Unbiased Multidimensional Analysis Reveals Novel Principles of Cortical Interneuron Synaptic Organization

Patrick Daniel Dummer

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2024

Abstract

Unbiased Multidimensional Analysis Reveals Novel Principles of Cortical Interneuron Synaptic Organization

Patrick Daniel Dummer

Neurons display exquisite specificity in synaptic connectivity, but we lack a complete understanding of neuronal connectivity and the rules that govern it. A major impediment to addressing this question lies in the vast diversity of neurons, the small size and large number of synapses formed by any given neuron over a wide territory, and the need to study these connections in intact tissue. We therefore developed an image-based tool to assess synaptic specificity in tissue sections and dissociated culture. We focused on three interneuron subpopulations that target distinct subcellular regions of the post-synaptic cell: soma-targeting basket cells (BCs), axon initial segment (AlS)-targeting chandelier cells (ChCs), and distal dendrite-targeting somatostatin cells (SstCs). Using mouse dissociated cortical culture as a starting point, we built a machine learning (ML) based image processing and analysis pipeline to classify individual presynaptic boutons at scale. Supervised ML classification revealed similar subcellular targeting profiles for these interneuron populations in slice and culture, indicating that targeting is primarily regulated by cell intrinsic

programs. We also observed a remarkable target-dependent laminar organization *in vivo*. An unsupervised ML analysis using the same input data not only identified the same three canonical targeting classes, but also revealed that these classes are comprised of multiple subpopulations. In slice, these synaptic subpopulations displayed distinct laminar organization. In dissociated culture, two soma-targeting synaptic subpopulations mapped to target cells with different cellular profiles. The six dendrite-targeting synaptic subpopulations were found at increasing distances from target soma, suggesting molecularly distinct proximal, medial, and distal dendritic compartments in culture. Tracking subtype targeting across axonal branches of individual neurons indicated that SstCs and BCs utilize distinct targeting strategies in culture that accord with established findings *in vivo*. In sum, our synaptic analysis pipeline revealed novel synaptic subpopulations in interneurons. Further analysis uncovered novel aspects of interneuron synaptic biology that, remarkably, are retained in culture.

Table of Contents

List of Charts, Graphs, Illustrations vii

List of Tables xi

Acknowledgments xii

Dedication xiii

Chapter 1: Introduction 1

1.1. Synopsis 1

1.1.1. Observation 1

1.1.2. Hypothesis 1

1.1.3. Methods 2

1.1.4. Results 3

1.2. Preface 4

1.2.1. A connectivity-centric overview of neurodevelopment 4

1.2.2. Organization of this introduction 8

1.1 Cortical Development 10

1.2.3. Cortical Interneuron development 12

1.2.4. Laminar distribution of cortical neuron types and their connections 24

1.2.5. Regional distribution of cortical neuron types and their connections 30

1.3. Subcellular compartmentalization: cellular and molecular models 44

1.3.1. Axonal subcompartments 47

1.3.2. Dendritic subcompartments 48

1.3.3. Cellular and molecular mechanisms of compartmentalization 51

1.3.4. Summary of subcellular compartmentalization 54

1.4. General principles of synaptic development 56

1.4.1. Pre-contact: a question of spatial proximity 56

1.4.2. Contact and beyond 67

1.4.3. Post-contact models of synaptic specificity 86
1.5. Targeting specificity *ex vivo* 102
1.5.1. Specificity and exclusivity of cortical interneuron connectivity *in vivo* 103
1.5.2. Dissection: isolation and control 104
1.5.3. Dissociation: a bit of chaos to disrupt intrinsic spatial organization 107
1.5.4. Summary of *ex vivo* targeting specificity 111
1.6. Summary of principles 112
1.7. Conclusion and hypothesis 113
Chapter 2: Materials and methods 133
2.1. Sample preparation and image acquisition 133
2.1.1. Animal husbandry 133
2.1.2. Genotyping 133
2.1.3. Tamoxifen Induction 134
2.1.4. Perfusions and tissue preparation 134
2.1.5. Cortical dissection and dissociation 135
2.1.6. Tissue culture 136
2.1.7. Immunostaining 137
2.1.8. Fluorescence microscopy 138
2.2. Sample Nomenclature 138
2.3. Image processing 139
2.4. Custom Code 140
2.4.1. Metadata extraction 140
2.4.2. Stitching 141
2.4.3. File Grouping 141
2.4.4. Image standardization 142
2.4.5. Image Correction 142
2.4.6. Pixel classification 143
2.4.7. Image concatenation 144
2.4.8. Object segmentation 144

2.4.9. Class mapping 144
2.5. Object classification 145
2.5.1. Supervised object classification 145
2.5.2. Unsupervised training by autoencoder embedding 145
2.5.3. Unsupervised clustering 146
2.6. Statistical analysis 147
2.6.1. Spatial clustering analysis 147
2.6.2. Statistical comparisons 148
Chapter 3: Image-based multidimensional analysis of synapses 150
3.1. Introduction 150
3.2. Experiment design 153
3.2.1. Initial question and impetus to create the pipeline 153
3.2.2. Labeling the presynaptic compartment: reporter selection 154
3.2.3. Labeling the target compartments: antibody selection 159
3.3. File organization 162
3.3.1. Nomenclature 162
3.3.2. Adding an input: standardization of metadata, files, and file structures 166
3.3.3. File Grouping and Data Organization 170
3.4. Image correction 173
3.4.1. X/Y correction 174
3.4.2. Batch correction 178
3.5. Pixel Classification 183
3.5.1. Presynaptic pixel classification 187
3.5.2. Target pixel classification 188
3.6. Object classification 190
3.6.1. Object Measurand Concatenation 191
3.6.2. Object segmentation 192
3.6.3. Object measurement 193
3.6.4. Supervised object classification 194

3.6.5. Supervised classifier validation 196
3.6.6. Ability to detect changes in subcellular targeting 198
3.6.7. Unsupervised object classification 200
3.6.8. Object class mapping 205
3.7. Conclusions 206
Chapter 4: Principles of cortical interneuron synaptic organization 209
4.1. Introduction 209
4.2. Experiment design 211
4.3. Supervised classification of cortical interneuron boutons 217
4.3.1. GABAergic interneuron subcellular targeting is recapitulated in dissociated culture 217
4.3.2. GABAergic interneuron boutons are highly organized within the cortex 221
4.4. Unsupervised clustering of cortical interneuron boutons 229
4.4.1. Developing a stable, unsupervised model of interneuron subcellular targeting in culture 230
4.4.2. Unsupervised classification of cortical interneuron boutons identifies novel subclasses 236
4.4.3. Dendrite-targeting bouton subclasses partition based on proximity to soma 240
4.4.4. Unsupervised soma-targeting boutons subtypes are associated with distinct target cell profiles 247
4.4.5. Basket and somatostatin cells display distinct axonal growth strategies in dissociated culture 250
4.4.6. Unsupervised classification in slice identifies bouton classes with distinct laminar distributions 254
4.5. Conclusions 257
Chapter 5: Discussion 260
5.1. Summary of methodological approach 260
5.1.1. A multidimensional imaging platform to assess synaptic connectivity 260

5.1.2. Comparison to other studies 262
5.1.3. Current and future imaging approaches to synaptic diversity 266
5.1.4. A hurdle to progress: batch-to-batch variation 269
5.2. Summary of biological findings 272
5.2.1. Insights into interneuron development from dissociated culture experiments 272
5.2.2. Persistence of interneuron identity 274
5.2.3. Models for cortical interneuron subcellular targeting 276
5.2.4. Redefining synaptic targeting classes 280
5.2.5. Cortical interneuron subcellular targeting in intact tissue 284
5.2.6. Interneuron targeting in the reeler cortex 287
5.2.7. Reframing synaptic specificity from studies in dissociated culture 289
5.2.8. A screening platform for neuronal connectivity 294
5.3. The synaptic potentiality landscape 298
5.3.1. Expanding the conceptual framework for synapse formation, diversity, and function 299
5.3.2. Topography of the synaptic potentiality landscape: reconciling paradoxes in neuronal connectivity 304
5.3.3. Defining synaptic state: conceptual and practical considerations 315
5.3.4. Biophysical assumptions and their implications for neuronal connectivity 317
5.3.5. Nothing is necessary, everything is sufficient: embracing a probabilistic, degenerate framework 321
5.3.6. Turning a heuristic into a predictive model 324
Tables 341
Table 1: Mouse lines used in this study 341
Table 2: Genotyping primers 342
Table 3: Tamoxifen induction of CreER driver lines 344
Table 4: Primary antibodies 345

Table 5: Secondary antibodies 346
Table 6: Saved source metadata fields by category, with data type and source 347
Table 7: Sizes of unsupervised bouton classes 351
Table 8: Permissions statements for reproduced figures. 352
Appendices 356
Appendix A: Buffer and media formulations: 356
Dissection and Dissociation media: 356
Tissue culture media formulations: 356
Tissue culture dish and plate preparation: 357
Perfusion and slice preparation buffers: 357
Appendix B: Sample nomenclature 359
General nomenclature structure for sample/image names: 359
Nomenclature Rules: 359
Field definitions: 360
Nomenclature for slides: 363
Automated Tile Grouping 364

List of Charts, Graphs, Illustrations

Unless noted in the legend, all figures are original works from the author. Permission statements for reproduced figures are only included in the figure legend if required by the reproduction agreement. The permission statements for all reproduced figures can be found in Table 8.

Figure 1.1 | GABAergic interneuron subcellular targeting specificity in vivo and model for interneuron specificty in culture. 2

Figure 1.2 | Stages of neurodevelopment (standard order). 5

Figure 1.3 | Stages of neurodevelopment (connectivity-centric view). 7

Figure 1.4 | Cortical development stages prior to circuit integration. 11

Figure 1.5 | The MGE lineage of cortical interneurons. 13

Figure 1.6 | Schematic summarizing the major MGE-derived cortical GABAergic interneurons based on their connectivity patterns. 15

Figure 1.7 | The developmental landscape reflects attractor dynamics. 19

Figure 1.8| The evolution of corticogenesis: inside-out birth and migration patterns yield ontogenicaly-related cortical columns. 25

Figure 1.9 | Lamination of the cerebral cortex. 27

Figure 1.10 | Regional organization of the mouse cortex. 31

Figure 1.11 | Spatial patterning of the cortex. 33

Figure 1.12 | Examples of topographic maps: the pathways transmitting tactile information from the muzzle to the cortex. 36

Figure 1.13| Independence of laminar and regional iden-tity: evidence from the reeler cortex. .. 40

Figure 1.14 | General localization of voltage-gated ion channels in a model neuron. 46

Figure 1.15 | Branch and sub-branch organization of morphologically defined pyramidal neuron dendritic domains. 48
Figure 1.16| Glutamatergic cell in the cortex have layered subcellular compartments. 51
Figure 1.17 | Mechanisms of axon and dendrite guidance. 57
Figure 1.18 | Lamina-specific innervation of hippocampal neurons by distinct inputs. 60
Figure 1.19 | Differential use of ephrin/Eph adhesion molecules in axon guidance during visual system development. 62
Figure 1.20 | Schematic Representation of the Bergmann Glial Fiber Scaffold and CHL1 in Directing Stellate Axons to Innervate Purkinje Dendrites. 65
Figure 1.21 | Schematic of the chemical synapse. 69
Figure 1.22 | Scale-free graph of the iPSD proteome. 73
Figure 1.23 | Visual representation of the synaptic adhesion molecule code. 77
Figure 1.24 | Subcellular specificity of cerebellar basket cell inhibitory connections to Purkinje cells (PkCs). 89
Figure 1.25 | Lamina-specific innervation of hippocampal CA3 neurons by mossy fibers (MF) is mediated by domain-restricted repulsive adhesion molecule interactions. 91
Figure 1.26| Regional regulation of synaptogenesis yields subcellular specificity. 93
Figure 1.27 | Remodelling of CF and PF synaptic connections onto PCs during postnatal cerebellar development. 95
Figure 1.28 | Molecular basis of PF-PC synapse formation and impact of targeted disruptions on PF and CF innervation domains. 98
Figure 1.29 | Model for transneuronal regulation of receptor localization by axonal netrin-G proteins. 101

Figure 3.1| A schematic displaying the five stages of the platform. 152
Figure 3.2 | Presynaptic cell labeling strategy 156
Figure 3.3 | Target Compartment labeling strategy. 161
Figure 3.4 | Input standardization and quality control checks. 169
Figure 3.5 | Generalized scene stitching 172
Figure 3.6 | Fast XY correction of tiled Z stacks. 175
Figure 3.7 | Sample-to-sample (batch) intensity correction. 179
Figure 3.8| Staining pattern of subcellular markers generally and at the AIS. 185
Figure 3.9 | Presynaptic pixel classifier. 188
Figure 3.10 | Target pixel classifier. 189
Figure 3.11 | Object classification schematic..... 191
Figure 3.12 | Object classification inputs and outputs. 195
Figure 3.13 | Supervised object classifier performance. 197
Figure 3.14 | IL13 shifts interneuron subcellular targeting preference from dendrite to soma. . 200
Figure 3.15 | supervised classification calls as autoencoded parameters rendered into UMAP space. 202
Figure 3.16 | Supervised classification calls as autoencoded parameters rendered into UMAP space..... 202
Figure 3.17 | Unsupervised classification reveals novel bouton subtypes of canonical targeting classes. 204
Figure 4.1 | GABAergic interneuron subcellular targeting specificity in vivo and model for interneuron specificty in culture. 209

Figure 4.2 | Presynaptic labeling of distinct interneuron populations using mouse driver lines and the Ai34 reporter. 212
Figure 4.3 | Target compartment labeling and bouton targeting classification strategies..... 216
Figure 4.4 | Recapitulation of interneuron subcellular targeting in a dissociated culture system. 220
Figure 4.5 | Spatial organization of the target determines interneuron bouton subtype laminar and areal enrichment, but not subtype targeting specificity. 224
Figure 4.6 | Generation of a model for interneuron bouton subcellular targeting in dissociated culture by unsupervised classification. 232
Figure 4.7 | Unsupervised classification of cortical interneuron boutons in culture identifies canonical interneuron bouton classes and reveals novel bouton subclasses. 238
Figure 4.8 | Unsupervised dendrite-targeting bouton subtypes in culture follow a proximal-to-distal spatial organization relative to the target soma..... 244
Figure 4.9 | Unsupervised soma-targeting bouton subtypes are associated with distinct target cell profiles: evidence of preferential cellular targeting in culture. 248
Figure 4.10 | Bouton subtype sequentiality reveals distinct targeting strategies for Basket and Sst cells in culture. 252
Figure 4.11 | Unsupervised classification of cortical interneuron boutons in slice identifies bouton subclasses with distinct laminar distributions. 256
Figure 5.1 | Waddington's visualization of the potential energy landscape. 299
Figure 5.2 | The complex system of interactions underlying the potential energy landscape. ... 304

List of Tables

Table 1: Mouse lines used in this study 341

Table 2: Genotyping primers 342

Table 3: Tamoxifen induction of CreER driver lines 344

Table 4: Primary antibodies 345

Table 5: Secondary antibodies 346

Table 6: Saved source metadata fields by category, with data type and source 347

Table 7: Sizes of unsupervised bouton classes 351

Table 8: Permissions statements for reproduced figures. 352

Acknowledgments

Science thrives on collaboration, and this thesis stands as a testament to the collective efforts of many individuals who supported and contributed to this research journey. I am deeply grateful to each of you for your invaluable contributions.

First and foremost, I extend my heartfelt appreciation to my collaborators and colleagues whose expertise and dedication enriched every aspect of this study. In particular, I would like to thank Sakib Hossain, Dylan Lee, Runsheng Wang, Melissa McKenzie Campbell, and Vilas Menon. Your insights, discussions, and willingness to challenge ideas have been instrumental in shaping the outcomes presented here.

I owe a debt of gratitude to my mentor Edmund Au and my thesis committee, Lori Zeltser, Franck Polleux, and Clarissa Waites for their guidance, encouragement, and unwavering belief in my abilities. Your wisdom and scholarly advice have been pivotal in shaping my academic and professional growth.

To my students and volunteers who generously contributed their time and effort to this study, I am deeply thankful for your willingness to be part of our research. Your participation has been invaluable in enhancing our understanding in this field, and you helped me grow as a mentor.

I am also grateful to my family and friends for their understanding, encouragement, and unwavering support during this challenging yet rewarding journey. Your belief in me sustained my motivation and kept me grounded throughout.

Lastly, I acknowledge the funding agencies whose financial support made this research possible. Your investment in scientific exploration plays a crucial role in driving innovation and discovery.

This thesis is a collective effort, and I am humbled and honored to have worked alongside such dedicated individuals. Although the authorship of this document bears my name, its accomplishments are the result of our shared commitment to advancing knowledge and understanding. I have used the first-person plural throughout this document to honor these contributions.

Dedication

To my beloved family, my parents Daniel Dummer & Lynda Tandanand, Kate Sinclair Dunn & Bill Dunn, my brother Joseph Dummer, my fiancée Lydia Booz, and soon-to-be mother-in-law Cynthia Hoffman:

Your unwavering support, love, and encouragement have been my guiding light throughout this academic journey. Your sacrifices and belief in me have made this achievement possible. I am eternally grateful for everything you have done.

To my respected mentors, Dr. Edmund Au, Dr. Jeffrey Kopp, and Dr. David Pickar:

Your guidance, wisdom, and expertise have shaped my intellectual growth and shaped the course of my research. I am deeply indebted to you for believing in my potential and challenging me to reach higher.

To my cherished friends, my chosen family, you know who you are:

Thank you for standing by me, understanding my occasional absences, and always being there with a listening ear and words of encouragement. Your friendship has enriched my life beyond measure.

To the mice who gave their lives for this study:

Your contribution to advancing scientific knowledge is immeasurable. Your sacrifice has not gone unnoticed, and it is with solemn gratitude that I acknowledge the role you played in our research journey. May the science gods honor your sacrifice with unending toys and chow in the great beyond.

This thesis is dedicated to each of you with profound gratitude and love.

Chapter 1: Introduction

1.1. Synopsis

1.1.1. Observation

Neurons display exquisite specificity in synaptic connectivity, but how this specificity arises is largely unknown. A major impediment to addressing this question lies in the large number of synapses formed by any given cell, their small size, and the need to study these connections in intact tissue. *In vivo*, distinct GABAergic cortical interneuron populations preferentially target specific subcellular domains of the target cell (Figure 1.1, A). Somatostatin cells (Sst, blue) target distal dendrites, basket cells (BC, red) target soma and proximal dendrites, and chandelier cells (ChC, green) target the axon initial segment.

1.1.2. Hypothesis

We hypothesized that cortical interneuron subcellular targeting specificity is largely independent of the broader cortical environment (laminar and areal positioning, afferent connections, etc.), and as such would be recapitulated in a dissociated culture setting (Figure 1.1, B). Note that recapitulation implies not only that targeting specificity is intrinsically determined by the presynaptic cell, but also that compartmentalization of the target molecules in the correct subcellular regions is intrinsically determined by the target cell.

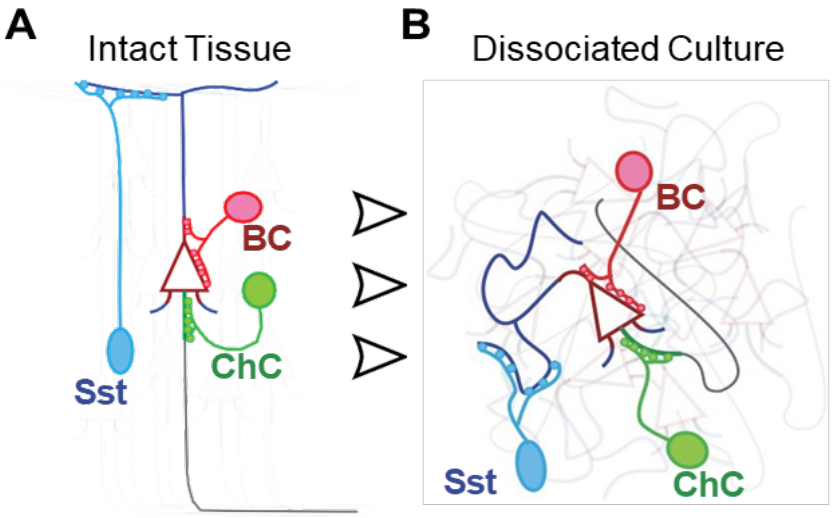

Figure 1.1 | GABAergic interneuron subcellular targeting specificity in vivo and model for interneuron specificty in culture.
A) In vivo, distinct GABAergic cortical interneuron populations target specific subcellular domains of the target cell. Somatostatin cells (Sst, blue) target distal dendrites, basket cells (BC, red) target soma and proximal dendrites, and chandelier cells (ChC, green) target the axon initial segment.
B) We hypothesized that interneuron subcellular targeting specificity is largely independent of the broader cortical environment (laminar, positioning, afferent connections, etc.), and as such would be recapitulated in a dissociated culture setting. Note that recapitulation implies not only that targeting specificity is intrinsically determined by the presynaptic cell, but also that compartmentalization of the target molecules in the correct subcellular regions is intrinsically determined by the target cell.

1.1.3. Methods

We developed an image-based tool to assess synaptic specificity in tissue sections and dissociated culture. We focused on three interneuron subpopulations that target distinct subcellular regions of the post-synaptic cell: soma-targeting basket cells (BCs), axon initial segment (AlS)-targeting chandelier cells (ChCs), and distal dendrite-targeting somatostatin cells (SstCs). Using dissociated mouse primary cortical culture as a starting point, we built an image processing and machine learning (ML) based analysis pipeline to classify individual presynaptic boutons at scale. We genetically labeled the presynaptic boutons of these interneuron populations using the same

fluorescent reporter, and visualized the distinct target compartments by antibody staining. Boutons can be classified using a supervised ML approach for hypothesis testing or an unsupervised ML approach for discovery. Adapting this pipeline to intact tissue allowed us to use the interneurons' targeting in their native context as a standard for direct, quantitative comparison with their targeting in a dissociated culture. Thus we were able to quantitatively assess interneuron targeting recapitulation within a spatially randomized environment. Spatial analysis of bouton identities from unsupervised classification provided insights into self organization and principles of interneuron axonal pathfinding.

1.1.4. Results

Supervised classification revealed similar subcellular targeting profiles for these interneuron populations *in vivo* and *in vitro*, indicating that interneuron subcellular targeting is primarily regulated by cell intrinsic programs. Unsupervised classification using the same input data not only identified the same three canonical targeting classes, but also revealed that these classes are comprised of multiple subpopulations. Bouton classes identified by both supervised and unsupervised classification displayed laminar and regional enrichment in intact tissue slices. In dissociated culture, two soma-targeting synaptic subpopulations mapped to target cells with different cellular profiles. The six dendrite-targeting synaptic subpopulations were found at increasing distances from target soma, suggesting molecularly distinct proximal, medial, and distal dendritic compartments in culture. Tracking subtype targeting across axonal branches of individual neurons indicated that SstCs and BCs utilize distinct targeting strategies in culture that accord with established findings *in vivo*. In sum, our synaptic analysis pipeline revealed novel synaptic subpopulations in interneurons. Further analysis uncovered novel aspects of interneuron synaptic biology that, remarkably, are retained in culture.

