

# Polymer physics-based classification of neurons

Kiri Choi, Won Kyu Kim, and Changbong Hyeon\*

*Korea Institute for Advanced Study, Seoul 02455, Korea*

## Abstract

Recognizing that diverse morphologies of neurons are reminiscent of structures of branched polymers, we put forward a principled and systematic way of classifying neurons that employs the ideas of polymer physics. In particular, we use 3D coordinates of individual neurons to calculate the form factor,  $F(q)$ , a Fourier transform of density-density correlation of particles comprising an object of interest. For a polymer-like object consisting of  $n$  monomers,  $F(q)$  scales with the wavenumber  $q$  as  $F(q) \sim q^{-\mathcal{D}}$  at an intermediate range of  $q$ , where  $\mathcal{D}$  is the fractal dimension or the inverse scaling exponent ( $\mathcal{D} = \nu^{-1}$ ) characterizing the geometrical feature ( $r \sim n^\nu$ ) of the object over a length scale  $r (= 2\pi/q)$ .  $F(q)$  can be used to describe a neuron morphology in terms of its size ( $R_n$ ) and the extent of branching quantified by  $\mathcal{D}$ . By assessing the similarity between  $F(q)$ s calculated for a set of neurons, we tackle the neuron classification problem. Our  $F(q)$ -based classification, applied to publicly available neuron datasets from three different organisms, not only complements other classification methods but also offers a physical picture of how the dendritic and axonal branches of an individual neuron fill the space of dense neural networks inside the brain.

---

\*hyeoncb@kias.re.kr

## INTRODUCTION

Form determines function in biology. Neurons, which are the basic signaling units of the nervous system, are not an exception [1]. Since Ramón y Cajal provided evidence that this general hypothesis is also at work in neurons [2], the problem of classifying neurons based on their morphology has been a subject of considerable interest in brain research. Along with a causal relationship for the correlation between the neuronal morphologies and spiking patterns of electrophysiological recordings [3], the morphological detail of a neuron has been suggested as one of the key determinants of physiology and functional differentiation of neurons [4–8], engendering a number of morphology-based classification methods [9–23].

Neurons, demonstrating a variety of elaborate arborization and branching patterns, are considered fractal objects, which has led several studies to calculate fractal dimensions to characterize neurons by the box-counting method [7, 24–28]. As a physical object, many neurons entangled and filling the space inside the brain are reminiscent of a solution or a melt of branched polymers. Since structures of polymer chains in various solvent conditions or polymer concentrations have long been investigated with scattering experiments [29, 30], we put forward calculating the form factor  $F(q)$  to quantify the structures of individual neurons, and then grouping the neurons based on the similarity between the calculated  $F(q)$ s.

Here we apply our  $F(q)$ -based analysis of neurons to publicly available datasets reconstructing neuron morphologies from (i) *C. elegans* nervous system [31], (ii) *Drosophila* olfactory projection neurons [13], and (iii) the mouse primary visual cortex (V1) neurons in the Allen Cell Type database [11]. The  $F(q)$ -based analysis leverages the full 3D coordinates of neurons reconstructed from images without resorting to any prior knowledge of the neuron morphology. The outcomes from our classification method are found comparable to other existing analyses; yet, they offer morphological insights into neurons in the language of branched polymers under various conditions.

## THEORETICAL BACKGROUND

### Two length scales in neuron morphology

Individual neurons tend to feature densely branched structures, and hence their morphological characteristics can be analyzed in terms of polymer physics describing the conformation of

branched polymers. Since branched polymers are composed of multiple sections of linear polymers, two distinct ranges of length scale ( $r$ ) are of relevance in our discussion (Fig. 1A) [32]: (i) The first range is defined at  $\bar{b} < r < \bar{l}$ , where  $\bar{b}$  is the mean segment size defined between the data points in the reconstructed neurons and  $\bar{l}(= N_l^{-1} \sum_i^{N_l} l_i)$  is the average length of linear sections (branches) between branch points or from one branch point to an axonal or dendritic tip. There is a great variation in length ( $l_i$ ) (Fig. 1B). The neurons over this scale are effectively described as a linear polymer. If the flexibility of a polymer in this length scale is large, one could divide the polymer chain into multiple segments of Kuhn length; yet we have found that neurons in this length scale are typically stiff, lacking flexibility. (ii) The second range is defined at  $\bar{l} < r < R_n$ , where  $R_n$  is the average size of the neuron, which is tantamount to the gyration radius ( $R_g^2 = \frac{1}{2N^2} \sum_{i=1}^N \sum_{j=1}^N (\mathbf{r}_i - \mathbf{r}_j)^2$ ) for a globular object (i.e.,  $R_n = R_g$ ). In this range, the structural features of a neuron can be best represented through those of branched polymers.

## Form Factor

Regardless of whether a given object is a polymer chain or a neuron, if the object is composed of  $N$  monomers whose relative displacement in space is given as  $\mathbf{r}_i - \mathbf{r}_j$  ( $i, j = 1, 2, \dots, N$ ), the corresponding form factor  $F(q)$  [33–36] can be calculated using

$$F(q) = \frac{1}{N} \left\langle \sum_i^N \sum_j^N e^{-i\mathbf{q} \cdot (\mathbf{r}_i - \mathbf{r}_j)} \right\rangle = \frac{1}{N} \sum_i^N \sum_j^N \frac{\sin(qr_{ij})}{qr_{ij}}, \quad (1)$$

where  $\mathbf{r}_i$  is the position vector of the  $i$ -th segment (or monomer), and  $\mathbf{q}$  is the scattering wave vector. Averaging over the angle between  $\mathbf{q}$  and  $\mathbf{r}_i - \mathbf{r}_j$  leads to the second expression with  $q = |\mathbf{q}|$  and  $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ .  $F(q)$  corresponds to the Fourier transform of the pair correlation function of the density field of monomers at two positions,  $\rho(\mathbf{r}_\alpha) = \sum_i \delta(\mathbf{r}_\alpha - \mathbf{r}_i)$  ( $\alpha = 1, 2$ ). With the density field  $\rho(\mathbf{r}_\alpha)$ , Eq. 1 is expressed in the form of  $F(q) = \frac{1}{N} \int \int \langle \rho(\mathbf{r}_1) \rho(\mathbf{r}_2) \rangle e^{-i\mathbf{q} \cdot (\mathbf{r}_1 - \mathbf{r}_2)} d\mathbf{r}_1 d\mathbf{r}_2$ , which can be further cast into  $F(q) = \frac{1}{N} \int \int \langle \rho(\mathbf{r}_{12}) \rho(\mathbf{r}_2) \rangle e^{-i\mathbf{q} \cdot \mathbf{r}_{12}} d\mathbf{r}_{12} d\mathbf{r}_2 = \int p(\mathbf{r}) e^{-i\mathbf{q} \cdot \mathbf{r}} d\mathbf{r}$ . Here, the term  $\langle \rho(\mathbf{r}_{12}) \rho(\mathbf{r}_2) \rangle$  is the joint probability that one monomer is located at  $\mathbf{r}_{12} (\equiv \mathbf{r}_1 - \mathbf{r}_2)$  relative to another monomer at position  $\mathbf{r}_2$ , which is equivalent to the conditional probability of the monomer at  $\mathbf{r}$  given that another monomer is at the origin ( $\mathbf{r}_0$ ) multiplied by the mean local monomer density, i.e.  $\langle \rho(\mathbf{r}_{12}) \rho(\mathbf{r}_2) \rangle = \langle \rho(\mathbf{r}|\mathbf{r}_0) \rangle \rho(\mathbf{r}_0) = p(\mathbf{r}) \rho(\mathbf{r}_0)$ .

For  $\bar{l} \ll r \ll R_g$  with  $r \equiv |\mathbf{r}|$ ,  $p(r)$  corresponds to the density of monomers inside a volume of a radius  $r$ , i.e.,  $p(r) \simeq n/r^d$  where  $n$  is the average number of monomers inside the volume  $\sim r^d$  at  $d$  dimensions. There is a scaling relationship between  $n$  and  $r$ ,  $r \sim n^\nu$  with  $\nu$  the size scaling exponent (or the inverse fractal dimension  $\mathcal{D} = 1/\nu$  that satisfies  $n \sim r^\mathcal{D}$ ), which reflects how the  $n$  monomers comprising the object are organized over the space defined by the length scale  $r$  [37]. The dimensional analysis with  $p(r) \sim r^{1/\nu-d}$  and  $d\mathbf{r} \sim r^{d-2}dr$  yields [33]

$$F(q) \sim \int d\mathbf{r} p(\mathbf{r}) e^{-i\mathbf{q} \cdot \mathbf{r}} \sim q^{-1/\nu} \sim q^{-\mathcal{D}}. \quad (2)$$

Depending on the wavenumber in the range of either  $2\pi/\bar{l} < q < 2\pi/\bar{b}$  or  $\pi/R_n < q < 2\pi/\bar{l}$  which translates into the physical length scale via  $r = 2\pi/q$ ,  $F(q)$  reveals different structures of neurons in terms of the scaling exponent  $\nu$  (or  $\mathcal{D}$ ).

On the other hand, for  $qR_g \sim 1$ , corresponding to the Guinier regime,

$$\begin{aligned} F(q) &\sim \int d\mathbf{r} p(\mathbf{r}) \left[ 1 - i(\mathbf{q} \cdot \mathbf{r}) - \frac{1}{2}(\mathbf{q} \cdot \mathbf{r})^2 + \dots \right] \\ &= N \left[ 1 - \frac{q^2 \langle R_g^2 \rangle}{3} \dots \right] \simeq F(0) e^{-\frac{q^2 \langle R_g^2 \rangle}{3}}. \end{aligned} \quad (3)$$

The radius of gyration of an object,  $R_g^2 = \frac{1}{2N} \int \mathbf{r}^2 p(\mathbf{r}) d\mathbf{r}$ , can be obtained from the slope of  $\log F(q)$  vs  $q^2$  at small  $q$ , such that  $\langle R_g^2 \rangle = -3 \lim_{q \rightarrow 0} \frac{\partial \log F(q)}{\partial q^2}$ .

In what follows, we review the scaling exponents of a linear or a branched polymer chain under various conditions [32, 33, 36].

## A polymer chain in dilute solution

According to the basic theory of polymer physics, the size ( $R_F$ , Flory radius) of a polymer chain scales with its length  $N$  as  $R_F \sim N^\nu$ , where the scaling exponent  $\nu$  is determined by the solvent quality. In three dimensions ( $d = 3$ ), a single flexible linear polymer adopts swollen conformations in a good solvent with  $\nu = 0.588$  but collapses to a globular form in a poor solvent with  $\nu = 1/3$  [33]. At the  $\Theta$ -condition, which is a tri-critical point between a good and poor solvent condition where the attraction and repulsion at the level of two-body interaction compensate each other, a flexible polymer adopts conformation like those of ideal polymer that obeys the scaling law of  $R_F \sim$

$N^{1/2}$  [33, 38, 39]. Meanwhile, for a stiff polymer characterized by a large persistence length, one finds  $R_F \sim N$  with  $\nu = 1$  regardless of the solvent quality.

