PARVALBUMIN-POSITIVE INTERNEURON MORPHOLOGIES IN MOUSE PRIMARY SOMATOSENSORY AND VISUAL CORTICES

SINGLE-CELL MORPHOLOGIES OF PARVALBUMIN-POSITIVE INTERNEURONS IN THE MOUSE PRIMARY SOMATOSENSORY AND VISUAL CORTICES

By Maheshwar Panday , Honours B.Sc. A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the Requirements for the Degree Master of Science

> McMaster University © Copyright by Maheshwar Panday, November 2024

MASTER OF SCIENCE (2024)

McMaster University

(Neuroscience)

Hamilton, Ontario

TITLE : Single-Cell Morphologies of Parvalbumin-Positive Neurons AUTHOR : Maheshwar Panday, Honours BSc (McMaster University) SUPERVISOR : Dr. Kathryn M. Murphy NUMBER OF PAGES: xii, 139

Lay Abstract

Cells in the brain perform well-defined functions, and are highly connected to their structures. Patterns in their structure, function or molecular composition have served to group cells together into cell types. Classically, approaches to defining cell types relied heavily on cell structure, their appearance as viewed under the microscope. However, these approaches were entirely descriptive, and alluded to complex shape properties that could not be quantified at the time. Modern approaches to defining cell types classify cells based on the expression measurements of thousands of genes per cell. Conversely, approaches to analyze and classify cells by their size and shapes have remained comparatively simple, in terms of a single measurement, and have not kept pace with more modern, molecular techniques. With an increased push for computational methods in image analysis, a tool to quantify cell morphologies is needed. In this thesis I have developed a tool that obtains 97 measurements that describe the size and shape of single cells for a large number of cells. This tool was applied to Parvalbumin-positive (PV+) interneurons, a well-defined subclass of neurons involved in sensory development, and disease.

Abstract

Classic approaches to defining cell types have looked to describe the elaborate size and shape characteristics of single cells.Classic approaches to neuroanatomy have remained a largely descriptive gold standard. However, to keep pace with the advances made by single-cell technologies, suitable data structures and analysis methods are needed to align the types of insights afforded by these two major techniques.

Using open-access in situ hybridization (ISH) data from the Allen Institute Mouse Neuroanatomy database, and the Allen Cell Types database , we developed a method to quantify and characterize the cell morphologies of PV+ interneurons in the mouse primary somatosensory (S1) and visual (V1) cortices. Using a custom image analysis pipeline in CellProfiler, we obtained high-dimensional, cell morphology data for thousands of single-labelled cells. We defined and characterized cell morphologies using Robust-Sparse K-Means Clustering (RSKC) and the denSNE dimensionality reduction algorithm to determine the salient morphology features that distinguish among groups of labelled cells. We show that our approach can integrate data from ISH and whole-cell fill, establishing the multimodal potential of our pipeline.Furthermore, we apply our pipeline to probe the morphological correlates of transcriptomic classes of PV+ interneurons establishing the potential for morphological differences highlighting different transcriptomic classes. Our pipeline identifies morphological complexity comparable to the complexity seen in single-cell transcriptomics studies. Insights from our work overcome the limitations in how cell morphologies are presented, using a quantitative, data-driven and unsupervised approach.

Acknowledgements

I would like to express my sincere thanks and appreciation to my supervisor, Dr. Kathryn Murphy for her steadfast guidance, patience, and support throughout my course of study. Her enthusiasm and passion for neuroscience continues to inspire me, and I am grateful for the countless, invaluable lessons I have learned throughout my study. I would also like to thank my supervisory committee members: Dr. Katrina Choe, Dr. Benicio Frey, and Dr. Pratheepa Jeganathan for their thoughtful insights and helpful comments in the completion of my research and thesis. I would also like to thank Ms. Sandra Murphy for her thoughtful, and supportive answers to my often never-ending lists of questions.

I would like to thank the students of the Visual Neuroscience Lab. I have learned so much from all of you on this journey. A special thank you to Leanne Monteiro for her support, assistance and encouragement.

Lastly, I would like to express a deep appreciation and gratitude to my parents: Nandrani and Motilall Panday, and to my aunt and uncle: Sonia and Bharrat Matadial. Your support and guidance has made this journey possible, and memorable. For always believing in me, I am grateful beyond words. Je vous apprécie énormément.

Table of Contents

Lay Abstract	ii
Abstract	iii
Acknowledgements	iv
Table of Contents	V
List of Figures and Tables	vi
List of Abbreviations	ix
Declaration of Academic Achievement	xi
Chapter 1. General Introduction	1
1.1 High-Dimensional Analyses of Cell Morphology	4
1.2 Parvalbumin-positive (PV+) interneurons	
1.3 High-Dimensional Workflows for Cell Morphologies	
Chapter 2. Parvalbumin-Positive Interneuron Morphologies in the Mouse Primary Vis Somatosensory Cortices	sual and 15
Abstract	
2.1 Introduction	17
2.2 Methods	
Animals	19
Assessing Label Quality	
Datasets	
Image Processing and Analysis	
Cleaning and Extracting Cell Morphology Data	
Converting to Standard Cortical Thickness	
Removing Non-Cell Shaped Particles	
Identifying Cortical Layers	
Cluster Analysis of Cell Morphologies	
Preparing the Data Frames of Features and Cell Identifiers	
Selecting the number of clusters	
Robust-Sparse K-Means Clustering	29
Dimensionality Reduction and Visualization	
Characterizing Morphological Clusters	
Cluster Analysis of the PV+ Cell Dataset	
Laminar Analysis of PV+ Morphology Clusters	
Identifying Laminar Patterns	
Cluster Analysis of the Whole-Cell Filled Dataset	
Clustering Cell Body Morphologies	

Clustering Dendritic Morphologies	33
Cluster Analysis of PV+ Associated Genes Dataset	34
Statistical Analyses	35
2.3 Results	35
Assessing ISH Label Quality Among Animals	35
Morphological Clustering of PV+ interneuron cell bodies	42
Visualizing PV+ Morphology Clusters	45
Cortical Area Distribution of PV+ Morphologies	52
Laminar Distribution of Morphology Clusters in S1 and V1	70
Comparing Cell Body Morphology and Dendritic Morphology	77
Cluster Analysis of Dendritic Morphologies	79
Comparing the Morphology of PV+ and PV+ Associated Gene Cells	90
2.4 Discussion	112
Chapter 3. General Discussion	
Aligning Single-Cell Transcriptomics and Morphology Data	118
Host Factors can Influence Cell Morphology Features	121
Conclusions and Future Research Directions	124
Chapter 4. References	126

List of Figures and Tables

Figure 1 - Fringe patterns for the first 21 Zernike moments	7
Figure 2 - Data quality assessment among animals	36
Figure 3 - Defining cortical layer boundaries in S1 and V1	38
Figure 4 - Size Distributions for PV+ interneurons in S1 and V1	40
Figure 5 - Estimation statistics comparing variation in PV+ soma size among cortical layers and areas.	42
Figure 6 - Unsupervised clustering on PV+ Interneuron morphologies.	44
Figure 7 - Visualizing RSKC morphology clusters for PV+ Cells	47
Figure 8 - Characterizing RSKC morphology clusters	49
Figure 9 - Characterizing size divisions and shape motifs among clusters	53
Figure 10 - Soma Area for Each Morphology Cluster	55
Figure 11 - Cell eccentricity among morphology clusters of PV+ interneurons	57
Figure 12 - Zernike Moment 2_2 values among morphology clusters of PV+ interneurons	59
Figure 13 - Zernike Moment 2_0 values among morphology clusters of PV+ interneurons	61
Figure 14 - Hu Moment 1 values among morphology clusters of PV+ interneurons	63
Figure 15 - Hu Moment 3 values among morphology clusters of PV+ interneurons	65
Figure 16 - Zernike Moment 0_0 values among morphology clusters of PV+ interneurons	67
Figure 17 - Zernike Moment 4_0 values among morphology clusters of PV+ interneurons	69
Figure 18 - Laminar distributions of PV+ Morphology Clusters in S1 and V1	71
Figure 19 - Proportions of PV+ cells by cortical layer for each morphology cluster.	72
Figure 20 - Characterizing the laminar distributions of PV+ cell morphologies.	76
Figure 21 - Filled cells have comparable size and shape parameters to ISH-labelled cells	78
Figure 22 - Clustering dendritic morphology features for PV+ filled cells	80
Figure 23 - Cell Reconstructions of Filled cells (Dendritic Morphologies A and B)	81

Figure 24 - Cell Reconstructions of Filled cells (Dendritic Morphology C)	83
Figure 25 - Clustering cell body morphologies of PV+ filled and ISH-labelled cells	85
Figure 26 - Morphological properties of filled cells integrate along the morphology gradients established by ISH-labelled cells	87
Figure 27 - Morphological phenotypes for clusters of PV+ ISH cells and Whole-Filled cells	88
Figure 28 - Correspondence among cell dendritic and cell body morphologies	90
Figure 29 - ISH samples for PV+ and Associated Genes	91
Figure 30 - Clustering PV+ gene morphologies in primary somatosensory cortex (S1).	93
Figure 31 - Clustering PV+ gene morphologies in primary visual cortex (V1)	95
Figure 32 - Visualising PV+ gene morphologies for S1	97
Figure 33 - Visualising PV+ gene morphologies for V1	99
Figure 34 - Characterizing morphology clusters for PV genes in S1	102
Figure 35 - Characterizing morphology clusters for PV+ genes in V1	104
Figure 36 - Expected proportions of PV-associated genes for S1 and V1.	106
Figure 37 - Size divisions and shape motifs for PV+ gene morphology clusters in S1	107
Figure 38 - Size divisions and shape motifs for PV+ gene morphology clusters in V1	108
Figure 39 - Laminar profiles for PV-associated genes : genes by morphology cluster for S1	111
Figure 40 - Laminar profiles for PV-associated genes : genes by morphology cluster for V1.	112
Table 1 - Dataset composition summary	21
Table 2 - Cell Counts for Cell Morphology Datasets Prepared in this Study.	25
Table 3 - 97 Morphometric Features for Cluster Analysis	27
Table 4 - Dendritic Morphology Features for Whole-Cell Filled Cells.	33
Table 5 - Summary Statistics for PV+ Cell Size by Cortical Layer for S1 and V1	41

List of Abbreviations

S1	Primary Somatosensory Cortex
V1	Primary Visual Cortex
ISH	In Situ Hybridization
PV	Parvalbumin
PV+	Parvalbumin-Positive
PNN	Perineuronal Net
E/I Balance	Excitation/Inhibition Balance
Т-Туре	Transcriptomic Type
P56	Postnatal Day 56
FIJI	Image J (FIJI Is Just ImageJ)
RGB	Red-Green-Blue
RSKC	Robust, Sparse K-Means Clustering
k	Number of clusters
t-SNE	t-Distributed Stochastic Neighbour Embedding
denSNE	Density Preserving t-distributed stochastic neighbour embedding
scRNASeq	single-cell RNA Sequencing
Calb1	Calbindin 1
Bdnf	Brain-Derived Neurotrophic Factor
Cacna2d3	Calcium Channel, Voltage-Dependent, alpha2/delta subunit 3
Col25a1	Collagen, Type XXV, alpha 1
Etv1	ETS variant 1
Fxyd6	FXYD domain-containing ion transcport regulator 6
Tac1	Tachykinin 1
Thsd7a	Thrombospondin, Type I, Domain Containing 7A
Tpbg	Trophoblast Glycoprotein
Kit	Kit Oncogene

Krt12	Keratin 12
Snca	Synuclein, Alpha
Nfib	Nuclear Factor I/B
Ndnf	Neuron-derived neurotrophic factor
Cux2	Cut-like Homeobox 2
Rorb	RAR related orphan receptor B
Deptor	DEP domain-containing mTOR-interacting protein
Foxp2	Forkhead Box Protein 2
Ctgf	Connective Tissue Growth Factor

Declaration of Academic Achievement

Chapter 2 of this thesis is in preparation for publication. This study was a collaboration between Dr. Kathryn Murphy and myself. I led the collection, analysis, and interpretation of data, and the writing for the manuscript. Assistance with analyses were provided by Leanne Monteiro.

Chapter 1. General Introduction

Defining cell types is a persistent, central question in cell biology. Having motivated decades of research, classifying cells into cell types has encompassed both classic and modern approaches. (Dance, 2024; Mao and Staiger, 2024). In the brain, classic approaches have relied heavily on describing cells in terms of their morphologies(Cajal, 1995; Freemon, 1996), whereas modern approaches expand on these classic morphology descriptions, integrating cell morphology with additional modalities such as their locations in situ, transcriptomic profiles, or electrophysiology. (Tasic et al., 2016; Zeng and Sanes, 2017; Gouwens et al., 2019; Peng et al., 2021; Zeng, 2022; Mao and Staiger, 2024). The integrative approach taken by modern techniques highlights a central problem in cell type classification studies: specifying the set of characteristics that reliably distinguish among different groups of cells. This is crucial in defining cell types that are both holistic, and biologically meaningful (Zeng, 2022). However, classic approaches have inferred differences in function from differences in cell morphology, and have served as the foundation for cell classification (Cajal, 1995).

Notably, work by Santiago Ramon Y Cajal defined cell types based on their appearance. This seminal work was a comprehensive morphological profile of several brain and spinal cord structures. The rich qualitative descriptions of the shapes served as a crucial foundation for parameterizing the diversity of brain cells in terms of their appearance. Across several cortical areas, Cajal provided elaborate descriptions of the cytoarchitecture of the layers, defining in precise terms the differences in cell morphology that distinguish the cortical areas from each other. For example, in the visual cortex, Cajal describes the morphologies of giant pyramidal cells, a morphological subtype of pyramidal cells :

"Their pyramidal cell body has an unusually wide base to its height. Second, an even more characteristic feature is the two, three, or more thick dendrites that arise from the base and extend strictly parallel to the cortical surface. They sometimes may be even

longer than the apical dendrite and they branch along the way, forming essentially tangential bundles."

This description of giant pyramidal cells in terms of their cell body and dendritic structure is strikingly different from that of "polygonal" cells:

"Other neurons include large to medium-sized polygonal cells with a thick, more or less tangential axon. The latter generates an arborization of very long oblique to tangential branches whose terminal ramifications appear to form around giant pyramidal cells. These nests are similar to those described earlier for the cortex in general."

From these structural differences, Cajal also posited that different cell types perform different functions because of their different structures. He also raised the notion that the different structures themselves are highly optimized for the functions they perform (Cajal, 1995). Cajal's work has established morphology as a classic and central hallmark of cellular diversity in the brain. These analyses have helped establish cells as the basic units of the nervous system and profiled the properties of single cells as a guiding principle in reconciling cellular diversity to understand cellular identities in the brain.

Newer approaches leveraging single-cell transcriptomics have revolutionized how we define and understand cell types by grouping single cells into classes by their transcriptomic landscapes (Tasic et al., 2016; Zeng and Sanes, 2017). This approach leverages high-dimensional transcriptomics data, where the expression of thousands of genes can be quantified for a single cell. These sparse data matrices can be analyzed to simultaneously identify and group common transcriptomic patterns into cell types (Haque et al., 2017). These transcriptomically defined cell types (T-types) identify common molecular signatures among different subsets of glutamatergic, GABAergic, and non-neuronal cells (Tasic et al., 2018). Because they are defined in terms of patterns of gene expression, T-types capture additional cellular diversity that was previously

inaccessible. This additional cellular diversity has also been able to distinguish among cortical areas, with V1 being most distinct from other areas (Jorstad et al., 2023).

Cajal's seminal work has identified elaborate shapes and described morphological differences that have not been robustly quantified or accessed through quantitative tools. T-types have proven to be a critical and pivotal paradigm shift in accessing and reconciling cellular diversity. However, insights into morphological complexity still need to be elucidated. To achieve this, high-dimensional, cell morphology data need to be obtained, and methods to analyze these data need to be comparable to methods applied to single-cell transcriptomics datasets (Colombo et al., 2022; Kamat et al., 2024). This will provide a comparable morphological profile of sufficient depth to capture subtle but characteristic morphological differences among different cell types.

1.1 High-Dimensional Analyses of Cell Morphology

Approaches to defining and classifying cell morphologies have mainly remained qualitative and, consequently, have yet to be able to capture morphology on a scale comparable to the molecular diversity uncovered in single-cell transcriptomics studies. From a classical perspective, Cajal's elaborate drawings capture more than the sizes of single cells, and his descriptions of cell shapes allude to subtle, nuanced shape characteristics that could only be described in qualitative terms. With his description of size and shape, he defined cell morphologies found across cortical layers and brain areas (Cajal 1995). However, approaches to quantifying those shapes remained elusive, in favor of semi-quantitative approaches combining readily measurable parameters like cell size (Lewis and Lund, 1990; Group et al., 2008; Lingley et al., 2018). Despite his work revealing tremendous morphological diversity and complexity among cells in the brain, morphologically-defined cells are typically defined in terms of size and

a limited range of shape parameters. Single-cell transcriptomics, conversely, characterizes single cells with several features. These datasets are often very large, consisting of gene expression for thousands of genes, quantified for hundreds of thousands of cells (Chen et al., 2019). By leveraging high-dimensional data, modern single-cell transcriptomics has captured the fine molecular diversity among cells by leveraging the gene expression levels of thousands of genes for each cell (Haque et al., 2017). To address the problems with defining and classifying cell morphologies, a comparable data structure is needed to characterize morphology in terms of several parameters for a sufficiently large number of individual cells. The selected parameters must preserve cell morphology's salient and defining aspects to provide structurally meaningful information (Carpenter et al., 2006).

Cell morphologies have garnered much attention, but in the brain, this has remained largely unexplored across most cell types, with the majority of studies having been conducted in microglia. As brain-resident immune cells, microglia are dynamic surveillants of their local microenvironments, and their morphologies change in response to their functional state(Nimmerjahn et al., 2005). Consequently, microglial morphologies are both dynamic and highly variable and characterizing their morphologies has faced comparable challenges when working with classical neuroanatomical data(Guillot-Sestier et al., 2021). This has motivated the use of high-dimensional data, and unsupervised machine or statistical learning approaches to identify patterns among microglial morphologies (Leyh et al., 2021; Kim et al., 2024). This paradigm shift has identified robust morphological sex differences in microglia in response to aging and in animal models of neurodegeneration(Guillot-Sestier et al., 2021; Colombo et al., 2022). These morphological sex differences were only readily discernible from high-dimensional analyses, and underscore the structural and functional correlates of increased microglial

activation and reactivity in females, which serves as preliminary causal evidence for sex differences in the prevalence of Alzheimer's disease in females(Colombo et al., 2022). The work done to characterize and ultimately describe progressive morphological changes in microglia clearly demonstrate the value, or applying suitable analytical approaches to cell morphology, in capturing several salient properties from a large number of morphology features. These insights serve as a critical motivation to explore other brain-resident cell types and characterize their morphological properties to begin to understand what and how changes in cell morphology are associated with development and disease.