1.2. Preface

1.2.1. A connectivity-centric overview of neurodevelopment

Neurons display exquisite specificity in synaptic connectivity, forming highly stereotyped connections between pre- and postsynaptic cells. There are multiple classes of synaptic specificity, which are generally defined by the target. A given presynaptic cell can target:

- specific brain regions (regional specificity)[1-5],
- specific laminae within these regions (laminar specificity)[6-11],
- specific cells within these laminae (cellular specificity)[5, 6, 12-16], and even
- specific subcellular regions of these target cells (subcellular specificity)[17-22].

In other words, synaptic specificity is generally classified by the interaction between a presynaptic cell and its postsynaptic target, a two-body problem. This 'two-body' formulation, however, glosses over another important actor in the developmental process: the microenvironment. These stereotyped connections do not arise fully formed, but develop over time, unfolding over a series of spatially and temporally coordinated developmental stages within a dynamic microenvironment[15, 21-26].

Generally, neurodevelopment is split into the following stages (Figure 1.2):

1. Neuronal Birth
2. Migration
3. Axodendritic polarization [1]
4. Neurite elaboration
5. Synaptic target recognition
6. Synaptogenesis
7. Synapse maintenance/elimination

[1] We are using the term 'axodendritic polarization' to refer to the developmental process wherein a neuron specifies dendrites and axon as separate compartments. This step is frequently referred to as 'neuronal polarization' or simply 'polarization,' and generally occurs between migration and neurite elaboration. As will be discussed, however, different polarization and compartmentalization events occur in all the above-mentioned developmental stages. Directed migration, for example, requires cellular polarization. For this reason, it is clearer to refer to this stage as 'axodendritic polarization' and reserve the terms 'polarization' and 'compartmentalization' for general cellular and molecular processes wherein a chemical seeding event leads to localized molecular enrichment zones, and thus a polarized subcellular region/compartment.

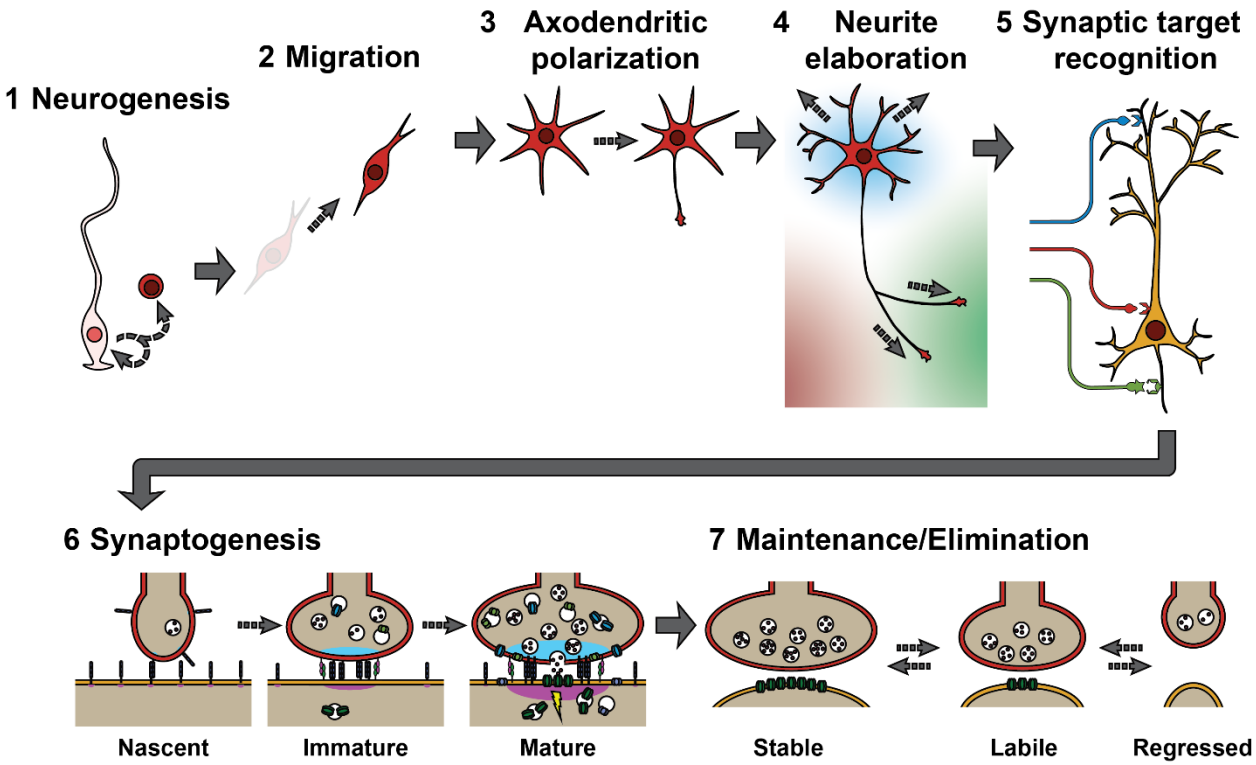

Figure 1.2 | Stages of neurodevelopment (standard order).
The neuronal development is typically partitioned into seven sequential stages: 1) neurogenesis, 2) migration, 3) axodendritic polarization, 4) neurite elaboration (which includes axonal and dendritic growth, branching, and guidance), 5) synaptic target recognition, 6) synaptogenesis, and 7) synapse maintenance and elimination (which includes activity-dependent synaptic plasticity and synaptic pruning).

Although these stages are presented in a specific order, it is important to recognize that 1) this is not necessarily the developmental trajectory for every neuronal connection, 2) these stages are dynamic, and often overlap, and 3) distinct neuron types rely more heavily on certain stages than others as they form stereotyped connections. Cortical pyramidal neurons, for instance, tend to follow the trajectory as presented, with a heavy bias on axon pathfinding for 'proper' regional connectivity. Martinotti interneurons, on the other hand, leave a trailing process as they migrate through their target region to their final settling location[27]. In other words, Martinotti cells couple

synaptic targeting, axodendritic polarization, neurite elaboration, and migration into a single developmental process. For this reason, from a connectivity-centric perspective, the best means of organizing neurodevelopment is not through these standard developmental stages, but rather by focusing on the three principal actors (presynaptic cell, target cell, and the environment), and how they interact to form this stereotyped connectivity. From this perspective, synaptic target recognition is a critical event in the development of neuronal connectivity. During this stage, neurons select specific synaptic partners out of all possible partners[2] (reviewed in [28-31]), a process that involves physical contact between putative partner cells. This physical interaction is significant: successful targeting establishes a prolonged conduit of transcellular communication, and therefore a central cell-cell interaction that did not exist previously. Since the operative cellular interactions prior to contact are so distinct from those where contact *is* established, it is conceptually helpful to distinguish stages prior to target selection from those after (Figure 1.3).[3]

Pre-contact, the presynaptic and postsynaptic cells develop independently along parallel paths (Figure 1.3, *left*). From a teleological standpoint, the problem can be reduced to 1) decreasing physical proximity between potential partners such that contact can occur within critical temporal windows, and 2) identifying the 'appropriate' partners for synapse formation. Thus, the problem at this stage is primarily spatiotemporal in nature: having the right molecules at the right place at the right time. This problem is solved by neuronal migration, neurite elaboration, and subcellular compartmentalization. Here, the primary interaction is between the cell and its microenvironment,

[2] Although partner selection is often presented as a presynaptic cell 'choosing' a postsynaptic partner, at a molecular level synaptic target recognition is a bidirectional process. Ascribing 'choice' to one partner or the other is therefore a narrative construct. We discuss current models for this process in greater detail below.

[3] If we were to provide a terminology to distinguish the stages before from those after contact, 'pre-synaptic' and 'post-synaptic' would be an obvious choice. Alas, they're taken. *Antesynaptic*, from Lt. *ante-*, 'before,' and *metasynaptic* from Gk. *meta-*, 'after' would work, but is also a bit confusing. *Antesynaptic* is a bit too close to antiseptic, and *meta-* is now more commonly understood as *above*. Funnily enough, *tango* is latin for 'contact,' so *presynaptango* and *postsynaptango* would work. A fun little dance, but also a bit of a mouthful. For these reasons, we will refer to these as 'pre-contact' and 'post-contact.'

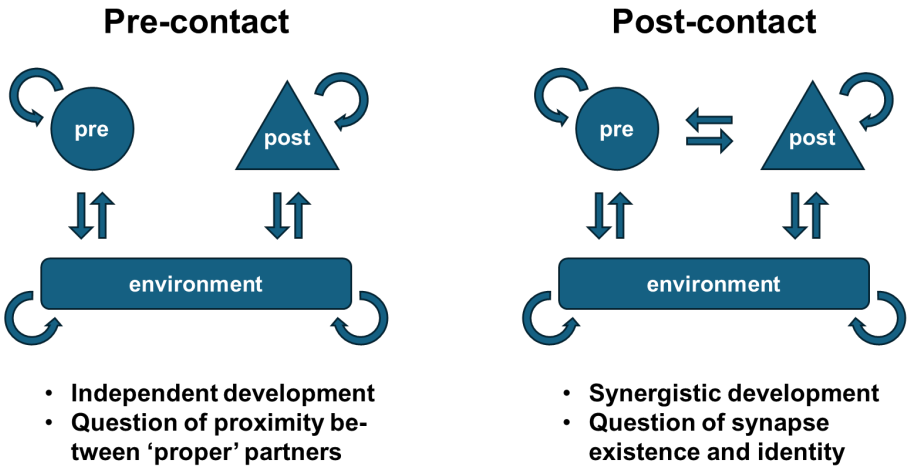

Figure 1.3 | Stages of neurodevelopment (connectivity-centric view).
The stages of neuronal development can be partitioned into stages prior to contact and synapse formation between the presynaptic and postsynaptic cells (*left*) and those after this connection has been established (*right*). During both stages, the presynaptic and postsynaptic cells interact with their respective microenvironments, and each of these actors undergoes intrinsic signaling. Pre-contact, the presynaptic and postsynaptic cells develop independently. From a teleological standpoint, the primary goal at this stage is to decrease proximity between 'proper' synaptic partners. Post-contact there is a sustained interaction between the presynaptic and postsynaptic cells. The cells can now interact causally and develop synergistically. Teleologically, the primary question at this stage is synapse existence and identity.

which functions as a framework, or molecular coordinate system, for these processes. This environmental framework can act permissively, as a passive bystander within which pre-established cell intrinsic programs simply unfold, or instructively, inducing the cell (or subcellular compartment) to take on a specific identity. Throughout this process, the neurons are sensing their environment to identify 'appropriate' synaptic partners out of all possible partners. These distinct interactions imply specific mechanisms for establishing and regulating specific synaptic connections, which we will discuss below.

Post-contact, the interactions change (Figure 1.3, *right*). In addition to the previously mentioned cell-environment interactions, there is also direct, transsynaptic signaling between presynaptic and postsynaptic cells. This change in interactions marks an important departure from the

prior state, introducing distinct mechanisms for regulating synaptic connectivity. Following contact, the problem can be reduced to an existential question for the synapse itself: formation, stabilization, growth, or elimination. The first step is identifying whether the contacted cell is an 'appropriate' synaptic partner. Positive identification initiates synaptogenesis, a process of transsynaptic communication and subcellular compartmentalization that will define the electrical, morphological, and molecular characteristics of the synapse itself. That is, post-contact the primary question is concerned with synapse existence and identity.

1.2.2. Organization of this introduction

The primary question we ask in our research is the extent to which a precisely patterned environmental framework is required for cortical interneuron subcellular connectivity. In order to cover this large and complex topic, we will first define the principle actors in our research by discussing the cerebral cortex and its development. This will allow us to introduce the cells of interest, their developmental trajectories, and their connectivity. It will also allow us to discuss the environmental framework within which these connections typically arise.

Since our research focuses on *subcellular* targeting specificity, subcellular compartmentalization is an important topic to expound upon. Moreover, a careful reading of the above discussion of synaptic connectivity will note that subcellular compartmentalization is a common feature of both the pre- and post-contact stages of neurodevelopment. Indeed, it is a common feature of cell biology. We will therefore discuss this subject in depth, specifically from the framework of the three acting parties: presynaptic cell, postsynaptic cell, and the cellular microenvironment.

Having covered all the relevant components, we will then discuss current models for how *specific* connections between them arise. We will primarily focus on subcellular specificity, but include other forms of synaptic targeting specificity when relevant. We will discuss these models

by focusing on the three acting parties. As noted above, the cellular and molecular mechanisms underlying subcellular targeting specificity can be partitioned by pre-contact and post-contact processes. First, we discuss pre-contact mechanisms, and principally axon guidance to specific subcellular targets. Next discuss we discuss current models of synaptic target recognition, synapse structure, and synaptogenesis. Finally, we discuss post-contact models of synaptic specificity. It is important to note that the pre- and post-contact mechanisms are not mutually exclusive in determining the final, stereotyped connectivity. The developmental trajectory of a neural connection can harness both classes of mechanisms to arrive at a specific, stereotyped connectivity. Different model systems often display a bias toward one or the other. We will therefore order the models based on the relative emphasis on pre- and post-contact mechanisms in determining subcellular targeting specificity.

We end with a discussion of relevant *ex vivo* experiments that help with background context for the experiments undertaken here. We tackle this from a conceptual standpoint, distinguishing the dissection and dissociation stages, and how they can be utilized to ask scientific questions.

Finally, we will provide a brief summary of what was discussed and a formalized hypothesis for our intervention.

1.1 Cortical Development

In the cerebral cortex, neurons are generally divided into two classes[32-40]: 1) glutamatergic excitatory neurons, which project their axons long distances and form synapses on target cells in other cortical regions, other parts of the brain, other parts of the central nervous system (CNS), and even outside of the CNS, and 2) GABAergic inhibitory neurons, which project locally and primarily form synapses on target cells within the same cortical region, often restricting their axonal trees to specific subregions, targets within a specific layer of the subregion, and even specific subcellular regions of the target cell.

Glutamatergic and GABA-ergic cells have distinct developmental origins and trajectories[23, 39, 41-43] (Figure 1.4). Glutamatergic neurons are born locally, within the cortical region that they will inhabit[40]. They originate either directly or indirectly from asymmetric division of radial glial cells in the ventricular zone at the base of the nascent cortex. They then climb the radial glial fiber past their earlier born cousins until they reach the cortical surface, where they settle[40, 43, 44]. Since different types of glutamatergic neurons are born in a well-defined sequence, this "inside-out" birth/migration pattern leads to the characteristic layering pattern of the cortex (reviewed in [23, 41]). Once settled, the excitatory neurons send out their axonal projections, synapsing onto distant targets (reviewed in [45]).

GABAergic cells have a very different origin. Whereas glutamatergic projection neurons are born locally within the cortex, cortical GABAergic neurons are born subcortically, in embryonic structures: the medial ganglionic eminence (MGE) or the caudal ganglionic eminence (CGE)[39, 41, 46, 47]. They then undergo a long, tangential migration into the cortex along two streams until they

reach the cortical region they will inhabit[46-48] (Figure 1.4). Once in the appropriate region. they migrate radially into the layer where they will settle, send out their axonal projections, and synapse

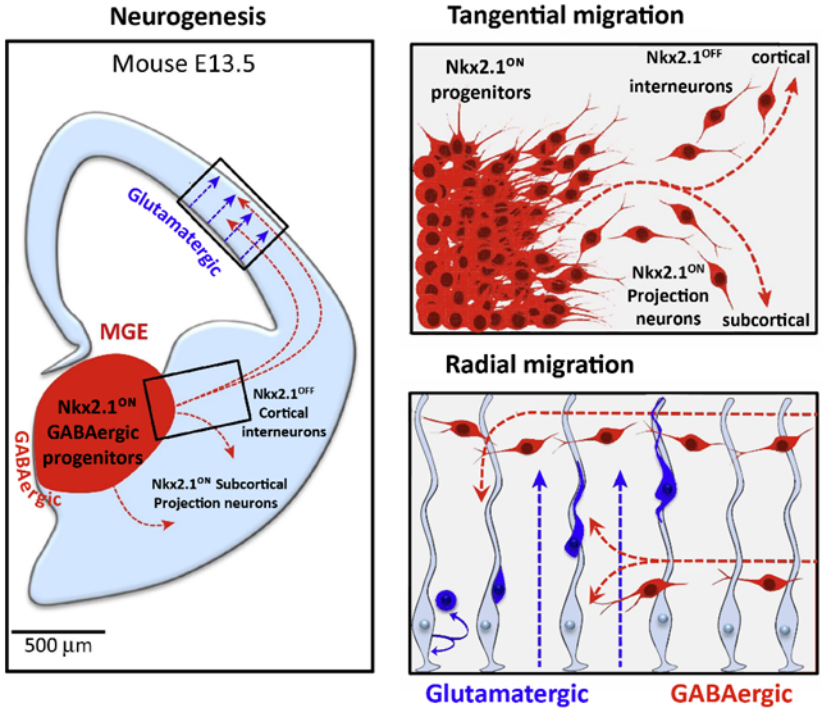

Figure 1.4 | Cortical development stages prior to circuit integration.
A) Neurogenesis occurs via asymmetric division of radial glial progenitors near the ventricle. Glutamatergic (blue) and GABAergic (red) neurons are born in spatially distinct regions of the embryonic telencephalon. Glutamatergic neurons are born within the cortex itself, whereas GABAergic neurons are born subcortically in the ganglionic eminences. Medial ganglionic eminence (MGE) derived cortical interneurons from Nkx2.1-expressing progenitors. **B)** MGE-derived neurons fated for subcortical regions maintain Nkx2.1 expression, whereas those fated for the cortex do not. The latter migrate tangentially into the cortex along two streams. **C)** Glutamatergic neurons (blue) migrate radially along the radial glial fiber to the surface of the cortex. Later born cells migrate past earlier born ones before settling, yielding an 'inside-out' laminar organization. After migrating tangentially into the cortical region they will inhabit, GABAergic neurons (red) migrate radially before settling into specific layers. Figure adapted from Tyson & Anderson, 2014.

onto appropriate target cells[47, 49]. These developmental trajectories occur at approximately the same time, with GABAergic interneuron entry into the cortex being slightly delayed[37, 40, 48, 50-55]. This delay is noteworthy, however, since it means that there is already a nascent structure and extended identity to the cortex when interneurons settle into place, send out projections, and synapse with their targets.

First, we discuss the diversity of cortical interneurons, including a brief summary of their molecular expression profiles and connectivity. Interneuron connectivity is covered in greater

depth later. Next, we examine cortical interneuron differentiation. We use this framework to define the context of our primary question: what role does a precisely patterned environment play in shaping cortical interneuron identity and connectivity? Given this framework, we next discuss the organization of the cortex. We begin by discussing its laminar organization, from both the cellular and synaptic standpoints. Next, we use glutamatergic neuron regional differentiation as a framework to discuss the complex and dynamic molecular environment into which interneurons migrate and form connections. We then use cortical regional connectivity to discuss how activity shapes regional identity. Finally, we examine the relationship between cortical laminar and regional organization, specifically focusing on what occurs when lamination is disrupted. We conclude with a brief summary of this section.

1.2.3. Cortical Interneuron development

1.2.3.1. Diversity of cortical interneurons

Cortical interneurons are an incredibly diverse population. Molecular expression data from single cell RNA sequencing (scRNAseq) experiments places estimates at roughly 60-120 distinct cell types depending on resolution, regions sampled, and dataset size[35, 56-59].[4] Despite this incredible diversity, these cell types can be organized hierarchically based on expression patterns, thus forming a cell type taxonomy that generally tracks with developmental trajectory. As noted above, the primary division in the cortex is between glutamatergic and GABAergic cells. The GABAergic population can be broken down into two distinct neighborhoods (or groups), which reflect their different developmental origin, namely the caudal and medial ganglionic eminences (CGE and

[4] For the sake of consistency, we use Yao et al., 2021 as a reference, which is the most the most comprehensive dataset to date and places the total number of cortical and hippocampal GABAergic interneurons at 119.

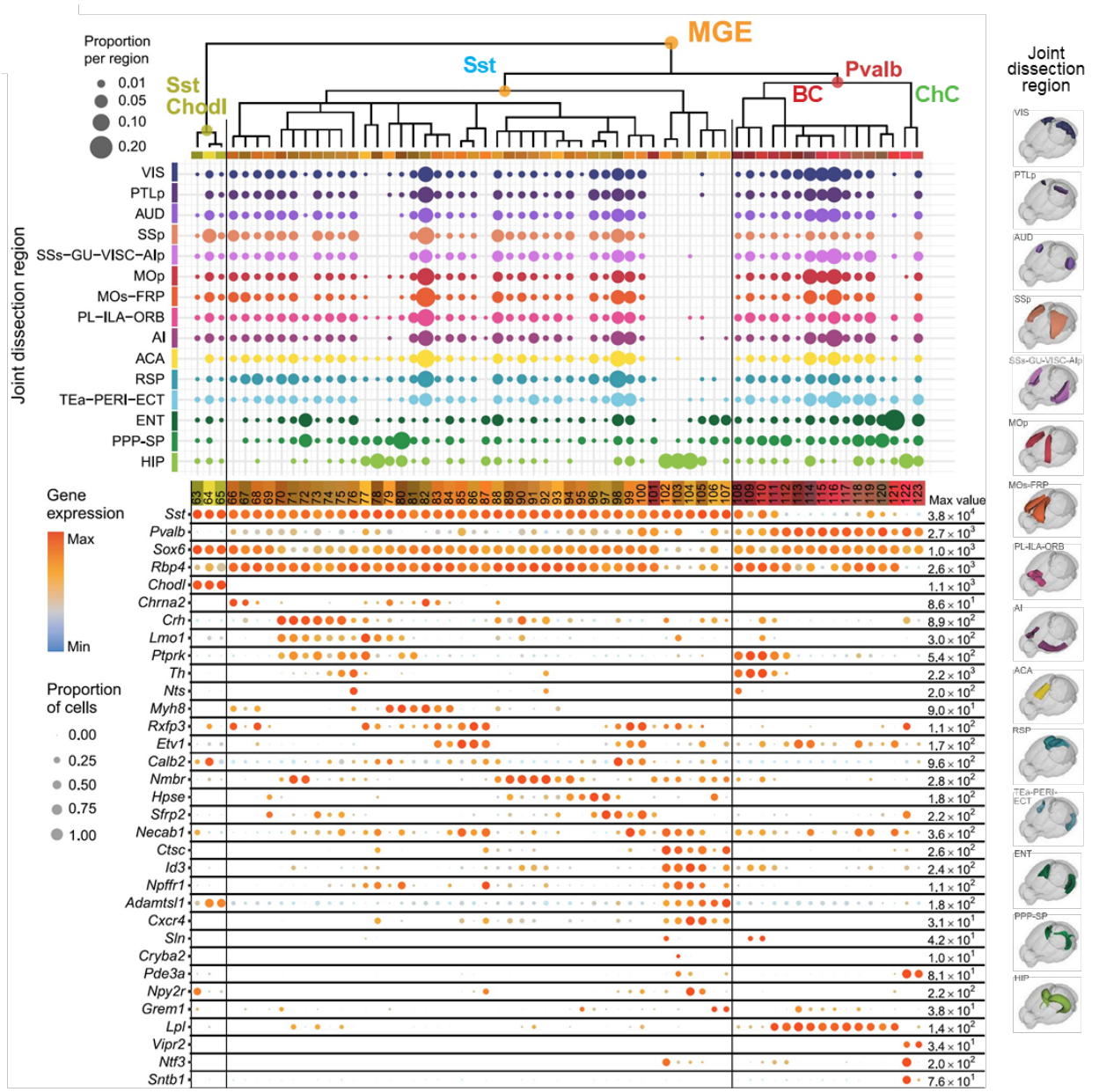

Figure 1.5 | The MGE lineage of cortical interneurons.
Top: A dendrogram of the MGE lineage of cortical interneurons based on scRNAseq expression profiles. *Middle*: Regional enrichment of specific subtypes. *Bottom*: gene expression levels of selected markers. Sst and PV (*Pvalb*) are at the top of the expression data. The BC and ChC lineages from the PV class are denoted within the dendrogram. Sst are largely dendrite-tareting interneurons and Sst-Chodl cells are GABAergic projection neurons. MGE-derived non-PV axo-axonic cells cluster with the CGE popula-tion and are not shown. *Right*: Overview of sampled brain regions rendered in Allen CCFv3. Region abbreviations can be found at https://doi.org/10.1016/j.cell.2021.04.021. Adapted from Yao et al., 2021.

