., VENKATESH, B. G., NARAYANA, V. K., KVASKOFF, D., HO, A., SULLIVAN, R. K., WINDELS, F., SAH, P. & MEUNIER, F. A. 2021. Saturated free fatty acids and association with memory formation. *Nat Commun,* 12**,** 3443.

WATTEYNE, J., PEYMEN, K., VAN DER AUWERA, P., BORGHGRAEF, C., VANDEWYER, E., VAN DAMME, S., RUTTEN, I., LAMMERTYN, J., JELIER, R., SCHOOFS, L. & BEETS, I. 2020. Neuromedin U signaling regulates retrieval of learned salt avoidance in a *C. elegans* gustatory circuit. *Nat Commun,* 11**,** 2076.

Supplementary Figures

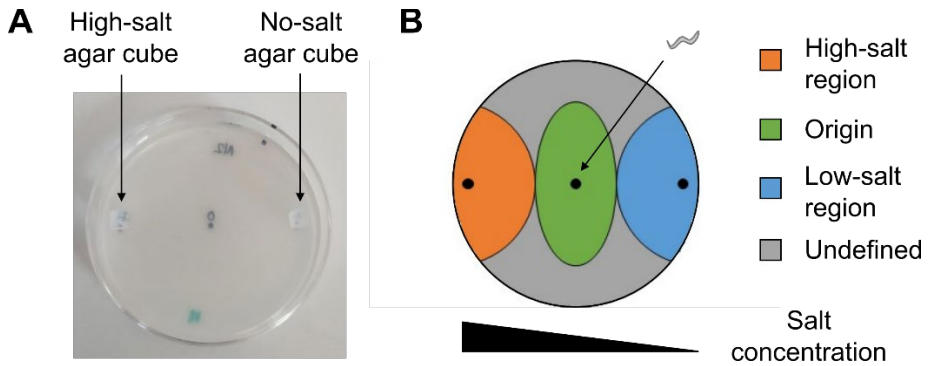
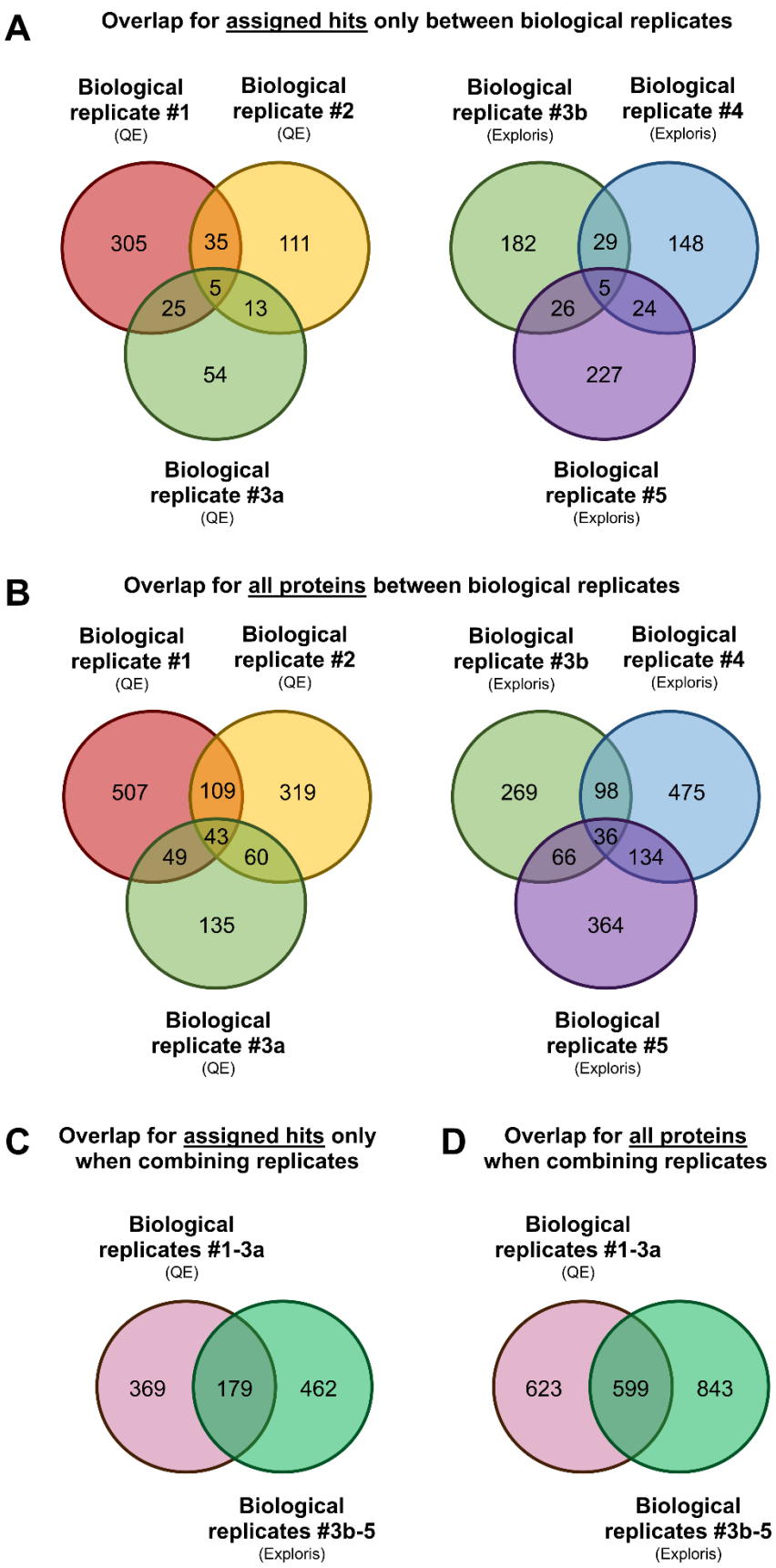


