VNC-Dist: A machine learning-based tool for quantification of neuronal positioning in the ventral nerve cord of *C. elegans*

Saber Saharkhiz[1,3,4], Mearhyn Petite[3], Tony Roenspies[3], Theodore Perkins[2,4*], Antonio Colavita[1,3*]

1. Department of Cellular and Molecular Medicine, Faculty of Medicine, University of Ottawa, Ottawa, ON, Canada
2. Department of Biochemistry, Microbiology and Immunology, Faculty of Medicine, University of Ottawa, Ottawa, ON, Canada
3. Neuroscience, Ottawa Hospital Research Institute, Ottawa, ON, Canada
4. Regenerative Medicine, Ottawa Hospital Research Institute, Ottawa, ON, Canada

***Corresponding authors**: Antonio Colavita, colavita@uOttawa.ca; Theodore Perkins, tperkins@ohri.ca

Abstract

The ventral nerve cord (VNC) of newly hatched *C. elegans* contains 22 motoneurons organized into three distinct classes: DD, DA, and DB, that show stereotypical positioning and arrangement along its length. The VNC represents a genetically tractable model to investigate mechanisms involved in neuron sorting and positioning. However, accurately and efficiently mapping and quantifying all motoneuron positions within large datasets is a major challenge. Here, we introduce VNC-Dist, a semi-automated software toolbox designed to overcome the limitations of subjective motoneuron positioning analysis in microscopy. VNC-Dist uses an annotator for neuron localization and an automated contour-based method for measuring the relative distances of neurons along the VNC based on deep learning and numerical analysis. To demonstrate the robustness and versatility of VNC-Dist, we applied it to multiple genetic mutants known to disrupt neuron positioning in the VNC. This toolbox will enable the acquisition and analysis of large datasets on neuronal positioning, thereby advancing investigations into the cellular and molecular mechanisms that control neuron positioning and arrangement in the VNC.

Introduction

Fluorescent reporters have become invaluable tools for investigating nervous system connectivity and function in *C. elegans*[1]. The adult hermaphrodite has 302 neurons, organized into sensory neurons, interneurons, and motoneurons, with most cell bodies clustered in the nerve ring, retrovesicular ganglion (RVG), preanal ganglion (PAG), and along the ventral nerve cord (VNC)[2], [3], [4]. The position, morphology, and connectivity of different sets of these neurons can be readily visualized using distinct colours of cell-specific fluorescent markers[5]. However, accurately and reproducibly identifying and quantifying cells in densely packed neuronal regions remain challenging and time-consuming, particularly when analyzing large sample sizes or conducting high-throughput genetic screens[6], [7], [8].

At hatching, the *C. elegans* VNC contains 22 motoneurons: 6 DDs, 7 DBs, and 9 DAs, organized in a stereotyped pattern and distribution to control and coordinate muscle contractions along the entire length of the animal[3], [9] (Fig. 1a,b). This arrangement is established during embryonic development through the directed movement and sorting of neuronal progenitors that form the VNC. Notably, this process shares underlying molecular and cellular mechanisms with neural tube

formation in vertebrates, suggesting that the VNC could serve as an anatomically simple and genetically accessible model to study the cell and molecular biology common to both[9], [10]. Achieving this requires distinct labeling of DD, DA, and DB motoneurons, along with a workflow that efficiently measures the position of each neuron and the overall pattern of all 22 motoneurons along the VNC.

With the advancement of deep learning pipelines in bioimage analysis, researchers can more efficiently analyze and interpret increasingly complex biological data[11], [12]. Additionally, progress in biology is correlated with how fast and efficiently we can conduct experiments and analyses. Therefore, establishing dedicated bioimage analysis support using quantitative reproducible methods within academic institutions is crucial for enhancing research quality and efficiency and significantly advancing scientific discovery in biology. Machine learning, particularly deep neural networks, enables predictive modeling through data-driven approaches that replicate human pattern recognition[13], making it ideal for automating the analysis of large and complex image datasets for more reproducible and efficient workflows[14], [15], [16]. In data-driven microscopy, computational methods optimize parameters like exposure time and illumination to enhance image quality, while ML algorithms handle tasks such as segmentation, classification, and quantification, facilitating high-throughput screening and rapid biological discoveries[17], [18]. Integrating prior knowledge with data is crucial for developing informed ML models that address real-world challenges by pre-learning relevant patterns[19]. In the particular case of image segmentation, training datasets typically include raw images paired with corresponding labels using a supervised learning approach. However, this method is time-consuming, as both dataset creation and neural network training can take several hours to days[20]. Therefore, zero-shot and few-shot learning—which enable models to perform tasks with little to no labeled data—provide alternative approaches[21], such as Meta AI's Segment Anything Model (SAM)[22], [23], [24].

Current generalist segmentation tools face significant limitations when it comes to segmenting worms from fluorescent micrographs[25], [26], [27], particularly in terms of performance and accuracy. Furthermore, current bioimage analysis tools for quantifying curved biological structures primarily rely on traced path analysis. This is exemplified by contour-based methods (e.g., Neurite Tracer[28], SNT[29]), which estimate curve lengths by summing Euclidean distances between consecutive pixels, providing only a discrete approximation of the arc length. Alternatively, pixel-based methods (e.g., skeletonization and NeuroQuantify[30]), which calculate lengths by summing distances between consecutive points along the skeleton, or graph-based approaches (e.g., NeuronJ[31]), where nodes represent pixels or regions and edges define potential paths, are also employed. However, these methods lack general applicability and are not easily adaptable to *C. elegans* bioimage analysis. Consequently, manual tools (e.g., the segmented line tool in Fiji/ImageJ; Fig. 3a) remain widely used despite being time-consuming and less precise.

Visually detecting changes in VNC neuron positions is particularly challenging when comparing genetic mutants and within each genotype across different animals. This is because, although neurons in the VNC of wild type (WT) first larval stage (L1) *C. elegans* display a consistent 3D spatial distribution along the anterior–posterior (AP) axis (Fig. 1b), these patterns become unpredictable in mutants[9] (Figs. 2a–e). Here, we present VNC-Dist, a high-throughput semi-automated pipeline, designed to overcome the challenges of distance measurement, specifically tailored for labs studying *C. elegans* to quantify neuron positioning along the VNC, where 'distance' is defined as the cumulative arc length along the VNC's traced path represented by its parametric curve (VNCL1; Fig. 1a)[32]. This pipeline enables biologists to quickly and reliably analyze neuron positions from 3D fluorescent images using batch analysis, allowing for more accurate measurements while significantly reducing the time and effort compared to conventional manual methods. We demonstrate the

pipeline's robustness and versatility by applying it to various genetic backgrounds, including *vang-1/Van Gogh*, *prkl-1/Prickle*, and *sax-3/Robo* mutants, that are known to cause VNC assembly defects, in both single and compound mutants[9]. This Python-based toolkit uses SAM for automated worm segmentation, OpenCV for shape extraction and contour detection, and a computational algorithm based on numerical analysis techniques such as adaptive Gaussian quadrature to measure distances[33].

1. Methods

1.1 *C. elegans*

Worms were cultured at 20°C. The N2 genetic background was used for all strains. We used the following alleles and transgenes organized by linkage group (LG): LGX: *vang-1(tm1422)* and *sax-3(zy5)*. LGII: *unc-4(zy123[unc-4::mNG])* and *zySi2[unc-30p::mcherry::H2B::unc-30]*. LGIII: *vab-7(zy137[vab-7::mNG::3xFlag])* and *vab-7(zy142[vab-7::mNG::T2A::mScarlet-I::H2B])*. LGIV: *prkl-1(ok3182)*.

1.2 Microscopy and image acquisition

All images of VNC neurons were acquired using a Zeiss Axio Imager M2 fluorescent widefield microscope equipped with a 40×/1.4 NA oil immersion objective. Images were saved as .czi files in a four-dimensional format (CZYX: 2 channels, 11–17 z-slices, 1216 × 1920 pixels). Channels and z-stacks were merged using maximum intensity projection (MIP) and presented as 8-bit images. Imaging resolutions were optimized for neuron diameters—mScarlet (0.258 µm), mNG (0.225 µm), mCherry (0.266 µm)—with an overall merged image resolution of approximately 0.266 µm and an isotropic voxel size of 0.146 µm. For sample preparation, newly hatched L1 stage worms were mounted on 2% agarose pad slides containing 0.1 µg/µL of the paralytic Levamisole in M9 buffer.

1.3 DD, DA, DB fluorescent reporter strain construction

zySi2 [unc-30p::mCherry::H2B::unc-30 3'UTR + unc-119(+)] was made using Mos1-mediated single-copy insertion (MosSCI)[34]. The MosSCI LGII ttTi5605 targeting and selection plasmid (pAC661) was constructed using Gibson Assembly to combine four PCR fragments: a ~2.5 kb *unc-30* promoter from N2 genomic DNA, a ~1.25 kb mCherry::H2B cassette from pOD2046 (Addgene #89367), a 410 bp *unc-30* 3'UTR from N2 genomic DNA, and a ~7.4 kb MOS targeting and selection cassette from pCFJ151. A sequence file for pAC661 is included in the supplemental materials.

CRISPR/Cas9 gene editing, utilizing the self-excising selection cassette (SEC) method, was used to knock-in mNG at the C-terminus of the endogenous *unc-4* and *vab-7* genes[35]. Guide sequences for C-terminal insertion were identified using the IDT guide design tool at www.idtdna.com/site/order/designtool/index/CRISPR_SEQUENCE. Sequences were cloned into the Cas9–sgRNA plasmid pDD162 (Addgene #47549) using the NEB Q5 Site-Directed Mutagenesis Kit. *unc-4* and *vab-7* homology arms were cloned into the mNG-SEC plasmid pDD268 (Addgene #132523). Primers are listed in Supplemental Table S1.