MGE, respectively). The majority of cortical interneurons derive from the MGE (~70% of all in-

terneurons, 62 subtypes[5])[60]. Since this population is the focus of our work, we will only discuss this branch of GABAergic neurons (Figure 1.5). MGE-derived neurons can be partitioned into PV-expressing (~60%, 16 subtypes) and Sst-expressing (~40%, 45 subtypes) interneurons, though there is a small population (<10%, ~6 subtypes)[6] that expresses both[33, 61] or neither (Unknown %, 5 subtypes)[15, 31, 35, 37, 59, 60].

Subcellular and laminar targeting are often used as a classification method following this division because 1) it tracks with other classification methods such as marker expression, intrinsic physiology, and morphology[53], and 2) these well-studied characteristics are only now being mapped onto the more recently recognized taxonomic subtypes[30, 36, 62]. A summary of the major classes of PV and Sst neurons is shown in Figure 1.6.

The PV population accounts for the majority of MGE interneurons. All PV-expressing interneurons are fast-spiking cells, though there are subtle differences in physiological profiles between different PV subtypes[63]. The PV+ population is generally divided into three distinct cell classes based on targeting profile[15, 59, 61, 64]. The largest of these classes is basket cells (BCs), which primarily target soma and proximal dendrite targeting. BCs account for 70-80% of all MGE-derived interneurons (11 subtypes). The next best-studied group consists of chandelier cells (ChC), (10%, 1-2 subtypes). ChCs target the axon initial segment (AIS) of the postsynaptic cell, forming characteristic 'cartridges' that resemble candlesticks. These cells tend to inhabit L2/3, but can be found in deeper layers[62]. The remaining PV cells are translaminar (presumed dendrite-targeting) cells

[5] It is essential to note that these percentages can vary between different studies and may be influenced by factors such as brain region, age, and experimental techniques used for quantification. This applies not only to the percentage of MGE-derived neurons, but also to all that follow within this section.

[6] The percent of total MGE interneurons expressing both PV and Sst is from Lee et al., 2010. The number of subtypes that expressing both PV and Sst is based on data from the Sst-IRES-Cre;Pvalb-T2A-Dre mouse published in Yao et al., 2021. Additional data from this study show the Sst-IRES-Cre and Pvalb-T2A-Dre drivers are likely leaky. The total number of interneuron subtypes labeled by this mouse is 20, and includes subtypes not labeled by the Pvalb-IRES-Cre mouse alone. We therefore only included cell types that accounted for at least 5% all labeled cells from this mouse and were witnessed in both the Pvalb-IRES-Cre and Sst-IRES-Cre mice.

(10%, unknown subtypes). These cells likely express both PV and Sst[9, 37], but are generally poorly studied, in part due to technical difficulties with combinatorial labeling strategies[65].

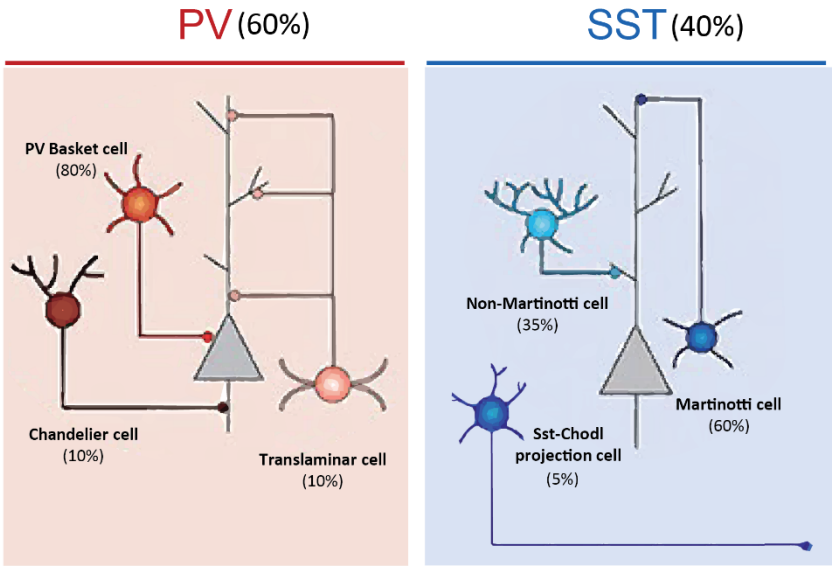

Figure 1.6 | Schematic summarizing the major MGE-derived cortical GABAergic interneurons based on their connectivity patterns. *Left*, the PV population is comprised of soma-targeting basket cells, AIS-targeting chandelier cells, and translaminar cells that (presumably) target dendrite. *Right*, the Sst population is primarily dendrite targeting. Martinotti cells (MCs) target pyramidal neuron distal dendrites in L1, whereas non-Martinotti cells (NMCs) do not. The Sst-Chodl population are projection neurons that target extracortical regions. These cells comprise a small subset of the total Sst population. Figure adapted from Llorca & Deogracias, 2022.

The Sst+ population comprises approximately 40% of MGE-derived cortical neurons. Despite being a minority population overall, it accounts for the majority of MGE interneuron diversity (45 subtypes). Notwithstanding this heterogeneity, all Sst cells types primarily target dendrite. Sst cells are generally divided into three distinct cell classes based on their targeting profile: Martinotti Cells (MCs), Non-Martinotti Cells (NMCs), and long-range projection cells (Sst-Chodl)[12, 31, 63, 66-68]. The best studied class is MCs (~60%, 11 subtypes), which reside in deeper layers and form translaminar connections onto pyramidal neuron distal dendritic tuft in L1[68]. Recent studies have delineated MCs into three distinct groups based on their morphology and settling position. These consist of L2/3 prototypic MCs (2 subtypes), L5 'fanning-out' MCs (≥4 subtypes), and L5 'T-shaped' MCs (5 subtypes)[5, 9, 65]. NMCs are a diverse set (~35%, 15 subtypes) of locally-projecting cortical interneurons ,that do not target L1[12, 31, 65]. These cells often inhabit L4 and L5 and can

display region-specific targeting patterns. Sst-Chodl cells (4.5%, 2 subtypes) are long-range projecting neurons that target extracortical regions[31, 33, 53, 61]. Sst-Chodl cells are not technically interneurons, but we include them here for completeness. The morphology and connectivity of the remaining Sst populations (17 subtypes) have not been mapped. Based on cell expression profile similarities, most will likely fit into the MC or NMC classes.

As noted above, there is a minority population of MGE-derived neurons that express neither PV nor Sst. At least one subtype within this population resembles chandelier cells. These cells targets AIS and form characteristic cartridges, but tend to inhabit deeper cortical layers and do not have a fast-spiking pattern[12, 35, 37, 53]. Since canonical chandelier cells are fast-spiking PV cells that inhabit L2/3, we refer to all AIS-targeting cells as 'axo-axonic,' and reserve the term 'chandelier cell' PV-positive population.

To summarize, MGE interneurons can be broadly divided into PV and Sst cells, with a small amount of overlap. PV consist of soma-targeting BCs, AIS-targeting ChCs, and (poorly-studied) translaminar cells that likely also express Sst. All cells within the Sst population preferentially target dendrite. They are distinguished as L1-targeting MCs, NMCs that do not target L1, and Sst-Chodl that project outside of the cortex[31, 69, 70]. Thus, despite the incredible diversity of MGE interneurons, there are fundamental and consistent organizing principles.

1.2.3.2. 1.1.1.2 Cortical interneuron specification and differentiation

The dominant model of cortical interneuron differentiation was proposed by Fishell and Kepecks, 2020[71]. They propose that interneuron differentiation proceeds through a three-stage process. We have added an additional stage (0) to define the ground state prior to specification. Here is a summary of these stages:

0. **Progenitor domain (unspecified):** Interneurons are born from cycling progenitors within the MGE and CGE. Although a small number of regionally-expressed genes can be detected within the proliferative zones at different times, cardinal class identities are generally not apparent at this stage. Each ganglionic eminence does, however, have a unique signature[72].
1. **Cardinal Specification:** Longitudinal genome analyses reveal that interneuron subtype identities emerge at a transcriptional level around the time interneurons become postmitotic. Four major cardinal classes are identified based on the expression of specific markers: PV, SST, VIP, and Reelin. Additionally, minor cardinal types are recognized, suggesting a complex diversity within interneurons. An interneuron's cardinal class can be predicted based on where and when it was generated. This stage limits the potential subtypes the cell can differentiate into based on a general class identity.
2. **Definitive Specification:** This stage involves the migration and settling of interneurons, where extrinsic cues shape subtype identity. Migration pathways are hypothesized to be generic (MZ or SVZ), while cues within layers or areal territories refine interneurons to specific local environments. Local cues encountered during migration and settling influence interneuron morphology, connectivity, and cell fate. Identity is progressively specified during subsequent developmental stages based on an interplay between intrinsic expression profile and extrinsic cues. Ever finer limits are placed on its fate, until subtype identity is irreversibly determined at the end of this stage.
3. **State Specification:** Interneurons adjust their function based on local circumstances. Transcriptomic analyses reveal discrete transcriptional states within interneuron subclasses. Some interneurons undergo lasting but reversible changes in transcriptional states based

on local circuit demands. Interneurons incorporate into circuits, adapting morphology and intrinsic properties according to local circuit demands. Extrinsic cues, possibly activity-coupled processes, influence synaptic properties and gene expression profiles, ensuring interneuron plasticity and functional integration within cortical networks.

This model is based on two extreme paradigms for conceptualizing how the cellular environment influences cellular identity.

- **Progenitor specification:** the environment simply potentiates an already intrinsically determined fate. Fate is specified by cell intrinsic programs.
- **Progressive specification:** the environment acts instructively on a more plastic identity to determine the cell's ultimate fate. Fate is specified by the environment.

To be reductionist, the above model places a dynamic bridge between these two extremes, leaving a bit of wiggle room at either end. By 'wiggle room' we just mean that the authors do not completely embrace either of these extreme positions. Cardinal classification *is* an intrinsically determined fate, but it is not *rigidly* determined. There is a fair amount of plasticity, as outlined by the Definitive Specification stage. But this plasticity operates within specific bounds: the ultimate fate is limited to a designated set of subtypes. Similarly, as outlined by the Definitive Specification stage, the authors propose that interneurons never acquire a rigid fate. Their ability to reversibly adjust to local circumstances proves that they are plastic. But again, this plasticity operates within bounds: once a specific subtype identity is assigned, it is unlikely to be abandoned. In other words, there is a gradual narrowing of potential states over the course of differentiation. This is the 'dynamic bridge,' namely, an attempt to incorporate Waddington's epigenetic landscape into a real-world model of cellular differentiation.

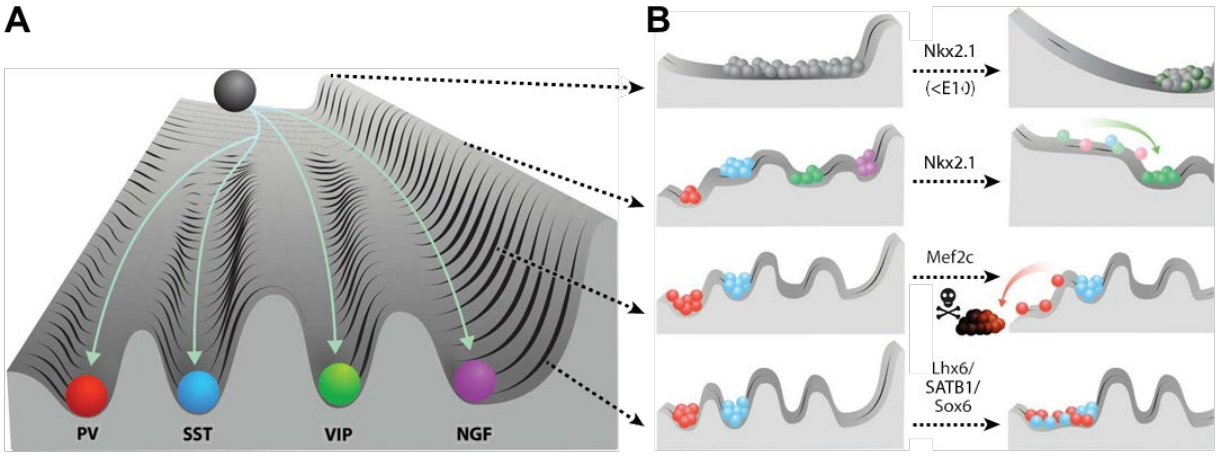

Figure 1.7 | The developmental landscape reflects attractor dynamics.
A) Diagram showing the landscape of development as the energy function of an attractor gene regulatory network, with cells rolling down through bifurcating valleys. At the bottom, the basins of attraction provide robustness to external perturbations and confer distinct stability properties, depending on the height of the energy barrier between interneuron subtypes. **B)** Schematic showing how distinct transcription factor manipulations generate distinct development landscapes, reducing barriers between attractor states and/or making them unstable. Abbreviations: NGF, neurogliaform; PV, parvalbumin; SST, somatostatin; VIP, vasoactive intestinal polypeptide. From Fishell & Kepecs, 2020.

The epigenetic landscape is a conceptual framework used to describe the development and differentiation of cells during embryogenesis[71, 72] (Figure 1.7, A). It visualizes the process of cellular differentiation as a metaphorical landscape, where cells are represented as marbles rolling down hills and valleys. The landscape consists of ridges, valleys, and basins, with each representing different developmental pathways or cell fates. The hills and valleys represent the stability of cellular states, with valleys being stable and representing differentiated cell types, while hills represent unstable states. The concept highlights how cells can transition between different states or developmental pathways, influenced by various factors such as genetic regulation, environmental cues, and cellular interactions. Waddington proposed that during development, cells undergo a process called canalization, where they become increasingly constrained towards specific devel-

opmental trajectories as they differentiate[55, 70]. The epigenetic landscape provides a visual metaphor for understanding how cells make decisions during development and how these decisions are influenced by both intrinsic and extrinsic factors.

Given this framework, the question becomes, 'what does the epigenetic landscape *actually* look like during this protracted developmental process?' That is, which interneuron features are determined at which developmental stage? Based on comparative analysis of cellular expression profiles across development, it very likely resembles the dendrogram presented in Figure 1.5[25, 35, 37, 56-58, 73, 74]. There are three major outstanding questions[7], which, admittedly, the interneuron field is working very hard to address:

- How does this dendrogram map onto the developmental timeline?
- How are these molecular expression profiles related to interneuron features?
- What shapes this landscape? (and by extension, the probability for an interneuron to adopt specific features?). There are two aspects to this question based on the intrinsic/extrinsic dichotomy:
    - How do intrinsically expressed molecules shape this landscape?
    - How does the cellular environment shape this landscape?

The question addressed in this study fits best within the third camp, but really touches on all three. Currently, our toolbox for addressing these questions consists of (among others) single cell gene expression profiles, genetic manipulations, and transplantation studies. Examples of insights into the cell intrinsic component of this question using genetic manipulations are shown in Figure 1.7, B. Our study is primarily interested in the role of the environment.

[7] This list is almost certainly incomplete.

Historically, transplantation has been the primary tool for addressing this question. Specifically, transplant experiments demonstrate the impact of local cues on interneuron migration and fate determination. The primary approach here is heterotopic transplantation, where cells or tissues are removed from their original location (donor site) and implanted into a different location (host site). We can therefore gain insights into what is known about this question from heterotopic transplantation studies. For cortical interneurons, this typically consists of transplanting from cortex to hippocampus, and vice versa. When this is done using a driver line that labels the entire MGE lineage, the proportionate composition of donor cell identities tended to match that of the host region[75]. Thus, the final environment appears to play a major role in determining interneuron identity. Of note, this conclusion was based on staining for PV, Sst, or nNOS of genetically labeled donor tissue in an unlabeled host. nNOS is a marker of the Sst Chodl class, so this approach assessed the proportionate composition of three cardinal classes within the MGE lineage. These findings were supported by electrophysiology profiles. Connectivity was not investigated, so we are blind to this feature. These data support a model wherein the cellular microenvironment is in part responsible for sculpting the final interneuron composition, diversity, maturation, and properties. But what does it tell us about interneuron differentiation? At face, these data appear to support that interneurons are very plastic: the cells intrinsically adjust their identity in response to the new environment. But we should proceed with caution. As noted by the authors, there are several mechanisms that could account for these findings. These data could be explained by a selective differentiation model, wherein multipotent progenitors give rise to specific cardinal classes, such that a generically-fated PV or Sst interneuron acquires its terminal fate through inductive environmental cues. According to this model, subtype specificity is determined after these cells

settle into their final location. Alternatively, these data could be due to selective survival of interneurons subtypes based on input connectivity[37, 76]: different host regions may be capable of supporting different loads of interneuron subtypes (discussed below). This possibility is made even more plausible by the observation that although the cortex and hippocampus display differences in GABAergic subtypes[37], they are not identifiable using these molecular markers. This alternate model implies that interneuron subtype identity is already determined by the time they reach the cortex (or hippocampus). The new environment does not change cellular identity, it simply selects different subpopulations for survival.

Heterotopic transplantation experiments of chandelier cells by Ishino and colleagues partly address this question[77]. Here, the authors used different genetic approaches to selectively label cortical (cChCs) and hippocampal (hChCs) chandelier cells within their progenitor domains. They then transplanted labeled donor MGE (for cChCs) into host hippocampus or labeled donor hippocampus (for hChCs) in host cortex. In both cases, the transplanted cells retain AIS specificity but adopt different morphological, physiological, and molecular expression profiles that match ChCs native to the host region. The authors conclude that local cellular environment shapes phenotypic properties, in accordance with a progressive specification model. Like the previously-described experiment, the hChC hippocampus-to-cortex experiments suggests that terminal features are determined by the terminal environment. The cChC MGE-to-hippocampus experiment provides much more insight. Here, the labeled cells are extracted from their progenitor region and transplanted into a host region that was never a migratory destination. This skips the intervening tangential migration step entirely, excluding any effects this intermediate environment may have on cellular differentiation. What can we conclude from this experiment? It sheds light on which fea-

tures are acquired during cardinal specification, which are impacted by the local cellular environment after settling, and the relative importance of the environment during tangential migration in acquiring these features. For chandelier cells, the preference to target AIS is determined at birth. It is a cardinal feature, and is not acquired during tangential migration. In contrast, features associated with synaptic functioning, like electrophysiological profile, are determined by the terminal environment. Similarly, morphology (beyond targeting preference) is not a cardinal feature: specific morphologies are adopted to suit the terminal environment, based on the spatial organization of its targets. These studies suggest that regional positioning of interneurons is an important factor in specifying certain aspects of identity, as the environment in part sculpts cellular identity through a process of progressive specification. It will be interesting to see whether this is a general principle of interneurons.

Although the cortex and hippocampus display region-specific GABAergic cell types[51], these cells are closely related. Both regions have evolved to support similar interneuron subtypes. These environments are populated by similar glutamatergic (pyramidal) cell types and have a similar structural organization. Moreover, interneurons tend to position themselves 'appropriately' near their synaptic targets[78, 79], even when the these targets are inappropriately positioned[73, 76] or the interneurons are transplanted[40, 43]. In other words, the local cues within the microenvironment of these different brain regions may not actually be so different. We may be witnessing subtle differences in cellular identity because there are only subtle differences between these environments. In the following subsections on cortical development we will discuss the cortex with an aim to further define this 'environment,' with a particular focus on its spatial organization.

1.2.4. Laminar distribution of cortical neuron types and their connections

Our primary question concerns the impact of the local environment on cortical interneuron connectivity. At its core, this cortical environment is the cellular and molecular framework within which cortical interneurons differentiate and form their connections. It is therefore important to discuss how the cortex is organized and how cortical interneurons are distributed within it. In general, the cortex can be broken into two distinct axes of organization: laminar (or layered) organization and regional (or areal) organization. We discuss both in turn.

1.2.4.1. 1.1.2.1 Corticogenesis and laminar identity

The level of spatiotemporal organization within the cortical environment is exemplified by the developmental trajectory of glutamatergic neurons during corticogenesis (reviewed in [22, 23, 80-85]) (Figure 1.8). Glutamatergic neurons are born through asymmetric division of radial glial progenitor cells (RGCs) in the ventricular zone (VZ) of the nascent cortex. These RGCs send a radially oriented glial fiber to the cortical surface, thereby connecting the basal and superficial surfaces of the nascent cortex through a cellular scaffold (Figure 1.8, A). Newly born daughter neurons utilize this glial fiber to migrate up to the marginal zone (MZ) near the cortical surface. Cajal-Retzius neurons (CRNs) are among the earliest born neurons generated during this process [85-87]. They migrate to the surface and settle in the MZ in response to signals from the meninges[43],[8] then secrete specific molecules that chemically polarize the cortical sheet along the superficial-to-deep axis. These chemical gradients serve as spatial cues for the ordered migration and compartmentalization of subsequent neurons[83, 88-94]. We discuss the role of reelin, an ECM protein secreted by CRNs, in cortical spatial patterning in greater depth below. Suffice it to say, these gradients serve

[8] Non-cortical CRNs also migrate into the cortex, settling in the MZ beside cortical CRNs.