Figure S1: Design of chemotaxis assay plates to quantify salt chemotaxis behaviours. (A) Method used to generate a salt concentration gradient (0-200 mM NaCl) on chemotaxis assay plates: 5 mm cubes of no-salt/salt-deficient agar with or without high-salt (200 mM NaCl) were placed on top of salt-deficient agar, 5 mm from the plate edge (6.0 cm diameter). Cubes were incubated overnight at 22 °C, or incubated sequentially for (1) 5.5 h at 37 °C, (2) 15 min at 15 °C, and then (3) 0.5-2 h at 22 °C. Each cube was replaced with a 1 µL spot of 0.1 M sodium azide <2 min before use. **(B)** Template used to quantify salt chemotaxis on chemotaxis assay plates: regions of high salt (orange) and low salt (blue) were defined as a circle (diameter = 2.4 cm). These regions were cut-off on their longitudinal axis since they were centred 0.5 cm from the edge of each plate. Worms were transferred onto the region of origin (green; 2 cm short axis and 2.4 cm long axis). Points where agar cubes or worms were transferred onto are represented by a small black circle.


A
Overlap for assigned hits only between biological replicates
Biological replicate #1 (QE)
Biological replicate #2 (QE)
Biological replicate #3a (QE)
305
35
111
5
25
13
54
Biological replicate #3b (Exploris)
Biological replicate #4 (Exploris)
Biological replicate #5 (Exploris)
182
29
148
5
26
24
227
B
Overlap for all proteins between biological replicates
Biological replicate #1 (QE)
Biological replicate #2 (QE)
Biological replicate #3a (QE)
507
109
319
43
49
60
135
Biological replicate #3b (Exploris)
Biological replicate #4 (Exploris)
Biological replicate #5 (Exploris)
269
98
475
36
66
134
364
C
Overlap for assigned hits only when combining replicates
Biological replicates #1-3a (QE)
Biological replicates #3b-5 (Exploris)
369
179
462
D
Overlap for all proteins when combining replicates
Biological replicates #1-3a (QE)
Biological replicates #3b-5 (Exploris)
623
599
843

1

Figure S2: Assessing overlap for proteins detected in trained worms across all biological replicates. (A-B) Numbers represent (A) assigned proteins only or (B) all (assigned and unassigned) proteins detected with the Q Exactive (QE) (left) or the Exploris (right) in 'TurboID, trained' animals only; these proteins were identified by subtracting proteins also in 'non-transgenic, trained' and/or in 'TurboID, control' worms. Six protein lists are compared here since one replicate was run in both mass spectrometers. **(C-D)** Lists containing proteins specific to trained animals for each biological replicate were compared based on the mass spectrometer used for (C) assigned proteins only and (D) all proteins.