Shape metrics have been developed to quantify various objects and are commonly used in optics and computer vision. For example, Zernike moments are a set of radial trigonometric functions in polar coordinates that can quantify a range of shapes. First developed by Fritz Zernike, these moments have been used in optics as a basis for quantifying and subsequently correcting aberrations in lenses and other optical systems (Hague et al., 2017; Niu and Tian, 2022). Since their inception, Zernike moments have been applied broadly across several clinical and biological fields. Their principal application in ophthalmology has been to quantify and correct corneal aberrations. In image analysis and computer vision, Zernike moments have often been applied as kernels that filter shape properties in the underlying image that align with the distributions of pixel intensities described by their given Zernike polynomials (Teague, 1980; Boland et al., 1998; Pincus and Theriot, 2007; Alizadeh et al., 2016). As seen in Figure 1, Zernike moments capture different shape characteristics associated with single cells. Lower-order Zernike moments (shown in the upper rows of Figure 1) capture shape characteristics such as circularity, elongation, and the major axis of elongation that are readily discernible by the eye. In contrast, higher-order Zernike moments (shown in the lower rows of

Figure 1 capture subtle variations in symmetry and features, such as protrusions or bumps around the edge of an object. Applying Zernike moments in image analysis and computer vision established that shape is a high-dimensional characteristic of an object and can be decomposed into subtle, latent components. For the analysis of cell morphology, this suggests that subtle characteristics of cellular shapes can serve as distinguishing factors among different morphological types of cells (Prasad and Alizadeh, 2019; Hagemann et al., 2021).

Zernike moments can reduce the dimensionality of images while still being able to completely reconstruct objects in the image from sufficiently many Zernike moments (Pincus and Theriot, 2007). In this way, Zernike moments are a compact basis system that captures cell shapes and describes nuanced but marked shape characteristics that are otherwise lost with reductive parameters (Pincus and Theriot, 2007). Furthermore, Zernike moments have been described as orthogonal over the unit circle, eliminating redundancy (Niu and Tian, 2022). Zernike moments are also favored for their size, scale and rotation invariance. These properties allow Zernike moments to quantify elaborate shape characteristics, like symmetry, which are not readily described in terms of basic shape parameters such as circularity and elongation (Teague, 1980). These properties make Zernike moments amenable to quantifying shapes and providing high-dimensional shape data for single cells.



Figure 1 - Fringe patterns for the first 21 Zernike moments arranged in rows by radial order, and columns by azimuthal order. Higher radial orders are towards the bottom of the pyramid. Image from : Nschloe, CC BY-SA 4.0 <https://creativecommons.org/licenses/by-sa/4.0>, via Wikimedia Commons from : https://commons.wikimedia.org/wiki/File:Zernike polynomials with read-blue cmap.png

Zernike moments have been used to quantify cell morphology. Several groups have leveraged Zernike moments' sensitivity and optimal properties to distinguish between healthy and cancerous cells across various cell and tissue types. For example, Zernike moments captured progressive morphological changes with advancing cancer stages as a morphological correlate of disease state (Tahmasbi et al., 2011; Alizadeh et al., 2016; Prasad and Alizadeh, 2019). These cell-type classification studies highlight the importance of subtle shape characteristics and information value in differentiating cell states. Further, Zernike moments have been applied to neuronal morphologies to identify morphological patterns associated with neurodegeneration. Hagemann and colleagues have found that Zernike moments quantify morphological differences related to the progressive neurodegenerative phenotypes in human samples and rodent models of ALS (2021). This supports the argument that subtle morphological characteristics can have a large-scale functional impact. By mapping morphological change throughout disease progression, Zernike moments are thought to give a phenotypic readout of the cell state as degeneration progresses. These findings reinforce the importance of subtle shape characteristics and that they can carry relevant biological information about the cells under study (Hu and Murphy, 2004; Hagemann et al., 2021).

1.2 Parvalbumin-positive (PV+) interneurons

Approximately 80% of cortical neurons are excitatory, and 20% are inhibitory (Rubenstein and Merzenich, 2003). Inhibitory interneurons are critical in gating information and modulating activity through cortical circuits (Kepecs and Fishell, 2014). PV+ interneurons are the most abundant class of inhibitory interneurons in the cortex. They derive from the medial ganglionic eminence and are involved in various neural functions (Druga et al., 2023). PV+ interneurons are abundant in primary sensory cortical areas, and their maturation is linked to the timing of critical periods for experience-dependent plasticity in both primary visual and somatosensory cortices (Morales et al., 2002; Lo et al., 2017). As integral regulators of experience-dependent plasticity, PV+ interneurons contribute to regulating experience-dependent plasticity, as their high-frequency firing rates act as timestamps , identifying coherent activity among cortical circuits (Cardin et al., 2009) . This refinement tunes circuit activity to produce the precise topographical maps involved in processing and spatially resolving sensory stimuli (Hensch et al., 1998; Morales et al., 2002; Levelt and Hübener, 2012; Lo et al., 2017). PV+ interneurons have uniquely structured dendritic and axonal arbors, allowing a single PV+ interneuron to make pre- and post-synaptic contacts with many local excitatory neurons. This allows a single PV+ interneuron to control and respond to inputs from several glutamatergic neurons (Fries et al., 2007; Levelt and Hübener, 2012). Furthermore, electrical and chemical synaptic coupling among PV+ interneurons forms larger networks where PV+ interneurons not only gate excitatory activity but also synchronize large pools of glutamatergic neurons and organize their synchronous firing to produce gamma oscillations, a high-frequency oscillation (30-100 Hz) that is involved in sensory processing and learning from sensory input (Galarreta and Hestrin, 2002). PV+ interneurons are necessary for establishing and maintaining the gamma rhythm because of the extensive networks of neural circuits they form. Specifically, inactivating PV+ interneurons resulted in loss of gamma oscillations, whereas PV+ interneuron activation orchestrated synchronous gamma frequency activity in pyramidal cells, indicating that PV+ interneuron activity and their network structure carries a large influence in directing synchronous, high-frequency activity (Sohal et al., 2009).

Moreover, in the barrel cortex, gamma frequency activity of PV+ interneurons gates excitatory activity, refining and modulating sensory processing (Cardin et al., 2009). It is also through this network-level structure that PV+ interneurons contribute to the large-scale refinement of circuits in primary sensory cortical areas associated with critical period closure through filtering sensory information in a spike-timing dependent manner (Galarreta and Hestrin, 2002; Fries et al., 2007; Levelt and Hübener, 2012). PV+ interneuron recruitment is critical to perceptual coding in V1, sharpening feature selectivity and discrimination between different visual stimuli (Lee et al., 2012). Work by Lee and colleagues demonstrates that PV+

interneurons, as a specific subtype of GABAergic interneuron, play crucial roles in visual processing and perception.

PV+ interneurons are preferentially enwrapped by perineuronal nets (PNNS), a lattice-like networks of extracellular matrix proteins surrounding the cell body (Seeger et al., 1994; Sorg et al., 2016; Fawcett et al., 2019; Carceller et al., 2020, 2023). These structures develop around the cell body coincident with the closure of the critical period and act as a brake on plasticity by restricting the lateral mobility of receptors and excessive synaptic connections (Sorg et al., 2016). The link proteins in the PNN connect the rigid lattice to the cell body, so the PNN can exert mechanical forces that influence the shape of the cell body (Yue, 2014). However, previous work investigating the influence of PNNs on cell morphology has yet to be conclusive in describing the impact of the PNN on cell body morphology (Sigal et al., 2019). In addition to the potential influence of PNNs on PV+ interneuron morphologies, PNNs are a transiently stable structure with variable density. They can be remodeled in response to sensory experience, attention and learning (Sanchez et al., 2023). In sensory areas, their presence around PV+ interneurons has often been described as activity-dependent, and their presence is thought to confer a neuroprotective effect to prevent metabolic stress and act as a cation sink (Brückner et al., 1993). Together, the role of the PNN suggests that PV+ interneuron morphologies are dynamic, and variations in their morphologies occur during development, aging, and disease could be linked to the presence or absence of PNNs.

PV+ interneurons are an important player in maintaining the structure and function of cortical circuits. Consequently, their dysfunction at both structural and functional levels are broadly implicated in several brain disorders. From a structural perspective, PV+ interneurons

form extensive synaptic contacts with local pyramidal cells. Their synaptic connectivity is highly organized, and refined in an activity-dependent manner (Levelt and Hübener, 2012). Conversely, in animal models of autism spectrum disorder, PV+ interneurons have excessive dendritic arborization, which is typically associated with more juvenile developmental stages (Janickova et al., 2020). Additionally, perineuronal nets (PNNs), an extracellular matrix preferentially associated with PV+ cells, plays a role in maintaining the structural and functional integrity of PV+ interneurons(Sorg et al., 2016; Fawcett et al., 2019). Degradation, and loss of PNNs has been linked to loss of PV+ inhibition across the brain, combined with persistent, dystrophic PV+ cell morphologies (Yang et al., 2023). However, structural abnormalities in PV+ interneurons have often been elucidated alongside functional abnormalities. PV+ interneurons have a higher excitatory drive, and higher frequency firing rates compared to other interneurons(Gulyás et al., 1999). Together with their elaborate synaptic connectivity, these functional factors establish PV+ interneurons as centrally important regulators of excitation/inhibition (E/I) balance, and coordinators for gamma oscillations in the cortex, which are involved in sensory perception, and memory(Fries et al., 2007; Cardin et al., 2009). In several brain disorders, disruptions to gamma rhythms accompany cognitive deficits in schizophrenia (Marín, 2024), runaway excitation in epilepsy and seizure disorders and memory deficits in Alzheimer's disease (Leitch 2024; Smeralda et al., 2024). Taken together, structural and functional factors contribute to the common patterns of PV+ interneuron dysfunction seen across several brain disorders.

PV+ interneurons are classically identified from their immunoreactivity to the PV protein, thereby labelling all PV+ interneurons homogeneously, often describing them as a single entity (Celio and Heizmann, 1981; Lewis and Lund, 1990). The impact of PV+ morphology function needs to be better understood because, despite their involvement in a broad range of functions,

PV+ interneurons are homogeneously classified (Celio and Heizmann, 1981; Kepecs and Fishell, 2014; Song et al., 2023). Subsequently, analyses of PV+ interneuron morphologies from classic anatomical modalities have mainly remained descriptive, with qualitative descriptions of shape and limited, quantitative approaches using cell size (Blümcke et al., 1990; Brederode et al., 1990; Group et al., 2008; Lingley et al., 2018). This has led to a primarily descriptive characterization of PV+ cells (Group et al., 2008).

Conversely, modern single-cell transcriptomics studies have further elucidated the extensive diversity among PV+ interneurons. A single, classically homogeneous cell class has been stratified into ten molecularly distinct subtypes (Tasic et al., 2016; Gouwens et al., 2019). Single-cell transcriptomics has elucidated molecularly distinct subclasses of PV+ interneurons that are defined by the genes they express by quantifying several genes per cell. This enables a more precise definition of PV+ cells beyond a single marker gene or protein (Tasic et al., 2016; Mayer et al., 2018; Chartrand et al., 2023). This also enables a more nuanced definition of different types of PV+ cells defined by their different transcriptomic landscapes (Tasic et al., 2016, 2018; Gouwens et al., 2019, 2020). These transcriptomic types, or "T-types," have captured unique electrophysiological properties and laminar distribution patterns specific to specific subsets of PV+ interneurons (Tasic et al., 2018)). T-types have revealed novel insights into cell types in the brain. Notably, T-types have identified V1-specific cell types, suggesting that the transcriptomic landscapes of cortical areas give rise to different populations or proportions of cell types that contribute to their functions (Jorstad et al., 2023).

Taken together, single-cell transcriptomics reveal additional molecular complexity beyond the singular marker genes and proteins used to define cell classes in classic neuroanatomical approaches. This supports the notion that morphological complexity exists beyond the simple,

semi-quantitative descriptive approaches used to characterize morphologically distinct cell types using classical neuroanatomical data. It is unknown whether morphological differences can recapitulate the unique transcriptomic organization of different cortical areas. This would require a comparable analytical framework that aligns cell morphology data with single-cell transcriptomics.

1.3 High-Dimensional Workflows for Cell Morphologies

Differences in cell morphologies have been the classic, longstanding factors contributing to reconciling cellular diversity in the brain. These differences have served as crucial in the first definitions of cell types and allude to subtle differences that define entire cell types. However, analyses of cell morphologies have yet to keep pace with the high-dimensional analysis approaches applied in single-cell transcriptomics studies. With a lack of appropriate, high-dimensional quantification, subtle but characteristic morphological differences still need to be discovered. High-dimensional, cell morphology data need to be acquired to address this knowledge gap, and a suitable, data-driven analytical framework needs to be leveraged to align morphological data with single-cell transcriptomics data.

Primary research objective: to develop a data-driven, high-dimensional analysis pipeline for a cell morphology dataset to characterize neuronal morphologies in the cortex.

Specific Research Aims :

- To acquire a large cell morphology dataset comparable in size and dimension to single-cell transcriptomics datasets, I will use neuroanatomical images of labelled neurons in the visual and somatosensory sensory cortex of the mouse and a modern image analysis toolbox to quantify many size and shape parameters.
- To develop a data-driven analysis pipeline to quantify neuronal morphologies, I will use the size and shape parameters as the input to a high-dimensional analysis pipeline to identify clusters of morphological subtypes of PV+ neurons and explore the features and laminar locations of the different subtypes.
- 3. To determine whether differences in cell morphology properties recapitulate the differences seen among different PV+ transcriptomic types, I will apply my morphology analysis pipeline to a series of ISH images labelling PV and genes associated with PV+ T-types. This will compare how morphologies for cells labelled for different genes align with both PV+ interneurons and whether the combinations of genes across these morphologies align with gene distributions seen in PV+ T- types.

Chapter 2. Parvalbumin-Positive Interneuron Morphologies in the Mouse Primary Visual and Somatosensory Cortices

Abstract

Classical neuroanatomical approaches have presented Parvalbumin-positive (PV+) interneurons in a homogeneous manner, with descriptive, semi quantitative approaches to defining their morphologies. Consequently, cell morphology data do not align with the high-dimensional structures used in single-cell transcriptions, which have identified additional molecular complexity among PV+ cells. The contrast in data structures between these two modalities underscored the importance of treating cell morphologies in a comparably robust, and data-driven manner. We quantified 97 morphometric features for several cell morphology datasets consisting of tens of thousands of cells, and identified more than 10 different morphologically defined PV+ cell types, which is comparable to the diversity summarized in single-cell transcriptomic taxonomies. We connected high-dimensional cell morphologies to patterns in dendritic arbors from whole-cell fill data, and also explored morphological differences across a range of cells labelled for genes associated with PV+ T-types. Our approach captured subtle, but characteristic morphological differences among PV+ interneurons identified from classic neuroanatomical data Our results motivate further investigations into the fine structural details of PV+ interneurons to elucidate and characterize the structural correlates of PV+ dysfunction in disease, or tracking morphological changes in development and plasticity.

2.1 Introduction

Parvalbumin-positive (PV+) interneurons are the brain's most abundant subclass of GABAergic interneurons. Recent single-cell RNAseg studies have identified multiple sub-categories of molecularly distinct PV+ neurons in the cortex. In contrast, PV+ interneurons are often categorized anatomically as a homogeneous group of neurons based on their immunoreactivity to the PV protein (Celio and Heizmann, 1981; Brederode et al., 1990; Rio et al., 1994) despite their unique and extensive dendritic and axonal arbor morphology differences (i.e., basket and chandelier cells) (Lewis and Lund, 1990; Wang et al., 2002; Jiang et al., 2015). (Galarreta and Hestrin, 2002). This gap between the single-cell RNA and morphological definitions of this important inhibitory interneuron class is partly due to the qualitative and semi-quantitative approaches used to study the morphology and anatomical location of PV+ neurons. PV+ neurons are intimately involved in critical period closure, spike-time-dependent plasticity, and regulation of the excitatory-inhibitory balance (Levelt and Hübener, 2012). Furthermore, changes to PV+ neurons are linked to a variety of neurological and psychiatric disorders including epilepsy, schizophrenia, autism as well as neurodegenerative disorders (Wöhr et al., 2015; Janickova et al., 2020; Ruden et al., 2021; Leitch, 2024) Here, we present a new approach to studying PV+ morphology that uses multiple measurements of morphological features and high-dimensional analyses to quantitatively categorize the organization of these cells in the mouse somatosensory and visual cortex.

PV+ interneurons are implicated in a broad range of complex processes. As fast-spiking interneurons, they deliver and maintain highly temporally resolved inhibitory tone, to pyramidal cells. In this way, PV+ interneurons act as cellular timekeepers, gating the flow of information through cortical circuits (Cardin et al., 2009). Forming elaborate networks of synaptic

connections, PV+ interneurons also control the synchronous firing of excitatory neurons, and are necessary for setting and maintaining gamma oscillations across larger cortical domains (Sohal et al., 2009; Levelt and Hübener, 2012). This establishes a centrally important role for PV+ cells in modulating sensory cortical circuitry in response to experience. Structurally, PV+ interneurons also influence plasticity through structural interactions through both myelin and associations with the extracellular matrix. Myelin surrounding cortical PV+ interneurons forms shorter, "patchy" internodes that are more dynamic in response to sensory experience, and remodel readily in response to abnormal experience (Micheva et al., 2018; Yang et al., 2020). PV+ interneurons also associate with elaborate extracellular matrix structures, perineuronal nets, that act as a structural brake to plasticity, by restricting the lateral mobility of receptors, and anchoring synaptic connections (Levelt and Hübener, 2012; Sorg et al., 2016).

Cell morphology has been the classic approach to reconciling cellular diversity in the brain. From Cajal's foundational work in the nineteenth century, there have been a range of extensive shapes that have been described in elaborate but qualitative terms. (Cajal, 1995). This has persisted in classic anatomical investigations of PV+ cells, where their morphologies were defined in terms of descriptions of their appearance together with their laminar locations in the cortex (Mikkonen et al., 1997; Zhu et al., 2018; Song et al., 2023). This gap in knowledge stems from a widening technical gap surrounding how to analyze single cells in anatomical techniques. Emerging approaches to quantifying and studying cell morphologies have identified subtle changes in cell shape that are imperceptible in conventional image analyses, carry salient, biologically relevant information about cell states in disease and quantitatively define disease progression at a single-cell level (Alizadeh et al., 2016; Hagemann et al., 2021). As single-cell transcriptomics have clearly demonstrated specific transcriptomic types of PV+ interneurons are

uniquely implicated in neurodegenerative disorders (Gabitto et al., 2024), it remains unknown whether there are specific PV+ interneuron morphologies that are indicative of an increased susceptibility to disease. To obtain these insights, comparable approaches to data acquisition and analysis are needed to align classic neuroanatomical data with single-cell transcriptomic approaches (Gliko et al., 2024).