A co-CRISPR approach[36] was used to swap the 3xFlag sequence in *vab-7(zy137[vab-7::mNG::3xFlag])* with a *T2A::mScarlet-I::H2B* cassette to generate *vab-7(zy142[vab-*

7::mNG::T2A::mScarlet-I::H2B]). A homologous repair plasmid (pAC798) was made containing 5' and 3' homology arms PCR amplified from *zy132* and a *T2A::mScarlet-I::H2B* cassette ordered as an IDT gBlocks gene fragment. An sgRNA targeting a site within the 3xFlag sequence was generated using pDD162 as described[35]. Primers are listed in Supplemental Table S1 and a sequence file for pAC798 is included in the supplemental materials.

1.4 CPU and GPU computing

The VNC-Dist pipeline for analyzing neuron positions along the AP axis of the VNC requires first the modified deep learning model SAM to segment 2D merged and collapsed images of worms from noisy fluorescent micrographs. This model, with approximately 636.12 million parameters, runs on GPUs through the Compute Canada Graham HPC cluster to segment all images in batch. The subsequent analysis, which is CPU-based, can be performed on a standard local computer.

1.5 Python programming

All components of our toolVNC-Dist are implemented in Python, with core functions built upon widely recognized open-source libraries (i.e. SciPy[37], https://scipy.org/; NumPy[38], https://numpy.org/ for numerical analysis; scikit-image[39], https://scikit-image.org/; Pillow[40], https://python-pillow.org/; OpenCV[41], https://opencv.org/ for image processing; pandas, https://pandas.pydata.org/; Matplotlib, https://matplotlib.org/ for output visualization), whereas others were developed by us to perform specific tasks during the analysis. This toolbox can also be used as a GUI version on various microscopy platforms, particularly Napari[42] (https://napari.org/stable/), providing easy access for biologists to perform their analyses without the need for programming knowledge.

1.6 Fiji/ImageJ tools for scoring and annotating

Unlike the conventional Scorer method, which uses the Segmented Line tool in Fiji/ImageJ (Fig. 3A) to measure the neuron distances manually, we utilized Fiji/ImageJ Multi-point tool (Annotator) to extract the (x, y) coordinates of all 25 neurons, along with the nerve ring and the rectum, from 3D Z-stack images (Fig. 3B). These coordinates were saved as .csv files. By combining these data with segmented worm images, our workflow calculates the relative distance along the VNC from anterior to posterior, tracing the worm's ventral curve (VNCL1) (Fig. 1a).

1.7 SAM in VNC-Dist

We integrated the zero-shot Segment Anything Model (SAM), using the ViT-H model type developed by Meta AI, into VNC-Dist without additional training (Fig. 6a) . The SAM mask generator was configured with parameters: 2 points per side prompt to guide the attention towards worm, prediction IoU threshold of 0.98, stability score threshold of 0.98, crop layers of 1, downscale factor of 1 (no downsampling), and a minimum mask region area of 99,000 pixels to discard smaller regions. We processed batches of 2D merged and collapsed micrographs (.tif) as input to SAM. Initial masks generated by SAM were refined by binarizing, inverting to highlight the foreground, and applying smoothing filters (Gaussian blur, median filter, bilateral filter), followed by morphological operations (erosion with a 19×19 kernel and closing with a 21×21 kernel). We created RGBA masks by setting

the background to fully transparent, the segmented worm to opaque black, and adding a 25-pixel transparent border around the mask to enhance the representation of the segmented worm. The refined segmented worms were saved as .png files named to match their corresponding images. Given the computational intensity of this part of VNC-Dist , we executed it on GPUs to enhance efficiency.

1.8 Automated contour-based algorithm in VNC-Dist

The structure of a key part of VNC-Dist pipeline, automated contour-based algorithm, is shown in Fig. 5 in the form of pseudocode, ranging from input which includes the segmented worm images (.png) from SAM (Fig. 6a) and neuron coordinates (.csv) from Fiji/ImageJ Annotator (Fig. 3B) to procedures like image processing, spline fitting, neuron projection, and relative distance calculation, to output as annotated images to visualize the projected neurons along the VNCL1 spline (.png) and normalized relative distances from SABVL to rectum (.csv).

In the spline fitting process of the automated contour-based algorithm, the spline function S(t) smooths a given contour by fitting cubic splines to its x and y coordinates separately. The process begins by parameterizing the contour with evenly spaced t-values between 0 and 1, corresponding to the number of contour points, while preserving the original order provided by OpenCV Python package. Typically, OpenCV orders contour points in a counterclockwise direction, starting from a specific point such as the topmost pixel. Using these t-values, the function constructs two cubic splines for the x and y dimensions with a smoothing factor of 900, balancing smoothness and fidelity to the original contour. These splines are then evaluated at each t-value to generate the fitted x and y coordinates. Finally, the spline function S(t) provides a smooth and continuous representation of the original contour for further analysis or visualization.

1.9 Statistics

To test whether the WT samples were normally distributed, we performed three steps: box plot visualization, Q–Q plot analysis, and the Shapiro–Wilk test.

Independent two-sample t-tests were performed to compare relative distance between WT and mutant groups, using Levene's F test to evaluate variance equality. Based on the results, either Student's t-test (for equal variances) or Welch's t-test (for unequal variances) was applied. Cohen's d was calculated to determine effect sizes, with features showing effect sizes ≥ 0.7 and $p < 0.05$ considered significant.

2. Results

2.1 A two-fluorophore palette for distinct labeling of DD, DA, and DB neurons.

In newly hatched *C. elegans*, VNC motoneurons exhibit a stereotypical 3D spatial distribution along the AP axis (Fig. 2f–h). However, in genetic mutants such as *vang-1(tm1422)*, *prkl-1(ok3182)*, *sax-3(zy5)*, *vang-1(tm1422); prkl-1(ok3182)*, and *sax-3(zy5) vang-1(tm1422)* these patterns change unpredictably, resulting in inconsistent neuronal distributions within each genotype across different animals (Fig. 2a–e).

To distinctly label the DD, DA, and DB motoneuron classes, we used their class-specific expression of transcription factors at the L1 stage: *unc-30* in DD1–6, *unc-4* in DA1–9 and the SAB

neurons, and *vab-7* in DB1–7[43], [44], [45]. We therefore used a combination of red and green fluorophores to distinctly label the nuclei of each motonuron class. MosSCI was used to insert a single copy of an *unc-30* promoter-driven *mCherry::H2B* transgene into the *ttTi5605* site on chromosome II. CRISPR/Cas9 gene editing was used to endogenously tag the C-terminus of *unc-4* with mNeonGreen and the C-terminus of *vab-7* with a dual-color tag, *mNeonGreen::T2A::mScarlet-I::H2B*. The T2A self-cleaving peptide[46] allows for co-expression of the mNG and mScarlet markers. mScarlet-I and mCherry were fused to histone H2B to target them to nuclei. In addition to DAs, *unc-4* is also expressed in the three SAB neurons, SABVL, SABVR, and SABD, located in the RVG[45]. When combined, this reporter system enables visualization of DD neurons in red, SAB and DA neurons in green, and DB neurons in yellow.

2.2 Reference points for VNC termini

Next, we sought to establish consistent reference points to define the anterior and posterior ends of the VNC across different genotypes for better distance comparisons. In WT worms, the VNC extends approximately 151 µm from the nerve ring at the anterior—which encircles the central region of the pharyngeal isthmus—to near the rectum at the posterior, with motoneurons distributed along its length[47] (Fig. 1a). While the rectum provides a consistent posterior landmark, the nerve ring position is known to vary among mutant genotypes[48], making it unreliable as an anterior reference point. In addition, the nerve ring is marked in each strain, situated between the anterior and posterior pharyngeal ganglia cell body clusters when the red signal channel is enhanced (Fig. 4a).

To address this, we performed correlation network analysis on WT (n = 42), *sax-3* (n = 61), *vang-1* (n = 63), and *prkl-1* (n = 37) mutants, using neurons as nodes and significant relative distances ($p < 0.05$, $r > 0.7$) as edges (Fig. 4a). This analysis revealed that most neurons shift relative to SABVL, the most anterior neuron in the RVG, in mutants. However, when normalized to nerve ring, the strong correlation among VNC and RVG neurons indicated that nerve ring is an unsuitable reference point, as its position shifts in tandem with these neurons. This finding aligns with previous study on sax-3 mutants as well[48]. In contrast, SABVL maintained a relatively consistent position across all genotypes (Fig. 4c). Therefore, we established SABVL as a reliable anterior reference point for comparing motoneuron positions along AP axis of VNC across different genotypes in *C. elegans* at L1.

2.3 VNC neuronal positioning pipeline design

VNC-Dist, a machine learning-based tool for quantification of VNC neuron positioning, consists of two main parts. The first part employs GPU-accelerated batch processing to segment worms in micrographs using the state-of-the-art Segment Anything Model (SAM). The second part involves CPU-based computational analysis of relative distance along the VNC curve from anterior to posterior, utilizing reference points assigned at the VNC ends—SABVL at the anterior and rectum at the posterior. To validate the relative distance measurements, the toolkit outputs images that visualize the projection of neurons onto the ventral side of the worm (VNCL1), displaying the t-values of each projected neuron. This allows us to assess whether the ordering of t-values correctly reflects the anterior-to-posterior sequence of neurons, in alignment with the reference points (Figs. 5i, ii, and 7). Additionally, VNC-Dist offers batch processing capabilities.

