



Published in final edited form as:

Nat Methods. 2019 January ; 16(1): 126–133. doi:10.1038/s41592-018-0233-6.

A Compressed Sensing Framework for Efficient Dissection of Neural Circuits

Jeffrey B Lee^{#1}, Abdullah Yonar^{#1,†}, Timothy Hallacy⁴, Ching-Han Shen^{2,3}, Josselin Milloz^{2,3}, Jagan Srinivasan⁵, Askin Kocabas^{2,6}, and Sharad Ramanathan^{1,2,3,4,7,†}

¹School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138

²FAS Quantitative Biology Initiative, Center for Brain Science, Harvard University, Cambridge, MA 02138

³Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138

⁴Biophysics Program, Harvard University, Cambridge, MA 02138.

⁵Biology and Biotechnology, Worcester Polytechnic Institute, Worcester, MA 01605

⁶Department of Physics, Koç University, Sariyer, Istanbul 34450, Turkey

⁷Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA 02138.

These authors contributed equally to this work.

Abstract

A fundamental question in neuroscience is how neural networks generate behavior. The lack of neuron subtype specific genetic tools makes it challenging to determine the roles of individual subtypes in behavior. We describe a compressed sensing based framework to rapidly infer candidate neurons controlling behaviors with much fewer measurements than previously thought possible by exploiting non-specific genetic tools. We tested this framework by inferring

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:http://www.nature.com/authors/editorial_policies/license.html#terms

† Corresponding authors: sharad@post.harvard.edu, yonar@g.harvard.edu.

Author contributions

J.B.L. designed and performed experiments, analyzed data and developed the tracking microscope and wrote the manuscript. A.Y. generated transgenic lines, designed and performed experiments, analyzed data and performed simulations, developed the structural illumination for the microscope, and wrote the manuscript. T.H. performed and analyzed experiments. C.H.S. generated transgenic lines, performed and analyzed experiments. J.M. generated transgenic lines, performed and analyzed some of the Arch inhibition experiments. J.G. supervised the annotation of the expression patterns of different promoters in the different neuronal subtypes. A.K. developed the tracking microscope, performed and analyzed experiments. S.R. designed and supervised the project and wrote the manuscript.

Reporting Summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Competing interests

Authors declare no competing interests.

Code availability

Custom codes implementing the Compressed sensing analysis and the statistical analysis of the measurement matrix are available from Code Ocean platform (DOI: 10.24433/CO.020095e2-4067-4d44-9ea4-30d55309dda9 and DOI:10.24433/CO.3bc0d23b-b316-474d-9ab2-076e1deac88a).

Data availability

The data that support the findings presented in this study are available from the corresponding authors upon reasonable request.

interneuron subtypes regulating the speed of locomotion of the nematode *Caenorhabditis elegans*. We validated the inferences using a novel real time stabilization microscope for accurate long-time, high magnification imaging and targeted perturbation of neural activity in freely moving animals. We show that a circuit of three interconnected interneuron subtypes, RMG, AVB and SIA control different aspects of the speed as the animal navigates the environment. Our work suggests that compressed sensing approaches can be broadly used to identify key nodes in complex biological networks.

Main

In multiple studies on nematodes, fruit fly, zebrafish, and mouse, the number of key neurons that regulate specific behaviors have been found to be a small fraction of the total number of neurons in the nervous system^{1–10}. Conventional methods for the identification of this small fraction (Fig. 1a,b) require the interrogation of the different subtypes of neurons in the neural network one at a time¹¹. However, genetic and molecular approaches to implement this search are challenging due to the lack of unique genetic markers and promoters specific to only one neural subtype both in *C.elegans*¹² and in other animals^{13–15}. Even if specific genetic tools did exist, the measurements would have to be made one neuron sub-type at a time. Compressed sensing techniques from statistics^{16,17} have been extremely successful in both efficient sampling and accurate inference of the small fraction of key features in data^{18–21}. If only a small fraction of the neurons in a neural network are essential for specific behaviors as suggested by experimental and theoretical observations^{1–10}, compressed sensing methods could provide an alternative strategy for rapidly prioritizing the contribution of the different neurons towards a behavior, allowing for a more focused and sophisticated analysis of the key neurons. Indeed, we show that an experimental framework based on compressed sensing allows us to identify key sets of neurons controlling speed of locomotion of *C. elegans*, despite the lack of promoters to target individual neuronal subtypes.

The first concept behind compressed sensing is to make measurements in an incoherent basis: defined as a basis in which the data are not sparse²². In the context of neural circuits, this implies that perturbations of neural activity should not target one neuron subtype at a time to match measured responses to a small number of key neurons. Instead groups of neuron subtypes should be perturbed simultaneously (over a total of at least $\sim \log N$ groups, where N is the total number of neuron subtypes in the network). Following theorems by Candes and Tao^{16,22,23}, these groups need not be carefully chosen, and the method will work for most arbitrary choices, even with very different numbers of neuron subtypes in each group. The second concept of compressed sensing is that the data from as few as $\sim n \log N$ such measurements can be analyzed using L1-norm regularization to find the small number n of key neuron subtypes driving the behavior ($n \ll N$). This is much less than the N measurements needed using a conventional approach that uses unique promoters to address one neuronal subtype at a time.

Results

Compressed sensing framework to identify essential interneurons controlling speed in *C. elegans*

The lack of specific promoters and mutations that affect individual interneuron subtypes in *C. elegans* has hindered efforts to systematically uncover the roles of these neurons in different behaviors. We thus used *C. elegans* as a system in which to test a framework based on compressed sensing to systematically identify the key interneuron subtypes controlling speed. We first designed an incoherent measurement matrix by leveraging non-specific promoters. We selected 27 promoters, primarily expressed in interneurons, with characterized expression patterns in the literature (Supplementary Table 1). Each of these promoters drove expression in anywhere from 2 to 15 neurons. Together, these promoters were active in 88 of the 118 neuron-types in the animal, including 56 types of interneurons, 22 types of sensory neurons, and 10 types of motor neurons (Fig. 1c, Supplementary Table 1)^{12,24,25}. These 27 promoters generate an incoherent 27×88 dimensional measurement matrix \mathbf{M} , where each row of this matrix corresponds to the promoter identity and each column to the neuron-type. If promoter drives expression in neuron type j , $\mathbf{M}_{ij} = 1$, and else $\mathbf{M}_{ij} = 0$ (Fig. 1c). To experimentally implement the measurement matrix, we constructed transgenic *C. elegans* lines in a *lite-1* mutant background²⁶, with each line driving the expression of archaerhodopsin-3 (Arch)^{27,28} under the control of one of the 27 promoters. Although most of our lines showed expression patterns, in a defined set of neurons, consistent with the literature, a few promoters (Supplementary Fig. 1) were significantly different. We used our observed expression patterns in these instances to correct the measurement matrix (see Methods).

To measure the locomotory phenotypes of the 27 lines, animals were grown on bacteria with all trans-retinal (ATR) and then placed 1.5 cm away from the center of a bacterial lawn about ~0.3 mm in radius on a standard agarose plate while being exposed to $5\text{mW}/\text{mm}^2$ of 525nm (green) light to inhibit all neurons expressing archaerhodopsin-3 (Supplementary Fig. 2a,b; see Methods for details of the experiment and the choice of light power level²⁶). Using multi-particle tracking, we measured the speed of transgenics as they performed chemotaxis (Supplementary Fig. 2b,c and Supplementary Fig. 3). Of the 27 transgenic lines, five (*Prom_{fjp-2l}*, *Prom_{npr-4}*, *Prom_{odr-2b}*, *Prom_{sra-1l}*, *Prom_{dop-2}*) showed speed distributions that were significantly different from that of the control animals (see Methods) as scored by the KL-divergence (Fig. 1d). Thirteen lines showed changes in reversal frequency or a lower arrival fraction on the bacterial lawn than the control (Supplementary Fig. 4) despite not showing any phenotype in speed changes. Thus, while the opto-genetic perturbations are effective in many of these different lines, the ability to control speed is not broadly distributed among all the neurons of the nervous system.

Next, we used these 27 distinct sets of measurements to find the contribution of each of the 88 neuron types towards the observed phenotype. We formulated the problem as an underdetermined set of equations $\mathbf{M}\vec{\omega} = \vec{P}$, where \mathbf{M} is the measurement matrix, $\vec{\omega}$ is the vector of the neuron weights, and \vec{P} is the phenotype vector. The measured fractional

reduction in the mean speeds of the 27 lines compared to the wild type gave us the phenotype vector \vec{P} (Supplementary Fig. 2d). To solve this underdetermined set of equations, we imposed a sparsity constraint that only a small fraction of all the neuron subtypes must be essential for controlling speed. Since we did not know how many neurons were key to controlling speed, we cannot constrain the sparse solution to contain a specific number of neurons. We thus used Lasso regression^{29,30}, imposing sparsity of the solution by constraining the sum of the absolute values of the individual weights of the neurons (*i.e.*, the L1 norm of \vec{w}). Minimizing the sum of the mean squared error ($\chi^2 = (\mathbf{M}\vec{w} = \vec{P})^2$) and the L1-norm of the weight vector ($\|\vec{w}\|_1$) multiplied by a Lagrange multiplier λ , *i.e.*,

$(\mathbf{M}\vec{w} = \vec{P})^2 + \lambda \|\vec{w}\|_1$ imposes a sparsity constraint on the solution. Tuning the sparsity

parameter λ allows one to change the sparsity of the solution: increasing the value of λ leads to fewer neurons being used to fit the data. Thus, using Lasso we obtained solutions for the neuron weights with λ spanning over three orders of magnitude. We also calculated the errors in the estimated neuron weights (see Methods). Over a wide range of values of the sparsity parameter, the χ^2 error of the fit of the weight vectors to $\mathbf{M}\vec{w} = \vec{P}$ remained stable, and our analysis yielded three key interneurons with large weights: AVB, RMG and SIA (Fig. 1e). AIY appeared in the solution but had smaller weights. While the weights of AIY increased for large values of λ , in this regime, the fit of the inferred weight vectors to $\mathbf{M}\vec{w} = \vec{P}$ deteriorated rapidly (Fig. 1e). Due to error propagation, AIB, AIZ, AVA, AVK, RIA, RIC, RID, RIF, RMD, SIB, BAG, CEP, IL2, RIM, and URA appeared with small weights in a small fraction of the solutions (Fig. 1f).

