Synaptic polarity of the interneuron circuit controlling *C. elegans* locomotion

Franciszek Rakowski[a], Jagan Srinivasan[b], Paul W. Sternberg[b], and Jan Karbowski[c,d,*]

[a] *Interdisciplinary Center for Mathematical and Computational Modeling, University of Warsaw, 02-106 Warsaw, Poland*

[b] *California Institute of Technology, Division of Biology, Pasadena, CA 91125, USA*

[c] *Institute of Applied Mathematics and Mechanics, University of Warsaw, 02-097 Warsaw, Poland*

[d] *Institute of Biocybernetics and Biomedical Engineering, Polish Academy of Sciences, 02-109 Warsaw, Poland*

Abstract

C. elegans is the only animal for which a detailed neural connectivity diagram has been constructed. However, synaptic polarities in this diagram, and thus, circuit functions are largely unknown. Here, we deciphered the likely polarities of 7 pre-motor neurons implicated in the control of worm's locomotion, using a combination of experimental and computational tools. We performed single and multiple laser ablations in the locomotor interneuron circuit and recorded times the worms spent in forward and backward locomotion. We constructed a theoretical model of the locomotor circuit and searched its all possible synaptic polarity combinations in order to find the best match to the timing data. The optimal solution is when either all or most of the interneurons are inhibitory, which suggests that inhibition governs the dynamics of the locomotor interneuron circuit. From the five pre-motor interneurons, only AVB and AVD are equally likely to be excitatory. The method used here has a general character and thus can be also applied to other neural systems consisting of small functional networks.

Keywords: C. elegans; Locomotory interneurons; Synaptic polarity; Locomotion; Neural circuit; Laser ablations.

∗ Corresponding author at: jkarbowski@duch.mimuw.edu.pl

Introduction

Caenorhabditis elegans nematode worms possess a very small nervous system composed of only 302 neurons connected by about 5000 chemical synapses and 3000 gap junctions [1]. Because of its smallness a precise map of neural connections was possible to construct [1, 2]. This places *C. elegans* in a unique position among all other animals [3], for which we have at best only rudimentary connectivity data to test various concepts regarding neural wiring and function [2, 4, 5, 6, 7, 8]. However, despite this achievement we still have a very limited knowledge about the nature of most of the worm's synaptic connections, i.e. whether they are excitatory or inhibitory.

The simplicity of the *C. elegans* nervous system does not preclude these worms from executing various non-trivial behaviors such as locomotion, feeding, mating, chemotaxis, etc [9, 10]. To understand the neural basis of these behaviors requires some information not only about the pattern and strength of the connections but also about the type of their synapses. The same neural circuit can perform different functions depending on the signs of synaptic polarities it contains. Specifically, circuits in which excitatory synapses dominate can sometimes become epileptic. On the other hand, networks with only inhibitory connections could be silent, and therefore in many situations useless. Thus, it may seem that some sort of an intermediate regime is necessary for a proper functioning of the nervous system [11]. For example, it was proposed that mammalian cortical networks operate in a dynamic state in which excitation is effectively balanced by inhibition [12, 13], although anatomical number of excitatory connections dominates over inhibitory in the cerebral cortex [14]. For nematode worms, a similar issue has been

addressed only sporadically. On a modeling level, in the context of a tap withdrawal circuit [15], and experimentally, studying genes that influence the ratio of excitatory to inhibitory signaling [16]. We think that this topic deserves more attention both theoretical and experimental, if we are to understand the functioning of worm's circuits [17, 18, 19].

Movement direction in *C. elegans* is governed by 5 distinct locomotory command interneurons (AVB, PVC, AVA, AVD, AVE), each in two copies (left and right). All of these 10 interneurons directly connect a downstream group of dorsal and ventral body wall excitatory motor neurons [20]. The topology of connections between the command interneurons is well known [1, 2], however, their synaptic polarities are not. Conventional thinking is that AVB and PVC control forward motion, while AVA, AVD, AVE control backward motion [20]. This reasoning is based on the fact that the former interneurons connect mainly motor neurons of type B (experimentally shown to be critical for forward locomotion [21]), whereas the latter interneurons connect mostly type A motor neurons (ruling the backward motion [21]). However, this simple locomotory picture, relying on a single neuron function doctrine may turn out to be too simplistic. Indeed, many laser ablation experiments show that removal of both AVB and PVC reduces forward motion, but does not abolish it completely (see below). Similarly, worms lacking the "backward" interneurons AVA, AVD, and AVE exhibit a comparable frequency of reversals as intact wild type worms [22]. Moreover, the major backward interneuron AVA makes also connections (both synaptic and gap junctions) with the forward B motor neurons [2]. Thus, perhaps the decision to move in a particular direc-

tion is generated by a collective activity of all command interneurons, rather than by an activity of a particular interneuron or a particular connection.

Our aim is to investigate the problem of decision making for movement direction in *C. elegans* on the level of its interneuron network. The main question we pose is how two antagonistic behaviors, i.e. forward and backward motions, can be controlled by the same circuit of mutually coupled pre-motor interneurons. A strictly related to this question is the problem of synaptic polarities of these interneurons. Specifically, by applying structural perturbations to the circuit we want to determine, using mathematical modeling, which combination of synaptic polarities gives the best match to the experimentally observed locomotory output of *C. elegans*. This knowledge allows us to answer a question about a relative influence of inhibition and excitation in the command interneuron circuit. Moreover, this approach provides an insight about a degree of interneuron collectiveness in choosing the direction of motion.

Results

The command interneuron circuit for *C. elegans* locomotion

To simplify data analysis and mathematical modeling we grouped the left and right members of each locomotor circuit neuron as one model neuron. Thus, in our circuit controlling worm's motion there are 5 command interneurons, one distinguished upstream polymodal sensory interneuron called ASH, and a modulatory neuron DVA (Fig. 1A). A recent study [23] indicates that DVA plays a role of a sensory neuron in locomotion. We included this neuron explicitly in the circuit, since it has direct connections with body wall motor neurons [2], similar to 5 command interneurons. Because of this similarity, we want to investigate whether DVA can serve additionally as a command interneuron.

The neurons in the circuit are connected either by chemical synapses (of unknown polarity) or by gap junctions. The strength of the connection between two arbitrary neurons is proportional to the number of anatomical contacts between them determined from the empirical data [2]. Additionally, each pre-motor interneuron receives an input from upstream (mostly head) neurons, which can be either weak or strong (this is variable in the model). The equations describing activities of all neurons are presented in Materials and Methods.

In an intact circuit for wild type (WT) worms, there is some average typical distribution of interneuron activities that makes worms to move more frequently forward than backward (roughly in 3 to 1 proportion). We wanted to perturb the system and investigate its corresponding output by performing laser ablations of selected neurons in

the locomotory circuit. We reasoned that a gradual removal of neurons from this circuit (Fig. 1B; see Materials and Methods) would not only affect its physical structure, but also would redistribute the remaining neurons activities, which in turn, should modify the worm's locomotory behavior. In particular, the ratio of times spent in forward and backward motion should change, and this should correlate with the changes in the activity output of this circuit. Thus, associating experimental times of forward and backward locomotion with theoretical outputs of the circuit model for every ablation type can allow us to predict synaptic polarities of the interneurons.

Experimental results

We performed single, double, triple, and quadruple ablations in the interneuron circuit. In total, we generated 17 types of ablation and recorded corresponding mean times the worms spent in forward (T_f) and backward (T_b) motion, as well as in stopped phase (T_s). These experimental results are presented in Table 1. From all the ablations executed, only extrusion of ASH and PVC neurons increase the time spent in forward locomotion in relation to wild type. This is an indication that these particular interneurons have a definitely negative influence on the forward direction, and for that reason are likely to be inhibitory (see below). All other removals have a detrimental effect on the duration of forward motion, even those traditionally associated with backward motion (AVA, AVD). In particular, the ablation of AVA has the most dramatic effect on T_f, leading to its 10 fold reduction in comparison to wild type, although the reversal frequency increases only mildly by a factor of 2. Moreover, the AVA ablated worms, including their multiple ablations, spent a lot of time not moving (stopped phase), far

more than WT and worms with other types of ablation. For instance, for the combined ablation ASH+AVA+AVB, we find the largest stopped mean time of 1.16 sec.

Worms with multiple ablations reverse roughly as frequently as worms with single ablations (Table 1). Generally, the frequency of direction reversals does not correlate well with the average time worms spent in forward motion (Table 1). For example, worms with killed ASH reverse approximately as often as worms with removed AVD, despite the fact that ASH worms spent 3 times more time in forward motion. Similarly, worms with almost equal T_f (about 0.9 sec), i.e. AVB+PVC and AVA+AVB+PVC, differ in reversals by a factor of 2.5.

An interesting result is that ablating the modulatory neuron DVA causes a sharp decline in the forward motion timing in comparison to WT, and a two-fold increase in reversals (Table 1). Also, the combined ablations of DVA with PVC and AVB show a similar property. This clearly suggests that this neuron has a significant influence on the interneuron circuit output, which could be more than just its sensory modulation.

