Behavioral screening of conserved RNA-binding proteins reveals CEY-1/YBX RNA-binding protein dysfunction leads to impairments in memory and cognition

Ashley N Hayden[1,2], Katie L Brandel[1,2], Paul R Merlau[2], Priyadharshini Vijayakumar[2], Emily J Leptich[1,2], Edward W Pietryk[2,3], Elizabeth S Gaytan[2,4], Connie W Ni[2,5], Hsiao-Tuan Chao[1,3,6,7,8], Jill A Rosenfeld[3,9], Rachel N Arey[2,10]*
*Corresponding Author: rachel.arey@bcm.edu

1. Department of Neuroscience, Baylor College of Medicine, Houston, TX 77030
2. Center for Precision Environmental Health, Baylor College of Medicine, Houston, TX, 77030
3. Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, 77030
4. Postbaccalaureate Research Education Program, Baylor College of Medicine, Houston, TX, 77030
5. Department of Neuroscience, Rice University, Houston, TX 77005
6. Department of Pediatrics, Division of Neurology and Developmental Neuroscience, Baylor College of Medicine, Houston, TX, 77030
7. Cain Pediatric Neurology Research Foundation Laboratories, Jan and Dan Duncan Neurological Research Institute, Texas Children's Hospital, Houston, TX, 77030
8. McNair Medical Institute, The Robert and Janice McNair Foundation, Houston, TX, 77030
9. Baylor Genetics Laboratories, Houston, TX 77021

Abstract

RNA-binding proteins (RBPs) are key regulators of translation and plasticity in the nervous system, which are processes required for learning and memory. Indeed, RBP dysfunction has been linked to a wide range of neurological disorders where cognitive impairments are a key symptom. However, the human genome encodes nearly 2,000 RBPs, many of which have yet to be characterized with regards to neurological phenotypes like associative behaviors. To address this, we used the model organism *C. elegans* to perform a targeted screen assessing the role of 20 conserved RBPs in learning and memory. We identified 16 associative memory regulators, eight of which are novel. Intriguingly, three novel memory regulators are in the *C. elegans* Y-Box (CEY) RNA-binding protein family: *cey-1*, *cey-2*, and *cey-3*. We determined that *cey-1* is the closest ortholog to the mammalian Y-Box RNA-binding (YBX) proteins and that CEY-1 is specifically required in the nervous system for the ability to remember. Additionally, CEY-1 is a memory promoting molecule, as increasing CEY-1 only in the nervous system improves memory. We also examined human copy number variation datasets and found that copy number deletions in the human *YBX1* and *YBX3* may be associated with neurological symptoms including intellectual disability, which mirror our findings in *C. elegans.* We identified a similar association with a rare, mono-allelic *YBX3* variant in individuals with neurological findings. Strikingly, introducing this predicted deleterious *YBX3* variant into the endogenous *cey-1* locus caused memory deficits in the worm. Our study highlights that high-throughput approaches in the worm reveal novel and conserved regulators of memory and identified a potential new source of rare neurological disease linked to RBP dysfunction.

Keywords: *C. elegans,* learning and memory, RNA-binding protein, Y-Box RNA-binding protein, variant modeling

Introduction

Careful regulation of RNA localization, translation, and stability is essential for the proper functioning of all tissues but is especially important in the nervous system. Neurons are composed of subcellular compartments, specifically somatic, axonal, and dendritic regions. Each region possesses a specialized and functionally distinct proteome (1). Moreover, these subcellular compartments must respond dynamically, and often uniquely, to external stimuli. A prominent example of this is local mRNA translation in response to plasticity-inducing stimuli, which is necessary for learning and memory (2,3). In this instance, both the pre- and post-synaptic compartment can undergo local protein synthesis in response to the same plasticity-inducing stimulus (1,4), yet mRNAs translated in response to that stimulus are compartment-specific, likely due to distinct mRNA regulators.

RNA-binding proteins (RBPs) are involved in each step of RNA regulation in neurons, controlling target RNA splicing, polyadenylation, localization, translation, and stability (5). As

such, their proper function is essential for neuronal plasticity and its associated behaviors – especially learning and memory. The importance of understanding the role of RBPs in the nervous system is underscored by a growing body of evidence that their dysfunction results in neurological disease, including disorders ranging from neurodevelopmental disorders to neurodegenerative disease (6–8). A frequent symptom of RBP-associated disorders are behavioral abnormalities and cognitive impairments, including Fragile X Messenger Ribonucleoprotein (FMRP) in Fragile X syndrome involving cognitive impairment and mobility deficits, *PUMILIO1 (PUM1)* associated spinocerebellar ataxia 47 involving seizures, *Quaking (QKI)* in schizophrenia, and both *Heterogeneous Ribonucleoprotein K (HNRNPK)* in Au-Kline syndrome and SON in Zhu-Tokita-Takenouchi-Kim (ZTTK) syndrome involving intellectual disability (9–13). In fact, RBPs outweigh all other protein types in their prevalence of predicted deleterious variation in Mendelian disorders, especially those that are diseases of the nervous system (7). However, despite their importance, many RBPs remain uncharacterized in the nervous system. Barriers remain to systematic characterization of RBPs, as such studies remain relatively slow in mammals and the human genome encodes nearly 2000 RBPs, most of which that remain to be comprehensively studied (14,15). Moreover, studying the tissue-specific function of RBPs is even more difficult as it requires challenging and costly genetic approaches. However, these issues can be circumvented in simpler model organisms like the nematode worm *Caenorhabditis elegans (C. elegans)* that allow for low-cost, rapid screening of molecules in a tissue-specific manner.

C. elegans have a rich history of use in high-throughput approaches and led to the discovery of many new regulators of neurological phenotypes, including molecularly conserved behaviors (16–18). Indeed, the use of the worm resulted in the discovery of many conserved regulators of disease, including those involving cognitive deficits (19–24). The *C. elegans* genome encodes almost 900 RBPs, over 80% of which have mammalian orthologs (25), and can be studied using thousands of publicly available mutants, including those that allow for neuron-specific RNAi for rapid tissue-specific genetic screening (26,27). Importantly, *C. elegans* offers the advantage of using these genetic tools to investigate the role of RBPs in memory and cognition, as they form molecularly conserved associative memories, including short-term (STM), intermediate-term (ITM), and long-term memories (LTM) (20,28–32). Given its wealth of molecular tools, ease of genetic screening, and molecularly conserved memory, *C. elegans* is an excellent model system to define the functions of neuronal RBPs in learning and memory *in vivo*.

Here, we have used the worm to perform a neuron-specific, targeted screen to discover conserved RBPs that control memory. Altogether, we find that 80% of our screened RBPs are memory regulators, eight of which are novel memory molecules. Importantly, our approach reveals a family of RBPs that are practically uncharacterized in their nervous system function and are essential for associative memory ability in the worm: the *C. elegans* Y-Box (CEY) RNA-binding protein family.

Of this family, we determine that CEY-1 is most closely related to mammalian YBXs and requires further study to understand the potentially conserved role of YBXs in the nervous system. We find that CEY-1 acts specifically in the nervous system to promote memory. We discovered that this role in memory is likely conserved, as human copy number losses of *YBX* genes may be associated with neurological symptoms, most commonly intellectual disability. We similarly describe one rare heterozygous variant in the human *YBX3* gene identified in two unrelated individuals with neurological symptoms. Importantly, introduction of the conserved *YBX3* variant into the endogenous *cey-1* locus causes memory deficits, supporting the idea that deleterious *YBX3* variants may contribute to human disease.

Altogether, we uncovered new associative memory regulators in an unbiased screen of conserved neuronal RBPs. We describe in detail a novel associative memory regulator, CEY-1/YBX, and highlight *C. elegans* as a discovery organism for associative memory regulators and potential disease-causing RBPs.

Results

Targeted screen identifies novel RBPs that regulate learning and memory

The *C. elegans* genome, which exhibits a high degree of conservation with mammals, encodes 887 predicted RBPs (33,34); thus, the worm is an excellent system to rapidly screen RBPs for conserved roles in memory. We prioritized candidate RBPs based on the following criteria to increase the likelihood of identifying memory regulators: 1) nervous system expression based upon transcriptomic studies, 2) evolutionary conservation with mammals, and 3) detection in synaptic regions across species, as RBPs are known to regulate plasticity at the synapse. By cross-referencing adult only, neuron specific transcriptomic datasets (24,35) as well as orthology tools including Wormcat and OrthoList2 (36,37), we identified 550 neuronally expressed RBPs with mammalian orthologs. After filtering this list based upon detection in *C. elegans* and mammalian synapses (35,38–42), 313 RBPs met screening criteria. Of those, we proceeded with the top 20 most synaptically-enriched RBPs based on our previous work identifying the *C. elegans* synaptic transcriptome (38,43) for behavioral study.

We assessed the roles of the RBPs in memory by performing adult-only RNAi in neuronally RNAi-sensitized animals *(punc119::sid-1)* (27) and measuring the effects of knockdown on various well-characterized, molecularly conserved associative behaviors. These behaviors included learning (0 min after training), protein synthesis-independent

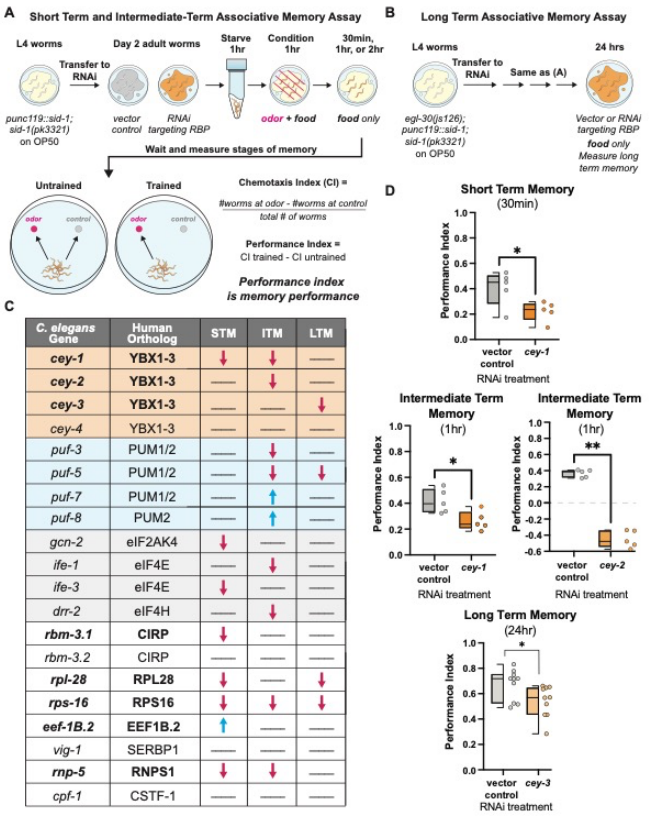
short term memory (STM, 30 min after training), and translation-dependent intermediate term memory (ITM, 1 hr after training) (**Figure 1A**) (29,30)**.** Effects of RBP knockdown on transcription and translation-dependent long term memory (LTM) were assessed in RNAi-sensitized animals with an additional *egl-30* gain of function mutation *(egl-30(js126), punc119::sid-1))* that increases $G_{\alpha q}$ signaling and confers the ability to form a LTM that is identical to wild-type animals with only one round of training, making them amenable to high-throughput screening (**Figure 1B**) (44).

