High-throughput analysis of dendritic and axonal arbors reveals transcriptomic correlates of neuroanatomy

Olga Gliko[1,*], Matt Mallory[1], Rachel Dalley[1], Rohan Gala[1], James Gornet[2], Hongkui Zeng[1], Staci Sorensen[1], and Uygar Sümbül[1,*]

[1]Allen Institute, Seattle WA, USA

[2]California Institute of Technology, Pasadena, CA, USA

[*]Correspondence: olgag@alleninstitute.org, uygars@alleninstitute.org

Abstract

Neuronal anatomy is central to the organization and function of brain cell types. However, anatomical variability within apparently homogeneous populations of cells can obscure such insights. Here, we report large-scale automation of neuronal morphology reconstruction and analysis on a dataset of 813 inhibitory neurons characterized using the Patch-seq method, which enables measurement of multiple properties from individual neurons, including local morphology and transcriptional signature. We demonstrate that these automated reconstructions can be used in the same manner as manual reconstructions to understand the relationship between some, but not all, cellular properties used to define cell types. We uncover gene expression correlates of laminar innervation on multiple transcriptomically defined neuronal subclasses and types. In particular, our results reveal correlates of the variability in Layer 1 (L1) axonal innervation in a transcriptomically defined subpopulation of Martinotti cells in the adult mouse neocortex.

Introduction

The shape of dendrites and axons, their distribution within the neuropil, and patterns of their long-range projections can reveal fundamental principles of nervous system organization and function. In the cortex, much of our understanding depends on the anatomical and functional descriptions of cortical layers. Yet, the origin and role of morphological and molecular diversity of individual neurons within cortical layers beyond broad subclass identities is poorly understood, in part due to low sample numbers. While molecular profiling techniques have recently improved by orders of magnitude, anatomical characterization remains time consuming due to continued reliance on (semi-)manual reconstruction.

Improvements in the throughput of the Patch-seq technique[1–8] have enabled measurement of electrophysiological features, transcriptomic signatures, and local morphology in slice preparations for thousands of neurons in recent studies.[5,6] In these repetitive experiments where maintaining a high throughput is a primary goal,[9] the brightfield microscope's speed, ease of use, and ubiquity make it an attractive choice to image local neuronal morphology despite its limitation in imaging resolution. Therefore, while the experimental steps are streamlined, morphological reconstruction remains a major bottleneck of overall throughput, even with state-of-the-art semi-manual tools.[5]

A rich literature exists on automated segmentation in sparse imaging scenarios. However, these methods typically focus on high-contrast, high-resolution images obtained by optical sectioning microscopy (i.e., confocal, two-photon, and light-sheet),[10–18] and are not immediately applicable to brightfield images because of the significantly worse depth resolution and the complicated point spread function of the brightfield microscope. Moreover, segmentation of full local morphology together with identification of the axon, dendrites, and soma has remained elusive for methods tested on image stacks obtained by the brightfield microscope.[19–22] Therefore, we first introduce an end-to-end automated neuron reconstruction pipeline (Figure 1a) to improve scalability of brightfield 3D image-based reconstructions in Patch-seq experiments by a few orders of magnitude. We note that our primary goal is not to report on a more accurate reconstruction method *per se*. Rather, we aim to demonstrate how automated tracing, with its potential mistakes, can be leveraged to rigorously address certain scientific questions by increasing the throughput. To this end, we select a set of brightfield images and use the corresponding manually reconstructed neuron traces to assign voxel-wise ground truth labels (axon, dendrite, soma, background). Next, we design a custom deep learning model and train

it on this curated ground truth dataset to perform 3D image segmentation using volumetric patches of the raw image as input. We implement fully automated post-processing steps, including basal vs apical dendrite separation for pyramidal cells, for the resulting segmentations to obtain annotated traces for each neuron. We compare the accuracy of these automated traces with a held out set of manual reconstructions, based on geometrical precision, a suite of morphometric features, and arbor density representations derived from the traces.

We utilize this pipeline to reconstruct a large set of neurons from Patch-seq experiments, and use the transcriptomic profiles captured from the same cells to systematically search for gene subsets that can predict certain aspects of neuronal anatomy. The existence of a hierarchical transcriptomic taxonomy[23] enables studying subsets of neurons at different levels of the transcriptomic hierarchy. At the finest scale of the hierarchy (transcriptomic types or "t-types"[5]), we study seven interneuron types and focus on a transcriptomically defined sub-population of L1-projecting, *Sst* gene expressing neurons (Sst cells) that correspond to Martinotti cells (e.g., [24]). While previous studies have elucidated the role of Martinotti cells in gating top-down input to pyramidal neurons via their L1-innervating axons,[25,26] the wide variability in the extent of L1 innervation behind it is not well understood. Our results suggest transcriptomic correlates of the innervating axonal mass, which may control the amount of top-down input to canonical cortical circuits. Our approach represents a general program to systematically connect gene expression with neuronal anatomy in a high-throughput and data-driven manner.

Results

As the first step to automate the reconstruction of in-slice brightfield images of biocytin-filled neurons, we curate a representative set of 51 manually traced neurons. We develop a topology-preserving variant of the fast marching algorithm[13] to generate volumetric labels from manual traces (Figure 1b). We train a convolutional neural network (U-Net)[16,27–29] using image stacks and labels as the training set and employing standard data augmentation strategies to produce initial segmentations of neuronal morphologies (Figure 1b,c). While knowledge of axonal vs. dendritic branches informs most existing insight, their automated identification poses a challenge due to the limited field-of-view of artificial neural networks. We find that image and trace context that is in the vicinity of the initial segmentation is sufficient to correct many axon vs dendrite labeling mistakes in an efficient way because this effectively reduces the problem to

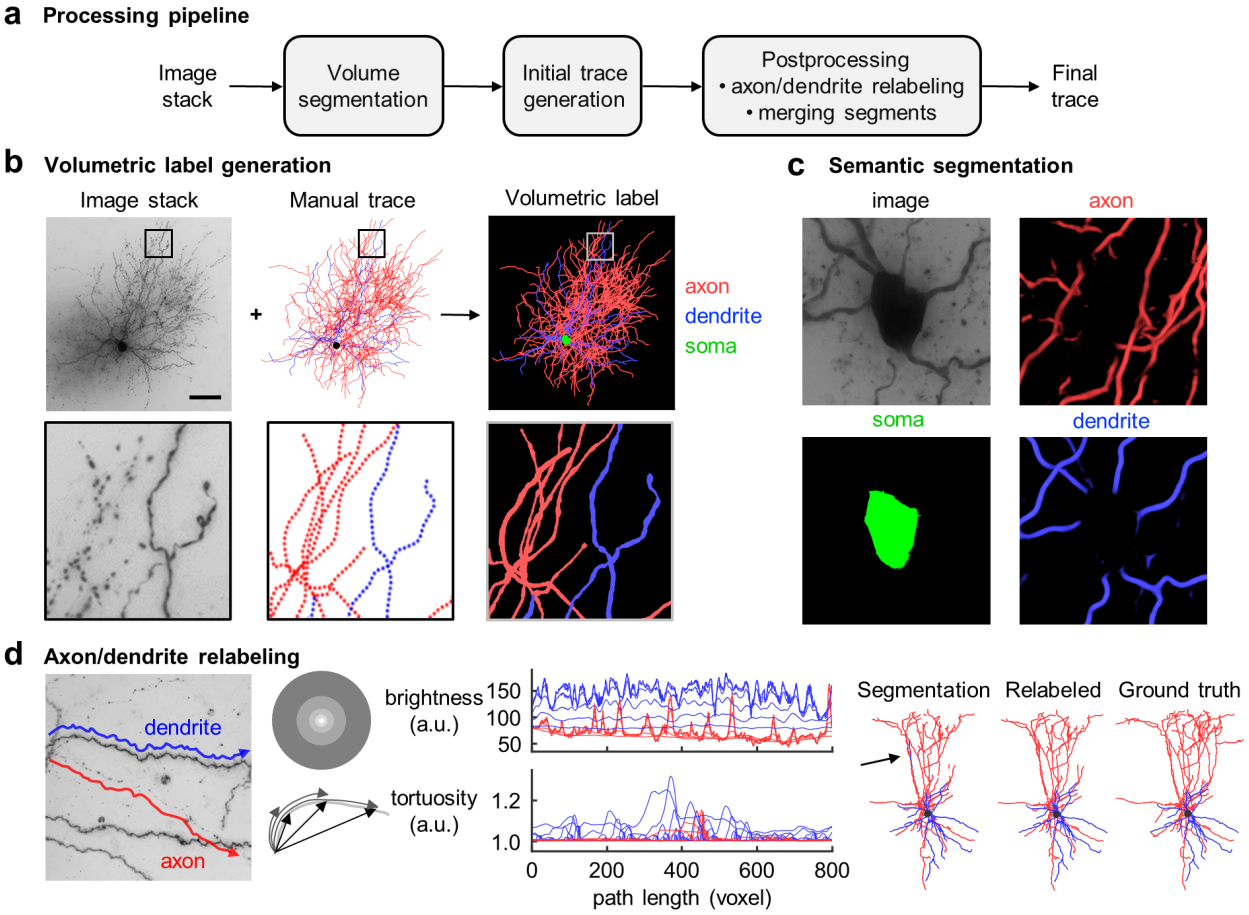

Figure 1: **Neuron reconstruction pipeline for in-slice bright-field images of biocytin-filled neurons.** **a**, Processing pipeline. Convolutional neural network (CNN) segmentations of 3D image stacks are post-processed by custom machine learning tools to produce digital representations of neuronal morphologies. **b**, Topology preserving fast marching algorithm generates the volumetric label from raw image stack and manual skeletonization. Dendrites (blue), axons (red), soma (green) are separately labeled to train a supervised CNN model. Scale bar, 100 μm. **c**, Semantic segmentation provides accurate soma location and boundary. **d**, Axon/dendrite relabeling. A neural network model predicts node labels from multiple image brightness and trace tortuosity features based on local contexts of different size along the initial trace. (left, example image of dendrite and axon segments; middle, corresponding feature plots; right, automated traces of test neuron with/without relabeling vs manual trace). Arrow indicates nodes mislabeled by segmentation and corrected during post-processing.

a single dimension, i.e., features calculated along the 1D initial trace (Figure 1d). We further algorithmically post-process the segmentations to correct connectivity mistakes introduced by the neural network and obtain the final reconstruction of axons and dendrites (Figure 1d, Figure 2a, Figures S1-S11, Methods). The overall pipeline produces neuron reconstructions in the commonly-used swc format[30] from raw image stacks at a rate of ∼100 cells/day using 16 GPU cards (Methods). Importantly, this represents two orders of magnitude improvement in speed over semi-manual segmentation[5] with one anatomist. We have so far processed the cells reported in Ref.,[5] which mapped neurons to an existing taxonomy of transcriptomic cell types[23] and introduced a transcription-centric, multimodal analysis of inhibitory neurons. We have also processed a set of ∼700 excitatory neurons which are analyzed in Ref.[31] After quality control steps (Methods), we focus on a set of 813 interneurons for further study in this paper.

The proposed pipeline produces end-to-end automated reconstructions in single-cell imaging scenarios. However, in practice, neurons are patched near each other to increase the throughput of physiological and transcriptomic characterization. The resulting image, which typically centers on the neuron of interest, can therefore contain neurites from other neurons. Neurites from off-target neurons within the image stack cannot be properly characterized because they rarely remain in the field of view. As part of algorithmic post-processing and quality control, disconnected segments are removed automatically when they remain relatively far from the cell of interest (Methods). When multiple neurons are patched in close proximity, quality control by a human is needed to check for and remove nearby extraneous branches. To ensure integrity of presented results with minimal manual effort, if quality control suggests the existence of nearby branches, the cell is not used if it was not manually traced before. If manual trace already exists, we simulate manual branch removal based on a mask obtained from the manual trace (Methods). We report quantification results separately for reconstructions obtained with/without nearby branch removal.

We evaluate the quality of automated traces by comparing them to the manual traces which we regard as the ground truth. To compare a pair of automated and manual traces, we perform a bi-directional nearest-neighbor search to find correspondence nodes in both traces within a certain distance.[13] A node in the automated trace that has (does not have) a corresponding node in the manual trace is referred to as a true (false) positive node, and a node in the manual trace that does not have a corresponding node in the automated trace is referred to as a false negative node. We calculate this metric separately for axonal and dendritic nodes, as well as for all nodes regardless of the type, and compute corresponding precision, recall, and f1-score (harmonic mean of the precision and recall) values. These metrics indicate how

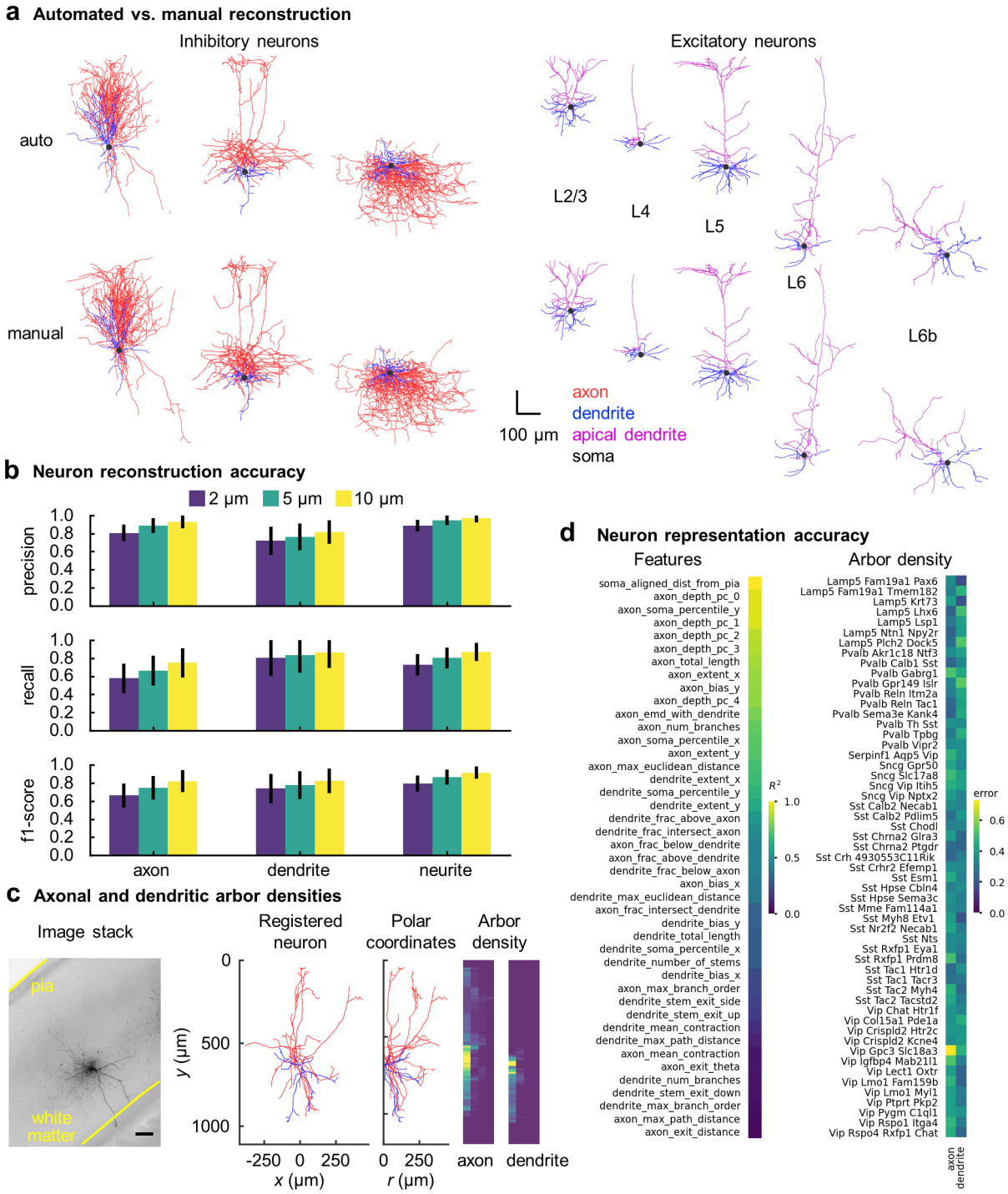

Figure 2: **Assessing quality of automated reconstructions.** **a**, Automated and manual traces of example test neurons (left, inhibitory neurons; right, excitatory neurons – apical and basal dendrites are assigned for excitatory cells.) **b**, Neuron reconstruction accuracy. Precision, recall, and f1-score values are calculated by comparing automated and manual trace nodes within a given distance (2, 5 and 10 μm). Mean values over 347 cells (error bars: standard deviation) are shown for axonal, dendritic and neurite (combined axonal and dendritic) nodes. **c**, Generation of 2D axonal and dendritic ADRs. Scale bar, 100 μm. **d**, R^2 values (left) and average root-mean-squared error (right) between automatically vs. manually generated features (left) and ADRs for each t-type (right). ADRs are normalized to have unit norm.

well the automated trace captures the layout of the axon/dendrites/neurites in a reconstructed neuron. Figure 2b and Table S4 display that, at a search radius of $10\mu m$, the mean f1-score is above 0.8 for both axonal and dendritic morphologies. Therefore, we expect these automatically generated traces to perform comparably to their manually generated counterparts in analyses that do not require a resolution better than $10\mu m$, as we demonstrate below.

To further assess the similarity in the arbor layout and other aspects of morphology that are not captured by the node correspondence study, we use standard morphometric features. We find that while many features summarizing the overall morphology can be accurately predicted, features related to the topology of arbor shapes, such as maximum branch order, are prone to mistakes (Figure 2d).

While the quantitative analyses described above both suggest that automated reconstruction succeeds in broadly capturing neuronal morphology, including separation of axonal vs dendritic branches, they also demonstrate that important differences nevertheless remain between automated and manual traces. Therefore, to robustly analyze anatomical innervation patterns against potential topological mistakes introduced by the automated routine, we develop a 2D arbor density representation (ADR)[12,32–35] of axons and dendrites registered to a common laminar axis defined by cortical pia and white matter boundaries. Here, the vertical axis represents the distance from pia and the horizontal axis represents the radial distance from the soma node (Figure 2c). Note that this 2D representation still requires 3D imaging because many branches become undetectable in 2D projected images due to noise (e.g., Figure 1b). Moreover, standardizing the orientations of the brain and the tissue slice is challenging in high-throughput experiments so that the rotation around the laminar axis would be hard to control in 3D representations.

At the level of transcriptomic types, the ADRs calculated from automated reconstructions appear similar to those calculated from manual segmentations based on the root-mean-squared difference between them (Figure 2d). To better quantify this similarity, we compare the performance of the ADR against that of morphometric features[36] by training classifiers to predict t-types and subclasses (Sst, Pvalb, Vip, Sncg, Lamp5).[5] We find that the ADR is not statistically significantly worse than the morphometric features in terms of classification accuracy (Boschloo's exact test, asymptotically exact harmonic mean of p-values over multiple runs:[37] $p = 0.81$ for t-types, $p = 0.26$ for subclasses, Methods), consistent with Ref.[38] (Figure 3a,b).

We also test robustness against imperfections due to fully automated tracing by comparing the

classification accuracy obtained from automated tracing versus manual tracing on the same set of cells. End-to-end automation appears to perform similarly as manual tracing in cell type prediction based on ADRs (Figure 3c,d) and morphometric features (Figure S12). Here, an ultimate comparison might be between automatically generated ADRs vs manually generated morphometric features. Boschloo's test finds no significant difference between the two in t-type classification (harmonic mean $p = 0.72$) but the $\sim 5\%$ advantage of manual morphometric features in subclass classification is significant with harmonic mean $p = 0.01$.