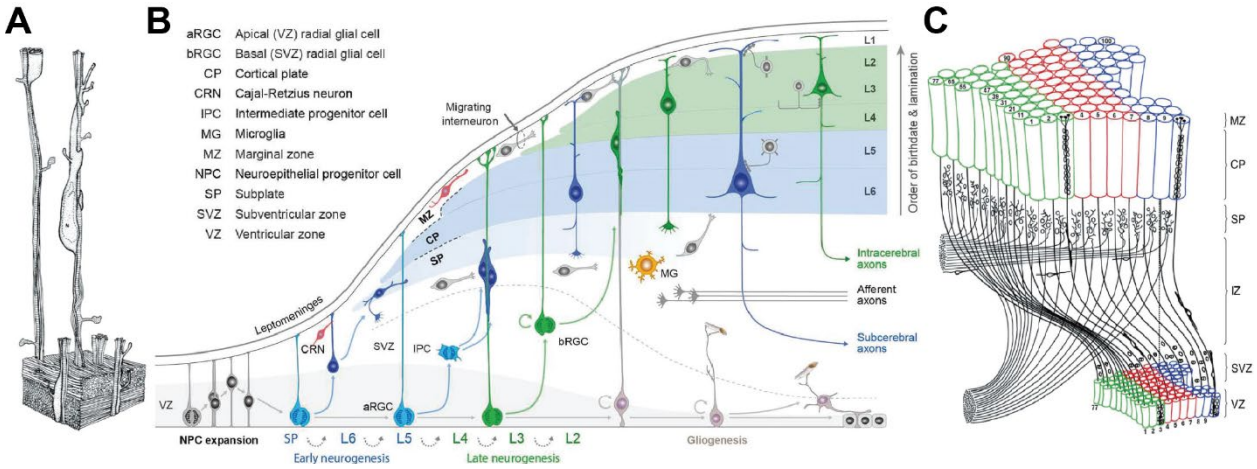

Figure 1.8| The evolution of corticogenesis: inside-out birth and migration patterns yield ontogenicaly-related cortical columns.

A) Three-dimensional reconstruction of postmitotic neurons migrating along radial glial fibers. Radial glial cells serve as both neural progenitors and scaffolds for migrating neurons during corticogenesis. Neuronal daughter cells migrate along radial glial fibers to reach their final positions in the cortex. As a result, neurons generated from the same radial glial cell tend to occupy the same radial column, contributing to the organization of cortical layers. **B)** Illustration of the dynamics of major developmental events and diversity of progenitors involved in the development of cerebral neocortex. Neurons are generated in the ventricular zone at the base of the cortex and migrate outward towards the pial surface, forming distinct layers in an inside-out sequence. The earliest-born neurons form the deepest layers of the cortex, while later-born neurons migrate past them to occupy more superficial layers. This sequential migration results in an organized cortical architecture, with layers arranged according to the birthdate of neurons. **C)** Representation of the radial unit hypothesis, which proposes that this process of birth at the ventricular zone (VZ) and radial migration toward the marginal zone (MZ) yields ontogenically-related columns of glutamatergic neurons. Adapted from Kolk & Rakic, 2022.

as critical cues for coordinating glutamatergic neuron migration. Newly-born neurons will migrate up the radial glial fiber past earlier-born neurons before they detach, and settle near the MZ. Since the identities of the neuronal daughter cells proceed in a largely predictable sequential order, this highly coordinated birth-and-migration process ultimately leads the inside-out organization of the cortex, where earlier born cells inhabit deeper layers and later born cells inhabit more superficial ones (Figure 1.8, B). These neurons, then, form a (largely) rigidly ordered column of cells, with neurons of distinct identities inhabiting distinct regions along the radial axis. According to the

radial unit hypothesis (Figure 1.8, C), since daughter neurons are most likely to utilize the glial fiber of their mother cell, all of the cells in any given cortical column tend to be ontogenically related[40, 43]. In addition, glutamatergic neurons within a column tend to form connections with other neurons in the column, forming a functional unit within the cortex. Thus, precisely-patterned spatial and temporal cues lead to a highly organized cortical environment. Tangentially migrating GABAergic interneurons enter into this environment as it is forming, taking up specific laminar positions after migrating radially into these locations. Thus, the cortical environment within which both GABAergic interneurons and glutamatergic projection neurons terminally differentiate and establish their connections is highly structured.

1.2.4.2. 1.1.2.2 Laminar distribution of cortical neuron types and their connections

A hallmark of cortical structure is its stratification into layers. The cortex is conventionally divided into six distinct layers (L1-L6) plus the white matter (WM). These layers are organized sequentially from superficial to deep. Due to its small size, L2 is often grouped with L3 as L2/3 in mouse. Thus, the standard order is: L1 > L2/3 > L4 > L5 > L6.

As a result of their 'inside-out' development pattern, glutamatergic neurons are rigidly stratified by cellular identity[22, 23, 33, 80-83, 95]. Thus, cortical layering is generally based on the stereotyped organization of these cells rather than their GABAergic counterparts. Historically, two methods have been used to define these layers: resident cellular identity and stereotyped connectivity. Regarding the first method, early studies based laminar distinctions on differences in soma densities and morphologies (Figure 1.9, left panel)[22]. This general schema has subsequently been supported by numerous studies showing distinct molecular expression patterns across these laminae[33, 37, 95, 96]. Regarding connectivity, each of the different cortical layers displays distinct, stereotyped connectivity patterns (Figure 1.9, right panel)[6, 97]. In addition to the long-range afferent and efferent

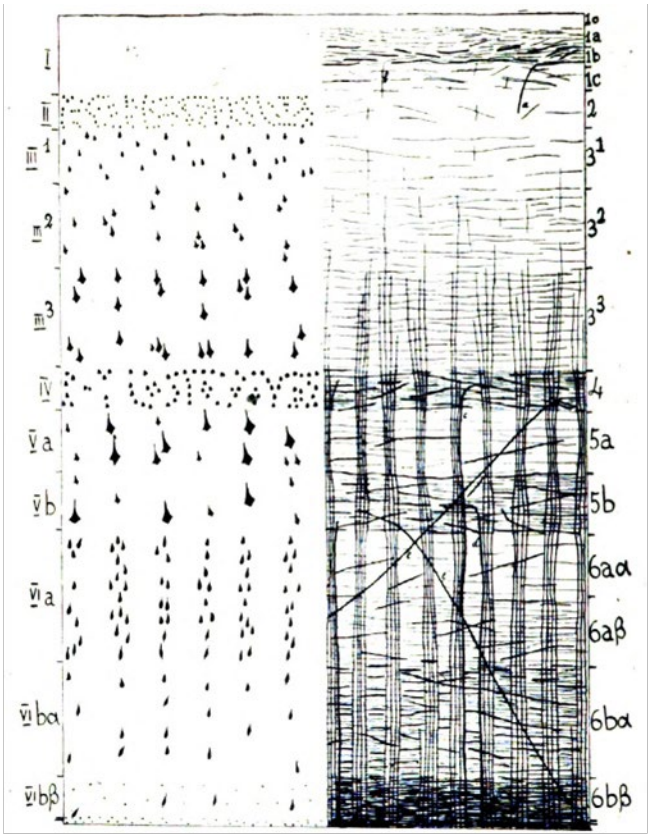

Figure 1.9 | Lamination of the cerebral cortex. The cortex is divided into six primary layers based on cellular identity (*left*) and connectivity (*right*) patterns. There is evidence of sublamination in almost every cortical layer, as demonstrated by the layer designations. These patterns ultimately result from the inside-out birth and migration pattern of excitatory glutamatergic neurons, which comprise the majority of cortical cells. Figure from Vogt & Vogt, 1919. *Reproduced with permission from Springer Nature*.

projections noted below, the glutamatergic cells within each of these layers also form stereotyped local translaminar connections, which result in intracolumnar microcircuits[60, 98-100]. As with cellular identity, the connectivity patterns tend to be rigidly stratified based on the stereotyped positioning and morphologies of glutamatergic neurons. Neurons have complex arbors that often extend beyond their layer of origin. As discussed below, since these other layers comprise distinct molecular environments, this often results in these cells having molecularly distinct subcellular compartments at different layers. Thus, the presynaptic and postsynaptic cellular identities of synapses located at a specific layer do not necessarily correlate with the most prominent soma at that layer: a presynapse in L1 does not necessarily target or derive from an L1 resident cell. Similarly, a neuron in L5A does not necessarily target other cells (or even target compartments) in L5A. Thus, laminar connectivity patterns and laminar cellular identity should be viewed as orthogonal methods of defining the cortical landscape.

Here is a simplified summary of the cortical layers:

- **L1** is the most superficial layer of the cortex. It is referred to as an 'acellular' layer due to the relative sparseness of resident cell bodies. It is comprised almost entirely of a complex network of neuronal connectivity. These connections are between apical dendritic tufts from pyramidal cells in L2-6 and both glutamatergic excitatory and GABAergic inhibitory axons[6, 12, 22, 35, 100-102].. The few resident cells in L1 are CGE-derived GABAergic inter-neurons[102, 103].
- ,**L2/3** contains pyramidal neurons (PyNs) that project to intratelencephalic sites, mediating intracortical and cortico-striatal communication streams[104, 105].
- **L4** is generally comprised of non-pyramidal glutamatergic neurons do not have dendrites that project into L1. These cells receive afferent connections from the thalamus and project intracortically. Cellular identity of L4 is known to differ between cortical areas[22, 35]. L4 is absent from the prefrontal cortex[95].
- **L5** is a pyramidal layer with cells that generally have large cell bodies and project axons extracortically. L5 is routinely divided into L5A and L5B based on differences in cellular expression patterns and connectivity[22, 106, 107].
- **L6** is highly diverse, with distinct pyramidal neurons defined by their dendritic arbors, inverted pyramidal cells, and oddly-shaped non-pyramidal cells[22, 106]. L6 is sometimes subdivided into L6A and L6B[108]. The excitatory neurons in L6 form both intra- and extracortical projections. It is generally associated with corticothalamic and colossal projections[109]. In addition to the long-range projections noted here, the glutamatergic cells within each of these layers also form local translaminar connections within the same cortical column.

- **WM** is the deepest layer of the cortex, and is sometimes referred to as subcortical. It is largely acellular, consisting primarily of myelinated axons and sparse synaptic connections[101, 108]. These axons are either afferents from extracortical regions of the CNS, pyramidal cell intracortical axons between cortical regions, or extracortically-targeting efferents[35]. The few resident cells in WM tend to be highly specialized glutamatergic and GABAergic neurons[22, 101, 109-111].

It is important to note, however, that this L1-L6 laminar nomenclature represents the top level of a hierarchically organized distribution of glutamatergic cell bodies in the cortex (discussed in [95, 100]). Taking into account subtle differences in cellular identity and neuronal connectivity has led to the identification of sublaminar divisions in almost all cortical layers[100, 112]. Indeed, such sublaminar divisions were recognized over a century ago, as illustrated in Figure 1.9. This conclusion is supported by both methodologies mentioned above. For example, a wide body of literature makes a strong argument for sublamination of L1 based on differential molecular expression[5, 95, 100, 113-115] and connectivity[33, 53] patterns.

Although the glutamatergic cells of the cortex are rigidly stratified, GABAergic inhibitory neurons are not. Instead, they tend to be interspersed in a salt and pepper fashion, with different interneuron types enriched in predictable manner across different layers[35, 37, 53, 77, 116, 117]. In general, interneurons tend to position themselves in an inside-out manner like their glutamatergic counterparts: early-born cell types in deeper layers and later-born ones in more superficial layers. A minority population of very-late-born interneurons breaks this general rule, tending to be positioned in deeper layers[52]. Interneuron laminar positioning occurs after the corresponding glutamatergic cells have settled into their respective laminar position and begun to differentiate[52, 54, 117]. Remarkably, interneurons tend to position themselves near their synaptic partners[54, 78, 118]. This is

true not only in the wild-type cortex, but also in the 'inverted' reeler cortex, where interneurons will also be positioned in a generally inverted manner[22, 37, 40], suggesting signaling from target cells drives interneuron positioning. Given the difficulty of separating interneurons from their targets *in vivo*, it is unclear whether this proximal positioning is important in shaping interneuron identity and connectivity, as discussed below.

Overall, this discussion of cortical laminar identity exemplifies the highly patterned molecular framework of the cortex in the radial dimension. In addition, these studies suggest a more refined scheme will likely have to be developed to take into account these more subtle sublaminar differences in cellular and synaptic identity[37]. As discussed above, the laminar connectivity patterns of glutamatergic neurons are well established. The spatial distribution of cortical interneuron synapses and their identities, however, is not. Given the sparse cellular distribution of cortical interneurons, a fine-mapping of interneuron synaptic identity within this laminar environment is of great interest.

1.2.5. Regional distribution of cortical neuron types and their connections

Based on the observation that distinct cortical regions are associated with distinct sensory, motor, and behavioral functions (Figure 1.10)[119], one might assume that distinct functional regions have similarly distinct cellular compositions. In other words, that regional differences in function arise from regional differences in cell types. Large-scale single cell molecular expression profiling of adult mice tells a different story, however[120]. It is true that there are a few examples of highly region-specific glutamatergic cell types, particularly in the retrosplenial and anterior cinglulate areas. But this is more of an exception than a rule. Most glutamatergic cell types are shared across multiple cortical regions, following a continuous distribution in the anterior-posterior axis

and a more graded distribution in the medial-lateral axis[121]. Remarkably, GABAergic interneurons show even fewer regional differences (Figure 1.5). Interneuron cell types tend to be evenly distributed throughout all cortical areas. Although there are a few cases of regional enrichment, there are no region-specific GABAergic cell types[9]. As discussed below, however, the relatively minor differences in expression patterns in adult mice belie the remarkably complex and dynamic regional diversity during development[122]. Thus, although cell type distribution may be relatively similar across adult cortical regions, minor and transient differences in expression patterns are responsible for the regional connectivity patterns.

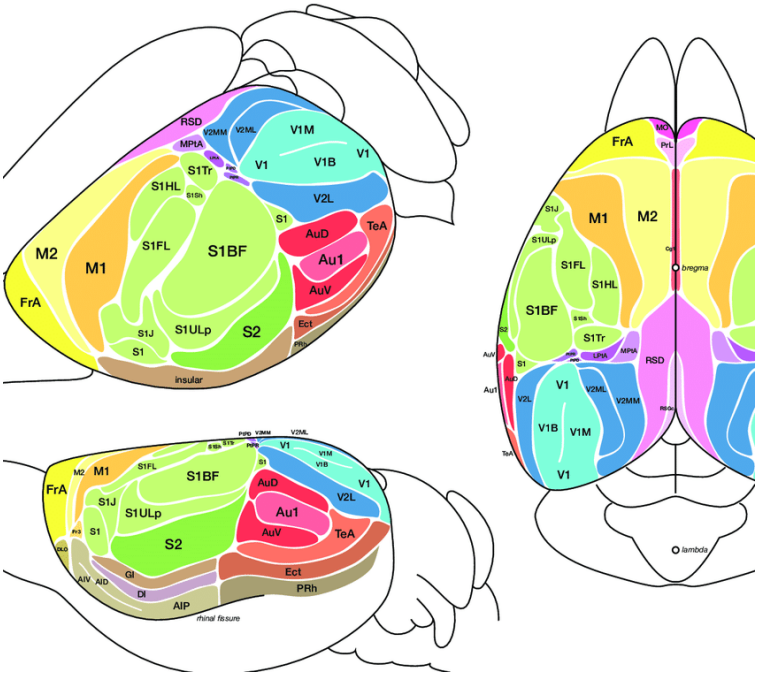

Figure 1.10 | Regional organization of the mouse cortex. Motor regions are in yellow and orange, somatosensory regions are in green, auditory in red, and visual regions in blue. Figure from Kirkcaldie, The Mouse Nervous System, 2012.

1.2.5.1. Areal identity in the cortex and the establishment of a molecular coordinate system

There are two hypotheses for how these regional difference arise (reviewed in [123]). The 'Protomap' hypothesis postulates that progenitors are spatially pre-patterned, thereby spawning region-

[9] The few instances of region-specific GABAergic interneuron types witnessed by Yao et al. are accounted for by differences between hippocampus and isocortex, where cell types present in the hippocampus were not present in the isocortex. The authors did not identify any region-specific interneuron cell types when comparing between cortical regions, and all interneuron cell types present in the cortex were also found in the hippocampus.

specific neurons with distinct molecular, physiological, and connectivity profiles. In contrast, the 'Protocortex' hypothesis postulates that the cortex is initially a blank slate of cellular identity that develops area-specific features in response to extrinsic signals, with a primary emphasis on thalamic afferents. Reality is more nuanced, with data indicating that elements of both are at play. To simplify, the current model proposes that extrinsic factors pattern a 'protomap' onto an otherwise homogenous cortical sheet, which is then further refined by thalamic afferents. To understand this process, it is important to recognize that the properties that define regional identity emerge gradually over development, with different region-specific features emerging at different developmental stages. Early stages of cortical development function largely according to a 'protocortex' model, whereby extrinsic factors (mainly soluble patterning factors) induce an early 'protomap' onto otherwise multipotent cells (Figure 1.11, A). Different morphogens gradients arise from different extracortical patterning centers[123-125]. These morphogens specify broad-brush regional identity by inducing graded expression patterns of transcription factors that drive regional identity programs[121, 123, 126, 127] (Figure 1.11, B). This induced regional identity in turn leads to the graded expression patterns across different regions. When expressed extracellularly, the graded pattern of cellular identity is mirrored within the extracellular space. In accordance with the chemoaffinity hypothesis[126], this yields a molecular coordinate system within the environment. As discussed below, this molecular framework is utilized for migration, axon guidance, and synaptic targeting[128, 129]. But it also reenforces and progresses differences in regional identity. Specifically, transcription factors and morphogens have evolved to reciprocally induce and repress each other. Thus, the extrinsically patterned cortical cells themselves express morphogens, yielding cortex-intrinsic autocrine and paracrine signaling that furthers patterning[122, 130, 131]. These positive and negative feedback and feedforward loops maintain and evolve regional identity. It is important to note that

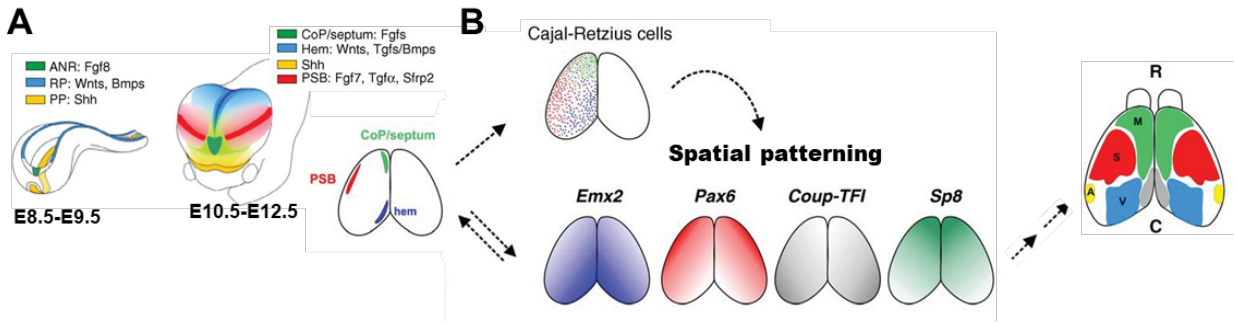

Figure 1.11 | Spatial patterning of the cortex.
The initial regional differences in cortical identity emerge from patterning factors (morphogens) expressed at extracortical organizing centers (reviewed by O'Leary et al., 2007). **A)** Early in development (E8.5-9.5), these morphogens includ BMPs and Wnts from the posteriomedial cortical the hem, FGFs from the anterior neural ridge, and Shh from a ventral domain. At E10.5-E12.5, FGFs continue to be expressed anteriorly in nested domains at the commissural plate (CoP) and septum (green). **B)** These morphogen gradients induce specific transcription factor programs in the cortical glutamatergic neuronal precursors, thereby leading to the distinct, graded expression patterns of identity-associated transcription factors along the anterior-posterior and lateral-medial axes within the developing cortex. These graded expression patterns in turn lead to graded expression of molecules that underly regional connectivity patterns, including intracortical, afferent and efferent targeting. The afferent and efferent connectivity in turn defines the information inputs and outputs, and so also any functional association with a region. Of note, region-specific gene expression in the cortex generally develops normally in the absence of thalamic afferents (Miyashita-Lin et al., 1999), though there is a loss in the sharp borders between regions (Vue et al., 2013). The same is true of the olfactory sensory map in the olfactory bulb, which generally develops normally in the absence of afferent glutamatergic projection neurons or GABAergic interneurons, suggesting this principle is more general than just the cortex. Thus, early embryonic patterning factors define the regional differences in the gene expression landscape prior to thalamocortical innervation, thereby establishing regional connectivity patterns and functional associations. Diagram adapted from Borello and Pierani, 2010.

the expression axes that account for the regional framework operate largely in the anterior-posterior (A-P) and lateral-medial (L-M) axes. In other words, they overlap with but are largely independent from those that account for cortical lamination along the superficial-to-deep axis[122, 132] (discussed in greater detail below).

The impact of morphogen signaling on cortical patterning is clearly shown by studies on FGF8. This morphogen is normally expressed from a region in the anterior telencephalon and is

in part responsible for patterning the Emx2 and Coup-TF1 transcription factor gradients by repressing their expression[121, 124, 132, 133]. Investigating the role of FGF8 in patterning of the somatosensory cortex, Fukuchi-Shimogori & Grove showed that increasing the FGF8 signal shifts the area boundaries posteriorly, whereas decreasing it shifts it anteriorly. Remarkably, ectopic FGF8 expression posterior to the original source leads to partial area duplication of the somatosensory cortex[134]. This experiment showcases the power of patterning centers and morphogen gradients to specify tissue identity.

Given the self-reenforcing signaling cascades and extreme impact of morphogens, one might expect that once regional identity is specified, identity is set. This is not the case. Rather, the 'protomap' is quite plastic. This is demonstrated by now classic surgical experiments. Transplanting donor visual cortex into host somatosensory cortex yields tissue with all the hallmarks of native somatosensory cortex: barrel formation and thalamocortical connections expected for barrel cortex[135]. That is, the tissue loses its original identity and takes on the identity of the host region. Similarly, surgically removing an entire cortical region before thalamic innervation will lead to a proportionally smaller, but complete and properly oriented thalamic innervation pattern of the remaining cortex[123].

Here we describe how morphogens and transcription factors induce regional identities, leading to graded expression patterns across different cortical areas. These experiments showcase how cortical identity is initially specified, maintained, and restabilized. A critical aspect of this system is the extracellularly expressed molecules, which yield a patterned extracellular environment. Previously we focused on how interactions between morphogens and transcription factors determine regional identity. Now we will discuss how these graded expression patterns relate to areal connectivity.

1.2.5.2. Topographic maps and activity-dependent refinement of regional identity

In the last section we described a molecular coordinate system within the extracellular environment that parallels the graded pattern of cellular identity. As discussed in greater detail below, these axes of molecular expression underly regional targeting specificity. The graded identity patterns result in graded expression of axon guidance and synaptic adhesion molecules that underlie areal targeting specificty[1, 136]. Once these regional afferent connections are established, they become the primary extrinsic factor shaping regional identity. This occurs through a process of activity-associated cell survival and circuit refinement through anterograde transsynaptic signaling. We will discuss the mechanisms of synaptic specificity in great depth below. Here we focus on how these connections shape regional identity.