Of the 20 RBPs tested, 80% control at least one type of memory (**Figure 1C**). Two classes of proteins known to control memory across species had multiple members represented in our screen candidates: the Pumilio and FBF RBP family (*pumilio1/2* orthologs *puf-3*, *puf-5*, *puf-7*, and *puf-8*) as well as four proteins involved in translation initiation (eIF4/2α orthologs *gcn-2, ife-1, ife-3,* and *drr-2*). Pumilios have been linked to both LTM and working memory in *Drosophila melanogaster* and mice (45,46), and we have previously described their importance in *C. elegans* memory (38,43). Similarly, eukaryotic translation initiation machinery eIF4E and eIF2α control memory formation, STM, and LTM in mice (47,48). However, to the best of our knowledge the requirement for this molecular machinery in olfactory associative memory was not previously demonstrated in the worm. Validating these two classes of known memory regulators strengthens the likelihood that novel molecules identified in our screen have a conserved function in higher organisms. Indeed, we found that 50% of memory regulating RBPs in our screen were not previously associated with memory phenotypes, and are therefore novel, bolded in **Figure 1C**. Importantly, no learning deficits were detected after knockdown of any RBP in our screen (**Supplemental Figure 1**), suggesting the deficits we identified are specific to memory and are not due to broad behavioral or neurological disruption. For all behavior timepoints where RNAi significantly altered memory, see **Supplemental Figure 2.**

Identification of CEY family as novel memory regulators

Strikingly, all members of the *C. elegans* Y-Box (CEY) RBP family (*cey-1-4),* orthologs of the mammalian *Y-box binding* (*YBX*) family, met our screening criteria despite being uncharacterized in learning and memory behavior. The CEY RBPs regulate fertility in the worm (49,50), and mammalian YBX proteins are translational regulators involved in cancer metastasis and cell cycle progression (51–54). Here we found that three of the four CEY proteins screened regulate memory, with each CEY controlling its own molecularly distinct form of memory (**Figure 1D**); *cey-1* is required for STM and ITM, *cey-2* is required for ITM, *cey-3* is required for LTM, and *cey-4* knockdown had no detectable effect on behavior. To our knowledge, this is the first documented behavioral phenotype regulated by CEY RBPs.

As there are no reported neuronal phenotypes associated with the CEY RBPs, we further examined their expression pattern based on previous transcriptomic profiling datasets. Single-cell RNA-seq data generated from neurons isolated from L4 larval animals revealed broad expression of *cey-1* and *cey-4,* while *cey-2* is only detected in eight neurons (**Supplemental Figure 3**)(55,56). *cey-3* is not detected in any neurons at L4 but is detected in adult neuron-specific datasets (24,40), highlighting the importance of compiling multiple transcriptomic datasets to choose our screen candidates.



C. elegans Gene	Human Ortholog	STM	ITM	LTM
cey-1	**YBX1-3**	↓	↓	—
cey-2	**YBX1-3**	—	↓	—
cey-3	**YBX1-3**	—	—	↓
cey-4	YBX1-3	—	—	—
puf-3	PUM1/2	—	↓	—
puf-5	PUM1/2	—	↓	↓
puf-7	PUM1/2	—	↑	—
puf-8	PUM2	—	↑	—
gcn-2	eIF2AK4	↓	—	—
ife-1	eIF4E	—	↓	—
ife-3	eIF4E	↓	—	—
drr-2	eIF4H	—	↓	—
rbm-3.1	**CIRP**	↓	—	—
rbm-3.2	CIRP	—	—	—
rpl-28	**RPL28**	↓	—	↓
rps-16	**RPS16**	↓	↓	↓
eef-1B.2	**EEF1B.2**	↑	—	—
vig-1	SERBP1	—	—	—
rnp-5	**RNPS1**	↓	↓	—
cpf-1	CSTF-1	—	—	—

Figure 1: Unbiased screen reveals novel associative memory regulators. (**A**) Short term/intermediate term associative memory (STM/ITM) assay workflow using adult-only RNAi knockdown(29,30). Adult worms are starved for 1 hour, conditioned for 1 hour, and tested immediately for 1x learning (0 min) via chemotaxis assays, or transferred onto holding plates with OP50 for 0.5, 1, or 2 hours after conditioning, then tested for STM and ITM changes. Memory performance is calculated by comparing the trained and untrained chemotaxis indices. (**B**) The long-term associative memory (LTM) assay workflow is identical to that for STM/ITM except worms have an additional *egl-30(js126)* mutation that allows for LTM formation after just 1x training (44). Worms are tested immediately for 1x learning (0 min) via chemotaxis assays or transferred onto holding plates with RNAi for 16-24 hours after conditioning, then tested for LTM changes. (**C**) Results of unbiased screen of 20 RBPs reveals eight novel associative memory regulators. All RBPs screened are shown and include the *C. elegans* gene, its mammalian ortholog, and any STM, ITM, or LTM changes. Blue arrows indicate enhanced memory while red arrows indicate decreased memory; lines indicate memory did not change. Novel memory regulators not reported in the literature are bolded. For full screen results, see **Supplemental Figures 1-2.** (**D**) Each member of the CEY RBP family plays a novel, unique role in memory. Adult-only RNAi knockdown of *cey-1* in neuronally-sensitized worms causes STM and ITM deficits, *cey-2* knockdown causes ITM deficits, and *cey-3* knockdown causes LTM deficits. Box and whisker plot: the center line denotes the median value (50th percentile) while the box contains the 25th to 75th percentiles. Whiskers mark the 5th and 95th percentiles.

Similar to their *C. elegans* orthologs, the mammalian YBX proteins are relatively understudied with regards to behavioral phenotypes or their function in the adult nervous system, though YBX1 is reported to be involved in neural stem cell development (57). Similar to *cey-1*, *YBX1* is broadly expressed throughout the nervous systems of humans, adult macaques, and rats. In these mammals, *YBX1* is expressed in memory-regulating regions including the hippocampus, parahippocampal cortex, and dentate gyrus (58). *YBX3* is also broadly expressed, including in memory-regulating regions, while *YBX2* is thought to be restricted to non-neuronal tissues (59,60). Importantly, YBX1 is reported to bind to plasticity-associated mRNAs, including *GluR2* mRNA and *CaM1* mRNA, in an activity-dependent manner, making the YBXs promising candidates as novel essential memory molecules in mammals (61). Moreover, both YBX1 and YBX3 proteins are detected in mouse hippocampal post-synaptic densities, suggesting that they have the potential to regulate local translation in a brain region critical for memory. (62). In non-neuronal tissues, CEYs and YBX proteins appear functionally conserved; both worm and mammalian proteins regulate polysome formation (49,63,64), suggesting that we can study conserved functions of this class of RBP in the worm.

To strengthen the translatability of our findings by focusing on conserved functions, we decided to study the *C. elegans* ortholog(s) that are the most closely related to the YBX proteins (53,54,57,65,66). To this end, we performed phylogeny and homology analysis using MegaX and DIOPT (67,68) and found that *cey-1* is the primary *C. elegans YBX* ortholog, closer in similarity to the *YBXs* than other CEY family members (**Figure 2A, 2B**). For example, the linker regions (NC9 and CC13) and Cold-Shock Domain (CSD) of CEY-1 are extremely similar to those of the YBX proteins (approximately 97% similar) (**Figure 2C**) (69,70). At both the N- and C-termini of the CEY-1/YBX proteins are moderately conserved (approximately 60%) low complexity regions, otherwise known as intrinsically disordered regions, responsible for protein-protein binding and stress granule formation (64,71).

Based on gene expression data, the expression pattern of CEY-1 and the mammalian YBX proteins is similarly broad across the nervous system (55,58–60,72–74). Though previous studies using promoter-fusion approaches corroborate the transcriptomic evidence that *cey-1* is broadly expressed in the nervous system (49), we sought to verify that the protein is indeed expressed in neurons. We crossed the previously published *cey-1::GFP* promoter-reporter system with a marker of neuronal nuclei (*unc-11(prom8)::2xNLS::TagRFP*) and validated their reported neuronal expression (**Supplemental Figure 4**). We confirmed these findings by verifying the neuronal localization of endogenous CEY-1 by crossing animals expressing an endogenously mNeonGreen-tagged CEY-1 *(mNG::CEY-1)* with *unc-11(prom8)::2xNLS::TagRFP* expressing animals (**Figure 2D,E)**.

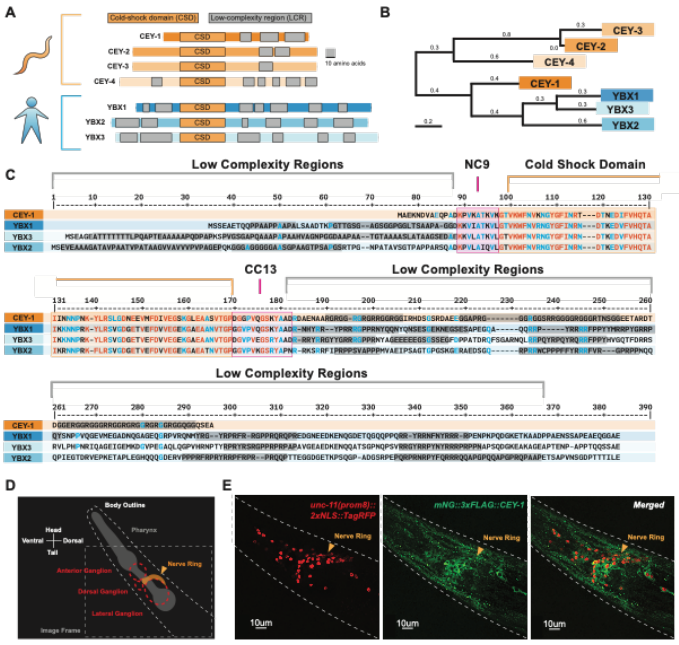


Figure 2: CEY-1 is the closest ortholog to the human Y-box binding proteins(**A**) Diagram of the conserved protein domains of the *C. elegans* CEY RBPs and human YBX RBPs. *C. elegans* have four CEY RNA-binding proteins, CEY-1, CEY-2, CEY-3, and CEY-4, that are orthologs of the three mammalian Y-box binding proteins, YBX1, YBX2, and YBX3. All CEY/YBX RNA-binding proteins have a conserved cold-shock domain (CSD) that mediates RNA binding as well as one or more low complexity regions (LCRs) thought to be important for protein-protein interactions. Grey scale bar is 10 amino acids. (**B**) Phylogenetic tree generated using MegaX[4,5] reveals that CEY-1 is more like the human YBX proteins than the other CEY RNA-binding proteins are. (**C**) Protein sequence alignment of the CEY and human YBX binding proteins made with MultiAlin[6] showing organization of protein domains. Highly conserved residues are depicted in red and blue. The cold-shock domains are highlighted in orange, the NC9 and CC13 linkers are highlighted in pink, and the low complexity regions are highlighted in grey. Domains were annotated according to SMART(128). (**D**) Diagram of microscopy images shown in (E). *C. elegans* head is labeled including the location of the pharynx and main neuronal ganglia/nerve rings. (**E**) CEY-1 is primarily localized to the cytoplasm at baseline conditions. Representative image of Day 2 adult worms with RFP-labeled neuronal nuclei *(unc-11(prom8)::2xNLS::TagRFP*) and an endogenously mNeonGreen-tagged CEY-1 *(mNG::CEY-1).*

CEY-1/YBX acts in neurons to promote memory

Despite its broad expression in the nervous system, we wanted to determine if *cey-1* was expressed in neurons involved in learning and memory. Because the *C. elegans* nervous system is invariant (White et al., 1986; Cook et al., 2019), the identities of specific neurons involved in a variety of associative behaviors are known. Based upon previously generated L4 single-cell RNA-seq data, we find that *cey-1* is indeed detected in all canonical memory-associated neurons, including those associated with gustatory plasticity, thermotaxis, and olfactory memory (**Figure 3A**) (44,75–81). To

confirm the finding from our initial screen that reduction of *cey-1* causes memory deficits, we examined the behavior of animals bearing two different loss-of-function *cey-1* alleles. Both animals exhibited severe deficits in learning and, as a result, inability to form a memory (**Figure 3B**).

