4
Control of Neuronal Network in *Caenorhabditis elegans*

Caenorhabditis elegans, a soil dwelling nematode, is evolutionarily rudimentary and contains only ~300 neurons which are connected to each other via chemical synapses and gap junctions. This structural connectivity can be modeled as a graph of nodes (neurons) connected via edges (synapses). Humans have sought for various means for better understanding of nervous system and its control. Many methods have been developed for identification of specific brain regions, which when controlled by external input can lead to control over the state of the system. Hence to gain insight on the type of control achieved in nervous system we implemented the notion of structural control from graph theory for the study of *C. elegans* neuronal network. 'Driver neurons' which are critical for achieving full control over the network were identified using controllability analysis. We studied phenotypic properties of these neurons which are referred to as 'phenoframe' as well as the 'genoframe' which represents their genetic correlates. The driver neurons were found to be primarily motor neurons located in the ventral nerve cord and contributing to biological reproduction of the animal. Identification of driver neurons and its characterization, presented in this chapter, adds a new dimension in controllability of *C. elegans* neuronal network. Our studies suggest the importance of driver neurons and their utility to control the behaviour of the organism.

Controllability naturally raises two key questions: what are the points of control and what is to be controlled. Determination of control points in a network has been attempted with the help of various graph theoretical measures such as degree, betweenness centrality, closeness [Y. Tang et al., 2012]. The idea of control of brain states is aligned with the studies on control of behaviour (state) of an organism by identifying and controlling a few important regions (nodes) via external inputs (impulses of electric or magnetic fields). From a connectionist paradigm, brain could be thought of as a network of neurons, a complex dynamical system, the state of which is to be controlled. Network control has been studied as 'structural control' which could be achieved with the help of 'driver nodes/neurons'. It has been proposed that networks possessing cacti structure (without having inaccessible nodes or dilations) are controllable [Ching-Tai Lin, 1974]. A structural network with linear time invariant dynamical system could be represented as Eq.(4.1), where $x(t)$ represents the state of the system at time t, A is the state matrix, B input matrix and $u(t)$ is input signal.

$$x(t) = Ax(t) + Bu(t), \tag{4.1}$$

The state of such a system is proven to be controllable only if it possess full rank [Kalman, 1963; D. G. Luenberger, 2002].

4.1 PHENOFRAME

C. elegans is a model biological organism whose neuronal network is fully charted [White et al., 1986]. Its rudimentary nervous system, consisting of 302 neurons, is central to processing of complex information from sensory organs and that pertaining to behaviour and memory [Ardiel and Rankin, 2010]. The neurons are divided into various subtypes and are classified based on their functional roles, location within the body of the animal and span of the axons. According to functional roles, neurons are primarily of three types viz. sensory neurons, motor neurons and inter neurons. Sensory neurons pick up external signals to which the animal responds by sending

motor signals to effector organs through motor neurons which connect to command inter-neurons on dendritic side and neuro-muscular junction on the axonal side [Von Stetina, Treinin, and Miller, 2006]. Motor neurons are distributed mainly over the ventral nerve cord (VNC) with ganglia at each end [White, Southgate, Thomson, and Brenner, 1976], some of which extend their processes circumferentially to form a dorsal nerve cord (DNC) as shown in Figure 4.1. Both VNC and DNC control locomotion of the animal.

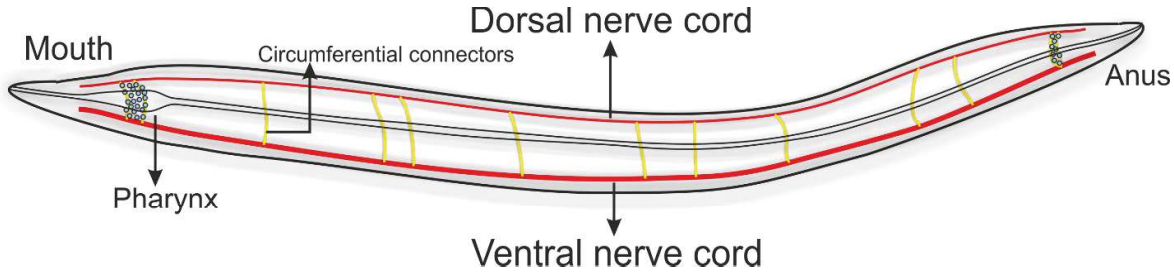

Figure 4.1 : Diagrammatic representation of *C. elegans* nervous system. The gastrointestinal tract lies in the middle of the body. Pharyngeal and circumferential ring neurons (yellow) are responsible for communication between dorsal nerve cord and ventral nerve cord.

In accordance with definition of driver nodes, these critical neurons are expected to control the state of neuronal network when provided with external input. To investigate this state space and what kind of changes one can bring by controlling driver neurons in *C. elegans* state, we examined phenotypic properties of these neurons. Study of properties such as location, functional type and span of neurons provided us with the potential functional association of driver neurons. Further we investigated specific biological functions underlying these neurons with the help of gene ontological enrichment studies.