For  $\bar{l} < r < R_n$ , the neuron morphology can be studied by considering the structure of a branched polymer. In general, the Flory free energy of a polymer consisting of  $N$  monomers with the statistical segment of size  $b$  can be written as

$$\beta F(R) \approx \frac{R^2}{\langle R_0^2 \rangle} + \frac{B_2}{2!} \langle \rho^2 \rangle R^d + \frac{B_3}{3!} \langle \rho^3 \rangle R^d + \dots . \quad (4)$$

The first term  $\langle R_0^2 \rangle$  is the mean square size of an ideal polymer:  $\langle R_0^2 \rangle \simeq b^2 N$  for linear polymers, whereas  $\langle R_0^2 \rangle \simeq b^2 N^{1/2}$  for randomly branched ideal polymers [36, 40]. The first term constrains the size of the polymer purely elastically so that any attempt to elongate the polymer engenders a restoring force of  $-kR$  with the spring constant of  $k \simeq 1/\langle R_0^2 \rangle$ . The remaining terms arise from the interaction between the monomers whose effective density inside the pervaded volume of a polymer of size  $R$  is given as  $\langle \rho \rangle \simeq N/R^d$ . The second and third terms are due to two-body and three-body interactions.  $B_2$  is the second virial coefficient, which decides the strength of the two-body interaction;  $B_2 > 0$  is for the interaction between monomers in good solvent conditions, making the two-body interaction repulsive. Under a simplifying assumption of zero correlation ( $\langle \rho^n \rangle = \langle \rho \rangle^n$ ), the Flory free energy up to the two-body interaction term is written as

$$\beta F(R) \approx \frac{R^2}{b^2 N^{1/2}} + \frac{B_2}{2!} \frac{N^2}{R^d}, \quad (5)$$

which is minimized via  $\partial F(R)/\partial R|_{R=R_F} = 0$  with  $B_2 = v$ , yielding

$$R_F \simeq b \left( \frac{v}{b^d} \right)^{\frac{1}{d+2}} N^{\frac{5}{2(d+2)}}. \quad (6)$$

Thus, the size scaling exponent of a randomly branched polymer in a good solvent at  $d = 3$  is  $\nu = 1/2$  ( $\mathcal{D} = 2$ ) [32].

Under the  $\Theta$ -solvent condition, the second virial coefficient vanishes ( $B_2 = 0$ ), and the Flory radius of a branched polymer is determined from the balance between the elastic free energy and the three-body interaction term with  $B_3 = w$  in Eq.4 as

$$R_F \simeq b \left( \frac{w}{b^{2d}} \right)^{\frac{1}{2(d+1)}} N^{\frac{7}{4(d+1)}}. \quad (7)$$

Thus, the size scaling exponent of a randomly branched polymer under  $\Theta$ -solvent condition at  $d = 3$  is  $\nu = 7/16$  ( $\mathcal{D} = 1/\nu = 2.29$ ) [32].

TABLE I: Scaling exponent  $\nu$  of polymer chain under various conditions.

Type	Solvent property	$\nu$	$\mathcal{D}$
Linear polymer in dilute solution	Poor	1/3	3
	$\Theta$ -condition	1/2	2
	Good	0.588	1.70
Branched polymer in dilute polymer solution	Ideal	1/4	4
	$\Theta$ -condition	7/16	2.29
	Good	1/2	2
Branched polymer in semidilute polymer solution	Semi-dilute (good solvent)	0.395	2.53
Branched polymer in dense polymer solution	Melts	1/4	4

### A polymer chain in semi-dilute/dense polymer solution.

When the concentration of polymer solution increases, the individual chains start overlapping, influencing the chain conformation of neighboring polymers, which defines the overlap volume fraction  $\Phi_c \sim Nb^3/R_F^3 \sim N^{1-3\nu}$ . In the semi-dilute regime ( $\Phi_c \lesssim \Phi \ll 1$ ), it is no longer discernible whether two spatially adjacent monomers are from the same chain or different chains, and the original strength of the intra-polymer two-body interaction ( $B_2 = v$ ) is screened to yield  $B_2 \simeq v/N^{1/2}$ . In this case, the Flory free energy is written as

$$\beta F(R) \approx \frac{R^2}{b^2 N^{1/2}} + \frac{1}{2} \left( \frac{v}{N^{1/2}} \right) \frac{N^2}{R^d}, \quad (8)$$

and  $\partial F(R)/\partial R \Big|_{R=R_F} = 0$  yields the Flory radius of

$$R_F = b \left( \frac{v}{b^d} \right)^{\frac{1}{d+2}} N^{\frac{2}{d+2}}. \quad (9)$$

Thus, the size scaling exponent of a randomly branched polymer in a semi-dilute polymer solution under good solvent condition is  $\nu = 2/(d+2) = 2/5 = 0.4$ . This number is close to the exponent  $\nu = 0.395$ , the inverse fractal dimension of percolating clusters at percolation threshold ( $p = p_c$ ),  $\mathcal{D} = 1/\nu = 2.53$ , estimated from computer simulations [41].

For dense melts of branched polymers ( $\Phi_c \ll \Phi \approx 1$ ), the interactions between polymer segments are fully screened. In this case, individual branched polymers behave as if they are effectively in an *ideal* condition, giving rise to the scaling relationship of  $R_F \sim N^{1/4}$ , and interesting scaling  $F(q) \sim q^{-4}$  emerges in the range of intermediate  $q$  ( $\pi/R_n < q < 2\pi/\bar{l}$ ). Since the linear sections of neurons ( $\bar{b} < r < \bar{l}$ ) are stiff and effectively rigid rod-like, the length of the statistical segment (Kuhn length) is  $\sim \bar{l}$ , and the number of statistical segments comprising the branched polymer is  $\bar{b}N/\bar{l}$ . Therefore, the size of polymer scales as  $R_F \simeq \bar{l}(\bar{b}N/\bar{l})^{1/4} = \bar{b}(\bar{l}/\bar{b})^{3/4}N^{1/4}$ . The fractal dimension of  $\mathcal{D} = 4$  may sound geometrically impossible; however, an object with fractal dimension  $\mathcal{D} = 4$  is still permissible in the range of  $\bar{l} < r < R_F$  because the stiff linear section creates extra space at  $\bar{b} < r < \bar{l}$  [36].

Although it is not directly relevant to our problem that discusses branched polymer-like neurons comprised of rigid segments with large persistence length, a scaling behavior of  $F(q) \sim q^{-4}$  could also emerge in the Porod scattering  $F(q) \sim q^{-(6-\mathcal{D}_s)}$  at large  $q$  regime ( $q \lesssim b^{-1}$ ) [42–44] when the surface of a compact globule formed by a *flexible* chain under a poor solvent condition, contributing to the scattering intensity, is smooth and has a surface fractal dimension of  $\mathcal{D}_s = 2$ .

The scaling exponents (fractal dimensions) for different types of polymers are summarized in Table I.

## RESULTS

Here we present our  $F(q)$ -based analysis on the datasets of neuron morphology reconstructions from three distinct organisms with an increasing degree of complexity (Fig. 2): (i) the *C. elegans* nervous system, (ii) projection neurons in the *Drosophila* olfactory system, and (iii) mouse primary visual cortex neurons.

### Neurons in *C. elegans*

*C. elegans* is one of the organisms whose neural connectivity has been fully mapped. Morphological reconstruction of the *C. elegans* nervous system is available as a part of the OpenWorm initiative which created a complete connectome of the organism [31] (see Fig. 2A). *C. elegans* has 302 neurons with their morphologies and inter-neural connectivity being fully specified [45, 46].

Because of its relative simplicity, it is an ideal organism to explore the basis of neural dynamics and brain function [47–49]. Partial, but more detailed reconstructions are also available for small parts of the sensory neurons [50–52]. Complex neural behaviors of both associative and non-associative learning are realized through these neurons [46] which are functionally classified into sensory, motor, interneurons, polymodal, and unknowns [31, 46, 53, 54] (Fig. 3). Respective circuits for specific behaviors such as locomotion and chemosensory responses are well-documented as well [46, 48, 51, 55].

The overall architecture of neurons in the *C. elegans* nervous system largely conforms to the organism's tubular body shape, which is reflected in the  $F(q)$  plots of large neurons classified into clusters  $C_1^{CE}$ ,  $C_2^{CE}$ , and  $C_3^{CE}$ . The neurons in these clusters exhibit  $\mathcal{D} \approx 1$  at the wavenumber smaller than the dips formed at  $q \approx 2\pi/d_C (\approx 0.13 \mu\text{m}^{-1})$  (dotted lines in Fig. 4A) where  $d_C (\approx 47.9 \mu\text{m})$  is the average diameter of the cross-section [56, 57]. The 2/3 of neurons ( $N = 198$ ) come in a pair projecting on the left and right sides of the body, and the rest ( $N = 104$ ) are largely confined either in the head or in the tail region.

The  $F(q)$ -based clustering groups 302 neurons into five clusters based on their overall morphological features (Figs. 4 and 5). (i) The neurons involved with the interneurons ( $N = 34$ ) and motor function ( $N = 37$ ), which span across the entire body, are grouped into the cluster  $C_1^{CE}$  ( $N = 85$ ). The rest of the cluster is composed of 6 sensory neurons, a polymodal neuron, and 6 unknowns. (ii) The cluster  $C_2^{CE}$  is made of 2 PVD neurons (PVDL and PVDR), which are known to be responsible for the sensory detection of drastic changes in mechanical forces, osmolarity, and temperature [58–60]. They feature a dendritic morphology of exceptional branching patterns repeating across the entire body (Fig. 4B). The  $F(q)$  of the PVD neurons scales  $F(q) \sim q^{-2}$ , exhibiting the fractal dimension of  $\mathcal{D} = 2$  at  $2\pi/d_C < q < 2\pi/\bar{l}$  (Fig. 4A), indicating that the morphology of PVDs with intricate branching pattern ( $C_2^{CE}$  of Fig. 4B) is similar to the configuration of a self-avoiding branched polymer in dilute solution. (iii) The neurons in the cluster  $C_3^{CE}$  ( $N = 44$ ), structurally similar to those in  $C_1^{CE}$  but are shorter in branch length ( $\bar{l}$ ), are mostly motor neurons ( $N = 32$ ). The cluster also contains a small number of sensory ( $N = 5$ ) and interneurons ( $N = 7$ ). (iv) The clusters  $C_4^{CE}$  and  $C_5^{CE}$  are comprised of small to mid-sized neurons mainly confined around the head region. Unlike the other neuron clusters, the neurons grouped in cluster  $C_4^{CE}$  ( $N = 100$ ) are diverse in terms of neuron type composition, containing 44 sensory neurons, 22 interneurons, 22 motor neurons, 6 polymodal neurons, and 6 unknowns. The cluster  $C_5^{CE}$  ( $N = 71$ ) consists of 4

sensory neurons, 33 interneurons, 26 motor neurons, and 4 polymodal neurons. The clusters  $C_4^{\text{CE}}$  and  $C_5^{\text{CE}}$  can be further classified into smaller clusters by repeating the application of the same clustering algorithm to each cluster ( $C_4^{\text{CE}}$  and  $C_5^{\text{CE}}$  in Fig. 5A, and Figs. 5B and 5C).

Taken together, it is found that the motor and interneurons grouped in  $C_1^{\text{CE}}$  and  $C_3^{\text{CE}}$  are branchless and extend across the entire body of *C. elegans*. In contrast, the neurons with sensory function (PVDS) grouped in  $C_2^{\text{CE}}$  are characterized by the pattern of extensive branching. Although the neurons classified in the clusters  $C_4^{\text{CE}}$  and  $C_5^{\text{CE}}$  are still in a mixture of many neurons with different functional types, their sizes (contour lengths) are smaller than those in  $C_1^{\text{CE}}$ ,  $C_2^{\text{CE}}$  and  $C_3^{\text{CE}}$ . Overall, the  $F(q)$ -based clustering, which captures the morphological features of neurons, has grouped the neurons in *C. elegans* nervous system into their respective functional type to a first approximation.