We tested the robustness of the results. Using either an L1-norm minimization alone or a non-linear phenotype vector yielded results consistent with those above: a core set of three neurons (AVB, RMG and SIA) along with few other neurons with small weights (Supplementary Fig. 5a-d). We further tested that our results were robust to corruption of the original measurement matrix to account for potential variations in the levels of archeorhodopsin expression (Supplementary Fig. 6 and 7; Methods) and the sensitivity of our answers to the chosen promoters by changing the number of promoters from 22 to 32. (Supplementary Fig. 8; Methods). Next, we determined the accuracy of our results by estimating the false positive and false negative discovery rates (Supplementary Fig. 9; Methods). The number of false negative interneurons using our measurement matrix with 27 promoters was on average less than one for up to 6 essential neurons (Supplementary Fig. 9a,f). Finally, we computationally tested the efficacy of our method to analyze non-linear deep learning networks which suggested that in such networks³¹, our methods has the potential to identify the key neurons. This is consistent with the efficacy of Lasso regression in the presence of nonlinear distortion³². (Methods, Supplementary Fig. 10).

Stabilization microscope for long-time calcium imaging and optogenetic manipulation on freely moving animals

To validate the solutions inferred using Lasso, we sought to accurately image and perturb neural activity in freely behaving animals. Imaging neural activity in a moving animal at

high magnification is challenging despite impressive progress in recent works^{33–36}. To make such imaging possible, these efforts^{33–36} restricted animal movement so that they are not completely free to move (restricted at least by a glass coverslip for accurate z tracking), hence potentially affecting their speed of movement. Furthermore, these papers imaged at light levels of at least 14 mW/mm² to track the animal for five minutes by which time, the effects of phototoxicity are evident³⁵. To overcome these limitations, we built a new real-time image-stabilization microscope with the capacity to perform imaging at high magnifications and at less than a tenth of the light power levels than previously possible^{33–35} to image animals for an hour without any sign of phototoxicity. Our stabilization microscope is capable of accurately imaging fluorescence signals from both the soma and processes of multiple neurons in different z-planes as the animal moves freely on an agar plate over ten centimeters (Fig. 2a). We achieved this level of performance by stabilizing a marker neuron (AWC^{ON}, expressing mKOrange; Fig. 2b, top-right) within our field of view in x, y and z to micron accuracy (Fig. 2b, top right, Supplementary Video 1) and its angular movement to within 45 degrees as the animal crawled freely (Fig. 2c, $\sigma = 22.3$ degrees; Supplementary Video 2) through a 7 millisecond control loop. We achieved z tracking without a coverslip in a freely moving animal by imaging the fluorescent marker neuron simultaneously at two focal planes (Fig. 2b, left, Supplementary Video 3; see Methods). Rotational stabilization was achieved using a dove prism (Supplementary Video 4) mounted on a rotation stage. Having stabilized the marker neuron accurately, we used a tunable liquid lens to simultaneously image other neurons expressing GCaMP³⁷, whose fluorescence changes based on the calcium levels in the cell, at different z planes (Supplementary Fig. 11 and Supplementary Video 5). We calibrated the accuracy and performance of this microscope to image from individual neurons (Fig. 2d-f and Supplementary Fig. 12a). Using a DLP projector, we could target light on a specific individual neuron in an animal in which multiple neurons expressed Arch, allowing for specific inhibition (Fig. 2f, Supplementary Fig 12b and Supplementary Video 6). The ability to simultaneously monitor and perturb neural activity with micron precision in animals moving freely several centimeters on an agar plate allowed us to uncover the roles of individual neuron subtypes in controlling speed (see Methods and Supplementary Note 1 for details of the microscope).

Targeted inhibition of single neurons SIA, RMG, and AVB reduces speed of locomotion

Using our microscope, we tested our inference that SIA, RMG and AVB control speed. We selectively inhibited SIA in a *Prom^{dop-2}::Arch* line using targeted illumination with a randomized duty cycle to accurately infer the effects of the inhibition³⁸. We observed a consistent decrease in speed every time SIA was illuminated (Fig. 3a). To quantify this effect, we normalized the speed of animal using the speed during light-off intervals as a baseline (Fig. 3b). While control worms fed no ATR did not show speed reduction, selective SIA inhibition showed significant reduction (32%) in speed (Fig. 3b). Similarly, targeted inhibition of RMG (*Prom^{flp-21}::Arch*) and AVB (*Prom^{sra-11}::Arch*) showed significant reduction in speed (59% and 33% respectively) compared to the no ATR controls (Fig. 3c-f and Supplementary Fig. 13a), validating the inference methods.

Along with certain sensory neurons (Supplementary Note 2), the interneurons AIY and RIB have been implicated in the literature as being important for speed control³⁹. Of the

interneurons, AIY was assigned a low weight through compressed sensing and RIB was not identified. Further while the command interneuron AVB have been shown to control forward movement and is identified above, the command interneurons PVC^{40,41} as well as AVA, AVE and AVD which control reversals were not. To test the validity of our findings we inhibited these neurons individually using our microscope. Inhibition of AIY or RIB did not show a speed phenotype (Fig. 3g). Further, while the inhibition of command interneurons (AVA, AVD and AVE) showed a suppression of reversals, they did not affect speed distribution (Supplementary Fig. 13b-d)⁴². In addition, we selectively inhibited the activity of 10 neural subtypes that had small or no weights in our compressed sensing solution. Of the neurons we tested (AIB, RIM+RIC, IL2, CEP, BAG, URX, AVE, RIM, MI, PVQ), only URX showed small but significant change (17% increase) in speed with inhibition.

Calcium imaging reveals distinct roles of SIA, RMG and AVB in controlling speed

To further understand how these neurons control speed during foraging behavior, we used our microscope to measure calcium activity in different neurons as freely moving animals searched for food on agar plate. Calcium activity in URX which showed a slight speed change upon inhibition, as well as the other neurons (AIY, CEP, IL2, BAG, RIM) whose inhibition in previous experiments resulted in no change in speed, showed no significant correlation with speed (Fig. 4a and Supplementary Fig. 14; see Methods for details of the cross-correlation analysis). Again, calcium activity in RIB which was implicated in the literature to control speed did not show any correlation with speed (Fig. 4a). Similarly, while the calcium activity in the command interneuron AVA correlated with reversals, it did not correlate with speed (Fig. 4a; Supplementary Fig. 13d). Thus, the interneurons in the literature being identified as unimportant for speed regulation by compressed sensing are not false negatives in the context of chemotaxis.

In contrast to these neurons above, activity patterns in the three neuron types RMG, SIA, and AVB, which we identified through our analysis as speed-controlling neurons, showed a significant correlation (Fig. 4a and Supplementary Fig. 12). These results together with the results from the targeted inhibition of these neurons (Fig. 3) validated our compressed sensing solutions.

We next investigated the activity patterns of SIA, RMG and AVB to understand their role in controlling speed. SIA calcium dynamics in animals, freely searching for food (Fig. 4b, Supplementary Fig. 15a,b and Supplementary Video 7) showed an autocorrelation time of ~50 seconds (Fig. 4c and Supplementary Fig. 15c). Further, the correlation between SIA activity and speed increased with the window size of low-pass filter beyond 1 minute, plateauing at window sizes over 3 minutes (Fig. 4d and Supplementary Fig. 15d). The slow modulation in SIA activity and the frequency dependent cross correlation analysis suggests that SIA controls speed modulations on the timescale of minutes (Supplementary Fig. 15b). In contrast, significant correlations of RMG calcium activity with speed arose solely due to pausing events (Fig. 4e,f, Supplementary Fig. 15e and Supplementary Video 8); when pausing events were removed computationally, correlations between RMG activity and speed vanished (Fig. 4g). During pausing, RMG activity decreased, but then increased again as the animal began moving, which in conjunction with our inhibition experiments suggests

that RMG neurons control whether the animal pauses or moves (Supplementary Fig. 13a). Consistent with AVB's role as a forward command interneuron as described in literature^{40,41}, we observed that AVB activity decreases whenever the animal makes a reversal and increases during the onset of transition from backward to forward motion (Fig. 4h,i, Supplementary Fig. 15f and Supplementary Video 9). In addition to these dynamics, we also observed that AVB activity decreases during pausing (Fig. 4h, Supplementary Fig. 15c,d).

Discussion

As the nematode *C.elegans* animal searches for food in its environment, the spatial profiles of odors are translated into temporally changing signals by ciliated sensory neurons exposed at the nose tip. These signals are then processed by the interneurons of the animal to determine exploration strategies. The translation of the spatial information into a time series of odor signals is determined by the speed of the animal. Here, we identified the three interneuron types, RMG, SIA and AVB that control the nematode's speed. These three neuron types form a circuit with RMG synapsing onto both AVB and SIA²⁴ (Fig 4j). Our analysis revealed that RMG, an inter-motor neuron that synapses heavily onto head muscles, determines when the animal pauses during chemotaxis. RMG has previously been implicated in controlling how active the animal is in the context of different O₂ concentrations^{43,44} and has been termed a 'hub' interneuron involved in signal integration⁴⁵. As predicted from previous work⁴⁰, we observed that forward command interneuron AVB activity correlates with the initiation of forward movement. However, we also discovered a slower AVB dynamic that correlates with pausing. We found that SIA activity modulates speed continuously. SIA neurons were originally identified as interneurons²⁴ but they also innervate neck muscle groups¹¹. In this circuit of RMG, SIA and AVB neuron types, our data together suggest that RMG acts a switch to determine whether the animal moves or not, AVB acts as a rectifier, to determine whether the animal moves forward or not, and SIA modulates the speed continuously over timescales of a minute (Fig. 4j). Our work raises the interesting question of how inputs into this circuit, including from dopaminergic neurons, which is upstream of RMG and SIA, play a role in modulating circuit activity and the resulting speed modulations during chemotaxis (Fig. 3g, Fig. 4a and Supplementary Fig. 14a).