Beyond the comparative aspect, these results demonstrate a correspondence between gene expression and the anatomy of local arbors as represented by the proposed registered 2D ADRs, which agrees with previous findings with morphometric features for these cells.[5] (subclass accuracy of $\sim 79\%$ vs random at 20%, most abundant label at 47%; t-type accuracy of $\sim 45\%$ vs random at $\sim 2\%$, most abundant label at $\sim 8\%$.) Even when the assignment is incorrect, the dominance of the entries around the main diagonal in Figure 3 demonstrates that cells are rarely assigned to transcriptomically far-away clusters based on the ADR or morphometric features. Note that the rows and columns of these confusion matrices are organized based on the reference taxonomy to reflect transcriptomic proximity (Figures S13 and S14). Therefore, the relative inaccuracy at the t-type level could be attributed to aspects of morphology not captured by the ADR or morphometric features (e.g., synapse locations), or other observation modalities (e.g., physiological, epigenetic) being key separators between closely related t-types.

Having established that registered 2D ADRs are as successful as a standardized, rich set of morphometric features in predicting transcriptomic identity and that ADRs can be generated in a fully automated manner from raw images with only mild loss in performance, we aim to uncover more explicit connections between gene expression and anatomy as captured by the ADR. Since layer-specific axon and dendrite innervations are prominently and reliably captured by the ADR, we study their transcriptomic correlates. We treat the search for genes that are predictive of laminar innervation strength (neurite length innervating a given layer) as a sparse regression problem[39,40] (Methods), and focus on 7 t-types whose morphologies are well sampled in our dataset with the help of automated reconstruction (3 Sst, 2 Pvalb, 2 Lamp5 types). That is, we aim to uncover minimal gene sets whose expression can predict the amount of axonal and dendritic innervation of individual laminae as well as the locations of the soma and centroids of the axonal and dendritic trees along the laminar axis. Throughout, we control the false discovery rate (FDR) by applying multiple testing correction (Methods). Tables 1 and S1-S3 summarize these results. We observe that no single anatomical feature is significantly predictable from gene expression for all inhibitory t-types and every studied

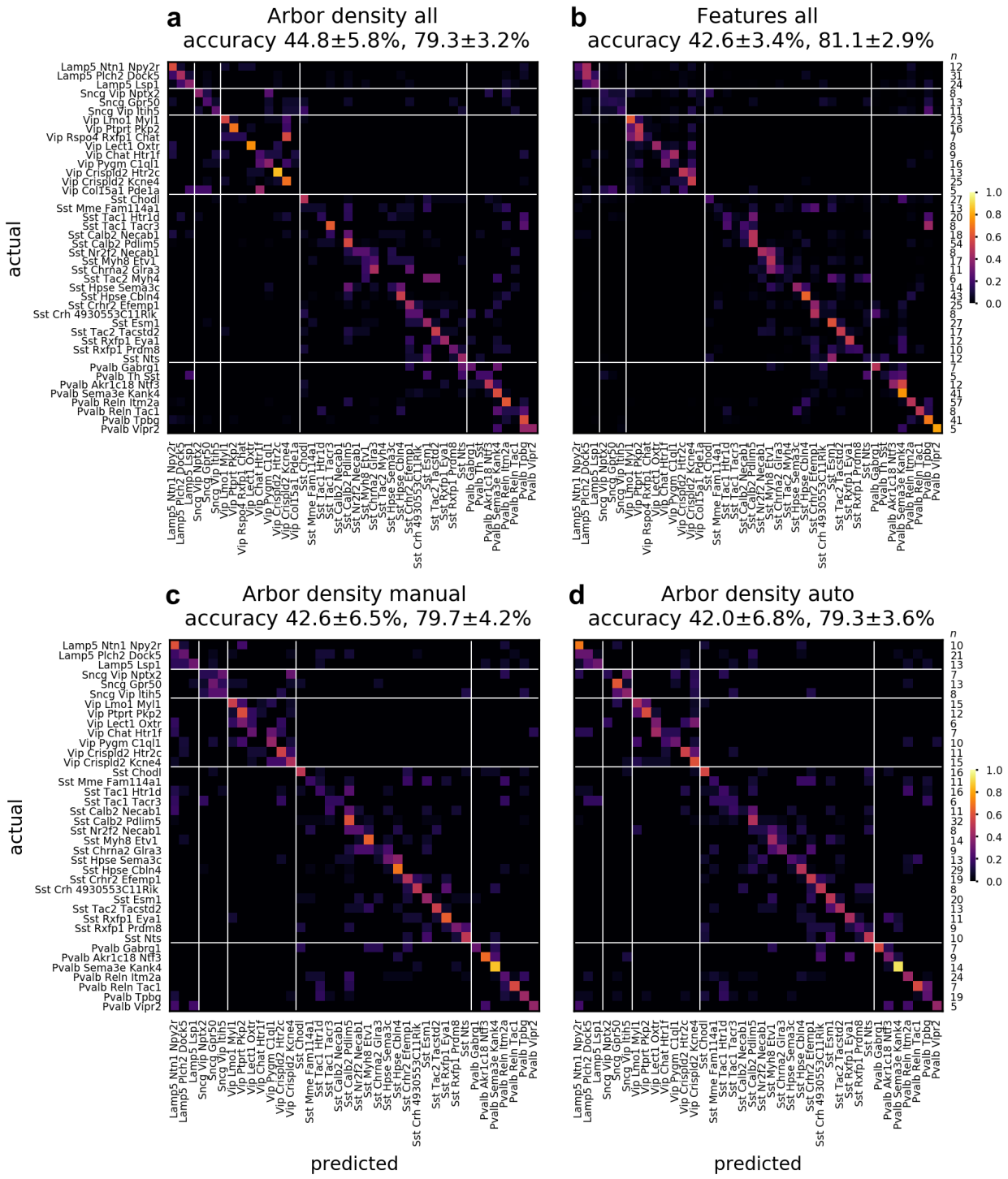

Figure 3: **Comparison of cell type classification accuracy based on the ADR vs a set of classical morphometric features.** Confusion matrix for the classification of 42 t-types based on ADRs (**a**) and morphometric features (**b**), using a combination of 246 automatically and 501 manually reconstructed cells. Confusion matrix for the classification of 38 t-types based on ADRs, using 488 manually (**c**) and automatically (**d**) reconstructed cells. Accuracy values reported in the headers refer to mean ± s.d. of the overall t-type and t-subclass classifiers, respectively, across cross-validation folds. Rightmost columns list the number of cells in each t-type (n).

t-type has at least one significantly predictable anatomical feature. (Only the L4 dendritic innervation strength is significantly correlated with gene expression for cells of type Sst Chodl.) Perhaps more interestingly, we find that the sets of laminae innervation-predicting genes within transcriptomically defined subclasses and t-types are highly reproducible (Table S3) and almost mutually exclusive(Figure 4g). These observations support a connectivity-related and discrete organization of cortical cells.[41] They also put forth a related question: can gene expression further predict innervation strength of a single layer in a continuum?

	Sst Calb2 Pdlim5	Sst Hpse Cbln4	Sst Chodl	Pvalb Reln Itm2a
L1 axon	**0** / **0.40**	1.000	0.637	**0.023** / 0.13
L2/3 axon	**0.013** / **0.30**	1.000	1.000	**0** / **0.35**
L4 axon	0.689	0.332	0.607	**0** / **0.35**
L5 axon	**0.033** / 0.15	**0.013** / **0.25**	1.000	0.102
L1 dendrite	**0.042** / 0.17	1.000	0.689	0.088
L2/3 dendrite	**0.033** / 0.15	1.000	0.332	**0.013** / **0.27**
L4 dendrite	0.697	**0** / **0.40**	**0.013** / **0.40**	0.058
L5 dendrite	**0** / 0.19	0.102	0.393	**0** / **0.25**
soma depth	**0** / **0.26**	**0.023** / 0.21	0.246	**0** / **0.34**
axon centroid	**0.023** / 0.16	0.210	0.058	**0** / **0.34**
dendrite centroid	0.150	**0.023** / **0.26**	0.096	**0** / **0.39**
	Pvalb Tpbg	Lamp5 Lsp1	Lamp5 Plch2 Dock5	
L1 axon	**0.033** / 0.21	**0.013** / **0.56**	**0.042** / **0.40**	
L2/3 axon	**0.023** / **0.27**	**0.042** / **0.52**	0.323	
L4 axon	0.216	1.000	0.351	
L5 axon	0.058	0.135	0.081	
L1 dendrite	**0** / **0.42**	0.283	**0** / **0.40**	
L2/3 dendrite	**0.042** / 0.21	1.000	1.000	
L4 dendrite	0.074	0.393	1.000	
L5 dendrite	**0.013** / **0.26**	1.000	1.000	
soma depth	**0.013** / **0.29**	**0.023** / **0.48**	0.067	
axon centroid	0.058	**0.013** / **0.44**	0.074	
dendrite centroid	**0.013** / **0.38**	**0.023** / **0.46**	0.074	

Table 1: Statistical significance and effect size values for predicting anatomical features from gene expression via sparse linear regression for five different cell types. For each entry, the FDR-corrected p-value as calculated by a non-parametric shuffle test is listed. If the value is considered statistically significant at $p \leq 0.05$, the R^2 value is also displayed (p / R^2). p-values less than or equal to 0.05 and R^2 values larger than or equal to 0.25 are shown in bold. A p value of 0 indicates that the calculated p value is less than 0.001, the sensitivity of the shuffle test, and less than 0.013 after FDR correction.

To elucidate this question, we choose a transcriptomically defined subpopulation that is well-sampled in our dataset with the help of automated reconstruction, produces a large effect size in the gene regression study (Table 1), and has been a source of confusion due to its anatomical variability: Sst Calb2 Pdlim5 neurons[23] represent a transcriptomically homogeneous subset of Martinotti cells, which are inhibitory neurons with L1-innervating axons that gate top-down

input to pyramidal neurons.[25,26] However, the amount of axon reaching L1 varies widely across cells.[5] Tables 1 and S3 show that a small set of genes, including genes implicated in synapse formation, cell-cell recognition and adhesion, and neurite outgrowth and arborization,[42–44] can nevertheless predict the L1-innervating skeletal mass of neurons belonging to this homogeneous population ($R^2 = 0.40$, $p < 0.001$, non-parametric shuffle test). Since the somata of this population are distributed across L2/3, L4, and L5 (Figure 4b), one potential explanation for this result is that gene expression is correlated with the overall depth location of the cells rather than L1 innervation strength in particular (Figure 4c). Therefore, we repeat the sparse regression study after removing the piecewise linear contribution of soma depth to L1 innervation (linear fit and subtraction for only the L1 innervating subpopulation because the relationship is trivially nonexistent for the non-innervating subpopulation, Methods). We find that the expression levels of a small set of genes are still statistically significantly predictive of L1 innervation: $R^2 = 0.31$, $p < 0.001$, non-parametric shuffle test. (Repeating with a linear fit and subtraction for the whole population does not change the qualitative result: $R^2 = 0.30$, $p < 0.001$.)

Next, we obtain a comparative perspective on the L1 innervation result for the Sst Calb2 Pdlim5 subset of Martinotti cells by juxtaposing this result with that for the cells of the Lamp5 Lsp1 type. Somata of these cells are also distributed across multiple cortical layers and their axons have highly variable levels of L1 innervation. Sparse regression again succeeds in finding a small set of genes whose expression level can predict L1 innervation ($R^2 = 0.56$, $p = 0.01$, non-parametric shuffle test, Table 1). However, it fails to uncover a statistically significant gene set after removing the piecewise linear contribution of soma depth: $R^2 = 0.06$, $p = 0.26$, non-parametric shuffle test. (Repeating with a linear fit and subtraction for the whole population does not change the qualitative result: $R^2 = 0.08$, $p = 0.39$.) That is, in contrast to the Sst Calb2 Pdlim5 cells, soma depth almost completely explains the variability in L1 innervation for cells in the Lamp5 Lsp1 population (Figure 4e,f).

Lastly, we consider the possibility that the cells whose axons do not reach L1 are simply irrelevant for this study and bias the statistics. (Axons of 3 out of 52 cells in the Sst Calb2 Pdlim5 population, and 7 out of 22 cells in the Lamp5 Lsp1 population do not reach L1.) We repeat the above comparison after removing the cells whose axons don't reach L1 altogether from this study. Sparse regression still uncovers a statistically significant relationship between L1 innervation strength and a set of genes for the Sst Calb2 Pdlim5 population after removing the linear contribution of soma depth ($R^2 = 0.28$, $p < 0.001$). In contrast, it again fails to find a statistically significant relationship for the Lamp5 Lsp1 population ($R^2 = 0.05$, $p = 0.39$).

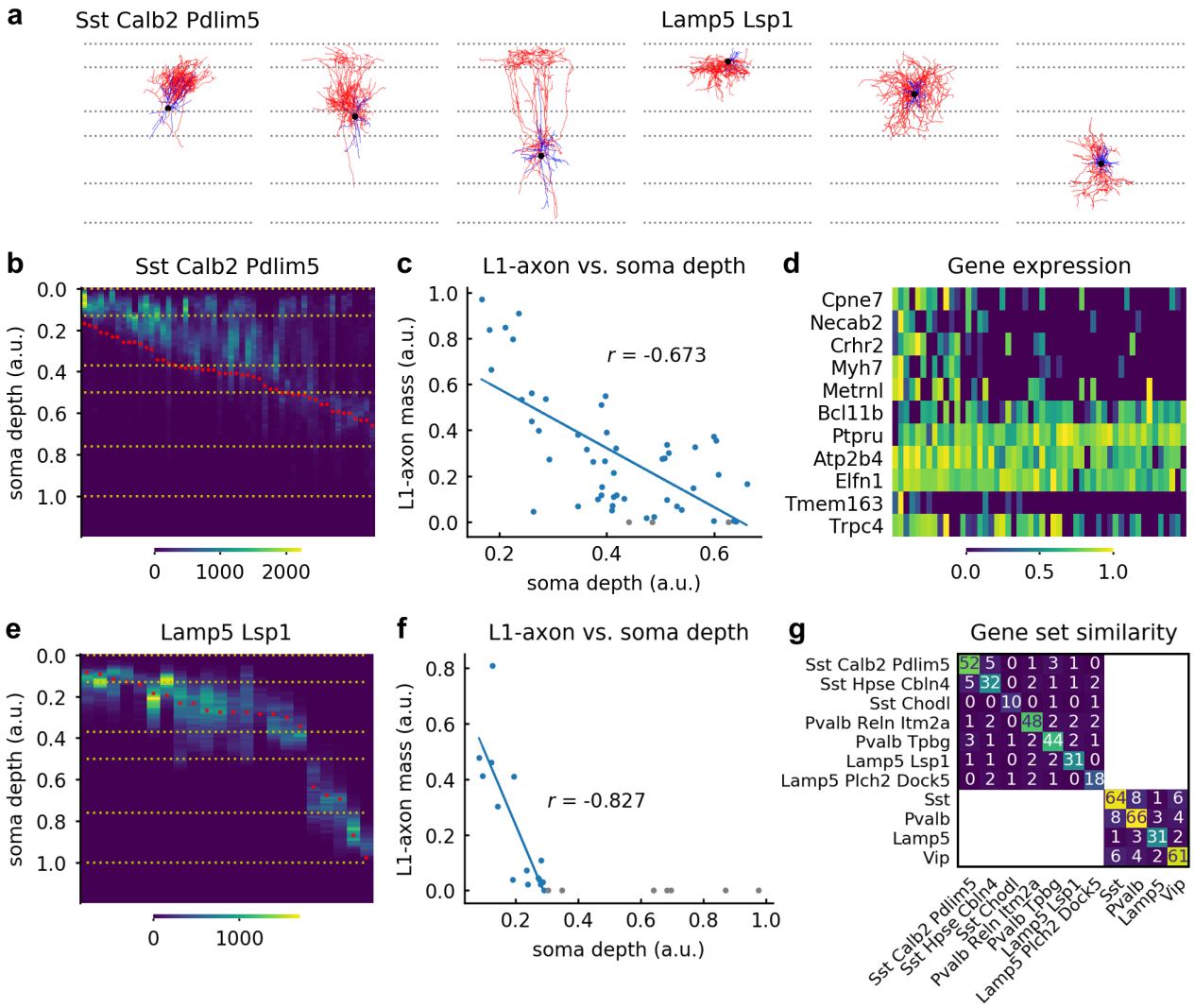

Figure 4: **L1 axonal innervation correlates with expression of subset of genes in Martinotti cells.** Example neurons of Sst Calb2 Pdlim5 and Lamp5 Lsp1 t-types (**a**). 1D axonal arbor density for the 52 cells in the Sst Calb2 Pdlim5 t-type (**b**) and the 22 cells in the Lamp5 Lsp1 t-type (**e**). (Yellow horizontal dashed lines and red dots indicate cortical layer boundaries and soma depth, respectively). Normalized L1-axon skeletal mass vs. soma depth for the Sst Calb2 Pdlim5 cells (**c**) and the Lamp5 Lsp1 cells (**f**). Lines fitted to cells with nonzero L1 innervation. Cells whose axons don't reach L1 are shown in gray. **d**, Gene expression vs. L1-axon skeletal mass for the genes selected by the sparse regression analysis. (L1-axon mass decreases from left to right.) **g**, Similarity matrix for the sets of laminae-predicting genes within transcriptomic types and subclasses. (See Table S3.) Each entry denotes the number of genes in the intersection between the corresponding row and column.

To summarize, while axons of Lamp5 Lsp1 cells appear to shift along the laminar axis according to their soma location within the cortical depth, soma location does not seem to dictate the axonal L1 innervation of Sst Calb2 Pdlim5 neurons, whose strength can nevertheless be predicted by gene expression. For both of these t-types, the automated reconstruction pipeline increased the sample size by more than 60% (Sst Calb2 Pdlim5: 63%, Lamp5 Lsp1: 69%), empowering the statistical analysis pursued here. Similarly, the sample counts for the t-types studied in Table 1 increased between 48% and 138%. (The increase over the whole dataset is 50%, from 543 to 813 cells.) Since t-types correspond to leaf nodes of the cell type hierarchy, their sample sizes are much smaller than the subclass-level counts. Therefore, automated reconstruction can be beneficial both by capturing more of the biological variability in single cell morphologies of populations at the finest level of transcriptomically defined taxonomies and by enabling cross-validation schemes similar to the ones pursued here.

Discussion

While classification of neuronal cell types is increasingly based on single cell and nucleus genomic technologies, characterization of neuron morphology – a classical approach – captures an aspect of neuronal identity that is stable over long time scales, is intimately related to connectivity and function, and can now be connected with genomic attributes through the use of simultaneous profiling techniques such as Patch-seq. Nevertheless, light microscopy based methods of neuronal reconstruction often inadequately reproduce the determinant attributes of morphological signature, especially in high-throughput settings. Here, we have presented an end-to-end automated neuronal morphology reconstruction pipeline for brightfield microscopy, whose simple setup supports flexible, single or multimodal, characterization protocols. We have also proposed an arbor density representation as a descriptor of cortical neuronal anatomy that is robust against noise in high-throughput imaging scenarios as well as mistakes of automated reconstruction.