The neurons from *C. elegans* neuronal network were classified on the basis of their anatomical locations (head, mid and tail), span (short and long), and function (sensory, motor and interneurons). The information of neuronal locations and span were obtained from White et al. [White et al., 1986], whereas that of functions were obtained from Hall and Russell [Hall and Russell, 1991]. Further, statistics of driver neurons were obtained in terms of their phenotypic properties (location, span and function) with the aim of characterizing them. Every neuron was thus characterized in terms of its 'phenoframe' which refers to a composite set of its phenotypic properties.

4.2 GENOFRAME

The functional association of gene products with the help of gene ontological studies can provide an insight to three biological domains: (a) biological processes, (b) cellular components and (c) molecular functions. Gene ontology (GO) enrichment process provides biological functions associated with the expression profile of a particular set of genes in the background of all the genes which are expressed in the organism. We performed a group level analysis using gene ontology enrichment to associate driver nodes expression profile to major biological functions.

To explore genetic underpinnings of driver nodes, we further created the gene co-expression network of *C. elegans* with the help of expression profiles of individual neurons. The expression data was obtained from WormWeb (www.wormweb.org), from where genes expressed in each group of neurons were collected. 297 neurons of *C. elegans* were classified into 116 groups on the basis of their expression profiles (Annexure B). Neurons within a group tend to have similar functional associations. *C. elegans* neurons were represented by 433 non-redundant genes, known to be expressed in neuronal cells. Towards creating a biologically

relevant network of genes underlying the neuronal architecture, a bipartite network of neuron-gene association was created. Further a weighted, unipartite neuronal gene co-expression network (GCN) was created (Figure 4.2). GCN depicts the expression profile similarity among neuronal genes, where every neuron was represented by a node and the number of genes commonly expressed in any two neurons is represented as a weighted edge.