Theoretical results

In our circuit model there are 7 neurons (5 pre-motor interneurons, DVA, and ASH), each of them can be either excitatory or inhibitory. Thus, there are $2^7 = 128$ possible copies of the circuit associated with synaptic polarities, i.e. the sign of ϵ_i (for $i = 1, ..., 7$). Two examples of the polarity copies are: (i) $\epsilon_{ASH} = -1$, $\epsilon_{AVB} = -1$, $\epsilon_{PVC} = -1$, $\epsilon_{DVA} = -1$, $\epsilon_{AVA} = -1$, $\epsilon_{AVD} = -1$, $\epsilon_{AVE} = -1$, and (ii) $\epsilon_{ASH} = 1$, $\epsilon_{AVB} = 1$, $\epsilon_{PVC} = 1$, $\epsilon_{DVA} = 1$, $\epsilon_{AVA} = 1$, $\epsilon_{AVD} = 1$, $\epsilon_{AVE} = 1$, which correspond to all inhibitory and all excitatory neurons, respectively. Additionally, each of the 6 pre-motor neurons

(excluding ASH) receives an upstream sensory input coming mostly form the head, except for PVC for which it comes predominantly from the tail. The sensory input is represented by a binary variable z_i that can have two values: 0 (weak input) and 1 (strong input). Consequently, every polarity copy can be found in additional $2^6 = 64$ activity configurations. This implies that, in total, our circuit model has $2^7 \cdot 2^6 = 8192$ distinct polarity-activity configurations.

For each possible configuration (i.e. synaptic polarity and an upstream input) of the circuit we performed 17 "computer ablations" analogous to the experimental ablations shown in Table 1, by setting $\epsilon_i = 0$ if the neuron with an index i was removed. This procedure removes all the connections (synaptic and electric) coming out of this ablated neuron. For each ablation we computed the fraction of time corresponding to forward motion, i.e. $T_f/(T_f + T_b)$. Thus, we generated 18 theoretical fractional times associated with every circuit configuration (17 types of ablation plus WT) and computed their Euclidean distance (ED) to the experimental values in Table 1 (see Materials and Methods for a more detailed description). The configuration with the smallest ED value corresponds to the optimal solution that predicts synaptic polarities and the pattern of the upstream input. Our strategy was to vary the level of this input, and for each level to find optimal values of synaptic and gap junction conductances (q_s, q_e), together with the sensory noise amplitude (η), which give the best fit of the theoretical $T_f/(T_f + T_b)$ to the experimental data. We noticed that we obtained better fits if we allowed the interneurons to receive a heterogeneous input from the upstream neurons.

We find that the best fit to the experimental data is achieved in the case when

all 7 neurons in the circuit are inhibitory (Tables 2-4). The winning synaptic polarity configuration is associated with the combination number 1, which gives the best (lowest) ED= 0.363, and the largest correlation with the data points, which is 0.743 (Table 2). The distribution of winning values of the upstream input z_i is nonhomogeneous, and nonzero only for AVB and PVC neurons, implying that much sensory excitation comes to the neurons controlling directly forward motion (Table 2). In Fig. 2 we display a comparison of theoretical and experimental values of $T_f/(T_f + T_b)$ across different ablations for this winning combination of synaptic polarities. In general, there is a good fit of the theoretical points to the data (correlation about 0.74).

The second place among all synaptic polarities is taken by the combination # 17, for which all neurons, except AVB, are inhibitory. This combination has ED value very close to that obtained by the winning combination # 1 (Table 2). Moreover, these two combinations appear the most often among the winners, also for other choices of parameters describing the sensory input (Tables 2-5). The third place is taken by a combination with the number 11, in which only two neurons are excitatory: AVD and DVA. This combination also appears quite often among the leading polarities, but its ED is in some separation from the two winning synaptic combinations 1 and 17.

For a comparison, the worst synaptic configurations had ED value at about 2.0, and were dominated by excitatory connections. The general trend is such that the lowest ED values are associated with mostly inhibitory circuits, while the highest ED's are related to mostly excitatory circuits.

To quantify the likelihood of a given synaptic polarity among leading combinations

we associate with each neuron a probability that it is inhibitory, for each input value σ and ASH activity level (see last columns in Tables 2-5). This probability is defined here as a fraction of times the $\epsilon = -1$ appears in the row for each neuron. One can notice that some of the locomotory interneurons, such as AVE, AVA and PVC, are inhibitory with probabilities that are close to 1, regardless of the input magnitude. The polymodal neuron ASH is also in this category. For the rest of the pre-motor interneurons these probabilities are not that high, but nevertheless are at least ≥ 0.5. We also computed an average probability that a given neuron is inhibitory across different input levels coming to the locomotory circuit (average values for all Tables 2-5). These average probabilities are: 1 for ASH, 0.875 for AVA, 0.5 for AVB, 0.5 for AVD, 0.938 for AVE, 0.656 for DVA, and 0.719 for PVC. Thus, from the whole group of 7 neurons investigated here and implicated in locomotion control, only AVB and AVD have about equal chances to be excitatory and inhibitory.

One might wonder if these probability estimates hold if we include more winning combinations, not just eight as in Tables 2-5. Including 20 leading combinations for the winning input values $\sigma = 8$ mV and $\kappa = 0.6$ (corresponding to Table 2), gives qualitatively a similar picture. Again, the neurons AVB and AVD have the highest likelihood of being excitatory, although inhibitory polarities of these interneurons have the lowest ED values (Fig. 3). Taken together, the above results strongly suggest that the majority of interneuron connections are inhibitory.

The Euclidean distance ED between theoretical and experimental fractions of time spent in forward locomotion depends in a non-monotonic manner on model free param-

eters $(\eta, \sigma, \kappa, q_s, q_e)$. In Fig. 4 we show ED as a function of sensory noise η. Either too strong or too weak sensory noise increases ED, and there is an optimal value of η for which ED has a minimum. A similar relationship emerges also for the rest of the free parameters. Typically, the parameters controlling the strength of the upstream input, i.e. σ and κ, should be in some intermediate range to reach a minimal value of ED. For instance, the winning polarity combinations for $\sigma = 4$ and $\sigma = 12$ mV have larger ED than that for $\sigma = 6$ or 8 mV (Tables 2-5). The same is true for the value of κ, characterizing ASH activity (see Materials and Methods), with the optimal $\kappa = 0.6$ for each σ level.

Discussion

Using a combination of experimental (laser ablations) and computational (circuit model and optimization) tools we were able to decipher the likely synaptic signs of interneurons composing the small network commanding *C. elegans* locomotion [1, 20]. It turns out that probably most of these neurons, i.e. synapses they sent, are inhibitory. In particular, the average probabilities that a particular interneuron is inhibitory are: 0.875 for AVA, 0.5 for AVB, 0.5 for AVD, 0.938 for AVE, 0.719 for PVC, and 0.656 for DVA.

Because of a suppressing nature, the command interneuron circuit must receive a sufficient amount of excitation from upstream (in large part sensory) neurons to be functional, i.e. to appropriately activate downstream motor neurons. Our computational results indicate that the behavioral data are best explained if the command circuit receives a mixed, heterogeneous input (denoted by X_α in our model equations). The best fit to the data is obtained if the largest excitation comes to forward interneurons AVB and PVC (Table 2). In this sense, the existence of sensory stimulation is an important factor for directional motion generation, which is in general agreement with the experimental findings [22, 24].

How are downstream motor neurons activated given an inhibitory nature of synapses in the pre-motor interneuron circuit? Recent calcium imaging studies show that AVA is active during backward locomotion [25, 26]. How can one explain this? The likely answer lies in the strong gap junction connections between AVA and A motor neurons, and between AVB and B motor neurons (Table 6). Specifically, during backward

movement AVA, due to its large sensory input, synaptically inhibits other interneurons including AVB, but at the same time excite downstream A motor neurons via strong electric coupling. Thus AVB sends less excitation to B motor neurons via its strong gap junctions than does AVA to A neurons. Consequently, the activity of A prevails over activity of B neurons (i.e. $E_b > E_f$ in our model), and the worm moves backward, even with inhibitory synapses in the locomotor circuit.

Asymmetry in forward and backward interneuron activities determine the likely direction of motion.

Our results indicate that a decision to move in a particular direction can be made on a small circuit level composed of the 6 pre-motor interneurons (including DVA). Specifically, the output from these interneurons is fed to the two types of body wall motor neurons (B and A) controlling forward and backward motions, whose relative asymmetric activities (E_f and E_b in our model) determine the likely direction of worm's movement. This simple decision making mechanism can explain 74 % of the correlations between the experimental data and computational results (see Table 2). Moreover, this behavioral picture is consistent with the findings in a recent experimental study [28], in which it is shown that the imbalance between activities of A and B motor neurons is a likely scenario for the selection of worm's motion direction. However, in contrast to these authors the imbalance between A and B neurons in our model is caused not so much by a strong AVA-A gap junction coupling, but by the asymmetric upstream excitatory input to command interneurons (in the winning combinations, AVB and PVC neurons receive much stronger upstream excitation than the rest of the circuit).

Qualitative interpretation of ablation data.