To determine if *cey-1* acts in the nervous system to control learning and memory, we next tested if nervous-system specific *cey-1* knockdown causes memory deficits by performing adult-only, neuron-specific knockdown of *cey-1* using worms only expressing *sid-1* in the nervous system *(punc119::sid-1; sid-1(pk3321)).* Contrary to the mutants used in our initial screen, which can have off-target effects in other tissues, these worms lack *sid-1* in all tissues other than the nervous system, leading to RNAi very specifically in the nervous system (27). Our findings replicated the initial results of our screen: *cey-1* knockdown decreases both STM and ITM but not learning (**Figure 3C**). Importantly, we detected no thrashing, motility, or butanone chemotaxis deficits (**Supplemental Figure 5**), suggesting the deficits we observed are memory-specific. We believe that by knocking down *cey-1* only in neurons and during adulthood, we are circumventing loss of *cey-1* during development, which may be causing the more severe effects on learning seen in *cey-1* mutants.

Elevated neuronal CEY-1 is sufficient to enhance memory

We next asked if restoring *cey-1* function in nervous system was sufficient to rescue learning and memory deficits observed upon loss of *cey-1*. We generated *cey-1(rrr12)* loss of function animals that also express a single-copy knock-in transgene driving expression of *cey-1* specifically in the nervous system *(rab-3p::cey-1::rab-3 3'UTR; cey-1(rrr12))* (**Figure 3D**). We found that neuron-specific *cey-1* rescue is sufficient to restore both the learning and memory deficits seen in *cey-1(rrr12)* mutants (**Figure 3E**). Importantly, neither the *cey-1* loss of function mutants nor the neuronal *cey-1* rescue worms have butanone sensing deficits or impaired motility, suggesting that phenotypes recorded are memory-specific (**Supplemental Figure 6**).

Thus far, our data suggested that *cey-1* acts in the nervous system to promote memory. To test this hypothesis, we determined whether increasing *cey-1* expression in the nervous system would have beneficial effects on behavior. A single round of food-butanone training results in memory that persists for no longer than two hours in wild-type animals, and we asked if additional neuronal *cey-1* increased the duration of this memory, as has been observed with other memory promoting genetic manipulations (28,29,43). Worms with (*rab-3p::cey-1::rab-3 3'UTR; cey-1(+))* showed enhanced memory (**Figure 3F**), as their memory persists longer than the average two hours, consistent with previous memory enhancers (28,29,43). Collectively, our results show that *cey-1* is sufficient to increase memory when expressed only in the nervous system, suggesting that *cey-1* is a memory enhancer.

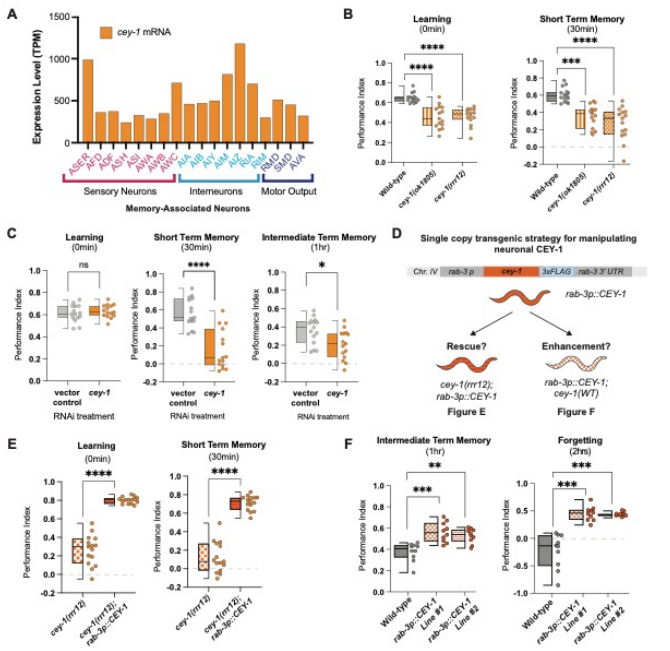


Figure 3: CEY-1 promotes memory in the *C. elegans* nervous system. (**A**) Single-cell RNA-seq data from L4 hermaphrodites show that *cey-1* is expressed in every canonical memory-associated neuron. (**B**) Loss-of-function mutants *cey-1(rrr12)* and *cey-1(ok1805)* have severe learning and memory deficits. Box and whisker plot as in Figure 1. n = 15 per genotype. ***p<0.001, ****p<0.0001. (**C**) Adult-only, neuron-specific knockdown of *cey-1* causes STM and ITM deficits. Box and whisker plot as in Figure 1. n = 15 per RNAi treatment. *p<0.05, ****p<0.0001. ns, not significant (p>0.05). (**D**) Diagram of strategy to determine whether neuronal *cey-1* is sufficient to rescue and enhance memory. Worms that express a single copy of *cey-1* only in the nervous system *(rab-3p::cey-1::rab-3 3'UTR; cey-1(rrr12))* were tested for phenotypic rescue in (E) while worms that have a single extra copy of *cey-1* only in the nervous system *(rab-3p::cey-1::rab-3 3'UTR; cey-1(+))* were tested in (F). (**E**) Single copy expression of *cey-1* in the nervous system *(rab-3p::cey-1::rab-3 3'UTR; cey-1(rrr12))* rescues loss-of-function learning and memory loss. Box and whisker plot as in Figure 1. n = 15 per genotype. ****p<0.0001. (**F**) Single copy overexpression of *cey-1* in the nervous system is sufficient to enhance memory *(rab-3p::cey-1::rab-3 3'UTR; cey-1(+)).* Box and whisker plot as in Figure 1. n = 15 per genotype. **p<0.01,***p<0.001.

Copy number variant deletions in *YBX1* and *YBX3* may be associated with neurological symptoms, including intellectual disability

Human genomic datasets are increasingly linking RBP dysfunction to neurodevelopmental and neurological disorders, including those characterized by intellectual disability and cognitive dysfunction (6–8,82). Therefore, we used copy number variation (CNV) datasets to determine if loss of the *YBXs* may potentially be associated with neurological dysfunction in humans, which would mirror our findings in *C. elegans*. We used an open-access database, DECIPHER, that reports CNVs from 50,000 individuals (83). We focused specifically on CNVs in *YBX1* and *YBX3*, as *YBX2* is not expressed in the adult human brain (58,60).

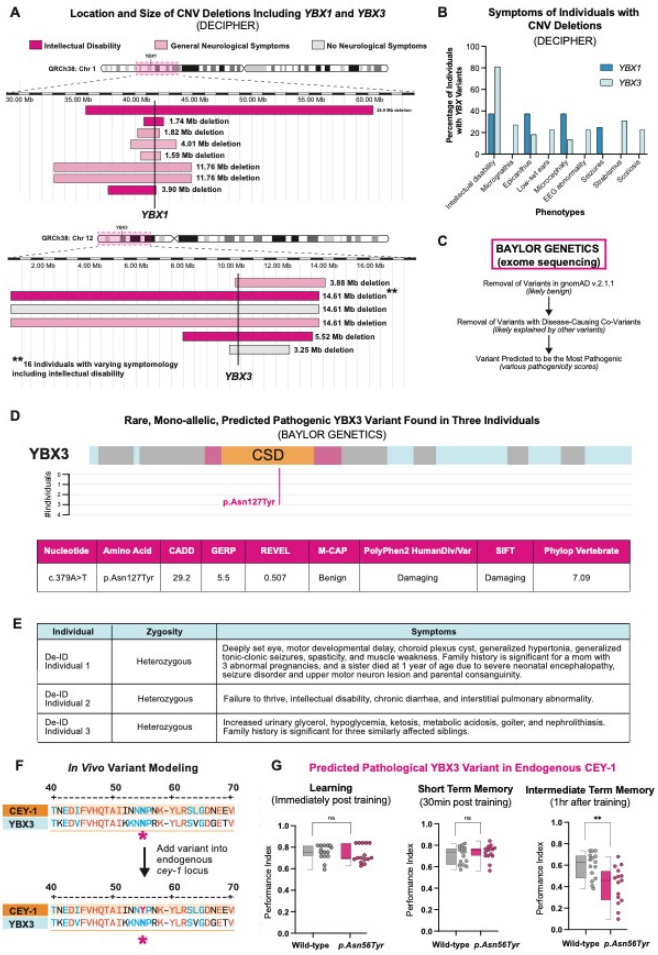
We found that many individuals with CNV deletions that include either *YBX1* or *YBX3* have neurological symptoms

(**Figure 4A**). The CNV deletions were variable in size, ranging from 1.59 Mb to 24.59 Mb. The most common symptom reported in individuals with CNV deletions including the *YBX*s is intellectual disability, as 37.5% of individuals with CNV deletions containing *YBX1* and 80% of individuals with CNV deletions in *YBX3* have intellectual disability (**Figure 4B**). Other common symptoms for individuals with *YBX1* CNV deletions include epicanthus, microcephaly, and seizures, while individuals with *YBX3* CNV deletions often exhibit strabismus and scoliosis. However, as is the case with the majority of CNVs, all CNV deletions contained other loci that may be contributing to these individuals' phenotypes. In the case of *YBX1*, the minimal overlap between CNV losses affecting the *YBX1* loci is 1.08 Mb containing 34 genes, including two loci that are predicted to be loss of function intolerant: forkhead box J3 (*FOXJ3*)(84) and solute carrier family 2 member 1 (*SLC2A1*) (85). *SLC2A1* is associated with autosomal dominant Dystonia 9 (86,87), GLUT1 deficiency syndrome 1 (88,89), GLUT1 deficiency syndrome 2 (90,91), Stomatin-deficient cryohydrocytosis with neurologic defects (92,93), and susceptibility to idiopathic generalized epilepsy 12 (94,95). In the case of *YBX3*, the minimal overlap between CNV deletions affecting the *YBX3* loci is 3.24 Mb containing 68 genes, including three loci that are predicted to be loss of function intolerant: ETS variant transcription factor 6 (*ETV6*)(96), low density lipoprotein receptor-related protein 6 (*LRP6*)(97), and dual-specificity phosphatase 16 (*DUSP16*)(98). However, unlike the loci surrounding *YBX1*, none of the loss of function intolerant loci around *YBX3* are known to cause monogenetic neurological disease. Summary statistics for DECIPHER data, including all individuals with CNV duplications and deletions as well as the DECIPHER phenotypic data for CNV deletions in *YBX1* and *YBX3* can be found in **Supplementary Figure 7** and **Supplementary Figure 8,** respectively.