In this study, we developed a sensitive, data-driven approach to define and characterize the cell morphologies of PV+ interneurons. We also applied an unsupervised, standardised approach to mapping the cortical layers. We robustly interrogate the morphology of the cell body, and explore the patterns and relationships that may be present with dendritic structure using an integrated dataset of ISH-labelled, and whole-filled cells. Finally, we explore the relationship between cell morphologies and transcriptomics by applying our pipeline to a collection of genes associated with PV+ transcriptomic types.

2.2 Methods

Animals

All anatomical data were retrieved from the Allen Brain Institute mouse ISH database (https://mouse.brain-map.org/) (Lein et al., 2007). The images in this database were from male C57BL6 mice at postnatal day 56 (P56). For the current study, PV+ data was collected from 6 mice but in the quality control step (described below), the images from one animal did not meet the criteria for inclusion, so the animal's data were not included in subsequent analyses.

Assessing Label Quality

For the dataset of PV+ cells, we assessed the label quality among the 6 animals. A single sample colorimetric ISH from S1 and V1 were processed and analyzed in FIJI. Images were first converted to grayscale, followed by a 50-pixel rolling ball background subtraction, before being binarized. In these processed images, the integrated density was obtained for all cells in the image. Estimation statistics were performed comparing the unpaired median differences for integrated density values among the six animals (using the dabestr package - version 0.3.0 in R) (Ho et al., 2019).

Datasets

A summary of the composition of each dataset is shown in Table 1.

Table 1 - Dataset composition summary

Dataset	Number of Animals	S1 cell count	V1 cell count	PV	Ndnf	Cux2	Rorb	Deptor	Foxp2	Ctgf	Cacna2d3	Calb1	Bdnf	Col25a1	Etv1	Fxyd6	Kit	Krt12	Nfib	Snca	Tac1	Thsd7a	Tpbg
PV+ cells	5	8823	5451	~																			
Laminar Marker Genes	6	1383	645		r	r	r	r	~	~													
Whole, Filled Cells	3	0	36	r																			
PV+ Cells and Associated Genes	13	2980	2162	r							r	r	r	r	7	r	~	r	r	r	~	r	r

Four datasets were constructed.

- 1. A dataset of ISH-labeled PV+ cells in S1 and V1.
- 2. A dataset of ISH-labeled cells for 6 laminar marker genes in S1 and V1.
- 3. A dataset of whole-cell filled PV+ cells in V1.
- 4. A dataset of ISH-labeled cells for 13 genes associated with PV+ transcriptomic types.

For the first three datasets, the colorimetric ISH and processed expression images were retrieved at full resolution (1 pixel = 1 μ m). The expression images were preprocessed by the Allen Brain Institute to remove the background (black) and color-code the cells by their integrated optical density, reflecting the density of the ISH labelling (blue = minimum integrated density, red = maximum integrated).

The first dataset of PV+ cells used 11 images from sagittal sections containing S1 and V1 for each of the 5 animals. The reference atlas (<u>https://atlas.brain-map.org/</u>) was used to identify the middle of S1 & V1, and 5 images medial and lateral to that image were added for a total of 11 images. Each sagittal image contained both S1 and V1. Sampling boxes from the pial surface to the white matter ($567\mu m X 1525\mu m$) for the colorimetric ISH and expression images were extracted from S1 and V1 using Adobe Photoshop 2024. Thus, S1 and V1 had a total of 11 ISH and expression image sampling regions extracted for analysis.

The second dataset used images from 6 laminar marker genes associated with a specific cortical layer (L1 - Ndnf, L2/3/4 - Cux2, L4 - Rorb, L5 - Deptor, L6A - Foxp2, L6B - Ctgf). One ISH and the corresponding expression image were taken from the middle of S1 and V1 for each gene, and a sampling box, as described above, was extracted from each cortical area.
The third dataset used images from S1 and V1 for 13 genes (Bdnf, Calb1, Cacna2d3, Col25a1, Etv1, Fxyd6, Kit, Krt12, Nfib, Snca, Tac1, Thsd7a, Tpbg) associated with PV+ T-types (Tasic et al., 2016). Among the marker gene list for PV+ T-types in Tasic et al., 2016, these 13 genes had good quality labelling in the ISH database. Furthermore, these genes also represent the multimodal PV+ transcriptomic types defined in Gouwens et al., 2020. Again, one ISH and the corresponding expression image from the middle of S1 and V1 were used for each of the 13 genes.

The fourth dataset used 36 whole-cell filled PV+ cells from the Allen Cell Types Database (https://celltypes.brain-map.org/data) (Hodge et al., 2019). The cells used for analysis were from the 5 PV+ transgenic lines (Htr3a-Cre_NO152|Pvalb-T2A-Dre, Pvalb-IRES2-Cre-Neo, Pvalb-IRES-CRE, Pvalb-T2A-CreERT2, PVALB-T2A-FIPO|Vipr2-IRES2-CRE) and were categorized as aspiny neurons with complete reconstructions. Cells were sampled from V1 (VISp) in layers $\frac{2}{3}$ to 6A (no cells were filled from layers 1 or 6B). Mice from both sexes aged P56 ± 3 days were in the database; however, the metadata did not specify the sex of individual mice (Hodge et al., 2019).

Since the whole-cell filled dataset did not include expression images, an image of each soma had to be extracted for analysis. Images of the 36 cells were downloaded from the database and opened in Adobe Photoshop 2024. The magic wand tool (default tolerance parameters) was used to identify and copy the cell body into a "simulated cortex". Two of the 36 cells had a dark halo around the cell body, so the levels tool was used to increase the contrast, and the soma was extracted with the magic wand. The simulated cortex contained the 36 cells at their laminar location identified in the Allen Cell Types Database.

Image Processing and Analysis

A custom analysis pipeline was created in CellProfiler (Carpenter et al., 2006; Stirling et al., 2021) to quantify the morphology of the cells from the PV+, PV+ associated genes and whole-cell filled datasets. The first step converted the RGB expression images to gravscale using the ColorToGray Module, where the red, blue, and green channels were given equal weights of 1. The grayscale images were binarized using the Threshold module and the global threshold strategy because the expression images had uniform backgrounds. Otsu's thresholding method and two-class thresholding were used to distinguish labelled cells and particles from the background. The thresholding was not smoothed and was not corrected. Next, the watershed algorithm was applied to segment the cells. The default watershed parameters were used (segmentation generated from a distance between adjacent objects, footprint of 8, and no downsampling). Finally, measurements of the cell morphologies were obtained using the MeasureObjectSizeShape module. This produced a data matrix where each row corresponds to a cell, and columns contain cell coordinates, metadata, and morphometric features. Also, an image of all the analyzed cells was generated. The unprocessed data file and CellProfiler-generated images were exported using the ExportToSpreadsheet and SaveImages modules, respectively.

Cleaning and Extracting Cell Morphology Data

To prepare the unprocessed data file for analysis, a two-step data-cleaning process was implemented in RStudio. First, the laminar location of all cells was converted into a standardized cortical thickness. Second, the data were filtered to remove particles that were either too small or irregularly shaped to be a cell (Kooijmans et al., 2020).

Converting to Standard Cortical Thickness

In mice, cortical thickness is relatively uniform; however, for the subsequent analyses, it was important to ensure that the laminar location of the cells was in standardized units as the relative depth from the pial surface to the white matter. The distance from the pial surface to the start of the white matter was measured at ten uniformly spaced intervals across the width of the image, and the y coordinate for each cell was transformed such that 0 was the pial surface and 1 was the bottom of layer 6B. Any cells with y coordinates outside the thickness of the cortex, either above the pial surface or below layer 6B, were removed.

Removing Non-Cell Shaped Particles

The data file was filtered in RStudio to include cells within the known size $(50-1500\mu m^2)$ and circularity for neurons (0.2-1.0) (Kooijmans et al., 2020). This step removed irregularly shaped particles and small or large artifacts in the expression images. The final cell counts for each of our datasets are shown in Table 2.

Dataset	S1 cell count	V1 Cell Count
PV+ Cell Dataset (14274 cells)		
PV	8823	5451
Laminar Marker Gene Dataset (2028 cells)		
L1 : Ndnf	47	10
L2/3/4 : Cux2	581	401
L4 : Rorb	515	71
L5 : Deptor	51	28

Table 2 - Cell Counts for Cell Morphology Datasets Prepared in this Study.

M.Sc.	Thesis	- M.	Panday;	McMaster	University -	Neuroscience
-------	--------	------	---------	----------	--------------	--------------

L6A: Foxp2	82	95
L6B: Ctgf	107	40
PV+ Genes Dataset (19416 cells)		
PV	8823	5451
Bdnf	210	177
Cacna2d3	495	336
Calb1	409	310
Col25a1	547	264
Etv1	157	155
Fxyd6	326	258
Kit	48	50
Krt12	62	62
Nfib	71	58
Snca	494	307
Tac1	20	28
Thsd7a	55	64
Tpbg	86	93
PV+ ISH and Whole-Cell Fill Datase (5487 Cells)	et	
PV+ ISH Cells (V1 Only)	-	5451
PV+ Filled Cells (V1 Only)	-	36

Identifying Cortical Layers

The dataset of laminar marker genes was used to define cortical layer boundaries in standard coordinates for S1 and V1. Laminar density profiles were calculated for each laminar marker gene in both S1 and V1 using the density() function in R. The density profiles for laminarly

adjacent pairs of genes (e.g., L4 and L5) were used to calculate the intersections of the two profiles. A function for the difference between the two density profiles was obtained by linear interpolation using the approx() function before solving it to find the intersections between the two linear density profiles. The intersections were calculated using the uniroot.all() function from the rootSolve package in R (Soetaert and Herman, 2009) and serve as putative layer boundaries for S1 and V1.

Cluster Analysis of Cell Morphologies

Preparing the Data Frames of Features and Cell Identifiers

We developed a sensitive, data-driven approach to analyze the cell morphology data and characterize morphology clusters. The first step was to separate the cell metadata (i.e., cortical area, animal ID, gene symbols) and X, and Y coordinates from the morphology measurements, creating 2 datafiles. Next, the redundant features (Central Moment 0_0, Spatial Moment 0_0, Inertia Tensor 0_1, Inertia Tensor 1_0) that duplicate cell area were removed. The remaining 97 size and shape morphometric features are listed in Table 2. The data frame used for the cluster analysis had a row for each cell and 97 columns of morphometric features.

 Table 3 - 97 Morphometric Features for Cluster Analysis

Morphometrics included in cluster analysis				
MaximumRadius	SpatialMoment_1_2	SpatialMoment_2_3		
MeanRadius	SpatialMoment_2_1	Zernike_3_3		
MinorAxisLength	HuMoment_2	Zernike_7_5		
EquivalentDiameter	Zernike_5_1	Zernike_5_5		
InertiaTensorEigenvalues_1	HuMoment_0	NormalizedMoment_1_1		

MinFeretDiameter	Zernike_1_1	Zernike_5_3
Area	Zernike_8_4	NormalizedMoment_2_0
MedianRadius	Zernike_3_1	Zernike_7_3
ConvexArea	SpatialMoment_0_3	CentralMoment_1_3
Perimeter	Zernike_8_6	Zernike_7_1
MaxFeretDiameter	InertiaTensor_1_0	NormalizedMoment_3_1
Eccentricity	CentralMoment_2_2	AspectRatio
MajorAxisLength	Zernike_8_0	Orientation
BoundingBoxArea	CentralMoment_1_1	NormalizedMoment_2_2
Zernike_0_0	Compactness	HuMoment_4
HuMoment_1	Zernike_8_8	NormalizedMoment_1_3
SpatialMoment_0_1	SpatialMoment_2_2	HuMoment_5
SpatialMoment_1_0	Zernike_6_0	NormalizedMoment_0_2
InertiaTensorEigenvalues_0	Solidity	NormalizedMoment_3_3
HuMoment_3	Extent	CentralMoment_0_3
InertiaTensor_1_1	SpatialMoment_1_3	CentralMoment_2_1
Zernike_4_0	Zernike_4_4	NormalizedMoment_2_1
Zernike_2_2	Zernike_9_7	CentralMoment_0_1
SpatialMoment_1_1	Zernike_6_6	NormalizedMoment_1_2
CentralMoment_2_0	Zernike_9_9	CentralMoment_1_2
InertiaTensor_0_0	Zernike_9_5	HuMoment_6
CentralMoment_0_2	Zernike_4_2	NormalizedMoment_0_3
SpatialMoment_2_0	Zernike_6_4	NormalizedMoment_3_0
Zernike_2_0	Zernike_7_7	CentralMoment_2_3
SpatialMoment_0_2	Zernike_6_2	CentralMoment_1_0
FormFactor	Zernike_9_3	NormalizedMoment_2_3
Zernike_8_2	Zernike_9_1	NormalizedMoment_3_2
		EulerNumber

Before clustering the morphometric data, a logarithmic transformation was applied to the seven Hu Moment Features (Hu Moments 0,1,2,3,4,5,6) to normalize the distribution. Then, all the data were z-scored to ensure that each feature had the potential to contribute equally to the cluster analysis.

Selecting the number of clusters

The elbow method was used to select the number of clusters by calculating the within-cluster sum of squares (WCSS) for a range of different numbers of clusters (k-values). The optimal k is identified as the point of maximal curvature, or "elbow", on the WCSS graph. The elbow was calculated using three different R packages and functions: the fviz_nbclust() function from the factoextra package (Kassambara and Mundt, 2020), the find_curve_elbow() function from the pathviewr package (Baliga et al., 2023), and the elbow function from the elbow package (Casajus, 2020) and compared among the three results. Agreement among two or more elbow plot methods confirmed the optimal k-value for cluster analysis.

Robust-Sparse K-Means Clustering

We applied the Robust-Sparse K-Means Clustering (RSKC) algorithm to cluster the morphometric data. This adaptive, iterative clustering algorithm is robust to outliers and does not create clusters with one or comparatively few members (Kondo et al., 2016). In addition, it provides feature weights of the relative contributions of each morphological feature to the clustering. The greater the contribution of a feature to the clustering, the higher the weight assigned from RSKC. The top morphological features were defined as those carrying at least 50% of the maximum weight given to a single feature. We performed RSKC cluster analysis using the rskc function from the RSKC package in R.

Dimensionality Reduction and Visualization

To reduce the dimensionality and visualize the clusters in the high-dimensional cell morphology space, we applied the density-preserving t-SNE (denSNE) algorithm. This is an optimized implementation of the t-SNE algorithm (Maaten and Hinton, 2008) that better preserves the density of data points around patterns in local and global structures in the high-dimensional space (Narayan et al., 2021). The RSKC clustering was analyzed by multiplying the z-scored morphometric feature data by the corresponding feature weights (Balsor et al., 2021). Then, the cluster content was inspected by color-coding the denSNE plots using the z-scored morphological values for the top-weighted features.

Characterizing Morphological Clusters

To quantify and characterize the representative morphological features in each cluster, the median z-scored values for the top-weighted morphology features were plotted in a heatmap to create a morphology phenotype matrix. The matrix had a column for each cluster and rows for the top-weighted morphology features. The resulting heatmap was organized by unsupervised hierarchical clustering (Ward.D2 method) to group similar clusters and features. The morphology phenotype matrix was then used to identify which size and shape morphological features drove the clustering.

The median z-scored values for the top-weighted size and shape features were used to identify morphological motifs. The number of size and shape morphology motifs were determined by applying unsupervised hierarchical clustering (Ward.D2) and the elbow method as described previously.

Cluster Analysis of the PV+ Cell Dataset

The morphological analysis pipeline, as described above, was applied to our dataset of 14274 PV+ cells labelled by ISH. In this analysis, cells from V1 and S1 were analyzed together. After obtaining morphological clusters for PV+ cells, we assessed whether any morphology clusters were specific to either V1 or S1. We compared the proportions of cells from each cortical area for each morphology cluster to the global proportion of cells in this dataset (S1:V1 = 62%38%). We assessed these proportions using the chi-square test for goodness-of-fit.

Laminar Analysis of PV+ Morphology Clusters

The laminar location of the cells in each PV+ cluster was visualized by mapping the location of each cell in a cluster using their standardized XY coordinates. This provided a visualization of the laminar location of all the cells that could be used to define the laminar profile for each morphological cluster using data from S1 and V1 for all animals. Then, laminar density profiles for each cluster in S1 and V1 were calculated using the ggMarginal() function from the ggExtra package in R (Attali and Baker, 2023). Laminar profiles were calculated separately for morphology clusters in S1 and V1.

Identifying Laminar Patterns

To determine if PV+ morphology clusters had different laminar arrangements, the 130 laminar density profiles (13 morphology clusters x 2 cortical areas x 5 animals) clustered. Since the laminar boundaries were at different depths in S1 and V1 we performed a cubic piecewise spline fit to transform the Y coordinate of V1 cells into the S1 cortical space. This allows us to compare the laminar profiles in V1 and S1, accounting for the difference in cortical thickness

among the two cortical areas. To perform the interpolation, we first resampled the S1 cells, so there were the same number of cells from S1 and V1. Then, we performed the spline fit on the V1 cells. A single set of layer boundaries can be used by interpolating data from the two cortical areas into a common cortical space.

Unsupervised hierarchical clustering (Ward.D2 method) was used to identify laminar patterns. We started by constructing a dataframe of linear densities for each profile using the density() function in R. Next, we z-scored the linear density values so each profile had equal potential to impact the clustering. After determining the optimal number of clusters with the elbow method, we applied the Ward.D2 unsupervised hierarchical clustering to the z-scored linear density profiles for each morphology cluster from each cortical area for each animal. The average laminar profile for each cluster was calculated using all the laminar profiles within a cluster and visualizing those with a 95% confidence interval. We then assessed the area bias of each of these laminar patterns by comparing the proportion of laminar profiles from each pattern from each cortical area to the expected proportions (S1:V1 = 50%:50%). We assessed these proportions using the chi-square test for goodness of fit. We calculated the Shannon Entropy for each laminar pattern to quantify the morphological diversity in each laminar pattern.

Cluster Analysis of the Whole-Cell Filled Dataset

Clustering Cell Body Morphologies

As described above, the cluster analysis pipeline was applied to the whole-cell filled dataset by combining it with the PV+ data from V1. Morphological data were prepared as described previously, and clustering was performed on morphometric z-scores. We determined the optimal number of clusters using the elbow method, partitioned cells into clusters using the RSKC algorithm, and visualized the clustering using the denSNE algorithm. For this dataset, implementing the denSNE serves the additional purpose of identifying whether whole-cell filled cells integrate into the clusters formed with V1 PV+ cells. We constructed morphology phenotypes for each cluster, paying particular attention to the clusters that contained whole-cell filled cells.