Statistical analyses of the measurement matrix through corruption and changes in size allow one to determine the robustness, and false positive/false negative rates of the inferences. Such analyses also allow us to identify specific promoters that can be added to the measurement matrix to improve the quality and accuracy of the inferences (Supplementary Fig. 9). These statistical analyses on our experimental measurement matrix suggest that we could have missed one interneuron subtype in our set of key neurons controlling speed. Future experiments activating neurons using channel rhodopsin and analyzing the data using our framework will help identify additional neurons we could have missed.

In conclusion, a compressive sensing based framework that exploits non-specific genetic tools, in conjunction with the microscope and the set of archaeorhodopsin lines developed in this work should allow us to rapidly and comprehensively understand the neural circuits that

drive the behaviors of *C.elegans*. Similar experimental methods based on compressive sensing have the potential to discover the key nodes that control a phenotype in complex biological networks including the nervous systems of higher organisms as well as gene regulatory networks.

Methods

Strains

Strains were grown and maintained under standard conditions unless indicated otherwise. All experiments were done in *lite-1(ce314)* mutants to minimize the animal's sensitivity to blue light. All transgenic Arch lines were generated by injecting fused promoter::Archaeorhodopsin-3-tagRFP/GFP constructs. The details of all transgenic lines used in this study can be found in Supplementary Table 1 and Supplementary Table 2.

Promoter expression, neuron identification and measurement matrix

The neuronal expression pattern of the promoter::arch constructs was verified using high magnification (63X) fluorescence and DIC microscopy. Only lines showing a phenotype or those with expression patterns that were obviously different from that of the literature under a lower magnification fluorescence microscope were examined. In most instances, a combination of fluorescence-DIC co-localization and neuron process morphology was sufficient to determine the expression pattern. We imaged a minimum of 10 animals for each line and claimed that the line expressed those neurons reliably if they were observed in at least 75% of the animals. Promoters and lines that showed a higher degree of mosaicism was discarded. The measured expression patterns were used to generate the measurement matrix **M**.

Behavioral assay with optogenetic perturbations and image processing

All promoters were fused to Archaeorhodopsin-3 by fusion-PCR and injected into *C. elegans* using standard protocols to create transgenic animals. Transgenic animals were fed all-trans retinal (1mM), a cofactor required for rhodopsin activity, for 12+ hours prior to the behavioral assay. To minimize the effect of the mosaicism in the experiments, animals were carefully selected under fluorescent microscope to pick only the worms that express Arch in all of the identified neurons in each behavioral experiment. Only young adults from each transgenic line were used in the behavioral assay. 5 to 50 animals at a time were placed 1.5 cm away from the center of the bacteria lawn, which is about 3 mm in radius. The same number of worms per run was used for corresponding controls every time. The green light ($5\text{mW}/\text{mm}^2$) from a ScopeLED G250 used to activate arch covered >70% of the plate area with less than 15% variation in intensity as verified with a Thorlabs S120C photodiode. It has been reported that worms can behave normally for more than hour under $5.5\text{mW}/\text{mm}^2$ green light²⁶. Thus, we used $5\text{mW}/\text{mm}^2$ as the safest known high power level in our silencing experiments. A video camera was used to record the movements of the worms at 3–5 fps⁴⁶. Only data from ~70% of the plate area centered on the bacterial lawn was processed and analyzed. For each set of experiments performed on a day, a parallel set of positive and negative control experiments were performed under identical conditions. As positive controls, experiments were performed on animals with pan-neuronal expression of

Archeorhodopsin. Under illumination, these animals paralyze, thus validating our set up as well as the quality of all-trans retinal. As negative controls animals from the same line that were not fed all trans retinal (and were hence unaffected by green light) and *lite-1* animals were used. The same LED illumination was used for the control and the experiments.

Custom MATLAB software with elements from “The Parallel Worm Tracker” (Miriam Goodman Lab)⁴⁷ was used to segment and track the position of individual nematodes across camera frames. We tracked worms until they entered the bacterial lawn. To estimate sub-pixel and hence sub-resolution scale displacements, we binned the real space into discrete pixels based on pixel size, i.e. a non-zero displacement was only observed if the animal moved more than $2 \times$ pixel width, the distance between diagonally adjacent pixels. Such analysis is likely more accurate than arbitrary time averaging which is typically used and does not take spatial resolution into account. Under roughly 1x magnification, each animal was typically around 50–100 pixels in size. Speed was calculated by taking the time derivative of centroid position.

Data analysis for behavior screen

The speed of animals was measured until they reached the bacteria lawn. Speed naturally fluctuates during the movement of wild type animals, and so to determine whether neurons affect speed, one must measure the distribution of the speed over the course of the behavioral experiment. Therefore, for each transgenic line and corresponding control, speed data from at least 10 animals was analyzed to obtain the speed distribution for that line. To determine whether the transgenic line showed significant changes in speed, we measured the Kullback-Leibler (KL) divergence between the speed distribution of this line and the control experiment. In parallel we also measured the KL divergence between the 131 distinct control experiments in our data set to obtain the distribution of the intra-control KL divergence values. As long as the KL divergence between the speed distribution of the transgenic line and the control was above the 95th percentile of the intra-control KL divergence values, we assumed that the transgenic line showed a speed phenotype.

To infer key neurons robustly, we tested two different phenotype vectors based on KL-divergence: binarized and continuous. For continuous phenotype vector, mean speed reductions were calculated with respect to control. Promoters not distinguishable from control were taken to have no speed change (we further tested that this latter assumption did not affect the results). Using the phenotype vector \vec{P} , and the measurement matrix \mathbf{M} , both the neuron weights \vec{w} and confidence in these weights were calculated for each 88 neuron types. To calculate the confidence in the weights, errors in the measured mean speed changes were propagated onto inferred neuron weights with bootstrapping. We sampled phenotype vector 10,000 times from normal distribution with a mean as fractional mean speed changes and standard deviation as standard error on mean speed changes. We, then, used these phenotype vectors to calculate the weight distribution for each 88 neuron types. For these analyses, Lasso and LassoCV packages from Python machine learning library (scikit-learn v0.19.2) was used with the default tolerance value. Other optimization methods were also tested such as MATLAB lasso, convex optimization package (CVX package

(<http://cvxr.com/cvx>), MATLAB *linprog* (L1-norm minimization: $\min \|\vec{\omega}\|_1$ subject to $\mathbf{M}\vec{\omega} = \vec{P}$). They also yielded consistent results.

Artificial neural network

In order to test whether we could infer key neurons in a nonlinear neural network, we trained a fully connected feedforward neural network using deep learning algorithms to recognize handwritten digits (which is a standard task for deep neural networks)⁴⁸. This network classified digits as 0 to 9 from 28 by 28 size gray scale images (MNIST, <http://yann.lecun.com/exdb/mnist/>). The network consisted of the input layer of 784 neurons, 4 hidden layers each having 100 neurons and the output layer consist of 10 neurons signaling the output class of the digits (Supplementary Fig. 10a). Each neuron in the network was extremely nonlinear with a sigmoidal activation function (Supplementary Fig. 10b). We trained this network on a training dataset (60,000 images) using categorical-crossentropy as a loss function. During training to produce networks with a small number of key neurons and to promote sparse representation of the data, we added L1 constrain to activations of neurons. After training, when the network was shown examples of handwritten digits from the training dataset, a fraction of units(neurons) in the network showed activity for individual digit images while with white noise data, the activation was broad (Supplementary Fig. 10c). For simplicity, we focused on single hidden (intermediate) layer of the trained network in our analysis. Exploiting our computational access to the neurons in this network, we inhibited each neuron in a layer one at a time to measure the change in accuracy of handwriting recognition on a test dataset consisting of 10,000 images. These measurements quantified the weights of individual neurons in handwriting recognition and constituted the ground truth about the network (Supplementary Fig. 10d). We then tested whether through compressed sensing we could recover these weights and in particular identify neurons that play an essential role in digit recognition. To do so, we built a measurement matrix by generating 50 groups each containing 5 randomly chosen neurons. As in our *C.elegans* experiments, we inhibited all the neurons in a row of the measurement matrix and measured the phenotype (change in accuracy of handwriting recognition on the test dataset) (Supplementary Fig. 10e,f). Using Lasso, with this measurement matrix we could correctly identify the key neurons essential for digit recognition. (Supplementary Fig. 10g-i). We also determined how error rate and accuracy of inference changes with the size (number of rows) of the measurement matrix. We changed the number of rows (equivalent the number of independent measurements using different groups) from 25 to 60 to determine how performance of the method changes with the size of the measurement matrix (Supplementary Fig. 10j). These tests suggested that compressed sensing can be used to find essential neurons in this extremely nonlinear artificial neural network. This work was implemented using Keras v2.1.3 python deep learning library with TensorFlow backend.