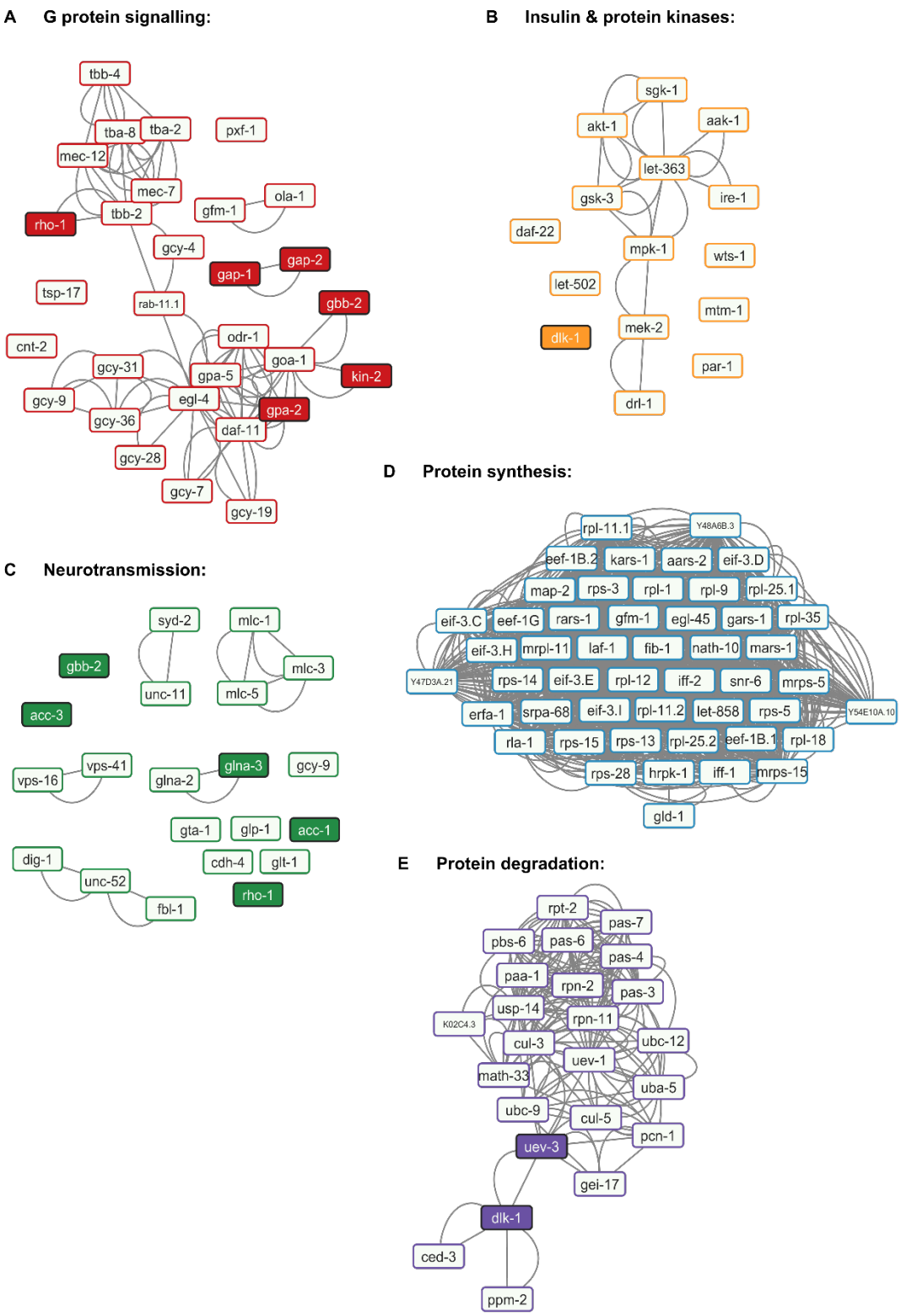
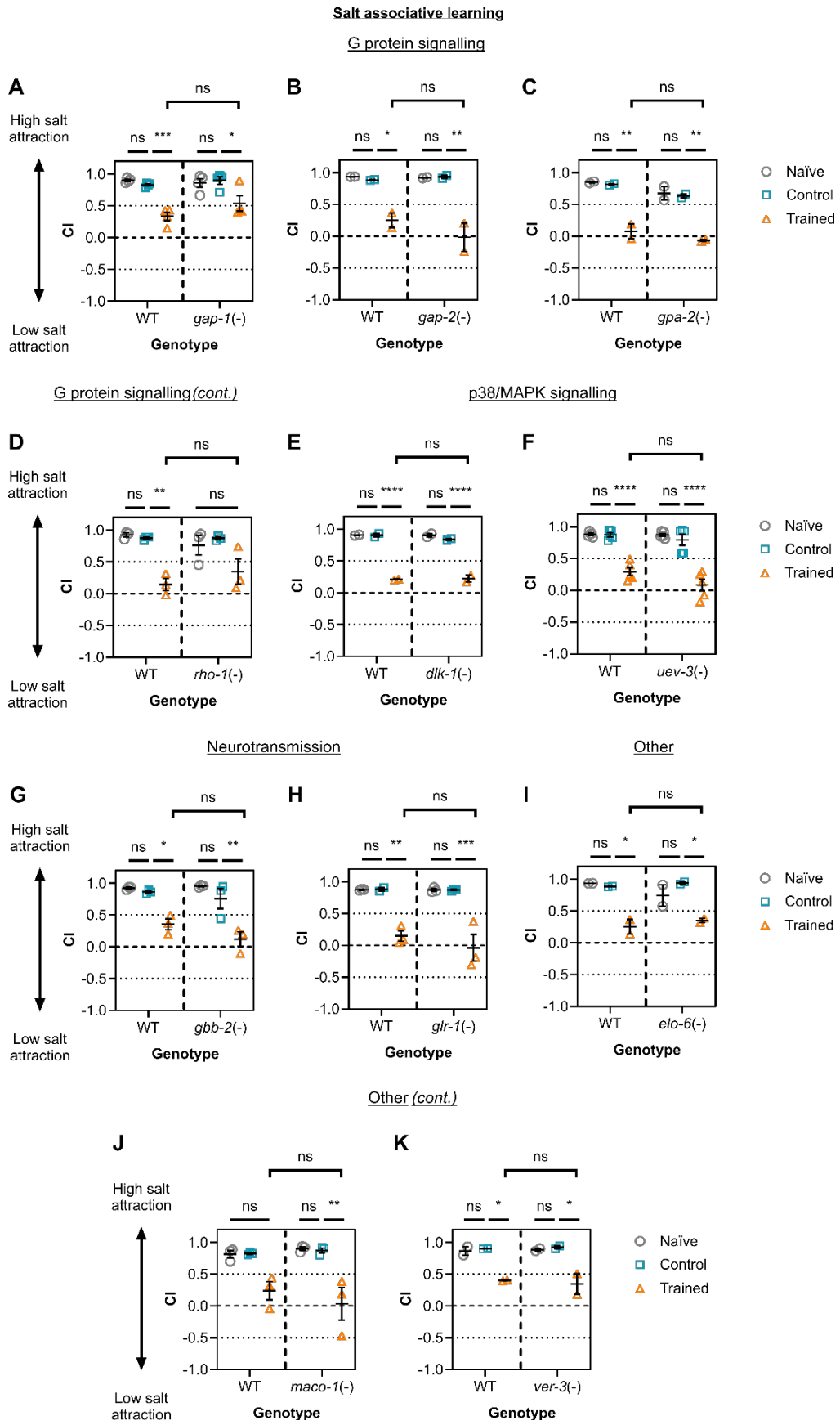