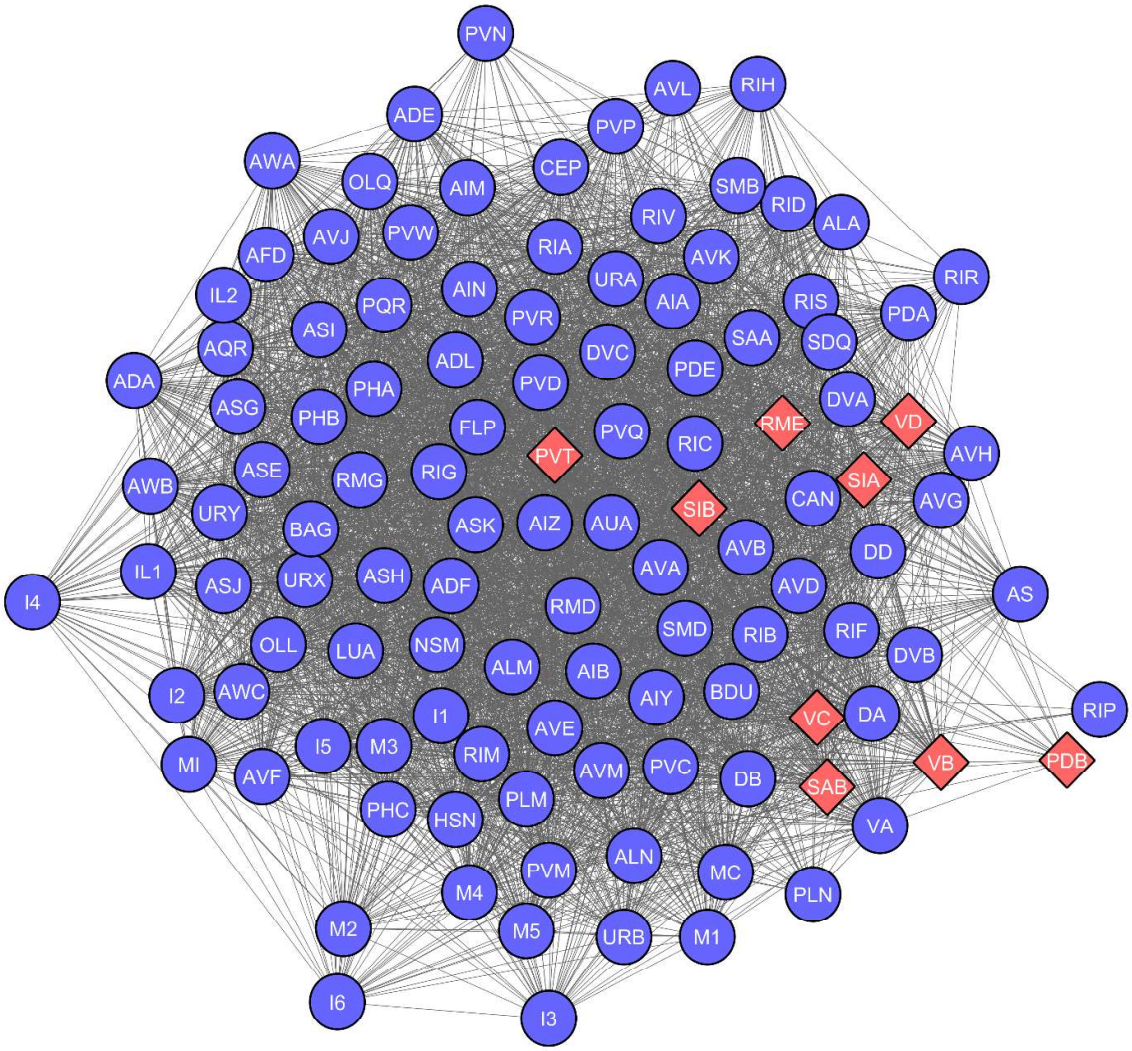

Figure 4.2 : The Gene Co-expression Network of CeNN. Each of the 116 neuronal groups is represented as a node. A weighted edge between any two neuronal groups represents the extent of gene co-expression. The shape and colour of the neuronal groups depict presence (Red Diamonds) or absence (Blue Circles) of driver neurons in them. The *C. elegans* co-expression based network is heterogeneous with 9 groups holding all the driver neurons, whereas rest of the 107 groups were devoid of driver neurons.

This gene co-expression network was further analysed and clustered using affinity propagation [Morris et al., 2011] clustering method in Cytoscape [Smoot, Ono, Ruscheinski, Wang, and Ideker, 2011], where affinity is based upon the shared number of genes between them.

4.2.1 Gene ontological enrichment

The importance of these clusters was found out by statistically analysing phenotypic properties of each cluster. From the clusters obtained we figured out the unique genes U_g which are expressed in each cluster. These unique genes were then analysed on the basis of gene ontological enrichment studies using Biological Networks Gene Ontology tool (BiNGO) [Maere, Heymans, and Kuiper, 2005] to find out the contribution of each cluster in specific biological

processes. Studies were performed to understand the uniqueness of specific genes expressed for each cluster on the basis of biological processes. Two types of gene ontological enrichment studies were performed to infer role of clusters in specific ontological processes. The first enrichment study of genes from each cluster was done against a background of all *C. elegans* genes, G (Eq.(4.2)). In a more refined enrichment study, the genes from the clusters were enriched against a subset of *C. elegans* genes that are expressed only in neurons, G_n (Eq.(4.3)).

$$Gene - enritchment \rightarrow U_g \subset G \quad (4.2)$$
$$Gene - enritchment \rightarrow U_g \subset G_n \quad (4.3)$$

4.2.2 Essentiality of genes

With the objective of associating genes central to driver neurons, we linked essentiality of genes to those obtained from GO enriched genes belonging to three major clusters of GCN. Essentiality of genes was attributed with the help of 'Database of essential genes' [Kamath et al., 2003; Luo, Lin, Gao, Zhang, and Zhang, 2014].

4.3 RESULTS

We perceive the *C. elegans* neuronal network as a complex adaptive system to look for mechanisms of its control. The worm has evolved to acquire the present neuronal architecture by adapting to biotic and abiotic stresses over a long period of time. Driver neurons is a novel concept associated with a subset of neurons which when driven by an external input allows one to control the state of the whole network. In this work, we aimed to identify driver neurons of *C. elegans* and to associate them with their phenotypic (phenoframe) as well as genotypic (genoframe) features. As part of phenoframe, we characterised driver neurons based on location, span and functional types. As part of genoframe, we identified the genetic underpinnings of the driver neurons. Number of driver neurons in *C. elegans* neuronal connectivity network are higher (16.5%) as compared with its degree distribution conserved (9.6%) and random network (0.3%) controls. Interestingly driver neurons tend to avoid hubs and have fewer synaptic connections (Figure 4.3).