Through the use of sparsity arguments and statistical testing, we demonstrated that this pipeline can help reveal relationships between gene expression and neuronal anatomy, where a large number of anatomical reconstructions enables accurate inference in the presence of a large gene set. As an application, we studied the correlation between gene expression and laminar innervation on a Patch-seq dataset of cortical neurons[5] and showed that the gene correlates of different innervation patterns have little overlap across transcriptomically

defined subpopulations. While the same program can potentially also address the relationship between morphological and electrophysiological properties of neurons, the accuracy of automated reconstructions should further improve for use in detailed compartmental models.[45]

Finally, we focused on axonal innervation of L1 by a transcriptomically defined subpopulation of Somatostatin-expressing Martinotti cells. We found that the innervation strength is relatively weakly correlated with soma depth for this cell type, but not all types. Moreover, a subset of genes can predict the remaining variability in the innervation strength after the effect of soma depth is removed, suggesting a control mechanism beyond simple shifting of the morphology within the cortical depth for this cell type. Considering that neurons in this population are thought to gate top-down input to cortical pyramidal neurons,[25,26] this result suggests tuning of innervation strength in a continuum within the discrete laminar organization of the mouse cortex,[46–49] potentially to improve task performance of the underlying neuronal network.

From a segmentation perspective, we believe our work represents a significant step forward as the first study to produce hundreds of automatically reconstructed morphologies obtained from the brightfield microscope (Fig. S1-S11). As demonstrated in the main text, these cortical neuron morphologies are statistically indistinguishable from their manually generated counterparts in certain aspects (e.g., cell type identification), but not in many others (e.g., arbor topology). Indeed, much further improvement is needed to achieve complete and accurate tracing of neurons, and the human should remain in the loop, at least for quality control. Nevertheless, advances in computer vision algorithms and computing infrastructure that can support complicated models and large datasets suggest that qualitative improvements may be within reach in the next few years.

References

[1] Marcela Lipovsek, Cedric Bardy, Cathryn R Cadwell, Kristen Hadley, Dmitry Kobak, and Shreejoy J Tripathy. Patch-seq: Past, present, and future. *Journal of Neuroscience*, 41(5):937–946, 2021.

[2] János Fuzik, Amit Zeisel, Zoltán Máté, Daniela Calvigioni, Yuchio Yanagawa, Gábor Szabó, Sten Linnarsson, and Tibor Harkany. Integration of electrophysiological recordings with single-cell rna-seq data identifies neuronal subtypes. *Nature biotechnology*, 34(2):175, 2016.

[3] Cathryn R Cadwell, Athanasia Palasantza, Xiaolong Jiang, Philipp Berens, Qiaolin Deng, Marlene Yilmaz, Jacob Reimer, Shan Shen, Matthias Bethge, Kimberley F Tolias, et al. Electrophysiological, transcriptomic and morphologic profiling of single neurons using patch-seq. *Nature biotechnology*, 34(2):199–203, 2016.

[4] Csaba Földy, Spyros Darmanis, Jason Aoto, Robert C Malenka, Stephen R Quake, and Thomas C Südhof. Single-cell rnaseq reveals cell adhesion molecule profiles in electrophysiologically defined neurons. *Proceedings of the National Academy of Sciences*, 113(35):E5222–E5231, 2016.

[5] Nathan W Gouwens, Staci A Sorensen, Fahimeh Baftizadeh, Agata Budzillo, Brian R Lee, Tim Jarsky, Lauren Alfiler, Katherine Baker, Eliza Barkan, Kyla Berry, et al. Integrated morphoelectric and transcriptomic classification of cortical gabaergic cells. *Cell*, 183(4):935–953, 2020.

[6] Federico Scala, Dmitry Kobak, Matteo Bernabucci, Yves Bernaerts, Cathryn René Cadwell, Jesus Ramon Castro, Leonard Hartmanis, Xiaolong Jiang, Sophie Laturnus, Elanine Miranda, et al. Phenotypic variation of transcriptomic cell types in mouse motor cortex. *Nature*, 598(7879):144–150, 2021.

[7] Brian R Lee, Agata Budzillo, Kristen Hadley, Jeremy A Miller, Tim Jarsky, Katherine Baker, DiJon Hill, Lisa Kim, Rusty Mann, Lindsay Ng, et al. Scaled, high fidelity electrophysiological, morphological, and transcriptomic cell characterization. *Elife*, 10:e65482, 2021.

[8] Lin Que, David Lukacsovich, Wenshu Luo, and Csaba Földy. Transcriptional and morphological profiling of parvalbumin interneuron subpopulations in the mouse hippocampus. *Nature Communications*, 12(1):1–15, 2021.

[9] Vivien Marx. Patch-seq takes neuroscience to a multimodal place. *Nature Methods*, 19(11):1340–1344, 2022.

[10] Duncan E Donohue and Giorgio A Ascoli. Automated reconstruction of neuronal morphology: an overview. *Brain research reviews*, 67(1-2):94–102, 2011.

[11] Engin Türetken, Fethallah Benmansour, and Pascal Fua. Automated reconstruction of tree structures using path classifiers and mixed integer programming. In *2012 IEEE conference on computer vision and pattern recognition*, pages 566–573. IEEE, 2012.

[12] Uygar Sümbül, Sen Song, Kyle McCulloch, Michael Becker, Bin Lin, Joshua R Sanes, Richard H Masland, and H Sebastian Seung. A genetic and computational approach to structurally classify neuronal types. *Nature communications*, 5(1):1–12, 2014.

[13] Rohan Gala, Julio Chapeton, Jayant Jitesh, Chintan Bhavsar, and Armen Stepanyants. Active learning of neuron morphology for accurate automated tracing of neurites. *Frontiers in neuroanatomy*, 8:37, 2014.

[14] Linqing Feng, Ting Zhao, and Jinhyun Kim. neutube 1.0: a new design for efficient neuron reconstruction software based on the swc format. *eneuro*, 2(1), 2015.

[15] Tingwei Quan, Hang Zhou, Jing Li, Shiwei Li, Anan Li, Yuxin Li, Xiaohua Lv, Qingming Luo, Hui Gong, and Shaoqun Zeng. Neurogps-tree: automatic reconstruction of large-scale neuronal populations with dense neurites. *Nature methods*, 13(1):51–54, 2016.

[16] James Gornet, Kannan Umadevi Venkataraju, Arun Narasimhan, Nicholas Turner, Kisuk Lee, H Sebastian Seung, Pavel Osten, and Uygar Sümbül. Reconstructing neuronal anatomy from whole-brain images. In *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, pages 218–222. IEEE, 2019.

[17] Johan Winnubst, Erhan Bas, Tiago A Ferreira, Zhuhao Wu, Michael N Economo, Patrick Edson, Ben J Arthur, Christopher Bruns, Konrad Rokicki, David Schauder, et al. Reconstruction of 1,000 projection neurons reveals new cell types and organization of long-range connectivity in the mouse brain. *Cell*, 179(1):268–281, 2019.

[18] Hanchuan Peng, Peng Xie, Lijuan Liu, Xiuli Kuang, Yimin Wang, Lei Qu, Hui Gong, Shengdian Jiang, Anan Li, Zongcai Ruan, et al. Morphological diversity of single neurons in molecularly defined cell types. *Nature*, 598(7879):174–181, 2021.

[19] Wenyun He, Thomas A Hamilton, Andrew R Cohen, Timothy J Holmes, Christopher Pace, Donald H Szarowski, James N Turner, and Badrinath Roysam. Automated three-dimensional tracing of neurons in confocal and brightfield images. *Microscopy and microanalysis*, 9(4):296–310, 2003.

[20] Hanchuan Peng, Zhi Zhou, Erik Meijering, Ting Zhao, Giorgio A Ascoli, and Michael Hawrylycz. Automatic tracing of ultra-volumes of neuronal images. *Nature methods*, 14(4):332–333, 2017.

[21] Zhi Zhou, Hsien-Chi Kuo, Hanchuan Peng, and Fuhui Long. Deepneuron: an open deep learning toolbox for neuron tracing. *Brain informatics*, 5(2):1–9, 2018.

[22] Dezhe Z Jin, Ting Zhao, David L Hunt, Rachel P Tillage, Ching-Lung Hsu, and Nelson Spruston. Shutu: open-source software for efficient and accurate reconstruction of dendritic morphology. *Frontiers in Neuroinformatics*, 13:68, 2019.

[23] Bosiljka Tasic, Zizhen Yao, Lucas T Graybuck, Kimberly A Smith, Thuc Nghi Nguyen, Darren Bertagnolli, Jeff Goldy, Emma Garren, Michael N Economo, Sarada Viswanathan, et al. Shared and distinct transcriptomic cell types across neocortical areas. *Nature*, 563(7729):72–78, 2018.

[24] Henry Markram, Eilif Muller, Srikanth Ramaswamy, Michael W Reimann, Marwan Abdellah, Carlos Aguado Sanchez, Anastasia Ailamaki, Lidia Alonso-Nanclares, Nicolas Antille, Selim Arsever, et al. Reconstruction and simulation of neocortical microcircuitry. *Cell*, 163(2):456–492, 2015.

[25] Masanori Murayama, Enrique Pérez-Garci, Thomas Nevian, Tobias Bock, Walter Senn, and Matthew E Larkum. Dendritic encoding of sensory stimuli controlled by deep cortical interneurons. *Nature*, 457(7233):1137, 2009.

[26] Benjamin Schuman, Shlomo Dellal, Alvar Prönneke, Robert Machold, and Bernardo Rudy. Neocortical layer 1: An elegant solution to top-down and bottom-up integration. *Annual Review of Neuroscience*, 44, 2021.

[27] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[28] Özgün Çiçek, Ahmed Abdulkadir, Soeren S Lienkamp, Thomas Brox, and Olaf Ronneberger. 3d u-net: learning dense volumetric segmentation from sparse annotation. In *International conference on medical image computing and computer-assisted intervention*, pages 424–432. Springer, 2016.

[29] Jan Funke, Fabian Tschopp, William Grisaitis, Arlo Sheridan, Chandan Singh, Stephan Saalfeld, and Srinivas C Turaga. Large scale image segmentation with structured loss based deep learning for connectome reconstruction. *IEEE transactions on pattern analysis and machine intelligence*, 41(7):1669–1680, 2018.

[30] Giorgio A Ascoli. Mobilizing the base of neuroscience data: the case of neuronal morphologies. *Nature Reviews Neuroscience*, 7(4):318–324, 2006.

[31] Staci A. Sorensen, Nathan W. Gouwens, Yun Wang, Matt Mallory, Agata Budzillo, Rachel Dalley, Brian R. Lee, Olga Gliko, Hsienchi Kuo, Xiuli Kuang, Rusty Mann, Leila Ahmadinia, Lauren Alfiler, Fahimeh Baftizadeh, Katherine S. Baker, Sarah Bannick, Darren Bertagnolli, Kris Bickley, Phil Bohn, Jasmine Bomben, Gabriella Boyer, Krissy Brouner, Alex Cahoon, Natalie Chen, Chao Chen, Kai Chen, Maggie Chvilicek, Forrest Collman, Tanya L. Daigle, Tim Dawes, Rebecca de Frates, Nick Dee, Maxwell DePartee, Tom Egdorf, Laila El-Hifnawi, Luke Esposito, Colin Farrell, Rohan Gala, Clare Gamlin, Amanda Gary, Jeff Goldy, Hong Gu, Kristen Hadley, Mike Hawrylycz, Alex Henry, Dijon Hill, Karla E. Hirokawa, Zili Huang, Katelyn Johnson, Zoe Juneau, Sara Kebede, Lisa Kim, Lauren Kruse, Changkyu Lee, Arielle Leon, Phil Lesnar, Anan Li, Yaoyao Li, Elizabeth Liang, Katie Link, Michelle Maxwell, Medea McGraw, Delissa A. McMillen, Alice Mukora, Lindsay Ng, Aaron Oldre, Daniel Park, Christina Alice Pom, Zoran Popovich, Lydia Potekhina, Ram Rajanbabu, Shea Ransford, Melissa Reding, Augustin Ruiz, David Sandman, Josh Sevigny, Lyudmila Shulga, La'Akea Siverts, Cliff Slaughterbeck, Kimberly A. Smith, Michelle Stoecklin, Josef Sulc, Susan M. Sunkin, Michael Tieu, Jonathan Ting, Jessica Trinh, Sara Vargas, David Vumbaco, Miranda Walker, Micheal Wang, Adrian Wanner, Jack Waters, Grace Williams, Julia A. Wilson, Wei Xiong, Ed S. Lein, Jim Berg, Brian E. Kalmbach, Shenqin Yao, Hui Gong, Qingming Luo, Lydia Ng, Uygar Sümbül, Tim Jarsky, Zizhen Yao, Bosiljka Tasic, and Hongkui Zeng. Connecting single neuron transcriptomes to the projectome in mouse visual cortex. *bioRxiv*, 2023.

[32] Armen Stepanyants, Patrick R Hof, and Dmitri B Chklovskii. Geometry and structural plasticity of synaptic connectivity. *Neuron*, 34(2):275–288, 2002.

[33] Hermann Cuntz, Friedrich Forstner, Alexander Borst, and Michael Häusser. One rule to grow them all: a general theory of neuronal branching and its practical application. *PLoS Comput Biol*, 6(8):e1000877, 2010.

[34] Marcel Oberlaender, Christiaan PJ de Kock, Randy M Bruno, Alejandro Ramirez, Hanno S Meyer, Vincent J Dercksen, Moritz Helmstaedter, and Bert Sakmann. Cell type–specific three-dimensional structure of thalamocortical circuits in a column of rat vibrissal cortex. *Cerebral cortex*, 22(10):2375–2391, 2012.

[35] Uygar Sümbül, Aleksandar Zlateski, Ashwin Vishwanathan, Richard H Masland, and H Sebastian Seung. Automated computation of arbor densities: a step toward identifying neuronal cell types. *Frontiers in neuroanatomy*, 8:139, 2014.

[36] NW Gouwens, SA Sorensen, J Berg, C Lee, T Jarsky, J Ting, SM Sunkin, D Feng, CA Anastassiou, E Barkan, K Bickley, N Blesie, T Braun, K Brouner, A Budzillo, S Caldejon, T Casper, D Castelli, P Chong, K Crichton, C Cuhaciyan, TL Daigle, R Dalley, N Dee, T Desta, SL Ding, S Dingman, A Doperalski, N Dotson, T Egdorf, M Fisher, Frates RA de, E Garren, M Garwood, A Gary, N Gaudreault, K Godfrey, M Gorham, H Gu, C Habel, K Hadley, J Harrington, JA Harris, A Henry, D Hill, S Josephsen, S Kebede, L Kim, M Kroll, B Lee, T Lemon, KE Link, X Liu, B Long, R Mann, M McGraw, S Mihalas, A Mukora, GJ Murphy, L Ng, K Ngo, TN Nguyen, PR Nicovich, A Oldre, D Park, S Parry, J Perkins, L Potekhina, D Reid, M Robertson, D Sandman, M Schroedter, C Slaughterbeck, G Soler-Llavina, J Sulc, A Szafer, B Tasic, N Taskin, C Teeter, N Thatra, H Tung, W Wakeman, G Williams, R Young, Z Zhou, C Farrell, H Peng, MJ Hawrylycz, E Lein, L Ng, A Arkhipov, A Bernard, JW Phillips, H Zeng, and C Koch. Classification of electrophysiological and morphological neuron types in the mouse visual cortex. *Nat Neurosci*, 22:1182–1195, Jul 2019.

[37] Daniel J Wilson. The harmonic mean p-value for combining dependent tests. *Proceedings of the National Academy of Sciences*, 116(4):1195–1200, 2019.

[38] Sophie Laturnus, Dmitry Kobak, and Philipp Berens. A systematic evaluation of interneuron morphology representations for cell type discrimination. *Neuroinformatics*, 18(4):591–609, 2020.

[39] Ming Yuan and Yi Lin. Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 68(1):49–67, 2006.

[40] Dmitry Kobak, Yves Bernaerts, Marissa A Weis, Federico Scala, Andreas S Tolias, and Philipp Berens. Sparse reduced-rank regression for exploratory visualisation of paired multivariate data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 2021.

[41] Emilia Favuzzi, Rubén Deogracias, André Marques-Smith, Patricia Maeso, Julie Jezequel, David Exposito-Alonso, Maddalena Balia, Tim Kroon, Antonio J Hinojosa, Elisa F. Maraver, et al. Distinct molecular programs regulate synapse specificity in cortical inhibitory circuits. *Science*, 363(6425):413–417, 2019.

[42] Akihiro Inoue and Joshua R Sanes. Lamina-specific connectivity in the brain: regulation by n-cadherin, neurotrophins, and glycoconjugates. *Science*, 276(5317):1428–1431, 1997.

[43] Alex L Kolodkin and Marc Tessier-Lavigne. Mechanisms and molecules of neuronal wiring: a primer. *Cold Spring Harbor perspectives in biology*, 3(6):a001727, 2011.

[44] Yu-Chi Sun, Xiaoyin Chen, Stephan Fischer, Shaina Lu, Huiqing Zhan, Jesse Gillis, and Anthony M Zador. Integrating barcoded neuroanatomy with spatial transcriptional profiling enables identification of gene correlates of projections. *Nature Neuroscience*, 2021.

[45] Wulfram Gerstner, Werner M Kistler, Richard Naud, and Liam Paninski. *Neuronal dynamics: From single neurons to networks and models of cognition*. Cambridge University Press, 2014.

[46] Kenneth D Harris, Hannah Hochgerner, Nathan G Skene, Lorenza Magno, Linda Katona, Carolina Bengtsson Gonzales, Peter Somogyi, Nicoletta Kessaris, Sten Linnarsson, and Jens Hjerling-Leffler. Classes and continua of hippocampal ca1 inhibitory neurons revealed by single-cell transcriptomics. *PLoS biology*, 16(6):e2006387, 2018.

[47] Geoffrey Stanley, Ozgun Gokce, Robert C Malenka, Thomas C Südhof, and Stephen R Quake. Continuous and discrete neuron types of the adult murine striatum. *Neuron*, 105(4):688–699, 2020.

[48] Timothy P O'Leary, Kaitlin E Sullivan, Lihua Wang, Jody Clements, Andrew L Lemire, and Mark S Cembrowski. Extensive and spatially variable within-cell-type heterogeneity across the basolateral amygdala. *Elife*, 9:e59003, 2020.

[49] Yeganeh Marghi, Rohan Gala, Fahimeh Baftizadeh, and Uygar Sumbul. Joint inference of discrete cell types and continuous type-specific variability in single-cell datasets with mmidas. *bioRxiv*, pages 2023–10, 2023.

[50] Leonard Kuan, Yang Li, Chris Lau, David Feng, Amy Bernard, Susan M Sunkin, Hongkui Zeng, Chinh Dang, Michael Hawrylycz, and Lydia Ng. Neuroinformatics of the allen mouse brain connectivity atlas. *Methods*, 73:4–17, 2015.

[51] Quanxin Wang, Song-Lin Ding, Yang Li, Josh Royall, David Feng, Phil Lesnar, Nile Graddis, Maitham Naeemi, Benjamin Facer, Anh Ho, et al. The allen mouse brain common coordinate framework: a 3d reference atlas. *Cell*, 181(4):936–953, 2020.

[52] James A Sethian. Fast marching methods. *SIAM review*, 41(2):199–235, 1999.

[53] Gilles Bertrand and Grégoire Malandain. A new characterization of three-dimensional simple points. *Pattern Recognition Letters*, 15(2):169–175, 1994.

[54] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[55] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[56] Stephen J Smith, Uygar Sümbül, Lucas T Graybuck, Forrest Collman, Sharmishtaa Seshamani, Rohan Gala, Olga Gliko, Leila Elabbady, Jeremy A Miller, Trygve E Bakken, et al. Single-cell transcriptomic evidence for dense intracortical neuropeptide networks. *Elife*, 8:e47889, 2019.

[57] Rohan Gala, Agata Budzillo, Fahimeh Baftizadeh, Jeremy Miller, Nathan Gouwens, Anton Arkhipov, Gabe Murphy, Bosiljka Tasic, Hongkui Zeng, Michael Hawrylycz, et al. Consistent cross-modal identification of cortical neurons with coupled autoencoders. *Nature Computational Science*, 1(2):120–127, 2021.