### Projection neurons in three neuropils of *Drosophila* olfactory system

The full *Drosophila* hemibrain connectome constructed from electron microscope images [13, 61] offers neuron morphologies at high resolution. The neurons constituting the second-order layer of the *Drosophila* olfactory system are made of three types of projection neurons, which include uniglomerular PNs (uPNs), multiglomerular PNs (mPNs), and local neurons (LNs). Among them, uPNs ( $N = 111$ ) are bundled together into on average 3 uPNs and comprise  $\sim 50$  distinct glomeruli, which receive signals from  $\sim 50$  different types of olfactory receptor neurons (ORNs) where a diverse array of chemical signals are encoded via the combinatorial coding [62–66]. The uPNs extend across the layer of second-order neurons, bridging between ORNs and higher olfactory centers (see Fig. 2B). The multiglomerular PNs (mPN) ( $N = 30$ ) extend over multiple glomeruli, and the local neurons (LNs) ( $N = 12$ ) are confined within the neuropil.

The PNs extend their dendritic and axonal branches and these branches are densely entangled to form three neuropils. Along the neural signal transmission pathway, the antennal lobe (AL) is the first neuropil, consisting of the  $\sim 50$  glomeruli and the axonal extensions of the uPNs in AL project to the other two neuropils, mushroom body (MB) calyx and lateral horn (LH), which are the sites that synapse with Kenyon cells and lateral horn neurons for learned and innate responses to olfactory signals, respectively [67–71]. Since neuron projections in each neuropil feature unique morphology which is discernible from one neuropil to another, we segmented the uPNs into the

parts contributing to each of the three neuropils [72] and analyzed them separately.

In AL,  $F(q)$ -based clustering results in seven clusters (Fig. 6), characterized by  $2 \leq \mathcal{D} \leq 4$  at  $\pi/R_n < q < 2\pi/\bar{l}$ . The clusters  $C_1^{\text{AL}}$ ,  $C_2^{\text{AL}}$ ,  $C_3^{\text{AL}}$ , and  $C_4^{\text{AL}}$ , which represent the major population of the 127 PN projections to AL, are characterized by the structure of a branched polymer in melts ( $\mathcal{D} = 4$ ), semi-dilute polymer solution ( $\mathcal{D} = 2.53$ ), and dilute polymer solution under  $\Theta$ -solvent ( $\mathcal{D} = 2.29$ ) or good solvent ( $\mathcal{D} = 2$ ) condition.

In MB calyx,  $F(q)$ -based clustering identifies six clusters (Fig. 7). The minor cluster  $C_1^{\text{MB}}$  is comprised of 3 regulatory neurons. These neurons are either an MB output neuron (MBON) or an MB-C1 (inhibitory interneurons) [13], whose shapes greatly differ from typical PNs. Remarkably, the  $F(q)$ -based clustering correctly separates them out. In the clusters  $C_2^{\text{MB}}$ ,  $C_3^{\text{MB}}$ ,  $C_4^{\text{MB}}$ , and  $C_5^{\text{MB}}$ , the PNs exhibit boutons of the size around  $3 - 6 \mu\text{m}$  [73] at the tip of protrusions stemming from the main axon (Fig. 7). KCs form claw-like projections and envelope boutons at the synaptic site [74]. The structural characteristic of the two-layered hierarchy manifests itself in the  $F(q)$  plots with  $\mathcal{D} \approx 2.29$ , indicating that axonal branches adopt the configurations of branched polymer in the  $\Theta$ -solvent at two different length scales, below and above the ‘plateau’ at  $q \approx 0.4 \mu\text{m}^{-1}$  ( $r \approx 16 \mu\text{m}$ ) (see the arrows pointing the plateau of  $C_2^{\text{MB}}$  and  $C_3^{\text{MB}}$  in Fig. 7). This trend is practically absent in clusters  $C_5^{\text{MB}}$  and  $C_6^{\text{MB}}$  which have only a small number of boutons in each PN.

In LH, there are six clusters (Fig. 8). Similar to  $C_2^{\text{MB}}$  and  $C_3^{\text{MB}}$  in MB calyx, the clusters  $C_1^{\text{LH}}$  and  $C_2^{\text{LH}}$  in LH are characterized by the  $F(q)$  displaying a ‘plateau’. Despite the overall similarity in the shape of  $F(q)$  in that a plateau is present at intermediate  $q$ , the details of  $F(q)$  and the actual PN morphologies in LH are slightly different from those in MB calyx (Fig. 8). The plateaux of  $C_1^{\text{LH}}$  and  $C_2^{\text{LH}}$  are identified at  $q \approx 0.5 - 0.6 \mu\text{m}^{-1}$  ( $r \approx 11 \mu\text{m}$ ). Furthermore, visual inspection of the morphologies indicates that smaller axonal branches are uniformly distributed over the main branches of PNs in LH. From  $C_1^{\text{LH}}$  to  $C_3^{\text{LH}}$ , the PNs change their morphologies from that of the melt of branched polymer ( $\mathcal{D} = 4$ ) to that of branched polymer in  $\Theta$ -solvent ( $\mathcal{D} = 2.29$ ) and to that of branched polymer in a good solvent ( $\mathcal{D} = 2$ ). In the remaining clusters,  $C_4^{\text{LH}}$ ,  $C_5^{\text{LH}}$ , and  $C_6^{\text{LH}}$ , PNs are small in size and relatively featureless ( $\mathcal{D} = 1$ ). The difference between the clusters stems from the overall size ( $R_n$ ).

In the second-order neurons of the *Drosophila* olfactory system, PNs with different morphology are mixed to constitute synaptically dense MB calyx and LH.

## Neurons in mouse primary visual cortex

The Allen Cell Type database contains full and partial reconstructions of neurons in the mouse primary visual cortex (V1), which plays a critical role in visual processing as the first cortical region to receive visual input [75]. Visual signals processed in V1 are sorted [76] and relayed to at least 18 different brain areas [77–79] which are deemed functionally distinct [80]. The V1 is divided into 6 functionally distinct layers [81]. For example, the layers 2 and 3 house many direction-selective oriented cells [82]. Layer 5 is known for mostly non-oriented cells with a large receptive field [82]. Diverse types of neurons with varying sizes spanning across multiple layers (see Fig. 2A) comprising V1 are rich in pyramidal cells, whose organization maps the orientation preferences [83]. Gouwens *et al.* performed hierarchical clustering on the Allen Cell Type database using a combination of morphometric features such as branch number and the total contour length of the neurons [11]. All neurons in the Allen dataset are categorized under two different dendrite types, either spiny or aspiny based on the existence of dendritic spines, which are small protrusions along the dendrite where synaptic inputs occur. Although individual neurons are labeled with dendrite type information, these minuscule structures are not represented in the morphological reconstructions. Here, instead of relying on morphometry, we calculated the  $F(q)$ s of V1 neurons by leveraging the 3D coordinates provided in the database to assess the similarity between them and repeated Gouwens *et al.*'s clustering procedure.

In comparison with Gouwens *et al.* who classified V1 neurons into 19 clusters for spiny neurons and 19 clusters for aspiny neurons [11], our  $F(q)$ -based clustering produced 11 and 8 clusters for spiny and aspiny neurons, respectively (Fig. 9 and Fig. 10). The difference in the number of clusters likely stems from the fact that information on relative soma depth, for instance, utilized in morphometry-based clustering, is not explicitly included in the  $F(q)$ -based clustering.

From the  $F(q)$  plots, it becomes clear that neurons are classified based on the fractal dimensions and the neuron size into several distinct subgroups. For spiny neurons, clusters  $C_1^{\text{spiny}}$ ,  $C_3^{\text{spiny}}$ ,  $C_6^{\text{spiny}}$ , and  $C_8^{\text{spiny}}$  all share  $\mathcal{D} = 2.29$  for  $\pi/R_n < q < 2\pi/\bar{l}$  (or  $\bar{l} < r < R_n$ ) (Fig. 9), which corresponds to the geometric characteristics of a branched polymer in  $\Theta$  solvent. The clusters  $C_2^{\text{spiny}}$ ,  $C_5^{\text{spiny}}$ , and  $C_9^{\text{spiny}}$  are featured with  $\mathcal{D} \approx 2.53$ , indicative of the configurations of branched polymer in semi-dilute solution. Morphologically, these clusters exhibit relatively denser dendrites with higher branchness than those clusters with  $\mathcal{D} = 2.29$  (Fig. 9). The clusters  $C_4^{\text{spiny}}$ ,  $C_7^{\text{spiny}}$ , and  $C_{11}^{\text{spiny}}$

display the same fractal dimension,  $\mathcal{D} = 2$ , at  $\pi/R_n < q < 2\pi/\bar{l}$ , indicative of the conformation of a self-avoiding branched polymer in dilute solution; yet differ from each other by their neuron size ( $R_n(C_4^{\text{spiny}}) > R_n(C_7^{\text{spiny}} > R_n(C_{11}^{\text{spiny}}))$ ). The cluster  $C_{10}^{\text{spiny}}$  displays  $2.53 < \mathcal{D} < 4$ , featuring unusually dense arborization, similar to that of a branched polymer in a melt. For aspiny neurons (Fig. 10), we find that the clusters  $C_1^{\text{aspiny}}$ ,  $C_3^{\text{aspiny}}$ , and  $C_5^{\text{aspiny}}$  are characterized with  $\mathcal{D} = 4$ ;  $C_4^{\text{aspiny}}$ ,  $C_7^{\text{aspiny}}$ , and  $C_8^{\text{aspiny}}$  with  $\mathcal{D} = 2$ ;  $C_2^{\text{aspiny}}$  with  $\mathcal{D} = 2.29$ ; and  $C_6^{\text{aspiny}}$  with  $\mathcal{D} = 2.53$ .

We notice that the spiny neurons are on average characterized with  $\bar{\mathcal{D}}^{\text{spiny}} \approx 2.37$ , whereas the aspiny neurons with  $\bar{\mathcal{D}}^{\text{aspiny}} = 2.55$ . Thus, the neuron mass density ( $\rho \sim n/r^3 \sim r^{\mathcal{D}-3} \sim n^{1-3/\mathcal{D}}$ ) is higher in aspiny neurons ( $\bar{\mathcal{D}}^{\text{spiny}} < \bar{\mathcal{D}}^{\text{aspiny}}$ ). It is noteworthy that the absence (or presence) of apical dendrites, whose size is comparable to the size of a neuron ( $R_n$ ), does not provide a significant contribution to the fractal dimension in the intermediate scale  $\pi/R_n < q < 2\pi/\bar{l}$ .

Compared with the morphology types defined by Gouwens *et al.*,  $F(q)$ -based clustering prioritizes the overall structural similarity dictated by  $\mathcal{D}$  and  $R_n$  (and more precisely by  $d_{ij}$  of Eq. 11). However, the two clustering outputs are highly associated, as the Pearson's  $\chi^2$ -test between the two clustering outputs results in very small p-values for both the spiny and the aspiny neurons ( $p = 9.368 \times 10^{-33}$  and  $p = 8.736 \times 10^{-21}$  respectively). The  $F(q)$ -based clustering correctly captures the morphological characteristics of the neurons and can be utilized as an alternative criterion for clustering.