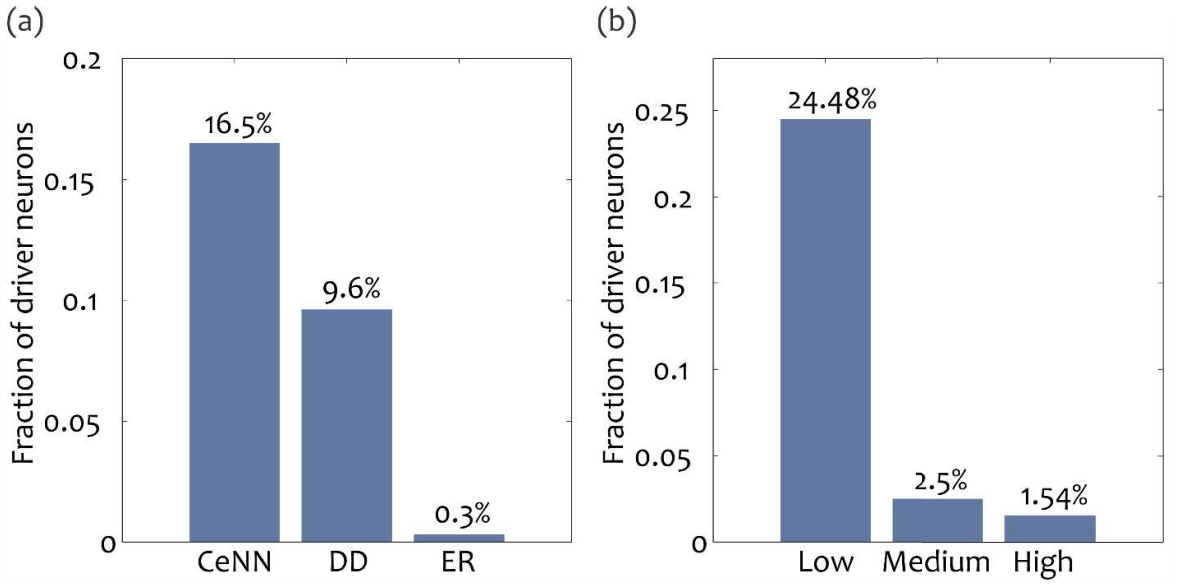

Figure 4.3 : Role of degree in specifying driver nodes. (a) Fraction of driver nodes as found in the CeNN, and their corresponding random counterparts: Degree Distribution conserved control (DD) and Erdos-Reney graph (ER). (b) Fraction of driver neurons with low, medium and high degree in *CeNN*. These results are consistent with the reports of Lui et al. [Y.-Y. Liu et al., 2011].

4.3.1 Phenoframe of driver neurons

The notion of driver neurons implies control over the state of network with minimum extent of external inputs. While we are acutely aware that the exact implication of control of neuronal system is not clear, we believe that the notion of driver neurons offers us quantitative metrics of assessing control in such a simple neuronal system. Study of phenotypic properties of neurons led us to the observation that the largest proportion (29.91%) of the driver neurons identified with maximum matching algorithm were associated motor activities, followed by inter neurons (11.8%) and sensory neurons (1.25%) as shown in Figure 4.4. We therefore conclude that motor neurons are the primary means of achieving desired state of neuronal activity in the phenoframe of *C. elegans*. This is an interesting observation given that one might be tempted to hypothesize that sensory neurons are critical for driving the state of the neuronal network.

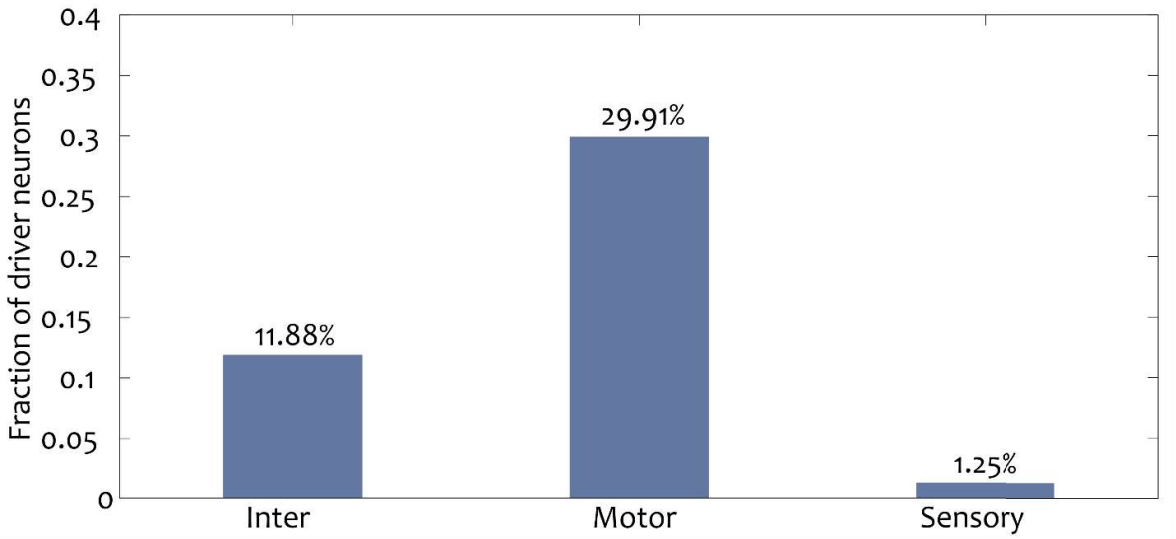

Figure 4.4 : Distribution of driver neurons across different types of neurons: Sensory, Motor and Inter. The fraction of driver neurons was computed for each class separately.