[58] Trygve E Bakken, Rebecca D Hodge, Jeremy A Miller, Zizhen Yao, Thuc Nghi Nguyen, Brian Aevermann, Eliza Barkan, Darren Bertagnolli, Tamara Casper, Nick Dee, et al. Single-nucleus and single-cell transcriptomes compared in matched cortical cell types. *PloS one*, 13(12):e0209648, 2018.

[59] Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288, 1996.

[60] Yoav Benjamini and Daniel Yekutieli. The control of the false discovery rate in multiple testing under dependency. *Annals of statistics*, pages 1165–1188, 2001.

[61] Skipper Seabold and Josef Perktold. statsmodels: Econometric and statistical modeling with python. In *9th Python in Science Conference*, 2010.

Methods

Dataset

The dataset profiling local morphology and transcriptome of GABAergic mouse visual cortical neurons was generated as part of the Patch-seq recordings described in Ref.[5] This dataset includes 2,341 cells with transcriptomic profiling and high resolution image stacks, where the brain sections were imaged on an upright AxioImager Z2 microscope (Zeiss) equipped with an Axiocam506 monochrome camera. Tiled image stacks of individual cells were acquired with a

63× objective lens (Zeiss Plan-Apochromat 63×/1.4 NA or Zeiss LD LCI Plan-Apochromat 63×/1.2 NA) at an interval of 0.28 μm or 0.44 μm along the Z axis. Individual cells were manually placed in the appropriate cortical region and layer within the Allen Mouse Common Coordinate Framework (CCF)[50,51] by matching the 20X image of the slice with a "virtual" slice at an appropriate location and orientation within the CCF. 1,259 cells were removed from the dataset either because they were mapped to nearby regions (instead of visual cortex) or because their images had incomplete axons. 543 of the remaining 1,082 bright-field image stacks of biocytin-filled neurons were reconstructed both manually and automatically, and this set is used for training and testing of the t-type classification algorithms and for error/R^2-value quantification. The remaining 539 cells were reconstructed only automatically. To ensure the quality of scientific results presented in Table 1 and Figure 4, we excluded 118 images with multiple neurons in the field of view and 151 images that failed at different stages of the pipeline (missing pia/white matter annotations, annotation-related failed upright transformation, reconstruction failing visual inspection) from those analyses. Finally, 16 cells in the manually and automatically reconstructed population and 20 cells in the automatically-only reconstructed population were not used for analyses involving t-types because these cells were deemed to not have "highly consistent" t-type mappings in Ref.[5]

Volumetric training data generation from skeletonized morphologies

Segmentation of neuronal morphologies from 3D image stack requires voxel-wise labels while manual reconstructions specified by traces only provide a set of vertices and edges corresponding to the layout of the underlying morphology. We developed a topology-preserving fast marching algorithm to generate volumetric ground truth using raw image stacks and manual traces by adapting a fast-marching based segmentation algorithm[13,52] initialized with trace vertices to segment image voxels. This segmentation should be consistent with the layout of morphology traces, without introducing topological errors (e.g., splits, merges, holes). We ensured this by incorporating simple point methods in digital topology[53] into the segmentation algorithm. (i.e., the proposal generated by fast marching is disallowed if the proposed voxel value flip changes the underlying topology.) We noticed that the soma can be incompletely filled by the fast marching algorithm. Therefore, we treated the soma region separately and used a sequence of erosion and dilation operations followed by manual thresholding to achieve complete labeling. Each voxel was labeled as axon, dendrite, soma, background.

Neural network architecture and training

We used a 3D U-Net convolutional neural network[28] to perform multi-class segmentation (i.e. each voxel is assigned a probability of belonging to classes specified in the label set). We trained two separate models using raw images and volumetric ground truth; one with 51 inhibitory neurons, and another with 75 excitatory neurons from mouse cortex. The U-Net architecture consists of a contracting path to capture context and a symmetric expanding path that enables precise localization and has been shown to achieve state-of-the-art performance in many segmentation tasks.[27,28] Building on previous work,[16] we developed an efficient Pytorch implementation that runs on graphical processing units (GPUs). To address GPU memory constraints, during each training epoch the training set was divided randomly into subsets of 3 stacks. Training proceeded sequentially using all subsets in an epoch, and data loading time did not exceed 10% of the total training time. We trained the model on batches of 3D patches ($128 \times 128 \times 32$ px^3, XYZ), which were randomly sampled from the subset of 3 stacks loaded into memory. All the models were trained using the Adam optimizer[54] with a learning rate of 0.1. Training with a GeForce GTX 1080 GPU for ~ 50 epochs took ~ 3 weeks. Since the neuron occupies a small fraction of the image volume, we chose patches that contained at least one voxel belonging to the neuron. To add salient, negative examples to the training set, we also included a number of patches with bright backgrounds produced by staining artifacts and pial surface. To enable the model to generalize from relatively small number of training examples and improve the segmentation accuracy, we augmented training data by 90° rotations and vertical/horizontal flips in the image (XY) plane.

End-to-end neuron reconstruction pipeline

An end-to-end automated pipeline combined segmentation of raw image stacks into soma, axon, and dendrite structures with post-processing routines to produce a swc file. Segmentation with trained models using a single high-end GPU takes ∼5.8 min per Gvoxel, or ∼186 min for average 32 Gvoxel image stack. Our pipeline had access to 16 GPU (NVIDIA Titan X) cards. Even though the model trained using inhibitory neurons generalized well on all types of neurons, we found that the model trained using excitatory neurons improved the axon/dendrite assignment accuracy on excitatory neurons. Segmentation was post-processed by thresholding the background, followed by connected component analysis to remove short segments, skeletonization, and converting to a digital reconstruction in the swc format.

Axon/dendrite relabeling

Since the initial segmentation by the UNet-based neural network assigns every foreground voxel one of {soma, axon, dendrite} based on local context defined by the patch size, it is prone to occasional errors, particularly in distinguishing between axon and dendrite. These errors propagate to the labeling of nodes in the tree representation. In order to improve this initial node labeling, we developed an error correction approach which utilizes the initial segmentation to make a decision based on a larger spatial context.

We trained a secondary neural network model to predict the labels based on features calculated using the raw image stack and the initial trace. First, for each connected component in the skeleton of the initial segmentation, we identified the longest path from the node closest to soma and calculated features on the node set defining that path: (i) The first 6 features are 1D arrays of image brightness values calculated for every node in the set at different spatial scales using spherical kernels of varying radii (1, 2, 4, 8, 16, 32). (ii) A second set of 6 features are 1D arrays of neurite tortuosity values calculated for every node in the set as ratios of path and Euclidean distances between each node and its n-th neighbor away from the soma (n = 4, 8, 16, 32, 64, 128). (iii) Two additional features are 1D arrays of node type of every node in the set and a single number representing distance of the closest node to the soma. For each 1D array, we used the first 2048 nodes and zero-padded shorter arrays to have a uniform array size.

The neural network architecture consists of three arms, two arms have two convolutional layers with 4 and 8 7x3 filters followed by 4x1 max pooling, a fully connected layer and a dropout layer each. These arms process feature sets (i) and (ii) above by stacking the sets of 6 1D arrays along a second dimension. The third arm processes feature set (iii) and has two convolutional layers with 4 and 8 7x1 filters followed by 4x1 max pooling, a fully connected layer and a dropout layer. The outputs of these three arms are concatenated with the 'distance to soma feature'. Finally, the concatenated hidden feature map is processed by two fully connected layers to produce a single scalar softmax output indicating the inferred label type. The network model was trained using examples from the training dataset of the semantic segmentation model, and is used to relabel the neuron traces during postprocessing. Averaging predicted label type over the 9 longest paths improved relabeling accuracy for connected components longer than 2048 nodes.

Connecting disconnected segments

Due to a combination of staining artifacts and limitations of brightfield imaging, the initial reconstruction is often characterized by multiple disconnected subtrees. We introduced artificial breaks to manual traces to train a random forest classifier to predict whether nearby pairs of connected components should be merged. First, we find all pairs of segment (subtree) end points that are located within a certain distance. Next, for each pair of end points, we calculate the Euclidean distance and four collinearity values which measure the segments' orientation relative to each other. Specifically, for each end point, we calculate two vectors representing segment end orientation at two different spatial scales (i.e., the orientation of the branch terminating at that end point). The collinearity values are the dot products of each of these vectors with the vector between end points of the pair. Finally, for each end point in the pair we also calculate the above features for the closest four other end points. As a result, for every pair of end nodes we have a total of 45 features. Only segments of the same type, axon or dendrite, are considered for merging.

Additional postprocessing

We passed the reconstructions through a series of additional post-processing for extraneous cell/artefact removal, down sampling, node sorting and pruning. We used quality control by a human to check for the presence of disconnected branches of extraneous cells that were not removed during postprocessing. We excluded samples that did not pass this quality control if they did not have a manual trace. For samples with a manual trace, we used a neighborhood of the manual trace to simulate manual removal of extraneous cells by masking with that neighborhood. We excluded these samples from reconstruction accuracy quantification and used them only for cell type classification. For excitatory neurons (Fig. 2a, and Ref.[31]), we trained a random forest classifier to identify apical dendrite segments in excitatory auto-trace reconstructions. The classifier was trained on geometric features that distinguish apical dendrite segments from the basal dendrite (e.g. upright distance from soma). This classifier achieved a mean accuracy of 85% percent across 10-fold cross validation.

Morphometric feature calculation

Reconstructions were transformed to an upright position that is perpendicular with respect to pia and white matter surfaces. Morphometric features as described in[36] were calculated using the skeleton_keys python package. Following Ref.,[36] z-dimension features were not included.

Arbor density generation of axons and dendrites

We represented axonal and dendritic morphologies as 2D density maps registered to a common local coordinate axis using the pia/white matter boundaries. First, we applied upright transform to the reconstructed neuron followed by the correction of z-shrinkage and the variable tilt angle of the acute slice.[36] Next, adapting our previous work[12,35] we conformally mapped pia and white matter boundaries to flat surfaces, calculated a nonlinear transformation on the whole tissue by a least-square fit to pia/white matter mappings, and applied this transformation to the morphology trace. Finally, we used the registered trace to generate a 2D density representation. The polar axes representing the cortical depth and the lateral distance from the soma (Figure 2c) make this representation invariant to rotations around the laminar axis. We calculated these maps separately for axons and dendrites. We downsampled these maps to 120px $\times$ 4px images with a pixel size of $8\mu m \times 125\mu m$ to be robust to minor changes in morphology. We normalized the intensity by the lateral area corresponding to each pixel so that each pixel value represents local arbor density.

Assessing neuron reconstruction accuracy

To assess the quality of automated neuron reconstructions, we used manual reconstructions as the ground truth. We quantified the correspondence of trace nodes, as described in the main text, to evaluate the accuracy of the trace layout. We calculated precision, recall and f1-score metrics at three distances (2, 5, and 10 μm), and reported mean and standard deviation values for the test set of 347 neurons (all samples that have manual traces excluding the ones used for training models or required masking). In addition, we evaluated the accuracy of neuron morphology representations, morphometric features and ADRs, for the same set. We reported the coefficient of determination R^2 for each morphometric feature. We calculated the average root-mean-squared error per t-type between normalized axon/dendrite ADRs derived from

automated and manual reconstructions.

Supervised classification

Supervised classification using morphometric features was performed by training a random forest classifier implemented in the scikit-learn Python package[55] using 5-fold cross-validation. This was repeated 20 times. Supervised classification using ADRs was done by training a feed-forward neural network classifier using our Pytorch implementation. The network architecture consists of a convolutional layer with 7x3x1 filters, a 4x1x1 max pooling layer, a convolutional layer with 7x3x1 filters, a 3x1x1 max pooling layer, a layer that concatenates the hidden features with the soma depth value, and a fully connected layer with the number of units corresponding to the number of classes. Each convolutional layer uses the rectified linear function as the non-linear transformation.

Since the depth locations of neurons vary within the cortex, we introduced a novel type of data augmentation based on simulation of cell type-dependent neuronal shift along the laminar axis. Namely, for each t-type we calculated the range of soma depth variations, and applied a random shift within that range to the input soma depth value, as well as the corresponding shift to the ADR intensity in the laminar direction. This cell type-dependent random shift of the ADR and the soma depth together with a modulation of the intensity values of the ADR improved the accuracy of classification.

The networks were trained using the cross-entropy loss function and the Adam optimizer with a learning rate of 0.001. Training using 10-fold cross-validation with GeForce GTX 1080 GPU for 50,000 epochs took $\sim$ 24 h. A set of 246 automatically and 501 manually reconstructed cells was used for training classifiers shown in Figure 3a and b, and a set of the same 488 automatically and manually reconstructed cells were used for Figure 3c,d and Figure S12. Both sets included only cells from t-types with at least 5 cells. Confusion matrices, mean and standard deviation of accuracy across cross-validation folds were reported.

We performed Boschloo's exact test on 2x2 contingency tables where columns/rows store total numbers of correct and incorrect predictions for two given classifiers. When comparing ADR-based to morphometric feature-based classifiers, we calculated the contingency table for each of the 20 repetitions used in the feature-based classifier study. We calculated the p-value of one-sided Boschloo's test to evaluate the null hypothesis of ADR-based accuracy being less

than feature-based accuracy. To aggregate the 20 p-values, we used Ref.[37] and the Python implementation at `https://github.com/benjaminpatrickevans/harmonicmeanp` to report the p-value of the asymptotically exact harmonic mean p-value test for t-type and subclass predictions.

Cophenetic agreement: To take the hierarchical organization of transcriptomically defined mouse cortical cell types[23] into account when evaluating the accuracies of the different classification tasks, we defined resolution index per cell as the scaled height of the closest common ancestor of assigned and predicted labels in the t-type hierarchical tree[56] (Figures S13 and S14). Accordingly, the resolution index for a correctly classified t-type ("leaf node" label) is 1. In the worst case, the closest ancestor for an assigned and predicted label can be the root node of the taxonomy, which corresponds to a resolution index of 0. We report the mean and s.e.m. values for this measure of cophenetic agreement between true and predicted assignments for each cell type in Figures S13 and S14.

Sparse feature selection analysis

Following Ref.,[57] a set of 1,252 genes were used for this analysis. This set was obtained by excluding genes if they satisfy any of the following criteria: (1) they are highly expressed in non-neuronal cells, (2) they have previously reported sex or mitochondrial associations, and (3) they are much more highly expressed in Patch-seq data vs. Fluorescence Activated Cell Sorting (FACS) data (or vice versa) and therefore may be associated with the experimental platform[5]. Further, we removed gene models and some other families of unannotated genes that may be difficult to interpret. We also used the β score, a published measure to evaluate the degree to which gene expression is exclusive to t-types,[58] to exclude genes expressed broadly across t-types. Gene expression values were CPM normalized, and then $\log_e(\bullet + 1)$ transformed for all the downstream analyses.

A set of 777 neurons was used for the feature selection analysis where automated reconstructions comprise $\sim 32\%$ of this set ($\sim 44\%$ of the subset of 7 t-types studied in Figure 4). Every neuron in the dataset was characterized by the expression levels of the set of 1252 genes, and their axonal and dendritic 1D arbor density representations were organized into two 120×1 vectors. For each neuron, axon/dendrite layer-specific skeletal masses normalized by total skeletal mass were calculated to quantify layer-specific innervation for axonal and dendritic morphologies, and axon/dendrite centroids were calculated to characterize laminar position of the morphology.

To select a small subset of genes that are responsible for the variability in individual anatomical features within each transcriptomic type or subclass, we solved the Lasso regression problem[59] (LassoCV command in the scikit-learn library[55]) for each anatomical feature using the cells in that transcriptomic set. We analyzed only the sets corresponding to the types and subclasses with at least 20 cells. Briefly, let K_t denote the number of cells of type or subclass t for which we have both anatomical features y_t (a $K_t \times 1$ vector) and gene expression values X_t (a $K_t \times N$ matrix). We solve

$$\text{minimize } \frac{1}{2K_t}\|y_t - X_t w_t\|_2^2 + \alpha\|w_t\|_1$$

by performing nested 5-fold cross-validation. For each cross-validation fold, we passed the training set into LassoCV which performed another splitting of the data to determine the hyperparameter α and the set of selected genes. We selected the 10 genes with maximum absolute weight values and calculated the coefficient of determination, R^2, for the test set. Finally, we selected the 10 most frequent genes across the 5 folds and calculated the mean test R^2 value. To evaluate statistical significance, we shuffled the rows of the gene expression matrix X_t 1000 times and used the same procedure to calculate mean test R^2 value for each shuffled run. We calculated the one-sided p-value as the fraction of shuffled runs with R^2 values greater than or equal to the true R^2. Finally, we performed multiple testing correction of p-values using Benjamini-Yekutieli method[60] to control the false discovery rate (multipletests command in the statsmodels library[61]). We report resulting p-values and test R^2 values for t-types in Table 1 and for t-types and subclasses in Tables S1 and S2.

Data availability

Transcriptomic and morphological data supporting the findings of this study is available online at `https://portal.brain-map.org/explore/classes/multimodal-characterization` ("Neurons in Mouse Primary Visual Cortex"). Additional dataset of automated morphological reconstructions is available at `https://github.com/ogliko/patchseq-autorecon`.

Code availability

Code pertaining to this study as well as the trained neural network model for automated segmentation are available at `https://github.com/ogliko/patchseq-autorecon` and `https:`

//github.com/rhngla/topo-preserve-fastmarching. Morphometric features are calculated using the skeleton_keys package at https://skeleton-keys.readthedocs.io.

Acknowledgements

We wish to thank the Allen Institute for Brain Science founder, Paul G Allen, for his vision, encouragement and support. We thank Michael Hawrylycz and Claire Gamlin for helpful feedback on the manuscript. This work was supported by the National Institute Of Mental Health of the National Institutes of Health under award numbers 1U01MH114824-01 and 1RF1MH128778-01. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Competing Interests

The authors declare no competing interests.