Our workflow quickly segmented worm images in batch through SAM. Neuron positions were then determined using Fiji/ImageJ Annotator (Fig. 3B) by selecting nuclei centers in 3D z-slices,

extracting x and y coordinates saved as .csv files (Fig. 6b). Contours of the segmented worm were extracted using OpenCV python package. To enable efficient traversal for a given t-value, the contour list C was rotated by an offset calculated as offset= $\lfloor n/2.31 \rfloor$. This value was determined based on the average size of the entire worm body contour at the L1 stage and the VNCL1 contour size on the ventral side, where n is the total number of contour points. The starting index k was determined by identifying the contour point ci closest to nerve ring (denoted as da1), minimizing $\|ci - da1\|$. Applying the shift k' = (k−offset) mod n ensured smooth wrapping of the contour points without data shuffling, thereby preventing neurons from being reordered along the VNC L1 spline (Fig. 6b). Subsequently, a parametric cubic spline function, S(t), was interpolated through the entire worm boundary to model its contours (Fig. 6c). Neurons, along with nerve ring and rectum, were first projected onto the VNCL1 spline using the shortest Euclidean distance (Fig. 6d, e). The distances of these projected neurons relative to the anterior reference point, SABVL, were then calculated along the VNCL1 spline using adaptive Gaussian quadrature. This approach was chosen because the actual VNC curve inside the worm body runs parallel to the VNCL1 spline, a segment of the S(t) spline on the worm's ventral boundary (Fig. 1a). Notably, this integration process for arc length calculation begins at the nerve ring, preceding the SABVL reference point, and extends to rectum following the last DA neuron (DA9). These distances were then normalized as percentages of the total VNC length between these reference points (SABVL as 0% to rectum as 100%) to obtain the relative distance for each neuron (Fig. 6c', i, ii).

2.4 Performance testing

To evaluate the efficiency of the VNC-Dist pipeline, we applied it to WT, PCP, and *sax-3* mutant images, analyzing over 280 images and more than 7,000 neurons. Fig. 7 presents three representative images of different mutants (*vang-1(tm1422); prkl-1(ok3182)*, *sax-3(zy5) vang-1(tm1422)*, and *prkl-1(ok3182)*) alongside the segmented worms and the corresponding neuron projection output.

As expected, the neuronal relative distances in WT exhibited normal distributions (Fig. 8c), confirming their suitability for parametric statistical analysis (Supplementary Figs. S1–S3). In addition, the analysis of neuron type arrangements revealed a consistent neuronal pattern across all animals (Fig. 8f).

On the other hand, the correlation network analysis revealed that all neurons along VNC could be categorized into three distinct clusters when normalized between SABVL (0%) and rectum (100%). The VNC neurons formed one cluster with minimal variation in the central region of the network, while the remaining two clusters comprised RVG and PAG neurons (Fig. 8b). These findings confirm that VNC neurons are anatomically distinct along the VNC.

2.5 Neuron displacement in PCP and *sax-3* mutants

To assess the impact of *vang-1*, *prkl-1*, and *sax-3* mutations on neuronal positioning, we quantified the mean relative distance of neurons in the following strains: *sax-3(zy5)*, *sax-3(zy5) vang-1(tm1422)*, *prkl-1(ok3182)*, *vang-1(tm1422)*, and *vang-1(tm1422); prkl-1(ok3182)*, using our semi-automated pipeline. Analysis of neuron positions revealed significant disruptions in the sorting patterns of neurons along the AP axis of VNC compared to WT worms. These findings align with previous results from manual measurements, which focused solely on DD motoneurons relative to DD1 using the Fiji/ImageJ Scorer tool (Fig. 3a)[9].

Mutant strains exhibited significant anterior shifts in neuronal positions, with the double mutant *sax-3(zy5) vang-1(tm1422)* showing the most pronounced displacement (Fig. 9a). Furthermore, neuron type arrangement analysis of all 22 motoneurons (including 6 DDs, 7 DBs, and 9 DAs) across each genotype revealed neuron type switches along the AP axis of VNC (Fig. 9b). Also, some neuron types overlapped at a single position, indicating a potential intercalation issue along the z-axis (Fig. 9b). These disruptions in neuronal positioning were inconsistent within each mutant genotype, as demonstrated by correlation network analysis (Fig. 4c, SABVL panels). This suggests unpredictable disruptions in anterior neuronal positioning and underscoring the critical roles of these genes in maintaining proper neuronal positions.

3. Discussion

Here we introduced VNC-Dist, a semi-automated tool for quantifying neuronal positions within the VNC of *C. elegans* at L1, providing precise reproducible measurements crucial for understanding neuronal organization and genetic influence on positioning. This high-throughput tool integrates deep learning with numerical analysis method to robustly quantify neuron spacing along the AP axis of VNC.

Our tool analyzed fluorescent micrographs from various mutant strains, including *sax-3*, *prkl-1*, *vang-1*, and their double mutant combinations, revealing distinct patterns of neuronal displacement. These findings underscore the critical roles of these genes in establishing the spatial order of motoneurons during VNC assembly. In addition, VNC-Dist confirmed that each motoneuron in the VNC of newly hatched wild type animals assumes a predictable position, validating the stereotypical pattern of neuronal spacing. Using correlation network analysis, we found that all motoneurons can be clustered into three groups: anterior RVG neurons, central VNC neurons, and posterior PAG neurons (Fig. 4c). Remarkably, correlation network analysis indicated that the VNC, RVG and PVG are distinct components as they are anatomically. This is consistent with the decreased variability we see in VNC neuron position when neurons are positioned relative to the most anterior and posterior neurons within the VNC and not to neurons in the RVG or PAG.

Our objective was to develop VNC-Dist to replace the time-consuming, labor-intensive, and error-prone manual distance measurements of VNC motoneurons. Recognizing that advances in biology often depend on the efficiency of experiments and analyses, we envision further enhancement of VNC-Dist into a fully automated system. Using state-of-the-art 3D cell and nuclei segmentation algorithms[49] (i.e. Cellpose[50], [51], [52], StarDist[53], [54]), future iterations will aim to identify all 22 motoneurons, including DA, DD, and DB types, thereby eliminating the need for manual annotation in Fiji/ImageJ to extract positional information. Furthermore, the segmentation component of VNC-Dist does not require training dataset generation or model training; however, it can be fine-tuned using the mask decoder of the SAM algorithm with human-verified ground truth generated from SAM to facilitate the segmentation of even completely challenging coiled worms. Consequently, VNC-Dist offers a powerful platform that combines mathematical precision with machine learning to produce accurate results in a time-efficient manner for advancing our understanding of molecular and cellular mechanisms governing neuronal positioning. The toolkit's GUI and code will be open-sourced on GitHub after publication in a peer-reviewed journal.

Acknowledgements

Some strains were provided by the Caenorhabditis Genetics Center, which is funded by the National Institutes of Health Office of Research Infrastructure Programs (P40 OD010440). The authors acknowledge the Cell Biology and Image Acquisition Core (RRID: SCR_021845) funded by the University of Ottawa, Ottawa, Natural Sciences and Engineering Research Council of Canada, and the Canada Foundation for Innovation. This work was supported by grants from the Canadian Institutes of Health Research (FRN 156160) and the Natural Resources and Engineering Research Council of Canada (2018-06790) to A. Colavita.

References

[1] H. Hutter, 'Fluorescent protein methods: strategies and applications', in *Methods in cell biology*, vol. 107, Elsevier, 2012, pp. 67–92.

[2] J. G. White, E. Southgate, J. N. Thomson, and S. Brenner, 'The structure of the ventral nerve cord of Caenorhabditis elegans', *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, vol. 275, no. 938, pp. 327–348, 1976.

[3] J. J. Smith *et al.*, 'A molecular atlas of adult C. elegans motor neurons reveals ancient diversity delineated by conserved transcription factor codes', *Cell Rep*, vol. 43, no. 3, 2024.

[4] S. J. Cook *et al.*, 'Whole-animal connectomes of both Caenorhabditis elegans sexes', *Nature*, vol. 571, no. 7763, pp. 63–71, 2019.

[5] S. R. Taylor *et al.*, 'Molecular topography of an entire nervous system', *Cell*, vol. 184, no. 16, pp. 4329–4347, 2021.

[6] D. Y. Sprague *et al.*, 'Unifying community-wide whole-brain imaging datasets enables robust automated neuron identification and reveals determinants of neuron positioning in C. elegans', *bioRxiv*, 2024.

[7] F. Long, H. Peng, X. Liu, S. K. Kim, and E. Myers, 'A 3D digital atlas of C. elegans and its application to single-cell analyses', *Nat Methods*, vol. 6, no. 9, pp. 667–672, 2009.

[8] M. Hailstone *et al.*, 'CytoCensus, mapping cell identity and division in tissues and organs using machine learning', *Elife*, vol. 9, p. e51085, 2020.

[9] P. K. Shah *et al.*, 'PCP and SAX-3/Robo pathways cooperate to regulate convergent extension-based nerve cord assembly in C. elegans', *Dev Cell*, vol. 41, no. 2, pp. 195–203, 2017.

[10] T. J. R. Frith, J. Briscoe, and G. L. M. Boezio, 'From signalling to form: the coordination of neural tube patterning', 2023.

[11] M. Mahmud, M. S. Kaiser, T. M. McGinnity, and A. Hussain, 'Deep learning in mining biological data', *Cognit Comput*, vol. 13, no. 1, pp. 1–33, 2021.

[12] J. G. Greener, S. M. Kandathil, L. Moffat, and D. T. Jones, 'A guide to machine learning for biologists', 2022. doi: 10.1038/s41580-021-00407-0.

[13] A. Tavanaei, M. Ghodrati, S. R. Kheradpisheh, T. Masquelier, and A. Maida, 'Deep learning in spiking neural networks', *Neural networks*, vol. 111, pp. 47–63, 2019.