## DISCUSSION

The  $F(q)$ -based clustering produces a comprehensible and systematic classification of neurons, complemented by the size ( $R_n$ ) and the fractal dimension ( $\mathcal{D}$ ) that quantifies the extent of the branching pattern of reconstructed neurons. We argue that different groups of neurons classified based on their morphology has functional significance. In particular, the branchness of neurons, quantified by  $\mathcal{D}$ , is of great significance to the firing activities of neurons, which should be determined by the densities of ion channels on the membrane surface. For a neuron characterized with  $\mathcal{D}$ , it is expected that the conductance ( $G$ ) per neuron is proportional to the total number of ion channels ( $n_{\text{ch}}$ ) or the capacity of the neuron dictated by the total contour length ( $L \sim R_n^{\mathcal{D}}$ ). If one imagines that a neuron is locally a cylindrical object with a cross-sectional diameter  $d(x)$  at a given position  $x$  along its contour [28] and that the density ( $\sigma_{\text{ch}}$ ) of ion channels on the membrane

surface is given, the conductance of ions across the membrane surface of a neuron is estimated as

$$G \sim n_{\text{ch}} \sim \sigma_{\text{ch}} \pi \int_0^L d(x) dx \sim \sigma_{\text{ch}} \pi R_n^{\mathcal{D}} \bar{d}. \quad (10)$$

where  $\bar{d} = \int_0^L d(x) dx / L$  is the average cross-sectional diameter of a neuron. A densely branched neuron characterized with a greater fractal dimension, say  $\mathcal{D} = 4$ , is expected to display a higher conductance than a branchless neuron with  $\mathcal{D} = 1$ . Functional differentiation of neurons, dictated by various firing patterns, has indeed been associated with a morphologically heterogeneous neuronal population [3].

The clustering based on topological characteristics (e.g. morphometrics) [11, 15, 17–19] is conceptually akin to our  $F(q)$ -based clustering. Although the morphometrics-based classification could be more comprehensive with a set of geometrical features selected, it requires a more extensive pre-processing as well as careful selection of feature sets to avoid introducing a bias or redundancy. Accompanied by electrophysiological measurement [10, 11] and stimuli response profiles [84], our  $F(q)$ -based classification, which employs the entire 3D image (coordinates) of neurons with no prior knowledge, holds good promise for a systematic study to elucidate the structure–function relationship in neuron, namely, the relationship between the neuron morphologies and the neural activities.

## METHODS

### Data preparation

- A total of 302 morphological reconstructions of neurons in the *C. elegans* nervous system were collected from the OpenWorm dataset [31]. The neuron type information available from the dataset was used to assign each neuron into one of the five functional categories (sensory, motor, interneuron, polymodal neuron, and others).
- For the *Drosophila* olfactory projection neurons, we use the TEMCA2 dataset [13]. A total of 162 *Drosophila* olfactory neurons are present in the dataset, which are extracted from the right hemisphere of the female *Drosophila*. In the dataset, three neurons (ID=203840, 2738003, 2738059) that are identified to be not PNs and six neurons (ID=1549518, 2738042, 2738083, 2738261, 2970058, 2970073) that did not project to any of the neuropils are dropped.

We removed artifacts in the data in which certain data points are dramatically offset from the expected trajectories due to misalignments of images when constructing the original morphological skeleton. These data points were identified by a geometrical criterion that the distances between three consecutive points should be smaller than  $10 \mu\text{m}$ . Next, we systematically segmented the neurons inside the three neuropils using the same process detailed in [72].

- The Allen Cell Type Database [11] contains 509 full and partial morphological reconstructions of neurons in the mouse primary visual cortex. Following Gouwens *et al.* [11], we classify the whole neurons into 234 spiny and 275 aspiny neurons. When comparing the clustering results, we noticed that Gouwens *et al.* used 461 morphological reconstructions which is smaller than what is currently available in the database. Gouwens *et al.* provided the exact cell IDs for 461 morphological reconstructions, but a subset of cell IDs do not match with those currently present in the database. Therefore, when performing  $\chi^2$ -independence tests, we only use neurons present in both the current Allen Cell Type Database and the list of neurons shared by Gouwens *et al.* A total of 430 neurons match this criteria (230 spiny and 200 aspiny).

In all datasets, the morphological information of each neuron is stored as a set of 3D coordinates with the connectivity specified in the parent samples. Complete reconstruction of neuron morphology is made by connecting data points based on their parent-child relationship. Reconstructed neuronal morphology is analyzed to identify branching points, where more than one child points identify as a parent. Tips of the neuron were defined if no other child points were present. In this study, a branch is defined as a set of points between two branching points or between a branching point and a tip.

### **$F(q)$ -based clustering of neurons**

After calculating the  $F(q)$  for each neuron, we cluster the set of neurons using a hierarchical clustering algorithm. We define the ‘distance’ between the  $F(q)$ s of  $i$ -th and  $j$ -th neurons by using the Euclidian distance (L2 norm),

$$d_{ij} = \sqrt{\sum_{n=1}^N |F_i(q_n) - F_j(q_n)|^2}. \quad (11)$$

Since our interest largely lies in the intermediate  $q$  regime where the features of branched polymers are observed,  $d_{ij}$ 's are calculated over the range of  $\pi/R_n < q < 2\pi/\bar{l}$ . Calculation of  $d_{ij}$ 's over all  $i$  and  $j$  pairs gives rise to a square distance matrix, to which the hierarchical clustering algorithm is applied. For the Allen Cell type dataset, we follow the same procedure as Gouwens *et al.*, who separated spiny and aspiny/sparsely spiny neurons before clustering [11]. Hierarchical clustering was carried out separately on spiny and aspiny neurons, and the clusters were collected based on the dynamic hybrid cut tree method [85]. The minimal cluster size of 4 was imposed for the mouse V1 in Gouwens *et al.* [11] and smaller clusters are automatically merged to a neighboring cluster. For the neurons in *C. elegans* and *Drosophila* olfactory system, we applied the same criterion of clustering as that of the mouse V1 neurons for consistency but did not impose the minimal cluster size restriction so that if necessary we can sample small clusters as well.

**Acknowledgments** We thank the Center for Advanced Computation in KIAS for providing computing resources.

**Funding** This study was supported by KIAS Individual Grants CG077002 (K.C.), CG076002 (W.K.K.), and CG035003 (C.H.).

**Code availability** Python scripts used for this manuscript are available at <https://github.com/kirichoi/FqClustering>

**Competing interests** The authors declare no competing interests.

---

- [1] E. R. Kandel, J. H. Schwartz, T. M. Jessell, S. A. Siegelbaum, and A. J. Hudspeth, eds., *Principles of Neural Science* (McGraw Hill, 2013), 5th ed.
- [2] S. R. y Cajal, *Histologie du système nerveux de l'homme & des vertébrés: Cervelet, cerveau moyen, rétine, couche optique, corps strié, écorce cérébrale générale & régionale, grand sympathique*, vol. 2 (A. Maloine, 1911).
- [3] Z. F. Mainen and T. J. Sejnowski, Nature **382**, 363 (1996).
- [4] A. Van Ooyen, J. Duijnhouwer, M. W. Remme, and J. van Pelt, Network: Computation in neural

systems **13**, 311 (2002).

- [5] Q. Wen, A. Stepanyants, G. N. Elston, A. Y. Grosberg, and D. B. Chklovskii, Proc. Natl. Acad. Sci. **106**, 12536 (2009).
- [6] R. A. van Elburg and A. van Ooyen, PLoS Comp. Biol. **6**, e1000781 (2010).
- [7] J. H. Smith, C. Rowland, B. Harland, S. Moslehi, R. Montgomery, K. Schobert, W. Watterson, J. Dalrymple-Alford, and R. Taylor, Scientific reports **11**, 1 (2021).
- [8] J. Woo, K. Choi, S. H. Kim, K. Han, and M. Choi, Front Biosci.-Landmark **27** (2022).
- [9] S. Latschus, D. Kobak, and P. Berens, Neuroinform. **18**, 591 (2020).
- [10] Y. Seki, J. Rybak, D. Wicher, S. Sachse, and B. S. Hansson, J. Neurophysiol. **104**, 1007 (2010).
- [11] N. W. Gouwens, S. A. Sorensen, J. Berg, C. Lee, T. Jarsky, J. Ting, S. M. Sunkin, D. Feng, C. A. Anastassiou, E. Barkan, et al., Nat. Neurosci. **22**, 1182 (2019).
- [12] J. M. Jeanne, M. Fişek, and R. I. Wilson, Neuron **98**, 1198 (2018).
- [13] Z. Zheng, J. S. Lauritzen, E. Perlman, C. G. Robinson, M. Nichols, D. Milkie, O. Torrens, J. Price, C. B. Fisher, N. Sharifi, et al., Cell **174**, 730 (2018).
- [14] G. S. Jefferis, C. J. Potter, A. M. Chan, E. C. Marin, T. Rohlfing, C. R. Maurer Jr, and L. Luo, Cell **128**, 1187 (2007).
- [15] H. B. Uylings and J. Van Pelt, Network: Computation in Neural Systems **13**, 397 (2002).
- [16] A. Tsiola, F. Hamzei-Sichani, Z. Peterlin, and R. Yuste, J. Comp. Neurol. **461**, 415 (2003).
- [17] R. Scorcioni, S. Polavaram, and G. A. Ascoli, Nature protocols **3**, 866 (2008).
- [18] Y. Lu, L. Carin, R. Coifman, W. Shain, and B. Roysam, Neuroinform. **13**, 47 (2015).
- [19] B. Mihaljević, P. Larrañaga, R. Benavides-Piccione, S. Hill, J. DeFelipe, and C. Bielza, BMC bioinformatics **19**, 1 (2018).
- [20] X. Vasques, L. Vanel, G. Villette, and L. Cif, Front. Neuroanatomy **10**, 102 (2016).
- [21] H. Zeng and J. R. Sanes, Nature Reviews Neuroscience **18**, 530 (2017).
- [22] A. P. Buccino, T. V. Ness, G. T. Einevoll, G. Cauwenberghs, and P. D. Häfliger, in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (IEEE, 2018), pp. 999–1002.
- [23] T. Zhang, Y. Zeng, Y. Zhang, X. Zhang, M. Shi, L. Tang, D. Zhang, and B. Xu, Sci. Rep. **11**, 1 (2021).
- [24] F. Caserta, W. Eldred, E. Fernandez, R. Hausman, L. Stanford, S. Bulderev, S. Schwarzer, and

H. Stanley, J. Neurosci. Methods **56**, 133 (1995).

[25] T. Smith Jr, G. Lange, and W. B. Marks, J. Neurosci. Methods **69**, 123 (1996).

[26] D. Ristanović, V. Nedeljkov, B. Stefanović, N. Milošević, M. Grgurević, and V. Štulić, Biological cybernetics **87**, 278 (2002).

[27] D. Ristanović, B. D. Stefanović, and N. Puškaš, Neurosci. Res. **84**, 64 (2014).

[28] M. Liao, X. Liang, and J. Howard, Proc. Natl. Acad. Sci. **118** (2021).

[29] M. Daoud, J. P. Cotton, B. Farnoux, G. Jannink, G. Sarma, H. Benoit, R. Duplessix, C. Picot, and P. G. de Gennes, Macromolecules **8**, 804 (1975).