Supplementary Material

	Sst Calb2 Pdlim5	Sst Hpse Cbln4	Sst Chodl	Pvalb Reln Itm2a	Pvalb Tpbg	Lamp5 Lsp1
L1 axon	**0** / **0.40**	1.000	0.637	**0.023** / 0.13	**0.033** / 0.21	**0.013** / **0.56**
L2/3 axon	**0.013** / **0.30**	1.000	1.000	**0** / **0.35**	**0.023** / **0.27**	**0.042** / **0.52**
L4 axon	0.689	0.332	0.607	**0** / **0.35**	0.216	1.000
L5 axon	**0.033** / 0.15	**0.013** / **0.25**	1.000	0.102	0.058	0.135
L6 axon	1.000	0.081	0.288	0.182	1.000	0.058
L1 dendrite	**0.042** / 0.17	1.000	0.689	0.088	**0** / **0.42**	0.283
L2/3 dendrite	**0.033** / 0.15	1.000	0.332	**0.013** / **0.27**	**0.042** / 0.21	1.000
L4 dendrite	0.697	**0** / **0.40**	**0.013** / **0.40**	0.058	0.074	0.393
L5 dendrite	**0** / 0.19	0.102	0.393	**0** / **0.25**	**0.013** / **0.26**	1.000
L6 dendrite	1.000	0.088	0.866	0.051	0.363	1.000
soma depth	**0** / **0.26**	**0.023** / 0.21	0.246	**0** / **0.34**	**0.013** / **0.29**	**0.023** / **0.48**
axon centroid	**0.023** / 0.16	0.210	0.058	**0** / **0.34**	0.058	**0.013** / **0.44**
dendrite centroid	0.150	**0.023** / **0.26**	0.096	**0** / **0.39**	**0.013** / **0.38**	**0.023** / **0.46**
	Lamp5 Plch2 Dock5	Sst	Pvalb	Lamp5	Vip	
L1 axon	**0.042** / **0.40**	**0** / **0.29**	**0** / **0.29**	**0** / **0.47**	**0** / 0.24	
L2/3 axon	0.322	**0** / **0.41**	**0** / **0.66**	0.813	**0** / **0.27**	
L4 axon	0.351	**0** / **0.44**	**0** / **0.40**	0.210	**0** / 0.18	
L5 axon	0.081	**0** / **0.33**	**0** / 0.21	1.000	**0** / 0.18	
L6 axon	0.540	**0** / **0.37**	**0** / **0.52**	1.000	0.813	
L1 dendrite	**0** / **0.40**	**0.023** / 0.01	**0** / **0.42**	**0** / **0.25**	1.000	
L2/3 dendrite	1.000	**0** / **0.37**	**0** / **0.66**	1.000	**0** / **0.48**	
L4 dendrite	1.000	**0** / **0.42**	**0** / **0.28**	1.000	**0** / 0.16	
L5 dendrite	1.000	**0** / 0.19	**0** / 0.24	1.000	**0** / **0.34**	
L6 dendrite	1.000	**0** / **0.44**	**0** / **0.59**	1.000	0.689	
soma depth	0.067	**0** / **0.48**	**0** / **0.70**	**0** / 0.24	**0** / **0.33**	
axon centroid	0.074	**0** / **0.53**	**0** / **0.72**	**0** / **0.25**	**0** / **0.29**	
dendrite centroid	0.074	**0** / **0.45**	**0** / **0.68**	**0** / **0.26**	**0** / 0.21	

Table S1: **Statistical significance and effect size values for predicting anatomical features from gene expression via sparse linear regression for seven different cell types and four subclasses.** For each entry, the FDR-corrected p-value as calculated by a non-parametric shuffle test is listed. (See Sparse feature selection analysis under Methods.) If the value is considered statistically significant at $p \leq 0.05$, the R^2 value is also displayed (p / R^2). p-values less than or equal to 0.05 and R^2 values larger than or equal to 0.25 are shown in bold.

	Sst Calb2 Pdlim5	Sst Hpse Cbln4	Sst Chodl	Pvalb Reln Itm2a	Pvalb Tpbg	Lamp5 Lsp1
L1 axon	**0** / **0.40**	0.481	0.090	**0.002** / 0.13	**0.003** / 0.21	**0.001** / **0.56**
L2/3 axon	**0.001** / **0.30**	0.232	0.306	**0** / **0.35**	**0.002** / **0.27**	**0.004** / **0.52**
L4 axon	0.100	**0.044** / 0.01	0.085	**0** / **0.35**	**0.027** / 0.06	0.312
L5 axon	**0.003** / 0.15	**0.001** / **0.25**	0.347	**0.012** / 0.07	**0.006** / 0.14	**0.016** / -0.45
L6 axon	0.251	**0.009** / 0.10	**0.037** / 0.11	**0.022** / -0.05	0.227	**0.006** / **0.32**
L1 dendrite	**0.004** / 0.17	0.895	0.100	**0.010** / -0.26	**0** / **0.42**	**0.036** / -0.03
L2/3 dendrite	**0.003** / 0.15	0.234	**0.044** / 0.05	**0.001** / **0.27**	**0.004** / 0.21	0.189
L4 dendrite	0.102	**0** / **0.40**	**0.001** / **0.40**	**0.006** / 0.09	**0.008** / 0.15	0.054
L5 dendrite	**0** / 0.19	**0.012** / 0.13	0.054	**0** / **0.25**	**0.001** / **0.26**	0.174
L6 dendrite	0.230	**0.010** / 0.15	0.130	**0.005** / 0.08	**0.049** / 0.01	0.177
soma depth	**0** / **0.26**	**0.002** / 0.21	**0.031** / 0.06	**0** / **0.34**	**0.001** / **0.29**	**0.002** / **0.48**
axon centroid	**0.002** / 0.16	**0.026** / 0.04	**0.006** / 0.26	**0** / **0.34**	**0.006** / 0.21	**0.001** / **0.44**
dendrite centroid	**0.018** / 0.18	**0.002** / **0.26**	**0.011** / 0.22	**0** / **0.39**	**0.001** / **0.38**	**0.002** / **0.46**
	Lamp5 Plch2 Dock5	Sst	Pvalb	Lamp5	Vip	
L1 axon	**0.004** / **0.40**	**0** / **0.29**	**0** / **0.29**	**0** / **0.47**	**0** / 0.24	
L2/3 axon	**0.042** / 0.08	**0** / **0.41**	**0** / **0.66**	0.120	**0** / **0.27**	
L4 axon	**0.047** / 0.01	**0** / **0.44**	**0** / **0.40**	**0.026** / 0.01	**0** / 0.18	
L5 axon	**0.009** / 0.07	**0** / **0.33**	**0** / 0.21	0.186	**0** / 0.18	
L6 axon	0.075	**0** / **0.37**	**0** / **0.52**	0.177	0.121	
L1 dendrite	**0** / **0.40**	**0.002** / 0.01	**0** / **0.42**	**0** / **0.25**	0.208	
L2/3 dendrite	0.218	**0** / **0.37**	**0** / **0.66**	0.510	**0** / **0.48**	
L4 dendrite	0.285	**0** / **0.42**	**0** / **0.28**	0.214	**0** / 0.16	
L5 dendrite	0.217	**0** / 0.19	**0** / 0.24	0.217	**0** / **0.34**	
L6 dendrite	0.292	**0** / **0.44**	**0** / **0.59**	0.160	0.099	
soma depth	**0.007** / 0.20	**0** / **0.48**	**0** / **0.70**	**0** / 0.24	**0** / **0.33**	
axon centroid	**0.008** / 0.23	**0** / **0.53**	**0** / **0.72**	**0** / **0.25**	**0** / **0.29**	
dendrite centroid	**0.008** / 0.16	**0** / **0.45**	**0** / **0.68**	**0** / **0.26**	**0** / 0.21	

Table S2: **Statistical significance and effect size values for predicting anatomical features from gene expression via sparse linear regression for seven different cell types and four subclasses – before FDR correction.** For each entry, original p-value as calculated by a non-parametric shuffle test is listed. If the value is considered statistically significant at $p \leq 0.05$, the R^2 value is also displayed (p / R^2). p-values less than or equal to 0.05 and R^2 values larger than or equal to 0.25 are shown in bold.

Population	Feature	Genes
Sst Calb2 Pdlim5	L1 axon	Cpne7 (5), Necab2 (4), Crhr2 (4), Myh7 (4), Metrnl (3), Bcl11b (3), Ptpru (3), Atp2b4 (2), Elfn1 (2), Tmem163 (2)
Sst Calb2 Pdlim5	L2/3 axon	Hs3st1 (5), Kcnq5 (5), Mb21d2 (5), Cacng5 (3), Adra1a (3), Fxyd7 (2), Pth2r (2), Agbl4 (2), Synpr (2), Syt10 (2)
Sst Calb2 Pdlim5	soma depth	Crhr2 (5), Car10 (4), Etv1 (4), Gap43 (2), Igfbp6 (2), Klk8 (2), Neto2 (2), Galnt14 (2), Sgcz (2), Syt2 (2)
Sst Hpse Cbln4	L5 axon	Timp2 (5), Rasgef1c (5), Slitrk1 (4), Medag (3), Tnfaip8l3 (3), LOC102632463 (3), Lrrc38 (2), Gprin3 (2), Kcnip4 (2), Zbtb20 (2)
Sst Hpse Cbln4	L4 dendrite	Hpse (5), Kcnab1 (5), Timp2 (4), Rasgef1c (4), Htr1a (3), Flt3 (2), Map3k5 (2), 1700016K19Rik (2), Pde9a (2), Igsf21 (1)
Sst Hpse Cbln4	dendrite centroid	Rasgef1c (5), Myh7 (4), Krt1 (4), Mme (3), Timp2 (3), Ptpro (3), Map3k5 (3), Fxyd7 (2), Nr4a2 (2), Gabra5 (2)
Sst Chodl	L4 dendrite	Col14a1 (5), 2010300C02Rik (4), Sema6d (4), Sema3a (3), Acbd4 (3), Dusp10 (3), Coro2a (3), Sema3e (2), Sel1l3 (2), Adcyap1r1 (2)
Pvalb Reln Itm2a	L2/3 axon	Ncam2 (5), Sertad1 (5), Cort (5), Lrrtm4 (5), Thsd7a (5), Galntl6 (4), Grm7 (4), Kcnd3 (4), Npas3 (2), Nell2 (2)
Pvalb Reln Itm2a	L4 axon	Cdh9 (5), Cort (5), Gabra4 (4), NetO1 (4), Cplx2 (4), Pdlim1 (3), Kcnk3 (3), Ephb1 (2), Ndst3 (2), Lxn (1)
Pvalb Reln Itm2a	L2/3 dendrite	Cntn5 (5), Vax1 (5), Brinp3 (5), Sertad1 (5), Ncam2 (5), Kcnd3 (5), LOC105244192 (5), Cdh4 (4), Gadd45a4 (4), Grik1 (2)
Pvalb Reln Itm2a	L5 dendrite	Npas3 (5), Galntl6 (5), Kcnip4 (5), Brinp3 (5), 9630002D21Rik (4), Ano3 (3), Grm7 (3), Adra1b (3), Cacna2d3 (2), Tox (2)
Pvalb Reln Itm2a	soma depth	Thsd7a (5), Npas3 (5), Brinp3 (5), Satb1 (5), Csgalnact1 (4), Sptbn5 (4), Ankrd34b (3), Grm7 (3), Ncam2 (3), Kcnip4 (3)
Pvalb Reln Itm2a	axon centroid	Npas3 (5), Thsd7a (5), Tox (4), Sptbn5 (4), Grm7 (4), Cdh22 (3), Grik1 (3), Opcml (3), Sertad1 (2), Rgs6 (2)
Pvalb Reln Itm2a	dendrite centroid	Galntl6 (5), Sptbn5 (4), Tox (4), Kcnip4 (3), Brinp3 (3), Opcml (3), Npas3 (3), Lrp1b (2), Grm7 (2), Kcnd3 (2)
Pvalb Tpbg	L2/3 axon	Poc1a (5), Cacna1g (5), Matn2 (3), 9630002D21Rik (3), Fam46a (3), 8030453O22Rik (3), Wipf3 (2), Sfrp2 (2), Necab2 (2), Map3k5 (2)
Pvalb Tpbg	L1 dendrite	Itpka (5), Nos1 (4), Nptx2 (4), Grhl1 (3), Lypd1 (3), Cnih3 (2), Igf1 (2), Tesc (2), Zar1l (2), Sel1l3 (2)
Pvalb Tpbg	L5 dendrite	9630002D21Rik (5), 8030453O22Rik (4), LOC105243542 (4), Slc10a4 (4), Sfrp2 (4), Calb1 (3), Cd24a (3), Cmya5 (3), Dgkb (2), Hs3st2 (2)
Pvalb Tpbg	soma depth	Col25a1 (4), 9630002D21Rik (3), BC048546 (3), Osbpl3 (3), Nxph1 (3), Myo1e (3), Itpka (2), Sfrp2 (2), Nr4a3 (2), LOC105243542 (2)
Pvalb Tpbg	dendrite centroid	9630002D21Rik (5), Col25a1 (5), Nxph1 (4), Klhdc8a (3), Sfrp2 (3), BC048546 (3), Osbpl3 (3), LOC105243542 (3), Dgkb (3), Itpka (3)
Lamp5 Lsp1	L1 axon	Dusp14 (5), March11 (5), Nptxr (4), Sstr4 (3), Rasal1 (3), Pou3f3 (3), Ndst4 (2), Glce (2), Fhod3 (2), Mgat4c (2)
Lamp5 Lsp1	L2/3 axon	Sertad1 (5), Sdk2 (4), Hunk (4), Sulf1 (4), Pbx3 (3), Teddm3 (2), Elavl2 (2), 3300002P13Rik (2), Rorb (1), Rnf152 (1)
Lamp5 Lsp1	soma depth	Adra1b (4), Sulf1 (4), Rnf152 (3), Pde1a (3), Dusp14 (2), Hdac9 (2), Ildr2 (2), LOC105243542 (2), Nog (2), Slc44a5 (2)
Lamp5 Lsp1	axon centroid	Rnf152 (5), Adra1b (4), Nog (3), Pde1a (3), Sulf1 (3), Mgat4c (2), Rap1gap2 (2), Dusp14 (2), LOC105243542 (2), Cplx3 (1)
Lamp5 Lsp1	dendrite centroid	Adra1b (5), Sulf1 (4), Rnf152 (4), Spats2l (3), Hdac9 (3), Pde1a (3), Nog (2), Brinp2 (2), Dusp14 (2), Ildr2 (2)
Lamp5 Plch2 Dock5	L1 axon	Zbtb20 (4), Fndc5 (4), Grin2a (4), Dner (4), Asap2 (3), Cntn3 (3), Slc5a5 (2), Sel1l3 (2), Rgs8 (2), Opcml (2)
Lamp5 Plch2 Dock5	L1 dendrite	Zbtb20 (5), Fndc5 (5), Camk1g (5), Ngf (3), Pradc1 (2), Tmod1 (2), Mitf (2), Thsd7a (2), Timp2 (1), Kcnip1 (1)
Sst	L1 axon	Pcsk5 (5), Myh7 (5), Mgat4c (5), Kctd8 (4), Olfm3 (4), Npy2r (4), Atp2b4 (4), Thsd7a (3), P2ry1 (3), Ajap1 (2)
Sst	L2/3 axon	Grik1 (5), Mgat4c (5), Thsd7a (5), Col11a1 (4), Rgs9 (4), LOC105244376 (4), Gda (3), Ajap1 (3), Col25a1 (2), Satb1 (2)
Sst	L4 axon	Tmem132d (5), Npnt (5), Hpse (5), Syndig1l (5), Lrrtm4 (4), Pde9a (4), Cort (3), March1 (3), Rerg (2), Kcnab1 (2)
Sst	L5 axon	Arc (5), Sema3e (5), Col11a1 (5), Grm1 (5), Spon1 (4), Olfm3 (4), Lypd6 (4), Myh7 (3), Mgat4c (3), Map3k5 (2)
Sst	L6 axon	Col6a1 (5), Dpp10 (5), Grik1 (4), Brinp3 (4), Car103 (4), Crym (4), Itga8 (4), Kirrel3 (3), Penk (3), Sp8 (3)
Sst	L2/3 dendrite	Rgs9 (5), Npy2r (5), Cbln4 (5), Mgat4c (5), LOC105244376 (5), Grm1 (5), Brinp3 (3), Myh7 (3), Thsd7a (2), Dab1 (2)
Sst	L4 dendrite	Syndig1l (5), Lypd6b (5), Hpse (5), Col11a1 (5), Rerg (4), Sema3e (4), Gabrg3 (4), Epha3(4), Grik1 (3), Pde9a (2)
Sst	L6 dendrite	Syndig1 (5), Col6a1 (5), Olfm3 (5), Kirrel3 (5), Sorcs3 (4), Crym (4), Ptpru (3), Dpp10 (3), Grik1 (3), Itga8 (2)
Sst	soma depth	Mgat4c (5), Olfm3 (5), Kirrel3 (5), Lypd6 (4), Brinp3 (4), Car10 (4), Spon1 (3), Grik1 (3), LOC105244376 (3), Ajap1 (2)
Sst	axon centroid	Kirrel3 (5), Thsd7a (5), Olfm3 (5), Car10 (5), Mgat4c (5), Syndig1 (4), Brinp3 (4), Grik1 (4), Unc5d (2), Unc5d (2)
Sst	dendrite centroid	Olfm3 (5), Mgat4c (5), Thsd7a (3), Spon1 (3), Rgs9 (3), Brinp3 (3), Myh7 (3), Car10 (3), Grik1 (3), Kirrel3 (2)
Pvalb	L1 axon	Igf1 (5), Col25a1 (5), LOC102636041 (5), 2010300C02Rik (4), Myh7 (4), C130074G19Rik (3), LOC102636700 (3), Itpka (3), Rgs14 (2), Cdh4 (2)
Pvalb	L2/3 axon	Grik1 (5), Thsd7a (5), Calb1 (5), Ncam2 (5), Galntl6 (5), Cnr1 (4), Pcp4 (4), Gpr83 (4), Kcnc2 (2), Col25a1 (2)
Pvalb	L4 axon	Lrrtm4 (5), Parm1 (5), Cntnap5c (4), Ntrk3 (4), Ddn (4), NetO1 (3), Lin7a (3), Cdh9 (2), Rasgrf1 (2), Gpr123 (1)
Pvalb	L6 axon	Chrm2 (5), Galntl6 (5), Sema3e (5), Prdm8 (4), Rasgrf1 (4), Zfp804b (3), Thsd7a (3), Kcnc2 (2), Lrp1b (2), Brinp2 (2)
Pvalb	L1 dendrite	Igf1 (5), Myh7 (5), Oprm1 (5), Rgs14 (4), Wnt10a (3), Igfbp6 (3), Kctd4 (3), Sstr2 (3), Itpka (2), Adcyap1 (2)
Pvalb	L2/3 dendrite	Thsd7a (5), Ncam2 (5), Grik1 (5), Galntl6 (5), Col25a1 (5), Cnr1 (5), Calb1 (5), Pcp4 (4), Cdh13 (4), Ntng1 (2)
Pvalb	L4 dendrite	Parm1 (5), Ntrk3 (5), Chrna7 (3), St6galnac5 (3), Cntnap5c (3), Gpr123 (3), Gria1 (2), Lrrtm4 (2), Col25a1 (2), Trpc3 (2)
Pvalb	L6 dendrite	Chrm2 (5), Sema3e (5), Rbp4 (5), Gpr83 (5), Lrp1b (4), Synpr (4), Galntl6 (4), Lmo3 (2), Fam19a5 (2), Cdh11 (2)
Pvalb	soma depth	Thsd7a (5), Chrm2 (5), Galntl6 (5), Gpr83 (4), Col25a1 (4), Myh7 (3), Fstl4 (3), Edil3 (3), Calb1 (3)
Pvalb	axon centroid	Thsd7a (5), Sema3e (5), Chrm2 (5), Col25a1 (5), Galntl6 (5), Edil3 (4), Kcnc2 (3), Fstl4 (3), Gpr83 (3), Lgals1 (2)
Pvalb	dendrite centroid	Thsd7a (5), Sema3e (5), Chrm2 (5), Galntl6 (5), Gpr83 (5), Col25a1 (4), Myh7 (3), Pcp4 (3), Edil3 (3), Lypd6b (2)
Lamp5	l1 axon	Opcml (5), Timp2 (5), Dner (5), Oxtr (3), Tshz3 (3), Rgs8 (3), Flt3 (2), Kcns3 (2), Tmem132d (2), Htr1f (1)
Lamp5	l1 dendrite	Timp2 (5), Rasgrf1 (3), Tnnt1 (3), Dner (3), Opcml (3), Plch2 (3), Zfp467 (2), Gabra5 (2), Lingo2 (2), Mitf (2)
Lamp5	axon centroid	Islr2 (4), Rgma (3), Cecr6 (3), Slc44a5 (3), Rtn4r (3), Rap1gap2 (3), Cplx3 (2), Dner (2), Oprm1 (2), Plch2 (2)
Lamp5	dendrite centroid	Slc44a5 (4), Glt8d2 (4), Timp2 (3), Cecr6 (3), Islr2 (3), Cplx3 (3), Rgma (3), Cdca7 (2), Rnf152 (2), Oprm1 (2)
Vip	L2/3 axon	Kcnmb4 (5), Elmo1 (5), Isoc1 (5), Kirrel3 (4), Cdh13 (4), Parm1 (4), Fndc5 (3), Whrn (3), Kit (3), Iqgap2 (2)
Vip	L2/3 dendrite	Synpr (5), Kirrel3 (5), Sorl1 (4), Sorcs2 (4), Cort (4), Gap43 (4), Eya4 (3), Slc10a4 (2), Chrm3 (2), Cxcl14 (2)
Vip	L5 dendrite	Synpr (5), Sfrp2 (5), Islr2 (5), Kirrel3 (5), Kit (4), Col24a1 (4), Megf10 (3), Lhx6 (3), Chrm3 (2), Erbb4 (2)
Vip	soma depth	Synpr (5), Ppapdc1a (5), Kirrel3 (5), Rln1 (4), Kit (3), Chrm3 (3), Megf10 (3), Whrn (2), Osbpl3 (2), Eya4 (2)
Vip	axon centroid	Synpr (5), Kirrel3 (5), Kit (5), Rln1 (4), Ppapdc1a (4), Sfrp2 (3), Whrn (2), Unc5a (2), Cort (2), Eya4 (2)

Table S3: **Gene sets selected via sparse linear regression for different cell types/subclasses and anatomical features.** Only statistically significant sets with $R^2 \geq 0.25$ are shown (See Supplementary Table 1). Numbers in parentheses denote the number of times the preceding gene was selected out of 5 cross-validation runs.