In neuroscience, a topographic map refers to a spatial representation of sensory information or neural activity within a specific region of the brain. These maps reflect the organization of sensory inputs or neural responses according to their location in the external environment or on the body. For example, in the visual system, topographic maps might represent the visual field, with neighboring neurons processing information from adjacent areas of space. Topographic maps are found throughout the nervous system, including the visual[101], auditory, olfactory[101], and somatosensory systems[101]. The best characterized topographic map is the mouse somatosensory system, and specifically the barrel cortex associated with whiskers on the snout (Figure 1.12, reviewed in [130]). Somatosensory signals from whisker-associated mechanosensors in the skin transmit sensory information indirectly from these peripheral sensors to the thalamus, which in turn forms excitatory, glutamatergic afferents onto the somatosensory cortex. These afferents are organized into spatially discrete, anatomic units called 'barrels,' with each barrel representing an individual whisker, and the barrel field as a whole forming an anatomical map that largely matches

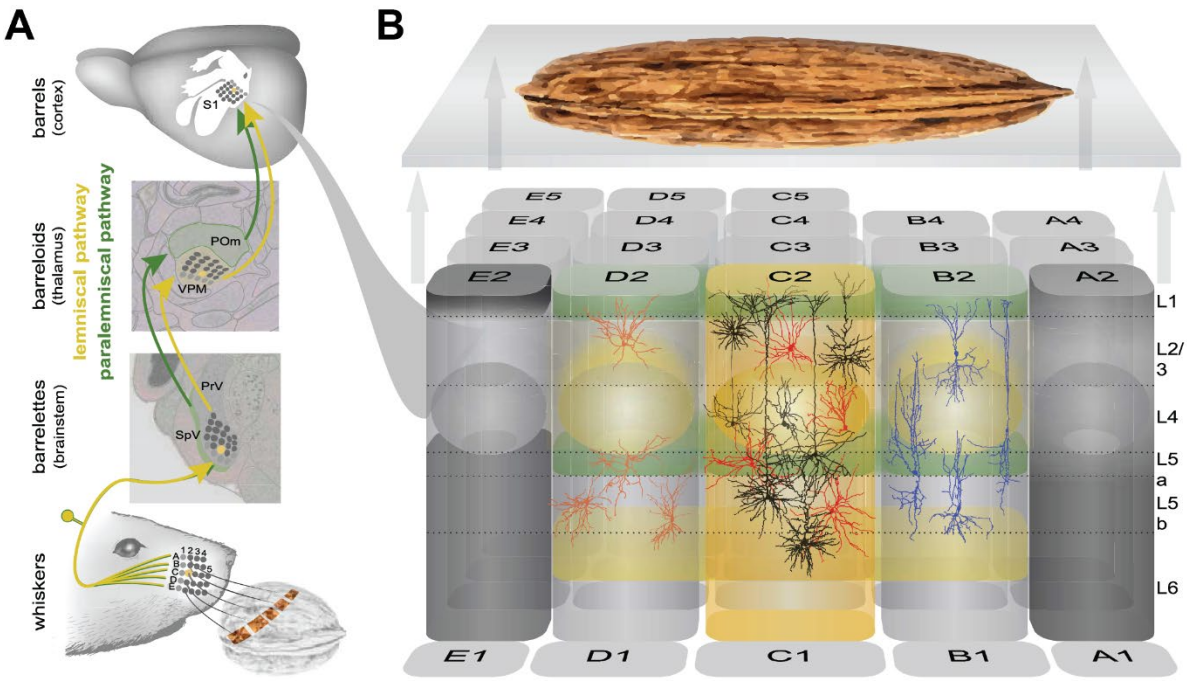

Figure 1.12 | Examples of topographic maps: the pathways transmitting tactile information from the muzzle to the cortex.
A) neurons from whisker follicles project to the brain stem, then branch into the lemniscal pathway (yellow) that innervates the PrV, and the paralemniscal pathway (green) that innervates the SpV. The lemniscal pathway displays a precise somatotopic organization at every stage, whereas the paralemniscal pathway lacks this detail. The lemniscal pathway relays through the VPM to reach the wS1 barrel cortex. The paralemniscal pathway relays through the POm to layer 4 of wS2. A higher order region of POm receives cortical input and projects to L1 and L5a of wS1 and wS2 of the cortex. **B)** A pseudo-3D visual representation of the barrel cortex (wS1), focusing on arc 2 and highlighting column C2 in yellow. Yellow shading indicates the termination layers of the lemniscal pathway, while green shading represents those of the higher-order pathway from POm. Gray shading on the left (E2) shows the density of inputs from the whisker motor cortex, while on the right (A2) it approximates the density of cholinergic fibers from the basal forebrain. The morphologies of different claases of cortical neurons involved in processing this information are depicted in specific columns. Column D2 shows Sst interneurons in orange. Column C2 shows PV interneurons in red and pyamidal neurons in black. Column B2 shows VIP-expressing interneurons in blue. Abbreviations: PrV, principal trigeminal nucleus; SpV, spinal trigeminal nucleus; VPM, ventral posteromedial thalamic nucleus; POm, posterior thalamic nucleus; wS1, primary whisker somatosensory barrel cortex; wS2 secondary somatosensory cortex. From Staiger, 2021. Reproduced with permission from the American Physiological Society.

the spatial organization of the whiskers themselves. This anatomical map generally matches a functional map, in which every barrel preferentially (but not exclusively) processes the sensory

information detected by its associated whisker[131]. The functional connectivity of this circuit is very easily mapped via whisker clipping experiments, where a mouse is allowed to explore a novel environment after clipping specific whiskers. Activated neurons in the cortex are then identified by c-fos staining, with the clipped whiskers leading to absence of c-fos in unactivated barrels[12, 101].

There are two primary thalamocortical afferent pathways into the barrel cortex: the lemniscal pathway (yellow) from the VPM targets L4 and L6, and the paraleminiscal pathway (green) from the pOM, which targets L1 and L5A. The lemniscal pathway is the best studied, and specifically the afferents from VPN onto L4 spiny stellate neurons. Of note, cortical interneurons are also involved in these barrels. L4 PV+ basket cells (PVBCs) and L4/L5B Sst+ non-martinotti cells (NMCs) form synapses within barrels, whereas L5A PVBC and L5A NMC[137] axons are confined to regions outside of the barrel[138]. These interneuron connectivity patterns and their ultimate spatial organization, in other words, are influenced by glutamatergic afferents from the lemniscal pathway.

The ultimate cellular composition of this brain region is shaped by activity-dependent cell death[74]. During development, neurons undergo a period of overproduction, generating more cells than ultimately needed. In the cortex, both GABAergic interneurons and glutamatergic pyramidal cells undergo extensive cell death during a postnatal critical period[75]. These neurons are selectively eliminated from the nervous system in response to activity levels. During a critical developmental window, neurons that receive insufficient synaptic input or fail to establish proper connections will undergo apoptosis. Conversely, neurons that are highly active and successfully integrate into functional circuits are more likely to survive. For cortical interneurons in particular, pyramidal cell input onto individual interneurons predicts their survival[137]. The level of neuronal activity

defines a lode for cell survival: neuronal survival within any given region is matched precisely to the level of activity within the region during this critical period. In other words, activity-dependent cell death plays a crucial role in tuning the cellular composition of brain regions to match activity requirements[101]. This leads to activity-dependent differences in cellular composition of different regions, and shapes neuronal circuits. By selectively eliminating neurons that are not effectively integrated into functional networks, this process contributes to the establishment of precise and efficient neuronal circuits essential for proper brain function and behavior[139].

In the barrel cortex, this has been shown experimentally by denervation or continuous trimming of whiskers during development[140], or more directly via genetic manipulations that disrupt activity[141-143]. Under these conditions, the spatial distribution of the thalamocortical afferents (TCAs) is generally preserved, but borders are poorly defined, and barrels show considerable overlap[144]. Anterograde (presynaptic to postsynaptic) signaling is responsible for the well-defined borders of barrels[144-146]. Specifically, this is an activity-dependent process: blocking neurotransmitter release or synaptic transmission leads to a loss of the sharp borders between barrels[146, 147]. Blocking activity specifically in the thalamus resulted in altered cortical lamination, specifically showing a contraction of L4 and expansion of L5[148], likely due to (in)activity-dependent cell death of L4 neurons, the primary target of the lemniscal pathway. Afferent connections and activity are important in shaping areal identity outside of the barrel cortex. For example, rerouting retinal inputs along auditory pathways leads to the development of orientation modules and long-range horizontal connectivity patterns that are reminiscent of the primary visual cortex[149]. Similarly, alternately increasing or decreasing the size of the dorsal lateral geniculate nucleus (dLGN), the major thalamic nucleus involved in the visual pathway, results in a corresponding change in the size of the primary visual cortex[150].

In this passage, we discussed how molecular expression patterns in the extracellular environment contribute to regional targeting specificity in the brain and, ultimately, regional identity. We discussed topographic maps, such as the barrel cortex in the somatosensory system, which represent spatial organization of sensory inputs, like whiskers in the case of the barrel cortex. The formation of these maps involves intricate processes including axon guidance, synaptic adhesion molecules, and activity-associated cell survival. The cellular composition of brain regions is shaped by activity-dependent cell death during development: neurons that fail to integrate do not survive.

1.2.5.3. Independence of Regional and Laminar identity in the cortex

Above we discussed cortical identity along two axes: regional identity across the cortical sheet along the A-P and M-L axes, and laminar identity along the superficial-to-deep (S-D) axis. As discussed above, topographic maps exist primarily along the A-P and M-L axes, with distinct inputs and outputs into the cortical column existing radially along the S-D axis. How important is the spatial organization of the cortex to neuronal identity?

These questions were tested using the *reeler* mouse, which have a spontaneous mutation in the *reln* gene required for proper glutamatergic 'inside-out' radial migration after birth. This in turn disrupts normal lamination, leading to region-dependent glutamatergic cell randomization and/or inversion, as shown by the staining for laminar fate markers (Figure 1.13)[89, 130, 131, 149, 150]. Within this context, glutamatergic cell identity is largely retained, with cellular identity markers having similar proportions in both the reeler and WT contexts. Thus, based on these markers, normal cortical organization in the radial dimension is not a prerequisite for normal glutamatergic cell differentiation. Of note, this contrasts sharply with the findings on cortical regional identity

discussed above, where neuronal fate is plastic and determined by relative positioning on the cortical sheet.

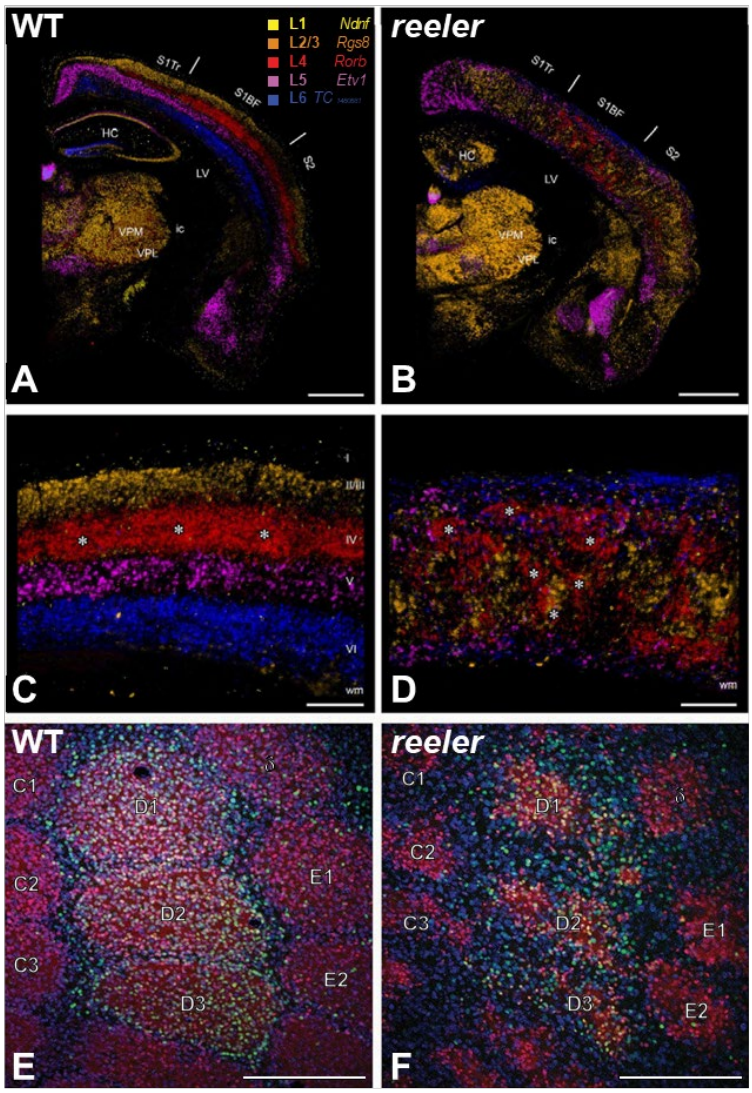

Figure 1.13| Independence of laminar and regional iden-tity: evidence from the reeler cortex.
A-D) Laminar fate markers show the dramatic disorganization of cortical layers in the *reeler* brain. Coronal hemisections through primary somatosensory (barrel) cortex of a wild type (***A, C***) and a *reeler* (***B, D***) mouse brain. Laminar fate markers (labeled in ***A***) have been stained by in situ-hybridization in serial sections, false color-coded and overlaid to obtain a comprehensive impression. The overview (***A, B***) shows that the general anatomical layout of subcortical nuclei and cortical areas is largely intact. A magnified view of the barrel field (***C, D***) shows how its normal laminar organization (***C***) is disrupted in the reeler mouse (***D***). Barrels are denoted by asterisks. (modified from Guy and Staiger, 2017, by permission of Oxford University Press). **E-F)** Normal organization and functioning of cortical regions in the reeler mouse. Neuronal activation in the barrel-related column of wild-type (***E***) and reeler (***F***) transgenic mice that label LIV targets of lemniscal afferents with tdTomato (red) after tactile exploration of a novel, enriched environment. The whiskers corresponding to the stimulated columns (D1-D3) were used during exploration, whereas those of the unstimulated columns (C1-3, d, E1-2) were clipped before exploration. Neuronal activity, visualized by c-fos staining (green), is higher in the stimulated columns than the unstimuated ones. (modified from Wagener et al., 2016). **Abbreviations:** HC, hippocampus; ic, internal capsule; LV, lateral ventricle; S1BF, barrel field; S2, secondary somatosensory cortex; S1Tr, primary somatosensory cortex, trunk region; VPL, ventral posterolateral thalamic nucleus; VPM, ventral posteromedial thalamic nucleus; wm, white matter. Barrels are labeled according to standard nomenclature. Scale bars: A, B 1000 µm; C-F 250 µm.

How are these axes related? This question was tested in the barrel cortex of the reeler mouse using multiple orthogonal methods to determine topographic map layout[130, 131]. These studies found that topographic map formation does not require correct lamination, though maps formed in a disorganized cortex are also disorganized. In other words, laminar identity along the radial S-D axis is independent of regional identity along the A-P and M-L axes. To be sure, barrels in the reeler cortex are disorganized. Rather than tightly packed and relegated to a single cortical layer, the barrels are extended and diffuse, spanning much of the S-D axis of the cortex. But connectivity is largely unaffected: the lemniscal afferents still target (normally) L4 Rorb+ spiny stellate neurons despite their mislocalization. Remarkably, when viewed tangentially from the cortical surface these barrels are properly positioned[130, 131] (Figure 1.13). The general gridded organization of barrels is largely intact, with barrel spatial organization matching that of their associated whiskers. Moreover, these now diffuse barrels still operate as functionally independent units[150-152]. These data indicate that normal, laminar organization of the cortex is not a prerequisite for target cell identity, proper afferent connectivity, or topographic map formation. Indeed, there is enough developmental plasticity and to compensate for the disorganized radial spatial organization of the reeler cortex.

1.2.5.4. Summary of cortical development

Cortical neurons exhibit remarkable diversity. These cells can be hierarchically organized based on expression patterns, reflecting developmental trajectories. The primary division is between glutamatergic and GABAergic cells, with the latter originating from the caudal and medial ganglionic eminences (CGE and MGE, respectively). The majority of cortical interneurons derive

from the MGE, which can be further classified into PV-expressing and Sst-expressing interneurons, with some expressing both or neither. These are further subdivided into distinct subtypes based on molecular expression patterns, morphology, physiology, and targeting profiles within the cortex. Subcellular targeting is a defining feature of interneuron subtypes: SstCs target dendrites, BCs target soma and proximal dendrite, and ChCs target AIS.

Interneuron differentiation has been hypothesized to follow a proposed three-stage process: Cardinal Specification, Definitive Specification, and State Specification. The metaphorical epigenetic landscape serves as a framework for understanding how cells transition between states during development. Ultimately, interneuron fate is determined by an interplay between intrinsic expression profiles and extrinsic (environmental) cues. Our study is primarily focused on how the environment shapes interneuron identity, and whether a precisely organized environment is required for proper subcellular targeting.

The cortex displays a remarkable level of spatiotemporal organization. Spatially, the cortex can be distinguished along two axes: laminar (superficial-to-deep) and regional (anterior-posterior and medial-lateral). Laminar organization involves the stratification of the cortex into layers, with distinct cellular and connectivity patterns. Cortical laminar organization results from the inside-out pattern of corticogenesis, where later born cells migrate past earlier born ones toward the cortical surface. This yields rigid stratification of glutamatergic neurons. GABAergic interneurons migrate tangentially into the cortex, then establish their final resting position by radial migration. This leads to GABAergic neurons being interspersed throughout the layers. Each layer has unique cellular compositions and connectivity patterns.

Regional identity is associated with different functional areas of the cortex. It is initially established through complex interactions between intrinsic genetic programs and extrinsic signaling

cues (morphogens). These interactions result in graded expression patterns of transcription factors that drive regional identity programs. Later in development, afferent, efferent, an intracortical connections further shape regional identity through activity-dependent death of unintegrated neurons. Thus, synaptic specificity and experience shape the cellular composition of cortical regions. This protracted developmental process results in topographic maps, which represent the spatial organization of sensory inputs as a spatially-organized map on the cortical sheet.

Cortical laminar and regional identity are differentially susceptible to spatially-organized environmental cues. Genetic loss of function, overexpression, transplantation, and surgical experiments reveal the remarkable plasticity of regional identity, and how it results from spatially-organized extracortical and intracortical cues. In contrast, studies on the reeler mouse reveal that glutamatergic neuron laminar identity is rigidly determined at birth and unperturbed by disorganization along the superficial to deep axis. The reeler mouse also demonstrates the independence of these two axes of identity.

1.3. Subcellular compartmentalization: cellular and molecular models

The primary question of this study is focused on neuronal subcellular targeting. Based on Sperry's chemoaffinity hypothesis, preferential subcellular targeting is associated with differential molecular compositions at different subcellular regions. Thus, to understand subcellular targeting, it is critical to understand subcellular compartmentalization.

Compartmentalization is a fundamental principle of biology[151, 153-159]. It is utilized to regulate local molecular concentrations, increase the efficiency of reactions, and isolate biological processes[159, 160]. Compartmentalization is initiated through chemical seeding events, wherein asymmetric local concentrations of molecules lead to positive feedback and feedforward loops of molecular recruitment. These processes culminate in molecularly distinct and highly specialized subregions of the cell. These subregions can be compartmentalized in the classical sense, with a three-dimensional structure wherein a semipermeable barrier separates a volume of the cell[151, 157, 160]. This form of compartmentalization is the basis for distinct subcellular organelles, and even the cell itself. Alternately, the compartment can be two-dimensional, as occurs with membrane-bound components[151, 160, 161]. In both cases, a diffusion barrier restricts molecular flow into and out of the region, leading to localized concentrations. These concentration changes can exist as highly discrete and abrupt changes, or as more diffuse molecular gradients. Moreover, it is important to recognize that polarization and compartmentalization are intrinsically linked[162]. The concentration asymmetry that arises due to the nucleation event is, by definition, polarizing. Finally, these subcellular compartments can become self-sustaining, meaning that the initiation event (and molecules involved) may not be required for compartment persistence.

Neurons are by far the most complex and compartmentalized cell type in the body[153, 157, 161, 163, 164]. The large number of spatially and molecularly distinct subcellular compartments results in

just as many incompletely isolated systems that operate pseudo-independently while also being connected to the whole. In other words, compartmentalization strikes a balance between autonomy and connectedness of a heterogenous system[154, 165-167]. But what does this compartmentalization actually look like? The primary division is between axons and dendrite, which contain distinct organelles and fulfill distinct roles in synaptic communication. As a general rule, dendrites are regions of synaptic input, whereas axons are associated with output[155]. Although the soma is not strictly a dendritic compartment, it contains the same organelles as dendrites and is a major target of synaptic input. It is thus both functionally and molecularly more similar to dendrites than axons. Each of these compartments contains morphologically and molecularly distinct sub-compartments. A schematic depicting the generalized subcellular localization of voltage-gated ion channels illustrates this point (Figure 1.14). On the presynaptic side, distinct biophysical organizations and electrochemical properties contribute to distinct functional roles in action potential initiation and propagation[155, 168]. On the postsynaptic side, many of these subcompartments are preferentially targeted by different presynaptic cell types. The relative spatial organization of these compartments synergizes with distinct electrochemical properties to yield functionally distinct responses to this differential input[97, 145, 163, 167, 169]. We will first discuss axonal and dendritic subcompartments,

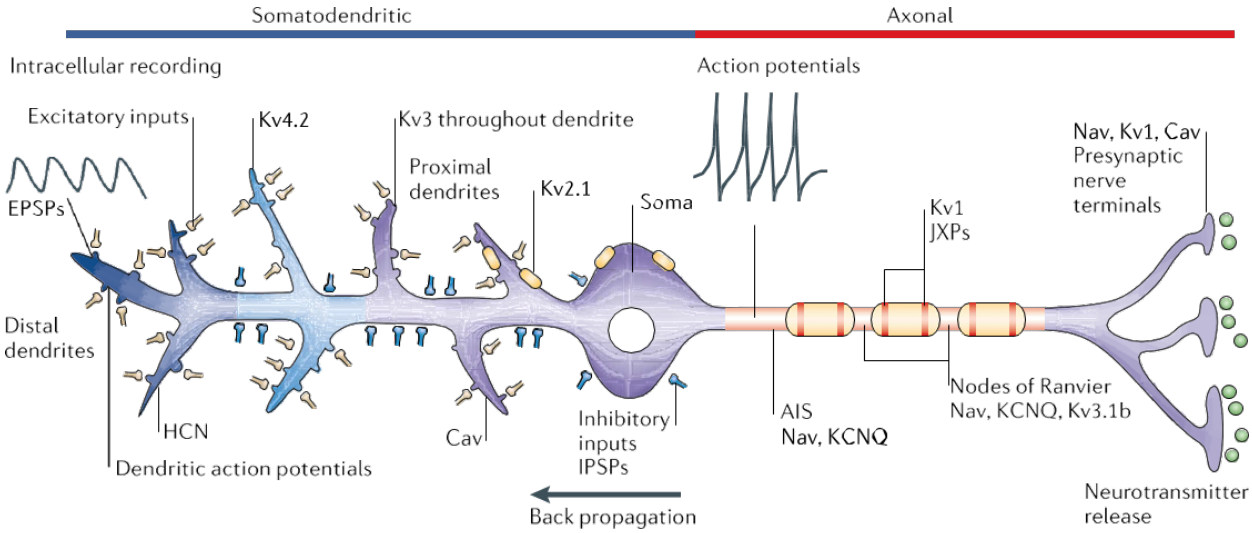

Figure 1.14 | General localization of voltage-gated ion channels in a model neuron.
In general, Nav channels are found in the axon initial segment (AIS), nodes of Ranvier and presynaptic terminals. Voltage-gated potassium Kv1 channels are found at the juxtaparanodes (JXPs) in adult myelinated axons and presynaptic terminals. The Kv channel KCNQ is found at the AIS and nodes of Ranvier, and Kv3.1b channels are also found at the nodes of Ranvier. Canonically, excitatory and inhibitory inputs (EPSPs and IPSPs — excitatory and inhibitory postsynaptic potentials; yellow and blue presynaptic nerve terminals, respectively) from the somatodendritic region spread passively to the AIS where action potentials are generated by depolarization, and travel by saltatory conduction to the presynaptic nerve terminals to activate voltage-gated calcium (Cav) channels that increase intracellular calcium levels, thereby triggering neurotransmitter release. Hyperpolarization-activated cyclic-nucleotide-gated (HCN) channels have a gradient distribution that increases in density from the soma to the distal dendrites (dark blue shading). Kv2.1 channels are found in clusters on the soma and proximal dendrites (light yellow ovals). Kv3 channels are found throughout the dendrite. Kv4.2 channels are located more prominently on distal dendrites (light blue shading). Kv channels in the dendrites contribute to controlling back propagation. Strong enough inputs in the dendritic region can generate dendritic action potentials. Dendritic Cav channels increase in density toward the proximal dendrites and the soma. Figure and legend from Lai & Jan, 2006. Reproduced with permission from Springer Nature.

then cover cellular and molecular models for how this compartmentalization arises, with a particular focus on compartmentalization of target compartments. Since the synapse is formed by a unique interaction, we will discuss it separately in its own section.