As part of phenoframe analysis, we further observed association of driver neurons on the basis of location of neuron and their span as shown in Figure 4.5. Driver neurons were populated in the middle of the body (29.7%) whereas head and tail regions had sparse representation; 9.09% and 6.25% respectively (Figure 4.5(a)). Across different length spans of neurons, 24.7% of all the short spanned neurons were driver neurons. On the other hand, among the long span neurons only 4.55% were driver neurons as depicted in Figure 4.5(b). Interestingly, these phenotypic properties are also shared by VNC neurons that are known to control activities of bodily movements of animal. This could be interpreted to state that driver neurons which share phenotypic markers with VNC neurons could have evolved to control the movement of the animal thus changing its physical behaviour (Figure 4.6).

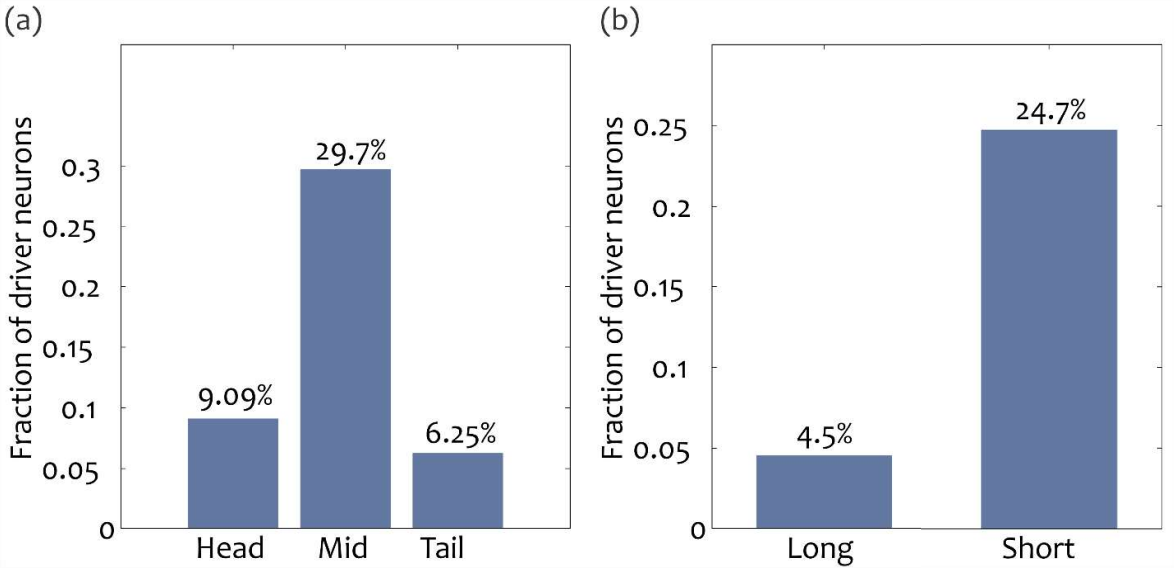

Figure 4.5 : Distribution of driver neurons across other phenotypic features (location and span). (a) Location of neurons within the body of the organism. (b) Span of neurons in accordance with the length of axons.

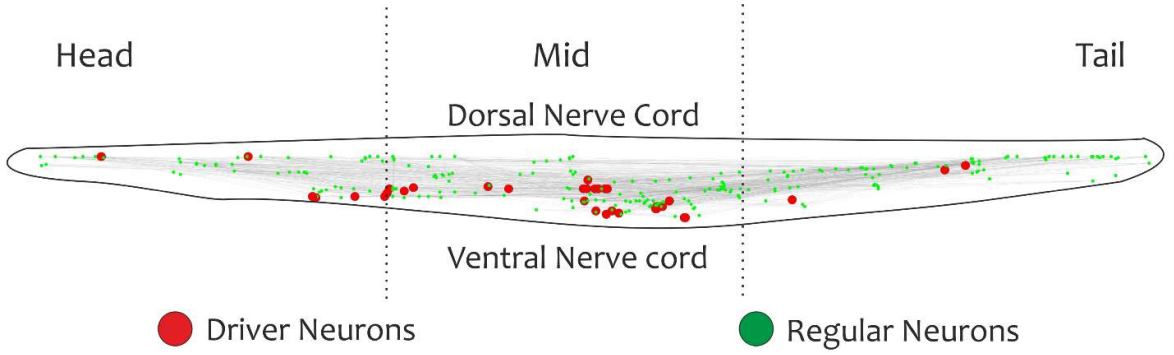

Figure 4.6 : Visualization of driver neurons in CeNN and distribution of driver neurons across the body. The neurons are arranged in accordance with Cartesian coordinates presented within the body. This figure clearly shows presence of driver neurons in the mid-ventral region of the organism.

4.3.2 Genoframe of driver neurons

Moving beyond phenotypic properties, we further explored the genetic association of driver neurons. Gene expression pattern of a cell plays major role in specifying its biological function. We aimed to identify genetic correlates of driver neurons, named as genoframe. Starting from the gene co-expression network (Figure 4.2), we identified clusters of neurons with the assumption that neurons belonging to the same cluster could be associated with similar functions (Figure 4.7). We obtained 6 major clusters. Interestingly three largest clusters comprised of groups of driver neurons. Further we identified unique genes that characterise each of the clusters. Our intention was to obtain ontological features that characterise driver neurons as well as group of neurons that comprise of driver neurons.