Precision	Axon		Dendrite		Neurite	
Distance, μm	raw	masked	raw	masked	raw	masked
2	0.81 ± 0.09	0.83 ± 0.08	0.72 ± 0.16	0.72 ± 0.16	0.89 ± 0.06	0.90 ± 0.06
5	0.89 ± 0.09	0.91 ± 0.07	0.76 ± 0.15	0.77 ± 0.15	0.95 ± 0.06	0.97 ± 0.04
10	0.93 ± 0.07	0.95 ± 0.05	0.81 ± 0.13	0.82 ± 0.13	0.97 ± 0.05	0.99 ± 0.02

Recall	Axon		Dendrite		Neurite	
Distance, μm	raw	masked	raw	masked	raw	masked
2	0.58 ± 0.16	0.58 ± 0.16	0.80 ± 0.20	0.80 ± 0.20	0.73 ± 0.12	0.73 ± 0.12
5	0.66 ± 0.17	0.66 ± 0.17	0.83 ± 0.20	0.83 ± 0.20	0.80 ± 0.12	0.80 ± 0.12
10	0.75 ± 0.16	0.75 ± 0.16	0.87 ± 0.18	0.87 ± 0.18	0.87 ± 0.10	0.87 ± 0.10

F1-score	Axon		Dendrite		Neurite	
Distance, μm	raw	masked	raw	masked	raw	masked
2	0.66 ± 0.13	0.67 ± 0.13	0.74 ± 0.16	0.74 ± 0.16	0.79 ± 0.09	0.80 ± 0.09
5	0.75 ± 0.13	0.75 ± 0.13	0.78 ± 0.15	0.78 ± 0.16	0.87 ± 0.08	0.87 ± 0.08
10	0.82 ± 0.12	0.83 ± 0.12	0.82 ± 0.14	0.83 ± 0.14	0.91 ± 0.07	0.92 ± 0.07

Table S4: **Neuron reconstruction accuracy.** Precision, recall, and F1-score values are calculated by comparing automated and manual trace nodes within a given distance (2, 5, and 10 μm). Mean ± s.d. of the values over 347 cells are reported for axonal, dendritic, and neurite (combined axonal and dendritic) nodes of raw and masked automated traces.

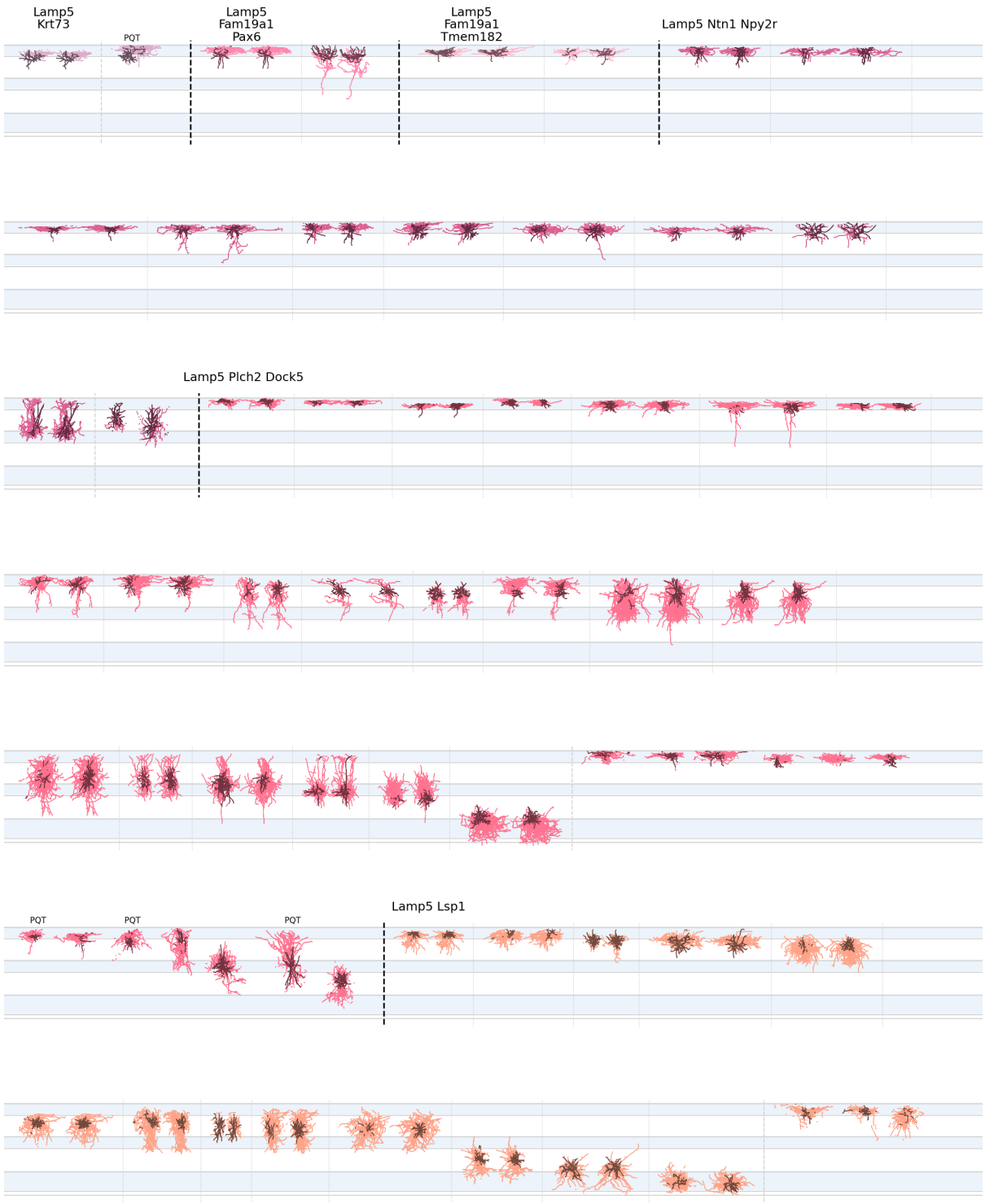

Figure S1: **Morphological reconstructions of inhibitory neurons ordered by t-type – 1 of 11.** 543 neurons have both automated and manual reconstructions, 270 - only automated ones. For each t-type, cells with both automated and manual reconstructions are shown first, separated by faint dashed lines, followed by cells that are reconstructed only automatically. Dendrites and axon are in darker and lighter colors, respectively. Best viewed digitally. PQT: poor quality transcriptomic characterization – not used for t-type related analyses. T: cells used in segmentation model training.

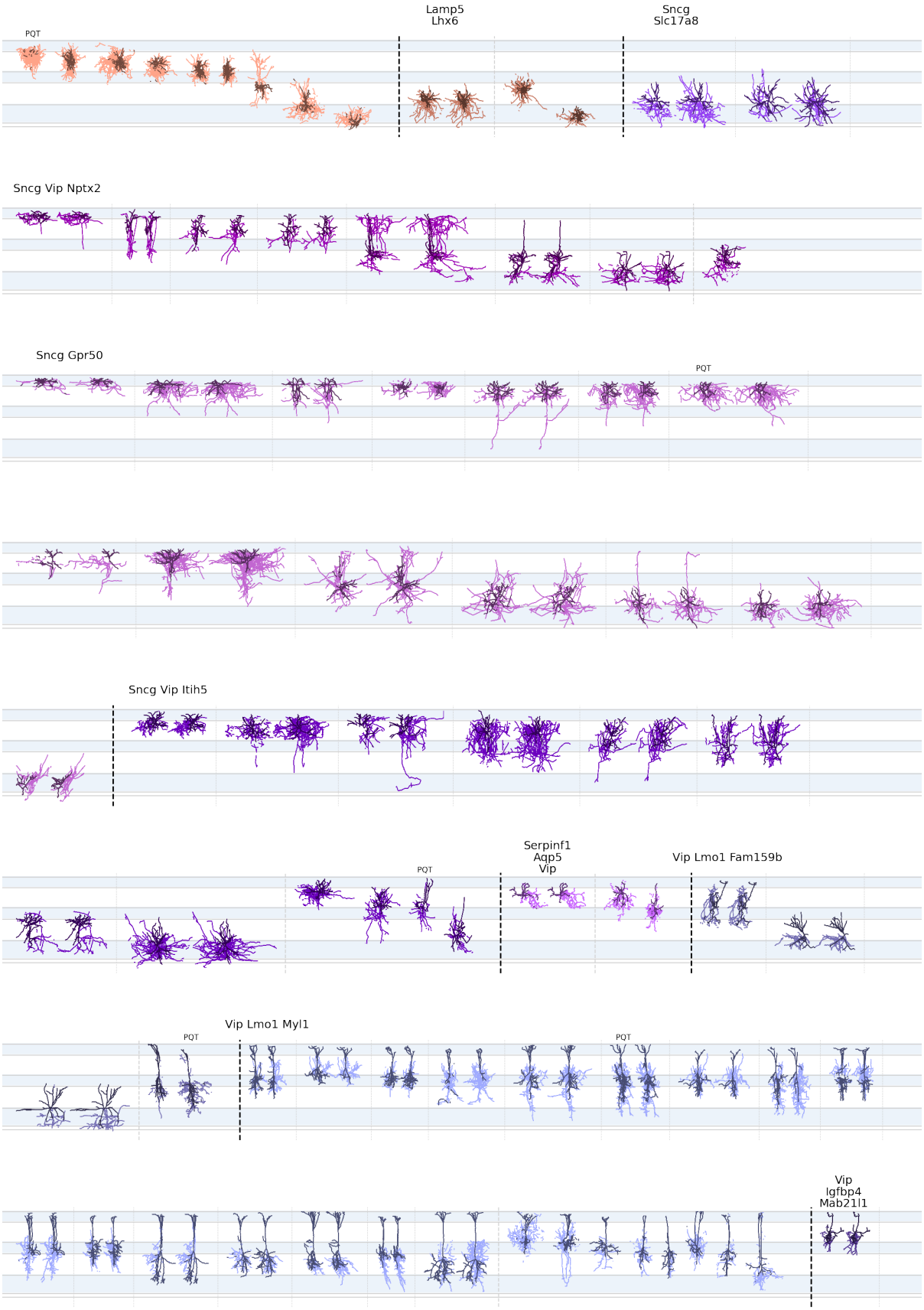

Figure S2: **Morphological reconstructions of inhibitory neurons ordered by t-type – 2 of 11.** Please refer to the caption of Figure S1.

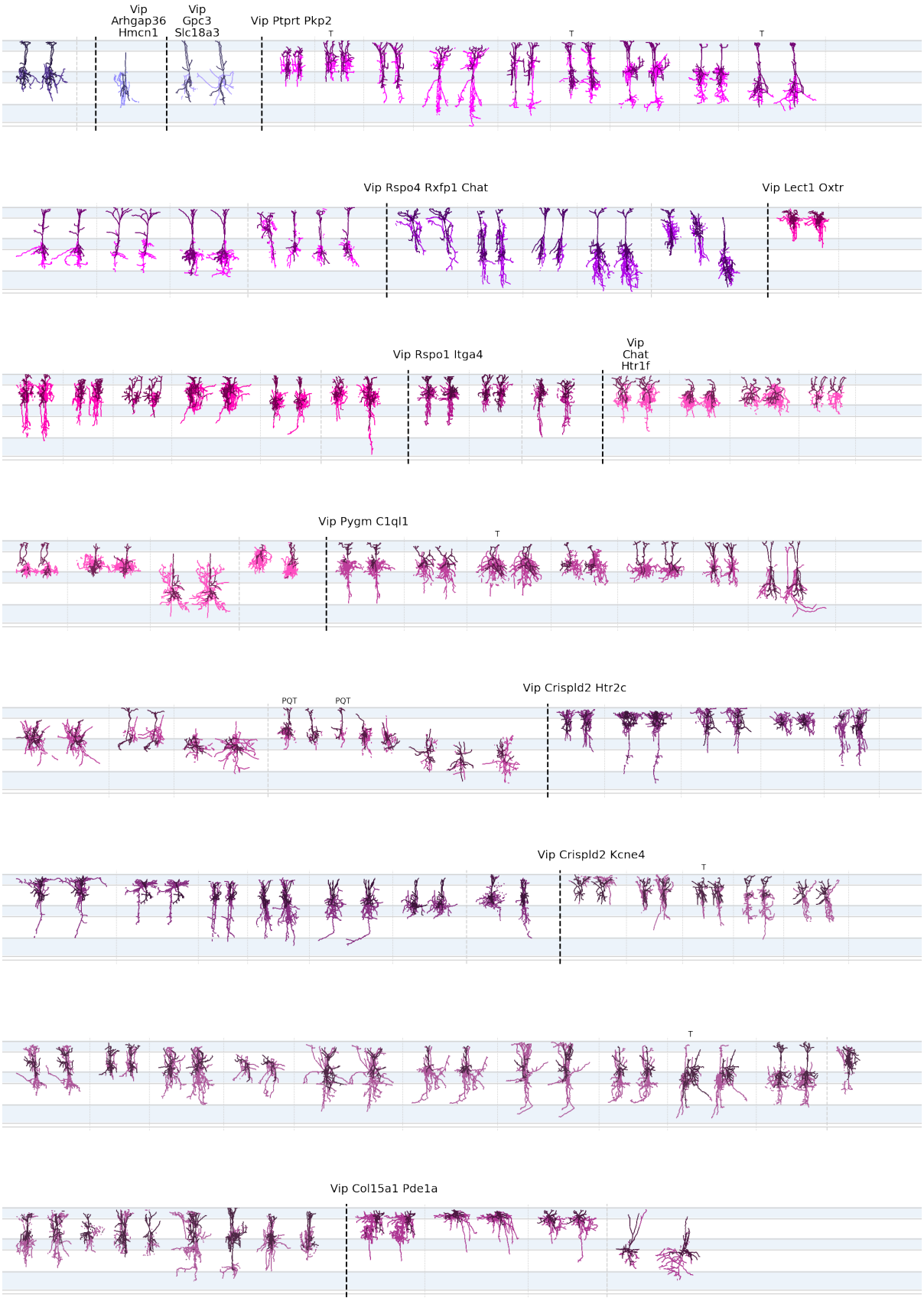

Figure S3: **Morphological reconstructions of inhibitory neurons ordered by t-type – 3 of 11.** Please refer to the caption of Figure S1.

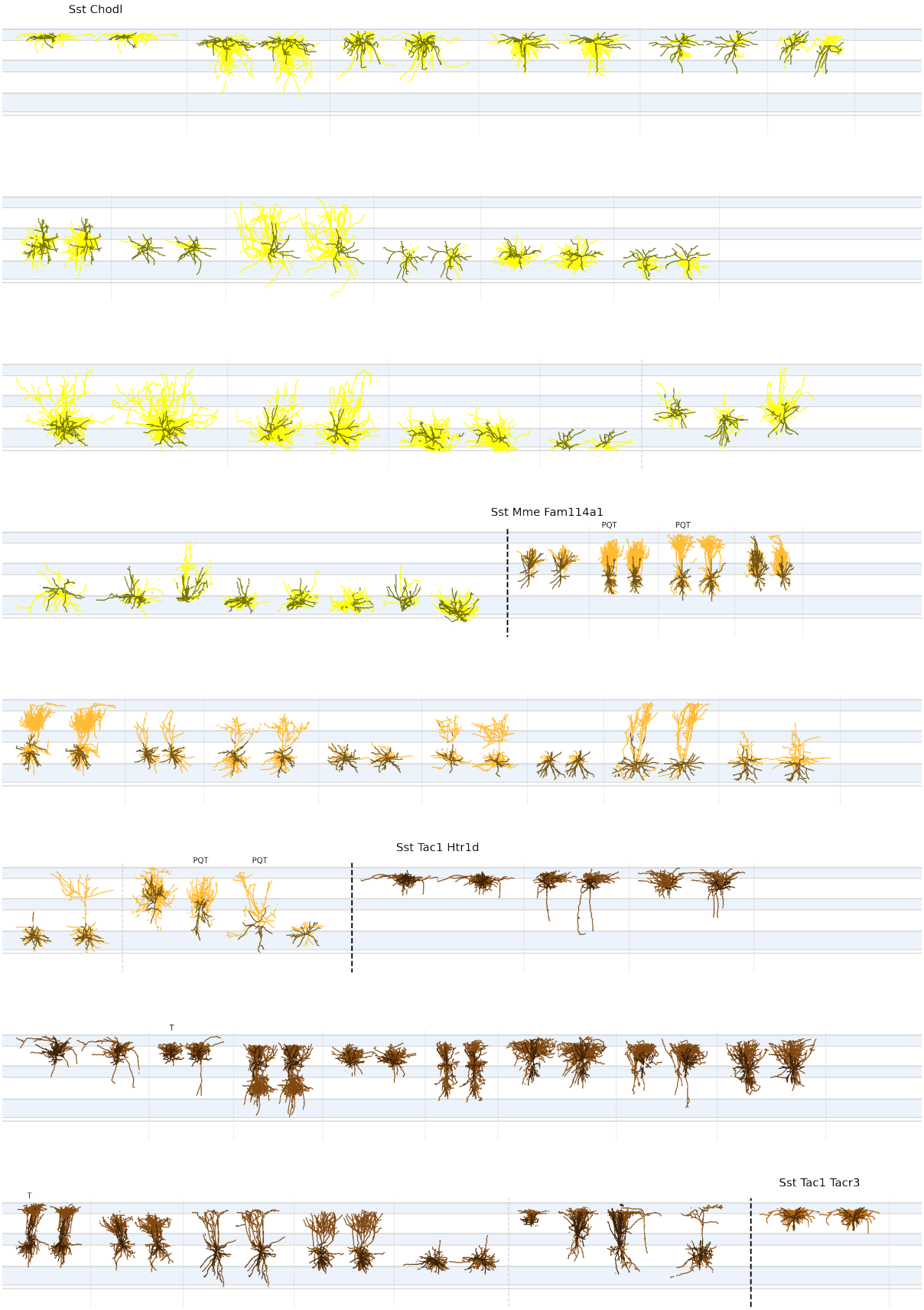

Figure S4: **Morphological reconstructions of inhibitory neurons ordered by t-type – 4 of 11.** Please refer to the caption of Figure S1.