Figure S3: Analysis of molecular functions of assigned hits in the learning proteome. Analyses using GO terms for molecular function were performed with ClueGO (version 2.5.10)

using Cytoscape (version 3.10.0). Protein-protein interaction (PPI) networks generated for proteins assigned the GO term **(A)** G protein signalling, **(B)** Insulin and protein kinases, **(C)** Neurotransmission, **(D)** Protein synthesis, and **(E)** Protein degradation. PPI interactions were exported from STRING (version 12.0) to be visualised with Cytoscape using the protein lists in **Table S2**. Highlighted proteins were tested in this study using appropriate genetic mutants. Known information on molecular function and predictions based on homology were also used for several nodes (Putrenko et al., 2005, Park et al., 2000, Gottschling et al., 2017, Gaudet et al., 2011, Hayashizaki et al., 1998, Lans et al., 2004). The consolidated data from GO term analyses can accessed on **Table S5**.


Salt associative learning
G protein signalling
A
B
C
High salt attraction
Low salt attraction
CI
Genotype
WT
gap-1(-)
gap-2(-)
gpa-2(-)
ns
Naïve
Control
Trained
G protein signalling (cont.)
p38/MAPK signalling
D
E
F
rho-1(-)
dlk-1(-)
uev-3(-)
Neurotransmission
Other
G
H
I
gbb-2(-)
glr-1(-)
elo-6(-)
Other (cont.)
J
K
maco-1(-)
ver-3(-)

1

Figure S4: Testing of learning proteome hits identified through mass spectrometry experiments. Worms with mutations targeting molecular components (in brackets) in the following pathways were tested for their salt associative learning capacity, compared to wild-type (WT) worms: **(A-D)** G protein signalling (*gap-1, gap-2, gpa-2,* and *rho-1*), **(E-F)** p38/MAPK signalling (*dlk-1* and *uev-3*), **(G-H)** Neurotransmission (*gbb-2* and *glr-1*), and **(I-K)** Other proteins of interest (*elo-6, maco-1,* and *ver-3*). Each data point represents the average chemotaxis index (CI) per biological replicate (three technical replicates per biological replicate, 11-540 worms per technical replicate). n = 2-5 biological replicates. Statistical analysis: Two-way ANOVA & Tukey's multiple comparisons test (**** ≤ 0.0001; *** ≤ 0.001; ** ≤ 0.01; * ≤ 0.05; ns = non-significant). Error bars = mean ± SEM.