Clustering Dendritic Morphologies

In addition to the 97 cell body morphology measurements for the 36 whole-cell filled cells, 20 dendritic morphology features were available to quantify the dendritic arbor morphologies. The features are listed in Table 4. We determined the optimal k using the elbow methods before applying the RSKC algorithm to z-scored measures of dendritic morphology and obtained weights for the 20 features. As there are fewer whole-cell filled cells, we used the .swc neuronal reconstructions for these neurons to visualize each cell in a cluster. Reconstructions were plotted using the Allen Institute's Cell Types Explorer Jupyter Notebook (unmodified original available from : <u>https://alleninstitute.github.io/AllenSDK/cell_types.html</u>), which we modified to automate and pseudo-color the reconstruction plots for improved visualization. We then visualized the correspondence between dendritic and cell body morphology clusters using the ggalluvial() package in R (Brunson, 2023).

Table 4 - Dendritic Morphology Features for Whole-Cell Filled Cells.

Morphometric			
number_nodes	overall_height		
total_length	average_fragmentation		

M.Sc. Thesis - M. Panday; McMaster University - Neuroscience

average_contraction	average_parent_daughter_ratio
average_diameter	overall_depth
total_surface	number_stems
number_tips	max_euclidean_distance
number_branches	max_path_distance
number_bifurcations	soma_surface
total_volume	average_bifurcation_angle_local
overall_width	max_branch_order

Cluster Analysis of PV+ Associated Genes Dataset

The cluster analysis pipeline was applied to the dataset of 13 genes associated with PV+ T-types. In this analysis, we treated S1 and V1 separately.

For each cortical area, RSKC was used to identify morphological clusters, denSNE was used to visualize the clustering and morphometric organization of the cell morphology space and morphology phenotypes for each cluster. We assessed the proportions of genes present in each morphology cluster to interrogate the relationship between cell morphologies and single-cell transcriptomics. We compared the within-cluster proportions to the global proportion of genes across the datasets for S1 and V1. We used the chi-square test for goodness of fit to determine if any clusters over- or under-represented the PV+ associated genes.

To explore the laminar distributions of PV+ associated genes, we calculated the laminar profiles of individual clusters for each gene. We calculated the profiles for clusters that had at least 10 cells from a gene and z-scored them. This revealed the morphological heterogeneity among groups of cells labelled for a single gene.

Statistical Analyses

Estimation statistics were used to compare morphological differences among clusters (dabestr package version 0.3.0) (Ho et al., 2019).

2.3 Results

Assessing ISH Label Quality Among Animals

The first step was to assess the quality of the ISH images among the animals to ensure comparable labelling so differences in label intensity did not impact the analyses. This quality control step was done by visual inspection and quantitative comparison of the distribution of integrated densities for the PV+ ISH-labelled cells from each animal.

Visually, the ISH labelling of animal 2293 in both S1 and V1 appeared lighter than the other five animals (Fig. 2A i-ii). Next, estimation statistics were used to quantitatively compare the integrated density of labelled cells among the animals (Fig. 2B). That analysis showed that the labelling intensity of cells from animal 2293 was less than the other animals and that the distributions from the other animals overlapped. Because the labelling intensity of animal 2293 was different, it was excluded from the subsequent analyses.



Figure 2 - *Data quality assessment among animals.* (A) Samples of In Situ Hybridization taken from (i) S1, and (ii) V1 labelling the PV gene. In all images, scale bar = $100\mu m$ (B) Distributions of integrated density measurements (product of cell area and mean gray value) for all cells in each sample from each animal. Single points are single cells, and the distribution is constructed from both S1 and V1. The median integrated density value is shown as a white dot beside each bee swarm, and the 95% confidence interval is shown as the black vertical bar beside each bee swarm. (C) Distributions of unpaired median difference for integrated density values among animals, using data from animal 2293 as a reference. The unpaired median difference is shown as a black dot, and the 95% confidence interval is shown as the vertical bar.

Defining cortical layers

To provide a high throughput analysis of PV+ somata in an anatomical context, we needed to define and map cortical layers an in a standardized and reproducible way. We used the laminarly restricted patterns for a panel of 6 laminar marker genes (L1 - Ndnf, L2/3/4 - Cux2, L4 - Rorb, L5 - Deptor, L6A - Foxp2, L6B - Ctgf) to construct a quantitative pipeline that would identify cortical layer boundaries in both V1 and S1.

We used the XY coordinates for each labelled cell in the ISH images shown in Figure 3 to identify layer boundaries. The coordinates were standardized among samples into a standard coordinate system, where the cortical depth ranged from 0 (pial surface) to 1 (bottom of L6B). We also used this standard coordinate system to identify subcortical cells (with a standard depth greater than 1) and excluded them from subsequent analyses.

The laminar profiles for each of the selected laminar marker genes were used to determine the layer boundaries in the standard cortical space. As expected, each laminar marker gene had a peak cell density in their respective layers. We determined the depth for layer boundaries in S1 (Fig.3A) and V1 (Fig. 3B) using the intersections of the laminar profiles for markers labelling pairs of adjacent cortical layers. These steps provided an unsupervised approach to identify the laminar boundaries in all of the animals.



Figure 3 - *Defining cortical layer boundaries in S1 and V1*. (A) upper panel : ISH for laminar marker genes extracted from S1, lower panel: laminar profiles for pairs of laminar marker genes that label adjacent layers. Laminar profiles are obtained in standard coordinates, where the depth ranges from 0 (pial surface) to 1 (bottom of L6B). The calculated intersections of the curves are taken as putative layer boundaries. (B): laminar boundaries calculated for V1. Upper panel : ISH for laminar marker genes. In all images, scale bar = $100\mu m$. Lower panel: calculated layer boundaries from laminar profiles in standard coordinates.

Laminar Cell Size Analysis of PV+ Cells

First, we tested whether the common approach, assessing cell size across the cortical layers, could identify laminar differences in PV+ cells. Globally, between S1 and V1, the distributions of cell sizes overlap as shown in Figure 4A (S1: Median = $203\mu m^2$, Mean = $200.97 \mu m^2$, 95% CI $186.59\mu m^2$]). PV+ cell sizes overlapped in layers $\frac{2}{3}$ to 6A. PV+ cells from Layer 1 in V1 (Median = 90 μ m², Mean = 114.27 μ m², 95% CI = [102.83 μ m² - 125.71 μ m²]) and S1 Layer 1 were smaller (Median = 149 μ m², Mean = 164.51 μ m², 95% CI = [154.55 μ m² - 174.47 μ m²]) compared to cells from other cortical layers. Cells in Layer 6B are smaller than cells from layers ²/₃-6A, but unlike Layer 1, the distributions overlap when comparing between S1 and V1 (Fig. 4B). The distributions of cell sizes for S1 and V1 are consistently overlapping for L2/3-L6A (Fig. 5) The summary statistics for PV+ cell size by layer for each cortical area are in Table 5. Following this typical approach to analyzing PV+ cells in the cortex, we found that PV+ cell sizes overlap across Layers $\frac{2}{3}$ -6A suggesting similar morphological organizations in those layers. However, this analysis only used cell size, whereas CellProfiler measured 97 size and shape parameters.



Figure 4 - *Size Distributions for PV+ interneurons in S1 and V1* (A) Raincloud plot comparing the global distributions of soma sizes of PV+ cells among S1 (purple) and V1(green). The global distribution is shown for each cortical area, above a horizontal boxplot. The median is shown above the boxplot at the median size. Cell size values are also plotted beneath the distribution curve, with single points as single cells. (S1: Median = 203μ m², Mean = 200.97μ m², 95% CI = [199.06 μ m²-202.91 μ m²], V1: Median =184 μ m², Mean =184.39 μ m², 95% CI = [182.18 μ m² - 186.59 μ m²]). (B) Raincloud plots comparing the distributions of PV+ soma size among cortical areas for each cortical layer separately. Summary statistics comparing between cortical areas for each cortical layer are shown in Table 5.

Cortical Area	Cortical Layer	Number of Cells	Median Cell Area (µm²)	Mean Cell Area (µm²)	95% CI (Lower Bound)	95% CI (Upper Bound)
S1	L1	345	149	164.51	154.55	174.46
V1	L1	139	90	114.27	102.83	125.71
S1	L2/3	808	215	203.12	196.19	210.04
V1	L2/3	745	192	191.63	185.32	197.94
S1	L4	2742	210	208.81	205.50	212.11
V1	L4	1258	201	199.94	195.25	204.64
S1	L5	3513	201	200.34	197.38	203.31
V1	L5	2379	178	179.31	176.18	182.43
S1	L6A	1059	218	211.16	205.22	217.10
V1	L6A	694	189.5	188.19	181.95	194.44
S1	L6B	356	138.5	147.36	139.17	155.54
V1	L6B	236	159.5	159.90	149.53	170.27

Table 5 - Summary Statistics for PV+ Cell Size by Cortical Layer for S1 and V1



Figure 5 - *Estimation statistics comparing variation in PV+ soma size among cortical layers and areas.* The unpaired median difference for 12 comparisons against a shared control are shown in the Cumming estimation plot. The reference distribution in this plot is all the data - representing the global distribution of cell size for the entire dataset. Cell size values are plotted on the upper axes as bee swarm plots. Purple bee swarms are from S1 and green swarms are from V1. In the bee swarm plots, single points are single cells. Beside each bee swarm, the median is shown as a white dot and the 95% confidence interval is plotted as a vertical black line. On the lower axes, the unpaired median differences are plotted as bootstrapped sampling distributions. The unpaired median difference is shown as a black dot and the 95% confidence interval is indicated by the vertical black line.

Morphological Clustering of PV+ interneuron cell bodies

To determine if using a rich set of cell size and shape measurements would identify morphological differences among PV+ cells, we constructed a large data matrix of 97 morphology features (columns) for each of the 14274 PV+ cells (rows). This produces

high-dimensional, cell morphology data for each PV+ cell in our dataset. The optimal number of clusters (k = 13) was found using the elbow method (Fig. 6A), and the data were clustered using the Robust, Sparse K-Means Clustering (RSKC) algorithm. The RSKC algorithm was selected because it is an iterative clustering algorithm that is robust to outliers. Additionally, RSKC provides feature weights which inform us of the relative contribution of each morphology feature to the overall clustering (Kondo et al., 2016)).

The feature weights from RSKC for the 13 morphology clusters of PV+ cells are shown in Figure 6B. Top-weighted features (>50% max feature weight) include a combination of size and shape parameters. Top-weighted size parameters include cell area, perimeter, diameters, axis lengths and radii. Among top-weighted shape parameters, both size-varying parameters that are proportional to the size of the cell (e.g., Spatial and Central Moments- which would also carry size information) and size-invariant parameters (e.g., Eccentricity, Zernike moments 0, 0, 2, 0, 2 2, ad 4 0 and Hu Moments 1 and 3) that describe the circularity, elongation and symmetry of the cell body. Morphology features carrying low weight include higher order Zernike, spatial and central moments that quantify more subtle differences in shape related to the symmetry of processes arranged around the cell body, and the Euler number, which describes the topology of the cell (i.e., the presence of holes or distensions of the cell body). Collectively, RSKC identified the morphology parameters that represent the largest variance among the cells. The distribution of RSKC weights describes morphology differences among cells in terms of their size, circularity, elongation, and symmetry of the cell body. Taken together, the feature weights from RSKC indicate that despite a consistent topological structure, there are aspects of cell size and shape that are sufficiently different to drive the separation of cells into distinct morphology clusters.



Figure 6 - Unsupervised clustering on PV+ Interneuron morphologies. (A) Plot of the within-cluster sum of squares against the number of clusters, k. The optimal k-value (k = 13) was determined by the elbow method and shown by the red line. (B) Weights for all 97 morphology features as determined by the robust, sparse K-means clustering (RSKC) algorithm. Feature weights are plotted in decreasing order from left to right. Top-weighted features (>50% of maximum weight attributed to a single morphology feature), are identified by the black bar.

Visualizing PV+ Morphology Clusters

To visualize the 13 morphology clusters, we applied the density-preserving t-SNE (denSNE) algorithm. For this analysis, the denSNE perplexity was set to 10, and the random number seed was 13579. The z-scored value for each morphometric feature was multiplied by the feature weight obtained from RSKC, and the denSNE algorithm was applied to these transformed values. This allows each feature to contribute to the dimensionality reduction in a manner consistent with their contribution to the clustering (Balsor et al. 2021). The denSNE plot was color-coded for the 13 PV+ cell clusters and showed a mosaic of distinct morphological clusters (Fig.7A).

We interrogated the distribution of features in the denSNE plot by color-coding the points using the 14 top-weighted size and shape features (Fig.7B). These visualizations showed organized, gradual shifts across the desSNE. For example, the Area of the cells shifted from right to left going from small (dark blue) to progressively larger (yellow) cells. Similar gradient patterns were found for the other morphological features; however, the orientation of the gradient was different for the shape features.

The cell shape gradients in the denSNE varied from top to bottom. They were roughly orthogonal to gradients for the size features (Fig.7B). Importantly, the intersection of the size and shape gradients helped to identify morphological differences among the clusters. For example,

clusters 1, 2 and 3 were small cells on the right side of the denSNE but were identified as different shapes and spread out across the vertical dimension of the denSNE plot.

We mapped the gradients of shape information and found measures of circularity increase from top to bottom, but measures of elongation increase from bottom to top in the denSNE plot (Fig.7B bottom row). Taken together, morphology clusters reflected the combination of size and shape features, which progressed orthogonally across the cell morphology space. Furthermore, the intersection of these gradients would determine how the morphology space is partitioned into clusters.



Figure 7 - *Visualizing RSKC morphology clusters for PV*+ *Cells* : (A) Morphology space for PV+ cells from S1 and V1 visualized with density-preserving t-SNE (denSNE) (perplexity = 10, seed = 13579). The denSNE algorithm was applied to a transformed dataset where feature weights from RSKC were multiplied to the z-scored morphology feature measurements. Single points are single cells and the denSNE plot is coloured by cluster labels, with cluster centroids marked with a cross. Cluster numbers reflect the progression of median cell sizes (cluster 1 = lowest median size, and Cluster 13 = highest median cell size). 4 cells were sampled randomly from each morphology cluster and are shown in the inset boxes surrounding the denSNE plot. (B) Gradients of selected top-weighted morphology features mapped onto morphology space, visualized using the denSNE algorithm. The colour progression from navy to yellow indicates increasing feature values for each morphology feature. Size features are shown in the top row and shape features are shown in the bottom row.

Characterizing PV+ Cluster Morphologies

The denSNE plots (Fig. 7) indicate that the size and shape features grouped the cells into morphological clusters. However, a quantitative description of each cluster's morphology features is needed to describe the characteristic PV+ soma phenotype for each cluster. We used the median z-scored of the top-weighted features for each of the 13 clusters to create a cluster-by-feature heatmap (Fig.8). The order of the morphology clusters (columns) and features (rows) was determined by hierarchical clustering. The size features ordered the 13 morphology clusters from left to right, with the clusters containing the smallest cells on the left and the largest on the right. The shape features further ordered the morphology clusters into interdigitated columns where clusters of similar-sized cells were separated based on the shape features. For example, morphology clusters 3 and 4 had a similar pattern of blue and green color-coded cells for the size features but a different pattern for the shape features.



B S1:V1 Proportions by Cluster (Global = 62%:38%)



Figure 8 - *Characterizing RSKC morphology clusters* (A) cluster morphology phenotypes describing each cluster's median size and shape characteristics. Phenotypes were constructed using the top-weighted morphology features, as shown in Figure 5B. Clusters are ordered and named according to the progression of median sizes (Cluster 1 = lowest median size, Cluster 13 = highest median size). The heatmap rows were organized using a row-wise dendrogram, using the Ward.D2 algorithm for hierarchical clustering. The first branch point on the row-wise dendrogram indicates the separation of top-weighted morphology features into size (towards the top of the heatmap) or shape features (towards the bottom of the heatmap). (B) Proportions of cells from each cortical area (S1 = purple, V1 = green) for each morphology cluster. P-values beneath each pie chart are used for the chi-square test, which compares the areal proportion swithin each cluster to the global areal proportion in the dataset. The overall proportion of cells from each cortical area in this dataset is S1 = 62% and V1 = 38%

We applied hierarchical clustering to the size and shape features separately within the phenotypes to determine how many different size divisions and shape motifs are represented in the 13 morphology clusters. The dendrograms for the size and shape clusters are shown in Figure 9. The clusters are separated coarsely into four major size divisions that we annotate as small, medium, large, and extra large. The clusters also separate into three prominent shape motifs that reflect fusiform, pyriform, and round cell shapes. We present here the characteristics of each shape motif.

Shape Motif 1 - fusiform cell bodies: clusters 1, 2, 4, 7, 12 \rightarrow These cells are elongated, as evidenced by their higher median eccentricity and second-order Zernike Moments. With increasing cell size, cells move from being more obliquely elongated to being vertically elongated- indicated by the progressive increase in median Zernike_2_0 (vertical astigmatism) and decrease in median Zernike_2_2 (oblique astigmatism). Cluster 2 was defined by a more extreme version of this shape motif - with more prominent elongation, resulting in sharply lower median measures of circularity (Fig.7A). These clusters occupy adjacent regions of the denSNE

visualization of the cell morphology space (Fig.7). These clusters occupy a region of the morphology space defined by extreme shape characteristics, but still are organized along the size gradient. These clusters had higher median measures of elongation relative to clusters in other shape motifs (Figs. 11-13)

Shape Motif 2 - pyriform cell bodies: clusters 3, 6, 9, 10, 13 \rightarrow These cells are defined by intermediate median measures of cell circularity and elongation (Fig.8A, Figs. 11-17). These cells have intermediate shape characteristics - they are not as elongated as cells in shape motif 1 and not as circular as cells in shape motif 3. Cells in shape motif 2 were round on one side and taper to a point. Their median z-scored shape measurement values were closer to 0 than those in the other shape motifs. These clusters were found in the middle of the morphology space, in between regions occupied by clusters with the other two shape motifs (Fig. 6). Relative to the other shape motifs, pyriform cells had lower measures of elongation than fusiform cells (Figs.11-13), and lower measures of circularity than round cells (Figs.14-17).

Shape Motif 3 - round cell bodies: clusters 5, 8, 11 \rightarrow Cells in these clusters are more circular with high median measures of circularity (Zernike_0_0, Hu Moments 1 and 3, Zernike_4_0) and lower measures of elongation. The largest cells, cluster 11, had radially symmetrical bumps around the perimeter of the soma that probably reflect processes emerging from the soma. The median Zernike_4_0 was high in cluster 11 and measured both circularity and fourfold radial symmetry. As processes around the perimeter of the soma will tend to display greater symmetry in more circular somata. These clusters were visualized across the bottom of the morphology (Fig.7). Estimation statistics performed on measures of circularity identify clusters 5, 8, and 11 as having the highest median circularities of all clusters (Figs. 14-17).