[30] M. Daoud and G. Jannink, J. Phys. **37**, 973 (1976).

[31] G. P. Sarma, C. W. Lee, T. Portegys, V. Ghayoomie, T. Jacobs, B. Alicea, M. Cantarelli, M. Currie, R. C. Gerkin, S. Gingell, et al., Philosophical Transactions of the Royal Society B **373**, 20170382 (2018).

[32] M. Daoud and J. Joanny, J. Phys. **42**, 1359 (1981).

[33] P.-G. De Gennes, *Scaling concepts in polymer physics* (Cornell university press, 1979).

[34] A. Y. Grosberg, A. R. Khokhlov, H. E. Stanley, A. J. Mallinckrodt, and S. McKay, Computers in Physics **9**, 171 (1995).

[35] M. Doi, S. F. Edwards, and S. F. Edwards, *The theory of polymer dynamics*, vol. 73 (oxford university press, 1988).

[36] M. Rubinstein, R. H. Colby, et al., *Polymer physics*, vol. 23 (Oxford university press New York, 2003).

[37] L. Liu and C. Hyeon, Biophys. J. **110**, 2320 (2016).

[38] B. Duplantier, J. Chem. Phys. **86**, 4233 (1987).

[39] L. Liu, P. A. Pincus, and C. Hyeon, Nano Lett. **19**, 5667 (2019).

[40] B. H. Zimm and W. H. Stockmayer, J. Chem. Phys. **17**, 1301 (1949).

[41] D. Stauffer and A. Aharony, *Introduction to percolation theory* (Taylor & Francis, 2018).

[42] P. Debye, H. Anderson Jr, and H. Brumberger, J. Appl. Phys. **28**, 679 (1957).

[43] H. D. Bale and P. W. Schmidt, Phys. Rev. Lett. **53**, 596 (1984).

[44] P.-z. Wong and A. J. Bray, Phys. Rev. Lett. **60**, 1344 (1988).

[45] J. G. White, E. Southgate, J. N. Thomson, S. Brenner, et al., Philos Trans R Soc Lond B Biol Sci **314**, 1 (1986).

[46] N. Chatterjee and S. Sinha, Prog. Brain Res. **168**, 145 (2007).

- [47] J. White, Trends in Neurosciences **8**, 277 (1985).
- [48] L. R. Varshney, B. L. Chen, E. Paniagua, D. H. Hall, and D. B. Chklovskii, PLoS Comp. Biol. **7**, e1001066 (2011).
- [49] T. A. Jarrell, Y. Wang, A. E. Bloniarz, C. A. Brittin, M. Xu, J. N. Thomson, D. G. Albertson, D. H. Hall, and S. W. Emmons, Science **337**, 437 (2012).
- [50] Y. Salzberg, C. A. Díaz-Balzac, N. J. Ramirez-Suarez, M. Atreed, E. Tecle, M. Desbois, Z. Kaprielian, and H. E. Bülow, Cell **155**, 308 (2013).
- [51] M. W. Moyle, K. M. Barnes, M. Kuchroo, A. Gonopolskiy, L. H. Duncan, T. Sengupta, L. Shao, M. Guo, A. Santella, R. Christensen, et al., Nature **591**, 99 (2021).
- [52] C. A. Díaz-Balzac, M. Rahman, M. I. Lázaro-Peña, L. A. M. Hernandez, Y. Salzberg, C. Aguirre-Chen, Z. Kaprielian, and H. E. Bülow, Curr. Biol. **26**, 2379 (2016).
- [53] E. L. Tsalik and O. Hobert, J. Neurobiol **56**, 178 (2003).
- [54] J. Y. Cho and P. W. Sternberg, Cell **156**, 249 (2014).
- [55] T. Wakabayashi, I. Kitagawa, and R. Shingai, Neuroscience research **50**, 103 (2004).
- [56] S. M. Maguire, C. M. Clark, J. Nunnari, J. K. Pirri, and M. J. Alkema, Current Biology **21**, 1326 (2011).
- [57] D. G. Andrews, Biology Methods and Protocols **4**, bpz020 (2019).
- [58] J. C. Way and M. Chalfie, Genes & development **3**, 1823 (1989).
- [59] M. Chatzigeorgiou, S. Yoo, J. D. Watson, W.-H. Lee, W. C. Spencer, K. S. Kindt, S. W. Hwang, D. M. Miller III, M. Treinin, M. Driscoll, et al., Nature neuroscience **13**, 861 (2010).
- [60] A. Mohammadi, J. Byrne Rodgers, I. Kotera, and W. S. Ryu, BMC neuroscience **14**, 1 (2013).
- [61] L. K. Scheffer, C. S. Xu, M. Januszewski, Z. Lu, S.-y. Takemura, K. J. Hayworth, G. B. Huang, K. Shinomiya, J. Maitlin-Shepard, S. Berg, et al., Elife **9**, e57443 (2020).
- [62] L. Buck and R. Axel, Cell **65**, 175 (1991).
- [63] B. Malnic, J. Hirono, T. Sato, and L. B. Buck, Cell **96**, 713 (1999).
- [64] Q. Gao, B. Yuan, and A. Chess, Nat. Neurosci. **3**, 780 (2000).
- [65] A. Couto, M. Alenius, and B. J. Dickson, Curr. Biol. **15**, 1535 (2005).
- [66] J. H. Bak, S. J. Jang, and C. Hyeon, PLoS Comp. Biol. **14**, e1006175 (2018).
- [67] G. Heimbeck, V. Bugnon, N. Gendre, A. Keller, and R. F. Stocker, Proc. Natl. Acad. Sci. **98**, 15336 (2001).

- [68] M. Heisenberg, *Nat. Rev. Neurosci.* **4**, 266 (2003).
- [69] Y. Aso, D. Hattori, Y. Yu, R. M. Johnston, N. A. Iyer, T.-T. Ngo, H. Dionne, L. Abbott, R. Axel, H. Tanimoto, et al., *Elife* **3**, e04577 (2014).
- [70] D. Vasmer, A. Pooryasin, T. Riemensperger, and A. Fiala, *Front. Behavioral Neurosci.* **8**, 174 (2014).
- [71] J. N. Schultzhaus, S. Saleem, H. Iftikhar, and G. E. Carney, *J. Insect Physiol.* **98**, 29 (2017).
- [72] K. Choi, W. K. Kim, and C. Hyeon, *bioRxiv* (2022).
- [73] N. J. Butcher, A. B. Friedrich, Z. Lu, H. Tanimoto, and I. A. Meinertzhagen, *J. Comp. Neurol.* **520**, 2185 (2012).
- [74] E. Gruntman and G. C. Turner, *Nat. Neurosci.* **16**, 1821 (2013).
- [75] D. J. Felleman and D. C. Van Essen, *Cerebral cortex* (New York, NY: 1991) **1**, 1 (1991).
- [76] L. L. Glickfeld, M. L. Andermann, V. Bonin, and R. C. Reid, *Nat. Neurosci.* **16**, 219 (2013).
- [77] S. W. Oh, J. A. Harris, L. Ng, B. Winslow, N. Cain, S. Mihalas, Q. Wang, C. Lau, L. Kuan, A. M. Henry, et al., *Nature* **508**, 207 (2014).
- [78] Y. Han, J. M. Kebschull, R. A. Campbell, D. Cowan, F. Imhof, A. M. Zador, and T. D. Mrsic-Flogel, *Nature* **556**, 51 (2018).
- [79] J. A. Harris, S. Mihalas, K. E. Hirokawa, J. D. Whitesell, H. Choi, A. Bernard, P. Bohn, S. Caldejon, L. Casal, A. Cho, et al., *Nature* **575**, 195 (2019).
- [80] Q. Wang, O. Sporns, and A. Burkhalter, *J. Neurosci.* **32**, 4386 (2012).
- [81] A. Resulaj, *Front. Neural Circuits* **15** (2021).
- [82] M. Hübener, *Curr. Opin. Neurobiol.* **13**, 413 (2003).
- [83] W. Ji, R. Gämänu, P. Bista, R. D. D'Souza, Q. Wang, and A. Burkhalter, *Neuron* **87**, 632 (2015).
- [84] S. E. de Vries, J. A. Lecoq, M. A. Buice, P. A. Groblewski, G. K. Ocker, M. Oliver, D. Feng, N. Cain, P. Ledochowitsch, D. Millman, et al., *Nat. Neurosci.* **23**, 138 (2020).
- [85] P. Langfelder, B. Zhang, and S. Horvath, *Bioinform.* **24**, 719 (2008).

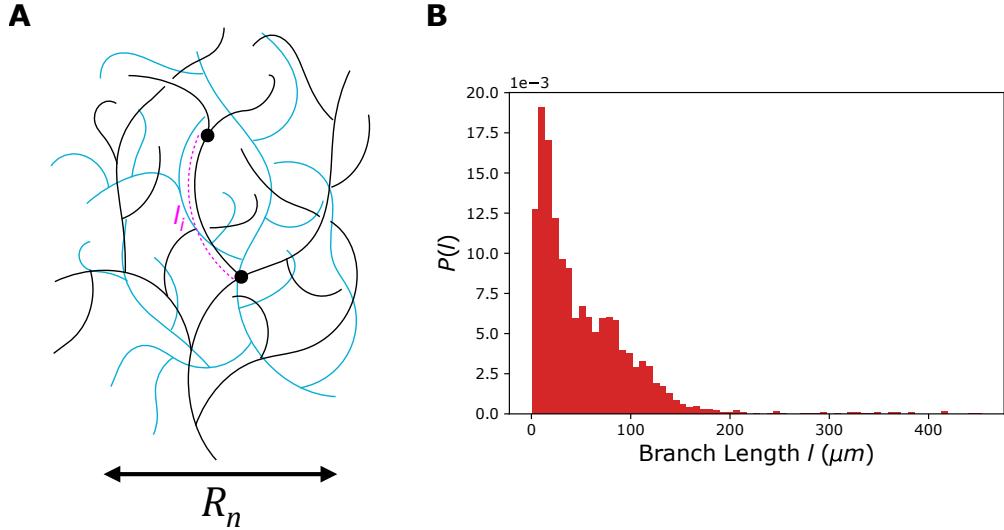


FIG. 1: (A) Two different range of length scales associated with typical neuronal architecture mapped on a structure of branched polymer: (i)  $\bar{b} < r < \bar{l}$  and (ii)  $\bar{l} < r < R_n$ . The black filled circles depict the branching points.  $l_i$  is the contour length of a branch, which is the segment between two branching points or between a branching point and a tip.  $\bar{b}$  is the average size of segment defined between two consecutive data points along the neuron branch,  $\bar{l}(= N_l^{-1} \sum_i^{N_l} l_i)$  is the average length of branches, and  $R_n$  is the size of neuron. (B) The distribution of the branch length ( $l_i$ ) of neurons in the cluster  $C_5^{\text{spiny}}$  from the mouse V1.