Statistical analysis of the measurement matrix

In order to evaluate the performance of our measurement matrix, we performed simulations. In these simulations we first assigned a randomly chosen set of n neurons (with n between 1 and 11 out of the total of 88 subtypes covered by the measurement matrix) as essential for behavior. We assigned weights ($\vec{\omega}_{real}$) to these essential neurons such that when each of

these essential neurons was inhibited in our simulation, the phenotype changed positively or negatively by at least 30% of the wild type value. We assigned 0 weights to non-essential neurons. We first assumed that when multiple neurons were inhibited simultaneously, the effects of these neurons on the phenotype added linearly. We used our measurement matrix to calculate the phenotype vector $\vec{P} = \mathbf{M}\vec{\omega}_{real}$. We then hid the identities of the essential neurons and attempted to infer these identities from this phenotype vector \vec{P} using compressed sensing. We used Lasso to infer the weights $\vec{\omega}_{inferred}$ of individual neurons. We compared $\vec{\omega}_{inferred}$ against $\vec{\omega}_{real}$ to obtain false positive and false negative rates. By performing this computational test thousand times for different randomly chosen sets of essential neurons, we determined the average false positive and false negative rates (Supplementary Fig. 9a). We then ascertained that the false positive/false negative rates were robust to how the phenotype vector was evaluated by using a non-linear function such that the combined phenotype of multiple neurons did not add linearly. We did so by evaluating the phenotype: $\vec{P} = sgn(\mathbf{M}\vec{\omega}_{real})$ where $sgn(x) = \{-1 \text{ if } x < 0; 0 \text{ if } x = 0; 1 \text{ if } x > 0\}$ and repeated the calculation of the false positive and false negative rates (Supplementary Fig. 9b). We further evaluated the false positive and false negative rates as the measurement matrix size changed from 22 to 32 rows (promoters) (Supplementary Fig. 9c-e) as well as how these rates depended on the total number of promoter constructs used in the experiments (Supplementary Fig. 9f).

Promoter removals and additions

We tested the robustness of the solution: RMG, SIA and AVB to the removal of promoters from and additions of promoters to the original set, and thus to changes in the measurement matrix. We used mean speed changes as our phenotype vector as before, but we did not do bootstrapping analysis with the uncertainty on the measurements this time for simplicity. First, we randomly removed from 1 to 5 promoters from our measurement matrix to create new measurement matrices. We next obtained the identities of the key neurons for a range of sparsity parameter for each of these new matrices (Supplementary Fig. 6a). In all instances, the inference of AVB, SIA and RMG was robust to the removal of promoters (Supplementary Fig. 6b). Nevertheless, upon removal of *Pmgl-1*, the solution contains AIA and decreases the weight of AVB (Supplementary Fig. 6b). This is because AIA is expressed in two promoters: *Pmgl-1* with no speed phenotype and *Psra-11* with significant speed phenotype. Therefore, removal of *Pmgl-1* introduces AIA as a neuron that could control speed. Once *Pmgl-1* and the fact that it does not show a speed phenotype is included, AIA is not contained in the sparse solution (Supplementary Fig. 6c). The sensitivity of the inference to perturbations of the measurement matrix allows us to determine the neurons whose inferences are most sensitive to the identity of the chosen promoters.

To test the robustness upon promoter additions, we generated 5 more transgenic lines expressing Arch in multiple neurons and measured speed phenotype in each of these lines. We added these five new promoters (*Podr-2(16)*, *Popt-3*, *Pntr-2*, *Psams-5*, *Ptdc-1*) to the original set to increase the size of our matrix to 32 promoters. The identity of the key

inferred neurons remained unchanged with these new behavioral data and the expanded measurement matrix (Supplementary Fig. 7).

Arch efficiency simulations

While the arch lines were imaged at high magnification to validate their expression pattern and only animals showing clear expression patterns of the fluorescent tag were picked for behavioral assays, errors in our inference due to variations in the levels of archeorhodopsin expression are possible. We further tested the robustness of our methods to corruption of the measurement matrix. Our original measurement matrix had an entry of 1 if a promoter drove expression in a neuron type and 0 otherwise. We corrupted 10% of the entries of our measurement matrix at random computationally, such that the 1s were changed with equal probability to a number between 0 and 0.5 (corresponding to randomly reducing or abrogating the expression of a subset of the neurons in a subset of the promoters). We thus generated 1000 distinct corrupted measurement matrices. We then used these corrupted measurement matrices to infer the key neurons from behavioral experiments to determine the statistical robustness of our inferences to reduction and abrogation of expression levels in subsets of neurons. In almost every instance, the solutions (RMG, SIA, AVB) were robust. There were two exceptions: when the corruption of the measurement matrix (Supplementary Fig. 8a) led to the abrogation of expression in RMG in the promoter *Pflp21* or in AVB in the promoter *Psra-11*, RMG or AVB were missed as key neurons respectively. This is because in our entire set of promoters, only *Pflp21* drives expression in RMG, and only *Psra-11* drives expression AVB. Such perturbations of level reduction in SIA in individual promoters did not affect the inference because multiple promoters in our arch lines show expression in SIA (Supplementary Fig. 8a,b).

Tracking (Stabilization) Microscope

Real-time tracking and image stabilization—Tracking and image stabilization are achieved through a combination of fast acquisition, computation, and mechanical feedback. Our system is able to image, process, and feedback to x, y, z and rotational controls, all within 7 milliseconds. In order to stabilize worm movement, motion in the x,y,z coordinates, and rotation resulting from head bending, must all be tracked simultaneously. Briefly, a green 543nm laser (<5mW/mm²) is used to excite the mKusabira-Orange (mKO) tagged marker neuron, AWC^{ON}. Fluorescence is collected by a 50X, 0.55 numerical aperture (Nikon 50X LU Plan ELWD 0.55NA) objective and passed through a dove prism mounted on a rotation stage. The dove prism utilizes internal reflections to rotate the image. The rotated beam is then split into two adjacent images that focus at slightly different z depths on an EMCCD - one in focus, the other slightly out of focus (Fig. 2b). This step, colloquially referred to as “dual-view”, was accomplished by inserting a slightly focusing lens in one of the otherwise optically equal length paths and is a modern interpretation of a classic technique used by H. Berg to track swimming *E. coli*⁴⁹. The pixels from the camera are then read serially onto the FPGA and processed immediately as they enter.

To facilitate fast and reliable real-time processing, Field Programmable Gate Arrays were used. FPGAs are parallel processing CPUs, that allow high speed synchronization of multiple simultaneous processes required for our tracking system. Positional x,y,z,r

information were simultaneously extracted from raw images in parallel, without any one process delaying another. The x,y position was calculated from the centroid of the cell body in the focused image. Z position was measured by calculating the difference in intensity between the focused and out-of-focus images. The angular orientation, θ , was calculated from the marker neuron's process which extends from the cell body to the animal's nose tip. The displacement from the target position was then passed to x,y,z, θ stages (x,y: piezo +servo, z: stepper, θ : servo) to mechanically move worm back to the target location. This was all done at 100 frames per second and less than 10 millisecond latency. As a result, we could stabilize the >250 μ m/sec movements of the marker neuron cell body to 1 μ m—less than its radius. Head bending motion artifacts were also reduced two-fold (Fig. 2b). Tracking and stabilization with 1-micron accuracy could be maintained over an hour and over centimeters of travel.

Dove prism for image rotation—We used a dove prism (Thorlabs PS992-A) mounted in a rotation stage to rotate the image and compensate for the worm's head bending since rapid computation by the FPGA was only possible using addition and subtraction operations and hence there was insufficient time to computationally rotate the image. Dove prisms have been used in astronomy and machine vision as 'beam rotators'⁵⁰. A challenge of aligning the dove prism is that its rotation axis must be collinear with that of the rotation stage. Small lateral or angular deviations result in image nutation, or 'rocking'. Fortunately, there are some established alignment procedures^{51,52}. We assembled an alignment apparatus using a combination of off-the-shelf stages. Then, we aligned the dove prism by minimizing the rotational displacement. More precisely, we measured the trajectory of a point-like fluorescent particle in our field of view as we rotated the dove prism and iteratively adjusted translations orthogonal to optical axis as well as the azimuth (tilt) to minimize the path total length and asymmetry as previously done in literature^{51,52}.

Z tracking from two focal planes—We tracked z position by comparing the fuzziness between two images of the marker neuron taken at different focal planes. The technique was first utilized by Berg in 1971⁴⁹ to track swimming *E. coli*. The fuzziness of an object, as measured by its intensity, does not give the absolute position as the fuzziness is symmetric above and below the focal plane. Comparing this to the image from a different focal plane, however, allows the exact z position to be calculated as the difference in intensity between the two images is monotonic near either focal plane. Berg explains, as the object moves toward one focal plane, "its image sharpens on [that focal plane] and fuzzes out more on [the other plane]"⁴⁹.

X, Y, Z, R stages—To get precision and speed, along with range, we stacked fast but small x,y piezo stages (PI) on long travel servomotor stages (Newport LTA-HS). Micron-scale position was read from the encoded stages onto an FPGA (National Instruments PXIe-7966R) as an analog voltage. A non-encoded linear stepper motorized stage (Thorlabs ZFS13B, Sparkfun EasyDriver) was used for controlling z position. The dove prism used for rotating the image was placed in a servo controlled rotation stage (Newport URB100C) which was set to track a voltage and corresponding angle using its own internal PID feedback. The servo stages were all driven by a stage controller (Newport XPS Controller)

and custom written .tcl scripts. All tracking movement signals were processed and sent by an FPGA. Custom built digital-to-analog and analog-to-digital converters (Harvard University Electronic Instrument Design Lab) were used wherever necessary.

LabVIEW FPGA for hardware control and real-time processing—Field

Programmable Gate Arrays (FPGAs) were used to control both tracking and imaging. Processes on FPGAs can run in parallel as opposed to sequentially on a CPU. This means that each process loop runs exactly at its specified rate no matter the loading on itself or other loops. We used a National Instruments 1483 Camera Link Adapter Module to interface the camera (Andor iXon Ultra 897) to the FPGA. The different stages have different sampling rates and because the FPGA can communicate with each of them independently, no process delays another.