The interesting experimental result concerning single-neuron ablations is that removal of certain interneurons causes an increase in forward motion timing, while removal of others leads to its dramatic decrease. Specifically, only killings of ASH and PVC neurons increase significantly the time T_f spent in forward locomotion. In relation to ASH, this suggests an important role of the sensory input. There are two surprises here. First, PVC was thought as promoting forward locomotion [20]. Second, given that the sensory ASH neuron makes only weak or intermediate synaptic connections with the command interneurons (with all "backward" interneurons and AVB; see Table 6), it should not have such a strong influence on the motion characteristics. The solution of these puzzles is that PVC and ASH are probably inhibitory, and in addition ASH should be highly depolarized in order to significantly downregulate the locomotory (mostly backward) interneurons via its synapses. The prediction that ASH is inhibitory is in contrast to an experimental finding, using calcium imaging technique, that ASH to AVA synapse is likely excitatory [26]. To resolve this discrepancy we note that in our model the upstream sensory input (X_α) is allowed to be only excitatory. However, it is also possible that it can have an inhibitory component, which was not included explicitly in our model. This can sometimes cause wrong predictions regarding the synapses associated with ASH.

Removal of AVA interneuron causes a large reduction in T_f, despite the fact that this neuron belongs to the "backward" locomotory circuit, and one might naively expect that it effectively prohibits forward motion. Moreover, single and multiple ablations

associated with AVA cause an increase in stopped time (Table 1). This suggests that removal of AVA decreases a difference between activities of forward and backward motor neurons, i.e. $E_f - E_b$ may become comparable with a threshold for movement initiation Δ (see Materials and Methods). This in turn may suggest that when AVA is absent, backward motor neurons are more active. Taken together, these results imply that the overall synaptic effect of AVA is most likely inhibitory.

The ablation results for the DVA neuron indicate that it plays a more significant role in the locomotory circuit than just its passive modulation. From Table 1 it is evident that killing DVA has one of the biggest impacts on T_f. Based on this, we hypothesize that DVA might serve also as a command neuron in generation of forward locomotion, which is a novel function for this neuron.

The case with multiple ablations is more complicated. These type of ablations do not have an additive property, i.e. removal of more neurons does not necessarily lead to a progressive drop in the forward motion timings. For example, double AVB and PVC ablation has $T_f = 0.91$ s, but additional removal of AVD actually increases T_f to 1.33 s (Table 1). The latter may seem paradoxical, however one has to remember that backward neurons do not act in isolation, but participate in the whole interneuron network activity, and thus indirectly also influence the forward motor neurons. Apart from that, interneurons interact among themselves both synaptically (nonlinear in nature) and via gap junctions (bidirectional in nature). This may additionally mask a single interneuron contribution to the locomotory output of the circuit. As a result it is very difficult to predict in advance the effect of any particular ablation on worm's locomotory

characteristics in the case of multiple ablations. For this, one needs to perform detailed computations on a network level, as was executed in this study.

Collective, mutually inhibitory interactions between command interneurons underlie *C. elegans* direction of locomotion.

These ablation results suggest that a picture in which a single neuron or a single connection control a specific behavior, advocated in several former studies [18, 20], may be oversimplified. Instead, our findings support the idea that behavioral (locomotory) output depends to a large degree on a collective activity of neurons comprising the "functional circuit" [24]. That is, the same neuron can participate to some extent in opposite behaviors. Obviously, some neurons or connections in the functional circuits may be more dominant than others for a particular behavior, but the presence or absence of a particular neuron in the circuit is generally not critical for its operation. This partial redundancy in neural function is probably evolutionary driven to ensure a robust circuit performance.

Similarly, none of the interneuron ablations, either single or multiple, abolishes the worm's movement or body oscillations completely. This clearly indicates that none of the interneurons alone is a Central Pattern Generator, which again speaks in favor of collective rather than individual interneuron dynamics as a determinant of locomotion.

Our main result that the pre-motor interneuron circuit has mainly inhibitory synapses is qualitatively similar to two earlier [15, 24] and one recent [27] study. Wicks et al [15] investigated computationally a tap withdrawal circuit in *C. elegans* and predicted that most interneurons composing it were inhibitory [15]. That study concluded that PVC

and AVD interneurons were probably excitatory. Our results suggest that AVD is equally likely to be inhibitory as excitatory, whereas PVC with high probability should be inhibitory. The possible sources of the discrepancy can be that Wicks et al [15] used an older incomplete connectivity data for the pre-motor interneurons [1], and did not include AVE neuron in their analysis. In another study, Zheng et al [24] hypothesized that the locomotory interneuron circuit should act as an inhibitory switch in order to explain qualitatively data on motion direction transitions. A recent experimental work [27] also suggests that the pre-motor interneurons should use inhibition as a main synaptic signaling.

An interesting question is which neurotransmitters mediate inhibitory interactions between interneurons. The most likely neurotransmitter between interneurons is glutamate. In mammalian brains, it is known to be exclusively an excitatory signal, since AMPA and NMDA postsynaptic receptors conduct mostly Na^+ and K^+ with an effective reversal potential around 0 mV. However, in *C. elegans* the situation is more complicated because locomotor interneurons contain apart from these receptors, also GluCl postsynaptic receptors [29]. These channels are gated by Cl^- (with large negative reversal potentials), and therefore mediate inhibition to postsynaptic cells [29]. Specifically, the currents associated with GluCl receptor have been observed in the AVA interneuron [30], and they may also exist in other interneurons.

Given these two types of postsynaptic receptors, it is possible that a single interneuron can have both excitatory and inhibitory synapses on distinct postsynaptic targets. In this case, the synaptic polarities associated with each neuron in our study have an

average character. More precisely, the determined probabilities that a given neuron is inhibitory are the fractions of inhibitory connections that the neuron makes with other postsynaptic neurons. Thus, for example, the inhibitory probability 0.5 found for AVB indicates that this neuron sends out about 50% inhibitory and 50% excitatory synapses.

Our computational model and its extension.

The theoretical approach in this paper blends a traditional neural network modeling with a probabilistic method for relating network activity to behavioral data. In particular, we envision the nematode locomotion as a three state system, one state for forward, second for backward motion, and third state for no motion. In this system there are transitions between the states that are caused by intrinsic relative activities of A and B motor neurons (E_b and E_f), as well as by a "sensory noise" (η in Eq. 20). Concerning the stochastic component, our approach is partly related to another phenomenological approach used in [31] for searching neuronal connectivity motifs governing spatial orientation. Note that many ablations in Table 1 have the ratio T_f/T_b close to unity, which in terms of our model implies that $(E_f - E_b)/\eta \ll 1$, i.e. for these ablations a stochastic influence of the environment is bigger than the relative activities of A and B motor neurons. This is interesting and shows that sensory noise gains in importance as we remove more neurons from our circuit. This may also suggest that some of the interneurons act as filters for the environmental noise.

Our combined approach allows us to achieve a concrete goal, which is the prediction of synaptic polarities for the well defined locomotory interneuron circuit. The knowledge of the probable synaptic polarities of the command interneurons may have a positive

impact on future modeling studies of *C. elegans* locomotion. We hope, that this will enable more realistic simulations of the neural dynamics that can extend the scope of testable predictions of the current locomotory models [32, 33, 34]. Our method of determining synaptic polarities, which combines structural perturbations with the computational modeling, is sufficiently general that can be also applied to other small functional neural systems in which synaptic polarities are unknown. However, it is important to keep in mind that our model, as every model describing biology, is subject to several assumptions (see the list in the Methods), and clearly has some limitations. The model does not include several subtle neurophysiological features. For example, a possibility that an individual neuron might release multiple neurotransmitters of similar importance, or that neuromodulators might provide an extra synaptic input throughout the network.

Materials and methods

The ethics statements does not apply to this study.

Collection of experimental data

Strain maintenance. For our automated locomotion analysis, we cultured *C. elegans* at 20 ^{o}C on NGM plates seeded with the OP50 using standard methods [35].

Automated worm tracking and data extraction. Worms tested by automated tracking were continuously cultured on E. coli OP50. For assaying various parameters of locomotion, 10 cm non-seeded NGM plates were used. These NGM plates used for recordings were equilibrated to 20 ^{o}C for 18-20 hours. After ablations of individual neurons, the worms were placed on plates with *E.coli* as a food source to recover. Ablated worms and mock controls were tested within 72 hours of the L4 molt. They were then transferred to assay plates containing no food. After 5 minutes of acclimatization on these plates, the worms were video taped for 5 minutes. Data presented in this paper represent the locomotory behavior of worms when they were exhibiting "area restricted search behavior". Data extraction and processing was done using image processing and analysis software as previously described [36]. From each video recording of 5 minutes, we used the middle 4 minutes, and used the software to derive values for times in forward and backward locomotion, as well as reversal frequencies. In our software, we used a velocity threshold for motion detection. Specifically, if the nematode velocity was below 0.05 mm/sec, we classified this as stopped time or "no motion". Every change of velocity direction that was above this value was classified as a reversal. The average time spent in the stopped phase is T_s. The time T_f the worm spent in forward motion

is defined as an average duration of time counted from a moment of moving forward to stopping. Similarly, the time T_b spent in backward motion is an average of times from the initiation of backward movement to stopping. Generally, because of many reversals, the sum of the times T_f, T_b, and T_s is much smaller than the recording time of 4 minutes. The numerical values of T_f, T_b, T_s provided in this paper are population averages. All incubations and recordings were done in a constant temperature room at 20 ^{o}C.