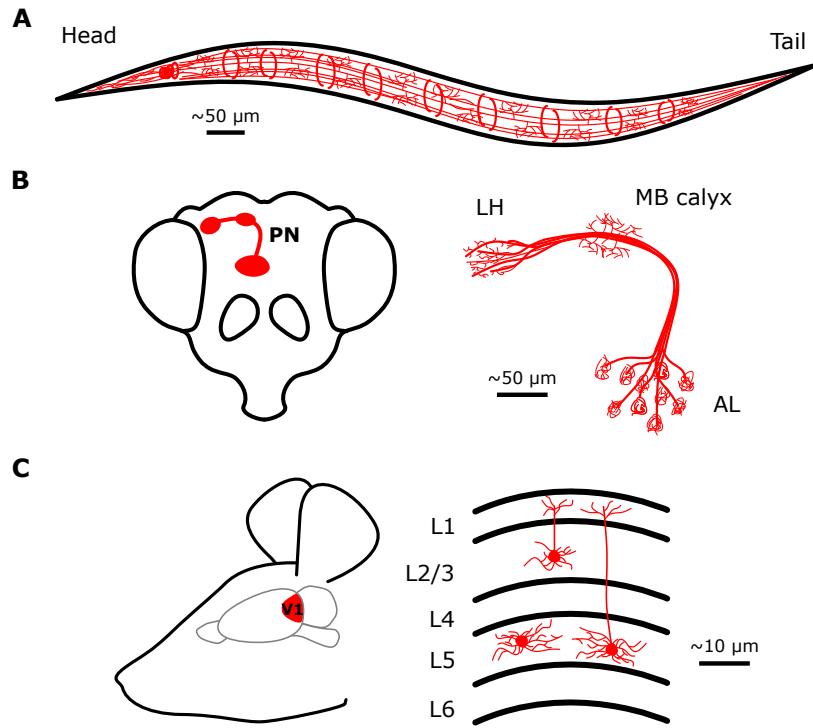


FIG. 2: Diagrams of three systems where  $F(q)$ -based analysis was made: (A) the *C. elegans* nervous system, (B) *Drosophila* olfactory projection neurons, and (C) Mouse primary visual cortex.

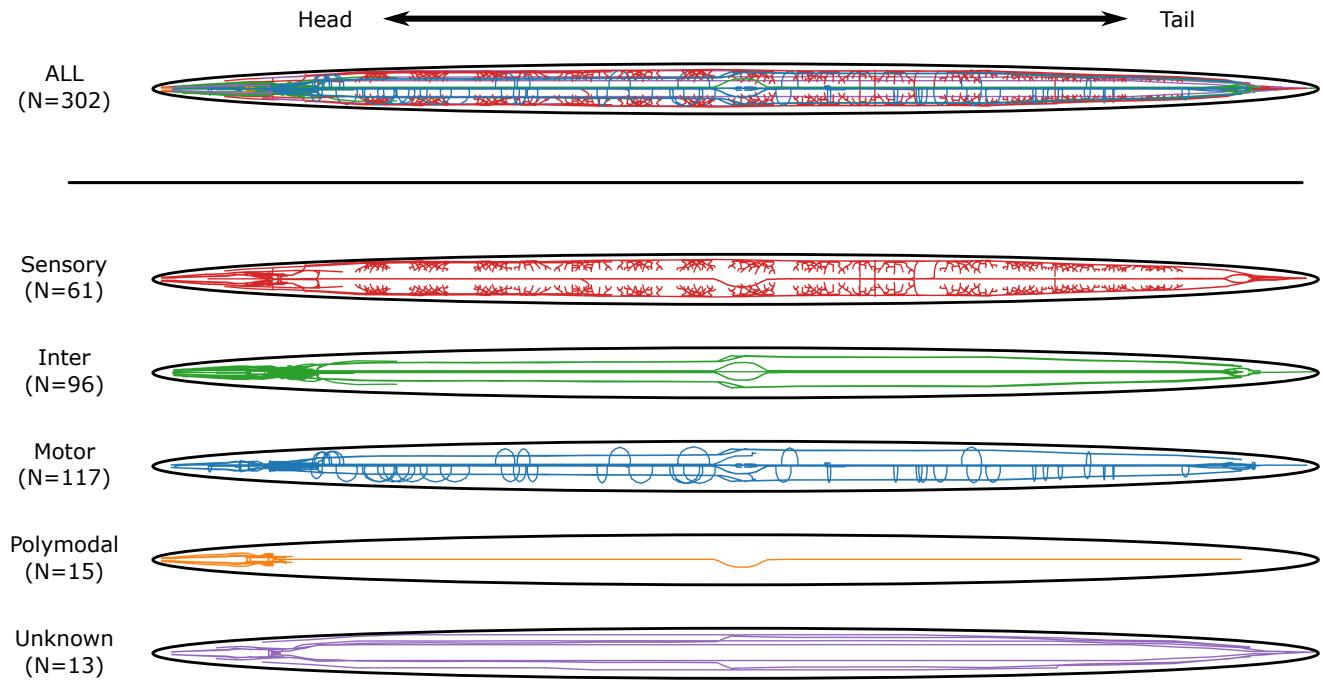


FIG. 3: Illustrations depicting the morphological reconstructions of neurons in *C. elegans*. Depicted at the top is the entire *C. elegans* nervous system. The nervous systems are redrawn for the neurons of different functional types.

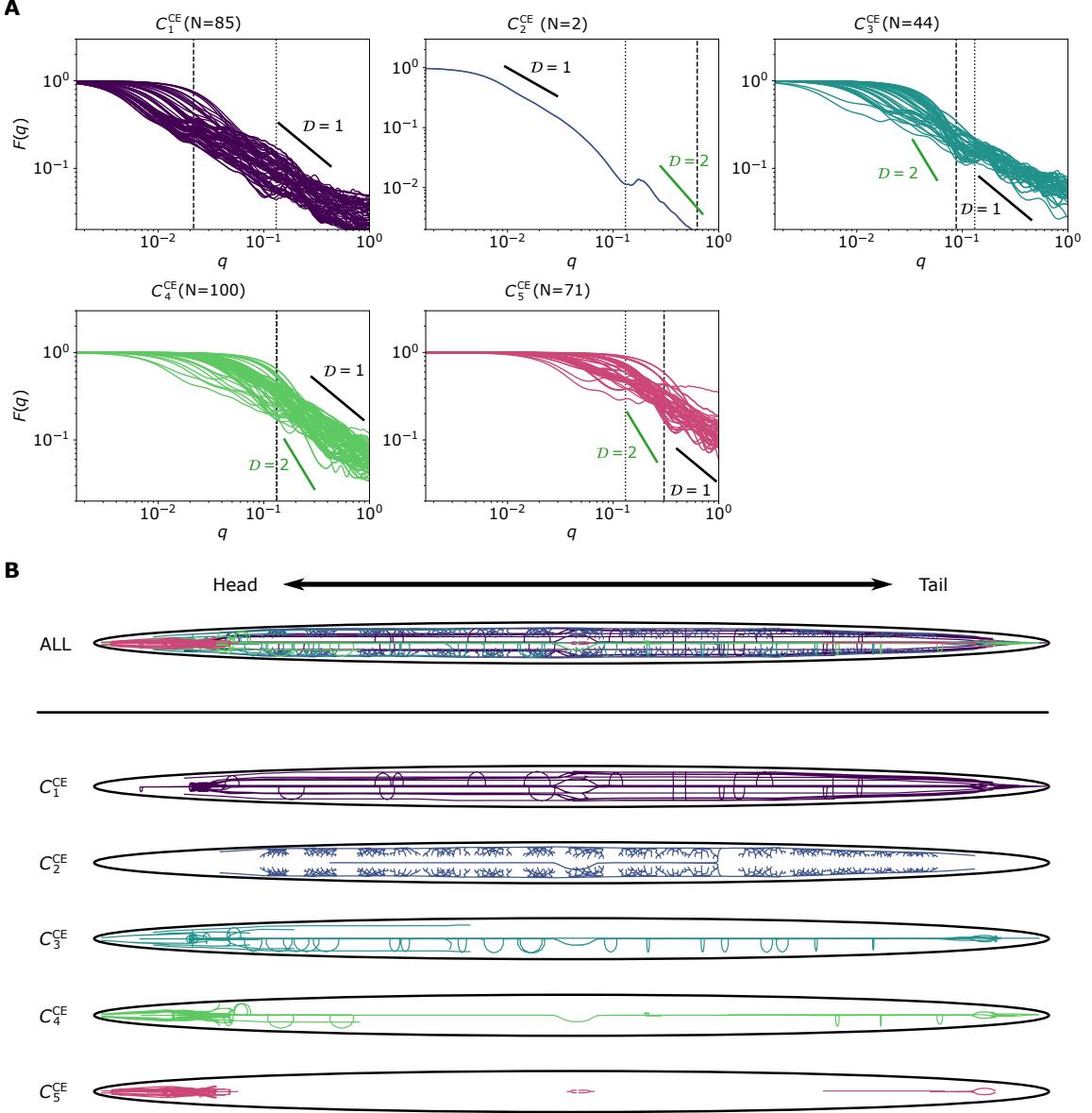


FIG. 4: (A)  $F(q)$  plots of neurons in the *C. elegans* nervous system grouped by the  $F(q)$ -based clustering result. The dotted lines denote  $q(= 2\pi/d_C)$  corresponding to the inverse scale of the average diameter ( $d_C$ ) of adult *C. elegans*. The dashed lines denote  $q(= 2\pi/\bar{l})$  correspond to the average branch size  $\bar{l}$  of neurons in the cluster. Reference fractal dimensions ( $D$ s) are highlighted in the  $F(q)$ -plots with the slopes. The clusters from  $C_1^{\text{CE}}$  to  $C_5^{\text{CE}}$  are indexed based on the average size ( $R_n$ ) of neurons in each cluster in a decreasing order. (B) Illustrations depicting the morphological reconstructions of neurons in *C. elegans* grouped by the  $F(q)$ -based clustering. The same colors are used as in (A) for each cluster.