1.3.1. Axonal subcompartments

Axonal subcompartments include the axon initial segment (AIS), the axonal shaft, nodes of Ranvier (for myelinated axons), and presynapse[145]. Each of these subcompartments is molecularly distinct and serves different functions. For our purposes we will focus on the AIS, since it is critical for establishing axodendritic polarity and is targeted by chandelier cells. We will not discuss the axonal shaft in depth, and instead note here that it can be viewed primarily as a conduit for signaling and molecular trafficking, and specifically conveyance of action potentials[162, 170]. We discuss the presynapse separately in a section on synaptogenesis.

1.3.1.1. Axon initial segment

The AIS is the site of action potential generation[162] and is critical for the establishment and maintenance of axodendritic polarization[171, 172]. The AIS is developmentally defined at the onset of axodendritic polarization by TRIM46 and AnkG clustering, which it retains throughout its existence[161, 172]. The axon initial segment is not only the initial spatial subcompartment of the axon, it is also developmentally primary: the clustering of AnkG and TRIM46 leads to formation of parallel microtubule arrays in the AIS and throughout the axon, a defining characteristic of the whole axonal compartment. Of note, without either AnkG or TRIM46 neurons do not undergo axodendritic polarization[171-173]. Other cytoskeletal components unique to or enriched in the AIS include βIV-Spectrin and TRIM46, which lead to a unique cytoskeletal actin filament formation and microtublue organization at this region[173]. These molecules directly and indirectly serve as a scaffolding for molecules localized to the AIS membrane, including ion channels, neurotransmitter receptors, and adhesion molecules[174]. Specific ion channels, including voltage-gated sodium (NaV) and potassium (KCNQ) family members, localize to the AIS and contribute to its distinct electrical properties as the site for action potential initiation. Specific adhesion molecules are also

enriched at this site, such as NF186 in cerebellar Purkinje cells, which has been shown both to define the local ECM[175] and to mediate specific synaptic targeting of this region[163, 176].

1.3.2. Dendritic subcompartments

The dendrite consists of many subcompartments[168, 169]. Dendritic subcompartments include the soma, proximal dendrites, medial, and distal dendrites, as well as dendritic spines. In cortical and hippocampal pyramidal cells, dendrites are further classified by morphology and branching pattern (Figure 1.15). Pyramidal neurons have multiple short basal dendrites, which project directly from the soma, and a singular apical dendrite, which extends from the cell body into the most superficial layer. The apical dendrite can be further broken down into multiple molecularly and morphologically distinct subdendritic compartments, such as the primary apical dendrite, apical obliques (which can be ordered hierarchically secondary, tertiary, etc. obliques), and the distal dendritic tuft. Basal dendrites can also be ordered hierarchically. Alternately, branches can be classified as primary, intermediate, or terminal, with distinct functional properties and synaptic inputs at each level[177, 178]. Finally, dendritic spines are morphologically distinct postsynaptic subcellular compartments that form in response to glutamatergic innervation.

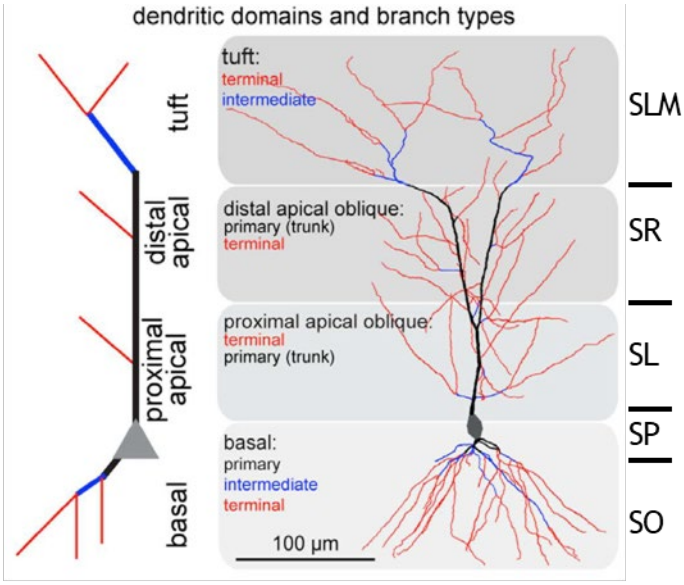

Figure 1.15 | Branch and sub-branch organization of morphologically defined pyramidal neuron dendritic domains.
A cartoon (*left*) and AT-reconstructed example of a CA1 pyramidal cell (*right*) depicting the breakdown of the arbor into dendritic domains (gray regions) and branch types (black, blue, and red). Approximate hippocampal laminar boundaries are indicated. From Bloss et al., 2016.

Although there is an extensive body of literature outlining the cellular and molecular mechanisms of dendritic spine formation, there is relatively little on these other subdendritic compartments. This is unfortunate for our purposes, since the majority of GABAergic subcellular targeting from the MGE lineage is onto these non-spine regions[155, 179]. For this reason, we discuss dendritic spines only in passing, and handle all of the other dendritic subcompartments as a group.

1.3.2.1. 1.2.2.1 Molecular enrichment patterns of dendritic subdomains

These dendritic compartments are not just morphologically distinct, they are also molecularly distinct[156, 180-183] (Figure 1.14). The pan-neuronal, canonical dendrite marker MAP2 is found in all dendritic compartment[184, 185], but not in axons. Dendritic spines, for instance, are enriched for F-actin[186, 187] as well as molecules rarely found outside these compartments, such as spinophilin and drebrin[156, 185-189]. Spine necks are also molecularly distinct from spine heads, with the former enriched for βIII-Spectrin[180] and the latter specific ion channels like Kv4.2, Kir3.1, and Kir3.2[190]. There are also specific markers for other dendritic subcompartments. Soma and proximal dendrite, for instance, are enriched for the ion channels TRPM4, Kv2.1, and Kv2.2[191-197], and the adhesion molecule AMIGO1[198]. In the cortex, L4 and L5 pyramidal cells show a polarized localization of a5 integrin, which is specifically localized to primary apical dendrites, but not other subdendritic regions[199]. Semaphorin 3F receptors are also exclusive to primary apical, but not apical oblique or tuft dendrites of hippocampal pyramidal cells[200]. NGL1 shows a similar primary apical dendritic localization pattern in L5 cortical pyramidal neurons[201], with specific enrichment of NGL1 in L4. Distal dendritic tuft is also enriched for specific molecules, such as HCN[202, 203] and Kv4 ion channels[200, 204] and their respective ancillary subunits[201, 205], cytoskeletal proteins[206, 207], cell surface receptors[208], and adhesion molecules[207, 208].

Importantly, apart from MAP2, the above molecules are generally not ubiquitously expressed in neurons. These enrichment patterns are not necessarily static, but can be dynamic, with enrichment of specific molecules occurring at different subcellular locations during different development stages. In the mouse visual cortex, for instance, the adhesion molecule Telencephalin is found predominantly in dendritic spines during early visual development and in dendritic shafts in mature mice[207]. There are also brain-region-specific expression patterns. L1CAM, for instance, is expressed in L1 (distal dendrites) of adult piriform and medial cortex, L5 of remaining neocortex[174, 198]. Moreover, although these subdomains can have highly defined borders, as in spines or primary apical dendrite[200], they can also exist as subcellular gradients, as is evident with the NF186 at the soma and AIS[205] or HCN channels in distal dendritic tuft[93]. Finally, the compartment-specific differences are not necessarily concentration-based. Cofilin, a cytoskeletal protein that functions in actin depolymerization by promoting disassembly of actin filaments, is phosphorylated in a reelin-dependent manner in the developing marginal zone, which will in turn become the distal dendritic tuft. This cofilin phosphorylation thus leads to distinct cytoskeletal dynamics within the distal dendrite[183, 209], likely acting as an integral step in defining this dendritic subdomain and establishing the observed molecular gradients. Again, with the notable exception of dendritic spines, the molecular profiles of dendritic subcompartments are generally poorly characterized. This is especially true of synaptic adhesion molecules, the subdendritic localization patterns of which are generally unknown. It is, however, reasonable to assume that the same subcellular enrichment patterns witnessed for ion channels or cell surface receptors are parallelled by adhesion molecules, as these proteins tend to form complexes and utilize overlapping mechanisms for subcellular localization.

1.3.2.2. Dendritic subdomains and lamination

These distinct apical and basal dendritic subcompartments align with specific laminae. In the hippocampus, which has a more simplified, less overlapping lamination pattern than the cortex, pyramidal cell basal dendrites populate the SO, the distal dendritic tuft corresponds to SLM, and each layer in between shows distinct dendritic branching patterns (Figure 1.15). This principle of subdendritic compartments of specific cells being enriched in specific laminae also holds true in the cortex, but is less clearly visible due to this lamination itself being layered (Figure 1.16). In the cortex, regions containing the medial or basal dendrite of one cell type are also populated by cell bodies from another cell type: basal dendrites often extend to the next-deepest layer, and apical compartments extend into more superficial layers until their termination as dendritic tuft in L1. Thus, layer-specific targeting in the cortex is often a reflection of combined cell-type- and sub-compartment-specific targeting.

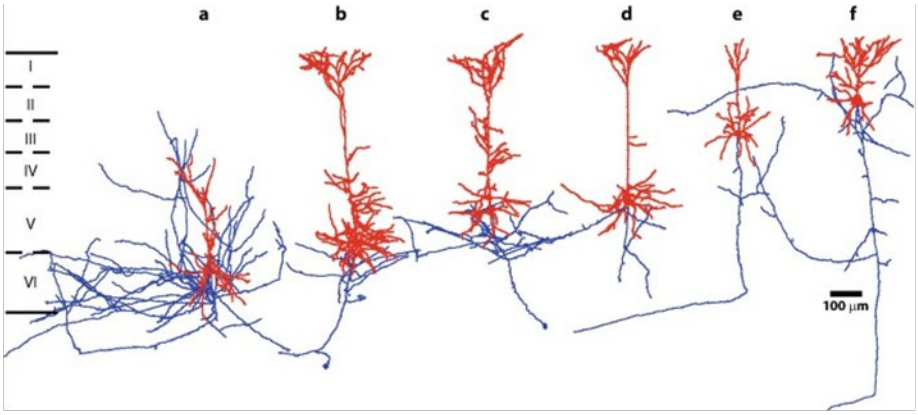

Figure 1.16| Glutamatergic cell in the cortex have layered subcellular compartments. Traces showing axonal (blue) and dendritic (red) arbors of individual glutamatergic neurons from different depths. Neurons ***a*** through ***f*** are from sequentially more superficial layers. Note the layering of basal and apical oblique dendrites across these cells. From Araya 2020, reproduced with permission from SNCSC.

1.3.3. Cellular and molecular mechanisms of compartmentalization

These localization patterns can arise through cell intrinsic programs or inductive environmental interactions. The subcellular domain can also specify the extracellular environment. Afferent

connections can be functional actors in the specification of domains. The molecular composition of subcellular domains can also be impacted by activity. We discuss all these here.

1.3.3.1. Cell intrinsic subcellular compartmentalization

A standard method for assessing whether subcellular compartmentalization is driven by cell-intrinsic programs or cell extrinsic ones is to plate isolated, dissociated cells in culture and assess whether or not the molecules compartmentalize in this context in the same way as in their native habitat. This approach has yielded considerable insights into cell intrinsic developmental programs, revealing that polarization and the initial stages of compartmentalization are intrinsically determined. Specifically, isolated pyramidal cells will polarize, specifying axon and dendrite as distinct compartments[203], as is apparent through the differential localization of the dendritic and axonal markers, MAP2 and TAU, respectively[210]. Moreover, given these conditions, neurons will cell intrinsically specify distinct axonal subcompartments, including the AIS, axonal shaft, and growth cones. Subdendritic compartments are also intrinsically determined, with isolated pyramidal cells showing distinct molecular compositions of soma and proximal dendrites, which are positive for Kv2.1 and Kv2.2, compared to more distal dendritic compartments, which are not[211]. Pyramidal cells intrinsically specify the basal and apical dendrites as a distinct compartments[212], with further differences between the primary apical dendritic process and oblique ones. Importantly, such intrinsic compartmentalization suggests that postsynaptic adhesion molecules that mediate subcellular targeting specificity can also be enriched at these subcellular locations via the same cell intrinsic programs.

1.3.3.2. 1.2.3.2 Specification of the extracellular environment by subcellular domains

Different subcellular domains can also differentially specify their extracellular environment. Perhaps the best characterized example of this phenomenon is the perineuronal nets (PNNs), a

specialized ECM that forms at the soma, proximal dendrites, AIS, and nodes of Ranvier[212], forming a mesh-like structure that interdigitates synapses. PNNs are formed via cell intrinsic localization of surface receptors, which in turn interact with and localize diffusely secreted extracellular matrix molecules[173], thereby forming an immobilized extracellular scaffold. Perineuronal nets are comprised of distinct molecules, including aggrecan, versican, neurocan, brevican, with distinct molecular compositions at distinct subcellular regions and cell types[212]. The PNN surrounding the AIS, for instance, is enriched for brevican[213], which is secreted by glial cells[151]. A similar mechanism occurs at the synapse, where pre- and postsynaptic receptors such as Nueruexins and GluDs, bind to diffuse extracellular molecules like cerebellins, thereby forming an extracellular scaffold that mediates transsynaptic signaling. We discuss this transsynaptic scaffold in greater depth in the section on synaptogenesis below.

1.3.3.3. 1.2.3.3 Specification of subcellular domains by inductive environmental factors

Despite this remarkable level of intrinsic programming, there are multiple instances where compartmentalization is driven by extrinsic factors. Environment-induced compartmentalization can occur via direct, adhesion molecule mediated cell-cell interactions, or via diffusible secreted molecules[200], cell-ECM[214, 215], and even in a ligand-independent manner via electrical field stimulation[200, 201]. Given our division of acting parties into presynaptic, postsynaptic, and environment, it is fitting to distinguish non-synaptic compartmentalization from that which occurs due to transsynaptic cross-talk. The latter, where transsynaptic signaling leads to formation of pre- and postsynaptic compartments, is a specific case of subcellular compartmentalization generally considered in the context of synaptogenesis. We therefore discuss it separately.

An example of subcellular compartment specification by non-synaptic communication occurs at the dendritic tuft. *In vivo*, this region shows a distinct enrichment of HCN[202, 203] and Kv4 ion

channels[200, 204] and their respective ancillary subunits[201, 205, 206], cytoskeletal proteins[207], cell surface receptors[208], and adhesion molecules[200]. Kupferman et al., showed that HCN molecules are not enriched in distal dendritic domains in isolated hippocampal neurons[216], indicating this localization pattern is driven by cell extrinsic mechanisms. Instead, proper HCN channel localization to distal dendrites is mediated by reelin, a large ECM glycoprotein secreted by Cajal Retzius cells in the superficial layers of the cortex and hippocampus (SLM in Hpc, L1 in Ctx). Reelin interacts with the cell surface receptors VLDR and ApoER2, which in turn leads to phosphorylation of the intracellular protein Dab1 and downstream signaling. Knocking down Dab1 after the (reelin-dependent) spatial organization of the hippocampus had already been established led to the loss of the graded localization pattern of multiple dendritic tuft proteins, but had no effect on dendritic proteins with a uniform distribution or overall dendritic arbor morphology[40]. Thus, the localization of these distal dendritic components is specifically regulated by cell-ECM interactions, where the ECM is differentially defined in specific laminae via 3rd party cells.

1.3.4. Summary of subcellular compartmentalization

This study focuses on neuronal subcellular targeting. A critical aspect of this process is subcellular compartmentalization. Compartmentalization, essential in biology, regulates localized molecular concentrations and enhances reaction efficiency. It involves chemical seeding events leading to molecularly distinct subregions within cells. Neurons exemplify complex compartmentalization. The primary division of neuronal compartments is between axons and dendrites. Axonal subcompartments include the axon initial segment (AIS) and presynapse, while dendritic subcompartments include soma, proximal and distal dendrites, dendritic spines, and postsynapses. Thus, compartmentalization of both the presynaptic and postsynaptic cells is critical in establishing stereotyped connections. We revisit synaptic compartmentalization below, and focused above on

compartmentalization of target compartments. These subcompartments can be defined both morphologically and molecularly, with different branching and molecular expression patterns characterizing different subcellular domains. Developmentally, compartmentalization is driven by cell intrinsic programs or inductive environmental factors. For example, reelin-mediated signaling regulates the localization of HCN channels in dendritic tufts, showing the impact of cell-ECM interactions. Overall, we highlights the dynamic interplay between intrinsic programs and extracellular factors in shaping neuronal compartments and the extracellular environment.

1.4. General principles of synaptic development

1.4.1. Pre-contact: a question of spatial proximity

Synaptic target selection requires a physical interaction between the presynaptic and postsynaptic cells. Since most neurons are born in distinct regions of the nervous system, prior to target selection the problem can be simplified into decreasing the spatial proximity between the presynaptic cell and its 'appropriate' target to the point of physical contact. As discussed in greater detail below, target selection is mediated by complementary adhesion molecules on the pre- and postsynaptic cells. Since molecular expression profiles are dynamic, this contact event often must occur within a defined developmental window. Simply put: right place, right time, and right equipment equals right connection. Many of these complementary adhesion molecules are recycled in different contexts, and so it is very common for cells in the 'wrong' place or at the 'wrong' time to have the 'right' equipment. Ectopic or 'inappropriate' localization of the growth cone among complementary but 'inappropriate' targets and ectopic localization of 'inappropriate' but complementary targets within the 'proper' path of the growth cone will lead to the formation ectopic connections. Such ectopic connections will change the flow of information, leading to altered information processing, with potentially disastrous consequences for the organism. Thus decreasing proximity between 'appropriate' partners while avoiding 'inappropriate' ones is critical for establishing stereotyped neuronal connectivity.

The nervous system has evolved multiple means of solving this problem, namely migration, neurite elaboration with directed growth, and subcellular compartmentalization (reviewed in [151,

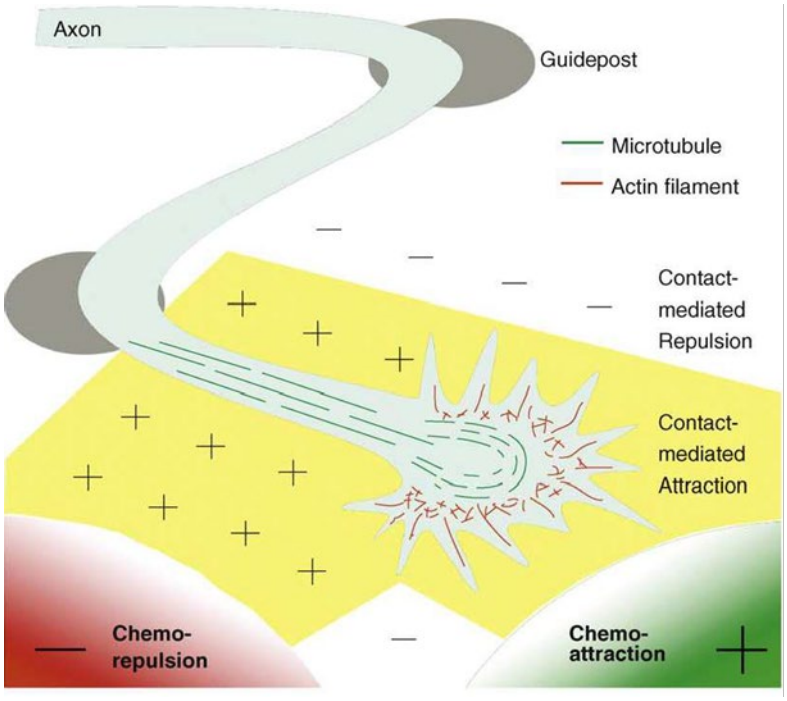

Figure 1.17 | Mechanisms of axon and dendrite guidance. A schematic of an axonal growth cone responding to guidance cues in its environment, resulting in a specified path of growth. Guideposts are discrete cells or cellular structures utilized by axons as spatial cues. The axon trajectory can be divided into multiple segments between guideposts. The guidance process between each guidepost is mediated by four types of forces, long-range chemoattraction or chemorepulsion, and short-range contact-mediated (haptotactic) attraction or contact-mediated repulsion. Note that although an axon is used as an example, the same mechanisms also apply to dendrites. Figure from Chen & Cheng, 2009.