Supplementary Tables

Table S1: List of *C. elegans* lines used in this study.

Table S2: Protein lists for all assigned hits detected in TurboID, trained worms. *C. elegans* peptides were assigned corresponding protein identities by MASCOT and then assigned protein lists were generated for each biological replicate by subtracting proteins also present in 'non-transgenic, trained' worms and/or in 'TurboID, control' worms within the same replicate. Each protein list in rows 9-378 are defined in this table by its corresponding biological replicate, the mass spectrometer used, and the inclusion of data from unbound and bound peptide samples. Peptide samples generated for biological replicate 3 were run in both mass spectrometers used in this study, the Q Exactive or the Exploris, and the corresponding protein list is named '3a' and '3b' respectively.

Table S3: Protein lists for all unassigned hits detected in TurboID, trained worms. Protein identities were determined through bulk BLAST searching *C. elegans* peptides detected by mass spectrometry, rather than MASCOT, and then unassigned protein lists were generated for each experimental group in each biological replicate. Proteins also in 'non-transgenic, trained' and/or 'TurboID, control' lists for the same biological replicate were subtracted from the corresponding 'TurboID, trained' list. Each protein list in rows 9-587 are defined in this table by its corresponding biological replicate, the mass spectrometer used, and the inclusion of data from unbound and bound peptide samples. Peptide samples generated for biological replicate 3 were run in both mass spectrometers used in this study, the Q Exactive or the Exploris, and the corresponding protein list is named '3a' and '3b' respectively.

Table S4: Cellular component GO terms associated with assigned hits in the learning proteome. All (1010) confident protein hits detected in TurboID, trained worms were separated into 10 clusters by k-means clustering on STRING (version 12.0) and GO term analysis was conducted for each cluster using ClueGO (version 2.5.10) on Cytoscape (version 3.10.0). GO terms were then sorted into 'plasma membrane & extracellular matrix (ECM)', 'cilia & dendrites', 'cell body & cytoplasm', 'nucleus', 'Golgi apparatus', 'cytoskeleton', 'mitochondria', 'axon & vesicles', 'pre-synapse', and 'other' (i.e., non-informative terms) categories due to recurring and/or similar terms. Colour-coding is based on the GO Group. Each category is separated by underlined and bolded titles corresponding to the cellular component in rows 1-285. Consolidated lists of genes/proteins are detailed from rows 287-465.

Table S5: Biological process and molecular function GO terms associated with assigned hits in the learning proteome. All (1010) confident protein hits detected in TurboID, trained worms were separated into 10 clusters by k-means clustering on STRING (version 12.0) and GO term analysis was conducted for each cluster using ClueGO (version 2.5.10) on Cytoscape (version 3.10.0). GO terms were then sorted into 'protein synthesis', 'protein degradation', 'insulin & protein kinases', 'G protein signalling', and 'neurotransmission' categories due to recurring and/or similar terms. Colour-coding is based on the GO Group. Each category is separated by underlined and bolded titles corresponding to the biological processes listed above in rows 1-109. Consolidated lists of genes/proteins are detailed from rows 111-162.

Table S6: Molecular pathways for proteins detected in TurboID, trained *C. elegans* only. Proteins unique to trained animals that express TurboID were identified by subtracting those also in 'TurboID, control' and/or 'non-transgenic, trained' protein lists. Gene names are listed with assigned hits in bold. Each gene/protein has been grouped based on 'Biological Process' GO term annotations provided by STRING (version 12.0) and information available on WormBase (version WS290) based on data accessed during September-November 2023.

Table S7 | Statistical analyses for learning assays performed in this study. Data from statistical analyses are separated by figure with appropriate underlined and bolded titles, including **Figure 1B** (red, rows 1-7), **Figure 4** (yellow, rows 9-40), **Figure 5** (green, rows 42-101), and **Figure S4** (blue, rows 103-189). Statistical analyses were performed on GraphPad PRISM 8 using two-way ANOVA and Tukey's multiple comparisons tests (alpha value = $p < 0.05$). Columns left-to-right list the groups being compared, the mean difference (Diff.), the 95% confidence interval of diff., whether the difference is significant, the degree of significance (**** ≤ 0.0001; *** ≤ 0.001; ** ≤ 0.01; * ≤ 0.05; ns = non-significant), and the adjusted p value.