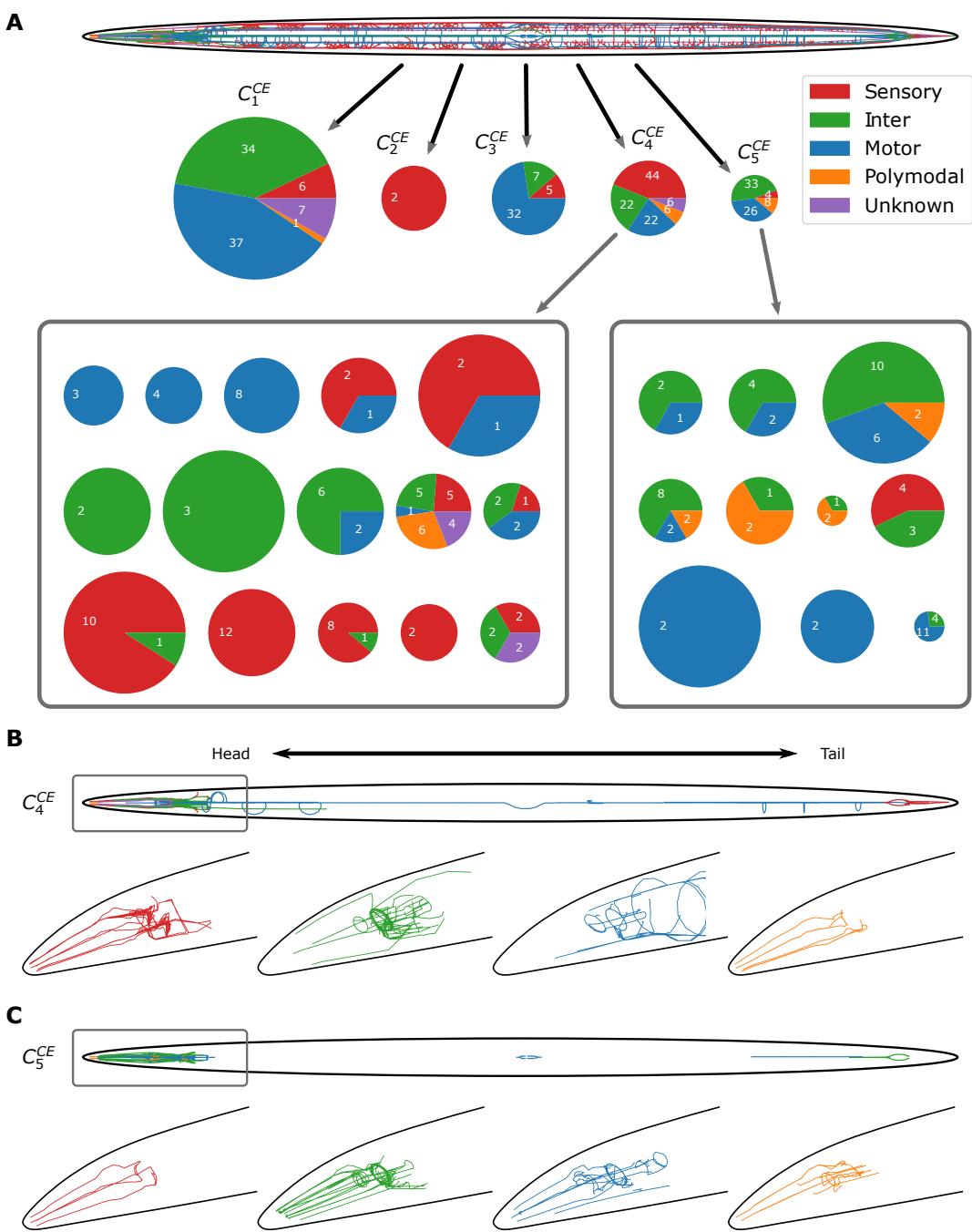


FIG. 5: (A) Pie charts depicting the functional composition of  $F(q)$ -based clusters in the *C. elegans* nervous system. The subclusters of  $C_4^{\text{CE}}$  and  $C_5^{\text{CE}}$  when the same clustering algorithm is applied repeatedly are shown inside the boxes. The size of the pie charts reflects the total contour length of neurons in each cluster. Illustrations depicting the morphological reconstructions of neurons in (B) the cluster  $C_4^{\text{CE}}$  and (C) the cluster  $C_5^{\text{CE}}$  with functional labeling. The enlarged 3D reconstruction of neurons in the head region, marked by gray boxes, is depicted below for each functional type.

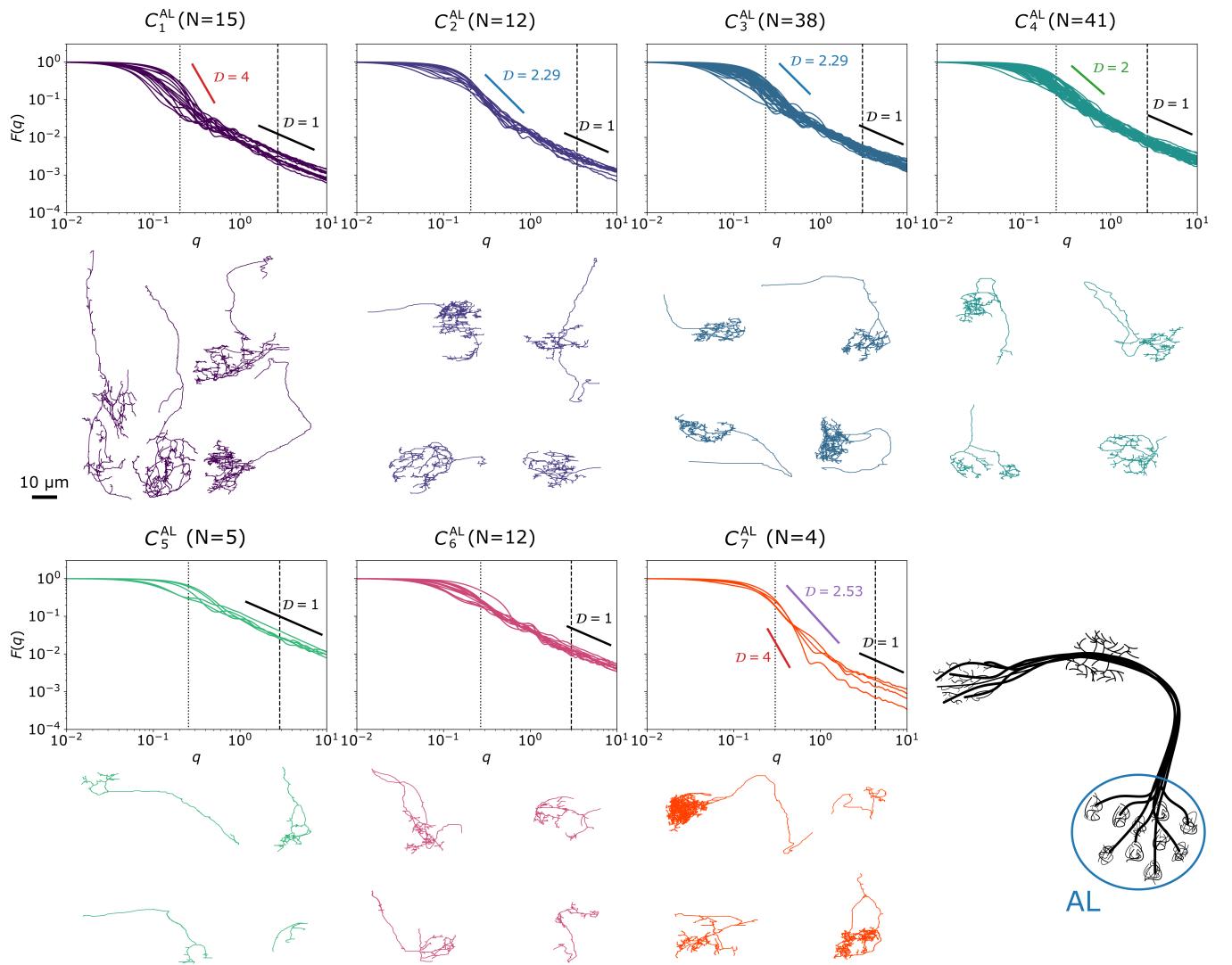


FIG. 6:  $F(q)$  plots of the PN projections to AL in the *Drosophila* olfactory system grouped by the  $F(q)$ -based clustering result. The dotted lines denote  $q$  corresponding to the inverse scale of neuron size ( $q = \pi/R_n$ ) of the cluster, and the dashed lines correspond to the average branch size  $\bar{l}$  of neurons in the cluster ( $q = 2\pi/\bar{l}$ ). Reference fractal dimensions ( $D$ s) are highlighted in the  $F(q)$ -plots with the slopes. The clusters from  $C_1^{\text{AL}}$  to  $C_7^{\text{AL}}$  are indexed based on the average size ( $R_n$ ) of neurons in each cluster in a decreasing order.

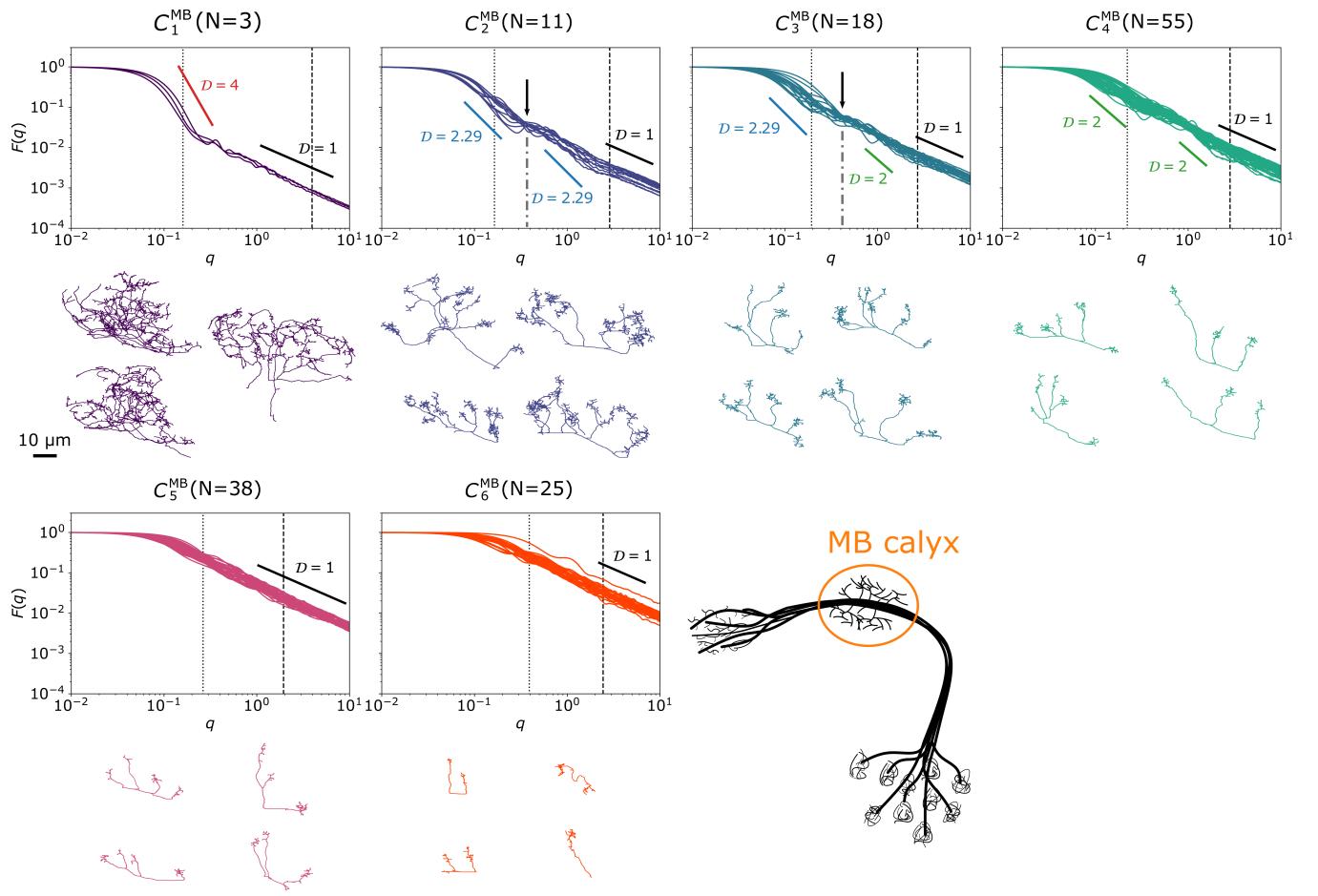


FIG. 7:  $F(q)$  plots of the PN projections to MB in the *Drosophila* olfactory system grouped by the  $F(q)$ -based clustering result. The locations of plateau in  $C_2^{\text{MB}}$  and  $C_3^{\text{MB}}$  at  $q \approx 0.4 \mu\text{m}^{-1}$  are indicated by the arrows. Other details are the same as Fig. 6.