[157, 163, 216, 217]). In all these developmental processes, this spatial organization is a strategy for limiting potential partners and optimizing the efficiency and probability of forming stereotyped, 'appropriate' connections. A newly born neuron must migrate from its place of birth to its final location, whereupon it undergoes a process of polarization and neurite elaboration, specifying axon and dendrite as distinct subcellular compartments. During this elaboration phase, the axons and dendrites arborize, orient themselves, and elongate, growing towards 'appropriate' regions in a process called axonal and dendritic pathfinding. The specific orientation and growth target are highly stereotyped and distinct for the axonal and dendritic growth cones, leading to stereotyped coverage patterns of these processes. Neurite guidance is mediated by molecular interactions between receptors on the growth cone and extracellular signals (Figure 1.17)[218]. These signals can act in a discrete or graded fashion, and can take the form of attractive or repulsive cues[1, 219]. They can be due to contact-mediated interactions with adhesion molecules or the extracellular matrix

that are stabilized in place (haptotactic interactions)[135, 220-222], or long-distance interactions via secreted molecules that diffuse through the environment (chemoattraction and chemorepulsion). Finally, non-target cells can act as scaffolds or guideposts for the growth cone during this guidance phase. Importantly, in an intact brain, these extrinsic molecular cues will often have highly stereotyped spatial organization and temporal expression windows, which facilitate neurite pathfinding and in turn decreases the probability for ectopic connections.

To summarize, prior to target selection, the main problem to forming stereotyped connections is decreasing spatial proximity between presynaptic growth cone and postsynaptic target. This process is solved through a two distinct but highly coordinated interactions, namely 1) the presynaptic cell interacting with its microenvironment, and 2) the postsynaptic target interacting with the target microenvironment. Here we are defining 'microenvironment' as anything other than the presynaptic cell and the postsynaptic target. Thus, although synaptic target selection and maturation are ultimately responsible for the specificity in these connections, the environment creates a framework for the precise assembly of neural circuits by regulating migration, neurite guidance, and subcellular compartmentalization. Finally, although we present the following models specifically from an *axon* guidance perspective, these mechanisms can also apply to *dendritic* guidance.

1.4.1.1. Axon guidance along diffusible molecular gradients in subcellular targeting

Diffusible molecules act to guide neurites and limit their extension, thereby functioning as a framework for directed axonal and dendritic outgrowth. These diffusible cues can take many forms, including Netrins[223], semaphorins[224-226], and slits[175, 225]. They can be attractive[223] or repulsive[224]. Moreover, the same molecule can act as a chemorepellent in one context and a chemoattractant in another. For example, in the cortex, Sema3A acts as a chemorepellant to pyramidal

cell axons[175] but a chemoattractant for their dendrites[227]. The polarized Sema3A expression gradient (it is more highly expressed in superficial layers) leads to the characteristic axodendritic orientation in the cortex, thereby limiting putative synaptic partners. The response is not compartment-specific: whereas Sema3A is a repulsive signal for cortical pyramidal cell axons, for cerebellar basket cell axons it acts as an attractive signal[228]. The effective intercellular communication range of these diffusible molecules is not constant, but is instead determined by the secretion and diffusion rates for each molecule, as well as the molecule's binding affinity to its receptor[229]. Assuming only entropic diffusion, Francis and Palsson calculated that the maximum effective distance for soluble chemotactic molecules, such as the neurotrophic factors BDNF and NGF or morphogens Shh and FGF8, is 25 cell diameters. In other words, this distance changes depending on the size of the subcellular compartment: if secreted from the soma, the range is approximately 525 um (range: 375-675 um)[175], but decreases significantly for thinner neurites. In many instances, the signaling molecule interacts with molecules in the extracellular matrix (ECM). If this interaction is reversible, such binding curtails the diffusion rate and thus the effective signal range. If it is irreversible, the molecule will act like a substrate-bound haptotactic cue. Moreover, the same molecule can function as a soluble, chemotactic cue, or as a substrate-bound, haptotactic cue, depending on the cellular microenvironment and posttranslational modifications[230].

The majority of the literature on axon guidance via diffusible molecules is associated with regional targeting of long range projection neurons. However, diffusible molecules are also important mediators of subcellular targeting[27, 198, 221]. The hippocampus has classically been used as a model system to study subcellular targeting due to its well-characterized connectivity and laminar organization (Figure 1.18). Here, perforant path afferents from the entorhinal cortex (EC) target the stratum luconosum-molecalare (SLM), which consists primarily of the distal dendritic tuft

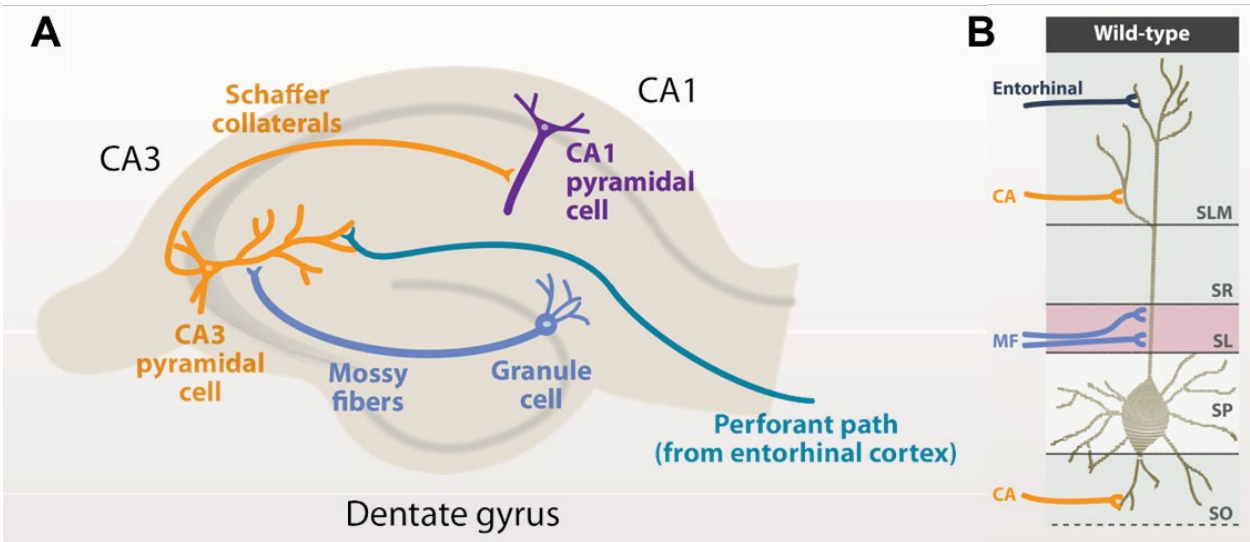

Figure 1.18 | Lamina-specific innervation of hippocampal neurons by distinct inputs.
A) Apical dendrites of pyramidal cells in CA1 and CA3 and of granule cells in the dentate gyrus receive inputs from distinct sources on discrete dendritic segments. **B)** In CA3, axons from entorhinal cortex, commisural/associational afferents (CA), and dentate gyrus [mossy fibers (MF)] innervate distinct dendritic segments in the stratum lacunosum-moleculare (SLM), stratum radiatum (SR), stratum lucidum (SL), and stratum oriens (SO) as shown. Adapted from Sanes & Yamagata, 2009.

of pyramidal neurons. In contrast, CA3 pyramidal neurons project ipsilaterally onto CA1 stratum radiatum (SR) and stratum oriens (SO), which correspond to medial apical (SR) and basal (SO) dendrite of the pyramidal neurons[198, 221]. Both targeting patterns are netrin-1 dependent. In netrin-1 deficient animals, the targeting region of EC afferents is expanded to include the SR, whereas CA3 axons targeting was expanded to include the SLM[217].

Another example of diffusible molecules regulating subcellular targeting is semaphorin-3F (Sema3F) in the cortex and hippocampus. Deletion of Sema3F or its receptor Npn2 led to increased spines in the proximal regions of the primary apical dendrite in L5 cortical pyramidal neurons and DG granule cells[218]. This spine phenotype was not witnessed in basal, apical oblique (secondary), or distal apical (tuft) dendrites, indicating that the increased excitatory inputs were specific to a single subcellular region (the primary apical dendrite). Importantly, using isolated hippocampal cells in culture, Tran et al., showed that Sema3F specifically bound to primary apical dendrites,

whereas Sema3A bound to basal as well as primary and secondary apical dendrites, indicating that the Sema3F cell surface receptors are cell intrinsically localized to primary apical dendrites, and that diffusible Sema3F negatively regulates synapse formation at this subcellular region.

1.4.1.2. Axon guidance and targeting by adhesion molecule gradients

Axon guidance can also be driven by adhesion molecules in the environment[219, 220, 231, 232]. Multiple families of adhesion molecules have been implicated in axon guidance, including membrane-bound semaphorins/plexins [233-236], Netrin-G/NGL[220, 237], Ephrin/Eph[238], DIPs/Dprs [10] [239-241], and the L1 family of adhesion molecules[3, 216] (reviewed in [220, 242-244]).

The best characterized of these is Ephrin/Eph signaling, which mediates regional targeting and topographic map formation (Figure 1.19). Classic examples occur in the visual system, where spatially patterned expression of different Ephs and Ephrins mediate key decision points for retinal ganglion cell (RGC) axons as they grow toward their targets in the thalamus and superior colliculus (SC; in non-mammalian vertebrates, optic tectum)[220, 245-249]. One of the major decision points along this route is whether or not to cross the midline (Figure 1.19, A). Axons that cross the midline ('contralateral axons') innervate the contralateral brain regions, whereas those that don't ('ipsilateral axons') innervate ipsilateral ones. This decision is mediated by repulsive interations between EphBs and eprhin-Bs at the optic chiasm (Figure 1.19, B). EphB-expressing RGC growth cones are repulsed by ephrin-B-expressing non-target cells at the optic chiasm. This diverts their trajectory toward target cells in the ipsilateral SC, the non-default state. In contrast, axons from RGCs that do not express EphB are unaffected, and proceed to the contralateral SC. The number of ipsilateral axons correlates with the extent of binocular vision, so a greater proportion of EphB expr-

[10] DIPs and Dprs are primarily studied in drosophila and do not have mammalian homologs. They are mentioned here because their role in axon guidance is particularly well-characterized.

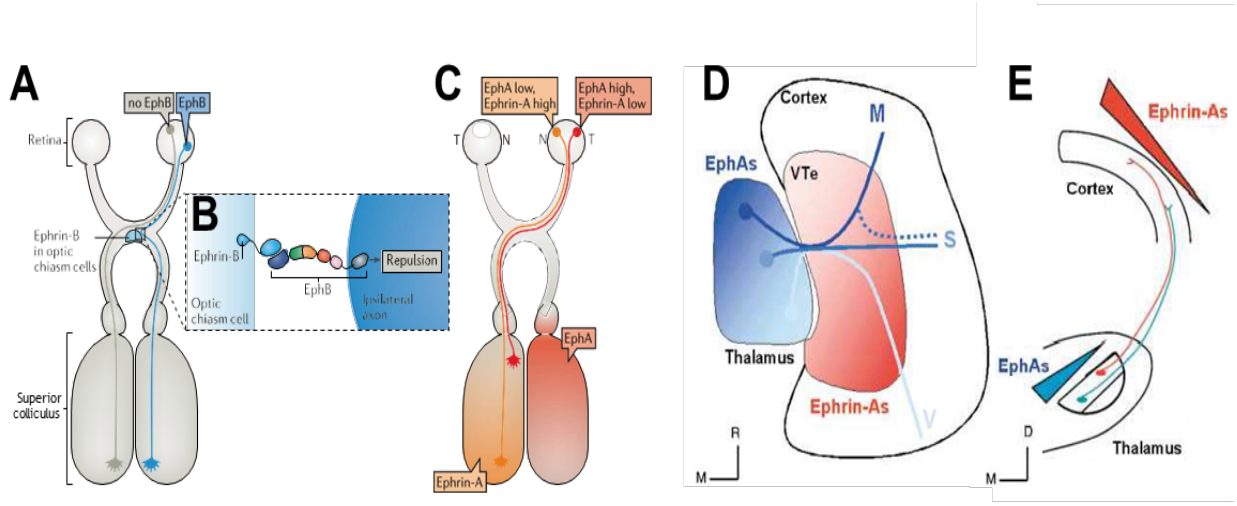

Figure 1.19 | Differential use of ephrin/Eph adhesion molecules in axon guidance during visual system development.

A-C) Ephrin and Eph adhesion molecules contribute to a spatial molecular framework that mediates axon guidance and regional connectivity in the visual system. **A)** Role of ephrin-B and EphB in determining whether a retinal axon crosses the midline. Retinal ganglion cell axons innervat a number of regions, including the superior colliculus (SC) (optic tectum in non-mammalian vertebrates). The decision to cross the midline is mediated by interactions between eprhin-Bs and EphBs at the optic chiasm. Those that cross the midline (so called 'contralateral axons', shown in grey) innervate the contralateral SC, whereas those that don't (so called 'ipsilateral axons', shown in blue) innervate the ipsilateral SC. **B)** Ipsilateral axons express EphB receptors and are repulsed by ephrin-B-expressing cells at the midline, thus preventing their crossing and driving directional growth toward the ipsilater SC. **C)** Eph/Ephrin signalling specifies the positioning of retinal axon termini within the SC. In retinal neurons, EphA displays graded expression: neurons in more temporal region of the retina (T) express high levels of EphAs (red), whearas those in the nasal region of the retina (N) express low levels (orange). Target cells in the SC express a compelemntary gradient of ephrin-As (orange): neurons in the anterior SC express low levels of ephrin-A whereas those in the posterior express high levels. There is also a countergradient of ephrin-As (red), which is formed by an inverse expression pattern: cells in the retina and SC that express high EphA levels express low ephrinA levels, and vice versa. Thus, more temporal axons (red) target more anterior regions of the SC, whereas more nasal axons (orange) target more posterior more anterior regions. These expression patterns results in a topographic map in the SC that reflects the spatial organization of the associated visual receptors in the retina. From Kania and Klein, 2016. *Reproduced with permission from Springer Nature.*

D, E) Role of ephrin-A and EphA in corticothalamic axon guidance. **D)** EphAs (blue) are expressed in a graded fashion in the thalamus, whereas ephrin-As (red) are expressed in a graded fashion in the ventral telencephalon (VTe), an intermediate area through which thalamocortical axons pass on their way to the cortex. These graded expression patterns in the thalamus and VTe guide thalamocortical axons in the tangential plane. Rostral-medial TC axons (dark blue) travel through the rostral part of the VTe and subsequently target the rostral part of the cortex (e.g.,

motor cortex, M), whereas axons from more caudal-lateral thalamic regions (light blue and grey) travel more caudally through the VTe, therby targeting more caudal cortical areas like the somatosensory (S) or visual (V) cortex. In the absence of proper Eph/Ephrin signaling, axons typically targeting medial regions like the motor cortex (M) regions shit to more caudal area like the somatosensory cortex. **E)** Graded Ephrin-A expression in the cortex (red) and EphA expression of the thalamus (blue) mediate thalamocortical axon guidance in the coronal plane. Axons from medial-ventral areas of the thalamus (green) express high levels of EphAs and therefore target more ventral-lateral regions of the cortex. In contrast, axons from dorsal-lateral regions of the thalamus (red) have lower EphA expression and therefore target more dorsal-medial cortical regions. Figures modified from Dufour et al., 2003.

essing RGCs will enlarge the field of visual overlap.

Eph/Ephrin signaling is also responsible for topographic map formation in the SC (Figure 1.19, C)[250-253]. In the retina, RGCs express EphAs and Ephrin-As along a medial-lateral gradient: more temporal neurons express higher EphAs levels and lower Ephrin-A levels, whereas more nasal neurons have the opposite expression pattern[250]. These expression patterns are mirrored in the SC target region along the anteroposterior axis: more anterior SC neurons express higher EphA and lower Ephrin-A levels, whereas more posterior neurons have the opposite expression pattern. Due to the mutually repulsive effects of the EphA/ephrin-A interaction, axons from more temporal regions of the retina are prevented from growing into more posterior regions of the SC, thus settling in more anterior regions. In contrast, axons from more nasal RGCs are repulsed from more anterior regions and therefore target more posterior areas of the SC. There is also an expression gradient of EphB and Ephrin-Bs in the orthogonal plane, thereby yielding a 3D coordinate system for axon guidance and regional targeting[1, 127, 252]. Such a molecular coordinate system was proposed by Sperry in his chemoaffinity hypothesis[1]. Of note, these patterns are controlled by relative, not absolute, levels of Eph and ephrin signaling[135]. These graded expression patterns result in a topographic map in the SC that matches the spatial organization of receptors in the retina.

Ephs and ephrins also play a role in axon guidance between the retina and thalamus, and again from the thalamus into the cortex. For the retinothalamic axons, graded expression of EphA in the retinal ganglion cell axonal growth cones is matched by similarly graded expression patterns of ephrin-A adhesion molecules in the target cells, which interact repulsively to guide retinothalamic afferents to the appropriate thalamic regions[135]. A similar mechanism is then again utilized by thalamocortical afferents to guide regional targeting in the cortex[1]. Here, thalamic neurons express EphAs in a graded fashion along the rostro-caudal (R-C) axis of the thalamus (Figure1.19, D). These afferents initially grow in a medio-lateral direction through the ventral telencephalon, which expresses a complementary but inverted R-C gradient of Ephrin-As, a repulsive adhesion molecule ligand of the EphA receptor, thereby guiding axonal growth into different rostral and caudal regions of the cortex by finely-tuned expression. Once these axons enter the cortex, they encounter another ephrin-A gradient, this time decreasing along the medio-lateral and dorso-ventral axes, which guides axons toward specific cortical regions[135]. Altering the EphA or Ephrin-A expression levels leads to mistargeting. For instance EphA overexpression in normally low-expressing regions leads to targeting of low (rather than high) ephrin-A expressing target regions, whereas deletion of EphA leads normally high EphA regions to target high (rather than low) ephrin-A expressing regions. Similar results occur with manipulation of Ephrin-A levels[254, 255]. Thus, adhesion molecules on axonal growth cones act as sensors for their environment, guiding axonal growth through a molecular coordinate system to their ultimate destination.

1.4.1.3. Axon guidance along cellular scaffolds

In the cerebellar cortex, Purkinje neurons (PkCs) are the sole efferent and thus form a central node for information processing and transmission. GABAergic stellate cells (SCs) innervate PkC

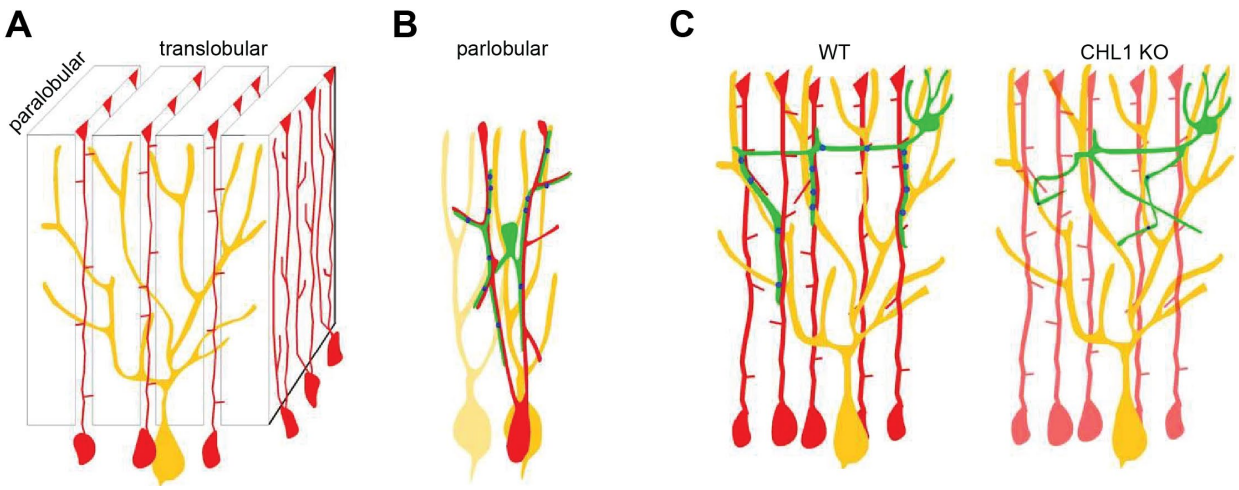

Figure 1.20 | Schematic Representation of the Bergmann Glial Fiber Scaffold and CHL1 in Directing Stellate Axons to Innervate Purkinje Dendrites.
A) In the cerebellar cortex, Purkinje dendrites (*yellow*) are restricted in the translobular plane. Each BG cell (*red*) gives rise to several ascending BG fibers, which extend in both the translobular and parlobular plane. These largely radial fibers from neighboring BG cells further aligned into thin walls, or palisades, in the parlobular plane, perpendicular to the Purkinje dendrites. As a consequence, several BG palisades cut across and impinge upon an individual Purkinje dendrite in a largely vertical orientation. **B)** A stellate axon (*green*) likely contacts segments of multiple intercalated Purkinje dendrites (parlobular view; neighboring Purkinje dendrites are represented in different shades of yellow). Stellate synapses (*blue dots*) are formed or stabilized at the interception of BG fibers and Purkinje dendrites. **C)** In the translobular plane in WT mice (left panel), stellate axons associate with and extend along BG fibers, and are thus organized into characteristic orientations and trajectories towards Purkinje dendrites. In CHL1-deficient mice (right panel), stellate axons can no longer associate with BG fibers, show aberrant orientation and trajectory, and are deficient in synapse formation and/or stability. Figure and legend from Ango et al., 2008.

distal dendrites. In contrast to the previously-discussed models of axon targeting over large, regional gradients, SC axons are guided towards Purkinje cell distal dendrites by a local cellular scaffold from astroglial cells. Specifically, SCs utilize Bergman Glial (BG) fibers to direct axon growth to PkC distal dendrites (Figure 1.20). When the adhesion molecule Close Homolog of L1 (CHL1) is conditionally deleted in BG fibers, SC targeting of Purkinje dendrites is also disrupted[254], indicating that SC axons utilize the BG fiber scaffold as an intermediate structure to target Purkinje cell dendrites. This three-cell process is a good example of how subcellular targeting specificity can be based on the structure of the local tissue environment. Most guidepost cells

are involved in axon guidance of long-range projection neurons[256], but this is an example of the involvement of guidepost cells in a local circuit: the presynaptic stellate cells are interneurons that residing in the most superficial molecular layer (ML) of the cerebellar cortex, whereas both the postsynaptic target (the PkC) and source of the scaffold (the BG cells) reside in the Purkinje layer, one layer below the ML. This is also important from a subcellular targeting standpoint since the Bergmann Glial fibers are thought to direct stellate cell axons to Purkinje Cell distal dendrites[25].