Laser ablations. For all species tested, we used the L1 larva stage for our ablations.

Description of the command interneuron circuit model

List of the assumptions used to construct the model. We make the following major assumptions in the theoretical model:

1) In the interneuron circuit left and right members of each interneuron are grouped as one interneuron.

2) Synaptic and gap junction strengths between any two neurons are proportional to the anatomical number of synapses and gap junctions between them.

3) Pre-motor neurons do not generate sodium-type action potentials but their activities are graded, as *C. elegans* genome lacks molecules coding for voltage-activated sodium channels [37]. This assumption is also consistent with electrophysiological observations in *C. elegans* and related nematodes [38, 39].

4) Each neuron releases a single neurotransmitter, or equivalently, there exists a dominant neurotransmitter type for each neuron.

5) Worm's movement direction is determined by a relative imbalance in the activities of excitatory motor neurons of type A and B, which is in agreement with recent experimental observations [28].
6) Behavioral output of the worm can be formally described in terms of a three-state model. The three states correspond to forward motion, backward motion, and stopped period. Each state has its probability of occurrence, which for forward and backward states is given by an exponential function of a difference between activities of type A and B motor neurons (see below).

Derivation of the model equations. Equations describing interneuron circuit responsible for forward-backward motion transitions are given below. This is a nonlinear model based on synaptic connectivity data from www.wormatlas.org (updated version of White et al [1] wiring diagram from [2]).

We start with a standard membrane equation describing the graded dynamics of neuron with an index i [40]:

$$CS\frac{dV_i}{dt} = -g_L S(V_i - V_r) - \sum_j g_{s,ij}\mathbf{H_0}(V_j)(V_i - V_{s,j}) - \sum_j g_{e,ij}(V_i - V_j) \quad (1)$$

where V_i is the voltage of neurons i, C is the membrane capacitance per surface area, S is the total surface area of neuron i, V_r is the resting voltage, g_L is the total membrane ionic conductance per surface area that is composed mainly of a constant leak current (typical

K^+ channel conductance is much smaller for voltages close to V_r), $g_{e,ij}$ is the gap junction conductance between neurons i and j. The symbol $g_{s,ij}$ denotes synaptic conductance coming from j presynaptic neuron with synaptic reversal potential $V_{s,j}$. The function $\mathbf{H_0}(V_j)$ is a nonlinear sigmoidal function characterizing synaptic transmission, and is given by

$$\mathbf{H_0}(V_j) = \frac{1}{1 + \exp[-\gamma(V_j - \theta_0)]}, \tag{2}$$

where θ_0 is the voltage threshold for synaptic activation, and γ is a measure of steepness of the activation slope.

Our goal is to write Eq. (1) in a more convenient form for the investigation of synaptic polarities. We assume that the resting potential V_r (when no synaptic or gap junction input is present) for *C. elegans* interneurons is -40 mV, which agrees with earlier suggestions [15], and it is close to a recent measurement (≈ -50 mV) in AIB neuron [22]. We want to expand the voltage in Eq. (1) around its resting value, i.e., we introduce $\Delta V_i \equiv V_i - V_r$. Let's denote by V_{ex} the reversal potential for excitatory, and by V_{in} the reversal potential for inhibitory synapses. The value of V_{ex} is around 0 mV (the current in excitatory synapses is mediated by Na^+, K^+, and partly by Ca^{++}). The value of V_{in} was reported between -70 mV [41] and -90 mV [42] (the current in inhibitory synapses is mediated by Cl^-). As an average for V_{in} we take -80 mV. Consequently, we obtain for excitatory synapses $V_i - V_{ex} = \Delta V_i + V_r - V_{ex} = \Delta V_i - 40$, and for inhibitory $V_i - V_{in} = \Delta V_i + V_r - V_{in} = \Delta V_i + 40$. The resulting average numerical factors in

both expressions have identical absolute values. Thus we can use an approximation: $V_i - V_{s,j} \approx \Delta V_i - \epsilon_j A$, where $A = 40$ mV, and ϵ_j characterizes the synaptic polarity of the presynaptic neuron j. The value of ϵ_j is either 1 for excitatory synapses or -1 for inhibitory. Taking the above into account and dividing both sides of Eq. (1) by $g_L S$, we can rewrite this equation as

$$\frac{C}{g_L}\frac{\Delta V_i}{dt} = -\left[1 + \sum_j \frac{g_{s,ij}\mathbf{H_0}(V_j)}{g_L S}\right]\Delta V_i + \sum_j \epsilon_j A \frac{g_{s,ij}}{g_L S}\mathbf{H_0}(V_j) - \sum_j \frac{g_{e,ij}}{g_L S}(\Delta V_i - \Delta V_j). \quad (3)$$

We can determine the strengths of synaptic and gap junction connections between any i and j interneurons by their anatomical numbers $N_{s,ij}, N_{e,ij}$, and maximal elementary conductances q_s, q_e, i.e., $g_{s,ij} = N_{s,ij} q_s$, and $g_{e,ij} = N_{e,ij} q_e$. The data for $N_{s,ij}$ and $N_{e,ij}$ are available from the data set in [2] (see Table 6). A typical range of conductances for chemical and electrical synapses is known from neurophysiology of other animals [42]. The leak conductance g_L is taken as $g_L^{-1} = 150\ k\Omega \cdot cm^2$ [15], which comes form the neurophysiological measurements in a related larger nematode *Ascaris* [38]. The surface area S of all interneurons is very similar and around $15 \cdot 10^{-6}$ cm^2 [1, 15], so we obtain $g_L S = 0.1$ nS, of which the inverse (i.e. $10^{10}\ \Omega$) is comparable to an experimental measurement of the total input resistance $\sim 0.5 \cdot 10^{10}\ \Omega$ [39]. Consequently, we can estimate the ratios present in Eq. (3) as: $A g_{s,ij}/(g_L S) = 400 N_{s,ij} q_s$, and $g_{e,ij}/(g_L S) = 10 N_{e,ij} q_e$, where q_s, q_e are expressed in nS. We checked that the term $\sum_j g_{s,ij}\mathbf{H}(V_j)/(g_L S)$, which

is associated with ΔV_i is generally much smaller than 1, since $\mathbf{H}(V_j) \ll 1$ for voltages not far away from V_r. Consequently, this term is neglected, which simplifies significantly the resulting equations for interneurons (see below). Thus, we can write Eq. (3) in an approximate form as:

$$\tau \frac{\Delta V_i}{dt} \approx -\Delta V_i + \sum_j \epsilon_j w_{ij} \mathbf{H}(\Delta V_j) - \sum_j g_{ij}(\Delta V_i - \Delta V_j), \tag{4}$$

where $\tau = C/g_L$ is the membrane time constant, w_{ij} is the synaptic coupling $w_{ij} = 400 q_s N_{s,ij}$, and g_{ij} is the gap junction coupling $g_{ij} = 10 q_e N_{e,ij}$. The function $\mathbf{H}(\Delta V_j)$ in Eq. (4) differs from the original function $\mathbf{H_0}(V_j)$ only by a substitution $\theta_0 \to \theta$, where $\theta = \theta_0 - V_r$, i.e.

$$\mathbf{H}(\Delta V_j) = \frac{1}{1 + \exp[-\gamma(\Delta V_j - \theta)]}. \tag{5}$$

It is important to keep in mind that synaptic polarities in Eq. (4) are determined simply by the signs of ϵ_j coefficients.