[14] E. Moen, D. Bannon, T. Kudo, W. Graf, M. Covert, and D. Van Valen, 'Deep learning for cellular image analysis', *Nat Methods*, vol. 16, no. 12, pp. 1233–1246, 2019.

[15] A. Hallou, H. G. Yevick, B. Dumitrascu, and V. Uhlmann, 'Deep learning for bioimage analysis in developmental biology', *Development*, vol. 148, no. 18, p. dev199616, 2021.

[16] L. von Chamier *et al.*, 'Democratising deep learning for microscopy with ZeroCostDL4Mic', *Nat Commun*, vol. 12, no. 1, p. 2276, 2021.

[17] L. Morgado, E. Gómez-de-Mariscal, H. S. Heil, and R. Henriques, 'The Rise of Data-Driven Microscopy powered by Machine Learning', *arXiv preprint arXiv:2401.05282*, 2024.

[18] M. Jan, A. Spangaro, M. Lenartowicz, and M. Mattiazzi Usaj, 'From pixels to insights: Machine learning and deep learning for bioimage analysis', *BioEssays*, vol. 46, no. 2, p. 2300114, 2024.

[19] H. Xu, Y. Chen, and D. Zhang, 'Worth of prior knowledge for enhancing deep learning', *Nexus*, 2024.

[20] A. Kalapos and B. Gyires-Tóth, 'Self-supervised pretraining for 2d medical image segmentation', in *European Conference on Computer Vision*, Springer, 2022, pp. 472–484.

[21] S. Kadam and V. Vaidya, 'Review and analysis of zero, one and few shot learning approaches', in *Intelligent Systems Design and Applications: 18th International Conference on Intelligent Systems Design and Applications (ISDA 2018) held in Vellore, India, December 6-8, 2018, Volume 1*, Springer, 2020, pp. 100–112.

[22] Nikhila Ravi, 'SAM 2: Segment Anything in Images and Videos', *ArXiv*, Aug. 2024.

[23] A. Kirillov *et al.*, 'Segment anything', in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023, pp. 4015–4026.

[24] J. Ma, Y. He, F. Li, L. Han, C. You, and B. Wang, 'Segment anything in medical images', *Nat Commun*, vol. 15, no. 1, p. 654, 2024.

[25] H. Blum, 'Biological shape and visual science (Part I)', *J Theor Biol*, vol. 38, no. 2, pp. 205–287, 1973.

[26] W. F. Marshall, 'Origins of cellular geometry', *BMC Biol*, vol. 9, pp. 1–9, 2011.

[27] A. C. Oates, N. Gorfinkiel, M. Gonzalez-Gaitan, and C.-P. Heisenberg, 'Quantitative approaches in developmental biology', *Nat Rev Genet*, vol. 10, no. 8, pp. 517–530, 2009.

[28] M. H. Longair, D. A. Baker, and J. D. Armstrong, 'Simple Neurite Tracer: open source software for reconstruction, visualization and analysis of neuronal processes', *Bioinformatics*, vol. 27, no. 17, pp. 2453–2454, 2011.

[29] C. Arshadi, U. Günther, M. Eddison, K. I. S. Harrington, and T. A. Ferreira, 'SNT: a unifying toolbox for quantification of neuronal anatomy', *Nat Methods*, vol. 18, no. 4, pp. 374–377, 2021.

[30] K. M. Dang *et al.*, 'NeuroQuantify–An image analysis software for detection and quantification of neuron cells and neurite lengths using deep learning', *J Neurosci Methods*, vol. 411, p. 110273, 2024.

[31] E. Meijering, M. Jacob, J. Sarria, P. Steiner, H. Hirling, and e M. Unser, ‘Design and validation of a tool for neurite tracing and analysis in fluorescence microscopy images’, *Cytometry A*, vol. 58, no. 2, pp. 167–176, 2004.

[32] S. Vincent and D. Forsey, ‘Fast and accurate parametric curve length computation’, *Journal of graphics tools*, vol. 6, no. 4, pp. 29–39, 2001.

[33] E. Süli and D. F. Mayers, *An introduction to numerical analysis*. Cambridge university press, 2003.

[34] C. Frøkjær-Jensen *et al.*, ‘Single-copy insertion of transgenes in Caenorhabditis elegans’, *Nat Genet*, vol. 40, no. 11, pp. 1375–1383, 2008.

[35] D. J. Dickinson, A. M. Pani, J. K. Heppert, C. D. Higgins, and B. Goldstein, ‘Streamlined genome engineering with a self-excising drug selection cassette’, *Genetics*, vol. 200, no. 4, pp. 1035–1049, 2015.

[36] J. A. Arribere, R. T. Bell, B. X. H. Fu, K. L. Artiles, P. S. Hartman, and A. Z. Fire, ‘Efficient marker-free recovery of custom genetic modifications with CRISPR/Cas9 in Caenorhabditis elegans’, *Genetics*, vol. 198, no. 3, pp. 837–846, 2014.

[37] P. Virtanen *et al.*, ‘SciPy 1.0: fundamental algorithms for scientific computing in Python’, *Nat Methods*, vol. 17, no. 3, pp. 261–272, 2020.

[38] C. R. Harris *et al.*, ‘Array programming with NumPy’, *Nature*, vol. 585, no. 7825, pp. 357–362, 2020.

[39] S. Van der Walt *et al.*, ‘scikit-image: image processing in Python’, *PeerJ*, vol. 2, p. e453, 2014.

[40] A. Clark, ‘Pillow (pil fork) documentation’, *readthedocs*, 2015.

[41] A. Mordvintsev and K. Abid, ‘Opencv-python tutorials documentation’, *Obtenido de https://media. readthedocs. org/pdf/opencv-python-tutroals/latest/opencv-python-tutroals. pdf*, 2014.

[42] G. J. Selzer, C. T. Rueden, M. C. Hiner, E. L. Evans III, K. I. S. Harrington, and K. W. Eliceiri, ‘napari-imagej: ImageJ ecosystem access from napari’, *Nat Methods*, vol. 20, no. 10, pp. 1443–1444, 2023.

[43] B. Esmaeili, J. M. Ross, C. Neades, D. M. Miller III, and J. Ahringer, ‘The C. elegans even-skipped homologue, vab-7, specifies DB motoneurone identity and axon trajectory’, 2002.

[44] Y. Jin, R. Hoskins, and H. R. Horvitz, ‘Control of type-D GABAergic neuron differentiation by C. elegans UNC-30 homeodomain protein’, *Nature*, vol. 372, no. 6508, pp. 780–783, 1994.

[45] D. M. Miller III and C. J. Niemeyer, ‘Expression of the unc-4 homeoprotein in Caenorhabditis elegans motor neurons specifies presynaptic input’, *Development*, vol. 121, no. 9, pp. 2877–2886, 1995.

[46] Z. Liu *et al.*, ‘Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector’, *Sci Rep*, vol. 7, no. 1, p. 2193, 2017.

[47] R. W. Ware, D. Clark, K. Crossland, and R. L. Russell, ‘The nerve ring of the nematode Caenorhabditis elegans: sensory input and motor output’, *Journal of Comparative Neurology*, vol. 162, no. 1, pp. 71–110, 1975.

[48] J. R. Kennerdell, R. D. Fetter, and C. I. Bargmann, ‘Wnt-Ror signaling to SIA and SIB neurons directs anterior axon guidance and nerve ring placement in C. elegans’, *Development*, vol. 136, no. 22, pp. 3801–3810, 2009.

[49] J. Ma *et al.*, ‘The multimodality cell segmentation challenge: toward universal solutions’, *Nat Methods*, pp. 1–11, 2024.

[50] C. Stringer, T. Wang, M. Michaelos, and M. Pachitariu, ‘Cellpose: a generalist algorithm for cellular segmentation’, *Nat Methods*, vol. 18, no. 1, pp. 100–106, 2021.

[51] M. Pachitariu and C. Stringer, ‘Cellpose 2.0: how to train your own model’, *Nat Methods*, vol. 19, no. 12, pp. 1634–1641, 2022.

[52] C. Stringer and M. Pachitariu, ‘Cellpose3: one-click image restoration for improved cellular segmentation’, *bioRxiv*, pp. 2022–2024, 2024.

[53] N. Dey *et al.*, ‘AnyStar: Domain randomized universal star-convex 3D instance segmentation’, in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2024, pp. 7593–7603.

[54] M. Weigert, U. Schmidt, R. Haase, K. Sugawara, and G. Myers, ‘Star-convex polyhedra for 3D object detection and segmentation in microscopy’, in *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, 2020, pp. 3666–3673.