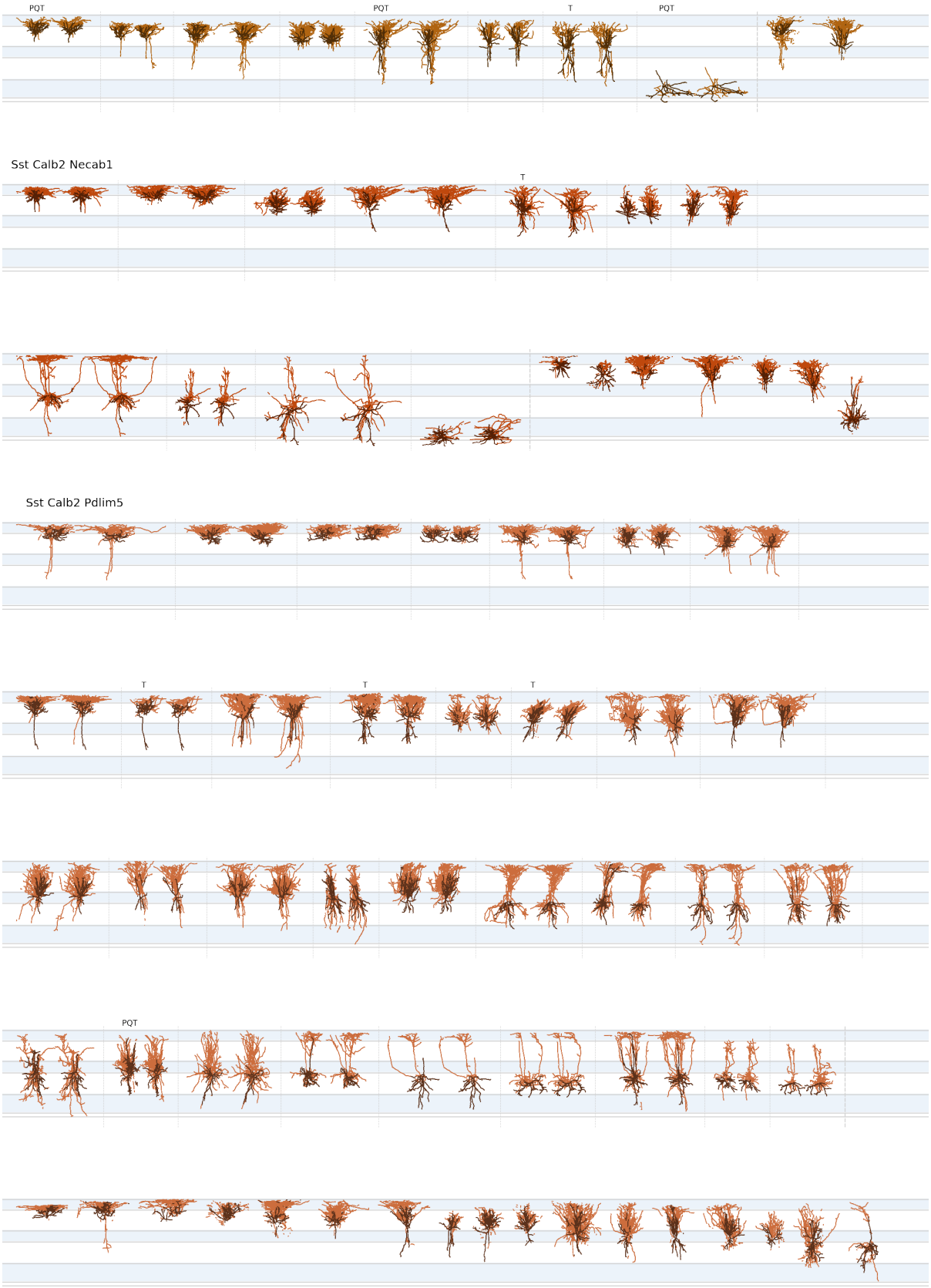

Figure S5: **Morphological reconstructions of inhibitory neurons ordered by t-type – 5 of 11.** Please refer to the caption of Figure S1.

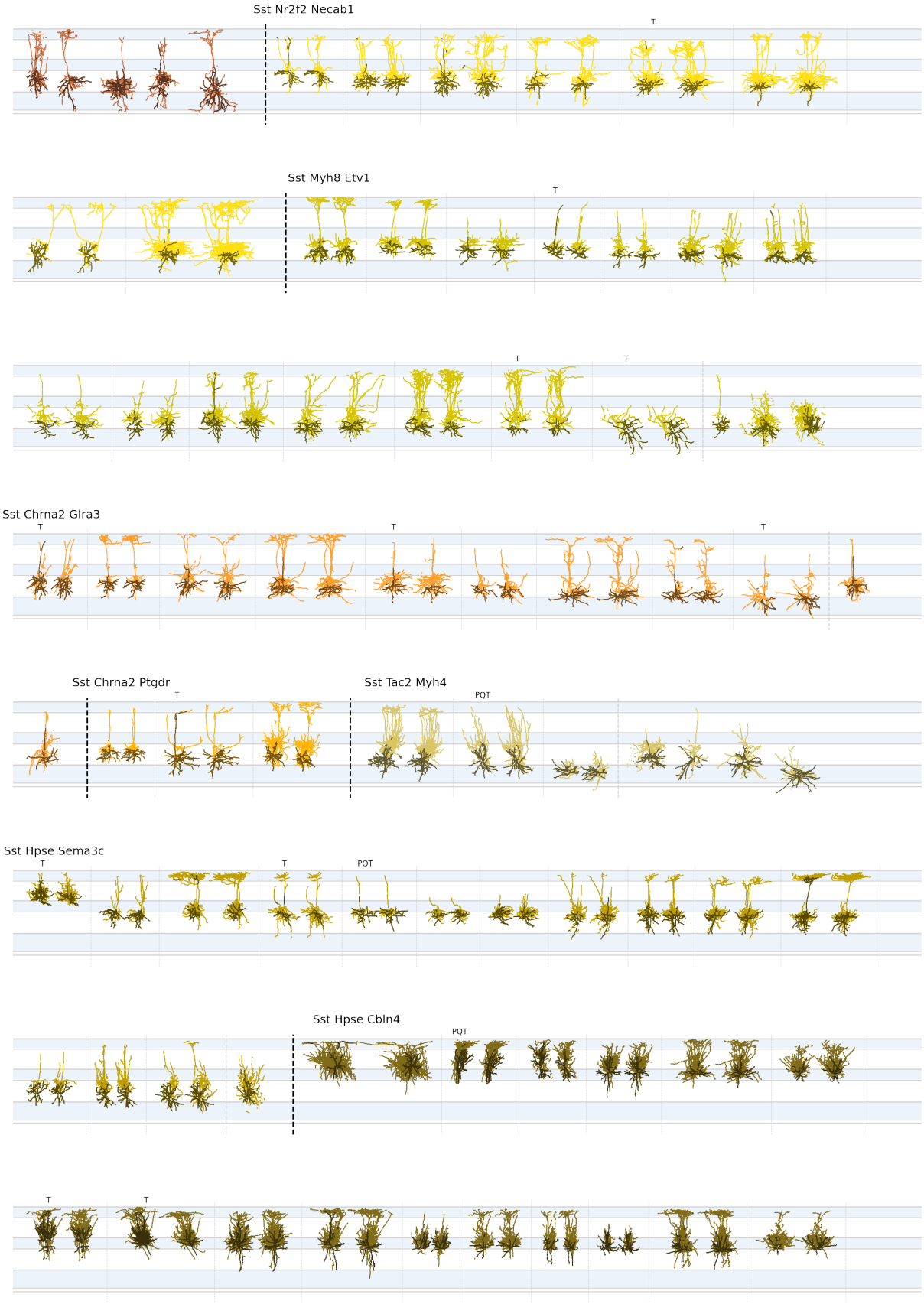

Figure S6: **Morphological reconstructions of inhibitory neurons ordered by t-type – 6 of 11.** Please refer to the caption of Figure S1.

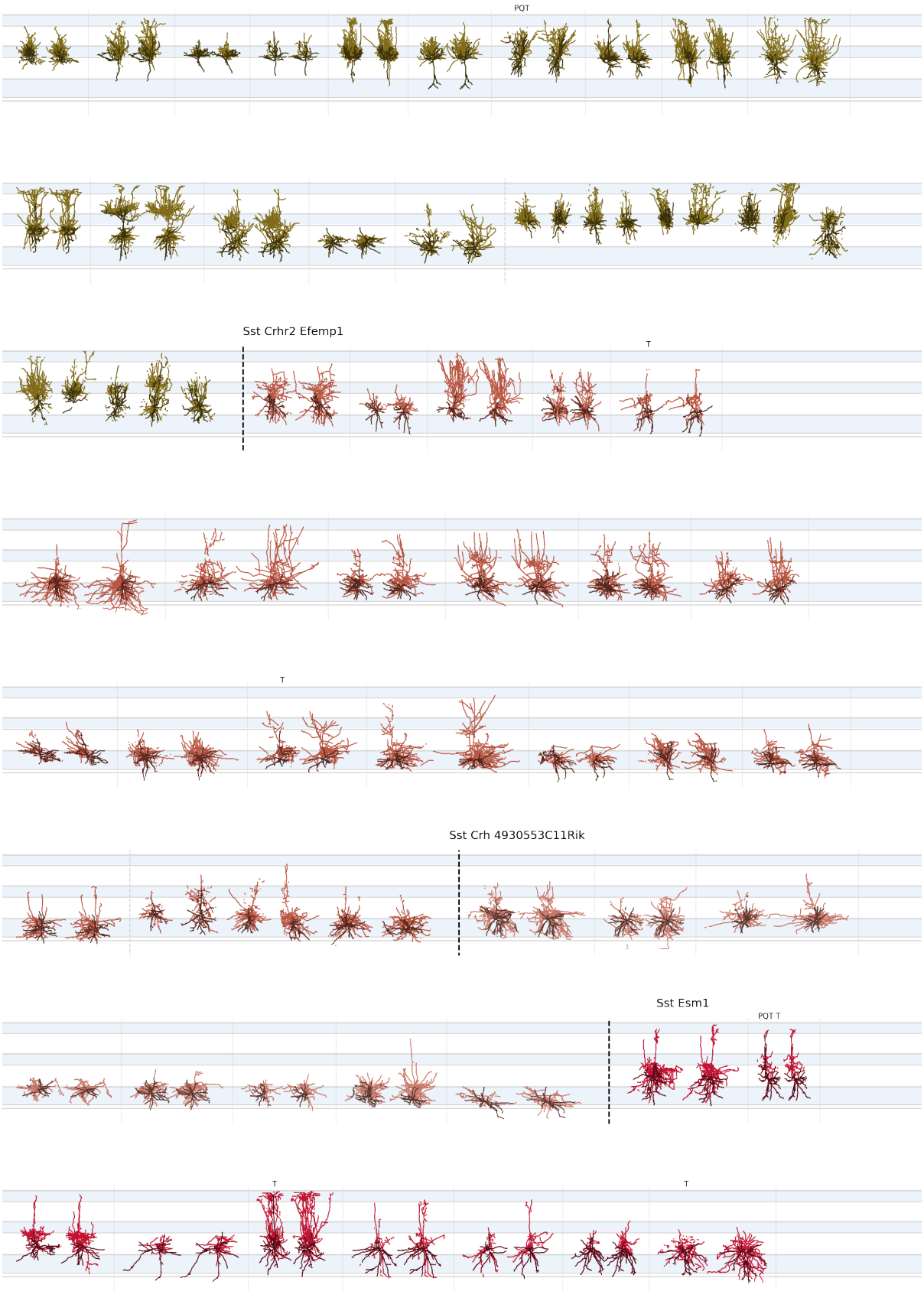

Figure S7: **Morphological reconstructions of inhibitory neurons ordered by t-type – 7 of 11.** Please refer to the caption of Figure S1.

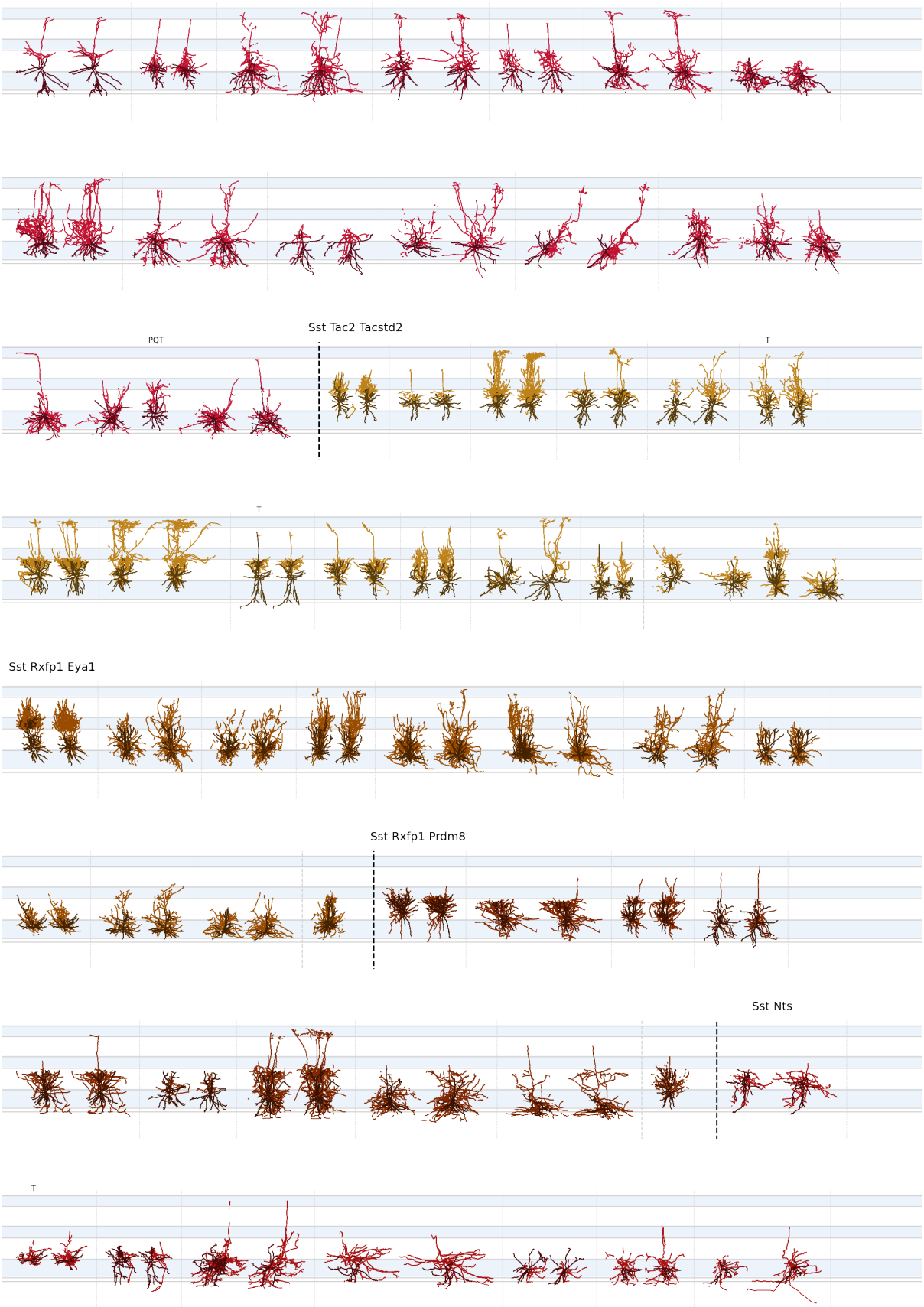

Figure S8: **Morphological reconstructions of inhibitory neurons ordered by t-type – 8 of 11.** Please refer to the caption of Figure S1.

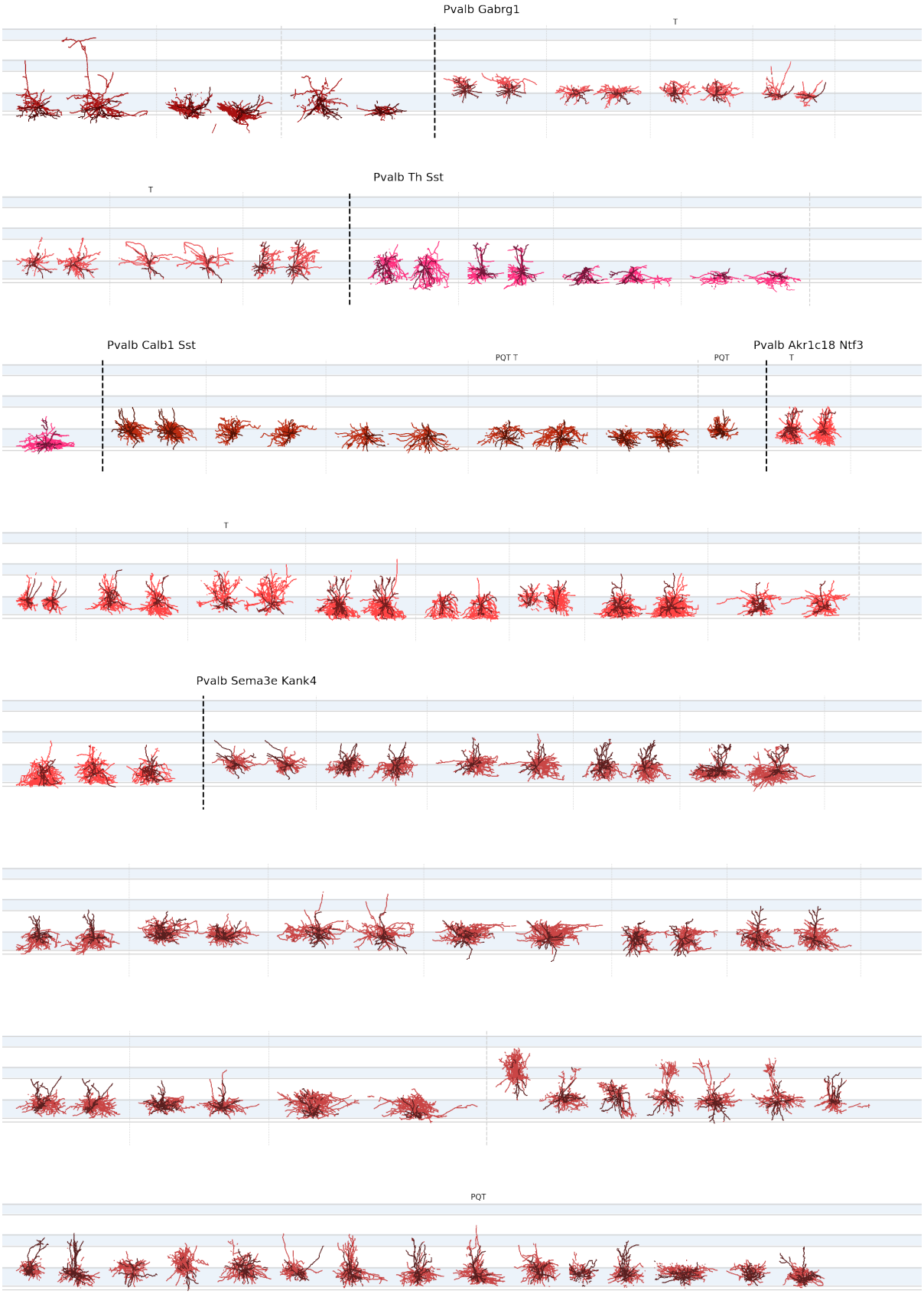

Figure S9: **Morphological reconstructions of inhibitory neurons ordered by t-type – 9 of 11.** Please refer to the caption of Figure S1.