FPGAs are programmed with The LabVIEW FPGA platform without using low-level hardware language. LabVIEW FPGA compiles code written in the standard LabVIEW interface into VHDL (Hardware Description Language). LabVIEW FPGA has the additional benefit of a PC interface which we used to not only to save images and data, but also to monitor the experiment in real-time and manually control the stages when necessary. Although data was saved on the PC where rates are variable, we minimized timing uncertainty by bundling timestamps from the FPGA clock with the data. All GCaMP, tracking, and wide-field images were synchronized using these timestamps.

The algorithms used to process images and track the worm are simple and fast. In the x,y,z tracking loop, 2 focal plane images of the marker neuron are first thresholded to remove all but the soma. The centroid of the soma from the in-focus image was calculated to obtain x and y positions. The difference in power, sum of all pixel intensities, between the in and out of focus images was used as a measurement for z position. In the tracking loop, a different threshold was used to remove background noise but keep the cell process. The angular orientation of process was determined by first fitting an approximate spline. The process was first split into roughly 8 segments by binning pixel rows roughly perpendicular to the process, calculating the centroid in each of them, and then drawing a line between centroids from adjacent bins. The angular orientations of these segments were then averaged to obtain the process' orientation. All of these position calculations occurred as pixels are serially read from the camera and completed by the time the last pixel in the image was read. This setup ensured the calculations did not add any latency, and their speeds were limited only by the camera's data transfer rate.

GCaMP imaging—Tracking and image stabilization allowed GCaMP to be imaged as if the worm were immobilized. A blue LED or DLP projector (Logic PD DLP LightCommander) was used to provide wide-field illumination for GCaMP imaging. 30 millisecond exposure time was used, which enabled us to use blue light levels around only $0.5\text{--}1\text{ mW/mm}^2$ —several times less than the reported dangerous level²⁶. This low blue light level is one of the reasons we have been able to image freely moving worms at high spatio-temporal resolution for an order of magnitude longer than other methods. Sampling rate was 15.625 frames per second. Stabilizing the image also enabled us to perform volumetric z scans as can be done on some commercial microscopes equipped with autofocus systems,

such as Zeiss Definite Focus and Nikon Perfect Focus, except that those systems track the coverslip surface rather than the sample. These commercial systems can compensate for temperature related z drift but cannot do volumetric scanning in a moving sample. Rather than changing the objective to sample distance to vary the focal plane, a liquid lens with an adjustable focal length (Optotune EL-10-30-C) was placed in a 1x relay system between the tube lens and GCaMP imaging camera (Hamamatsu ImagEM X2). An FPGA (National Instruments PXIe-7961) was used to acquire images from the camera and control the liquid lens. The focal length of the tunable lens is voltage controlled and can be varied between 100 to 200mm. In this configuration, the tunable lens can scan through the entire ~50 μm width of the worm without any variance in image quality or magnification⁵³. Including the 30ms exposure times, the entire worm (15 z-slices) could be scanned through in one second.

GFP control to characterize motion artifacts and sampling noise—We used GFP as a control to characterize remaining motion artifacts and noise in the microscope. A specific promoter (*P_{ttx-3}*) was used to drive GFP expression in a left right pair of interneurons: AIYR/L. Movement artifacts would have the same effect on measuring GFP as it would GCaMP. Similarly, readout and dark noise from the camera would be the same, as would shot noise given that the GFP and GCaMP intensities are comparable. GFP control shows that the error in the measurement, including all remaining motion artifacts and camera noise, was less than 18% as measured by the standard deviation (Fig. 2e). Smoothed GFP intensity time series over 3 seconds had standard deviation is less than 10% indicating that much of the noise we observe is at frequencies faster than 0.3Hz.

Structured illumination—The DLP projector was used for spatially structured illumination for targeted photo-activation. A series of ~ 75mm relay lenses were used to map the image of the DLP to the sample plane and appropriate excitation filters were used depending on the application. The DLP was controlled using a combination of software written in LabVIEW and its native software. In our configuration, the DLP is limited to refreshing at 30 fps through HDMI, but with more hardware-level control, this could be increased to 60+ fps.

Targeted inhibition of neurons

Animals expressing multiple neurons were fed with ATR overnight before tracking with our stabilization microscope. After tracking and stabilization, we located the position of the neuron of interest on the camera. A pattern of light was then projected from the DLP projector to selectively target the neuron of interest. In order to ensure that only the neuron of interest remained illuminated as the worm changed orientation, we used a custom labview script that computationally rotated the projected image in sync with the dove prism. This was made possible by aligning the center of the rotation of the projected pattern to the center of rotation of the dove prism. By following the entire process on camera, we were able to confirm the selective illumination of the neuron (Fig. 2f).

Targeted inhibition experiments were performed by pulsing structured illumination patterns while recording the animals speed. Speed changes due to inhibition shown in Fig. 3 were found by smoothing the speed traces with a ~1 sec low pass filter and then normalizing

speed during pulses (light ON periods) with the speed between pulses (light OFF periods). If OFF period average speed was less than 20 μ m/s, we discarded those cycles since it generates artifacts and introduce bias to the distribution. The results are not sensitive to 20 μ m/s threshold. Each track (data points) had 5 cycles on average. Transgenic lines used in targeted inhibition experiments and single neuron expressing lines related to Fig. 3 are listed in Supplementary Table 2.

Image analysis of GCaMP data

GCaMP intensity information was extracted using custom software written in MATLAB. Regions of interest and segmentation corresponding to different neurons in x,y are obtained from the maximum intensity projection. Because of the image stabilization and tracking, each neuron stays localized to a specific region of interest throughout the time series (Supplementary Fig. 11a,b). Within a region of interest that resolves neurons in x,y (Supplementary Fig. 11b), neurons at different z position could be identified from the distribution of maximum intensity pixel containing slice each stack since the intensity of a neuron increases towards the focused plane. The best focal plane was taken as that which maximizes intensity in the region of interest for specific neuron (Supplementary Fig. 11c,d).

GCaMP activity and speed correlation analysis

Cross correlation between calcium activity and speed of animals were obtained at different frequencies by applying low pass, high pass or band pass filters to both GCaMP and Speed to find the best significant correlation between two signals using Pearson correlation. GCaMP data was randomly shuffled 100 times and the correlation coefficient was calculated. The correlation with the shuffled data was used to measure if cross-correlation between the actual calcium activity and speed was significant or not. For instance, SIA experiments shown in Supplementary Fig. 15a and b, have significant correlation even when the smoothing window sizes are up to 3 minutes long (Supplementary Fig. 15g,h) while the significance of the correlation of other neurons deteriorated with the smoothing window size (Supplementary Fig. 15i,j).

Statistical information and reproducibility

Statistical test used in this study is written in corresponding figure legends. Other randomization and bootstrapping methods applied to data were explained in detail in the Methods section. Details on statistical test used in Fig. 3 can be found in the Supplementary Table 3. Experiments in Fig. 2b,c were repeated independently with similar results more than fifty times. Experiments in Fig. 2d,e were repeated three times independently with similar results. Targeted illumination control experiments as in Fig. 2f were repeated three times independently with similar results. Experiments in Fig. 3a-f were repeated independently with similar results 5, 5, 10, 5, 5 and 10 times respectively. Experiments in Fig. 4b were repeated nine times independently with similar results. Experiments in Fig. 4e,f were repeated four times independently with similar results. Experiments in Fig. 4h,i were repeated seven times independently with similar results. Experiments in Supplementary videos 7, 8 and 9 were repeated independently with similar results nine, four and seven times respectively.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank Niels Ringstad and members of the Ramanathan Lab for useful discussions and comments. This work was supported by an NIH Directors Pioneer award (DP1OD008197, S.R.) and a grant from the NIH (R01DC016058, J.S.). We also acknowledge support from the NSF-Simons Center for Mathematical and Statistical Analysis of Biology at Harvard University, supported by NSF grant no. DMS-1764269, and the Harvard FAS Quantitative Biology Initiative (S.R.).

References

1. Trent C, Tsung N & Horvitz HR Egg-Laying Defective Mutants of the Nematode *Caenorhabditis elegans*. *Genetics* 104, 619–647 (1983). [PubMed: 11813735]
2. Beverly M, Anbil S & Sengupta P Degeneracy and Neuromodulation among Thermosensory Neurons Contribute to Robust Thermosensory Behaviors in *Caenorhabditis elegans*. *J. Neurosci.* 31, 11718–11727 (2011). [PubMed: 21832201]
3. Luo L et al. Bidirectional thermotaxis in *Caenorhabditis elegans* is mediated by distinct sensorimotor strategies driven by the AFD thermosensory neurons. *Proc. Natl. Acad. Sci.* 111, 2776–2781 (2014). [PubMed: 24550307]
4. Zimmer M et al. Neurons Detect Increases and Decreases in Oxygen Levels Using Distinct Guanylate Cyclases. *Neuron* 61, 865–879 (2009). [PubMed: 19323996]
5. Zaslaver A et al. Hierarchical sparse coding in the sensory system of *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci.* 112, 1185–1189 (2015). [PubMed: 25583501]
6. White JQ & Jorgensen EM Sensation in a Single Neuron Pair Represses Male Behavior in Hermaphrodites. *Neuron* 75, 593–600 (2012). [PubMed: 22920252]
7. Tran DH, Meissner GW, French RL & Baker BS A Small Subset of Fruitless Subesophageal Neurons Modulate Early Courtship in *Drosophila*. *PLOS ONE* 9, e95472 (2014). [PubMed: 24740138]
8. Huang K-H, Ahrens MB, Dunn TW & Engert F Spinal Projection Neurons Control Turning Behaviors in Zebrafish. *Curr. Biol.* 23, 1566–1573 (2013). [PubMed: 23910662]
9. Capelli P, Pivetta C, Esposito MS & Arber S Locomotor speed control circuits in the caudal brainstem. *Nature* 551, 373 (2017). [PubMed: 29059682]
10. Milo R et al. Network Motifs: Simple Building Blocks of Complex Networks. *Science* 298, 824–827 (2002). [PubMed: 12399590]
11. Gray JM, Hill JJ & Bargmann CI A circuit for navigation in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. U. S. A.* 102, 3184–3191 (2005). [PubMed: 15689400]
12. Hobert O, Glenwinkel L & White J Revisiting Neuronal Cell Type Classification in *Caenorhabditis elegans*. *Curr. Biol.* 26, R1197–R1203 (2016). [PubMed: 27875702]
13. Jones AR et al. An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature* 489, 391 (2012). [PubMed: 22996553]
14. Liang AJ et al. Genome-wide atlas of gene expression in the adult mouse brain. *Nature* 445, 168 (2006). [PubMed: 17151600]
15. Murphey DK, Herman AM & Arenkiel BR Dissecting inhibitory brain circuits with genetically-targeted technologies. *Front. Neural Circuits* 8, (2014).
16. Candès EJ, Romberg JK & Tao T Stable signal recovery from incomplete and inaccurate measurements. *Commun. Pure Appl. Math.* 59, 1207–1223 (2006).
17. Donoho DL Compressed sensing. *IEEE Trans. Inf. Theory* 52, 1289–1306 (2006).
18. Lustig M, Donoho DL, Santos JM & Pauly JM Compressed Sensing MRI. *IEEE Signal Process. Mag.* 25, 72–82 (2008).
19. Satat G, Tancik M & Raskar R Lensless Imaging With Compressive Ultrafast Sensing. *IEEE Trans. Comput. Imaging* 3, 398–407 (2017).