Taken together, these data suggest that loss of function CNVs including the *YBX1* and *YBX3* loci may be associated with neurological dysfunction; however, the confounding variable of losing other loci prevents any direct association. As a result, we decided to model a rare single nucleotide variant in *YBX3* for biological function *in vivo*.

Introducing a predicted deleterious variant in *YBX3* into the endogenous *cey-1* locus causes memory deficits *in vivo*

While publicly available CNV datasets suggest that single-copy loss of *YBXs* may potentially contribute to neurological symptoms, the minimal overlap regions for CNV deletions containing *YBX1* and *YBX3* include loci beyond the gene of interest that are associated with monogenic neurological disorders in humans. Therefore, single-nucleotide variants (SNVs) in the genes of interest may better link deleterious variants in single genes directly to neurological disorders (10,12,99–101). For example, rare missense variants in PolyQ binding protein 1 (*PQBP1*) cause Renpenning



Nucleotide	Amino Acid	CADD	GERP	REVEL	M-CAP	PolyPhen2 HumanDiv/Var	SIFT	Phylop Vertebrate
c.379A>T	p.Asn127Tyr	29.2	5.5	0.507	Benign	Damaging	Damaging	7.09

Individual	Zygosity	Symptoms
De-ID Individual 1	Heterozygous	Deeply set eye, motor developmental delay, choroid plexus cyst, generalized hypertonia, generalized tonic-clonic seizures, spasticity, and muscle weakness. Family history is significant for a mom with 3 abnormal pregnancies, and a sister died at 1 year of age due to severe neonatal encephalopathy, seizure disorder and upper motor neuron lesion and parental consanguinity.
De-ID Individual 2	Heterozygous	Failure to thrive, intellectual disability, chronic diarrhea, and interstitial pulmonary abnormality.
De-ID Individual 3	Heterozygous	Increased urinary glycerol, hypoglycemia, ketosis, metabolic acidosis, goiter, and nephrolithiasis. Family history is significant for three similarly affected siblings.

Figure 4: Copy number variants in *YBX1* and *YBX3* found in individuals with neurological symptoms and *in vivo* modeling of a predicted pathogenic single nucleotide variant in YBX3. (**A**) Location, size, and neurological symptoms associated with copy number variant (CNV) deletions reported in DECIPHER. The exact location of *YBX1* and *YBX3* are shown with a bold vertical line. The area around the gene affected by the CNV deletions is highlighted with a pink box on their respective chromosome ideograms at the top of the figure. CNVs where the individual had neurological symptoms are in light pink, and CNVs where the individual had associated intellectual disability are in dark pink. (**B**) Intellectual disability is the most often associated symptom of patients with CNV deletions of *YBX1* or *YBX3* based on DECIPHER. (**C**) Workflow to examine single nucleotide variants (SNVs) that may be disease-linked using the postnatal exome sequencing database from Baylor Genetics. Several strategies were used to remove likely non-pathogenic variants, including removing cases with known co-variants and removal of variants found in the general population (gnomAD v2.1.1)(123). (**D**) Location and pathogenicity score of rare a heterozygous SNV in *YBX3* found in four individuals with indicated neurological symptoms. The variant we identified, p.Asn127Tyr, had the highest predicted pathogenicity using seven different metrics (see text). (**E**) Table of p.Asn127Tyr cases from Baylor Genetics de-identified exome sequencing dataset. Two of the three individuals have neurological symptomology reported, such as intellectual disability and seizures. (**F**) Location of p.Asn127Tyr variant shows highly conserved residues between CEY-1 and YBX3 and location of variant when introduced into the endogenous CEY-1 locus. (**G**) Introduction of p.Asn127Tyr into the endogenous *cey-1* locus causes ITM deficits. Box and whisker plot as in Figure 1. n = 15 per genotype. **p<0.01, ns, not significant (p>0.05).

syndrome, hallmarked by intellectual disability, microcephaly, short stature, and hypogonadism (102).

We received a deidentified dataset of *YBX* variants identified through exome and genome sequencing completed at Baylor Genetics (103) to identify potentially deleterious SNVs in the *YBXs*. To identify variants of interest, we curated rare, heterozygous variants in postnatal individuals, as these are the most likely to be deleterious (104,105). We first removed SNVs that were most likely not contributing to early developmental brain disorders by eliminating variants found in gnomAD v2.1.1 as well as cases where the subject's symptomology was explained by co-variants in the genome (**Figure 4C**). We then used seven different pathogenicity metrics to determine which variants were the most likely to be deleterious, including CADD (106,107), GERP (108), REVEL (109), M-CAP (110), PolyPhen 2 (111), SIFT (112), and Phylop Vertebrate (113). The variant with the highest overall pathogenicity scores was in *YBX3*: c.379A>T resulting in the amino acid change p.Asn127Tyr. The affected amino acid is located in the Cold-Shock Domain of the YBX3 protein and, significantly, this variant is not reported in gnomAD v.2.1.1 nor gnomAD v4, suggesting it is rare (**Figure 4D**). We found this variant in three different individuals, two of which present neurological symptoms such as intellectual disability and seizures (**Figure 4E**). The third individual without reported neurologic features may have been related to one of the other individuals as they shared the same last name, but their exact relation is unknown. If these individuals are indeed related, it may suggest variable expressivity of neurological symptomology.

The p.Asn127 residue is highly conserved across species, including between *C. elegans* CEY-1 and human YBX3 (**Figure 4F**). Therefore, we decided to test the potential significance of this variant *in vivo* in *C. elegans* by introducing it into the endogenous *cey-1* locus and examining its effect on memory. We used CRISPR to introduce the same mutation into the endogenous *cey-1* gene at the same conserved residue as in YBX3: p.Asn56Tyr (*cey-1 p.Asn56Tyr (p.Asn56YTyr)).* We then tested for memory deficits at learning, STM, and ITM. Remarkably, introduction of the p.Asn56Tyr variant into *cey-1* caused ITM deficits, suggesting the variant may have consequences on human neurological function (**Figure 4G**). Importantly, CEY-1 p.Asn56Tyr worms had no butanone sensing deficits or motility impairments (**Supplemental Figure 9**), suggesting that in this case, neurological dysfunction is memory-specific. By using the worm to model one predicted damaging variant in *YBX3*, our findings suggest that *YBX3* may be of interest for future exploration of human disease associations.

Discussion

C. elegans as a discovery platform for novel associative memory regulators

We performed an unbiased screen of 20 synaptically-enriched RBPs examining their roles in learning and memory and identified eight novel associative memory regulators, each controlling memory differently. While we focused on the CEY RBP family in this work, five other RBPs remain to be investigated. We found that *rbm-3.1* is required for STM, *rnp-5* is required for STM and ITM, *rpl-28* is required for STM and LTM, *rps-16* is required for STM, ITM, and LTM, and *eef-1B.2* suppresses STM. *rbm-3.1,* ortholog of CIRP, is reported to mediate alcohol-induced spatial memory deficits in mice (114) but has never been associated with memory independently of alcohol use prior to this study. In addition, *eef-1B.2* and *rnp-5* are associated with human neurological disease. EEF1B.2 causes a novel form of intellectual disability (115), while *rnp-5* ortholog RNPS1 is a risk factor for both intellectual disability and autism (116,117). To our knowledge, neither *rps-16* nor *rpl-28* have been studied in the nervous system of any organism.

Here, we demonstrated novel roles for eight RBPs in memory using *C. elegans*. Our results suggest that these RBPs play a conserved role in memory that remains understudied, and future studies should combine worm behavior with human genetic data to examine these and other uncharacterized genes. Taken together, our study highlights the utility of *C. elegans* as a high-throughput screening tool for novel associative memory regulators, including those that control different, molecularly distinct forms of memory.

Three of the four CEY RBPs control memory

We identified three members of the CEY RBP family, orthologs of the mammalian YBX RBPs, as novel memory molecules. We found that each CEY RBP has its own distinct role in memory: *cey-1* regulates STM/ITM, *cey-2* regulates ITM, and *cey-3* regulates LTM. Of note, *cey-4* does not appear to be required for memory ability but is expressed throughout the adult nervous system (55,73,74), and therefore likely contributes to other neuronal functions. The unique role of each CEY in memory is not explained by their expression patterns, as *cey-1* and *cey-4* are both broadly expressed in the nervous system while *cey-2* and *cey-3* are found in very few neurons (55,73,74). While evidence from the germline suggests the CEY RBPs are promiscuous in binding and are not affiliated with specific mRNA subsets (49), it is unclear whether there is more mRNA target specificity in the nervous system. Future studies examining how each CEY RBP controls the translational landscape of neurons are warranted to determine the independent mechanisms by which the CEYs proteins regulate memory.

CEY-1 is a novel neuronal memory promoting RBP

Our study of CEY-1, the closest *C. elegans* ortholog to the mammalian YBX proteins, revealed that it acts in the nervous system to promote memory, as neuron-specific rescue of CEY-1 ameliorates behavioral defects cause by the loss of *cey-1,* and an additional copy of CEY-1 only in the nervous system improves memory performance. Future investigations are needed to unveil the molecular mechanisms

by which CEY-1 regulates memory, but based on known CEY/YBX biology, we hypothesize that CEY-1 controls transcription and/or translation of mRNAs involved in memory.

YBX1 is reported to bind to plasticity-associated mRNAs in an activity-dependent manner *in vitro* (61), and both CEY-1/YBX proteins regulate polysome formation and translation (49,63,64). Furthermore, in non-neuronal tissues, YBX1 is reported to interact with known regulators of plasticity-dependent translation, specifically mTOR, AKT and translation initiation machinery (53,63,118); therefore, it is likely that CEY-1/YBX is involved in these processes as well. Unlike other RBPs linked to plasticity-dependent translation, such as FMRP (119), YBX proteins are relatively non-specific with regards to target mRNAs (64,120). It is unclear if a subset of molecules required for memory are regulated by YBX, or if it is involved in controlling broad machinery involved in plasticity and memory such as general protein synthesis. Future work will be necessary to determine the exact mechanisms by which CEY-1 promotes memory.

Single nucleotide variant in CEY-1 decreases intermediate term memory in *C. elegans*

Given the lack of experimental evidence linking the mammalian YBX proteins to memory, we delved into human variant datasets to explore potential associations between YBX dysfunction and neurological symptoms. We focused on *YBX1* and *YBX3* as they are expressed in the adult nervous system. We found that several CNVs involving the *YBX1* and *YBX3* loci, as well as one rare SNV in *YBX3*, were identified in individuals with neurological symptoms including intellectual disability. However, the CNV analysis is limited by other disease-associated loci of the minimal overlap regions involving *YBX1* and *YBX3* and the SNV analysis is limited by the sample size and available de-identified clinical findings.