Activity equations for interneurons. The dynamics of the interneurons in the locomotory circuit are based on the derived Eq. (4), and therefore can be represented by a following set of differential equations:

$$\tau\frac{dAVB}{dt} = -AVB + \epsilon_{ASH}w_{AVB,ASH}\mathbf{H}(ASH) + \epsilon_{PVC}w_{AVB,PVC}\mathbf{H}(PVC)$$
$$+\epsilon_{AVA}w_{AVB,AVA}\mathbf{H}(AVA) + \epsilon_{AVD}w_{AVB,AVD}\mathbf{H}(AVD)$$
$$+\epsilon_{DVA}w_{AVB,DVA}\mathbf{H}(DVA) + \epsilon^2_{AVB}\epsilon^2_{DVA}g_{AVB,DVA}(DVA - AVB)$$
$$+\epsilon^2_{AVB}g_{AVB,F}(E_f - AVB) + \epsilon^2_{AVB}g_{AVB,B}(E_b - AVB) + X_{AVB}, \quad (6)$$

$$\tau\frac{dPVC}{dt} = -PVC + \epsilon_{AVA}w_{PVC,AVA}\mathbf{H}(AVA) + \epsilon_{DVA}w_{PVC,DVA}\mathbf{H}(DVA)$$
$$+\epsilon_{AVD}w_{PVC,AVD}\mathbf{H}(AVD) + \epsilon_{AVE}w_{PVC,AVE}\mathbf{H}(AVE)$$
$$+w_{PVC,F}\mathbf{H}(E_f) + w_{PVC,B}\mathbf{H}(E_b)$$
$$+\epsilon^2_{PVC}\epsilon^2_{AVA}g_{PVC,AVA}(AVA - PVC) + \epsilon^2_{PVC}\epsilon^2_{DVA}g_{PVC,DVA}(DVA - PVC)$$
$$+\epsilon^2_{PVC}g_{PVC,F}(E_f - PVC) + \epsilon^2_{PVC}g_{PVC,B}(E_b - PVC) + X_{PVC}, \quad (7)$$

$$\tau\frac{dDVA}{dt} = -DVA + \epsilon_{PVC}w_{DVA,PVC}\mathbf{H}(PVC) + w_{DVA,F}\mathbf{H}(E_f)$$
$$+\epsilon^2_{DVA}\epsilon^2_{AVB}g_{DVA,AVB}(AVB - DVA) + \epsilon^2_{DVA}\epsilon^2_{PVC}g_{DVA,PVC}(PVC - DVA)$$
$$+\epsilon^2_{DVA}g_{DVA,F}(E_f - DVA) + X_{DVA}, \quad (8)$$

$$\tau\frac{dAVA}{dt} = -AVA + \epsilon_{ASH}w_{AVA,ASH}\mathbf{H}(ASH) + \epsilon_{AVB}w_{AVA,AVB}\mathbf{H}(AVB)$$
$$+\epsilon_{PVC}w_{AVA,PVC}\mathbf{H}(PVC) + \epsilon_{AVD}w_{AVA,AVD}\mathbf{H}(AVD)$$
$$+\epsilon_{AVE}w_{AVA,AVE}\mathbf{H}(AVE) + \epsilon_{DVA}w_{AVA,DVA}\mathbf{H}(DVA)$$
$$+w_{AVA,B}\mathbf{H}(E_b) + \epsilon^2_{AVA}\epsilon^2_{PVC}g_{AVA,PVC}(PVC - AVA)$$
$$+\epsilon^2_{AVA}g_{AVA,B}(E_b - AVA) + \epsilon^2_{AVA}g_{AVA,F}(E_f - AVA) + X_{AVA}, \qquad (9)$$

$$\tau\frac{dAVD}{dt} = -AVD + \epsilon_{AVB}w_{AVD,AVB}\mathbf{H}(AVB) + \epsilon_{PVC}w_{AVD,PVC}\mathbf{H}(PVC)$$
$$+\epsilon_{ASH}w_{AVD,ASH}\mathbf{H}(ASH) + \epsilon_{AVA}w_{AVD,AVA}\mathbf{H}(AVA)$$
$$+\epsilon_{AVE}w_{AVD,AVE}\mathbf{H}(AVE) + w_{AVD,B}\mathbf{H}(E_b) + X_{AVD}, \qquad (10)$$

$$\tau\frac{dAVE}{dt} = -AVE + \epsilon_{AVB}w_{AVE,AVB}\mathbf{H}(AVB) + \epsilon_{PVC}w_{AVE,PVC}\mathbf{H}(PVC)$$
$$+\epsilon_{DVA}w_{AVE,DVA}\mathbf{H}(DVA) + \epsilon_{ASH}w_{AVE,ASH}\mathbf{H}(ASH)$$
$$+\epsilon_{AVA}w_{AVE,AVA}\mathbf{H}(AVA) + X_{AVE}, \qquad (11)$$

where AVB, PVC, DVA, AVA, AVD, AVE are the relative activities of the corresponding interneurons with respect to their resting values. The parameters $w_{\alpha,\beta}$ are synaptic strengths, and $g_{\alpha,\beta}$ are gap-junction (electrical) couplings, where the subscripts α and β denote the above interneurons. As before, the symbol ϵ_α denotes synaptic polarity of the neuron α, and it assumes value 1 (if the neuron is excitatory), value -1 (if inhibitory), or 0 (if the neuron is absent because of the ablation). Note that gap junction couplings contain the ϵ_α^2 factors, which are either 1 (if the neuron α is present), or 0 (if it is removed from the network). The parameter X_α describes an input coming from the upstream neurons to the interneuron α. For all interneurons except PVC this input comes from the head neurons. It is represented by $X_\alpha = x_0 + \sigma z_\alpha$, where $x_0 = 2.0$ mV, σ is a variable parameter, and z_α is either 0 (weak input) or 1 (strong input). The parameter z_α, similar to ϵ_α, is unknown. We want to find both of them for each interneuron.

Equations (6-11) describing graded activities of neurons in *C. elegans* are of similar kind to those used before in [33] for analyzing forward locomotion.

The above pre-motor interneurons make synaptic and gap junction connections with downstream excitatory motor neurons. Two separate groups of these motor neurons generating forward and backward motion, called B and A respectively, directly connect locomotory muscles. The activities of excitatory motor neurons are given by:

$$\tau \frac{dE_f}{dt} = -E_f + \epsilon_{PVC} w_{F,PVC} \mathbf{H}(PVC) + \epsilon_{DVA} w_{F,DVA} \mathbf{H}(DVA)$$

$$+\epsilon_{AVA} w_{F,AVA} \mathbf{H}(AVA) + \epsilon_{AVB} w_{F,AVB} \mathbf{H}(AVB)$$
$$+\epsilon_{AVD} w_{F,AVD} \mathbf{H}(AVD) + \epsilon_{AVE} w_{F,AVE} \mathbf{H}(AVE)$$
$$+\epsilon^2_{AVB} g_{F,AVB}(AVB - E_f) + \epsilon^2_{AVA} g_{F,AVA}(AVA - E_f)$$
$$+\epsilon^2_{PVC} g_{F,PVC}(PVC - E_f) + \epsilon^2_{DVA} g_{F,DVA}(DVA - E_f), \tag{12}$$

and

$$\tau \frac{dE_b}{dt} = -E_b + \epsilon_{AVA} w_{B,AVA} \mathbf{H}(AVA) + \epsilon_{AVD} w_{B,AVD} \mathbf{H}(AVD)$$
$$+\epsilon_{AVE} w_{B,AVE} \mathbf{H}(AVE) + \epsilon_{AVB} w_{B,AVB} \mathbf{H}(AVB)$$
$$+\epsilon_{PVC} w_{B,PVC} \mathbf{H}(PVC) + \epsilon_{DVA} w_{B,DVA} \mathbf{H}(DVA)$$
$$+\epsilon^2_{AVA} g_{B,AVA}(AVA - E_b) + \epsilon^2_{AVB} g_{B,AVB}(AVB - E_b) + \epsilon^2_{PVC} g_{B,PVC}(PVC - E_b), \tag{13}$$

where E_f and E_b are the total relative activities of forward (type B) and backward (type A) motor neurons, respectively, measured from their resting voltages.

We solve Eq. (6-13) using a second order Runge-Kutta method. We are interested only in the steady-state activities of the interneurons and motor neurons. One can think about these activities as temporal averages over sufficiently long periods of time. This means that all fluctuations in the neural voltage that can be caused by an input from upstream networks are averaged over and thus neglected. This simplifying step significantly enhances the feasibility of the analysis. The steady-state values of E_f and E_b are inserted in Eq. (20); see below.

ASH neuron. From all upstream neurons we selected explicitly ASH neuron because of its polymodal sensory role. Specifically, it has been implicated in avoidance responses [43], which are associated with modulating locomotion. We do not write an explicit equation for ASH dynamics, because it receives a massive input form many other head neurons, of which we have no knowledge. Instead, we make computations for 4 different levels of ASH activity that we set by hand. We choose $ASH = \kappa\theta$, where $\kappa = 0.2$, 0.4, 0.6, or 0.75. The value of the normalized threshold θ is set to 40 mV.

Theoretical ablations. Ablations or removals of neurons in the model are performed by setting $\epsilon_{neuron} = 0$. For example, if we remove neuron AVB, then we put $\epsilon_{AVB} = 0$ in all equations above.

Relationship to the behavioral data. From the experimental part we have average times the worms spent in forward and backward locomotion, which we denote as T_f and T_b, respectively. These average times should be somehow related to the average activities of the two type of motor neurons, E_f and E_b. The precise relationship between them is unknown and probably complicated. Nevertheless, we can make some progress by using an analogy with statistical physics [44], and treating worm's locomotory behavior as a three state system influenced by both deterministic and stochastic factors. These three states correspond to forward movement, backward movement, and stopped time (no motion). There could be transitions between the states driven by sensory input from the environment (either deterministic or stochastic). We are interested only in "average" or steady states activities of the system. Our model is motivated in large

part by experimental results of Kawano et al [28]. In that study it was shown that *C. elegans* motion direction is determined by a relative activity of A and B motor neurons. In particular, it was suggested [28] that when forward motor neurons are much more active than backward (i.e. $E_f \gg E_b$), then there should be a high probability of finding the worm in the forward motion. Conversely, if the activity of backward motor neurons dominates over the activities of their forward counterparts (i.e. $E_b \gg E_f$), then there is a high chance that the worm moves backward. Thus, it appears that the sign of $E_f - E_b$ plays a key role in the choice of worm's motion direction. Moreover, one can expect that when the activities of both types of motor neurons are comparable (equal or almost equal), then *C. elegans* likely does not move.