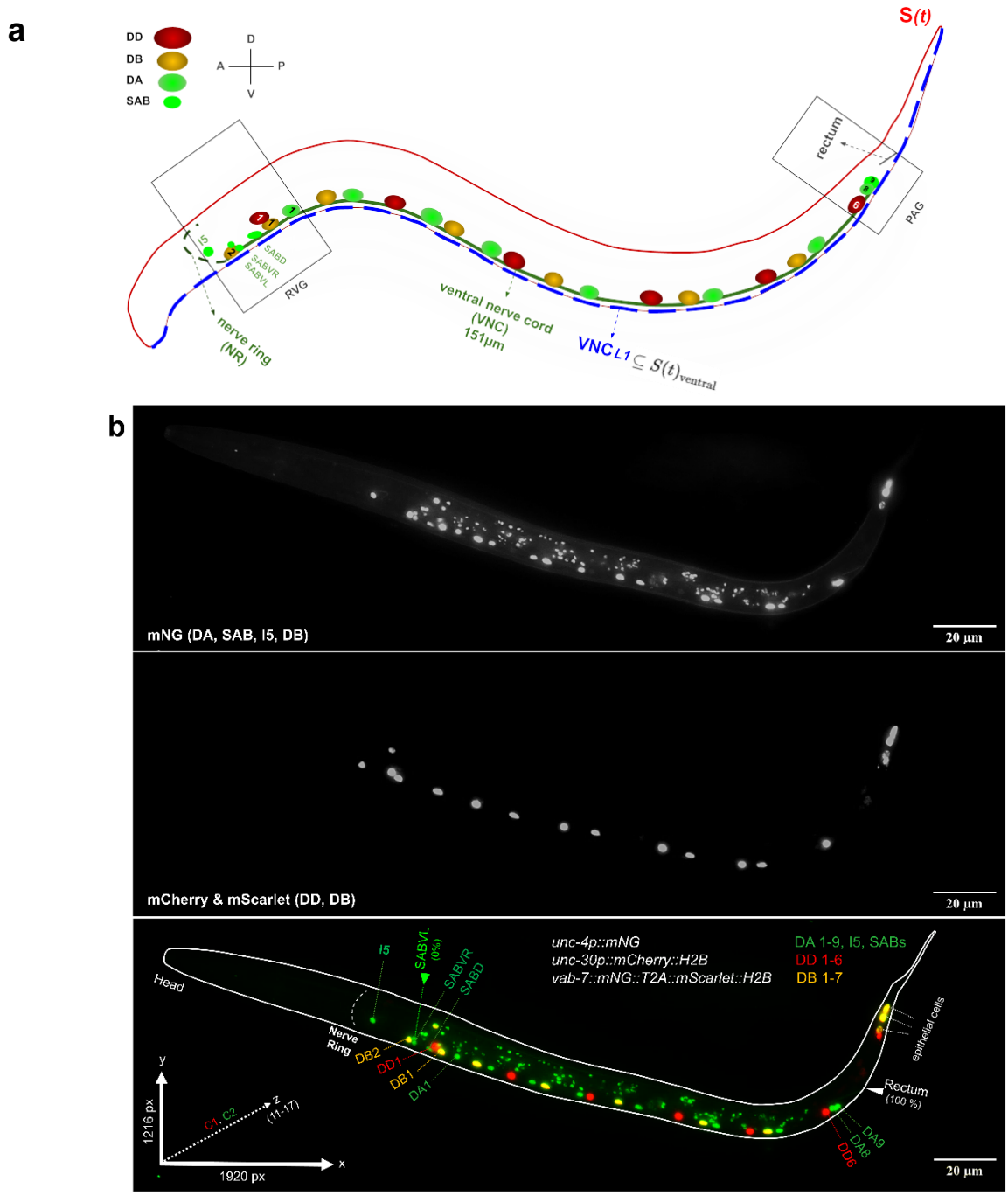

Fig. 1 | Nervous system of a newly hatched wild type *C. elegans* hermaphrodite

(a) Diagram of a wild type (WT) L1 *C. elegans* highlighting key nervous system structures: the nerve ring, ventral nerve cord (VNC), motoneurons, SAB neurons (SABVL, SABVR, SABD), and the I5 pharyngeal neuron. The retrovesicular ganglion (RVG) includes SAB neurons, DA1, DB1, DB2, and DD1; the preanal ganglion (PAG) comprises DA8, DA9, and DD6. The worm's body is modeled as a parametric spline function S(t) (red), with the VNCL1 spline (dashed blue) running along the ventral side of S(t), depicted in parallel to the VNC (green). **(b)** Fluorescent micrograph of the VNC expressing nuclear-targeted fluorescent reporters. In our triple reporter strain, mCherry labels DD neurons (red), mNG labels DA, SAB and I5 neurons (green), and the overlapping expression of mScarlet and mNG labels DB neurons (orange). In addition, epithelial cells in the tail region are labeled with the DB reporter. Channels were merged and optimal z-slices were collapsed through maximum intensity projection (MIP) for displaying all motoneurons along the anterior–posterior (AP) axis. Reference points are set at SABVL (0% anterior) and the rectum (100% posterior). For clarity, single-channel images for green and red are also provided. Scale bar, 20 µm (1 µm = 6.82 pixels).

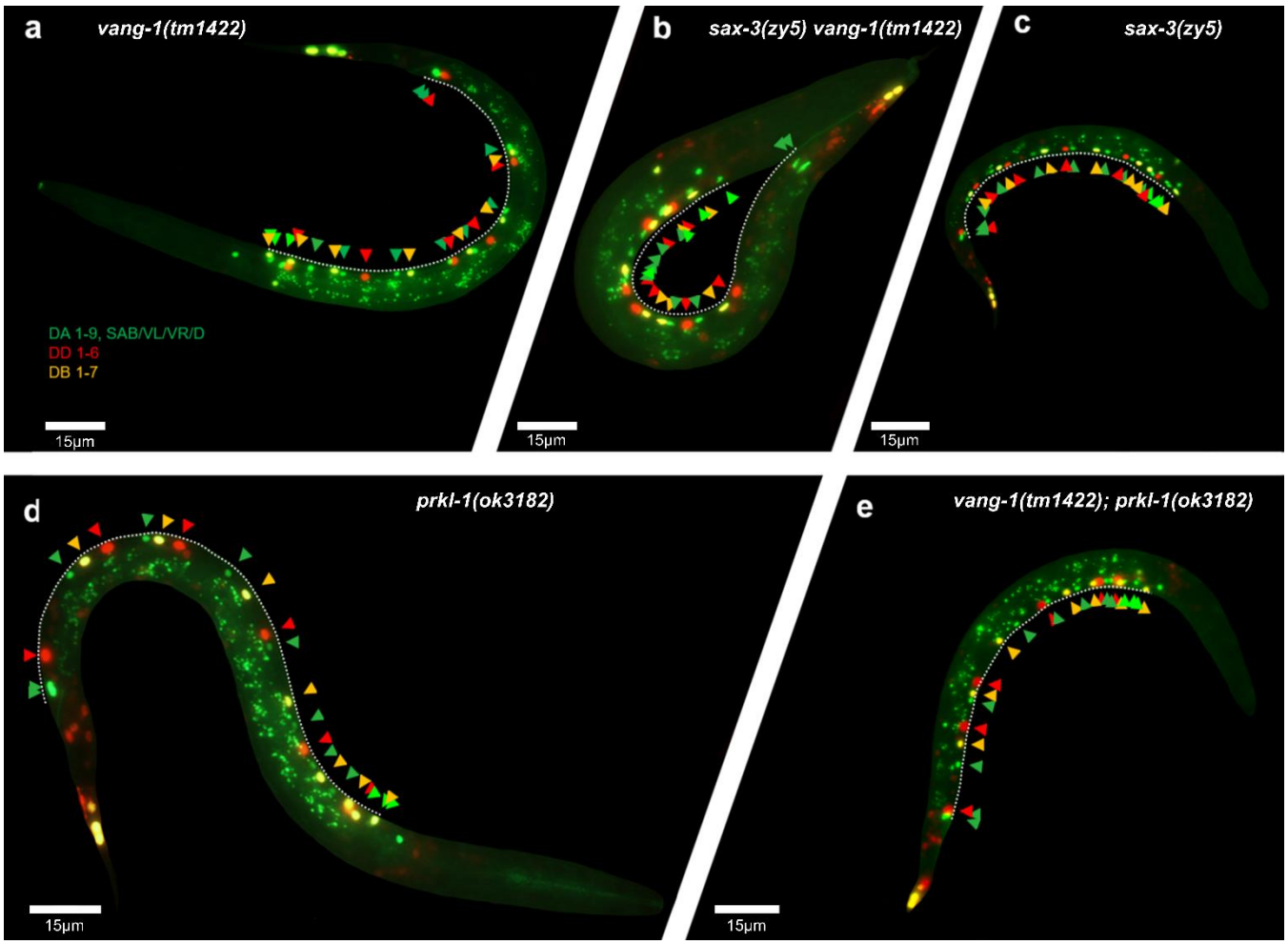

Fig. 2 | Neuronal positioning across mutant genotypes in *C. elegans*.

(a–e) Representative fluorescent images of *vang-1(tm1422)*, *sax-3(zy5) vang-1(tm1422)*, *sax-3(zy5)*, *prkl-1(ok3182)* and *vang-1(tm1422); prkl-1(ok3182)* mutants respectively. Green arrows: DAs and SABs; red arrows: DDs; orange arrows: DBs. Neuron sorting patterns along the AP axis of the VNC are not fully discernible; however, neuronal positions appear to shift and vary across genotypes. The dashed white line represents the VNC curve. Note: Worms were segmented from the background.