Cortical Area Distribution of PV+ Morphologies

We analyzed if any of the morphologies were more commonly found in S1 versus V1 by quantifying the proportion of cells from each area in a cluster (Fig. 7B), and assessed these differences with the chi-square test for goodness of fit. Among clusters 11-13, which have the largest median cell sizes across all clusters, more cells come from S1 than expected (global proportion S1:V1 = 62%:38%. Cluster 11: p = 5.546e-06, Cluster 12: p = 1.789e-09, Cluster 13 : p = 4.074e-14). Among clusters 3-6, which represent small and medium median cell sizes, more cells come from V1 than expected (global proportion S1:V1 = 62%:38%, Cluster 3 : p = 1.082e-06, Cluster 4 : p = 5.160e-06, Cluster 5 : p = 3.967e-04, Cluster 6 : p = 3.107e-04).

In contrast to measuring only the area of the cell body, the analysis of PV+ cells applied an unsupervised, data-driven approach to a high-dimensional representation of PV+ cell morphologies. The results of using our analysis pipeline indicate that measurements of both size and shape parameters are necessary to characterize the morphology of PV+ neurons fully. For example, we showed multiple shapes within each size division. Furthermore, the approach highlights the importance of using a large number of cells and morphology features to quantify the complexity of PV+ cell morphologies and gain insight into the distribution of PV+ cell morphology.



B Shape Motifs for PV+ Cell Morphology Clusters



Figure 9 - *Characterizing size divisions and shape motifs among clusters.* The median z-scores for top-weighted morphology features for each cluster were analyzed by unsupervised hierarchical clustering (Ward.D2) method, to identify the groups of clusters that share common size and shape characteristics.(A) Dendrogram for hierarchical clustering grouping individual morphology clusters into size divisions, based on the median values of top-weighted size features. (B) Dendrogram for hierarchical clustering grouping morphology clusters into shape motifs using the median z-scored values of top-weighted shape features. Cells fall into one of three motifs : fusiform, pyriform or round. The number of size divisions and shape motifs were determined using the elbow method.



Figure 10 - *Soma Area for Each Morphology Cluster*: Estimation statistics plots showing the distributions of z-scored cell body area for PV+ interneurons grouped by morphology cluster. The collection of plots shows comparisons among the clusters using each morphology cluster as a reference. Beeswarm plots of the z-scored cell body area are plotted on the upper axes of each plot. Beside each beeswarm the median z-scored cell size for each cluster is indicated as a white dot, and the 95% confidence interval about the median is shown by the black bar. The lower axes of each plot show the unpaired median difference plotted as a bootstrapped sampling distribution. The 95% confidence interval is shown as a black bar and the median is shown with a black dot.



Figure 11 - *Cell eccentricity among morphology clusters of PV+ interneurons*. Estimation statistics plots showing the distributions of z-scored cell body eccentricity for PV+ interneurons grouped by morphology cluster. The collection of plots shows comparisons among the clusters using each morphology cluster as a reference. Beeswarm plots of the z-scored cell eccentricity are plotted on the upper axes of each plot. Beside each beeswarm the median z-scored eccentricity for each cluster is indicated as a white dot, and the 95% confidence interval about the median is shown by the black bar. The lower axes of each plot show the unpaired median difference plotted as a bootstrapped sampling distribution. The 95% confidence interval is shown as a black bar and the median is shown with a black dot.


Figure 12 - Zernike Moment 2_2 values among morphology clusters of PV+ interneurons. Estimation statistics plots showing the distributions of z-scored Zernike 2_2 moment values for single PV+ interneurons grouped by morphology cluster. The collection of plots shows comparisons among the clusters using each morphology cluster as a reference. Beeswarm plots of the z-scored cell Zernike 2_2 are plotted on the upper axes of each plot. Beside each beeswarm the median z-scored Zernike 2_2 moments for each cluster is indicated as a white dot, and the 95% confidence interval about the median is shown by the black bar. The lower axes of each plot show the unpaired median difference plotted as a bootstrapped sampling distribution. The 95% confidence interval is shown as a black bar and the median is shown with a black dot.



Figure 13 - Zernike Moment 2_0 values among morphology clusters of PV+ interneurons. Estimation statistics plots showing the distributions of z-scored Zernike 2_0 moment values for single PV+ interneurons grouped by morphology cluster. The collection of plots shows comparisons among the clusters using each morphology cluster as a reference. Beeswarm plots of the z-scored cell Zernike 2_0 are plotted on the upper axes of each plot. Beside each beeswarm the median z-scored Zernike 2_0 values for each cluster is indicated as a white dot, and the 95% confidence interval about the median is shown by the black bar. The lower axes of each plot show the unpaired median difference plotted as a bootstrapped sampling distribution. The 95% confidence interval is shown as a black bar and the median is shown with a black dot.



Figure 14 - *Hu Moment 1 values among morphology clusters of PV+ interneurons*. Estimation statistics plots showing the distributions of z-scored Hu moment 1 values for single PV+ interneurons grouped by morphology cluster. The collection of plots shows comparisons among the clusters using each morphology cluster as a reference. Beeswarm plots of the z-scored cell Hu Moment 1 values are plotted on the upper axes of each plot. Beside each beeswarm the median z-scored Hu Moment 1 values for each cluster is indicated as a white dot, and the 95% confidence interval about the median is shown by the black bar. The lower axes of each plot show the unpaired median difference plotted as a bootstrapped sampling distribution. The 95% confidence interval is shown as a black bar and the median is shown with a black dot.



Figure 15 - *Hu Moment 3 values among morphology clusters of PV+ interneurons*. Estimation statistics plots showing the distributions of z-scored Hu moment 3 values for single PV+ interneurons grouped by morphology cluster. The collection of plots shows comparisons among the clusters using each morphology cluster as a reference. Beeswarm plots of the z-scored cell Hu Moment 3 values are plotted on the upper axes of each plot. Beside each beeswarm the median z-scored Hu Moment 3 values for each cluster is indicated as a white dot, and the 95% confidence interval about the median is shown by the black bar. The lower axes of each plot show the unpaired median difference plotted as a bootstrapped sampling distribution. The 95% confidence interval is shown as a black bar and the median is shown with a black dot.



Figure 16 - Zernike Moment 0_0 values among morphology clusters of PV+ interneurons. Estimation statistics plots showing the distributions of z-scored Zernike Moment 0_0 values for single PV+ interneurons grouped by morphology cluster. The collection of plots shows comparisons among the clusters using each morphology cluster as a reference. Beeswarm plots of the z-scored cell Zernike Moment 0_0 values are plotted on the upper axes of each plot. Beside each beeswarm the median z-scored Zernike moment 0_0 values for each cluster is indicated as a white dot, and the 95% confidence interval about the median is shown by the black bar. The lower axes of each plot show the unpaired median difference plotted as a bootstrapped sampling distribution. The 95% confidence interval is shown as a black bar and the median is shown with a black dot.



Figure 17 - Zernike Moment 4_0 values among morphology clusters of PV+ interneurons. Estimation statistics plots showing the distributions of z-scored Zernike Moment 4_0 values for single PV+ interneurons grouped by morphology cluster. The collection of plots shows comparisons among the clusters using each morphology cluster as a reference. Beeswarm plots of the z-scored cell Zernike Moment 4_0 values are plotted on the upper axes of each plot. Beside each beeswarm the median z-scored Zernike moment 4_0 values for each cluster is indicated as a white dot, and the 95% confidence interval about the median is shown by the black bar. The lower axes of each plot show the unpaired median difference plotted as a bootstrapped sampling distribution. The 95% confidence interval is shown as a black bar and the median is shown with a black dot.

Laminar Distribution of Morphology Clusters in S1 and V1

The spatial location of the cells in each cluster was mapped onto the standardized cortical space for S1 and V1. The laminar density profile was calculated for each cluster in each cortical area (Fig. 18). Visual inspection of these maps showed that none of the PV+ morphology clusters was laminarly restricted. Still, all of the clusters had a laminar location with peak PV+ cell density. Comparing the cortical areas, the same morphology cluster appeared to have different laminar patterns. For example, clusters 2 and 12 have both a supragranular and infragranular peak in S1, cluster 2 has a more broad, diffuse distribution across the layers in V1, whereas cluster 12 is mostly found in supragranular layers.

The laminar distributions of each morphology cluster in S1 and V1 are shown in Figure 19. We used chi-square analysis to test whether any clusters had more PV+ cells in a layer than expected. With the exception of clusters 4 and 10 in S1, and clusters 4, 7, and 10 in V1, all clusters had different proportions of cells by layer than expected. Notably, we found that in S1 and V1, cluster 2 had more cells from layers 1 and 6B, and cluster 12 had more cells from layers 4 and 5 than expected (S1 cluster 2: p= 3.112e-37, 12: p = 1.683e-16; V1 clusters 2: p = 3.396e-50, 12: p = 3.238e-11).



Figure 18 - *Laminar distributions of PV+ Morphology Clusters in S1 and V1* - For each cluster, cells (single points) are shown in standard coordinates in each cortical area. In a single plot, data from all 5 animals are shown. The layer boundaries are calculated from the analysis in Figure 3. Each cluster cell plot shows data pooled from all animals within a single cortical area. Beside each plot is the laminar profile for each morphology cluster, showing their distribution across the cortical layers for S1 and V1.

Laminar Descriptions of Morphology Clusters

A Primary Somatosensory Cortex (S1)

L4:23.08%

i) Global Laminar Distribution ii) Laminar distributions of S1 cells for each cluster





p = 0.1146

12

p = 3.238e-11

49.92% p = 1.021e-03

40.16%

p = 3.790e-03

11

L5:43.64%



34.48%

p = 0.6311

p = 1.072e-02

L2/3 L4 L5 L6A

L6B

Cortical Layer L1

45.83%

37.54%

p = 4.390e-03

p = 1.960e-08

Unsupervised Analysis of PV+ Cell Laminar Patterns

To quantify the laminar patterns of each PV+ cell morphology cluster, we applied unsupervised hierarchical clustering to a dataset of 130 laminar density profiles (13 morphology clusters x 2 cortical areas x 5 animals). The clustering identified 6 different laminar patterns (Fig. 20A). We interrogated each laminar cluster to determine the proportion of profiles from S1 vs V1 and the 3 cell shape motifs. Finally, we used a river plot to connect the cells in each of the 13 PV+ morphology clusters with their laminar profile cluster (Fig. 20B).

Laminar Pattern A - This laminar pattern describes a sharp increase in cell density across layers 2/3, before decreasing across layer 4 and increasing again in layer 5. This pattern is found slightly more in S1 compared to V1 (S1:V1 = 77.78%:22.22%, p = 9.588e-02). This pattern consists primarily of extra large fusiform cells (Cluster 12). This pattern is both morphologically restricted, and biased towards a single morphology (p = 4.62e-02), as it has a Shannon entropy score of 1.149.

Laminar Pattern B - This laminar pattern increases more gradually across layer $\frac{2}{3}$ and continues to increase across layer 4 to peak cell density in layer 5. This pattern is found slightly more in V1(S1:V1 = 34.38%:65.62%, p = 7.710e-02). Compared to S1, where large, fusiform PV+ cells populate supragranular layers (Fig. 8A), in V1, medium-sized pyriform cells populate these layers(Fig. 8B). Unlike pattern A, this laminar pattern is morphologically broad, with less morphological bias within the pattern (Shannon entropy = 2.326).

Laminar Pattern C - This laminar pattern describes low cell density across the supragranular layers - layers 1 and 2/3. Cell density increases across layers 4 and upper layer 5 before decreasing across the deeper sublaminae of layer 5. This pattern is found slightly more in

S1 than V1, but does not differ strongly from expected (S1: V1 = 56.52%, 43.48%), p = 5.316e-01. This pattern is the only one to consist mostly of round cells (43.5% round cells, p = 5.236e-02). Morphology clusters with medium-sized cells (Clusters 5, 6, and 7) are in this pattern. This pattern is also morphologically diverse with less morphology bias (Shannon Entropy = 2.178).

Laminar Pattern D - Compared to pattern C, which also has a peak in L5, pattern D has increased cell density across the supragranular layers, including layer 1. Cell density increases progressively along the depth of layer 4, with a peak towards the middle of layer 5. This pattern, in addition to the latter two, shows increased cell densities into L6A, which characterize the deep infragranular biases seen across some of the morphology clusters shown previously. This pattern consists primarily of pyriform cells, with slightly more than expected (52.2% pyriform cells, p = 4.000e-01). This pattern also consists of morphology clusters with mainly medium to large sizes.

Laminar Pattern E - This pattern, like pattern A, increases cell density across layers $\frac{2}{3}$, and 5. However, the increase in cell density is more gradual across the supragranular layers than in pattern A. In pattern E, cell density is sustained from layer $\frac{2}{3}$ across layer 4 before increasing again across layer 5. This pattern does not show a sharp decrease in cell density across layer 4 as seen in pattern A. This pattern is found slightly more in S1(S1:V1 = 55.56%, 44.44%, p = 6.374e-01) and consists mainly of small, fusiform cells (66.7% fusiform, p = 4.356e-02). With the second-lowest Shannon entropy of all the laminar patterns, this pattern is also morphologically biased and seen across fewer morphology clusters. As this pattern is dominated by small fusiform cells, cells from clusters 1 and 2 are more represented in this pattern than in others.

Laminar Pattern F - Like patterns D and E, pattern F gradually increases cell density across layers 2/3, 4 and 5 before decreasing in L6A. This pattern has the deepest peak cell density in Layer 5 of all the laminar patterns. Found slightly more in V1, this pattern has a similar proportion of fusiform, round and pyriform cells as the distribution of shape motifs in pattern C, with the greatest share of cells being round (40% round cells, p = 1.054e-01). This pattern has a similar composition of sizes compared to pattern C, which also consists mainly of medium cell sizes with a combination of all three shape motifs. However, this pattern consists of larger cells than pattern C's composition. This laminar pattern is also morphologically broad with comparably low morphological bias (Shannon entropy = 2.279).

We found weak laminar differences when we applied the typical method that uses only cell size to assess laminar differences (Fig. 4). In contrast, we identified robust differences in the laminar patterns of PV+ cell morphologies using density profiles and unsupervised clustering. Layers with the greatest PV+ cell density were found to have comparable, overlapping size distributions, indicating high similarity. The approach, however, identifies variations in shape among cells of similar sizes and leverages the variations in shape in addition to size differences to characterize morphologies of single PV+ cells. Although strong areal biases were only seen for extra-large cells with these more comprehensive morphology clusters, the laminar analysis of PV+ morphologies identified laminar patterns biased towards either V1 or S1. Our data-driven approach considers both the size and shape information of single PV+ cells. The spatial distributions of these morphology clusters follow characteristic patterns more representative of S1 and V1.



Figure 20 - *Characterizing the laminar distributions of PV+ cell morphologies*. (A) 130 laminar profiles (5 animals x 2 cortical areas x 13 morphology clusters) were clustered into 6 patterns using unsupervised hierarchical clustering (ward.D2 algorithm). Cells from V1 were interpolated into S1 cortical space using a cubic piecewise spline fit so a single-layer definition could be used. Single-coloured profiles represent the laminar profile for a single morphology cluster in one cortical area from one animal. The average curve for each laminar pattern is plotted as a thick black line. Top row: pie charts summarizing the distribution of laminar profiles from each cortical area in each laminar pattern (S1: purple, V1: green). The proportions are compared to the overall distribution of profiles in the dataset (S1: 50%, V1:50%). Lower row: pie charts summarizing the distribution of shape motifs in each laminar pattern (F = Fusiform, P = Pyriform, R = Round). (B) River plot showing the correspondence between the laminar patterns (left column) and the morphology clusters (right column). Size divisions for the morphology clusters are shown to the left of the river plot. Shannon entropy values were calculated for each laminar pattern and are shown to the right of the river plot.

Comparing Cell Body Morphology and Dendritic Morphology

We analyzed the soma and dendritic morphology of whole-cell filled PV+ cells using data from mouse V1 from the Allen Brain Institute (Hodge et al., 2019). The first step in integrating these data with the PV+ ISH dataset was to use estimation statistics to compare 3 size (Cell Body Area, Bounding Box Area, and Perimeter) and shape (Circularity, Aspect Ratio and Eccentricity) parameters from the whole-cell filled cells and ensure that they were not different from the V1 PV+ dataset (Fig. 21). None of those measured differed between the PV+ whole-cell filled and ISH-labelled cells so we integrate the morphometric data from those cells. Combining data from two modalities, we used this dataset to assess the relationship between dendritic and soma morphologies in V1. This dataset comprises 5451 cells labelled by ISH and 36 filled cells from mouse V1. For the 36 filled cells, 20 measurements quantifying the dendritic morphologies were obtained from the Allen database. We used RSKC to cluster the cells separately based on their dendritic and soma morphologies.



Figure 21 - *Filled cells have comparable size and shape parameters to ISH-labelled cells*. Gardner-Altman estimation plots showing the unpaired median difference between ISH labelled cells and filled cells. ISH-labelled cells are being used as a reference. Both groups are plotted on the left axes, and the unpaired median difference is plotted on the right axes as a bootstrapped sampling distribution. The median difference is plotted as a black dot, and the 95% confidence interval is indicated by the vertical black line (A) Area, (B) Circularity, (C) Bounding Box Area, (D) Aspect Ratio, (E) Perimeter, (F) Eccentricity

Cluster Analysis of Dendritic Morphologies

The 20 dendrite morphology measures were used to cluster the 36 filled cells. The elbow method determined that the optimal k=3 and RSKC was used to cluster the cells based on the dendritic morphology features (Fig. 22A). The top-weighted dendritic morphology features included the number of nodes, the total length of the dendritic arbour, the average contraction, average diameter, total size of the dendritic field, and the number of tips (Fig. 22B). These features describe the size of the dendritic field, together with large-scale shape properties related to the arrangements of dendrites within the field. The lowest-weighted features include the soma size, average bifurcation angle, and branch order. Although information about the cell body was present in this cluster analysis, the only parameter provided was the size. The clusters obtained here using dendritic morphology features (Figs. 23 and 24) are consistent with previously published data (Jiang et al., 2015).





Figure 22 - *Clustering dendritic morphology features for PV+ filled cells.* (A) Elbow plot showing the within-cluster sum of squares plotted as a function of the number of clusters (k). The optimal number of clusters (k=3) is shown in red. (B) Feature weights as determined by RSKC for the 20 dendritic morphology features. Features quantify dendritic arbour morphologies, and top weighted (>50% maximum weight assigned to a single feature) features are indicated with the black bar.