20. Shental N, Amir A & Zuk O Identification of rare alleles and their carriers using compressed sensing. *Nucleic Acids Res.* 38, e179–e179 (2010). [PubMed: 20699269]
21. Amir A & Zuk O Bacterial Community Reconstruction Using Compressed Sensing. *J. Comput. Biol.* 18, 1723–1741 (2011). [PubMed: 21999287]
22. Candes EJ & Wakin MB An Introduction To Compressive Sampling. *IEEE Signal Process. Mag.* 25, 21–30 (2008).
23. Candes EJ & Tao T Decoding by linear programming. *IEEE Trans. Inf. Theory* 51, 4203–4215 (2005).
24. White JG, Southgate E, Thomson JN & Brenner S The structure of the nervous system of the nematode *Caenorhabditis elegans*. *Phil Trans R Soc Lond B* 314, 1–340 (1986). [PubMed: 22462104]
25. Xu M et al. Computer Assisted Assembly of Connectomes from Electron Micrographs: Application to *Caenorhabditis elegans*. *PLOS ONE* 8, e54050 (2013). [PubMed: 23342070]
26. Edwards SL et al. A Novel Molecular Solution for Ultraviolet Light Detection in *Caenorhabditis elegans*. *PLOS Biol.* 6, e198 (2008). [PubMed: 18687026]
27. Dobry AS et al. High-performance genetically targetable optical neural silencing by light-driven proton pumps. *Nature* 463, 98 (2010). [PubMed: 20054397]
28. Husson SJ et al. Microbial Light-Activatable Proton Pumps as Neuronal Inhibitors to Functionally Dissect Neuronal Networks in *C. elegans*. *PLOS ONE* 7, e40937 (2012). [PubMed: 22815873]
29. Tibshirani R Regression shrinkage and selection via the lasso: a retrospective. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 73, 273–282 (2011).
30. Tibshirani R Regression Shrinkage and Selection via the Lasso. *J. R. Stat. Soc. Ser. B Methodol.* 58, 267–288 (1996).
31. Hu T & Chklovskii DB Reconstruction of Sparse Circuits Using Multi-neuronal Excitation (RESCUME). *Adv. Neural Inf. Process. Syst. NIPS* 22, (2009).
32. Thrampoulidis C, Abbasi E & Hassibi B LASSO with Non-linear Measurements is Equivalent to One With Linear Measurements in *Advances in Neural Information Processing Systems* 28 (eds. Cortes C, Lawrence ND, Lee DD, Sugiyama M & Garnett R) 3420–3428 (Curran Associates, Inc., 2015).
33. Shipley FB, Clark CM, Alkema MJ & Leifer AM Simultaneous optogenetic manipulation and calcium imaging in freely moving *C. elegans*. *Front. Neural Circuits* 8, (2014).
34. Piggott BJ, Liu J, Feng Z, Wescott SA & Xu XZS The Neural Circuits and Synaptic Mechanisms Underlying Motor Initiation in *C. elegans*. *Cell* 147, 922–933 (2011). [PubMed: 22078887]
35. Nguyen JP et al. Whole-brain calcium imaging with cellular resolution in freely behaving *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci.* 113, E1074–E1081 (2016). [PubMed: 26712014]
36. Faumont S et al. An Image-Free Opto-Mechanical System for Creating Virtual Environments and Imaging Neuronal Activity in Freely Moving *Caenorhabditis elegans*. *PLoS ONE* 6, e24666 (2011). [PubMed: 21969859]
37. Chen T-W et al. Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* 499, 295 (2013). [PubMed: 23868258]
38. Jazayeri M & Afraz A Navigating the Neural Space in Search of the Neural Code. *Neuron* 93, 1003–1014 (2017). [PubMed: 28279349]
39. Li Z, Liu J, Zheng M & Xu XZS Encoding of Both Analog- and Digital-like Behavioral Outputs by One *C. elegans* Interneuron. *Cell* 159, 751–765 (2014). [PubMed: 25417153]
40. Kawano T et al. An Imbalancing Act: Gap Junctions Reduce the Backward Motor Circuit Activity to Bias *C. elegans* for Forward Locomotion. *Neuron* 72, 572–586 (2011). [PubMed: 22099460]
41. Chalfie M et al. The neural circuit for touch sensitivity in *Caenorhabditis elegans*. *J. Neurosci.* 5, 956–964 (1985). [PubMed: 3981252]
42. Roberts WM et al. A stochastic neuronal model predicts random search behaviors at multiple spatial scales in *C. elegans*. *eLife* 5, e12572 (2016). [PubMed: 26824391]
43. Busch KE et al. Tonic signaling from O₂ sensors sets neural circuit activity and behavioral state. *Nat. Neurosci.* 15, 581–591 (2012). [PubMed: 22388961]

44. Laurent P et al. Decoding a neural circuit controlling global animal state in *C. elegans*. *eLife* 4, e04241 (2015).
45. Bargmann CI et al. A hub-and-spoke circuit drives pheromone attraction and social behaviour in *C. elegans*. *Nature* 458, 1171 (2009). [PubMed: 19349961]

Methods only references

46. Baek J-H, Cosman P, Feng Z, Silver J & Schafer WR Using machine vision to analyze and classify *Caenorhabditis elegans* behavioral phenotypes quantitatively. *J. Neurosci. Methods* 118, 9–21 (2002). [PubMed: 12191753]
47. Ramot D, Johnson BE, Jr TLB, Carnell L & Goodman MB The Parallel Worm Tracker: A Platform for Measuring Average Speed and Drug-Induced Paralysis in Nematodes. *PLOS ONE* 3, e2208 (2008). [PubMed: 18493300]
48. Lecun Y, Bottou L, Bengio Y & Haffner P Gradient-based learning applied to document recognition. *Proc. IEEE* 86, 2278–2324 (1998).
49. Berg HC How to Track Bacteria. *Rev. Sci. Instrum.* 42, 868–871 (1971). [PubMed: 4940742]
50. Shin CW, Inokuchi S & Kim KI Retina-like visual sensor for fast tracking and navigation robots. *Mach. Vis. Appl.* 10, 1–8 (1997).
51. Mao W Error and adjustment of reflecting prisms. *Opt. Eng.* 36, 3367–3372 (1997).
52. Sullivan DL Alignment of Rotational Prisms. *Appl. Opt.* 11, 2028–2032 (1972). [PubMed: 20119276]
53. David Giese J, Ford TN & Mertz J Fast volumetric phase-gradient imaging in thick samples. *Opt. Express* 22, 1152–1162 (2014). [PubMed: 24515075]