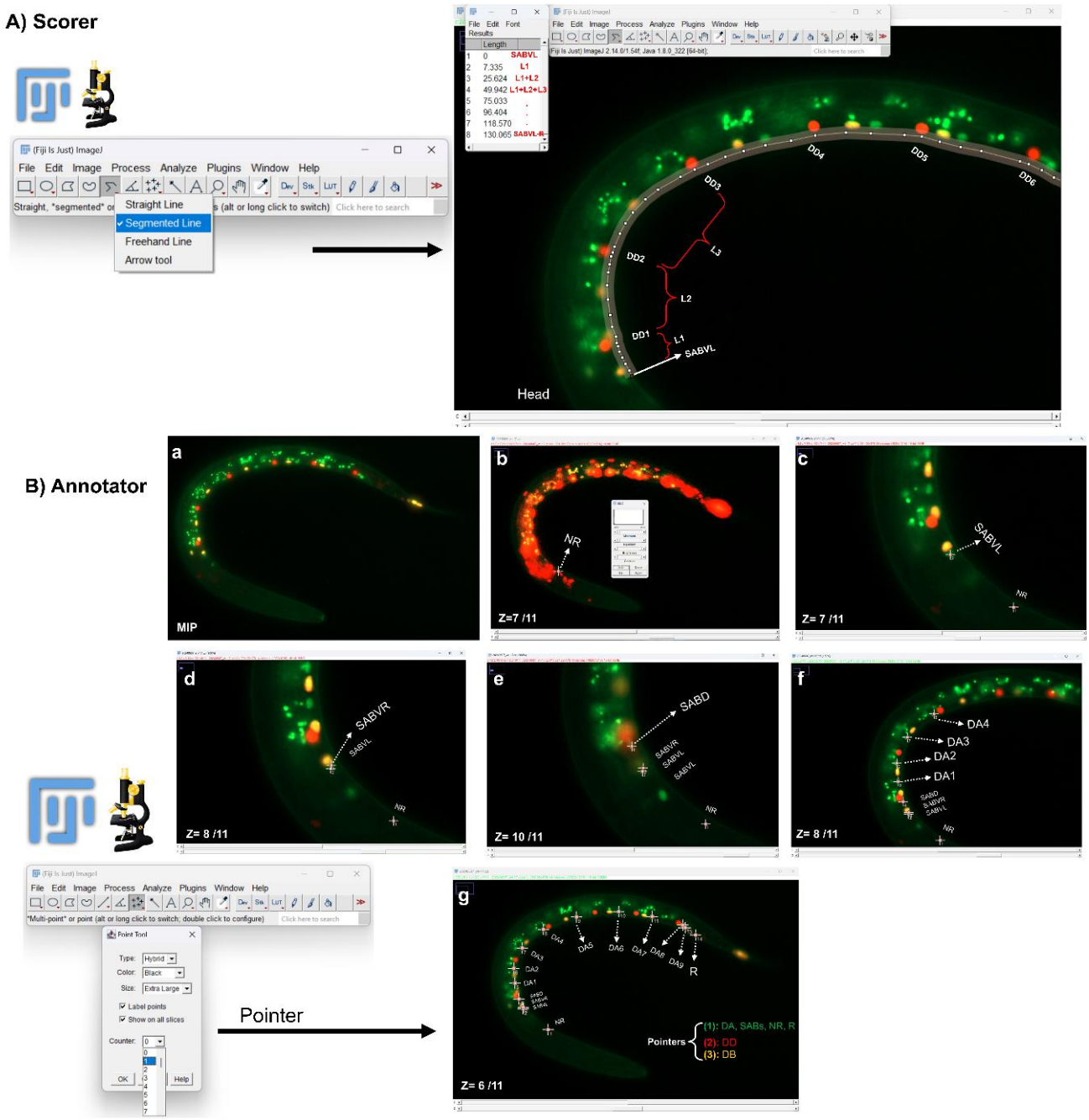

Fig. 3 | Fiji/ImageJ tools for scoring and annotating 3D micrographs.

(A) Manual measurement of relative distance of motoneurons using the Segmented Line tool for distance analysis. **(B)** Neuron annotation to obtain positional coordinates (x, y) using three different pointers (Counters in Fiji): 1 for DAs, SABs, and rectum; 2 for DDs; 3 for DBs. The Multi-point tool is used to annotate neurons, serving as a preparatory step for further precise computational distance analysis.

(a) Merged and collapsed image where some neurons may not be visible due to slice collapsing. **(b)** Annotation of the nerve ring (NR) by enhancing red channel signals between two red clusters. **(c–g)** Sequential annotation of neurons across z-slices: **(c)** SABVL at z = 7/11; **(d)** SABVR at z = 8/11; **(e)** SABD at z = 10/11; **(f)** DA1–DA4 at z = 8/11; **(g)** remaining DAs and rectum at z = 6/11. Annotations are performed by clicking on neurons across z-slices as they become visible. Text and arrow labels in dicate pointer (1) assigned to DAs, SABs, and rectum. The same annotation process is applied for DDs and DBs.

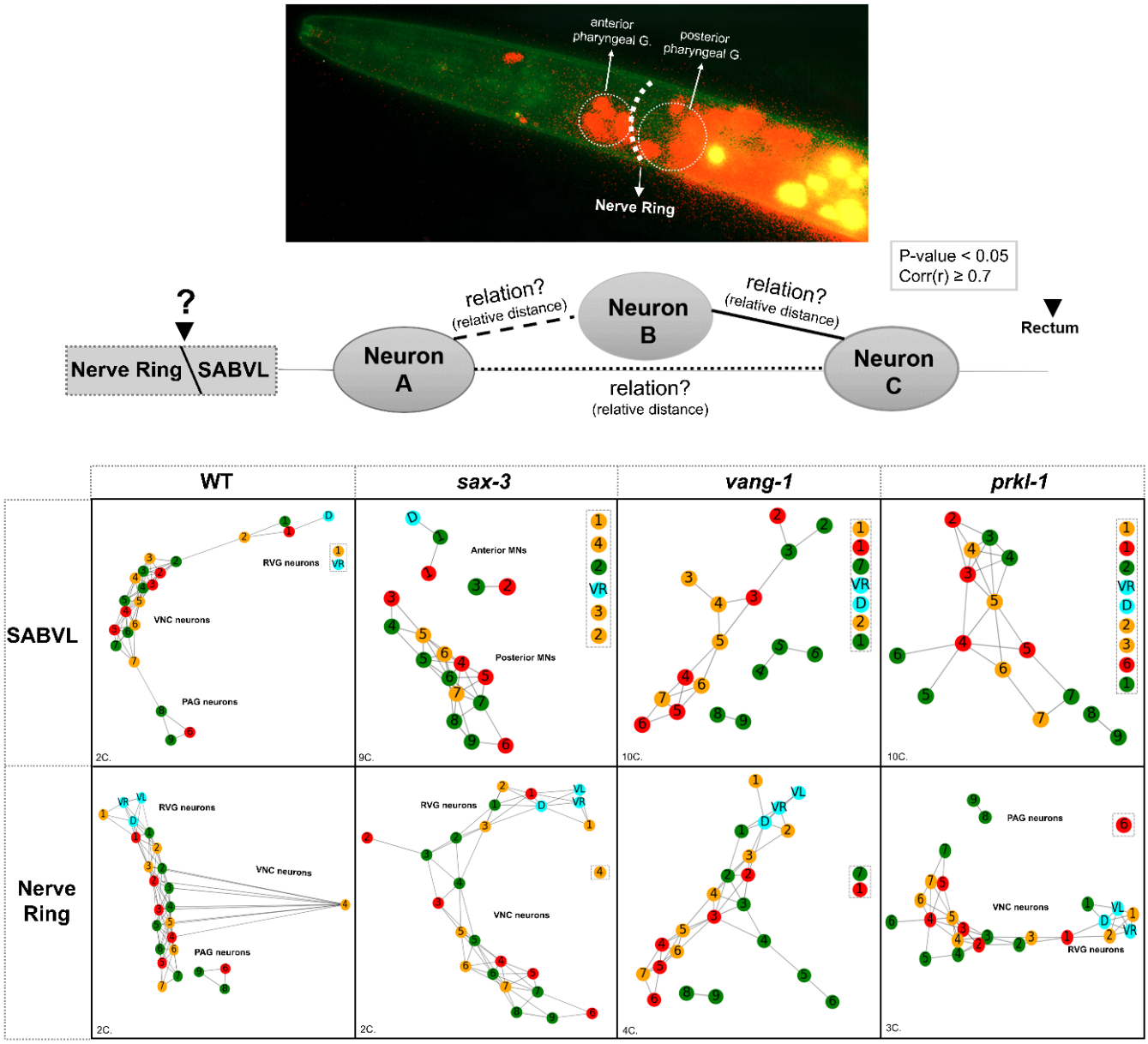

Fig. 4 | Correlation network of WT, PCP and *sax-3* mutants to validate the anterior reference point.

(a) Enhanced red channel intensity highlights the nerve ring in wild type, located between two prominent red clusters, the anterior and the posterior pharyngeal ganglia. **(b)** Diagram of correlation network analysis ($p < 0.05$, $r > 0.7$), where nodes represent neurons and edges represent significant relative distances. **(c)** Network components across genotypes. Fewer components are observed when using the nerve ring reference point, particularly in *sax-3* mutants, suggesting that the nerve ring is also shifting along with its adjacent anterior neurons. Numbers in the panels indicate the number of network components. Neuron types are color-coded: DAs in green, SABs in cyan, DDs in red, and DBs in orange. In WT with the SABVL reference point, networks display a biological pattern where VNC neurons are highly correlated centrally, while RVG neurons anteriorly and PAG neurons posteriorly show lower correlation with VNC neurons.

Algorithm Automated Contour-Based in VNC-Dist

Input:
- Segmented worm images as PNG **from SAM**
- CSV files with neuron (SAB, DA, DD, DB), neve ring and rectum coordinates **from Fiji/ImageJ Annotator**

Output:
- Annotated images with neuron projections as PNG
- CSV files with normalized neuronal relative distances.

Procedure:

1. Initialize application to load image and CSV folders.
2. For each image:
- **2.1.** Match the image to its corresponding CSV file.
- **2.2.** If CSV is found:
  - **2.2.1. Image Preprocessing:**
    - Convert the image to grayscale and threshold to binary.
    - Extract contours and select the largest contour.
  - **2.2.2. Spline Fitting**:
    - Fit a cubic spline to the largest contour points.
  - **2.2.3. Neuron Projection:**
    - Project the neurons (SAB, DA, DD, DB), nerve ring, and rectum from the CSV onto the fitted spline.
    - Reorder contour points starting from nerve ring.
  - **2.2.4. Relative Distance Calculation:**
    - Compute arc length between neurons along the spline.
    - Normalize neuronal relative distances from nerve ring or SABVL as 0% to rectum as 100%.
- **2.3.** Else, log a warning.