Figure 23 - *Cell Reconstructions of Filled cells (Dendritic Morphologies A and B)*. Whole cell reconstructions for each filled cell for the first two of three dendritic morphology clusters. Pale segments are the dendritic arbour, and dark segments are the axonal arbour. Truncated portions of arbors are indicated with a coloured dot. The black dot indicates the soma. Cells are numbered from 1-36 and are ordered by their cortical layer. Scale bars in all reconstructions = 100 microns. The different colors indicate the different dendritic morphology clusters (A = pink, B = gold).



Figure 24 - *Cell Reconstructions of Filled cells (Dendritic Morphology C)*. Whole cell reconstructions for each filled cell for the last of three dendritic morphology clusters. Pale segments are the dendritic arbour, and dark segments are the axonal arbour. Truncated portions of arbors are indicated with a coloured dot. The black dot indicates the soma. Cell numbers indicate their position in the cortical layers, lower numbers indicating upper layers and higher numbers indicating deeper layers. Scale bar for all images = 100 microns.

Next, cell body morphology measurements of the PV+ whole-cell filled and ISH cells from V1 and RSKC were applied to cluster those data. Eleven cell body morphology clusters were identified, and the PV+ whole-cell filled cells were found in 7 of those clusters (Figs. 25 & 26). The top-weighted features were similar to those found when clustering only the PV+ ISH-labeled cells and included a combination of size and shape parameters (Fig. 25B). Among the top-weighted shape features were measures of elongation (Eccentricity, Zernike Moments 2_0 and 2_2, and the aspect ratio) and circularity (Zernike Moments 0_0 and 4_0, together with Hu Moments 1 and 3).



Figure 25 - *Clustering cell body morphologies of PV+ filled and ISH-labelled cells*. (A) - Elbow plot showing the optimal number of clusters (k = 11), (B) Feature weights as determined by RSKC performed on 97 morphology features for each cell soma. Top-weighted features (>50% of the maximum weight assigned to a single feature) are indicated with the black bar.

The 36 filled cells were found in 7 clusters along PV+ ISH-labelled cells (Fig. 26). Similar to the analysis of PV+ ISH cells, the denSNE plot represented roughly orthogonal gradients of size and shape. The filled cells were small to medium size and distributed across the range of shape patterns (Fig. 26B). For example, filled cells found in clusters 3, 4, and 6 represent a band of medium-sized cells in the denSNE plot but cut across a range of cell shapes from round in cluster 4 (dark blue region in the denSNE for eccentricity) to more and more oval-shaped cells in clusters 3 and 6 (Fig. 26A).

The heatmap of the top-weighted features for the filled and ISH PV+ cells from mouse V1 provides a clear visualization that can be used to phenotype the clusters based on the morphological features that separate cells into clusters. For example, the greenish color-coding of the values for the size features was similar between clusters 3 and 4, but the shape features had opposite patterns. Eccentricity was higher and color-coded reddish for cluster 3 to indicate that these cells are round, while the color-code was bluish for cluster 4, indicating not round cell shapes (Fig. 27).



Figure 26 - *Morphological properties of filled cells integrate along the morphology gradients established by ISH-labelled cells.* (A) - denSNE plot of the PV+ cell body morphology space annotated by cluster assignment. Cluster numbers indicated the progression of median cell size (Cluster 1 = lowest median cell size Cluster 11 = highest median cell size). Single points are single cells. Fully opaque points are whole-filled cells, and translucent points are ISH-labelled cells. For morphology clusters with filled cells, the cell bodies of those filled cells are shown (B) denSNE plot from panel A, but annotated for six top-weighted size features (top row) and six top-weighted shape features (bottom row). The colour progression from dark blue to yellow indicates increasing feature values for each morphology feature.





89

size parameters towards the top of the heatmap, and shape parameters towards the bottom of the heatmap. Furthermore across the size gradient, repeated shape motifs can be seen.

We used a river plot to map the correspondence between the clustering of PV+ filled cells based on dendritic versus soma morphologies (Fig. 28). Dendritic morphology clusters did not restrict themselves to exclusive combinations of cell body morphology clusters. Dendritic morphology A is primarily seen across the smaller, round PV+ cells from deep layers, but smaller fusiform PV+ cells in the supragranular and granular layers. This dendritic morphology pattern was not identified in the cells from layer 5. Dendritic morphology B is morphologically broad across cell body morphology clusters 3, 4, 5, and 6. These cells encompass a range of sizes and shapes (small to medium sizes and all three main shapes - fusiform, pyriform and round). Despite various cell body morphologies, this dendritic morphology is concentrated across layer 5 cells. Dendritic morphology C is also very morphologically and laminarly broad, found across cell body morphology clusters 4, 5, 6, 7, and 10, with the greatest proportions of cells falling into cell body morphology clusters 5 and 6 (pyriform and fusiform cells respectively). This dendritic morphology pattern is more representative of cells with less rounded cell bodies and is found across the layers in larger proportions than other morphology clusters.



Figure 28 - *Correspondence among cell dendritic and cell body morphologies.* River plot showing the correspondence among dendritic morphology clusters (left), cell body morphology cluster (centre) and cortical layer (right) for the 36 filled PV+ cells. Cell body morphology cluster labels reflect their position along the size gradient defined across all morphology clusters.

Comparing the Morphology of PV+ and PV+ Associated Gene Cells

We compared the morphologies of PV+ cells with cells labelled for PV-associated genes from different T-types to determine if cell morphology could help identify PV+ T-types. The 13 PV-associated genes were selected from the list of marker genes for PV+ T-types defined in Tasic et al. 2016 and represent the transcriptomic types defined in Gouwens et al., 2020. The ISH for these selected genes had good-quality labelling in both S1 and V1. Although we could not perform the spatial transcriptomics needed to validate PV+ cells across the T-types, it is likely that these cells are from different PV+ T-types. The PV-associated cells overlap with the PV+ ISH cells in their laminar locations, however, some genes (Calb1, Etv1) display characteristic laminar restrictions, whereas PV+ cells were found across all cortical layers. (Fig. 29).



Figure 29 - *ISH samples for PV+ and Associated Genes*. Samples of ISH extracted from (A) S1 and (B) V1 labelling a panel of 14 genes associated with PV+ transcriptomic types. Each gene's ISH was extracted from a different animal, with S1 and V1 samples coming from a single section for each gene. The samples were cropped to the pial surface and leave a portion of the subcortical white matter visible beneath the deep limit of L6B. In all samples, only cortical cells were used in subsequent analyses. Pixel size : 1 pixel = 1 microns. Scale bar = 100 microns.

The morphology measurements from cells in S1 and V1 were clustered separately. In S1, 13 clusters were identified, and in V1, there were 10 clusters (Figs. 30A & 31A). The top-weighted features included size (area, perimeter, diameters, radii, axis lengths) and shape parameters (circularity – Zernike 0_0, Zernike 4_0, Hu Moments 1 and 3; elongation – eccentricity, Zernike 2_0, Zernike 2_2). Features with low weights in both cortical areas described the topology of the cells and higher-order shape moments indicated the presence of processes extending from the cell body (S1: Figure 30B, V1: Figure 31B). The larger number of clusters in S1 represented a greater range of cell sizes than V1.



A Determining the optimal number of clusters (k)





Figure 30 - *Clustering PV+ gene morphologies in primary somatosensory cortex S1.* Weights plot showing the distribution of weight across the 97 morphology features as determined by the RSKC algorithm. Features are arranged so weights appear in decreasing order in the weights plot. Top weighted features (>50% max weight) are indicated by the black bar.


A Determining the optimal number of clusters (k)

B Morphology feature weights determined by RSKC



Figure 31 - *Clustering PV+ gene morphologies in primary visual cortex V1.* (A) Elbow plot for the within cluster sum of squares as a function of the number of clusters (k). The optimal k = 10 is shown with the red line. (B) Weights plot showing the distribution of weight across the 97 morphology features as determined by the RSKC algorithm. Features are arranged so weights appear in decreasing order in the weights plot. Top weighted features (>50% max weight) are indicated by the black bar.

The denSNE plots of morphology clusters for the PV-associated genes in S1 (Fig. 32) and V1 (Fig. 33) had similar organizations to the plot of just PV+ cells labelled by ISH(Fig. 7). Both denSNE plots had roughly orthogonal gradients for size and shape features (Figs. 32B & 33B). Examples of the PV-associated cells in the clusters illustrate the range of cell sizes and shapes

(Figs. 32A & 33A).



Figure 32 - *Visualising PV+ gene morphologies for S1-* (A) Morphology space for S1 cells labelled for PV and associated genes visualized using density preserving t-SNE (denSNE). The denSNE algorithm was applied to a transformed dataset where the weights obtained from RSKC were multiplied to their corresponding feature z-scores. Single points are single cells and points are coloured by their cluster assignments. Cluster numbers reflect the progression of median cell sizes (Cluster 1 - lowest median cell size , Cluster 13 - highest median cell size). 4 cells were randomly sampled from each cluster and are shown in the inset squares surrounding the denSNE plot. (B) The denSNE plot in panel A was annotated for a series of six top-weighted size parameters (top row), and six top-weighted shape parameters (bottom row). The colour progression from dark blue to yellow indicates increasing feature values for each morphology feature.



Figure 33 - *Visualising PV+ gene morphologies for V1-* (A) Morphology space for V1 cells labelled for PV and associated genes visualized using density preserving t-SNE (denSNE). The denSNE algorithm was applied to a transformed dataset where the weights obtained from RSKC were multiplied to their corresponding feature z-scores. Single points are single cells and points are coloured by their cluster assignments. Cluster numbers reflect the progression of median cell sizes (Cluster 1 - lowest median cell size , Cluster 10 - highest median cell size). 4 cells were randomly sampled from each cluster and are shown in the inset squares surrounding the denSNE plot. (B) denSNE plot from panel A, but annotated for a series of six top-weighted size parameters (top row), and six top-weighted shape parameters (bottom row). The colour progression from dark blue to yellow indicates increasing feature values for each morphology feature.

Heatmaps of the morphology clusters and top-weighted features were plotted for each cortical area to identify the morphological features that differentiated among the clusters. The heatmaps had similar organizations in S1 and V1 and were consistent with the previous analyses. For example, clusters were ordered by cell size features and clusters of similar size were separated by the shape features (Figs. 34A & 35A). In both cortical areas, there were three shape motifs: fusiform, pyriform and round, similar to the PV+ cell dataset (Figs. 37B & 38B). In V1, 3 size divisions were identified: small, medium and large (Fig. 38A), whereas S1 had 4 sizes: small, medium, large and extra large (Fig. 37A).

We calculated the proportions of cells labelled for each of the PV-associated genes in S1 and to assess whether specific morphologies are specific to different groups of genes. To compare these to the expected distribution of cells labelled for PV-associated genes, we used the chi-square test for goodness of fit. Across all clusters, the proportions of genes are different from the expected proportions shown in figure 36. In both S1 and V1, the associated genes more readily identify morphologies of PV+ interneurons with smaller median cell sizes (Figs. 34B, 35B). PV-associated genes are more distributed across these clusters, whereas among clusters with larger median cell size, specific genes account for large proportions of the cells labelled for

PV-associated genes. Among the large and extra-large size divisions, cluster 9 was the only cluster without a largely different distribution of genes than expected. In V1, similar to S1, most of the PV-associated genes concentrate across the smaller size divisions. The medium and large size divisions display greater restriction to a smaller subset of genes, with the largest proportion of non-PV+ cells in V1 being labelled for Calb1 (p = 5.207e-09). Despite the greater representation of our 13-gene panel across the smaller size divisions, our panel of 13 genes does identify some larger cells among fewer of the genes, with the greatest proportions of cells being labelled for Cacha2d3 in S1 or Calb1 in V1.





A Cluster Morphology Phenotypes

Figure 34 - *Characterizing morphology clusters for PV genes in S1* - (A) Morphology phenotype heatmap for the morphology clusters determined for S1. Phenotypes are constructed using the median z-score values for the top-weighted morphology features for each cluster, and organized with a row-wise dendrogram. Cluster numbers reflect the progression of median cell sizes (Cluster 1 : lowest median cell size , Cluster 13 : highest median cell size). (B) Proportions of cells labelled for the 13 genes in each morphology cluster. The proportions were compared to the global proportion of cells labelled for each gene in the dataset. P-values were determined using the chi-square test to assess goodness of fit.



A Cluster Morphology Phenotypes

B Prorportions of genes by morphology cluster



Figure 35 - *Characterizing morphology clusters for PV*+ *genes in V1*. (A) Using the top-weighted features, a morphology phenotype was constructed for each cluster and organized by row-wise dendrograms. The clusters are ordered and numbered according to the progressions of median sizes (Cluster 1 : lowest median cell size, Cluster 10: highest median cell size). The first branch point on the row-wise dendrogram separates features into size and shape parameters. (B)The proportions of PV-associated genes are shown in pie charts for each morphology cluster. Proportions are compared against the global proportion of cells for each gene, p-values were determined using the chi-square test for goodness of fit .



Global Proportions of PV-Associated Genes

A Primary Somatosensory Cortex (S1)

Figure 36 - Expected proportions of PV-associated genes for S1 and V1. Pie charts showing the proportion of cells labelled for each PV-associated gene in (A) S1 and (B) V1. These are the proportions of cells taken across the entire dataset of PV-associated genes for S1, and V1 respectively. These are the comparators used in the chi-square tests presented in figure 35.







Figure 38 - *Size divisions and shape motifs for PV+ gene morphology clusters in V1.* The median z-scored values for the top-weighted size and shape parameters were clustered separately using unsupervised hierarchical clustering (Ward.D2 method) to identify the groups of clusters in V1 with similar sizes, and similar shape characteristics. Dendrograms showing the grouping of morphology clusters into (A) three size divisions, and (B) three shape motifs. The number of size divisions and shape motifs were determined using the elbow method.

To compare the laminar location of the different PV-associated gene in each morphology cluster, we mapped the profiles for each gene in the morphology clusters (Figs. 39 & 40). We only plotted data for genes where there were at least 10 cells in the morphology cluster labelled for that gene. Among the small size divisions, in S1, PV+ cells display a unimodal, infragranular peak across layers 5 and 6A. The medium and larger size divisions display a different distribution with two peaks. A smaller peak across the supragranular and granular layers (L2/3-4), and a larger peak across infragranular data. In V1, PV+ cells follow a similar distribution across layers among small and medium size divisions, the large size division displays a sustained, more uniform density of cells across the layers. These findings indicate that different PV+ morphologies are found across different layers, which alludes to transcriptomic differences among PV+ cells (Gouwens et al., 2019).

In both cortical areas, there is a greater representation of our panel of 13 genes across the small size divisions along with the smaller end of the medium size division (clusters 1-6 in S1 and V1). These distributions indicate a broad laminar distribution for these morphologies, however within each of these clusters, the laminar distributions of individual genes are separating across different groups of layers. For cluster 12 in S1, the laminar distribution of Cacna2d3 overlaps with the infragranular peak of PV+ cells (Fig. 39). Whereas in cluster 8, the distribution of PV+ cells and Cacna2d3+ cells appear out of phase with one another. In cluster 12, a subset of PV+ cells and Cacna2d3+ cells share a common morphology with comparable laminar distributions. This would allude to Cacna2d3+ cells in cluster 12 being of a specific PV+ T-type that is not shared with the Cacna2d3+ cells in cluster 8.

In V1, among the large size divisions, there are fewer genes represented across individual morphology clusters compared to small size divisions. This is consistent with what is seen in S1.

However, among the large size division, we can see the lack of alignment between PV and several of the associated genes. For example, in cluster 10, there is a low PV+ cell density across the supragranular layers, where Calb1+ cell density is high (Fig. 40). This combination of common morphology but lack of alignment with th PV+ laminar profile could suggest that the Calb1+ cells identified in cluster 10 do not share the transcriptomic correlates of the PV+ cells identified in the same cluster.

Notably, Etv1 was found across several morphology clusters, despite its characteristic restriction to layer 5 (Figs. 39 & 40). This finding establishes that several cell morphologies can exist among cells labelled for a single gene, and that for a single cortical layer, there is still a distribution of cell morphologies that are present.

Together, these analyses of PV-associated cells highlight that cell morphology can add another dimension to studying the diversity of PV+ cells. Our approach uncovers robust morphological complexity, and establishes that subtle morphological differences for a sufficiently large number of morphology parameters, for a sufficiently large number of cells, can yield pronounced differences in size and shape that distinguish among classically homogeneous cell classes.



Figure 39 - *Laminar profiles for PV-associated genes : genes by morphology cluster for S1.* In each morphology cluster, the Z-scored linear density profile was plotted along the standard cortical depth for each gene. In each morphology cluster, curves were plotted for a gene if at least 10 cells were labelled within a morphology cluster.



Figure 40 - *Laminar profiles for PV-associated genes : genes by morphology cluster for V1.* In each morphology cluster, the Z-scored linear density was plotted along the standard cortical depth for each gene. In each morphology cluster, singe curves were plotted for a gene if at least 10 cells were present within the morphology cluster.

2.4 Discussion

Our approach is a novel data-driven framework to analyze cell morphologies. By obtaining high-dimensional cell morphology data, our approach aligns analyses of neuroanatomical data with approaches applied in single-cell transcriptomics studies. Our image analysis procedure constructs a high-dimensional dataset that quantifies 97 morphometric parameters per cell. We leverage unsupervised clustering and dimensionality reduction to cluster cells based on patterns in their morphologies, and visualize these clusters in their high-dimensional morphology space. Our morphology data extend beyond classic morphological data in neuroanatomical studies, as we also quantify the shapes of single cells, and incorporate shape information together with size in defining our morphology clusters. Our inclusion of Zernike moments serve as an important shape parameter, as these moments capture complex shape features, including aspects of cell geometry and symmetry (Pincus and Theriot, 2007). Because of their invariance to scale and rotation, Zernike moments provide a basis to consistently quantify the morphologies of PV+ interneuron cell bodies, capturing salient shape features that differentiate among different groups of cells. Our use of the RSKC and denSNE algorithms served as powerful ways to both define morphology clusters, but also interrogate the relative importance and contribution of single morphometric features. RSKC is particularly powerful for our study as feature weights identify which morphometric features most strongly separate PV+ cells(Kondo et al., 2016). The weights were also important in characterizing the aspects of cell shape that further distinguish among cells of similar size. We incorporated the weights from RSKC into the denSNE algorithm, to visualize the clustering result but also how these morphometric parameters organize cells in their high-dimensional space (Balsor et al., 2021). By combining both high-dimensional data and an unsupervised cluster analysis approach,

we provide a more nuanced, reproducible characterization of PV+ interneuron morphologies, by quantifying aspects of cell morphology that have remained beyond the reach of classic, morphological analyses(Lingley et al., 2018; Bjerke et al., 2021). Our approach bridges classical neuroanatomical techniques, with modern data-driven analyses. This novel, and accessible approach allows us to construct comparable data structures and apply a highly sensitive approach to characterize single PV+ interneuron morphologies in mouse S1 and V1.