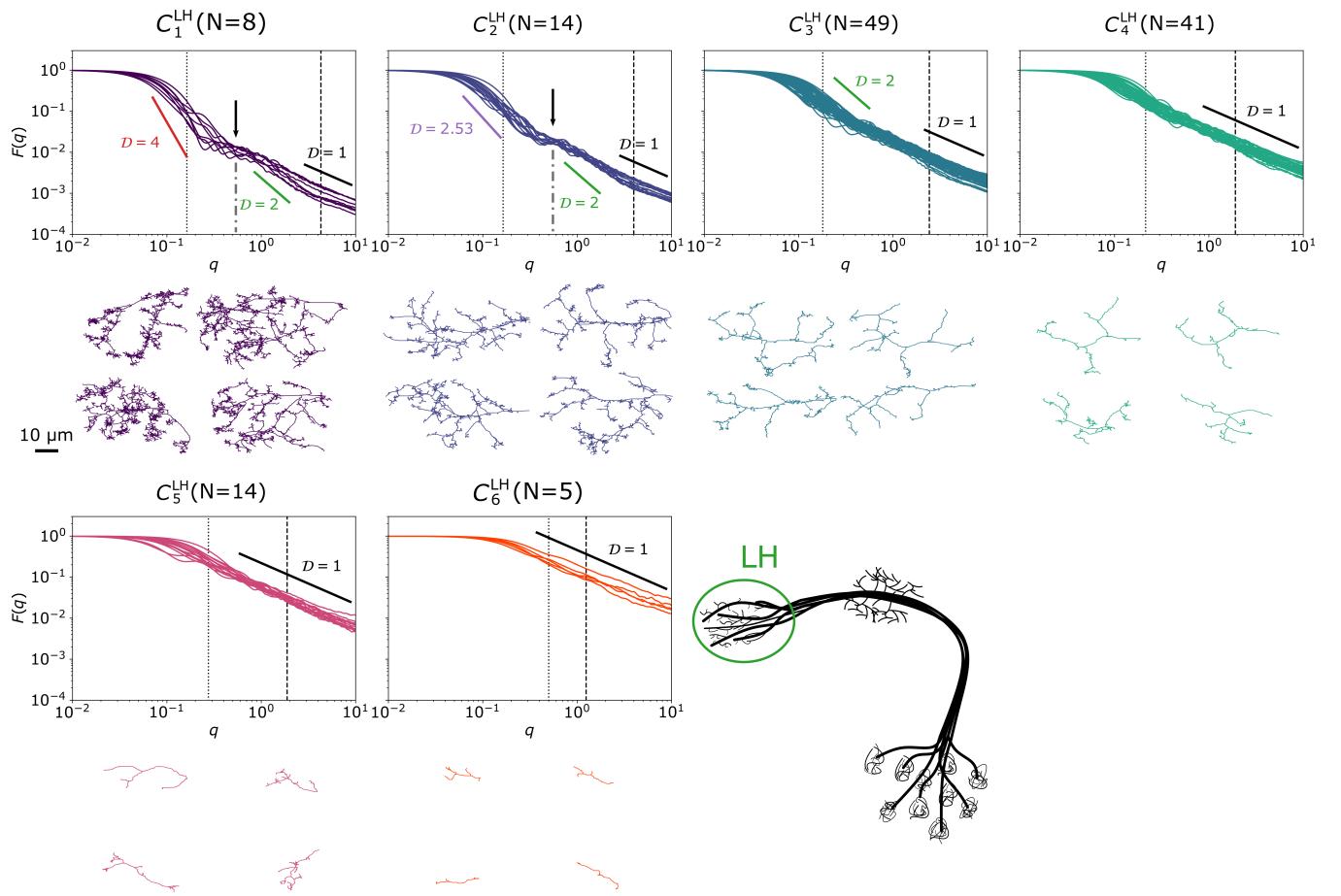


FIG. 8:  $F(q)$  plots of the PN projections to LH in the *Drosophila* olfactory system grouped by the  $F(q)$ -based clustering result. The locations of plateau in  $C_1^{\text{LH}}$  and  $C_2^{\text{LH}}$  at  $q \approx 0.5 - 0.6 \mu\text{m}^{-1}$  are indicated by the arrows. Other details are the same as Fig. 6.

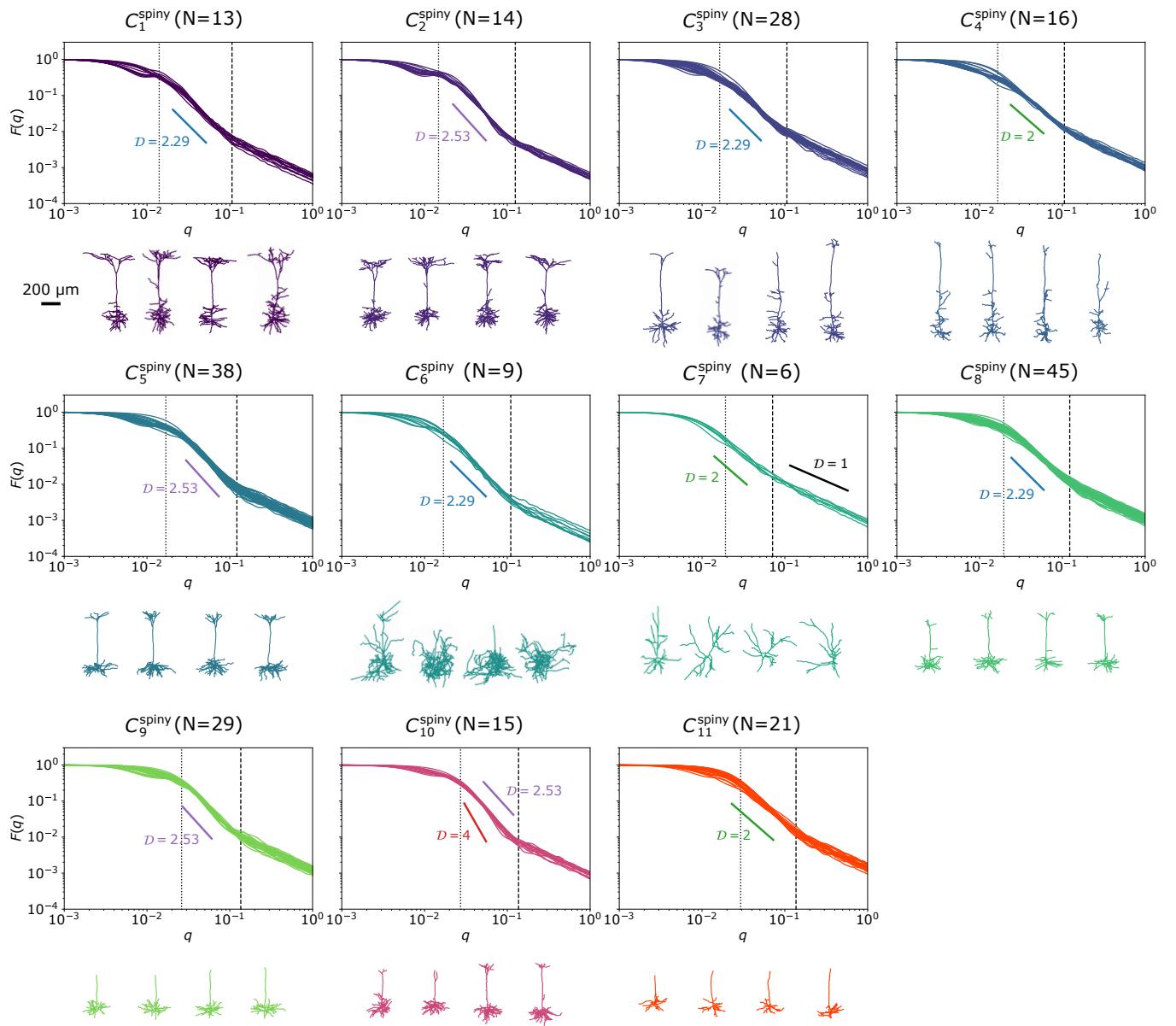


FIG. 9:  $F(q)$  plots of spiny neurons in the mouse primary visual cortex grouped by the  $F(q)$ -based clustering result. Other details are the same as Fig. 6.

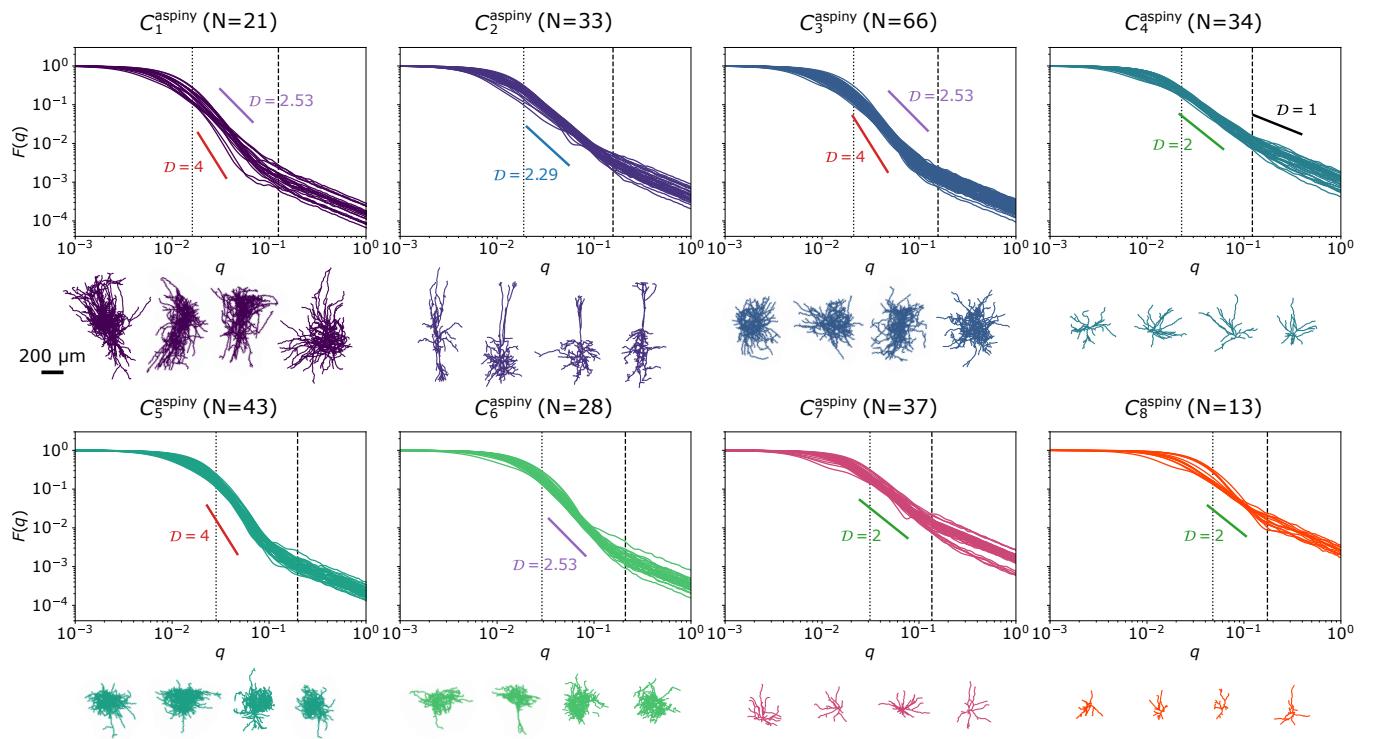


FIG. 10:  $F(q)$  plots of aspiny neurons in the mouse primary visual cortex grouped by the  $F(q)$ -based clustering result. Other details are the same as Fig. 6.