3. End the application when all images and CSVs are processed.

Fig. 5 | Pseudocode for the automated contour-based algorithm

The algorithm used in the VNC-Dist processes PNG images of segmented worms from SAM and corresponding CSV files containing neuron coordinates (SAB, DA, DD, DB), nerve ring and rectum from Fiji/ImageJ Annotator. It extracts the largest contour from each worm image, fits a cubic spline to this contour, and projects neuron, nerve ring and rectum coordinates onto the spline using the shortest distance. Using adaptive Gaussian quadrature, we calculated the normalized relative distances (%) for each neuron from anterior (SABVL, 0%) to posterior (rectum, 100%). The output includes annotated PNG images and CSV files of these distances. Relative distances can be normalized using either SABVL or the nerve ring as the anterior reference point.

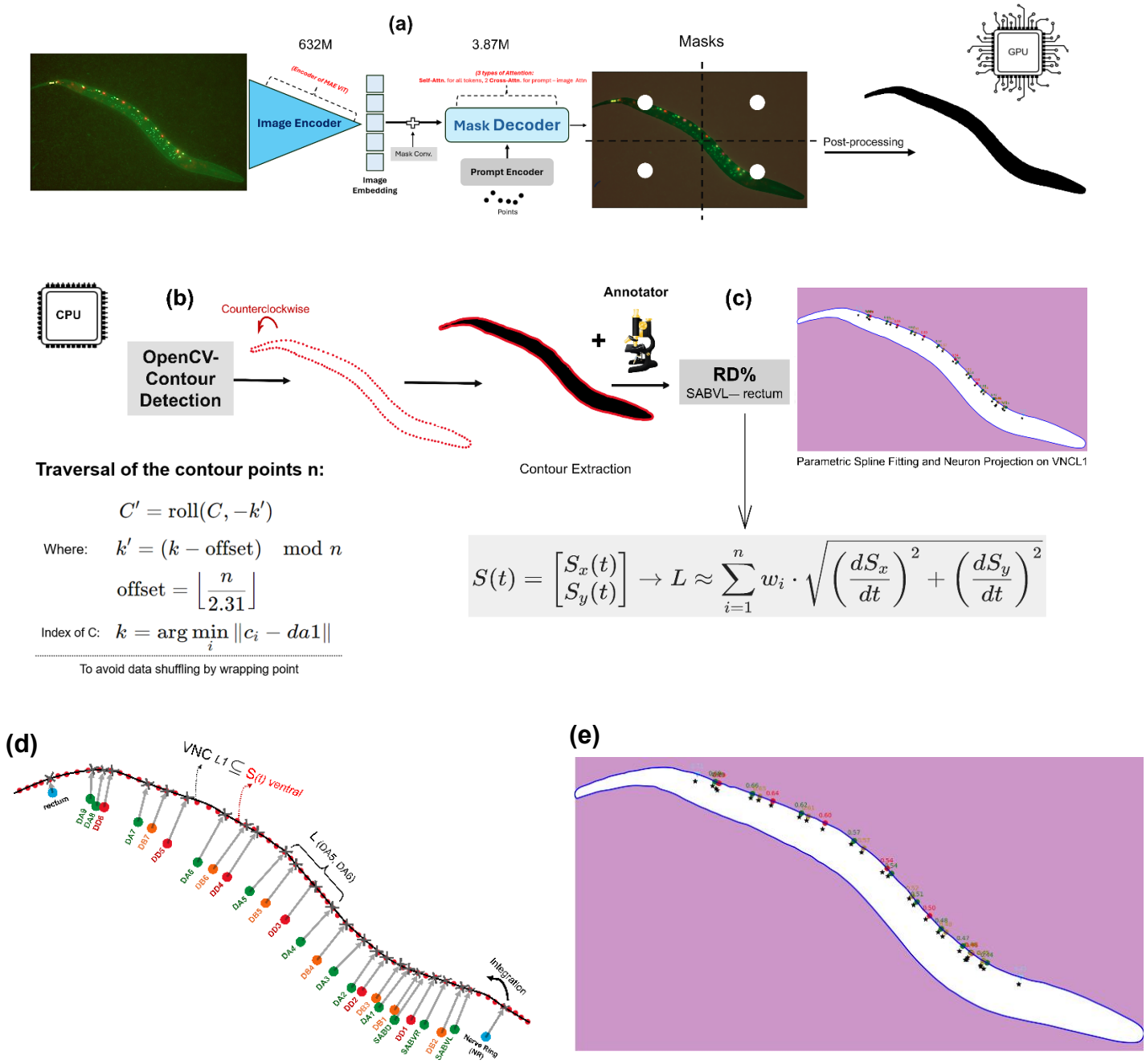

Fig. 6 | Overview of VNC-Dist pipeline for analyzing VNC neuron positioning.

(a) Fully automated worm body segmentation using the huge size of zero-shot Segment Anything Model (SAM) with a grid of four-point prompts, running exclusively on GPU, followed by post-processing to isolate the segmented worm. **(b)** Manual neuron annotation in Fiji/ImageJ to obtain (x, y) coordinates. **(c)** Contour extraction and spline fitting via adaptive Gaussian quadrature; neurons are projected onto the VNCL1 spline to calculate their relative distances (RDs) from anterior (SABVL: 0%) to posterior (rectum: 100%). **(d, e)** RD Calculation: Each neuron is projected onto the VNCL1 spline of the worm's ventral shape function. The parametric spline's t-values determine neuronal order from anterior to posterior, normalized with SABVL at 0% and the rectum at 100%.

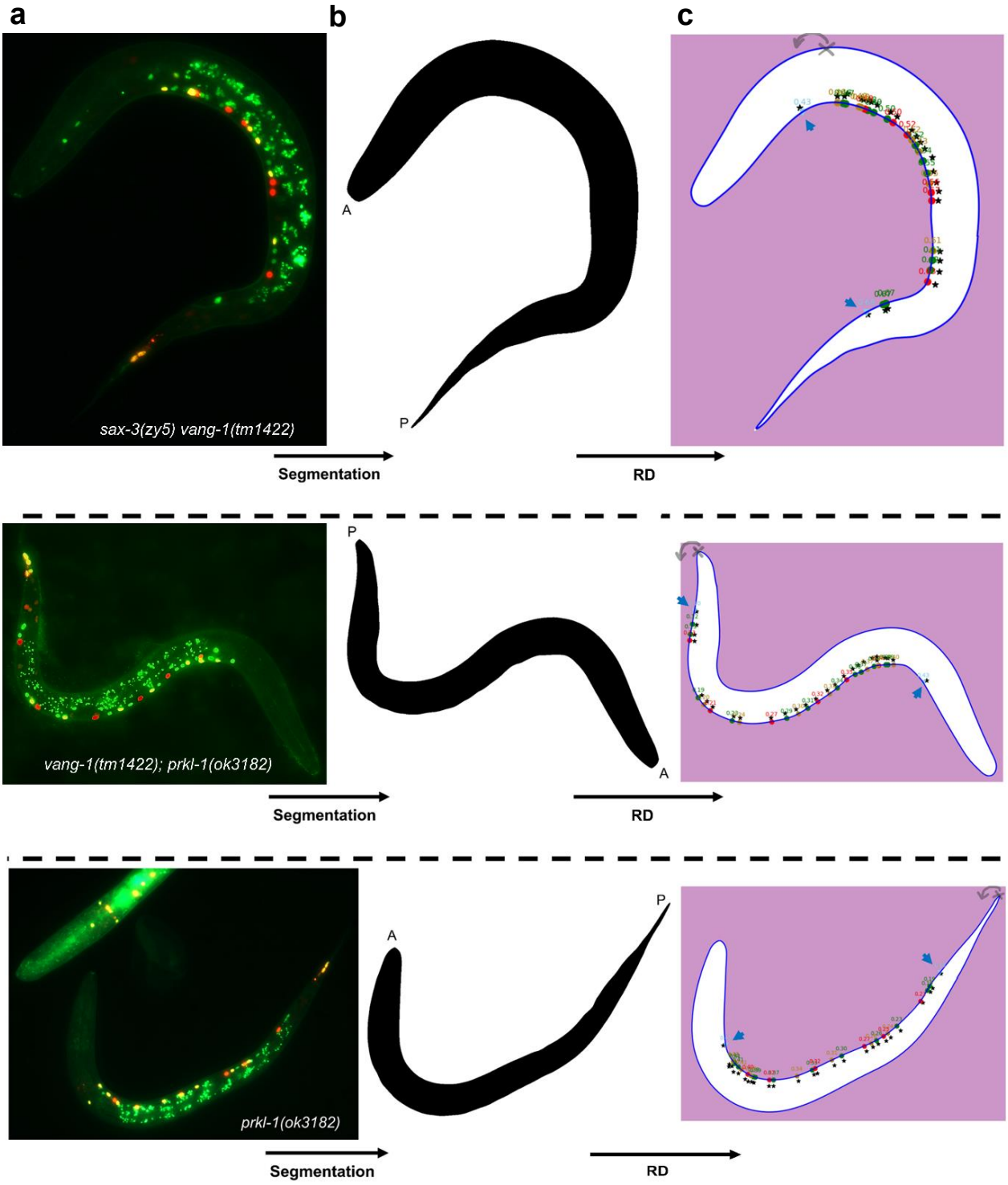

Fig. 7 | Neuron positioning in *C. elegans* mutants analyzed with VNC-Dist

Outputs from the VNC-Dist pipeline for mutants: (i) *sax-3(zy5) vang-1(tm1422)*, (ii) *prkl-1(ok3182)*, and (iii) *vang-1(tm1422); prkl-1(ok3182)*. **(a)** The raw 2D merged and collapsed fluorescent micrographs **(b)** underwent segmentation to extract the worm body contours. **(c)** The segmented images were further processed to calculate the relative distance (RD) of motoneurons along the AP axis of the VNC. The direction of contour detection proceeds from the topmost point counterclockwise (indicated by the grey arrow), ensuring that neurons remain in the correct sequential order without reordering. Blue arrows near the anterior (A) and posterior (P) ends indicate the nerve ring and rectum respectively.