To assess the biological significance of rare, mono-allelic SNVs in either *YBX1* or *YBX3*, we conducted *in vivo* modeling of the variant with the highest pathogenicity scores. The YBX3 p.Asn127Tyr variant was identified in three individuals, two of whom exhibit neurological symptoms. The p.Asn127Tyr variant is located within the YBX3 CSD at a highly conserved site (69,70,120). Introduction of the corresponding p.Asn56Tyr variant into the endogenous *cey-1* gene caused ITM deficits, suggesting that this SNV in *YBX3* may be of interest for future studies exploring the role of *YBX3* in human disease.

We believe that the observed ITM deficits may stem from dysregulation of mRNA translation by CEY-1. The CSD, responsible for RNA binding, often lacks a highly specific affinity for a particular RNA motif (70,120). As a result, proteins with CSDs are often considered master regulators of the mRNA landscape. It is likely that introduction of a variant into the CSD disrupts broad RNA regulation by CEY-1, including memory-regulating mRNAs. Regulation of target mRNAs by CEY-1 may be at the level of transcription, translation, or another mechanism, so it is unclear how exactly this variant affects CEY function. Though the mechanism is currently unknown, our data indicate that rare *YBX3* variants may be biologically significant and potentially affect neurologic function.

Foundations for mammalian *in vivo* studies of YBX proteins in plasticity and neurological disease

Studying the function of specific YBX proteins in mammalian models is challenging due to the potential for YBXs to genetically compensate for one another and due to the low survival rate of double knock-outs (57). Here, we report that either depleting CEY-1/YBX levels specifically in adulthood or introducing a point variant detected in three human individuals disrupts memory. Each of these manipulations is less likely to be lethal or induce genetic compensation than loss during development. Future studies employing similar approaches in mammalian models will be extremely valuable – conditional *YBX1* or *YBX3* deletion in adult animals is likely to provide insight into how they control memory and plasticity, while introducing rare human variants will potentially allow for more detailed modeling of human symptomology than is possible in the worm. Moreover, future studies will need to be considerate of loss of function mechanisms of variants. YBX1 is predicted to be loss of function intolerant (pLI = 1, gnomAD v4.0), suggesting haploinsufficient variants may be deleterious. Conversely, YBX3 is predicted to be loss of function tolerant (pLI = 0, gnomAD v4.0), suggesting haplosufficiency, and rare variants in YBX3 may be dominant negative or result in gain of function. Identification of more rare variants in the YBXs using comprehensive exome human datasets and modeling variants *in vivo* in mammals will enhance our understanding of mechanisms by which YBX dysfunction may play a role in human neurological disease.

Taken together, our results uncover the CEY-1/YBX RBPs as novel associative memory regulators and potentially new contributors to rare neurological disease. Our current findings are limited to the functional testing of a single variant in *YBX3* but studying this class of proteins further will provide new insights into the molecular mechanisms of memory and deepen our understanding of how disruption of RBPs contributes to disorders of the nervous system.

Methods

***C. elegans* maintenance:**

All strains were maintained at 20°C on 10cm plates made from standard nematode growth medium (NGM: 3g/L NaCl, 2.5g/L of Bacto-peptone, 17g/L Bacto-agar in milliQ water) or high nematode growth medium (HGM: 3g/L NaCl, 20g/L of Bacto-peptone, 30g/L Bacto-agar in milliQ water). After autoclaving and allowing molten agar to cool slightly, we added 1mL/L cholesterol (5mg/mL in ethanol), 1mL/L 1M $CaCl_2$, 1mL/L 1M

$MgSO_4$, and 25mL/L 1M potassium phosphate buffer (pH 6.0) (121). Experiments were performed using NGM plates seeded with OP50 *E. coli* as the food source for *ad libitum* feeding (121).

Hypochlorite population synchronization was performed by collecting eggs from gravid hermaphrodites via exposure to an alkaline-bleach solution (85mL water, 15mL sodium hypochlorite, 5mL 5M NaOH), followed by repeated washing of collected eggs in 1mL of M9 buffer (6g/L Na_2HPO_4, 3g/L KH_2PO_4, 5g/L NaCl and 1mL/L 1M $MgSO_4$ in milliQ water (121)).

Strains:
Wild-type: (N2 Bristol)
Mutants: NM1380(*egl-30(js126)),* LC108(*uIs69[punc-119::sid-1; pmyo-2::mCherry]),* TU3401*(usIS69[myo-2p::mCherry + unc119p::sid-1]; sid-1(pk3321))*, RAF5(*cey-1(rrr12)),* VC1310*(cey-1(ok1805)),* GLW51*(cey-1(utx43[mNG::3xFLAG::cey-1]) II),* OH13605*(otIs619[unc-11(prom8)::2xNLS::TagRFP] X)* were obtained from the Caenorhabditis Genetics Center (University of Minnesota, Minneapolis, MN).

Nuclear GFP reporter for *cey-1* #1087*(cey-1p::PEST::GFPH2B::cey-1u; unc-119 was* described previously (49) and generously gifted by Rafal Ciosk.

Strains made in collaboration with InVivo Biosystems include: COP2655*(knuSi967[(rab-3p::CEY-1::3xFLAG::rab-3u, unc-119(+))] IV; unc-119(ed3) III)* and COP2656*(knuSi968[(rab-3p::CEY-1::3xFLAG::rab-3u, unc-119(+))] IV; unc-119(ed3) III)* are two lines with the same genotype, COP2689(*cey-1(knu1233 [N56Y])).*

The following strains were generated by crosses: RNA1*(egl-30(js126); uIs69[(pCFJ90)myo-2p::mCherry + unc-119p::sid-1])* was made by crossing NM1380*(egl-30(js126))* with LC108*(uIs69[(pCFJ90)myo-2p::mCherry + unc-119p::sid-1]).* RNA20(*cey-1(utx43[mNG::3xFLAG::cey-1]) II; otIs619[unc-11(prom8)::2xNLS::TagRFP] X)* was made by crossing OH13605*(otIs619[unc-11(prom8)::2xNLS::TagRFP] X)* with GLW51*(cey-1(utx43[mNG::3xFLAG::cey-1]) II).* RNA26(*cey-1p::PEST::GFPH2B::cey-1u; unc-119(+); otIs619[unc-11(prom8)::2xNLS::TagRFP] X)* was made by crossing OH13605*(otIs619[unc-11(prom8)::2xNLS::TagRFP] X)* with *cey-1* #1087*(cey-1p::PEST::GFPH2B::cey-1u; unc-119(+)).* RNA40*(cey-1(rrr12); knuSi968[(rab-3p::CEY-1::3xFLAG::rab-3u, unc-119(+)] IV; unc-119(ed3) III)* was made by crossing RAF5(*cey-1(rrr12))* with COP2656*(knuSi968[(rab-3p::CEY-1::3xFLAG::rab-3u, unc-119(+))] IV; unc-119(ed3) III).*

RNAi treatment:
For adult-only RNAi, worms that are neuronally sensitive to RNAi, RNA1*(egl-30(js126); uIs69[(pCFJ90)myo-2p::mCherry + unc-119p::sid-1])),* or allow RNAi only in the nervous system, TU3401*(usIS69[myo-2p::mCherry + unc119p::sid-1]; sid-1(pk3321)),* were synchronized by bleaching. Then, at L4, worms were transferred from plates seeded with OP50 to plates seeded with RNAi (*HT115 E. coli)* for *ad libitum* feeding and allowed to grow for two days. For all RNAi experiments, standard NGM molten agar was supplemented with 1mL/L IPTG (isopropyl β-d-1-thiogalactopyranoside) and 1mL/L 100mg/mL carbenicillin.

Short-term and intermediate-term associative memory assays:
Wild-type, mutant, or RNAi-treated worms were trained and tested for short- and intermediate-term memory changes as previously described (28). Briefly, synchronized Day 2 adult worms were washed off plates with M9 buffer. Worms were then allowed to settle by gravity and washed twice more with M9 buffer to remove any bacteria. After washing, the worms were starved for one hour in M9 buffer. For 1x food-butanone pairing, hereby called conditioning, starved worms were transferred to 10cm NGM conditioning plates seeded with OP50 *E. coli* bacteria and with a total of 16μL of 10% butanone (Sigma Aldrich) diluted in ethanol streaked on the lid in a '#' shape for one hour. After conditioning, the trained population of worms were tested for chemotaxis to 10% butanone and to an ethanol control using standard, previously described chemotaxis conditions (76). Different stages of memory were tested by measuring chemotaxis of different subpopulations of worms at different timepoints at molecularly distinct stages of memory (29,30). These stages are immediately after training (0min, learning) or after being transferred to 10cm NGM plates with fresh OP50 for 30 minutes (short-term associative memory), 1 hour (intermediate term associative memory), or 2 hours (forgetting).
Chemotaxis indices for each timepoint were calculated as
Chemotaxis index =
$(\#worms_{butanone} - \#worms_{ethanol})/(total\#worms).$
Performance index is the change in the chemotaxis index after training relative to the untrained chemotaxis index, or
Performance index =
$Chemotaxis\ index_{trained} - Chemotaxis\ index_{untrained}$

Long-term associative memory assays:
As previously published, *egl-30(js126)* animals form a long-term memory after just one round of training (44). For LTM screening, we administered RNAi to adult, neuronally RNAi sensitive *egl-30(js126);punc119::sid-1* worms, performed S/ITM training (1 CS-US pairing), and measured learning immediately after training and long term memory 16-20 hours post-training. Chemotaxis index and performance index were calculated in the same manner as short term and intermediate term associative memory assays.

Baseline chemosensation assays:
Chemotaxis, or chemosensation experiments, were performed based on previously published assays (76). In brief, assays were

performed on unseeded 10cm NGMs. Two marks were made on the back of the plate on opposite sides of the plate, approximately 0.5cm from the edge. 1μL of sodium azide (Thermo Fisher) was placed on both spots and allowed to dry before adding 1μL of test odorant diluted in ethanol on one side and ethanol on the other. Odorants included 0.1% and 10% butanone (vol/vol), 0.1% nonanol (2-nonanone)(vol/vol), 1% isoamyl alcohol (vol/vol), 1% benzaldehyde (vol/vol), 10% pyrazine (weight/vol), and 1% diacetyl (vol/vol) (all from Sigma Aldrich). Worms were washed off their plates and subsequently washed three times with M9 buffer, then placed near the bottom center of the plate, equidistant between the two marks, and allowed to chemotax for an hour. Chemotaxis indices for each timepoint were calculated as

Chemotaxis index =

$(\#worms_{butanone} - \#worms_{ethanol})/(total\#worms).$

Thrashing assays:

Thrashing assays were performed as described previously (122). Briefly, an unseeded 35mm NGM agarose plate was filled with 1mL of M9 buffer solution. Next, 10-15 well-fed adult worms were transferred to an unseeded NGM agarose plate to remove bacteria. Worms were placed into M9 for 1 minute to acclimate, then thrashing behavior was recorded in 30 second increments per group of worms using a Nikon DS-Fi3 camera and NIS Elements. Thrashing was defined as two body bends of the worm: swinging its head and tail to the same side to form a C shape and then back to the initial position.