However, despite these strengths, there are important limitations that must be considered. All of the mice in the Allen Mouse ISH Database are male, so we cannot comment on potential sex differences in PV+ cell morphology. Moreover, our study compares between S1 and V1, and does not account for the presence for additional PV+ cell morphologies that could be present in other cortical or subcortical areas. These additional morphologies, which were not profiled in this analysis, may differ from the morphologies we identified here. As a result, our analysis does not necessarily represent a complete repertoire of brain-wide PV+ interneuron morphologies. In our integrated data analysis, we combined data from ISH and whole-cell fill. Despite the overlap of size distributions for cells obtained from either labelling technique, the filled cells we analysed did not represent the full range of sizes identified in the ISH data, as they mostly fell into clusters with larger median cell sizes. This is due in part to the low number of filled cells (n=36), and that small cells are more difficult to fill and record from (Ritzau-Jost et al., 2021). Consequently, it is unknown for the data obtained from the Allen Cell Types Database, whether smaller cells are underrepresented due to technical challenges with the whole-cell fills, or whether tissue processing steps affect the apparent sizes of cells across the full size range.

Investigating cell morphology has garnered attention, especially in other cell types where dynamic morphological change is associated with changes in functional state (Nimmerjahn et al., 2005). For example, in microglia, brain-resident macrophages, morphological changes associated with Alzheimer's disease have been shown to follow sexually dimorphic patterns, with different disease-related changes occurring in females than males (Guillot-Sestier et al., 2021). Additionally, in animal models of Alzheimer's disease, microglial morphological change has been shown to occur earlier in females (Colombo et al., 2022). In terms of our cluster analysis, our results provide intriguing insights into the subtle morphological differences among PV+ cells, however, it remains unknown at this time whether there are direct functional differences associated with the different morphologies we have uncovered.

We also explored the relationship between gene expression and cell morphology, to interrogate how single cell morphologies recapitulate the transcriptomic patterns seen among PV+ interneurons. It is important to recognize that our gene panel is limited as this panel does not represent the entire panel of marker genes associated with PV+ transcriptomic types (Tasic et al., 2016), these were the marker genes with good quality available in the Allen Institute mouse ISH database. As we found our 13-gene panel was more represented among smaller cell morphologies, an expanded gene panel for more of the marker genes stated across major cell type taxonomy studies would be needed to canvas the range of morphological types we have discovered. We found similar gene distributions in S1 and V1, and this suggests alignment with our first analysis of cell morphologies of PV+ interneurons from S1 and V1 together. This has identified no specific morphologies that distinguish V1 from S1, but that common morphologies occur in different proportions. As V1 has been identified as transcriptomically distinct from other cortical areas (Jorstad et al., 2023), our initial results suggest that there may be morphological

differences that distinguish V1 from other cortical areas. Although we did not observe distinct morphologies between S1 and V1, it is possible that either the distribution of PV+ morphological types in V1 differs from other cortical areas, or that certain morphologies outside of V1 are also present. This is an important avenue that our approach can assess directly, by analyzing data from several cortical areas, and comparing the distributions of different cell morphologies that are identified across other cortical areas. Capturing the distributions of cell morphologies across several cortical areas would also further align our approach with the brain-wide approaches used to profile the transcriptomic landscapes of several brain areas (Langlieb et al., 2023; Zhang et al., 2023).

Quantifying and analyzing cell morphologies has been established as a powerful quantitative tool for understanding how cell structures change in response to complex biological processes like aging, disease and dysfunction. cell morphology studies have revealed novel morphological biomarkers that serve as the structural correlates of age, sex, and disease-specific morphology patterns for different cell types (Prasad and Alizadeh, 2019; Hagemann et al., 2021; Colombo et al., 2022; Kamat et al., 2024). Assessments of morphological change of PV+ cells in disease needs to be further elucidated as current approaches have emphasised the dendritic arbour together with the spatial networks of axonal fibres (DeFelipe, 1999; Pierri et al., 1999; Song et al., 2023) , with less attention given to the size and shape of the cell body. Given the cell body's role in coordinating cellular energy and the high metabolic demand of fast-spiking PV+ interneurons, altered cell body morphologies could underscore the heightened vulnerabilities of PV+ interneurons in neuropsychiatric disorders where cell damage due to metabolic stress impacting PV+ interneurons (Whittaker et al., 2011; Steullet et al., 2017; Ruden et al., 2021). Our approach responds to the pertinent need for a more comprehensive approach to capturing

the morphological diversity of PV+ cells. Our ability to resolve and quantify the morphological patterns among the cell body is a powerful additional insight into PV+ interneuron structure and provides novel structural dimensions to explore in disease states. Our approach provides novel insight into the structural changes of PV+ cells across all major neuronal compartments, and has the potential to enrich our understanding of how morphological changes impact PV+ cell dysfunction in disease.

Taken together, we have developed and applied a novel approach to quantitative neuroanatomy that readily obtains high-dimensional, cell morphology data for large numbers of individual cells. The datasets we can obtain are comparable in structure and dimension to those generated in single-cell RNA sequencing studies and offer a powerful new platform to explore cell morphology at a scale that was previously unattainable from classic neuroanatomical image data. Classic approaches to defining cell morphologies rely heavily on expert annotations, which are largely descriptive, and often labor-intensive. Consequently, this approach becomes time-intensive and impractical if applied to larger datasets, in addition to issues with reproducibility across datasets. Our consistent image processing steps, automated morphology quantification, and unsupervised clustering techniques overcomes the subjectivity and time-intensity previously associated with analyses of cell morphology. Our use of open software platforms also helps make large-scale, high-throughput workflows accessible to a wider range of researchers. In sum, our analytical framework is a powerful tool to unlock cell morphologies at an unprecedented scale, with the potential to afford novel insight into the dynamic changes in cell morphology during development, disease progression, or in response to novel treatments.

M.Sc. Thesis - M. Panday; McMaster University - Neuroscience

Chapter 3. General Discussion

Aligning Single-Cell Transcriptomics and Morphology Data

Our approach to characterizing the cell morphologies of PV+ interneurons leverages a data structure comparable to those obtained in single-cell transcriptomics studies. With sufficiently many features for a comparably large number of cells, we achieve a suitably extensive morphological profile that describes single cell body morphologies in a comprehensive manner. We can recapitulate subtle morphological differences that identify morphological variation among single cells separating them into distinct morphological cell types. Whereas single-cell transcriptomic classifications assign T-type classifications by differences in gene expression (Haque et al., 2017), we define morphological types in terms of differences in single morphometric features.

Compared to classic approaches where cell morphologies are described entirely by an observer, and subsequently grouped based on qualitative observations (O'Kusky and Colonnier, 1982a, 1982b; Cajal, 1995), our approach treats cell morphology in an unbiased manner, defining empirical morphological types from quantitative morphology data, using an unsupervised clustering algorithm. Those morphological types can be defined along gradients of individual morphometric features. These served as organizing principles across the high-dimensional cell morphology space, where common shape motifs repeat across a range of sizes. This approach is a fundamental change to the ways in which cell morphologies are quantified and subsequently analyzed, to better reflect and capture the morphological diversity of individual cells.

The advent of single-cell transcriptomics has identified diverse transcriptomic landscapes across the brain, revealing area-specific cell type taxonomies (Siletti et al., 2023). In the human brain,

V1 was found to have a greater number of excitatory and inhibitory neuron types compared to other cortical areas, suggesting that this increased molecular diversity gives rise to more specialized cell types that are required for processing visual information (Jorstad et al., 2023). Conversely, cell morphologies have not been characterized so widely across the brain; with emergent studies focusing intently on microglia (Colombo et al., 2022; Kim et al., 2024). Whereas our approach does not define specific morphologies that are unique to either V1 or S1, our analysis did find morphologies in greater proportion in either V1 or S1, with the largest cells found more in S1. We would need to validate whether these area-biased morphologies are defined across other cortical areas first, to determine whether single cell morphologies can distinguish cortical areas.

Cell morphologies, from recent years, have focused intently on the structures of dendritic and axonal arbors, which help define the circuits neurons form, in terms of their connection and projection patterns(Peng et al., 2021). These structures have been described using high-dimensional data, and are more commonly used when integrating morphology with other single-cell data modalities (e.g., transcriptomics, or electrophysiology)(Gouwens et al., 2019, 2020; Lee et al., 2021; Chartrand et al., 2023). These multimodal approaches integrate structural and functional properties of single cells to define cell types that capture more diversity across several levels of organization. However, these studies often overlook soma morphologies, describing neuronal morphology in terms of only their arbors. This presents a unique problem for PV+ interneurons, where known, classically defined morphologies (i.e., basket and chandelier cells), possess different cell body morphologies. Together with elaborate descriptions of the synaptic contacts arranged around these cells, differences in cell body appearance played a large

role in defining these cells without complete visualization of their arbour structures (Lewis and Lund, 1990; Ribak et al., 1990). With our quantitative approach, a defined, morphological phenotype for PV+ basket and chandelier cell morphologies could be elucidated to identify potential basket and chandelier cells empirically, from classic neuroanatomical techniques. This can serve as a powerful approach to readily characterize potential morphological subtypes and map their spatial distributions across the cortical layers.

In sum, our approach for cell morphology analysis can capture morphological diversity in a manner consistent with the way we capture molecular diversity among single cells. We overcome the "classic" descriptive approaches in analyzing neuroanatomical data, by leveraging subtle, but meaningful morphometric parameters that capture several aspects of cell body size and shape. Importantly, this work identifies tremendous morphological complexity among PV+ interneurons that was previously indiscernible from analyses of classic neuroanatomical data. In turn, this work also presents PV+ interneurons in terms of their morphological complexity, overturning the long-held impression of homogeneity among PV+ interneurons when labelled in classic neuroanatomy (brederode et al., 1991). Consequently, this would suggest we can use cell morphologies in a comparable manner to single-cell transcriptomics; to understand and reconcile cellular diversity in complex processes, such as development, aging, disease. Leveraging a data-driven, morphology-based approach would probe the structural correlates of these complex processes, and afford a more nuanced understanding of single cell morphologies.

Host Factors can Influence Cell Morphology Features

Our approach provides a foundation for a more nuanced understanding of morphological differences at the level of single cells. This is a valuable extension of morphological analyses of single-cell data, which have readily been applied to understand morphological changes in aging and disease. As in neuroanatomy, histopathological investigations leverage limited sets of morphology features to define the morphology of healthy and diseased cells (Gil et al., 2002). Whereas this serves as a gold standard, the need for sensitive computational approaches to analyzing cell morphologies was also expressed (Gil et al., 2002; Carpenter et al., 2006). Our approach helps address this need, by providing a sensitive, data-driven framework to analyzing high-dimensional morphology data.

We were able to define subtle shape differences in terms of single features, which provides a precise, measurable series of parameters that quantify major differences in single cell morphology. We achieved this through the deliberate use of the RSKC algorithm. RSKC provides feature weights, which afford insight into the relative contributions of the different morphology features used in unsupervised clustering (Kondo et al., 2016; Balsor et al., 2021). Our approach was able to identify subtle shape differences that distinguish among cells of similar size. From the weights obtained by RSKC, we found these subtle shape differences are principally driven by differences in eccentricity, and low order Zernike Moments. These differences contribute to the elaboration of shape motifs that repeat across the increasing sizes. Similarly, work by Kamat and colleagues identified cell morphologies associated with both senescence and recovery following exposure to pharmacological interventions (Kamat et al., 2024). Whereas our work profiled the morphologies of PV+ interneurons from control animals

at a single age, work by Kamat and colleagues lends strong support to the possibility of identifying a core set of morphological parameters that define representative morphological changes among different groups of cells. Our work agrees with their central finding that high-dimensional cell morphologies capture subtle differences that are not readily discernible by eye (2024).

Taken together, the influence of host factors on single cell morphologies provides novel, nuanced insights into understanding aging, and disease states of single cells. Approaches leveraging high-dimensional cell morphology features uncover robust morphological differences in terms of age, disease state, or sex. These differences indicate cell morphology can capture biologically salient and relevant information about the functional states of single cells. Investigating the influence of host factors on PV+ interneurons is necessary to better understand their roles in development and disease. As PV+ interneurons have already been implicated in stress responses (Woodward and Coutellier, 2021), addiction (Fabian et al., 2024), and neurodegenerative disorders (Terstege and Epp, 2023) in a sexually dimorphic fashion. These sex differences in how PV+ interneurons are implicated in various diseases strongly motivates the application of our approach to characterizing sex differences in PV+ interneurons. As this is still largely unstudied among neuronal populations, with most of the work performed in microglia (Colombo et al., 2022; Guillot-Sestier et al., 2021).

Importantly, documented sex differences for PV+ interneurons are largely characterized at the cell-population or physiological level, overlooking how differences in single cell morphologies

differ from control states, and whether representative morphologies in PV+ interneurons are indicative of disease across various different disease states (Woodward and Coutellier, 2021; Terstege and Epp, 2023). In sum, single cell morphologies can capture salient features associated with aging and disease, which can enrich our understanding of PV+ interneuron involvement in the pathophysiology of several brain disorders. At a single-cell level, PV+ interneuron morphologies may recapitulate important characteristics that reflect the influence of specific host factors, or the progression of specific disease states (Hagemann et al., 2021; Kamat et al., 2024). These insights will help provide a more unified, holistic understanding of PV+ interneuron involvements in aging and across a range of brain disorders.

Conclusions and Future Research Directions

Our approach to characterizing cell morphologies of PV+ interneurons raises important open questions that stem from both insights into single-cell transcriptomics, and emerging studies of cell morphologies. By aligning the two modalities, it would be of value to pair the two modalities to assess how high-dimensional morphology features overlap with transcriptomic approaches to better ground the transcriptomic classifications in visible cell morphometric features. There has been recent work exploring how automated cell reconstructions for populations of transcriptomically diverse cell types, can be used to connect morphological properties with specific patterns of gene expression. However, despite being able to manually segment and identify the soma, this work focuses intently on arbour morphology. (Gliko et al., 2024). The approach leveraged in the work by Gliko and colleagues highlights the potential to integrate

soma morphologies with the morphologies of other cellular compartments. Our pipeline can serve as a platform to directly quantify and connect the soma and dendritic morphologies of filled cells from different transcriptomic types, providing a more direct connection between the soma and arbour morphologies for different transcriptomic types.

Our approach provides robust morphological profiles of the cell body, in a manner consistent with transcriptomic profiles of single cells (Haque et al., 2017; Chen et al., 2019). Second, complex processes like sensory cortical development occur along protracted timescales (Stiles and Jernigan, 2010; Rupert and Shea, 2022), and characterizing morphological changes at the level of single cells would afford insight into our understanding of typical and disordered development and the complement of morphologies that develop, together with the spatial arrangement of cell morphologies in the cortex. Third, because complex processes are shaped by the influence of host factors, for PV+ interneurons, there is value in characterizing sex differences in cell morphology. This would help characterize sexually dimorphic patterns of morphological change among PV+ interneurons or other cell types.

This work establishes cell morphologies as a critical component in reconciling the diversity of PV+ interneurons in the brain. The elaboration of distinct morphological types underscores additional complexity that has been previously overlooked in classic neuroanatomical data. By mirroring the data structure and analysis procedures used in single-cell transcriptomics datasets, our approach aligns analyses of cell morphology with current approaches to capture the diversity in single-cell transcriptomics. Our approach modernizes the classic practice of defining morphological cell types, with robust morphology quantification and a suitable high-dimensional

analysis framework..By leveraging high-dimensional morphology data, we define morphological cell types in a quantitative, unsupervised manner, in lieu of using elaborate qualitative descriptions to characterize cell shape. Our approach provides a novel, and robust approach to characterizing cell morphologies that elevate the power of classical techniques, together with the depth of knowledge that can be gained from classic anatomical data.

Chapter 4. References

Alizadeh E, Lyons SM, Castle JM, Prasad A (2016) Measuring systematic changes in invasive cancer cell shape using Zernike moments. Integr Biol 8:1183–1193.

Attali D, Baker C (2023) CRAN: Package ggExtra. Available at: https://cran.r-project.org/web/packages/ggExtra/index.html.

Baliga VB, Armstrong MS, Press ER (2023) CRAN: Package pathviewr. Available at: https://cran.r-project.org/web/packages/pathviewr/index.html.

Balsor JL, Arbabi K, Singh D, Kwan R, Zaslavsky J, Jeyanesan E, Murphy KM (2021) A Practical Guide to Sparse k-Means Clustering for Studying Molecular Development of the Human Brain. Front Neurosci 15:668293.

Bjerke IE, Yates SC, Laja A, Witter MP, Puchades MA, Bjaalie JG, Leergaard TB (2021) Densities and numbers of calbindin and parvalbumin positive neurons across the rat and mouse brain. iScience 24:101906.

Blümcke I, Hof PR, Morrison JH, Celio MR (1990) Distribution of parvalbumin immunoreactivity in the visual cortex of Old World monkeys and humans. J Comp Neurol 301:417–432.

Boland MV, Markey MK, Murphy RF (1998) Automated recognition of patterns characteristic of subcellular structures in fluorescence microscopy images. Cytometry 33:366–375.

Brederode JFMV, Mulligan KA, Hendrickson AE (1990) Calcium-binding proteins as markers for subpopulations of GABAergic neurons in monkey striate cortex. J Comp Neurol 298:1–22.

Brederode JFM van, Helliesen MK, Hendrickson AE (1991) Distribution of the calcium-binding proteins parvalbumin and calbindin-D28k in the sensorimotor cortex of the rat. Neuroscience 44:157–171.

Brückner G, Brauer K, Härtig W, Wolff JR, Rickmann MJ, Derouiche A, Delpech B, Girard N, Oertel WH, Reichenbach A (1993) Perineuronal nets provide a polyanionic, glia-associated form of microenvironment around certain neurons in many parts of the rat brain. Glia 8:183–200.

Brunson JC (2023) Alluvial Plots in ggplot2. Available at: https://cran.r-project.org/web/packages/ggalluvial/vignettes/ggalluvial.html.

Cajal SRY (1995) Histology of the Nervous System of Man and Vertebrates: Volume Two. New York: Oxford University Press. Available at: https://archive.org/details/histologyofnervo0002ramo/page/n7/mode/2up?view=theater.

Carceller H, Gramuntell Y, Klimczak P, Nacher J (2023) Perineuronal Nets: Subtle Structures with Large Implications. Neurosci 29:569–590.

Carceller H, Guirado R, Ripolles-Campos E, Teruel-Marti V, Nacher J (2020) Perineuronal Nets Regulate the Inhibitory Perisomatic Input onto Parvalbumin Interneurons and γ Activity in the Prefrontal Cortex. J Neurosci 40:5008–5018.