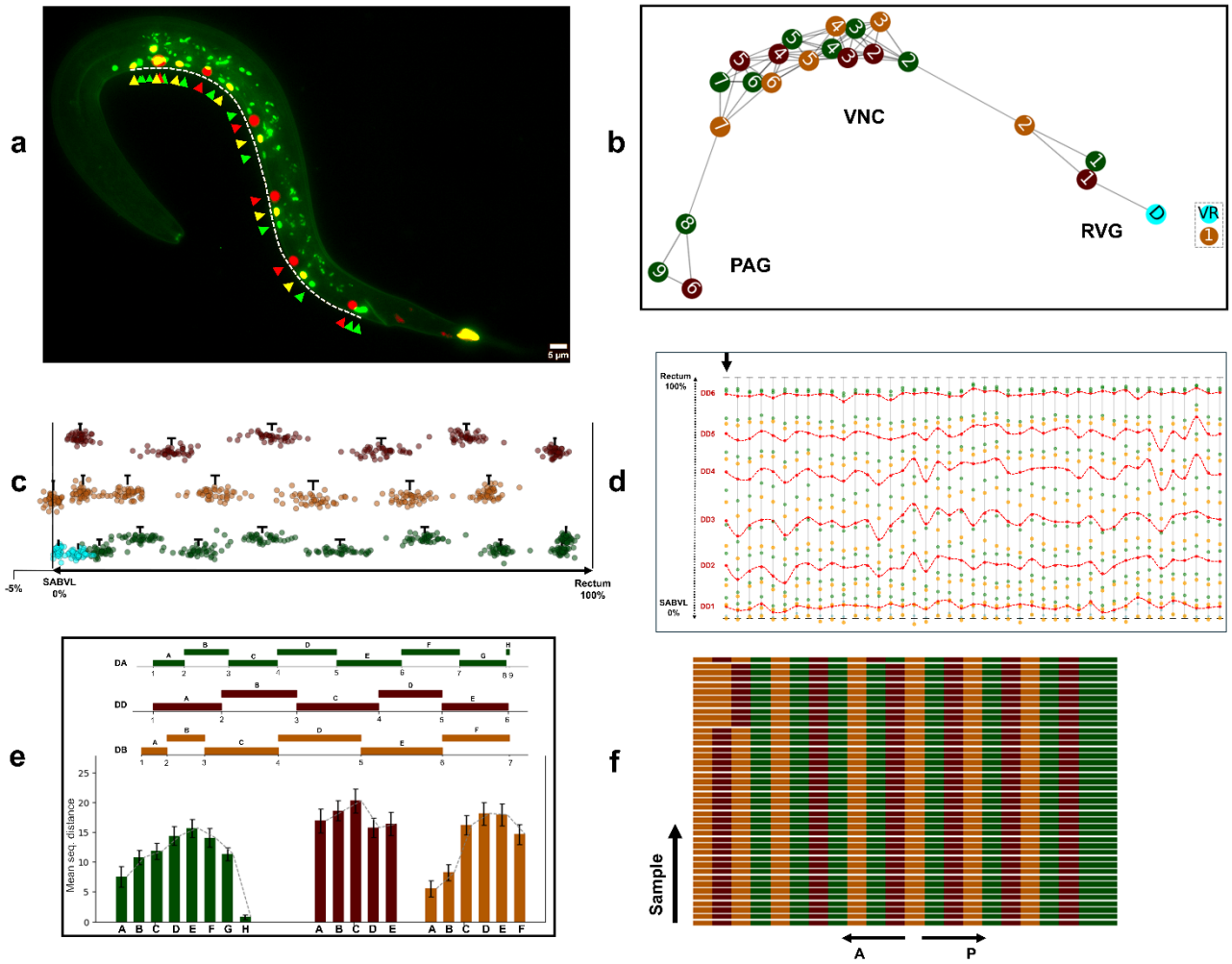

Fig. 8 | VNC Neuron positioning in L1 wild type *C. elegans*

(a) Merged and collapsed micrograph showing motoneurons along the AP axis (dashed line). Green arrows: DAs and SABs; red arrows: DDs; orange arrows: DBs. Scale bar, 5 µm. **(b)** Correlation network analysis grouping all motoneurons into VNC, RVG, and PAG neurons relative to SABVL (0%). **(c)** Dot plot of normalized relative distance of motoneurons from SABVL (0%) to rectum (100%) for n = 42. Error bars represent mean ± 95% CI. **(d)** Neuron positions per animal from SABVL to rectum, with mean positions indicated (black arrow). Neurons follow a consistent pattern, color-coded by type. **(e)** Sequential spacing of neurons along the AP axis, with segmented bar plots (top). Error bars represent mean ± SD. **(f)** Neuron type arrangement analysis of all 22 motoneurons from anterior to posterior. Neuron types: DA (green), DD (red), DB (orange), SAB (cyan).

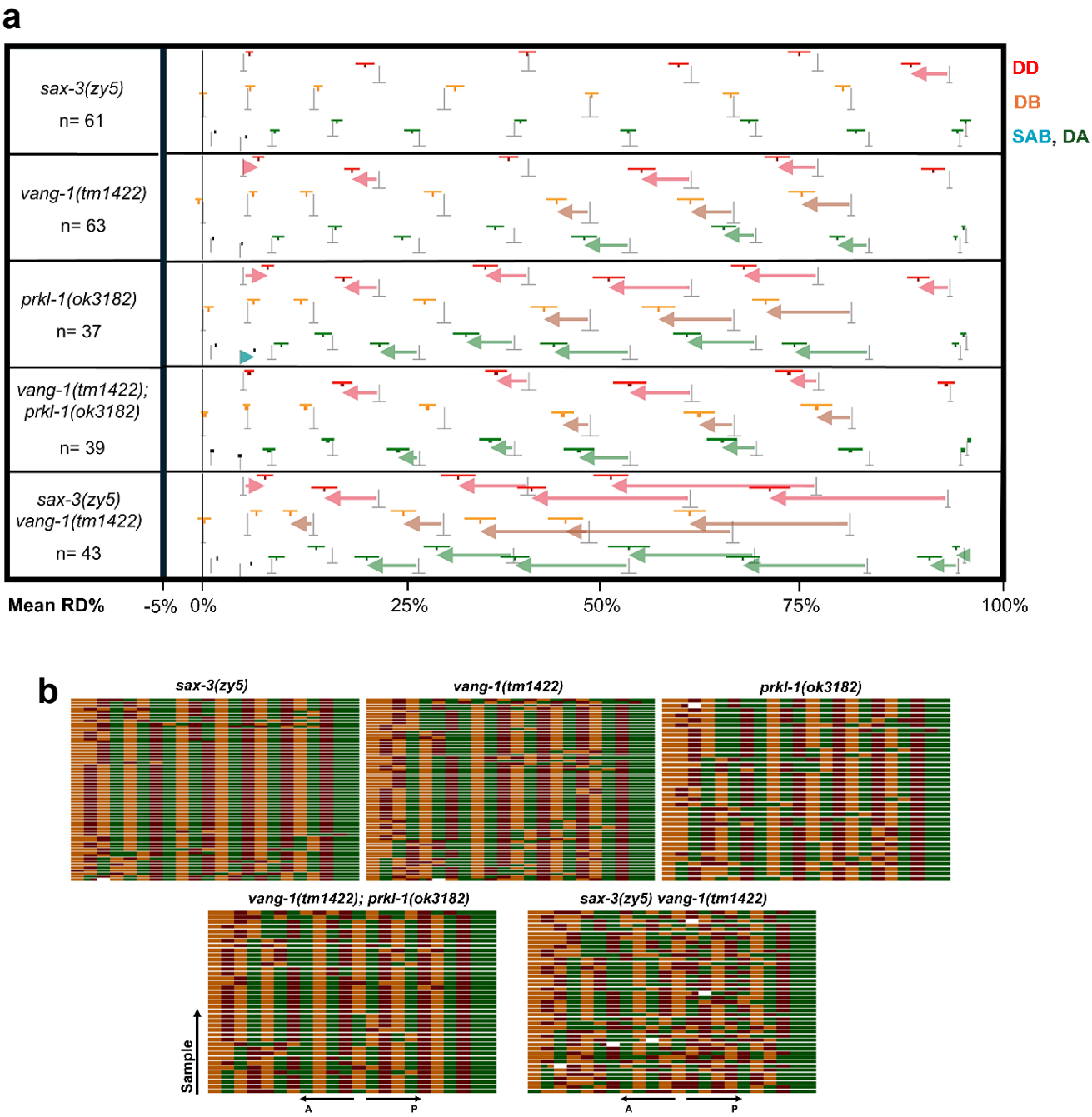

Fig. 9 | Quantification of neuron displacements in single and compound mutants via VNC-Dist.

(a) Mean relative distance of neuronal positions in *C. elegans* at L1 along the AP axis of VNC from SABVL (0%) to rectum (100%) are shown. Mean position of SAB (cyan), DD (red), DA (green), and DB (orange) neurons are compared to WT (gray). Significant anterior shifts ($p < 0.05$, Cohen's $d > 0.7$) are visualized as arrows from WT to mutant. Error bars represent mean neuron positions ± 95% confidence intervals (CI). **(b)** ranked positions of all 22 motoneurons from left (anterior) to right (posterior), highlighting how neuron types (DBs: orange, DAs: green, and DDs: red) shift or switch across different genotypes along the AP axis. Overlapping neuron positions in mutants are indicated by skipped white bars.