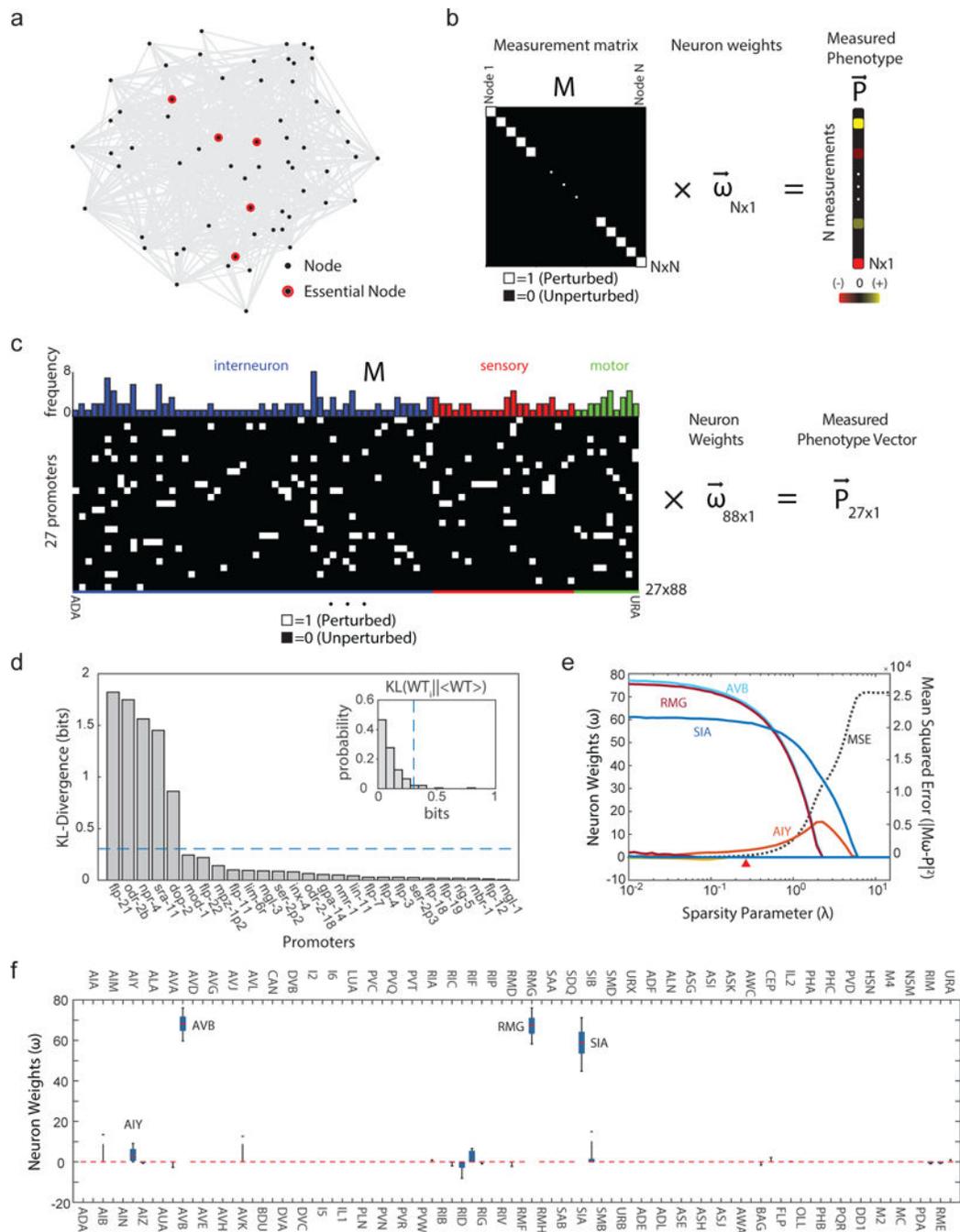


Fig. 1]. Compressed sensing framework allows identification of essential neurons controlling speed in *C. elegans*.

a, A network of nodes (black dots) with a small number of essential nodes for a phenotype (red), synaptic connections (grey lines). A node is either individual neuron or neuron subtypes. **b**, Identifying essential neurons by perturbing nodes one at a time is equivalent to solving matrix equation with $N \times N$ diagonal measurement matrix (diagonal entries of one for the perturbed neuron, zero otherwise) multiplying unknown neuron weight vector $\vec{\omega}$ set equal to measured phenotype vector \vec{P} . The number of measurements N linearly increases

with the size of the network. **c**, 27×88 measurement matrix \mathbf{M} generated from 27 transgenic lines expressing Archaeorhodopsin-3 under the control of non-specific promoters. Rows are promoters, columns are neuron types. Matrix has an entry 1 (white) if the promoter drives expression in that neuron type, else 0 (black). Bars above the measurement matrix shows number of times each neuron appears in the matrix (interneurons: blue, sensory neurons: green, motor neurons: red). Finding key neurons from 88 types using just 27 measurements requires solving underdetermined set of linear equations $\mathbf{M}\vec{\omega} = \vec{P}$. **d**, KL divergence between speed distributions of control and Arch line placed in descending order. Inset: variation of KL divergence in wild type replicates, blue dashed lines: 95th percentile of the distribution. **e**, Median neuron weights from 10,000 lasso regression solutions through bootstrapping for sparsity parameter spanning three orders of magnitude (see Methods). Red arrow: sparsity parameter picked for **(f)**. Dashed line: mean squared error $\chi^2 = (\mathbf{M}\vec{\omega} - \vec{P})^2$ of optimized solution for given sparsity parameters. **f**, Neuron weights distribution for each neuron types for sparsity parameter shown in **(e)**. Red line: median, box: 25 and 75, whiskers: 5 and 95 percentiles. The neurons arranged in the same order in **(c)**.

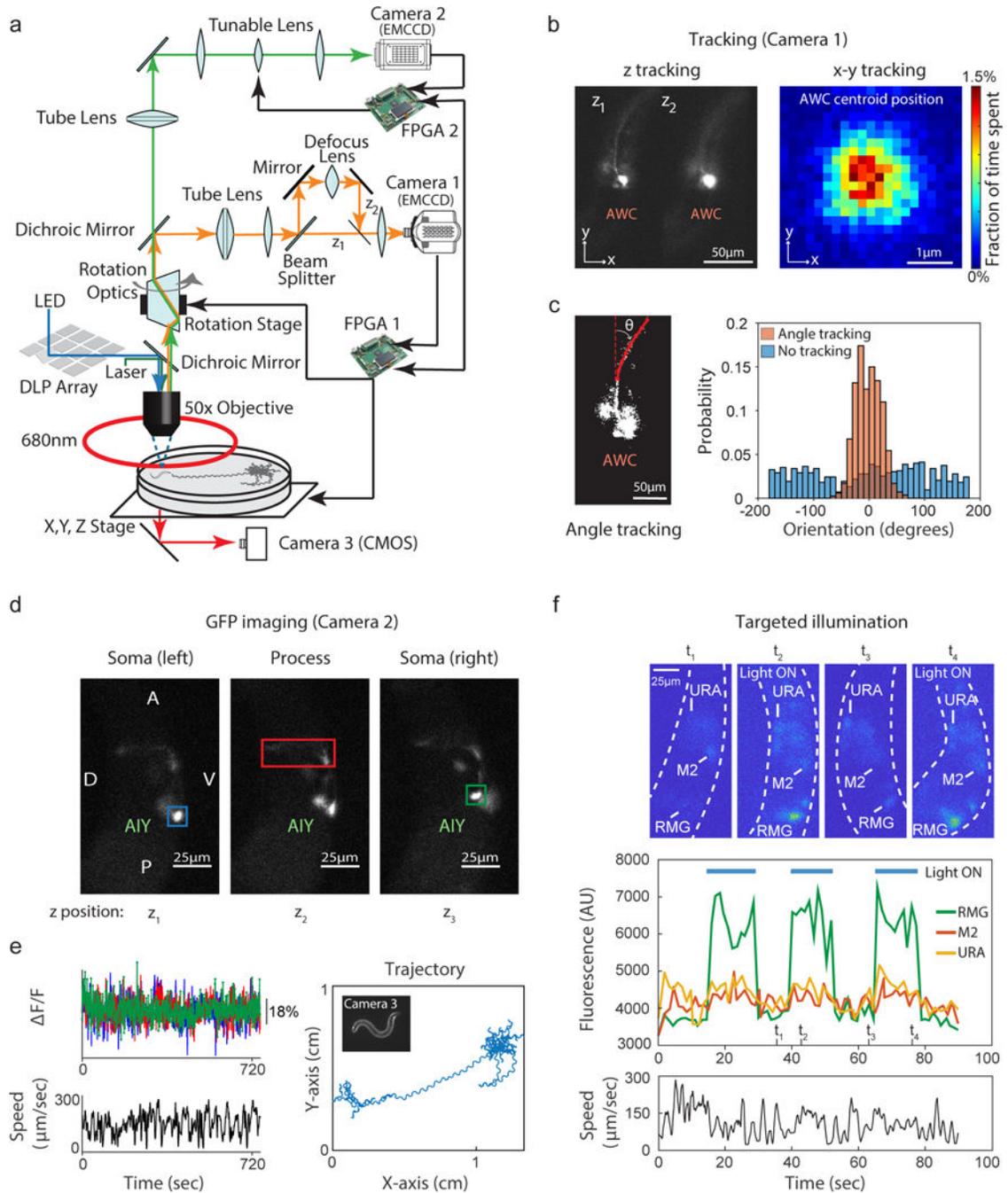


Fig. 2|. Stabilization microscope for long-time calcium imaging and optogenetic manipulation of freely moving animals.

a. Schematic of tracking and image stabilization microscope (orange: mKO emission, light green: GCaMP emission, blue: GCaMP excitation, dark green: mKO excitation). Data from Camera 1 is processed by FPGA1 to control x,y,z and rotation stages for simultaneous tracking. Camera 2 acquires GCaMP imaging data via FPGA2 that also controls tunable lens for z scan. Camera 3: low magnification animal behavior. DLP mirror array, controlled by a PC, is used to target specific neurons. **b.** Left, images of marker neuron AWC^{ON} at two z

planes in freely moving animal for z-tracking. Right, heat map of percentage of time spent by center of mass of marker neuron at each pixel for animal in (a) shows $1\mu\text{m}$ x and y tracking accuracy. **c**, Left, AWC^{ON} marker neuron thresholded image (from Camera 1) to track rotation angle (θ). Right, histogram of θ with (orange) and without (blue) stabilization using rotation optics. **d**, AIY::GFP images (Camera 2) at different z-positions through z-scan using liquid lens. Left to right: AIYL (blue), AIY process (red), AIYR (green). **e**, Quantification of noise. Left top, intensity data (AIYL: blue, AIYR: green, Process: red all expressing GFP) from freely moving animal shows 18% fluctuations in GFP fluorescence. Fluorescence signal changes above 18% from GCAMP determined statistically significant. Left bottom, speed of animal moving freely on agar plate during the experiment. Right, trajectory of animal. Inset image from Camera 3. **f**, Control for targeted illumination of neuron of interest. Top, *Pflp-21::GFP* images (Camera 2) at four time points with targeted light on or off. Middle, GFP intensities from RMG, M2 and URA neurons. Blue bar: the patterned light on. Bottom, speed of animal during targeted illumination.