These qualitative observations can be formulated mathematically in the following way. The probability of forward motion P_f can be written as

$$P_f = Z \exp[(E_f - E_b - \Delta)/\eta_0], \tag{14}$$

where Δ is some activity threshold for locomotion initiation, the parameter η_0 characterizes the level of noise in the system, and Z is the normalization factor. By symmetry considerations, we can write the probability for backward motion P_b as

$$P_b = Z \exp[(E_b - E_f - \Delta)/\eta_0]. \tag{15}$$

The choice of the exponential function in P_f and P_b is motivated by two major argu-

ments. First, with the exponentials both P_f and P_b are always increasing and positive functions of the arguments $E_f - E_b$ and $E_b - E_f$ for the whole range of their variabilities (from $-\infty$ to $+\infty$), which is generally not the case for other simple choices, in particular polynomials. For example, the choice $P_f \sim (E_f - E_b - \Delta)^n$, with n an even integer, is not satisfactory because the probability P_f is a non-monotonic function of its argument (it has a minimum for $E_f = E_b + \Delta$). Similarly, if the exponent n is an odd integer, then P_f becomes negative for $E_f < E_b + \Delta$, which is clearly wrong. A second argument in favor of the exponential function in the probabilities is the fact that many phenomena occurring in nature have a similar type of dependence (e.g. [44]).

The probability that the nematode does not move, i.e. it is in the stopped state, is

$$P_s = Z \cdot S_s\left(\frac{E_f - E_b}{\Delta}\right), \tag{16}$$

where $S_s(x)$ is some unknown function that should have the following properties. For $|x| \ll 1$, i.e. when $|E_f - E_b|$ is much smaller than the motion activation threshold Δ, the function $S_s(x) \gg 1$. For $|x| \gg 1$, i.e. when either E_f or E_b dominates, the worm should not be motionless, which corresponds to $S_s(x) \to 0$. Additionally, due to symmetry one should have $S_s(-x) = S_s(x)$. The form of the $S_s(x)$ function is however irrelevant for the kind of computations made in this paper (see below).

From the normalization condition for probabilities, $P_f + P_b + P_s = 1$, we can determine Z, which allows us to find explicit forms for the state probabilities. The latter read:

$$P_f = \frac{e^{(E_f - E_b - \Delta)/\eta_0}}{e^{(E_f - E_b - \Delta)/\eta_0} + e^{(E_b - E_f - \Delta)/\eta_0} + S_s} \tag{17}$$

$$P_b = \frac{e^{(E_b - E_f - \Delta)/\eta_0}}{e^{(E_f - E_b - \Delta)/\eta_0} + e^{(E_b - E_f - \Delta)/\eta_0} + S_s} \tag{18}$$

$$P_s = \frac{S_s}{e^{(E_f - E_b - \Delta)/\eta_0} + e^{(E_b - E_f - \Delta)/\eta_0} + S_s} \tag{19}$$

In a case when the activity of forward motor neurons dominates over the rest ($E_f \gg E_b + \Delta$), we have $P_f \to 1$, $P_b \to 0$, and $P_s \to 0$. For $E_f \approx E_b$, we have $P_f \approx P_b \ll 1$, and $P_s \to 1$.

On the other hand the probabilities P_f, P_b, and P_s are related to the times (T_f, T_b, T_s) the worm spends in the corresponding states. The average probability that *C. elegans* is in the forward state is $P_f = T_f/(T_f + T_b + T_s)$. Similarly, $P_b = T_b/(T_f + T_b + T_s)$, and $P_s = T_s/(T_f + T_b + T_s)$. Thus, the ratio T_f/T_b is equal to P_f/P_b. Combining the above equations, we obtain

$$T_f/T_b = \exp\left[(E_f - E_b)/\eta\right], \tag{20}$$

where $\eta = \eta_0/2$. Note that the ratio of times associated with forward and backward locomotion neither depends on the activation threshold Δ nor on the form of the $S_s(x)$ function. The ratio of times depends only on the difference in the activation of complementary motor neurons and the level of noise in the system. It is interesting to stress that the quantity of empirical interest, i.e. $T_f/(T_f+T_b)$ is equal to $[1+\exp((E_b-E_f)/\eta)]^{-1}$. The latter expression is known as a sigmoidal logistic function, and serves as a transfer function from neural activities to behavioral output.

The goal. We want to determine which combination of neuron polarities: ϵ_{ASH}, ϵ_{AVB}, ϵ_{PVC}, ϵ_{DVA}, ϵ_{AVA}, ϵ_{AVD}, ϵ_{AVE}, together with their corresponding upstream inputs z_α, yields the best fit to the experimental values of $T_f/(T_f+T_b)$.

Comparison of the theory with the data. In our model there are 8192 distinct combinations of synaptic polarities ϵ_α and the upstream inputs z_α, i.e., different configurations in which the circuit can be found. For each circuit configuration, we perform 17 interneuron ablations in our computer model, and compute theoretical values of $T_f/(T_f+T_b)$ for each ablation. Next, we compute an Euclidean distance ED of these values to the experimental values given in Table 1, according to the expression:

$$\mathrm{ED} = \left[\sum_{a=1}^{18}(R_{th}-R_{ex})_a^2\right]^{1/2}, \tag{21}$$

where $R = T_f/(T_f+T_b) \equiv (T_f/T_b)\,[1+T_f/T_b]^{-1}$, and the subscripts th and ex refer to

theoretical and experimental values of R. The subscript a refers to the ablation number, in the same order as in Table 1. In particular, $a = 1$ corresponds to the mock ablation, i.e. wild type (WT).

Values of the connectivity matrix. The strength w_{ij} of synaptic input to neuron i coming from neuron j is given according to Eq. (4) by the expression $w_{ij} = 400 q_s N_{s,ij}$, where $N_{s,ij}$ is the number of synaptic contacts of neuron i with presynaptic neuron j. The number $N_{s,ij}$ is determined as an arithmetic mean for the right- and left-hand side interneurons. As an example, the right AVB neuron receives an input from both right and left PVC neurons, of which we take an arithmetic mean. Similarly, the left AVB neuron receives an input from both right and left PVC, of which we again take an arithmetic mean. Next, we take an arithmetic mean of these two arithmetic means, and obtain a single value representing average number of synaptic contacts between presynaptic PVC and postsynaptic AVB ($N_{s,AVB,PVC}$). The strength of gap junction g_{ij} between neurons i and j is given by an analogous formula $g_{ij} = 10 q_e N_{e,ij}$, where $N_{e,ij}$ is the number of gap junction contacts between i and j (arithmetic mean of right and left interneurons). Empirical data for $N_{s,ij}$ and $N_{e,ij}$ were taken from the database in [2] and are presented in Table 6. Parameters q_s and q_e are taken in the range: $q_s = 0.1 - 0.6$ nS, and $q_e = 0.1 - 0.5$ nS [42].

Acknowledgments

We thank Cori Bargmann for comments on the manuscript.

References

[1] White JG, Southgate E, Thomson JN, Brenner S (1986) The structure of the nervous system of the nematode *Caenorhabditis elegans*. *Philos. Trans. R. Soc. Lond. B* **314**: 1-340.

[2] Chen BL, Hall DH, Chklovskii DB (2006) Wiring optimization can relate neuronal structure and function. *Proc. Natl. Acad. Sci. USA* **103**: 4723-4728.

[3] Varshney LR, Chen BL, Paniagua E, Hall DH, Chklovskii DB (2011) Structural properties of the *Caenorhabditis elegans* neuronal network. *PLoS Comput. Biol.* **7**: e1001066.

[4] Cherniak C (1994) Component placement optimization in the brain. *J. Neurosci.* **14**: 2418-2427.

[5] Kaiser M, Hilgetag CC (2006) Nonoptimal component placement, but short processing paths, due to long-distance projections in neural systems. *PLoS Comput. Biol.* **2**: e95.

[6] Karbowski J (2001) Optimal wiring principle and plateaus in the degree of separation for cortical neurons. *Phys. Rev. Lett.* **86**: 3674-3677.

[7] Karbowski J (2003) How does connectivity between cortical areas depend on brain size? Implications for efficient computations. *J. Comput. Neurosci.* **15**: 347-356.

[8] Chklovskii DB (2004) Synaptic connectivity and neuronal morphology: two sides of the same coin. *Neuron* **43**: 609-617.

[9] Hobert O (2003) Behavioral plasticity in *C. elegans*: Paradigms, circuits, genes. *J. Neurobiol.* 54: 203-223.

[10] de Bono M, Maricq AV (2005) Neuronal substrates of complex behaviors in *C. elegans*. *Annu. Rev. Neurosci.* 28: 451-501.

[11] van Vreeswijk C, Sompolinsky H (1996) Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science* **274**: 1724-1726.