Motility assays:

We measured motility as previously published (20) by finding the average percentage of worms remaining at the origin of the plate after a naïve (untrained) chemotaxis assay using an attractive concentration of butanone (0.1% vol/vol).

Confocal microscopy:

Worms were paralyzed with fresh 4% levamisole diluted in M9 buffer and imaged for up to 30 minutes once put onto the slide to preserve protein dynamics. In all experiments, imaging of Day 2 adult worms was performed on a Nikon Ti2E inverted microscope system with a W1 spinning disk confocal unit at 100x magnification. We used an excitation wavelength of 488nm for GFP and 561nm for RFP, used a pixel size of 0.11um/px, and z stacks with a z-step of 0.6nm. Images were processed in Nikon NIS Elements software. The same settings for laser power and detector gain were used for all genotypes.

Variant identification and pathogenicity analysis:

We received a deidentified dataset of human *YBX* gene variants identified through clinical exome and genome sequencing completed at Baylor Genetics (103) and removed any variants where symptomology listed could be explained by other genetic alterations found in the individual's genome. Next, we removed any variants found in gnomAD v.2.1.1, even if only one case was reported (123), as these variants would not likely cause severe disease if found in the general population. We then selected variants for potential *in vivo* modeling based on the presence of neurological symptoms, such as seizures, microcephaly, autism, and intellectual disability. Variants of interest were then assessed using multiple pathogenicity metrics to determine the final variant for *in vivo* testing: specifically, M-CAP (110), SIFT (112), PolyPhen 2 (111), GERP (108), CADD (106,107), REVEL (109), and Phylop Vertebrate (113). All metrics were found and visualized using the UCSC genome browser (124). We selected the variant with the highest overall predicted pathogenicity across all score metrics as opposed to any specific numerical cutoffs.

Phylogenetic analysis:

To analyze the evolutionary history and determine the similarity of RNA-binding proteins between phyla, we used maximum likelihood (ML) analysis with the JTT matrix-based model (125). We obtained protein sequences of the *C. elegans* CEY proteins as well as the human YBX proteins from UniProt (126) then used these sequences to generate a tree with the highest log likelihood score (-4517.57) using the Neighbor-Join and BioNJ algorithms to calculate pairwise distances from the JTT model. The tree was drawn to scale, with branch lengths measured in the number of substitutions per site and a 0.20 scale bar indicating the relative distance between nodes. All analyses were performed using MEGA X (67,127).

Orthology analysis:

Drosophila RNAi Screening Center Integration Ortholog Prediction Tool (DIOPT) is an online bioinformatics resource designed to facilitate the identification of orthologous genes across various species (68).

Statistical analysis

Statistical data is reported in the main text, figures, and tables as noted. Significance threshold of $p < 0.05$ was used. The symbols *, **, ***, and **** refer to $p < 0.05$, 0.01, 0.001, and 0.0001, respectively. For the comparison of performance indices between two behavior conditions (eg. vector control vs *cey-1* RNAi), a Mann-Whitney test comparing ranks was used because it does not assume normality. For comparison of performance indices between three or more groups (eg wild-type vs two different *cey-1* loss-of-function mutants), one-way analysis of variances followed by Bonferroni post hoc tests for multiple comparisons were performed. All experiments were repeated on separate days with separate populations to confirm reproducibility of results. Sample size n represents the number of chemotaxis assays performed for behavior, with each assay containing approximately 50-150 worms each.

Statistical analysis software
All statistics and code were run in GraphPad Prism 10, using standard toolboxes.

Acknowledgments: We thank the Ciosk lab for generously sharing the cey-1 #1087 strain; the CGC for strains; Ben Jussila and In Vivo Biosystems for strains and helpful discussion; Peter Boag for helpful discussion; and members of the Arey lab, particularly Katie L Brandel and Catherine Stuart for their detailed feedback on the manuscript. This study makes use of data generated by the DECIPHER community. A full list of centres who contributed to the generation of the data is available from https://deciphergenomics.org/about/stats and via email from contact@deciphergenomics.org. DECIPHER is hosted by EMBL-EBI and funding for the DECIPHER project was provided by the Wellcome Trust (WT223718/Z/21/Z).

Funding: ANH funded by a NINDS NRSA (F31NS129312) and Baylor Research Advocates for Student Scientists (BRASS). RNA is supported by a Glenn Foundation for Medical Research and AFAR Grant for Junior Faculty, the Whitehall Foundation, and an NIH Director's New Innovator Award (DP2NS132372). HTC's research effort is partially supported by The Robert and Janice McNair Foundation, Burroughs Wellcome Fund, Cain Pediatric Neurology Research Foundation Laboratories, Anne and Bob Graham, and the Child Neurology Foundation and Society. KBA is supported by a NSF GRFP, EWP is supported by a NIGMS NRSA, EJL is funded by a NIA NRSA. ESG is funded by BCM PREP. Some strains were provided by the *C. elegans* Genetics Center (CGC), which is funded by the NIH Office of Research Infrastructure Programs (P40 OD010440).

Author contributions: ANH and RNA conceptualized and designed experiments. ANH, RNA, JAR, and HTC analyzed human genetics datasets. ANH, KLB, PRM, PV, EJL, EWP, ESG, and CN performed experiments. ANH analyzed the data. ANH, RNA, JAR, and HTC wrote the manuscript. All authors edited the manuscript.

Competing interests: Authors declare no competing interests.

References

1. Hafner A-S, Donlin-Asp PG, Leitch B, Herzog E, Schuman EM. Local protein synthesis is a ubiquitous feature of neuronal pre- and postsynaptic compartments. Science. 2019 May 17;364(6441).

2. Costa-Mattioli M, Sossin WS, Klann E, Sonenberg N. Translational control of long-lasting synaptic plasticity and memory. Neuron. 2009 Jan 15;61(1):10–26.

3. Martin KC, Casadio A, Zhu H, Yaping E, Rose JC, Chen M, et al. Synapse-specific, long-term facilitation of aplysia sensory to motor synapses: a function for local protein synthesis in memory storage. Cell. 1997 Dec 26;91(7):927–38.

4. Akins MR, Berk-Rauch HE, Fallon JR. Presynaptic translation: stepping out of the postsynaptic shadow. Front Neural Circuits. 2009 Nov 4;3:17.

5. Schieweck R, Ninkovic J, Kiebler MA. RNA-binding proteins balance brain function in health and disease. Physiol Rev. 2021 Jul 1;101(3):1309–70.

6. Doxakis E. RNA binding proteins: a common denominator of neuronal function and dysfunction. Neurosci Bull. 2014 Aug;30(4):610–26.

7. Gebauer F, Schwarzl T, Valcárcel J, Hentze MW. RNA-binding proteins in human genetic disease. Nat Rev Genet. 2021 Mar;22(3):185–98.

8. Prashad S, Gopal PP. RNA-binding proteins in neurological development and disease. RNA Biol. 2021 Jul;18(7):972–87.

9. Lauriat TL, Shiue L, Haroutunian V, Verbitsky M, Ares M, Ospina L, et al. Developmental expression profile of quaking, a candidate gene for schizophrenia, and its target genes in human prefrontal cortex and hippocampus shows regional specificity. J Neurosci Res. 2008 Mar;86(4):785–96.

10. Au PYB, Goedhart C, Ferguson M, Breckpot J, Devriendt K, Wierenga K, et al. Phenotypic spectrum of Au-Kline syndrome: a report of six new cases and review of the literature. Eur J Hum Genet. 2018 Sep;26(9):1272–81.

11. Dingemans AJM, Truijen KMG, Kim J-H, Alaçam Z, Faivre L, Collins KM, et al. Establishing the phenotypic spectrum of ZTTK syndrome by analysis of 52 individuals with variants in SON. Eur J Hum Genet. 2022 Mar;30(3):271–81.

12. Quartier A, Poquet H, Gilbert-Dussardier B, Rossi M, Casteleyn A-S, Portes V des, et al. Intragenic FMR1 disease-causing variants: a significant mutational mechanism leading to Fragile-X syndrome. Eur J Hum Genet. 2017 Apr;25(4):423–31.

13. Gennarino VA, Palmer EE, McDonell LM, Wang L, Adamski CJ, Koire A, et al. A Mild PUM1 Mutation Is Associated with Adult-Onset Ataxia, whereas Haploinsufficiency Causes Developmental Delay and Seizures. Cell. 2018 Feb 22;172(5):924–936.e11.

14. Hentze MW, Castello A, Schwarzl T, Preiss T. A brave new world of RNA-binding proteins. Nat Rev Mol Cell Biol. 2018 May;19(5):327–41.

15. Olshansky SJ, Goldman DP, Zheng Y, Rowe JW. Aging in America in the twenty-first century: demographic forecasts from the MacArthur Foundation Research Network on an Aging Society. Milbank Q. 2009 Dec;87(4):842–62.

16. Kim KW, Tang NH, Piggott CA, Andrusiak MG, Park S, Zhu M, et al. Expanded genetic screening in Caenorhabditis elegans identifies new regulators and an inhibitory role for NAD+ in axon regeneration. Elife. 2018 Nov 21;7.

17. Poole RJ, Bashllari E, Cochella L, Flowers EB, Hobert O. A Genome-Wide RNAi Screen for Factors Involved in Neuronal Specification in Caenorhabditis elegans. PLoS Genet. 2011 Jun 16;7(6):e1002109.

18. Doitsidou M, Flames N, Topalidou I, Abe N, Felton T, Remesal L, et al. A combinatorial regulatory signature controls terminal differentiation of the dopaminergic nervous system in C. elegans. Genes Dev. 2013 Jun 15;27(12):1391–405.

19. McDiarmid TA, Belmadani M, Liang J, Meili F, Mathews EA, Mullen GP, et al. Systematic phenomics analysis of autism-associated genes reveals parallel networks underlying reversible impairments in habituation. Proc Natl Acad Sci USA. 2020 Jan 7;117(1):656–67.

20. Lakhina V, Arey RN, Kaletsky R, Kauffman A, Stein G, Keyes W, et al. Genome-wide functional analysis of CREB/long-term memory-dependent transcription reveals distinct basal and memory gene expression programs. Neuron. 2015 Jan 21;85(2):330–45.

21. Sohrabi S, Mor DE, Kaletsky R, Keyes W, Murphy CT. High-throughput behavioral screen in C. elegans reveals Parkinson's disease drug candidates. Commun Biol. 2021 Feb 15;4(1):203.

22. Vukojevic V, Gschwind L, Vogler C, Demougin P, de Quervain DJ-F, Papassotiropoulos A, et al. A role for α-adducin (ADD-1) in nematode and human memory. EMBO J. 2012 Mar 21;31(6):1453–66.

23. Fenyves BG, Arnold A, Gharat VG, Haab C, Tishinov K, Peter F, et al. Dual Role of an mps-2/KCNE-Dependent Pathway in Long-Term Memory and Age-Dependent Memory Decline. Curr Biol. 2021 Feb 8;31(3):527–539.e7.