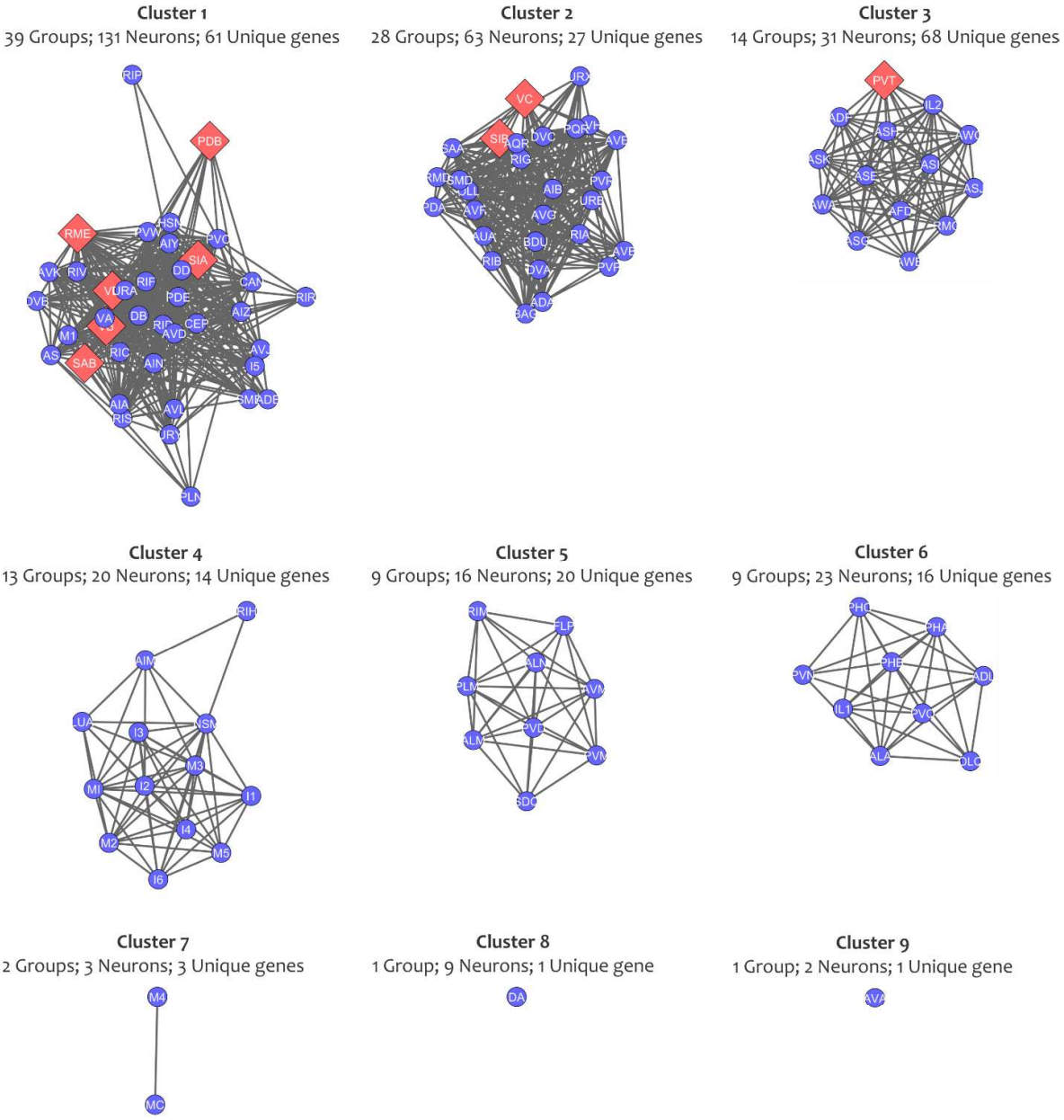

Figure 4.7 : Nine neuronal groups (clusters) obtained starting from GCN using affinity propagation algorithm [Morris et al., 2011]. The three largest clusters contain driver nodes (red diamonds) hinting at possible role they play. Cluster three forms a perfect clique with highest number of unique genes expressed within a cluster.