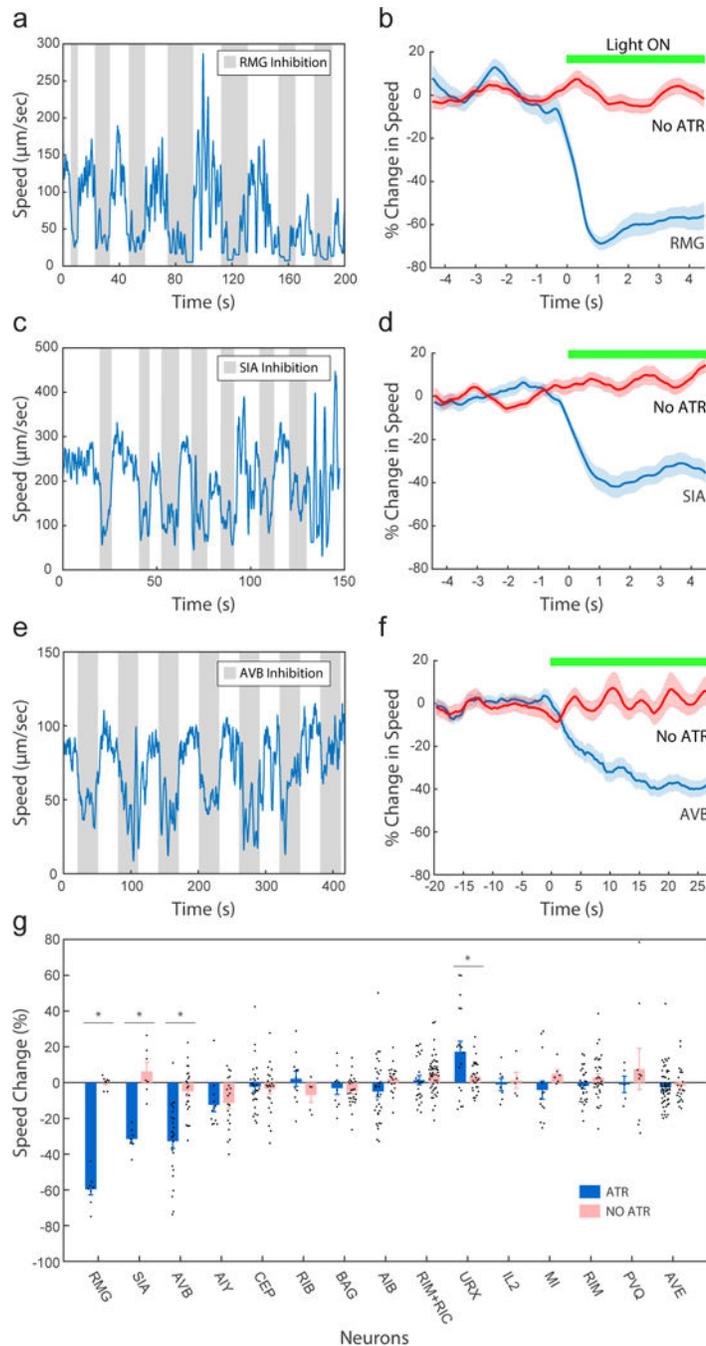


Fig. 3]. Targeted inhibition of single neurons SIA, RMG and AVB reduces the speed of locomotion.

a, Time trace of speed with targeted RMG inhibition using on-off 540nm light cycle. Gray bars: light on. **b**, Percentage speed change upon optical inhibition. Speed trace normalized to the light-off speed as light switches from off to on, inhibiting RMG (n= 71 ATR+, 86 ATR- cycles). Green bars: light on. **c,d**, SIA (n= 47 ATR+, 54 ATR- cycles). **e,f**, AVB (n= 153 ATR+, 122 ATR- cycles). **b,d,f**, Data mean (solid line) \pm s.e.m (shaded area). **g**, Percentage of speed change upon targeted inhibition of specified neurons (All n values in

Supplementary Table 3). Blue bars: animals fed with ATR; pink bars: NO ATR controls. Data shown as mean \pm s.e.m. (* RMG: $p=5.36E-11$, SIA: $p=2.23E-5$, AVB: $p=1.26E-6$, URX: $p=0.01$, two-sample, two-sided t-test). p-values for other neurons are bigger than 0.05 (see Supplementary Table 3).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

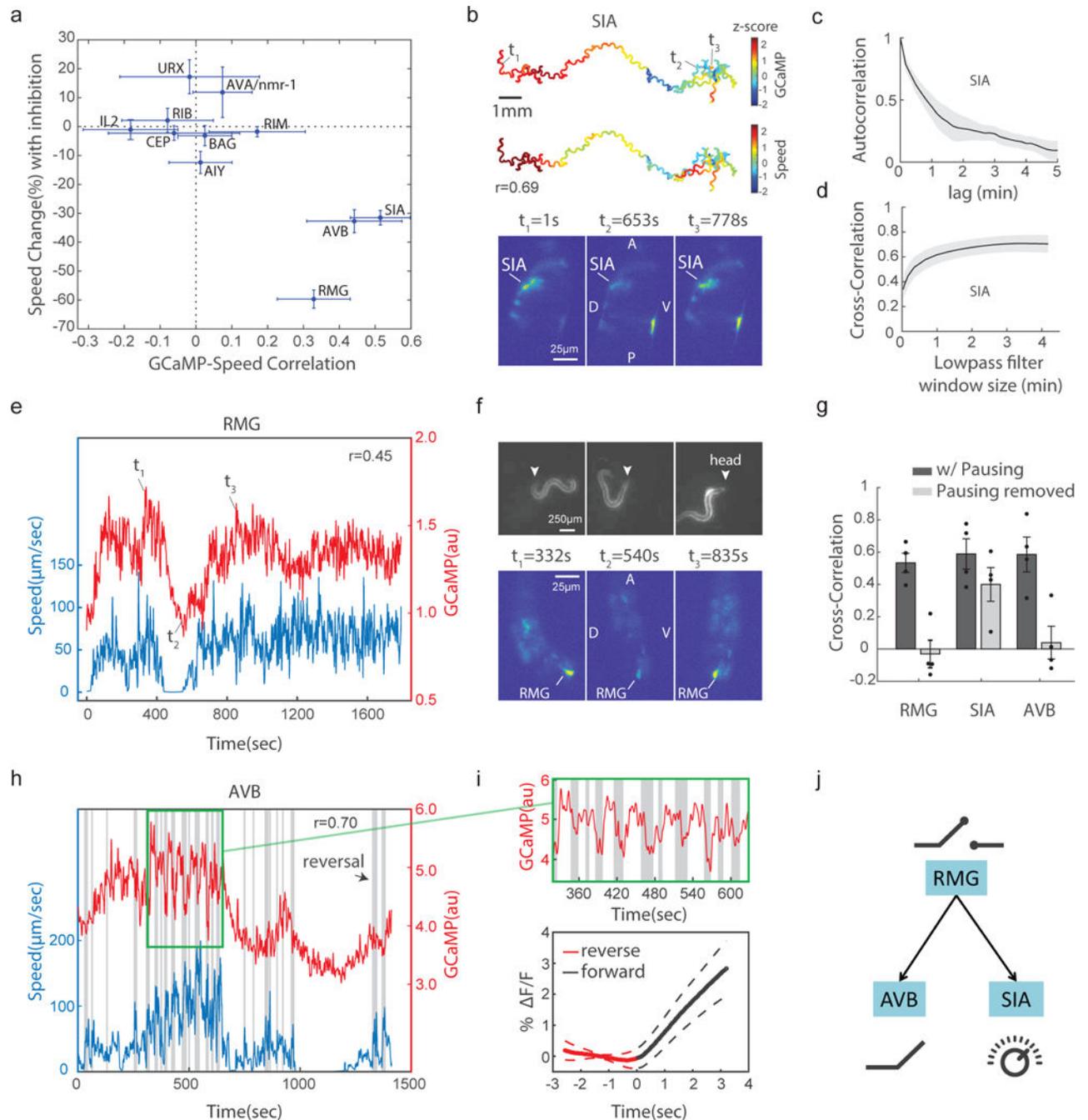


Fig. 4. Calcium imaging of SIA, RMG and AVB reveals distinct roles in controlling speed.
a. Cross-correlation between calcium imaging and speed versus the speed change upon inhibition for neuron types: RMG(n=10), SIA(n=9) AVB(n=7), AIY(n=7), CEP(n=5), RIM(n=9), RIB(n=7), IL2(n=5), URX(n=4), BAG(n=5), AVA(n=5) where n is the number of animals imaged for each neuron type. **b.** Trajectory of a worm color coded by SIA calcium activity (above) and the speed of the animal (below). High mag images (from camera 2 in Fig. 2) shown at three time points marked on the trajectory. **c.** Autocorrelation function for SIA activity shows the correlation time of ~1 min (n=9). **d.** Cross-correlation between speed

and SIA activity with low pass filter (n=9). **e**, Time trace of RMG activity (red) and speed (blue) for a typical animal. **f**, Neuron intensity images and worm body posture images at three different time points from RMG experiment in (**e**) showing of activity in RMG and corresponding snapshot of the animal. **g**, Cross-correlation between RMG, SIA, AVB (n=4 for each) and speed with pausing and after pausing regions are removed. **h**, Trace of AVB GCaMP (red) and speed (blue) for a typical freely behaving animal. AVB activity correlates both with pausing events and forward motion onset. Gray bars indicate reversal events. **i**, Top, zoom-in to show AVB activity increase with forward motion initiation. Bottom, quantification of the increase in AVB activity with forward motion onset over reversal cycles (n=63 reversals). Time zero is the end of reversal and beginning of forward motion. **a,c,d,g,i**, Data are shown as mean \pm s.e.m. **j**, RMG-SIA-AVB circuit controls speed. RMG synapses onto both AVB and SIA. RMG acts as switch controlling pausing, AVB acts as a rectifier controlling forward motion and SIA controls the value the speed continuously.