[12] Haider B, Duque A, Hasenstaub AR, McCormick DA (2006) Neocortical network activity in vivo is generated through a dynamic balance of excitation and inhibition. *J. Neurosci.* **26**: 4535-4545.

[13] Vogels TP, Sprekeler H, Zenke F, Clopath C, Gerstner W (2011) Inhibitory plasticity balances excitation and inhibition in sensory pathways and memory networks. *Science* **334**: 1569-1573.

[14] DeFelipe J, Alonso-Nanclares L, Avellano J (2002) Microstructure of the neocortex: Comparative aspects. *J. Neurocytology* **31**: 299-316.

[15] Wicks SR, Roehrig CJ, Rankin CH (1996) A dynamic network simulation of the nematode tap withdrawal circuit: Predictions concerning synaptic function using behavioral criteria. *J. Neurosci.* **16**: 4017-4031.

[16] Jospin M, Qi Y, Stawicki T, Boulin T, Schuske K, et al (2009) A neuronal acetylcholine receptor regulates the balance of muscle excitation and inhibition in *Caenorhabditis elegans*. *PLoS Biology* **7**: e1000265.

[17] Sengupta P, Samuel ADT (2009) *Caenorhabditis elegans*: a model system for systems neuroscience. *Curr. Opin. Neurobiol.* **19**: 637-643.

[18] Gray JM, Hill JJ, Bargmann CI (2005) A circuit for navigation in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. USA* **102**: 3184-3191.

[19] Ha H, Hendricks M, Shen Y, Gabel CV, Fang-Yen C, et al (2010) Functional organization of a neural network for aversive olfactory learning in *Caenorhabditis elegans*. *Neuron* **68**: 1173-1186.

[20] Chalfie M, Sulston JE, White JG, Southgate E, Thomson JN, Brenner S (1985) The neural circuit for touch sensitivity in *Caenorhabditis elegans*. *J. Neurosci.* **5**: 956-964.

[21] Haspel G, O'Donovan MJ, Hart AC (2010) Motorneurons dedicated to either forward or backward locomotion in the nematode *C. elegans*. *J. Neurosci.* **30**: 11151-11156.

[22] Piggott BJ, Liu J, Feng Z, Wescott SA, Xu XZS (2011) The neural circuits and synaptic mechanisms underlying motor initiation in *C. elegans*. *Cell* **147**: 922-933.

[23] Li W, Feng Z, Sternberg PW, Xu SXZ (2006) A C. elegans stretch receptor neuron revealed by a mechanosensitive TRP channel homologue. *Nature* **440**: 684-687.

[24] Zheng Y, Brockie PJ, Mellem JE, Madsen DM, Maricq AV (1999) Neuronal control of locomotion in *C. elegans* is modified by a dominant mutation in the GLR-1 ionotropic glutamate receptor. *Neuron* **24**: 347-361.

[25] Faumont S, et al (2011) An image-free opto-mechanical system for creating virtual environments and imaging neuronal activity in freely moving *Caenorhabditis elegans*. *PLoS ONE* **6**: e24666.

[26] Chronis N, Zimmer M, Bargmann CI (2007) Microfluidics for in vivo imaging of neuronal and behavioral activity in *Caenorhabditis elegans*. *Nature Methods* **4**: 727-731.

[27] Qi YB, Garren EJ, Shu X, Tsien RY, Jin Y (2012) Photo-inducible cell ablation in Caenorhabditis elegans using the genetically encoded singlet oxygen generating protein miniSOG. *Proc. Natl. Acad. Sci. USA* **109**: 7499-7504.

[28] Kawano T, Po MD, Gao S, Leung G, Ryu WS, et al (2011) An imbalancing act: gap junctions reduce the backward motor circuit activity to bias *C. elegans* forward locomotion. *Neuron* **72**: 572-586.

[29] Brockie PJ, Maricq AV (2006) Ionotropic glutamate receptors: genetics, behavior, and electrophysiology. *Wormbook*, ed. The C. elegans Research Community; http://www.wormbook.org.

[30] Mellem JE, Brockie PJ, Zheng Y, Madsen DM, Maricq AV (2002) Decoding of polymodal sensory stimuli by postsynaptic glutamate receptors in *C. elegans*. *Neuron* **36**: 933-944.

[31] Dunn NA, Conery JS, Lockery SR (2007) Circuit motifs for spatial orientation behaviors identified by neural network optimization. *J. Neurophysiol.* **98**: 888-897.

[32] Karbowski J, Cronin CJ, Seah A, Mendel JE, Cleary D, Sternberg PW (2006) Conservation rules, their breakdown, and optimality in *Caenorhabditis* sinusoidal locomotion. *J. Theor. Biol.* 242: 652-669.

[33] Karbowski J, Schindelman G, Cronin CJ, Seah A, Sternberg PW (2008) Systems level circuit model of *C. elegans* undulatory locomotion: mathematical modeling and molecular genetics. *J. Comput. Neurosci.* **24**: 253-276.

[34] Bryden JA, Cohen N (2008) Neural control of *Caenorhabditis elegans* forward locomotion: the role of sensory feedback. *Biol. Cybern.* **98**: 339-351.

[35] Brenner S (1974) The genetics of Caenorhabditis elegans. *Genetics* **77**: 71-94.

[36] Cronin CJ, Mendel JE, Mukhtar S, Kim YM, Stirbl RC, et al (2005) An automated system for measuring parameters of nematode sinusoidal movement. *BMC Genetics* **6**: 5.

[37] Bargmann CI (1998) Neurobiology of the *Caenorhabditis elegans* genome. *Science* **282**: 2028-2033.

[38] Davis RE, Stretton AOW (1989) Passive membrane properties of motorneurons and their role in long-distance signaling in the nematode *Ascaris*. *J. Neurosci.* **9**: 403-414.

[39] Goodman MB, Hall DH, Avery L, Lockery SR (1998). Active currents regulate sensitivity and dynamic range in *C. elegans* neurons. *Neuron* 20: 763-772.

[40] Ermentrout GB, Terman DH (2010) *Mathematical Foundations of Neuroscience.* Springer: New York.

[41] Purves D, et al (2008) *Neuroscience, 4th Ed.* Sinauer Assoc.

[42] Koch C (1998) *Biophysics of computation.* Oxford: Oxford Univ. Press.

[43] Kaplan JM, Horvitz HR (1993) A dual mechanosensory and chemosensory neuron in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. USA* **90**: 2227-2231.

[44] Gardiner CW (2004) *Handbook of Stochastic Methods.* Springer: Berlin, 3rd ed.

Figure Captions

Fig. 1

Schematic diagram of the interneuron locomotory circuit. (A) Intact circuit. ASH neuron is an upstream neuron that provides synaptic input to the locomotory interneurons. The output coming from the 6 neurons (5 interneurons and DVA) feeds the activities of motor neurons, represented by E_f (controlling forward motion) and by E_b (controlling backward motion). Synaptic connections are shown as solid arrows (blue), and gap junctions are represented by dashed lines (red). The magnitude of an arrow and the width of a dashed line are indicators of the strength of synaptic and gap junction connections, respectively. (B) An example of an ablated circuit, in which ASH and AVB neurons are removed. Note that this leads to the removal of all connections (synaptic and electric) coming out from these neurons. Such ablations not only change the circuit architecture but also modify its activity output.

Fig. 2

Comparison of theoretical relative times spent in forward locomotion with experimental values for the winning polarity combination # 1. Theoretical points are represented by triangles (red), and experimental by circles (blue).

Fig. 3

Distribution of synaptic polarities for each interneuron. The first 20 polarity combinations with the smallest Euclidean distance (ED) are shown, and they are associated with the optimal parameters given in Table 2. Note that the interneurons ASH, AVA, AVE, and PVC are inhibitory with a high probability. There are some nonzero likeli-

hoods that AVB, AVD, and DVA neurons are excitatory (especially AVB and AVD), although the smallest ED values are associated with negative polarities.

Fig. 4

Typical dependence of the Euclidean Distance (ED) on the noise factor η. ED has a minimum for some optimal η. Shown are combinations of synaptic polarities number 1 (solid line), 17 (dashed line), and 11 (dotted line), for $\sigma = 8.0$ mV and $\kappa = 0.6$, which correspond to the configurations with the smallest ED.

Table 1: Experimental data of the impact of neuron ablation on *C. elegans* locomotion. Shown are population average times and standard errors of the mean (SEM) for worms in forward and backward motion, and in stopped phase. The last column gives population averages of reversal frequencies with their SEM.