Cardin JA, Carlén M, Meletis K, Knoblich U, Zhang F, Deisseroth K, Tsai L-H, Moore CI (2009) Driving fast-spiking cells induces gamma rhythm and controls sensory responses. Nature 459:663–667.

Carpenter AE, Jones TR, Lamprecht MR, Clarke C, Kang IH, Friman O, Guertin DA, Chang JH, Lindquist RA, Moffat J, Golland P, Sabatini DM (2006) CellProfiler: image analysis software for identifying and quantifying cell phenotypes. Genome Biol 7:R100.

Casajus N (2020) Elbow Package (R) . Available at: https://nicolascasajus.fr/elbow/articles/introduction.html.

Celio MR, Heizmann CW (1981) Calcium-binding protein parvalbumin as a neuronal marker. Nature 293:300–302.

Chartrand T et al. (2023) Morphoelectric and transcriptomic divergence of the layer 1 interneuron repertoire in human versus mouse neocortex. Science 382:eadf0805.

Chen G, Ning B, Shi T (2019) Single-Cell RNA-Seq Technologies and Related Computational Data Analysis. Front Genet 10:317.

Colombo G, Cubero RJA, Kanari L, Venturino A, Schulz R, Scolamiero M, Agerberg J, Mathys H, Tsai L-H, Chachólski W, Hess K, Siegert S (2022) A tool for mapping microglial morphology,

morphOMICs, reveals brain-region and sex-dependent phenotypes. Nat Neurosci 25:1379–1393.

Dance A (2024) What is a cell type, really? The quest to categorize life's myriad forms. Nature 633:754–756.

DeFelipe J (1999) Chandelier cells and epilepsy. Brain 122:1807–1822.

Druga R, Salaj M, Al-Redouan A (2023) Parvalbumin - Positive Neurons in the Neocortex: A Review. Physiol Res 72:S173–S191.

Fabian CB, Jordan ND, Cole RH, Carley LG, Thompson SM, Seney ML, Joffe ME (2024) Parvalbumin interneuron mGlu5 receptors govern sex differences in prefrontal cortex physiology and binge drinking. Neuropsychopharmacology 49:1861–1871.

Fawcett JW, Oohashi T, Pizzorusso T (2019) The roles of perineuronal nets and the perinodal extracellular matrix in neuronal function. Nat Rev Neurosci 20:451–465.

Freemon FR (1996) Histology of the Nervous System of Man and Vertebrates. JAMA 275:493–493.

Fries P, Nikolić D, Singer W (2007) The gamma cycle. Trends Neurosci 30:309–316.

Gabitto MI et al. (2024) Integrated multimodal cell atlas of Alzheimer's disease. Nat Neurosci:1–18.

Galarreta M, Hestrin S (2002) Electrical and chemical synapses among parvalbumin fast-spiking GABAergic interneurons in adult mouse neocortex. Proc Natl Acad Sci 99:12438–12443.

Gil J, Wu H, Wang BY (2002) Image analysis and morphometry in the diagnosis of breast cancer. Microsc Res Tech 59:109–118.

Gliko O, Mallory M, Dalley R, Gala R, Gornet J, Zeng H, Sorensen SA, Sümbül U (2024) High-throughput analysis of dendrite and axonal arbors reveals transcriptomic correlates of neuroanatomy. Nat Commun 15:6337.
Gouwens NW et al. (2019) Classification of electrophysiological and morphological neuron types in the mouse visual cortex. Nat Neurosci 22:1182–1195.

Gouwens NW et al. (2020) Integrated Morphoelectric and Transcriptomic Classification of Cortical GABAergic Cells. Cell 183:935-953.e19.

Group PIN et al. (2008) Petilla terminology: nomenclature of features of GABAergic interneurons of the cerebral cortex. Nat Rev Neurosci 9:557–568.

Guillot-Sestier M-V, Araiz AR, Mela V, Gaban AS, O'Neill E, Joshi L, Chouchani ET, Mills EL, Lynch MA (2021) Microglial metabolism is a pivotal factor in sexual dimorphism in Alzheimer's disease. Commun Biol 4:711.

Gulyás AI, Megías M, Emri Z, Freund TF (1999) Total Number and Ratio of Excitatory and Inhibitory Synapses Converging onto Single Interneurons of Different Types in the CA1 Area of the Rat Hippocampus. J Neurosci 19:10082–10097.

Hagemann C, Tyzack GE, Taha DM, Devine H, Greensmith L, Newcombe J, Patani R, Serio A, Luisier R (2021) Automated and unbiased discrimination of ALS from control tissue at single cell resolution. Brain Pathol 31:e12937.

Haque A, Engel J, Teichmann SA, Lönnberg T (2017) A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications. Genome Med 9:75.

Hensch T, Stryker M, Kash S (1998) Local GABA circuit control of experience-dependent plasticity in developing visual cortex. Sci New York N Y 282:1504–1508.

Ho J, Tumkaya T, Aryal S, Choi H, Claridge-Chang A (2019) Moving beyond P values: data analysis with estimation graphics. Nat Methods 16:565–566.

Hodge RD et al. (2019) Conserved cell types with divergent features in human versus mouse cortex. Nature 573:61–68.

Hu Y, Murphy RF (2004) Automated interpretation of subcellular patterns from immunofluorescence microscopy. J Immunol Methods 290:93–105.

Janickova L, Rechberger KF, Wey L, Schwaller B (2020) Absence of parvalbumin increases mitochondria volume and branching of dendrites in inhibitory Pvalb neurons in vivo: a point of convergence of autism spectrum disorder (ASD) risk gene phenotypes. Mol Autism 11:47.

Jiang X, Shen S, Cadwell CR, Berens P, Sinz F, Ecker AS, Patel S, Tolias AS (2015) Principles of connectivity among morphologically defined cell types in adult neocortex. Science 350:aac9462.

Jorstad NL et al. (2023) Transcriptomic cytoarchitecture reveals principles of human neocortex organization. Science 382:eadf6812.

Kamat P, Macaluso N, Min C, Li Y, Agrawal A, Winston A, Pan L, Starich B, Stewart T, Wu P-H, Fan J, Walston J, Phillip JM (2024) Single-cell morphology encodes functional subtypes of senescence in aging human dermal fibroblasts. bioRxiv:2024.05.10.593637.

Kassambara A, Mundt F (2020) CRAN: Package factoextra. Available at: https://cran.r-project.org/web/packages/factoextra/index.html.

Kepecs A, Fishell G (2014) Interneuron cell types are fit to function. Nature 505:318–326.

Kim J, Pavlidis P, Ciernia AV (2024) Development of a High-Throughput Pipeline to Characterize Microglia Morphological States at a Single-Cell Resolution. eNeuro 11:ENEURO.0014-24.2024.

Kondo Y, Salibian-Barrera M, Zamar R (2016) RSKC: An RPackage for a Robust and Sparse K-Means Clustering Algorithm. J Stat Softw 72:1–26 Available at: https://www.jstatsoft.org/article/view/v072i05.

Kooijmans RN, Sierhuis W, Self MW, Roelfsema PR (2020) A Quantitative Comparison of Inhibitory Interneuron Size and Distribution between Mouse and Macaque V1, Using Calcium-Binding Proteins. Cereb Cortex Commun 1:tgaa068. Langlieb J, Sachdev NS, Balderrama KS, Nadaf NM, Raj M, Murray E, Webber JT, Vanderburg C, Gazestani V, Tward D, Mezias C, Li X, Flowers K, Cable DM, Norton T, Mitra P, Chen F, Macosko EZ (2023) The molecular cytoarchitecture of the adult mouse brain. Nature 624:333–342.

Lee BR et al. (2021) Scaled, high fidelity electrophysiological, morphological, and transcriptomic cell characterization. eLife 10:e65482.

Lee S-H, Kwan AC, Zhang S, Phoumthipphavong V, Flannery JG, Masmanidis SC, Taniguchi H, Huang ZJ, Zhang F, Boyden ES, Deisseroth K, Dan Y (2012) Activation of specific interneurons improves V1 feature selectivity and visual perception. Nature 488:379–383

Lein ES et al. (2007) Genome-wide atlas of gene expression in the adult mouse brain. Nature 445:168–176.

Leitch B (2024) Parvalbumin Interneuron Dysfunction in Neurological Disorders: Focus on Epilepsy and Alzheimer's Disease. Int J Mol Sci 25:5549.

Levelt CN, Hübener M (2012) Critical-Period Plasticity in the Visual Cortex. Annu Rev Neurosci 35:309–330.

Lewis DA, Lund JS (1990) Heterogeneity of chandelier neurons in monkey neocortex: Corticotropin-releasing factor-and parvalbumin-immunoreactive populations. J Comp Neurol 293:599–615.

Leyh J, Paeschke S, Mages B, Michalski D, Nowicki M, Bechmann I, Winter K (2021) Classification of Microglial Morphological Phenotypes Using Machine Learning. Front Cell Neurosci 15:701673.

Lingley AJ, Bowdridge JC, Farivar R, Duffy KR (2018) Mapping of neuron soma size as an effective approach to delineate differences between neural populations. J Neurosci Methods 304:126–135.

Lo SQ, Sng JCG, Augustine GJ (2017) Defining a critical period for inhibitory circuits within the somatosensory cortex. Sci Rep 7:7271.

Maaten LV der, Hinton G (2008) Visualizing Data using t-SNE. Journal of Machine Learning Research 9:2579–2605.

Mao X, Staiger JF (2024) Multimodal cortical neuronal cell type classification. Pflügers Arch - Eur J Physiol 476:721–733.

Marín O (2024) Parvalbumin interneuron deficits in schizophrenia. Eur Neuropsychopharmacol 82:44–52.

Mayer C, Hafemeister C, Bandler RC, Machold R, Brito RB, Jaglin X, Allaway K, Butler A, Fishell G, Satija R (2018) Developmental diversification of cortical inhibitory interneurons. Nature 555:457–462.

Micheva KD, Chang EF, Nana AL, Seeley WW, Ting JT, Cobbs C, Lein E, Smith SJ, Weinberg RJ, Madison DV (2018) Distinctive Structural and Molecular Features of Myelinated Inhibitory Axons in Human Neocortex. eNeuro 5:ENEURO.0297-18.2018.

Mikkonen M, Soininen H, Pitkänen A (1997) Distribution of parvalbumin-, calretinin-, and calbindin-D28k–immunoreactive neurons and fibers in the human entorhinal cortex. J Comp Neurol 388:64–88.

Morales B, Choi S-Y, Kirkwood A (2002) Dark Rearing Alters the Development of GABAergic Transmission in Visual Cortex. J Neurosci 22:8084–8090.

Narayan A, Berger B, Cho H (2021) Assessing single-cell transcriptomic variability through density-preserving data visualization. Nat Biotechnol 39:765–774.

Nimmerjahn A, Kirchhoff F, Helmchen F (2005) Resting Microglial Cells Are Highly Dynamic Surveillants of Brain Parenchyma in Vivo. Science 308:1314–1318.

Niu K, Tian C (2022) Zernike polynomials and their applications. J Opt 24:123001.

O'Kusky J, Colonnier M (1982a) A laminar analysis of the number of neurons, glia, and synapses in the visual cortex (area 17) of adult macaque monkeys. J Comp Neurol 210:278–290.

O'Kusky J, Colonnier M (1982b) Postnatal changes in the number of neurons and synapses in the visual cortex (area 17) of the macaque monkey: A stereological analysis in normal and monocularly deprived animals. J Comp Neurol 210:291–306.

Peng H et al. (2021) Morphological diversity of single neurons in molecularly defined cell types. Nature 598:174–181.

Pierri JN, Chaudry AS, Woo T-UW, Lewis DA (1999) Alterations in Chandelier Neuron Axon Terminals in the Prefrontal Cortex of Schizophrenic Subjects. Am J Psychiatry 156:1709–1719.

Pincus Z, Theriot JA (2007) Comparison of quantitative methods for cell-shape analysis. J Microsc 227:140–156.

Prasad A, Alizadeh E (2019) Cell Form and Function: Interpreting and Controlling the Shape of Adherent Cells. Trends Biotechnol 37:347–357.

Ribak CE, Nitsch R, Seress L (1990) Proportion of parvalbumin-positive basket cells in the GABAergic innervation of pyramidal and granule cells of the rat hippocampal formation. J Comp Neurol 300:449–461.

Rio J del, Lecea L de, Ferrer I, Soriano E (1994) The development of parvalbumin-immunoreactivity in the neocortex of the mouse. Dev Brain Res 81:247–259.

Ritzau-Jost A, Tsintsadze T, Krueger M, Ader J, Bechmann I, Eilers J, Barbour B, Smith SM, Hallermann S (2021) Large, Stable Spikes Exhibit Differential Broadening in Excitatory and Inhibitory Neocortical Boutons. Cell Rep 34:108612.

Rubenstein JLR, Merzenich MM (2003) Model of autism: increased ratio of excitation/inhibition in key neural systems. Genes, Brain Behav 2:255–267.

Ruden JB, Dugan LL, Konradi C (2021) Parvalbumin interneuron vulnerability and brain disorders. Neuropsychopharmacology 46:279–287.

Rupert DD, Shea SD (2022) Parvalbumin-Positive Interneurons Regulate Cortical Sensory Plasticity in Adulthood and Development Through Shared Mechanisms. Front Neural Circuits 16:886629. Sanchez B, Kraszewski P, Lee S, Cope EC (2023) From molecules to behavior: Implications for perineuronal net remodeling in learning and memory. J Neurochem 168:1854–1876.

Seeger G, Brauer K, Härtig W, Brückner G (1994) Mapping of perineuronal nets in the rat brain stained by colloidal iron hydroxide histochemistry and lectin cytochemistry. Neuroscience 58:371–388.

Sigal YM, Bae H, Bogart LJ, Hensch TK, Zhuang X (2019) Structural maturation of cortical perineuronal nets and their perforating synapses revealed by superresolution imaging. Proc Natl Acad Sci 116:7071–7076.

Siletti K et al. (2023) Transcriptomic diversity of cell types across the adult human brain. Science 382:eadd7046.

Soetaert K, Herman PMJ (2009) A Practical Guide to Ecological Modelling (Springer, ed). Available at: https://link.springer.com/content/pdf/10.1007/978-1-4020-8624-3.pdf.

Sohal VS, Zhang F, Yizhar O, Deisseroth K (2009) Parvalbumin neurons and gamma rhythms enhance cortical circuit performance. Nature 459:698–702.

Song C et al. (2023) Spatial Distribution of Parvalbumin-Positive Fibers in the Mouse Brain and Their Alterations in Mouse Models of Temporal Lobe Epilepsy and Parkinson's Disease. Neurosci Bull 39:1683–1702.

Sorg BA, Berretta S, Blacktop JM, Fawcett JW, Kitagawa H, Kwok JCF, Miquel M (2016) Casting a Wide Net: Role of Perineuronal Nets in Neural Plasticity. J Neurosci 36:11459–11468.

Steullet P, Cabungcal J-H, Coyle J, Didriksen M, Gill K, Grace AA, Hensch TK, LaMantia A-S, Lindemann L, Maynard TM, Meyer U, Morishita H, O'Donnell P, Puhl M, Cuenod M, Do KQ (2017) Oxidative stress-driven parvalbumin interneuron impairment as a common mechanism in models of schizophrenia. Mol Psychiatry 22:936–943.

Stiles J, Jernigan TL (2010) The Basics of Brain Development. Neuropsychol Rev 20:327–348. Stirling DR, Swain-Bowden MJ, Lucas AM, Carpenter AE, Cimini BA, Goodman A (2021) CellProfiler 4: improvements in speed, utility and usability. BMC Bioinform 22:433.

Tahmasbi A, Saki F, Shokouhi SB (2011) Classification of benign and malignant masses based on Zernike moments. Comput Biol Med 41:726–735.

Tasic B et al. (2016) Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. Nat Neurosci 19:335–346.

Tasic B et al. (2018) Shared and distinct transcriptomic cell types across neocortical areas. Nature 563:72–78.

Teague MR (1980) Image analysis via the general theory of moments*. J Opt Soc Am 70:920.

Terstege DJ, Epp JR (2023) Parvalbumin as a sex-specific target in Alzheimer's disease research – A mini-review. Neurosci Biobehav Rev 153:105370.

Wang Y, Gupta A, Toledo-Rodriguez M, Wu CZ, Markram H (2002) Anatomical, Physiological, Molecular and Circuit Properties of Nest Basket Cells in the Developing Somatosensory Cortex. Cereb Cortex 12:395–410.

Whittaker RG, Turnbull DM, Whittington MA, Cunningham MO (2011) Impaired mitochondrial function abolishes gamma oscillations in the hippocampus through an effect on fast-spiking interneurons. Brain 134:e180–e180.

Wöhr M, Orduz D, Gregory P, Moreno H, Khan U, Vörckel KJ, Wolfer DP, Welzl H, Gall D, Schiffmann SN, Schwaller B (2015) Lack of parvalbumin in mice leads to behavioral deficits relevant to all human autism core symptoms and related neural morphofunctional abnormalities. Transl Psychiatry 5:e525–e525.

Woodward EM, Coutellier L (2021) Age- and sex-specific effects of stress on parvalbumin interneurons in preclinical models: Relevance to sex differences in clinical neuropsychiatric and neurodevelopmental disorders. Neurosci Biobehav Rev 131:1228–1242.

Yang SM, Michel K, Jokhi V, Nedivi E, Arlotta P (2020) Neuron class–specific responses govern adaptive myelin remodeling in the neocortex. Science 370.

Yang T, Hu S, Chang W-C, Kao H-Y, Wang Y (2023) Perineuronal Nets Degradation and Parvalbumin Interneuron Loss in a Mouse Model of DEPDC5-Related Epilepsy. Dev Neurosci 44:671–677.

Yue B (2014) Biology of the Extracellular Matrix. J Glaucoma 23:S20–S23.

Zeng H (2022) What is a cell type and how to define it? Cell 185:2739–2755.

Zeng H, Sanes JR (2017) Neuronal cell-type classification: challenges, opportunities and the path forward. Nat Rev Neurosci 18:530–546.

Zhang M, Pan X, Jung W, Halpern AR, Eichhorn SW, Lei Z, Cohen L, Smith KA, Tasic B, Yao Z, Zeng H, Zhuang X (2023) Molecularly defined and spatially resolved cell atlas of the whole mouse brain. Nature 624:343–354.

Zhu Q, Ke W, He Q, Wang X, Zheng R, Li T, Luan G, Long Y-S, Liao W-P, Shu Y (2018) Laminar Distribution of Neurochemically-Identified Interneurons and Cellular Co-expression of Molecular Markers in Epileptic Human Cortex. Neurosci Bull 34:992–1006.