24. Yao V, Kaletsky R, Keyes W, Mor DE, Wong AK, Sohrabi S, et al. An integrative tissue-network approach to identify and test human disease genes. Nat Biotechnol. 2018 Oct 22;

25. Sharifnia P, Jin Y. Regulatory roles of RNA binding proteins in the nervous system of C. elegans. Front Mol Neurosci. 2014;7:100.

26. Kamath RS, Ahringer J. Genome-wide RNAi screening in Caenorhabditis elegans. Methods. 2003 Aug;30(4):313–21.

27. Calixto A, Chelur D, Topalidou I, Chen X, Chalfie M. Enhanced neuronal RNAi in C. elegans using SID-1. Nat Methods. 2010 Jul;7(7):554–9.

28. Kauffman AL, Ashraf JM, Corces-Zimmerman MR, Landis JN, Murphy CT. Insulin signaling and dietary restriction differentially influence the decline of learning and memory with age. PLoS Biol. 2010 May 18;8(5):e1000372.

29. Stein GM, Murphy CT. C. elegans positive olfactory associative memory is a molecularly conserved behavioral paradigm. Neurobiol Learn Mem. 2014 Nov;115:86–94.

30. Kauffman A, Parsons L, Stein G, Wills A, Kaletsky R, Murphy C. C. elegans positive butanone learning, short-term, and long-term associative memory assays. J Vis Exp. 2011 Mar 11;(49).

31. Rankin CH, Beck CD, Chiba CM. Caenorhabditis elegans: a new model system for the study of learning and memory. Behav Brain Res. 1990 Feb 12;37(1):89–92.

32. Stetak A, Hörndli F, Maricq AV, van den Heuvel S, Hajnal A. Neuron-specific regulation of associative learning and memory by MAGI-1 in C. elegans. PLoS One. 2009 Jun 24;4(6):e6019.

33. Anantharaman V, Koonin EV, Aravind L. Comparative genomics and evolution of proteins involved in RNA metabolism. Nucleic Acids Res. 2002 Apr 1;30(7):1427–64.

34. Matia-González AM, Laing EE, Gerber AP. Conserved mRNA-binding proteomes in eukaryotic organisms. Nat Struct Mol Biol. 2015 Dec;22(12):1027–33.

35. Kaletsky R, Yao V, Williams A, Runnels AM, Tadych A, Zhou S, et al. Transcriptome analysis of adult Caenorhabditis elegans cells reveals tissue-specific gene and isoform expression. PLoS Genet. 2018 Aug 10;14(8):e1007559.

36. Holdorf AD, Higgins DP, Hart AC, Boag PR, Pazour GJ, Walhout AJM, et al. WormCat: An Online Tool for Annotation and Visualization of Caenorhabditis elegans Genome-Scale Data. Genetics. 2020 Feb;214(2):279–94.

37. Kim W, Underwood RS, Greenwald I, Shaye DD. OrthoList 2: A New Comparative Genomic Analysis of Human and Caenorhabditis elegans Genes. Genetics. 2018 Oct;210(2):445–61.

38. Arey RN, Kaletsky R, Murphy CT. Nervous system-wide profiling of presynaptic mRNAs reveals regulators of associative memory. Sci Rep. 2019 Dec 30;9(1):20314.

39. Ostroff LE, Santini E, Sears R, Deane Z, Kanadia RN, LeDoux JE, et al. Axon TRAP reveals learning-associated alterations in cortical axonal mRNAs in the lateral amgydala. Elife. 2019 Dec 11;8.

40. Kaletsky R, Lakhina V, Arey R, Williams A, Landis J, Ashraf J, et al. The C. elegans adult neuronal IIS/FOXO transcriptome reveals adult phenotype regulators. Nature. 2016 Jan 7;529(7584):92–6.

41. Nijssen J, Aguila J, Hoogstraaten R, Kee N, Hedlund E. Axon-Seq Decodes the Motor Axon Transcriptome and Its Modulation in Response to ALS. Stem Cell Rep. 2018 Dec 11;11(6):1565–78.

42. Cajigas IJ, Tushev G, Will TJ, tom Dieck S, Fuerst N, Schuman EM. The local transcriptome in the synaptic neuropil revealed by deep sequencing and high-resolution imaging. Neuron. 2012 May 10;74(3):453–66.

43. Arey R, Kaletsky R, Murphy CT. Profiling of presynaptic mRNAs reveals a role for axonal PUMILIOs in associative memory formation. BioRxiv. 2019 Aug 13;

44. Arey RN, Stein GM, Kaletsky R, Kauffman A, Murphy CT. Activation of gαq signaling enhances memory consolidation and slows cognitive decline. Neuron. 2018 May 2;98(3):562–574.e5.

45. Dubnau J, Chiang A-S, Grady L, Barditch J, Gossweiler S, McNeil J, et al. The staufen/pumilio pathway is involved in Drosophila long-term memory. Curr Biol. 2003 Feb 18;13(4):286–96.

46. Dong H, Zhu M, Meng L, Ding Y, Yang D, Zhang S, et al. Pumilio2 regulates synaptic plasticity via translational repression of synaptic receptors in mice. Oncotarget. 2018 Aug 14;9(63):32134–48.

47. Costa-Mattioli M, Gobert D, Stern E, Gamache K, Colina R, Cuello C, et al. eIF2alpha phosphorylation

bidirectionally regulates the switch from short- to long-term synaptic plasticity and memory. Cell. 2007 Apr 6;129(1):195–206.

48. Gindina S, Botsford B, Cowansage K, LeDoux J, Klann E, Hoeffer C, et al. Upregulation of eIF4E, but not other translation initiation factors, in dendritic spines during memory formation. J Comp Neurol. 2021 Aug 1;529(11):3112–26.

49. Arnold A, Rahman MM, Lee MC, Muehlhaeusser S, Katic I, Gaidatzis D, et al. Functional characterization of C. elegans Y-box-binding proteins reveals tissue-specific functions and a critical role in the formation of polysomes. Nucleic Acids Res. 2014 Dec 1;42(21):13353–69.

50. Calculli G, Lee HJ, Shen K, Pham U, Herholz M, Trifunovic A, et al. Systemic regulation of mitochondria by germline proteostasis prevents protein aggregation in the soma of C. elegans. Sci Adv. 2021 Jun 25;7(26).

51. Budkina K, El Hage K, Clément M-J, Desforges B, Bouhss A, Joshi V, et al. YB-1 unwinds mRNA secondary structures in vitro and negatively regulates stress granule assembly in HeLa cells. Nucleic Acids Res. 2021 Sep 27;49(17):10061–81.

52. Kwon E, Todorova K, Wang J, Horos R, Lee KK, Neel VA, et al. The RNA-binding protein YBX1 regulates epidermal progenitors at a posttranscriptional level. Nat Commun. 2018 Apr 30;9(1):1734.

53. Wang J-Z, Zhu H, You P, Liu H, Wang W-K, Fan X, et al. Upregulated YB-1 protein promotes glioblastoma growth through a YB-1/CCT4/mLST8/mTOR pathway. J Clin Invest. 2022 Apr 15;132(8).

54. Mehta S, McKinney C, Algie M, Verma CS, Kannan S, Harfoot R, et al. Dephosphorylation of YB-1 is Required for Nuclear Localisation During G2 Phase of the Cell Cycle. Cancers (Basel). 2020 Jan 29;12(2).

55. Taylor SR, Santpere G, Reilly M, Glenwinkel L, Poff A, McWhirter R, et al. Expression profiling of the mature *C. elegans* nervous system by single-cell RNA-Sequencing. BioRxiv. 2019 Aug 16;

56. Liska D, Wolfe Z, Norris A. VISTA: visualizing the spatial transcriptome of the C.elegans nervous system. Bioinformatics Advances. 2023 Sep 27;3(1):vbad127.

57. Evans MK, Matsui Y, Xu B, Willis C, Loome J, Milburn L, et al. Ybx1 fine-tunes PRC2 activities to control embryonic brain development. Nat Commun. 2020 Aug 13;11(1):4060.

58. Unkrüer B, Pekcec A, Fuest C, Wehmeyer A, Balda MS, Horn A, et al. Cellular localization of Y-box binding protein 1 in brain tissue of rats, macaques, and humans. BMC Neurosci. 2009 Mar 26;10:28.

59. Wang Q, Ding S-L, Li Y, Royall J, Feng D, Lesnar P, et al. The allen mouse brain common coordinate framework: A 3D reference atlas. Cell. 2020 May 14;181(4):936–953.e20.

60. Regev A, Teichmann SA, Lander ES, Amit I, Benoist C, Birney E, et al. The human cell atlas. Elife. 2017 Dec 5;6.

61. Tanaka T, Ohashi S, Funakoshi T, Kobayashi S. YB-1 binds to GluR2 mRNA and CaM1 mRNA in the brain and regulates their translational levels in an activity-dependent manner. Cell Mol Neurobiol. 2010 Oct;30(7):1089–100.

62. Distler U, Schmeisser MJ, Pelosi A, Reim D, Kuharev J, Weiczner R, et al. In-depth protein profiling of the postsynaptic density from mouse hippocampus using data-independent acquisition proteomics. Proteomics. 2014 Nov;14(21–22):2607–13.

63. Nekrasov MP, Ivshina MP, Chernov KG, Kovrigina EA, Evdokimova VM, Thomas AAM, et al. The mRNA-binding protein YB-1 (p50) prevents association of the eukaryotic initiation factor eIF4G with mRNA and inhibits protein synthesis at the initiation stage. J Biol Chem. 2003 Apr 18;278(16):13936–43.

64. Skabkin MA, Kiselyova OI, Chernov KG, Sorokin AV, Dubrovin EV, Yaminsky IV, et al. Structural organization of mRNA complexes with major core mRNP protein YB-1. Nucleic Acids Res. 2004 Oct 19;32(18):5621–35.

65. Kretov DA. Role of Y-Box Binding Proteins in Ontogenesis. Biochemistry (Mosc). 2022 Jan;87(Suppl 1):S71–4.

66. Fotovati A, Abu-Ali S, Wang P-S, Deleyrolle LP, Lee C, Triscott J, et al. YB-1 bridges neural stem cells and brain tumor-initiating cells via its roles in differentiation and cell growth. Cancer Res. 2011 Aug 15;71(16):5569–78.

67. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol. 2018 Jun 1;35(6):1547–9.

68. Hu Y, Flockhart I, Vinayagam A, Bergwitz C, Berger B, Perrimon N, et al. An integrative approach to ortholog prediction for disease-focused and other functional studies. BMC Bioinformatics. 2011 Aug 31;12:357.

69. Yang X-J, Zhu H, Mu S-R, Wei W-J, Yuan X, Wang M, et al. Crystal structure of a Y-box binding protein 1 (YB-1)-RNA complex reveals key features and residues interacting with RNA. J Biol Chem. 2019 Jul 12;294(28):10998–1010.

70. Budkina KS, Zlobin NE, Kononova SV, Ovchinnikov LP, Babakov AV. Cold Shock Domain Proteins: Structure and Interaction with Nucleic Acids. Biochemistry (Mosc). 2020 Jan;85(Suppl 1):S1–19.