Phenoframe of clustered GCN shows that the neuron groups in cluster 1 contains the highest number of driver neurons which are preceded by cluster 2 and 3 as shown in Figure 4.7. The phenotypic properties of these clusters (especially cluster 1) resemble that of driver neurons (Figure 4.8). Among the non-trivial clusters comprising of neurons from multiple groups, Cluster 3 was predominantly represented by sensory neurons (Figure 4.8(a)). Cluster 2, Cluster 5 and Cluster 6 were also dominated by sensory neurons. In terms of neuronal positions, Cluster 3 was dominated by neurons that were located in head region (Figure 4.8(b)). Cluster 6 was composed of neurons located in head as well as tail regions, with no representation of mid neurons. Cluster 3 and Cluster 6 predominantly contained long neurons, whereas Cluster 1 and Cluster 2 were dominated by short neurons (Figure 4.8(c))

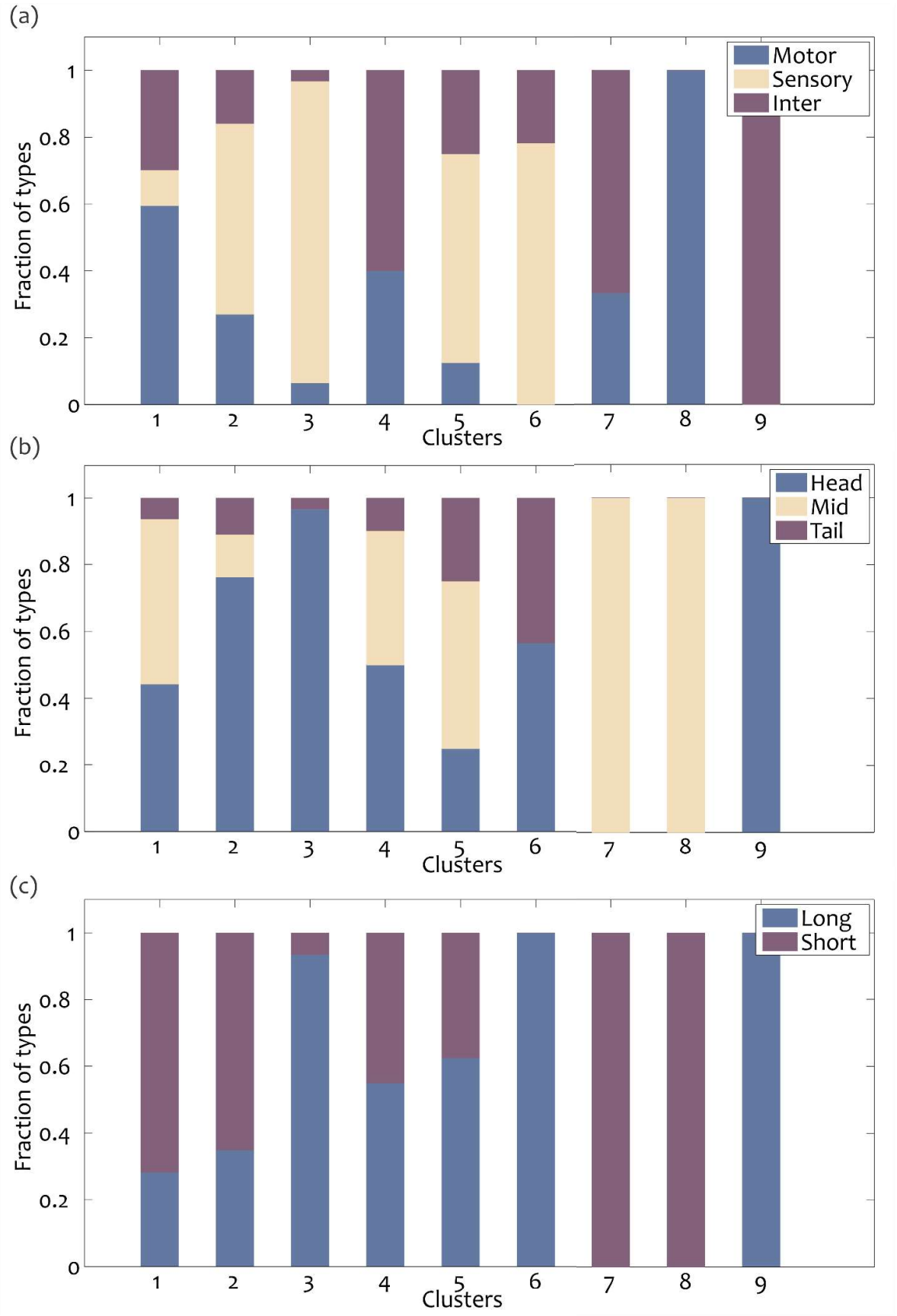

Figure 4.8 : Phenotypic distribution of neurons in the clusters of GCN in accordance with, (a) functional types, (b) location of the body, (c) Span of the neuronal axon.

Ontological studies of genes from driver neurons

We propose that the neuronal clusters containing driver neurons may play specific roles critical for evolutionary survival of the organism. To obtain ontologically relevant features, we performed gene ontological enrichment studies of the genes expressed in the clusters with driver neurons against the background of whole genome of *C. elegans*. We used three largest clusters with driver neurons for the purpose of GO enrichment studies.