Ablation type	Forward time T_f (sec)	Backward time T_b (sec)	Stopped time (sec)	Reversals (min^{-1})
mock ablated (WT, N=43)	8.98 ± 0.57	2.80 ± 0.27	0.26 ± 0.01	5.29 ± 0.27
ASH (N=14)	12.57 ± 1.67	0.93 ± 0.17	0.27 ± 0.01	3.79 ± 0.80
AVA (N=11)	0.71 ± 0.09	0.53 ± 0.04	0.60 ± 0.05	10.26 ± 0.56
AVB (N=8)	2.26 ± 0.40	2.14 ± 0.23	0.38 ± 0.02	6.10 ± 0.64
AVD (N=4)	4.23 ± 1.80	3.12 ± 0.36	0.31 ± 0.04	3.50 ± 0.31
DVA (N=22)	1.51 ± 0.18	1.23 ± 0.08	0.44 ± 0.02	10.04 ± 0.57
PVC (N=12)	11.99 ± 1.81	1.89 ± 0.39	0.29 ± 0.02	5.46 ± 0.74
ASH+AVA (N=7)	1.91 ± 0.42	0.85 ± 0.20	0.52 ± 0.06	5.18 ± 0.90
ASH+AVB (N=12)	2.05 ± 0.47	2.04 ± 0.43	0.42 ± 0.06	6.92 ± 1.08
AVA+AVB (N=9)	0.56 ± 0.14	0.46 ± 0.06	0.89 ± 0.17	10.09 ± 1.36
AVA+PVC (N=11)	4.09 ± 0.91	0.67 ± 0.14	0.37 ± 0.04	9.78 ± 0.71
AVB+PVC (N=5)	0.91 ± 0.24	1.19 ± 0.19	0.44 ± 0.08	14.96 ± 3.52
DVA+PVC (N=19)	2.18 ± 0.21	1.35 ± 0.08	0.40 ± 0.02	11.63 ± 0.57
ASH+AVA+AVB (N=8)	0.75 ± 0.24	0.52 ± 0.11	1.16 ± 0.26	6.47 ± 0.83
AVA+AVB+PVC (N=8)	0.93 ± 0.33	0.47 ± 0.12	0.87 ± 0.21	6.10 ± 0.87
AVB+AVD+PVC (N=5)	1.33 ± 0.31	0.94 ± 0.13	0.49 ± 0.07	11.21 ± 2.00
AVB+DVA+PVC (N=10)	1.90 ± 0.28	1.03 ± 0.14	0.40 ± 0.21	12.21 ± 1.53
AVA+AVB+AVE+PVC (N=10)	0.60 ± 0.21	0.39 ± 0.14	1.00 ± 0.12	8.66 ± 1.54

Table 2: The winning combinations of interneuron polarities (ϵ_i) for the upstream input: $\sigma = 8.0$ mV and $\kappa = 0.6$.

Neuron	Rank								inhibitory
	1	2	3	4	5	6	7	8	likelihood
ASH	-1	-1	-1	-1	-1	-1	-1	-1	1
AVA	-1	-1	-1	-1	-1	-1	1	1	3/4
AVB	-1	1	-1	1	1	-1	-1	-1	5/8
AVD	-1	-1	1	1	-1	-1	-1	-1	3/4
AVE	-1	-1	-1	-1	-1	-1	-1	-1	1
DVA	-1	-1	1	1	1	1	1	-1	3/8
PVC	-1	-1	-1	-1	-1	-1	-1	-1	1
Combination #	1	17	11	27	19	3	35	33	
ED	0.3625	0.3651	0.374	0.377	0.380	0.383	0.396	0.409	
Corr	0.7433	0.7417	0.722	0.717	0.740	0.746	0.690	0.731	

The optimal values of the parameters are: $\eta = 1.05$, $q_s = 0.1$ nS, $q_e = 0.1$ nS.

The winning combination # 1 receives the following input from the head: $z_{AVB} = 1$, $z_{PVC} = 1$, and $z_\alpha = 0$ for other interneurons. The last column provides estimates of probabilities that a given neuron is inhibitory.

Table 3: The winning combinations of interneuron polarities (ϵ_i) for the upstream input: $\sigma = 6.0$ mV and $\kappa = 0.6$.

Neuron	Rank								inhibitory
	1	2	3	4	5	6	7	8	likelihood
ASH	-1	-1	-1	-1	-1	-1	-1	-1	1
AVA	-1	-1	-1	-1	-1	-1	-1	-1	1
AVB	-1	1	-1	1	-1	1	1	1	3/8
AVD	-1	-1	1	1	-1	-1	1	1	1/2
AVE	-1	-1	-1	-1	-1	-1	-1	-1	1
DVA	-1	-1	1	1	1	1	-1	1	1/2
PVC	-1	-1	-1	-1	-1	-1	1	-1	7/8
Combination #	1	17	11	27	3	19	26	10	
ED	0.368	0.377	0.394	0.404	0.422	0.424	0.431	0.433	
Corr	0.734	0.722	0.687	0.665	0.692	0.669	0.627	0.637	

The optimal values of the parameters are: $\eta = 0.85$, $q_s = 0.1$ nS, $q_e = 0.1$ nS.

The winning combination # 1 receives the following input from the head: $z_{AVB} = 1$, $z_{PVC} = 1$, and $z_\alpha = 0$ for other interneurons. The last column provides estimates of probabilities that a given neuron is inhibitory.

Table 4: The winning combinations of interneuron polarities (ϵ_i) for the upstream input: $\sigma = 4.0$ mV and $\kappa = 0.6$.

Neuron	Rank								inhibitory
	1	2	3	4	5	6	7	8	likelihood
ASH	-1	-1	-1	-1	-1	-1	-1	-1	1
AVA	-1	-1	-1	-1	-1	-1	-1	-1	1
AVB	-1	1	-1	-1	-1	1	1	1	1/2
AVD	-1	-1	1	1	1	1	1	1	1/4
AVE	-1	-1	-1	-1	-1	-1	-1	-1	1
DVA	-1	-1	1	-1	-1	1	-1	-1	3/4
PVC	-1	-1	-1	-1	1	-1	1	-1	3/4
Combination #	1	17	11	9	10	27	26	25	
ED	0.414	0.431	0.453	0.461	0.465	0.471	0.476	0.487	
Corr	0.644	0.606	0.560	0.613	0.530	0.512	0.485	0.578	

The optimal values of the parameters are: $\eta = 0.7$, $q_s = 0.1$ nS, $q_e = 0.1$ nS.

The winning combination # 1 receives the following input from the head: $z_{AVB} = 1$ and $z_\alpha = 0$ for other interneurons. The last column provides estimates of probabilities that a given neuron is inhibitory.

Table 5: The winning combinations of interneuron polarities (ϵ_i) for the upstream input: $\sigma = 12.0$ mV and $\kappa = 0.6$.

Neuron	Rank								inhibitory
	1	2	3	4	5	6	7	8	likelihood
ASH	-1	-1	-1	-1	-1	-1	-1	-1	1
AVA	-1	-1	1	1	-1	-1	-1	-1	3/4
AVB	1	-1	1	-1	1	-1	-1	1	1/2
AVD	1	1	1	1	-1	-1	-1	-1	1/2
AVE	-1	-1	-1	-1	1	1	-1	-1	3/4
DVA	-1	-1	-1	-1	-1	-1	-1	-1	1
PVC	1	1	1	1	1	1	-1	-1	1/4
Combination #	26	10	58	42	22	6	1	17	
ED	0.383	0.386	0.390	0.392	0.393	0.394	0.397	0.398	
Corr	0.715	0.708	0.694	0.688	0.691	0.686	0.687	0.685	

The optimal values of the parameters are: $\eta = 0.85$, $q_s = 0.1$ nS, $q_e = 0.3$ nS.

The winning combination # 26 receives the following input from the head: $z_{AVB} = 1$, $z_{PVC} = 1$, and $z_\alpha = 0$ for other interneurons. The last column provides estimates of probabilities that a given neuron is inhibitory.

Table 6: Connectivity matrix for the command interneuron circuit. Shown are average anatomical number of synapses ($N_{s,ij}$) and gap junctions ($N_{e,ij}$) (in the brackets below synaptic contacts) between postsynaptic neuron i and presynaptic neuron j. Symbols F and B denote forward and backward motor neurons, respectively.

postsynaptic neuron	presynaptic neurons								
	ASH	AVA	AVB	AVD	AVE	DVA	PVC	F	B
AVA	1.75	–	6.75	15.75	10.5	2.0	5.0	–	0.25
	–	–	–	–	–	–	(2.5)	(3.5)	(25.5)
AVB	2.25	0.5	–	0.25	–	0.5	7.75	–	–
	–	–	–	–	–	(1.0)	–	(13.75)	(0.5)
AVD	3.0	1.0	0.75	–	0.25	–	3.25	–	0.25
	–	–	–	–	–	–	–	–	–
AVE	0.75	1.0	0.75	–	–	7.0	1.25	–	–
	–	–	–	–	–	–	–	–	–
DVA	–	–	–	–	–	–	2.0	0.5	–
	–	–	(1.0)	–	–	–	(0.5)	(0.5)	–
PVC	–	7.0	–	0.25	0.25	2.0	–	0.25	1.25
	–	(2.5)	–	–	–	(0.5)	–	(0.75)	(0.75)
F	–	2.5	0.25	0.25	0.25	6.5	–	–	–
	–	(3.5)	(13.75)	–	–	(0.5)	(0.75)	–	–
B	–	41.75	1.5	7.0	8.25	1.0	1.0	–	–
	–	(25.5)	(0.5)	–	–	–	(0.75)	–	–