71. Kretov DA, Curmi PA, Hamon L, Abrakhi S, Desforges B, Ovchinnikov LP, et al. mRNA and DNA selection via protein multimerization: YB-1 as a case study. Nucleic Acids Res. 2015 Oct 30;43(19):9457–73.

72. Eliseeva IA, Sogorina EM, Smolin EA, Kulakovskiy IV, Lyabin DN. Diverse Regulation of YB-1 and YB-3 Abundance in Mammals. Biochemistry (Mosc). 2022 Jan;87(Suppl 1):S48–167.

73. Taylor SR, Santpere G, Weinreb A, Barrett A, Reilly MB, Xu C, et al. Molecular topography of an entire nervous system. Cell. 2021 Aug 5;184(16):4329–4347.e23.

74. Hammarlund M, Hobert O, Miller DM, Sestan N. The cengen project: the complete gene expression map of an entire nervous system. Neuron. 2018 Aug 8;99(3):430–3.

75. Nishida Y, Sugi T, Nonomura M, Mori I. Identification of the AFD neuron as the site of action of the CREB protein in Caenorhabditis elegans thermotaxis. EMBO Rep. 2011 Jul 8;12(8):855–62.

76. Bargmann CI, Hartwieg E, Horvitz HR. Odorant-selective genes and neurons mediate olfaction in C. elegans. Cell. 1993 Aug 13;74(3):515–27.

77. Chou JH, Bargmann CI, Sengupta P. The Caenorhabditis elegans odr-2 gene encodes a novel Ly-6-related protein required for olfaction. Genetics. 2001 Jan;157(1):211–24.

78. Choi JI, Lee HK, Kim HS, Park SY, Lee TY, Yoon K-H, et al. Odor-dependent temporal dynamics in Caenorhabitis elegans adaptation and aversive learning behavior. PeerJ. 2018 Jun 12;6:e4956.

79. Yoshida K, Hirotsu T, Tagawa T, Oda S, Wakabayashi T, Iino Y, et al. Odour concentration-dependent olfactory preference change in C. elegans. Nat Commun. 2012 Mar 13;3:739.

80. Luo L, Gabel CV, Ha H-I, Zhang Y, Samuel ADT. Olfactory behavior of swimming C. elegans analyzed by measuring motile responses to temporal variations of odorants. J Neurophysiol. 2008 May;99(5):2617–25.

81. Kunitomo H, Sato H, Iwata R, Satoh Y, Ohno H, Yamada K, et al. Concentration memory-dependent synaptic plasticity of a taste circuit regulates salt concentration chemotaxis in Caenorhabditis elegans. Nat Commun. 2013;4:2210.

82. Nussbacher JK, Tabet R, Yeo GW, Lagier-Tourenne C.

Disruption of RNA metabolism in neurological diseases and emerging therapeutic interventions. Neuron. 2019 Apr 17;102(2):294–320.

83. Firth HV, Richards SM, Bevan AP, Clayton S, Corpas M, Rajan D, et al. DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. Am J Hum Genet. 2009 Apr 1;84(4):524–33.

84. Online Mendelian Inheritance in Man, O. MIM 616035. John Hopkins University; 2019 Jun.

85. Online Mendelian Inheritance in Man, O. MIM 138140. John Hopkins University; 2023 Dec.

86. Online Mendelian Inheritance in Man, O. MIM 601042. John Hopkins University; 2017 May.

87. Weber YG, Kamm C, Suls A, Kempfle J, Kotschet K, Schüle R, et al. Paroxysmal choreoathetosis/spasticity (DYT9) is caused by a GLUT1 defect. Neurology. 2011 Sep 6;77(10):959–64.

88. Online Mendelian Inheritance in Man, O. MIM 606777. John Hopkins University; 2016 Jun.

89. De Vivo DC, Trifiletti RR, Jacobson RI, Ronen GM, Behmand RA, Harik SI. Defective glucose transport across the blood-brain barrier as a cause of persistent hypoglycorrhachia, seizures, and developmental delay. N Engl J Med. 1991 Sep 5;325(10):703–9.

90. Online Mendelian Inheritance in Man, O. MIM 612126. John Hopkins University; 2023 Dec.

91. Münchau A, Valente EM, Shahidi GA, Eunson LH, Hanna MG, Quinn NP, et al. A new family with paroxysmal exercise induced dystonia and migraine: a clinical and genetic study. J Neurol Neurosurg Psychiatry. 2000 May;68(5):609–14.

92. Online Mendelian Inheritance in Man, O. MIM 608885. John Hopkins University; 2004 Aug.

93. Flatt JF, Guizouarn H, Burton NM, Borgese F, Tomlinson RJ, Forsyth RJ, et al. Stomatin-deficient cryohydrocytosis results from mutations in SLC2A1: a novel form of GLUT1 deficiency syndrome. Blood. 2011 Nov 10;118(19):5267–77.

94. Online Mendelian Inheritance in Man, O. MIM 614847. John Hopkins University; 2007 Aug.

95. Striano P, Weber YG, Toliat MR, Schubert J, Leu C, Chaimana R, et al. GLUT1 mutations are a rare cause of familial idiopathic generalized epilepsy. Neurology. 2012 Feb 21;78(8):557–62.

96. Online Mendelian Inheritance in Man, O. MIM 600618. John Hopkins University; 2021 Jan.

97. Online Mendelian Inheritance in Man, O. MIM 603507. John Hopkins University; 2022 Dec.

98. Online Mendelian Inheritance in Man, O. MIM 607175. John Hopkins University; 2009 Dec.

99. Lefebvre S, Bürglen L, Reboullet S, Clermont O, Burlet P, Viollet L, et al. Identification and characterization of a spinal muscular atrophy-determining gene. Cell. 1995 Jan 13;80(1):155–65.

100. Wang X, You B, Yin F, Chen C, He H, Liu F, et al. A presumed missense variant in the U2AF2 gene causes exon skipping in neurodevelopmental diseases. J Hum Genet. 2023 Jun;68(6):375–82.

101. Molitor L, Bacher S, Burczyk S, Niessing D. The molecular function of PURA and its implications in neurological diseases. Front Genet. 2021 Mar 11;12:638217.

102. Lenski C, Abidi F, Meindl A, Gibson A, Platzer M, Frank Kooy R, et al. Novel truncating mutations in the polyglutamine tract binding protein 1 gene (PQBP1) cause Renpenning syndrome and X-linked mental retardation in another family with microcephaly. Am J Hum Genet. 2004 Apr;74(4):777–80.

103. Yang Y, Muzny DM, Xia F, Niu Z, Person R, Ding Y, et al. Molecular findings among patients referred for clinical whole-exome sequencing. JAMA. 2014 Nov 12;312(18):1870–9.

104. Mani A. Pathogenicity of de novo rare variants: challenges and opportunities. Circ Cardiovasc Genet. 2017 Dec;10(6).

105. Järvelä I, Määttä T, Acharya A, Leppälä J, Jhangiani SN, Arvio M, et al. Exome sequencing reveals predominantly de novo variants in disorders with intellectual disability (ID) in the founder population of Finland. Hum Genet. 2021 Jul;140(7):1011–29.

106. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet. 2014 Mar;46(3):310–5.

107. Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M. CADD: predicting the deleteriousness of variants throughout the human genome. Nucleic Acids Res. 2019 Jan 8;47(D1):D886–94.

108. Davydov EV, Goode DL, Sirota M, Cooper GM, Sidow A, Batzoglou S. Identifying a high fraction of the human genome to be under selective constraint using GERP++. PLoS Comput Biol. 2010 Dec 2;6(12):e1001025.

109. Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, et al. REVEL: an ensemble method for predicting the pathogenicity of rare missense variants. Am J Hum Genet. 2016 Oct 6;99(4):877–85.

110. Jagadeesh KA, Wenger AM, Berger MJ, Guturu H, Stenson PD, Cooper DN, et al. M-CAP eliminates a majority of variants of uncertain significance in clinical exomes at high sensitivity. Nat Genet. 2016 Dec;48(12):1581–6.

111. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. Nat Methods. 2010 Apr;7(4):248–9.

112. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res. 2003 Jul 1;31(13):3812–4.

113. Pollard KS, Hubisz MJ, Rosenbloom KR, Siepel A. Detection of nonneutral substitution rates on mammalian phylogenies. Genome Res. 2010 Jan;20(1):110–21.

114. Jacob A, Ma Y, Nasiri E, Ochani M, Carrion J, Peng S, et al. Extracellular cold inducible RNA-binding protein mediates binge alcohol-induced brain hypoactivity and impaired cognition in mice. Mol Med. 2019 May 30;25(1):24.

115. Gong P, Liu J, Jiao X, Niu Y, Wang J, Wang X, et al. Novel biallelic loss of EEF1B2 function links to autosomal recessive intellectual disability. Hum Mutat. 2022 Mar;43(3):299–304.

116. Nguyen LS, Kim H-G, Rosenfeld JA, Shen Y, Gusella JF, Lacassie Y, et al. Contribution of copy number variants involving nonsense-mediated mRNA decay pathway genes to neuro-developmental disorders. Hum Mol Genet. 2013 May 1;22(9):1816–25.

117. Gonatopoulos-Pournatzis T, Wu M, Braunschweig U, Roth J, Han H, Best AJ, et al. Genome-wide CRISPR-Cas9 Interrogation of Splicing Networks Reveals a Mechanism for Recognition of Autism-Misregulated Neuronal Microexons. Mol Cell. 2018 Nov 1;72(3):510–524.e12.

118. Lyabin DN, Eliseeva IA, Ovchinnikov LP. YB-1 synthesis is regulated by mTOR signaling pathway. PLoS One. 2012 Dec 20;7(12):e52527.

119. Darnell JC, Van Driesche SJ, Zhang C, Hung KYS, Mele A, Fraser CE, et al. FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. Cell. 2011 Jul 22;146(2):247–61.

120. Kleene KC. Y-box proteins combine versatile cold shock domains and arginine-rich motifs (ARMs) for

pleiotropic functions in RNA biology. Biochem J. 2018 Sep 11;475(17):2769–84.

121. Brenner S. The genetics of Caenorhabditis elegans. Genetics. 1974 May;77(1):71–94.

122. Nawa M, Matsuoka M. The Method of the Body Bending Assay Using Caenorhabditis elegans. Bio Protoc. 2012;2(17).

123. Chen S, Francioli LC, Goodrich JK, Collins RL, Kanai M, Wang Q, et al. A genome-wide mutational constraint map quantified from variation in 76,156 human genomes. BioRxiv. 2022 Mar 21;

124. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, et al. The human genome browser at UCSC. Genome Res. 2002 Jun;12(6):996–1006.

125. Jones DT, Taylor WR, Thornton JM. The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci. 1992 Jun;8(3):275–82.

126. UniProt Consortium. Uniprot: the universal protein knowledgebase in 2023. Nucleic Acids Res. 2023 Jan 6;51(D1):D523–31.

127. Stecher G, Tamura K, Kumar S. Molecular Evolutionary Genetics Analysis (MEGA) for macOS. Mol Biol Evol. 2020 Apr 1;37(4):1237–9.

128. Letunic I, Bork P. 20 years of the SMART protein domain annotation resource. Nucleic Acids Res. 2018 Jan 4;46(D1):D493–6.