We identified genes in each cluster that are unique (specifically expressed in a cluster) to it. We expect that significant ontological terms thus obtained may help us to identify biological processes and molecular functions central to the control. Gene ontological study showed that Cluster 1 was mainly associated with (a) signaling process, (b) reproductive processes and (c) anatomical structural development of the animal, and could be relevant for controlling the movement of cells and associated processes. Gene ontological studies performed for Cluster 1 in the background of genes expressed exclusively in neuronal genes indicated an ontological association with reproduction and cellular localization. The ontological coordinates for Cluster 2 and 3 were mainly concerned with biosynthetic processes. Thus we surmise that few of the most biological processes relevant to *C. elegans* are captured through the studies performed with clustering of GCN and gene ontological enrichment studies. Controlling driver neurons may have an implicit effect on cellular localization (p-value = 5.61×10^{-4}) thus changing the anatomical structural development (p-value = 9.42×10^{-4}) of the organism which in turn may have direct correlation with the state of the neuronal network. Another observation that was made from GO enrichment studies is the involvement of multi-organism reproductive processes (p-value = 6.94×10^{-4}), which implies that these group of neurons can also control its reproductive behaviour and hence could be of potential relevance for evolution of the animal.

Genes ontologically relevant to driver neurons are not biologically essential

Constructing a network of gene co-expression networks and clustering them facilitated gauging the role of genes that are critically related with driver neurons. Table 4.1 lists these genes for each of the three major clusters of GCN. None of the neuronal genes including those critical for driver neurons were found to be biologically essential.

Table 4.1 : List of gene obtained within unique set of genes after gene ontological enrichment of each clusters.

	Genes
Cluster 1	ldb-1, jnk-1, daf-10, syd-2, pll-1, syd-1, lin-14, gsa-1, hbl-1, tol-1, lat-1, unc-10, unc-47, ast-1, syd-2, cat-2, unc-40, hbl-1
Cluster 2	gcy-9, ets-5, gcy-33, gcy-31, gcy-32, gcy-34, gcy-25, fax-1, gcy-36, gcy-37, daf-16, gpa-8
Cluster 3	daf-11, gcy-6, gcy-5, gcy-4, gcy-3, gcy-14, gcy-8, odr-1, gcy-20, gcy-15, gcy-7, sma-6, gcy-22, gcy-23, gcy-19, sad-1, gcy-27, nhr-69, odr-7, skn-1, dsc-1, trx-1, gpa-5, egl-30, odr-3, gpa-4

4.4 DISCUSSION

Neurons play a very vital role in controlling the functional and bodily behaviour of the animal. In early parts of evolution, individual neurons may have been associated with a specific function, but with time neurons may have started to decentralize functional tasks. The neuronal network of *C. elegans* is one such example where the body of ~1000 cells is governed by ~300 neurons. The map of *C. elegans* neuronal network is complete and almost static. Similar to most other real world networks, this network is characterised with scale free topology, high average clustering coefficient ($C = 0.172$), and low characteristic path-length ($L = 2.64$), making it a good candidate for the study of controllability [Watts and Strogatz, 1998].

As compared to other measures of centrality, the unmatched nodes are shown to be efficient in controlling the network state using structural control theorem by Liu et. al. [Y.-Y. Liu et al., 2011]. We utilised Hopcroft-Karp algorithm for finding unmatched nodes with computational complexity of $O(\sqrt{N} \times E)$ [Hopcroft and Karp, 1973], which makes its application on larger biological networked systems feasible. Our results highlight the possible role of connectivity of neurons, as the number of driver neurons in degree distribution preserved control (9.6%) is closer to that of *C. elegans* neuronal network (16.5%) in contrast of a random control (0.3%).

The neurons of *C. elegans* are divided into various classes on the basis of their association to function, position and length. Finding a few important target neurons that are critical for achieving full control over the behaviour of the animal takes us closer to its control mechanisms. The target neurons (also known as driver neurons) are defined in terms of structural controllability. Driver neurons were further classified based on their phenotype and interestingly, it was observed that a large number of such nodes are of motor type, and are located in the middle of the body with a short span. This gives us an insight into control mechanisms of the animal, thereby providing practical ways of controlling animal behaviour with the help of external stimulus. The phenoframe of *C. elegans* suggested the presence of these neurons in ventral nerve cord of animal.

Gene ontological enrichment studies bring out unique function of genes in the background of housekeeping genes. By performing enrichment studies of specific groups of neurons with high number driver neurons we conclude that the driver neurons are primarily associated with reproductive behaviour of *C. elegans*, as compared to other functional associations. This may also imply that genetic make-up of regular neurons is largely concerned with growth and development of the organism. Driver neurons on the other hand are related to passing the genetic message to its progeny.

Thus targeting external signal towards the VNC can bring changes in the animal reproductive behaviour which in turn is speculated to have effect on the entire behavioural state of the animal. Further the importance of motor memory is also reflected from the fact that most of the driver neurons are involved in motor functions. This gives us a scope for controlling the behavioural state of the organism by stimulating motor neurons.

…