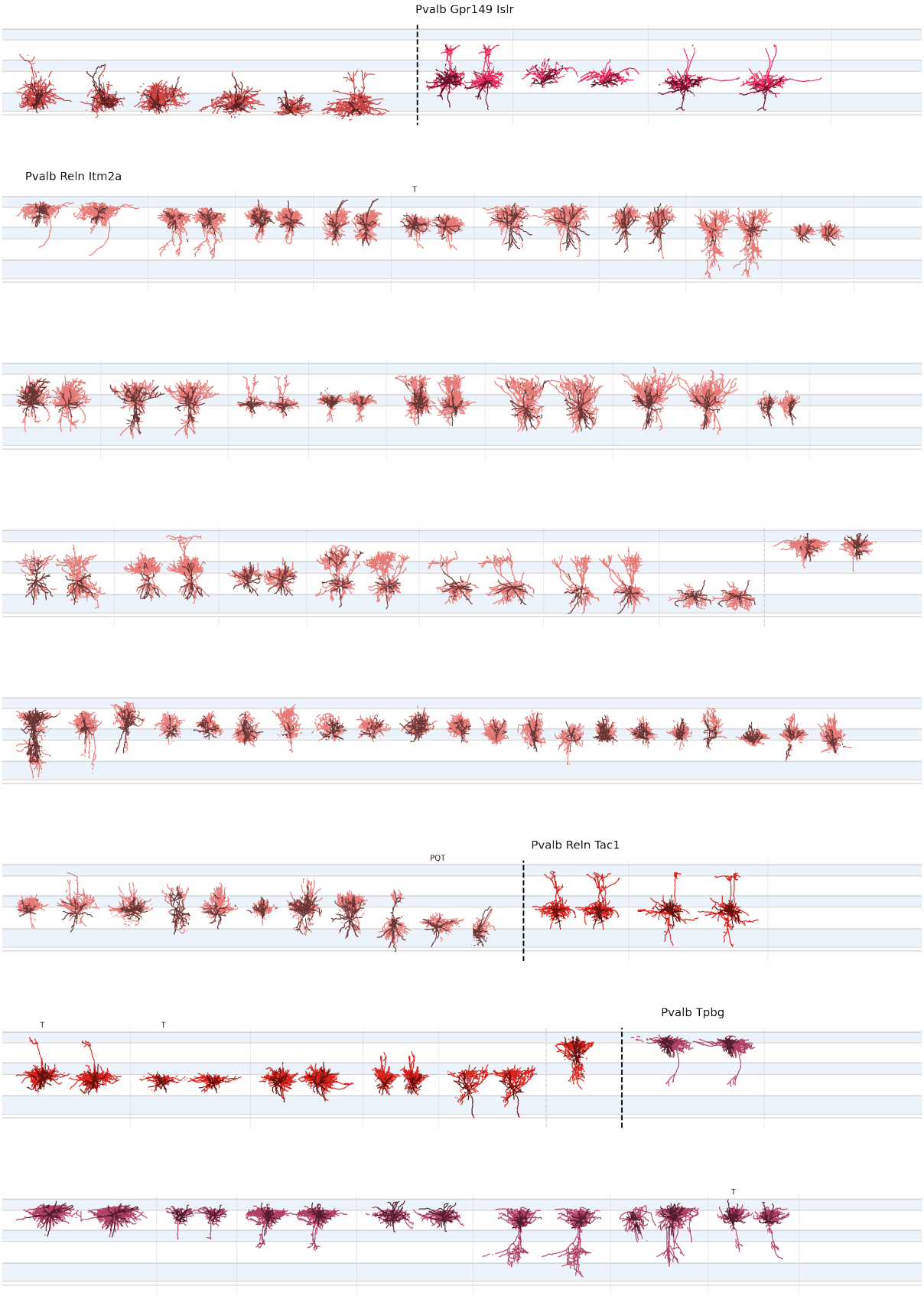

Figure S10: **Morphological reconstructions of inhibitory neurons ordered by t-type – 10 of 11.** Please refer to the caption of Figure S1.

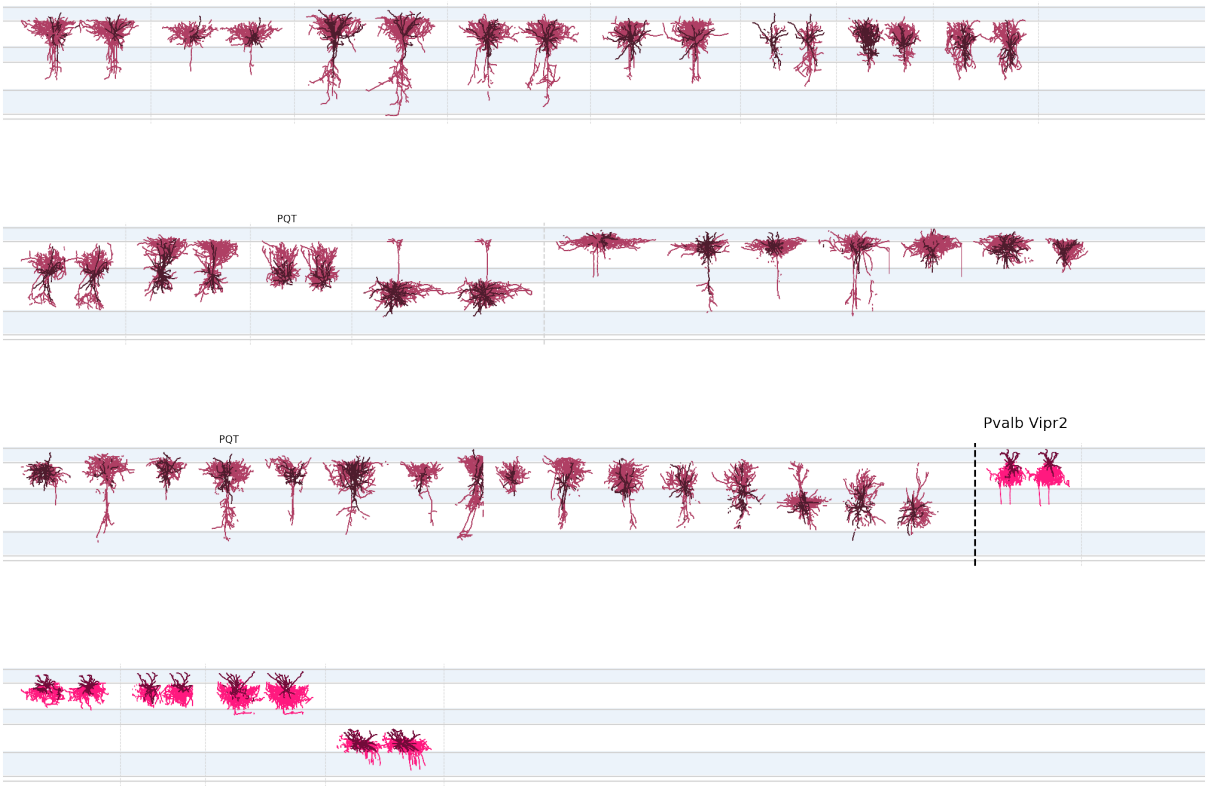

Figure S11: **Morphological reconstructions of inhibitory neurons ordered by t-type – 11 of 11.** Please refer to the caption of Figure S1.

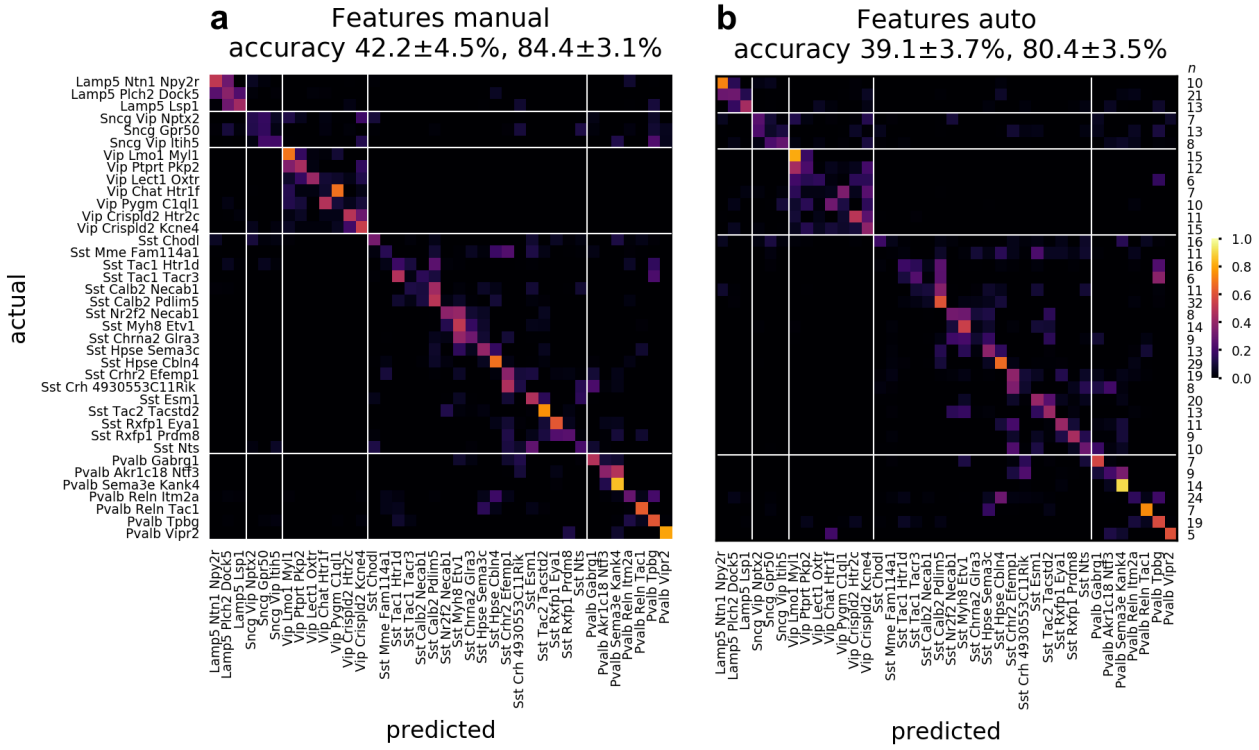

Figure S12: **Comparison of cell type classification accuracy based on a set of classical morphometric features using manually vs automatically reconstructed cells.** Confusion matrix for the classification of 38 t-types based on features, using 488 manually (**a**) and automatically (**b**) reconstructed cells. Accuracy values reported in the headers refer to mean ± s.d. of the overall t-type and t-subclass classifiers, respectively, across cross-validation folds. Rightmost column lists the number of cells in each t-type (n).

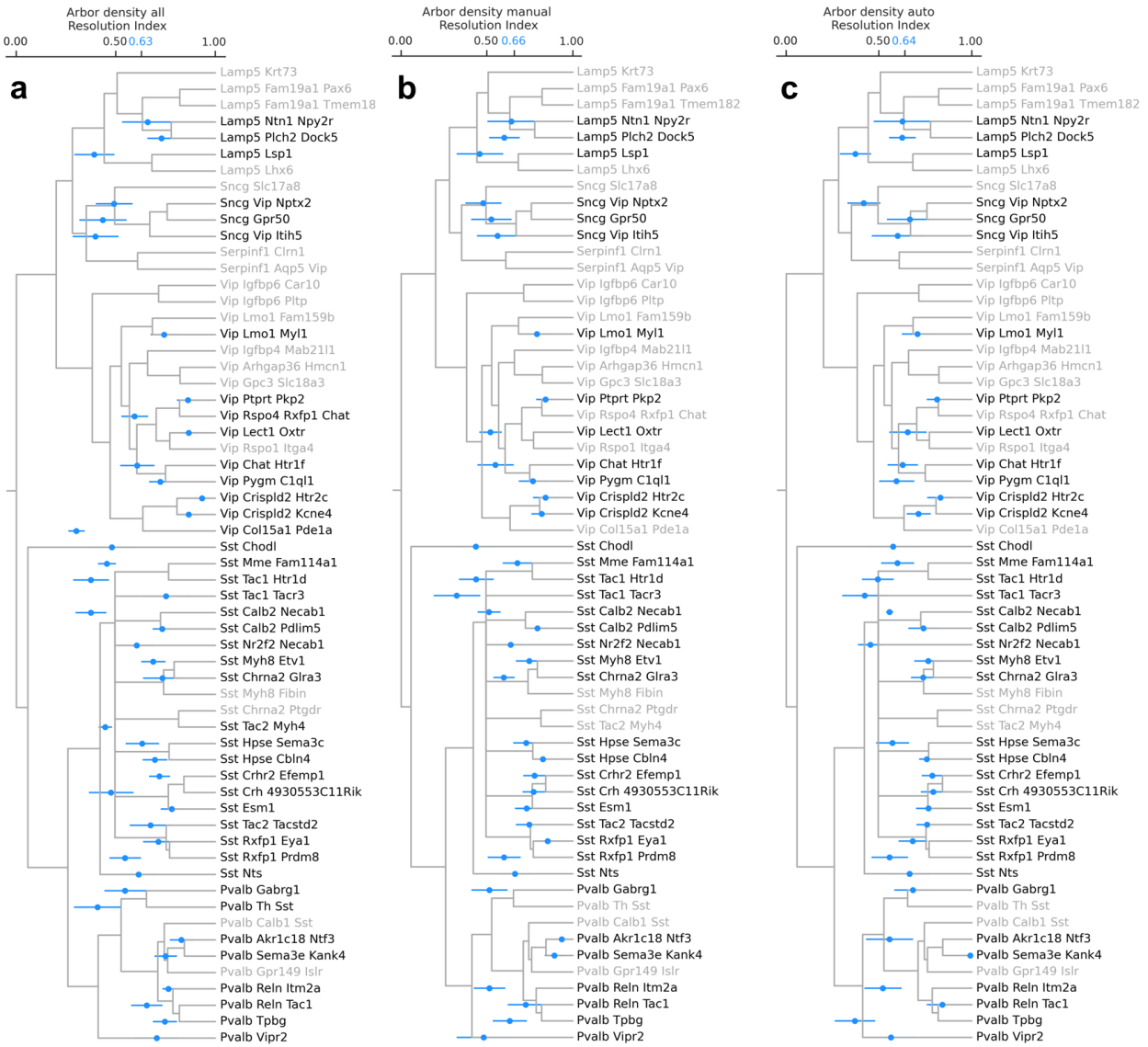

Figure S13: **Arbor density-based cell type classification performance using resolution index.** The depth of the Tasic et al. GABAergic neuron hierarchical tree is normalized to the 0-1 range, and resolution index for each cell is defined as the value of the closest common ancestor of the true and predicted leaf node labels. Perfect classification corresponds to resolution index of 1. Error bars indicate resolution index (mean ± SE over 10-fold stratified cross validation sets) for each cell type. The overall mean across cell types is indicated by blue on the y-axis. Cell types absent from the dataset are indicated by grayed out labels on the x-axis. Classification was performed with multi-layer perceptron classifiers using arbor density representation of all (**a**), manually (**b**) and automatically (**c**) reconstructed neurons as input. Resolution index values are not strictly comparable across panels because the numbers of cells and types are not identical. Left panel: 747 cells, 42 types; middle and right panels: 488 cells, 38 types. (See Figure 3.)

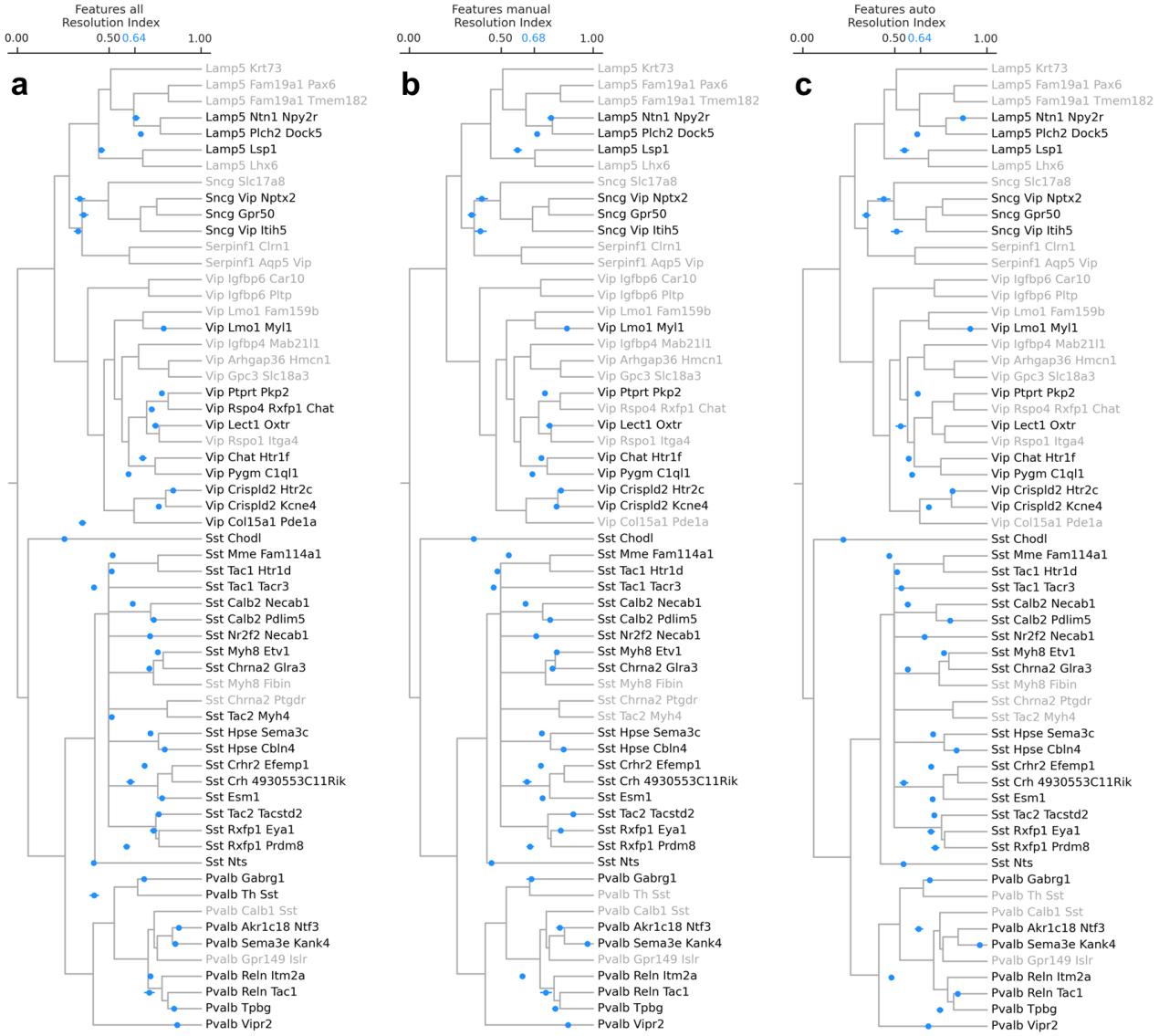

Figure S14: **Feature-based cell type classification performance using resolution index.** As before, error bars indicate resolution index (mean ± SE over 5-fold stratified cross validation sets repeated 20 times) for each cell type. The overall mean across cell types is indicated by blue on the y-axis. Cell types absent from the dataset are indicated by grayed out labels on the x-axis. Classification was performed with random forest classifiers using morphometric features of all (**a**), manually (**b**) and automatically reconstructed cells (**c**) as input. Resolution index values are not strictly comparable across panels because the numbers of cells and types are not identical. Left panel: 747 cells, 42 types; middle and right panels: 488 cells, 38 types. (See Figure 3.)