1.4.1.4. Coupling neuronal migration and axonal targeting

The prior examples of axon targeting all display how axon growth is used as a means of decreasing proximity between target and axonal growth cone. These models all accept that the presynaptic cell soma is settled in place at the time of axonal targeting. Rakic, 1971, proposed an alternate model for neurons to solve the pre-contact proximity question based on the development of cerebellar granule cells, namely that axonal targeting of specific lamina is accomplished by neurons leaving a trailing nascent axon during migration through the target region[174]. A similar model was recently proposed by Lim et al., to explain translaminar targeting of cortical interneurons like Sst+ Martinotti cells[254]. Since the latter is more relevant to this work, we will discuss it here. As mentioned above, cortical interneurons are born in subcortical regions and undergo a long tangential migration into the cortex to acquire regional positioning, then undergo a switch to radial migration that determines their laminar positioning. Cortical interneurons tend to migrate tangentially along two highly stereotyped streams, a superficial one within the marginal zone near the cortical surface, and a deeper one in the subventricular zone near the ventricle (see diagram in Figure 2). Lim et al., showed that L1-targeting Martinotti cells preferentially migrate tangentially through the superficial route and leave a trailing nascent axon in the marginal zone upon switching to radial migration. Since the marginal zone develops into L1, the Martinotti cell's axonal growth

cone is always within the target layer, even as the soma continues migrating into deeper layers. Subsequent axonal arborization and synaptic targeting within the nascent L1 then determines the ultimate connectivity of these cells within this layer. This model suggests that the primary guidance molecules dictating Martinotti cell targeting of distal dendritic tuft are actually those used during tangential migration, with those used during axon guidance itself providing finer specificity.

1.4.1.5. Summary of pre-contact methods

The process of synaptic development involves several key principles. Initially, it's about spatial proximity between presynaptic and postsynaptic cells, ensuring their physical interaction. Neurons undergo migration, polarization, and neurite elaboration, specifying axon and dendrite as distinct compartments. Axonal and dendritic pathfinding involves molecular interactions with extracellular signals, guiding growth cones towards appropriate targets. Diffusible molecules like Netrins and semaphorins play roles in regional and subcellular targeting, influencing neurite extension through attractive or repulsive cues. Adhesion molecules, such as Ephrins and Ephs, also guide axon targeting by mediating regional and topographic map formation. Cellular scaffolds, like Bergman Glial fibers, can direct axon growth towards specific targets. Neuronal migration can also contribute to axon targeting, with some neurons leaving a trailing axon during migration, ensuring their axon growth cone remains within the target layer. These interactions with the extracellular environment collectively ensure the precise assembly of neural circuits by regulating migration, axon guidance, and subcellular compartmentalization.

1.4.2. Contact and beyond

Synaptic target recognition is the process of selecting a specific postsynaptic partner out of all possible partners. This specificity can take the form of stereotyped regional, laminar, cellular, or

subcellular connectivity. These forms of connectivity are not mutually exclusive: specific neurons can show preferential targeting not only of specific cell types, but also specific subcellular regions on those targeted cells[14, 27, 169, 257, 258]. As noted above, this process is initiated by a physical interaction between the presynaptic cell and its putative partner. Generally, the cellular components are the axonal growth cone on the presynaptic cell and the target compartment of the postsynaptic cell. There are instances, however, where contact is initiated by a dendritic growth cone interacting with a presynaptic cell axon (reviewed in [163]). Although most synapses are axo-dendritic, there are many instances of axo-somatic and axo-axonic synapses. These classes can be broken down into even finer subclasses. Different cells will preferentially form axo-dendritic synapses onto different subdendritic regions, such as dendritic shafts, spines or sub-spine regions; onto different branches, such as basal vs apical dendrites; and at different distances from the target cell soma, such as proximal, medial, or distal dendrites, and primary vs oblique, and tuft[6, 259]. There are also distinct targeting patterns among cells that form axo-somatic synapses, with distinct enrichment patterns by PV basket cells[260, 261], CCK basket cells[174, 262], or mossy fibers targeting of the pyramidal cell soma, leading to giant synapses[16]. Axo-axonic connections include synapses onto the axon initial segment[263, 264], nodes of ranvier[265-267], and even synapses onto other presynaptic boutons[268]. Finally, although rare, there are also instances of dendro-dendritic synapses[269]. These connections are all mediated by adhesion molecules expressed by the presynaptic cell and postsynaptic cells.

In this section, we will first discuss the synapse and its components: the presynapse and postsynapse. We will discuss how these highly specialized subcompartments exhibit diversity, focusing primarily on GABAergic synapses. We then cover general principles of synaptic development, after the presynaptic and postsynaptic cells come into contact. Next, we discuss synaptic recognition and synaptogenesis, the de novo formation of synapses. We also cover current models

for how cell-cell contact initiates distinct synaptic identities. This background will set us up for the next section, wherein we discuss models for how subcellular targeting specificity can arise at the point of contact and beyond.

1.4.2.1. The Synapse

The chemical synapse is comprised of two asymmetric subcellular compartments from distinct cells tethered by intercellular adhesion molecules and separated by a 15-20 nm space (Figure 1.21). The presynaptic compartment is specialized for the release of neurotransmitter-containing vesicles, whereas the postsynaptic compartment is specialized for reception of this signal via membrane-bound, neurotransmitter-activated ion channels. As a general rule, presynaptic specializations are composed of canonical presynaptic components with minimal variation, whereas postsynaptic specializations are highly diverse, with components determined by the neurotransmitter receptor and little overlapping between different types[152, 268]. In order for synaptic transmission to be efficient,

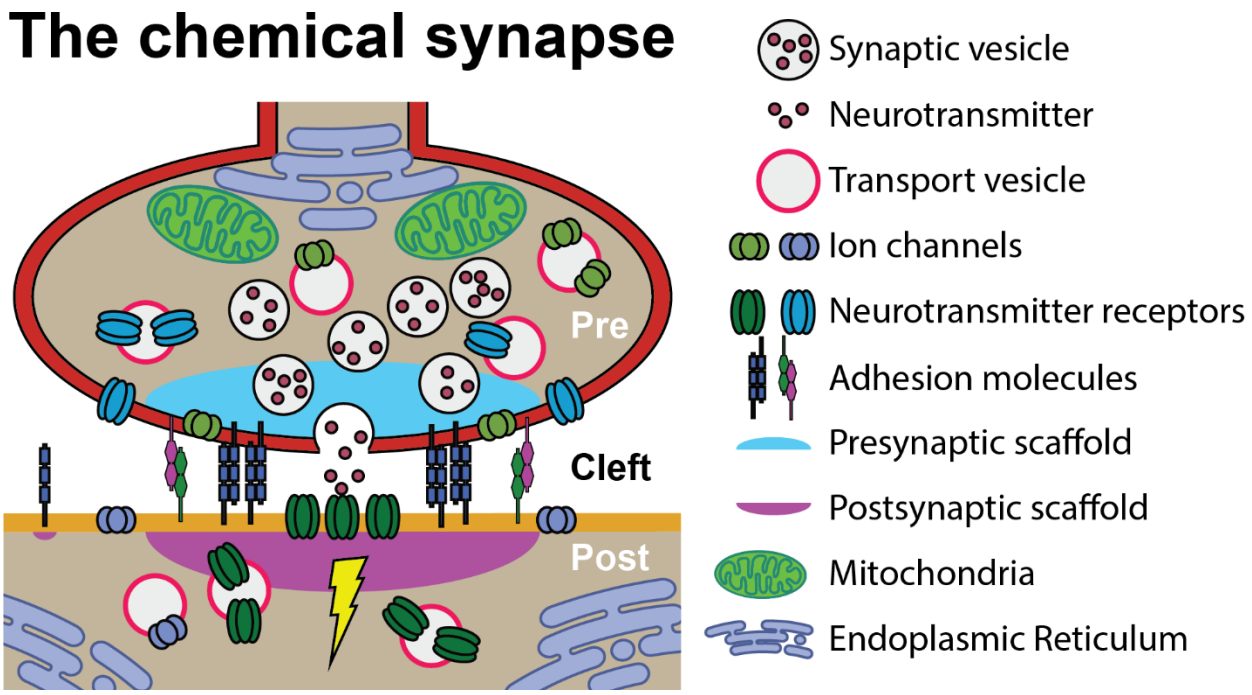

Figure 1.21 | Schematic of the chemical synapse.
A simplified schematic of the chemical synapse, showing important structural components. See text for details.

reliable, and rapid, neurotransmitter release sites on the presynaptic membrane must be precisely aligned with ionotropic receptors on the postsynaptic membrane. This has led to the hypothesis that synaptic adhesion molecules act as synaptic organizing molecules, whereby the trans-synaptic interaction initiates anterograde and retrograde localized signal transduction within the cytoplasm, thereby recruiting synaptic molecules to the contact sites[26]. Despite this appealing model, little is known about the cytoplasmic signals transduced by transsynaptic adhesion molecule interactions.

1.4.2.2. The Presynapse

The presynapse is the functional unit for anterograde transmission of synaptic signaling[270, 271]. Presynapses in the central nervous system are typically small, hemi-spherical shaped objects resembling buttons, hence the name 'bouton.' The presynapse is enriched for synaptic vesicles, which contain neurotransmitter, vesicle transport and release machinery, presynaptic scaffolding proteins, calcium buffering molecules, ion channels, neurotransmitter reuptake receptors, and synaptic adhesion molecules. Also enriched are ancillary support organelles such as mitochondria, endoplasmic reticulum, proteasomes, and lysosomes. Together, these molecules form a microenvironment that specializes in activity-dependent release of neurotransmitter-containing vesicles. As outlined below, presynaptic sites are specified by successful synaptic targeting and initiation of synaptogenesis. Cell-cell contact with a post-synaptic partner initiates retrograde signaling to induce recruitment of these specialized presynaptic molecules, leading to morphological and molecular profiles of presynaptic boutons, as outlined below. Although many of the developmental processes underlying presynapse formation have been characterized, how transsynaptic communication translates to cytoplasmic signaling that leads to recruitment of these components is not known.

Presynaptic diversity is evident in distinct morphological, molecular, and physiological profiles[268, 272-274]. As a general rule, all presynapses utilize the same basic cellular pathways, with slight variation in molecular components. All presynapses tend to use the same generalized release machinery to secrete neurotransmitter, regardless of which neurotransmitter is used[275]. The synaptic vesicle protein synaptophysin and the scaffolding molecules bassoon and piccolo, for instance, are expressed in most if not all presynapses of the CNS[276, 277]. Within the vesicle release machinery, the major source of variation is in the genetic isoforms utilized, which, as discussed below, will have functional consequences. Molecularly, the major source of presynaptic diversity derives from the neurotransmitter and its associated enzymes and transporters, which are expressed as a package deal. Specific neurotransmitter enzymes will be accompanied by associated transporters and reuptake receptors. For instance, GABAergic cells will express the GABA synthetic enzymes GAD65 and GAD67, as well as the vesicular GABA transporter VGAT for release, and the GABA transporter GAT1 responsible for reuptake. In contrast, glutamatergic cells express vesicular glutamate transporters (VGLUTs) for release and excitatory amino acid transporters (EAATs) for reuptake. Even when individual neurons are able to synthesize multiple neurotransmitters, these neurotransmitters are sorted into distinct presynaptic vesicles[278, 279] and presynaptic compartments[268, 277, 280] that are associated with distinct postsynaptic targets[262] (reviewed in [14]).

Morphologically, presynapses can vary widely in size and shape. In the CNS, presynaptic boutons are typically small and round. There are notable exceptions, however, such as mossy fiber synapses that connect dentate gyrus and CA3 neurons, which are 50-100 times larger than a typical asymmetric synapse, have a convoluted shape more reminiscent of the NMJ, and can contain over 30 separate vesicle release sites[281]. There is also size and shape diversity within the population of 'typical' CNS synapses, with distinct cell types displaying distinct synapse sizes[282]. Furthermore,

such diversity is evident even within the same cell, with differences in bouton size reflecting their vesicle number, physiological strength, and reliability[283-286]. Indeed, convincing arguments have been made that each individual synapse is unique, with its specific molecular composition and shape being a function of the presynaptic and postsynaptic cell types, its local environment, developmental stage, and history of activity[272].

Moreover, there is a wide body of literature on how specific molecular profiles lead to distinct functional properties[287, 288]. For example, presynapses of fast-spiking cells, such as PV basket cells, utilize synaptotagmins to mediate binding of synaptic vesicle to the active zone membrane, thereby enabling neurotransmitter release into the synaptic cleft. This binding is calcium-mediated: calcium binding leads to a conformation change in synaptotagmin that initiates vesicle binding to the membrane. Unlike other vesicle release proteins, synaptotagmins have particularly fast calcium binding and release kinetics, which enables rapid cycling between vesicle binding and release[272, 287, 289-294]. Fast-spiking basket cells also express the calcium buffering protein parvalbumin (PV) in their presynapses. Like synaptotagmins, PV has calcium release kinetics that enable rapid changes in calcium concentrations, thereby allowing for the fast-spiking profile characteristic of these cells[33, 295]. Combined, these evolutionarily conserved molecular mechanisms and expression patterns enable these cells to communicate rapidly with their postsynaptic partners. Other neuron types have evolved to express different molecular expression modules that result in distinct functional properties.

1.4.2.3. The Postsynapse

Postsynaptic specializations are highly diverse and specialized, with the utilized molecules closely coupled to the presynaptic neurotransmitter and their postsynaptic receptors. Very few components are shared between postsynaptic specializations for different neurotransmitters.

Postsynaptic specializations consist of elaborate, multimolecular complexes. Proximity ligation assays targeting inhibitory postsynaptic proteins revealed 181 proteins specifically enriched at inhibitory PSDs, greatly expanding the known proteins at the inhibitory PSD (Figure 1.22)[35]. These included trafficking proteins, cytoskeletal regulatory proteins, integral membrane proteins, and protein kinases/phosphatases. Many of these gene products have been implicated in neurological disorders. Of note, this network represents a combined average of all inhibitory synapses in the sample, and thus masks any diversity between these synapses.

The defining characteristic of postsynapses is the neurotransmitter receptor, which are modularly associated with other postsynaptic components such as scaffolding molecules, cytoskeletal

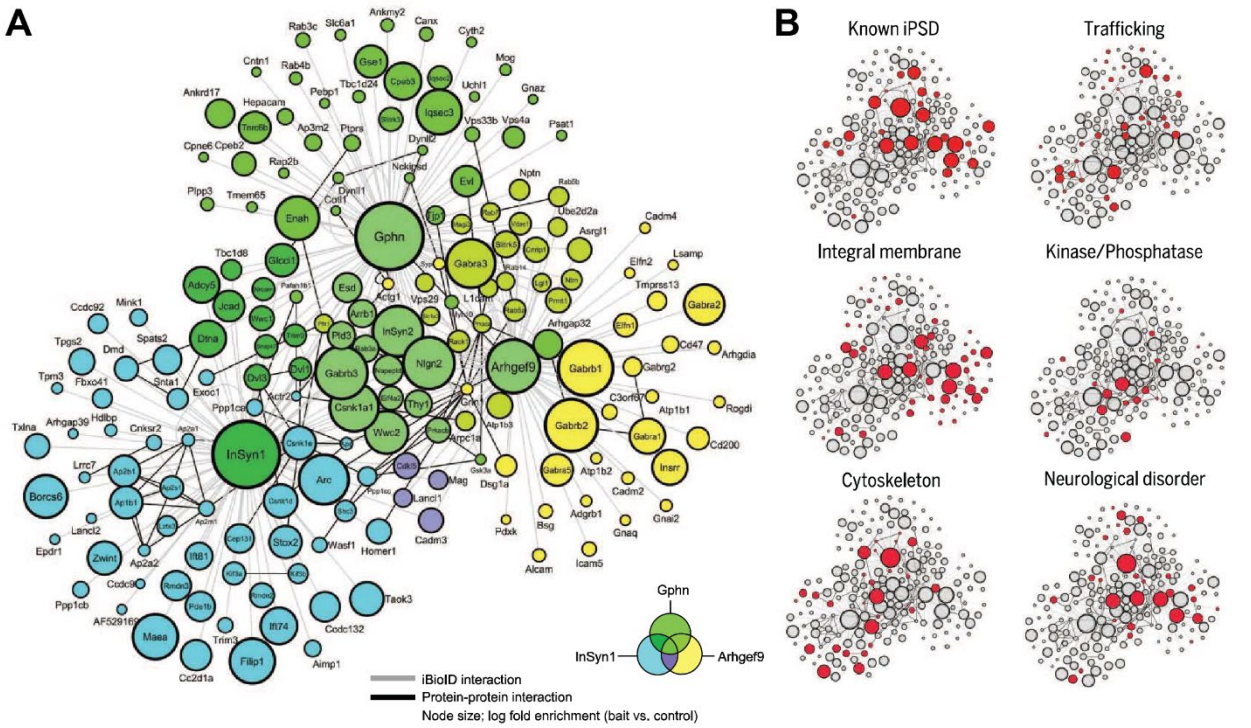

Figure 1.22 | Scale-free graph of the iPSD proteome.
A) InSyn1 (*blue*), gephyrin (*green*), and arhgef9 (*yellow*) BirA-dependent iBioID identify a rich network of known and previously unknown proteins enriched at the iPSD. Node titles correspond to gene name; size represents fold-enrichment over negative control. Edges are shaded according to the types of interactions (*gray*, iBioID; *black*, protein-protein interactions previously reported). **B)** Clustergram topology of iPSD proteins (*red*) in selected functional categories. Figure and legend from Uezu et al., 2017 . Reprinted with permission from AAAS.

adaptor proteins, kinases/phosphatases, and trafficking molecules. GABA has two receptor classes: $GABA_A$ receptors, the fast, synaptic receptors driven by neurotransmitter, and $GABA_B$ receptors, slow, metabotropic receptors where activation can persist for seconds and important for neuromodulation. We will only discuss $GABA_A$ receptors, which are hetero-pentamers consisting of distinct combinations of α, β, and γ subunits. The most common conformation is two αs, two βs, and one γ subunits. There are many subunits of each type, α1–6, β1–3, γ1–3. $GABA_A$Rs can be synaptic or extrasynaptic, with those containing α1–3 and γ2 subunits tending to be synaptic. Different cell types express distinct profiles of GABAAR subunits[37, 296-298] and expression patterns can differ by disease[298-300]. Different subunits confer distinct pharmacological profiles[301, 302], and tend to be enriched at different synapses and at different subcellular locations, though these localization patterns tend to differ by brain region[299]. α1-3 are all found at perisomatic synapses, but α1 mediates transmission from PV basket cells, whereas α2 and α3 mediate inputs from CCK basket cells[301]. α2 and α3 are also localized to the AIS and mediate chandelier synapses[303, 304], whereas pharmacological data suggests that α5 mediates synapses from Martinotti cell, which target distal dendritic tuft[232, 234]. α4 and α6 on the other hand, tend to be localized to extrasynaptic $GABA_A$ receptors. This diversity in $GABA_A$ receptor composition is contrasted with the observation that most if not all GABAergic postsynapses contain the scaffolding protein gephyrin.

1.4.2.4. Synaptic targeting recognition: selecting your partner from the crowd

Synaptic target recognition[11] is the process of selecting a specific synaptic partner out of all possible partners. It is thus primarily a process of cellular recognition. Strictly speaking, synaptic

[11] We use the term 'synaptic target recognition' or 'target recognition' here rather than the standard term 'synaptic targeting' to distinguish the molecular process of synaptic partner recognition/selection from 'targeting specificity,' the specific connections that arise through any of the pre- and/ or post-contact mechanisms discussed in section 1.4.

target recognition is mediated by cell adhesion molecules (CAMs) that are *a priori* localized at the cellular contact sites and interacting *in trans*. In other words, this stage of neuronal development involves a physical interaction between CAMs on the presynaptic cell and those of its putative partner.

Since partner selection is primarily a process of cellular recognition, the first step in this process can be boiled down to a question of molecular complementarity: these molecules can interact in an attractive or repulsive manner, thereby initiating downstream molecular cascades that increase or decrease the probability of synapse formation[305, 306]. It is important to note that molecular affinity is distinct from the interaction being attractive or repulsive[307]. Attractive and repulsive interactions both arise from high affinity binding; it is the cellular and molecular response that determines whether the interaction is attractive or repulsive. For repulsive interactions, molecular contact initiates downstream mechanisms that destabilize the interaction and lead to its collapse, thereby hindering synapse formation[29], whereas in attractive interactions the affinity either remains constant or is further increased, leading to a stable interaction that is conducive to synapse formation.

The simplest transsynaptic interaction involves an interaction between two molecules, one at the presynaptic membrane and the other on the postsynaptic membrane, interacting in trans (opposite cell surface). These interactions can be homophilic (same protein)[305, 308, 309], or heterophilic (different proteins) [238, 310-312]. Many transsynaptic interactions are mediated by multi-protein complexes, however, utilizing *cis* (same cell surface) interactions with other membrane-bound proteins[175, 313-315] and/or interacting with secreted adaptor protein intermediates[16, 175, 230, 305, 316] that stabilize the transsynaptic interaction. Finally, evolutionarily, most cell adhesion molecules evolved from two related protein folds, namely cadherin and immunoglobulin (Ig)-like domains,

and thus belong to one of these superfamilies[317]. Different neuron types express distinct combinations of synaptic adhesion molecules and form highly specific connections[34, 35, 37, 318, 319]. This observation has led to the proposition that there is a synaptic adhesion molecule code that underlies synaptic specificity[12] (Figure 1.23)[29, 305, 320]. Many families of adhesion molecules have been localized to the synapse (Figure 1.23, upper panel)[239], each of which comprise multiple members with distinct binding affinities. We show here the affinities for (*mammalian*) clustered protocadherins (Pcdh), (*drosophila*) DIPs & DPRs, and (*mammalian*) Type II cadherin families (Figure 1.23, lower panel)[237], all of which are strongly implicated in synaptic target recognition[238, 262, 307, 310, 319]. Indeed, that a synaptic adhesion molecule code underlies synaptic specificity has been partly validated by studies showing that interactions between specific synaptic adhesion molecules mediate specific instances of stereotyped cellular[16, 321] and subcellular[322-325] targeting. That said, data suggests that most of these adhesion molecules are not involved in target recognition. Complete deletion of presynaptic adhesion molecule families that form interaction hubs, like Neurexins[326] and PTPRs[239, 327] led to diverse, circuit-specific phenotypes that generally did not accord with these proteins (and by extension, their postsynaptic partners) playing an essential role in target recognition, or even a canonical role in synapse function. Relatively few of these protein

[12] The synaptic adhesion molecule code is a reformulation of Sperry's Chemoaffinity hypothesis that specifically focuses on the role of synaptic adhesion molecules in synaptic target recognition. Sperry's chemoaffinity hypothesis is more general, proposing that neurons form specific connections with their targets based on interactions with specific, genetically-encoded molecular cues. It does not make any claims as to the form of these molecular cues, i.e., whether they are adhesion molecules or diffusely secreted molecules, or the developmental stage within which they operate, i.e., it includes both target recognition and axon pathfinding. The latter has historically been the primary focus of research associated with the chemoaffinity hypothesis, and thus many erroneously equate it with axon pathfinding.

families have a canonical role in synaptic target recognition: the (*drosophila*) DIPs-DPRs[328], cadherins[310], Sidekicks[321, 329], L1CAMs[322], Latrophilin-Teneurin-FLRTs[232, 233, 323, 330-333], and (membrane-bound) Semaphorins-Plexins[234, 235, 334, 335]. Of these, only the Sidekicks, Latrophilins,

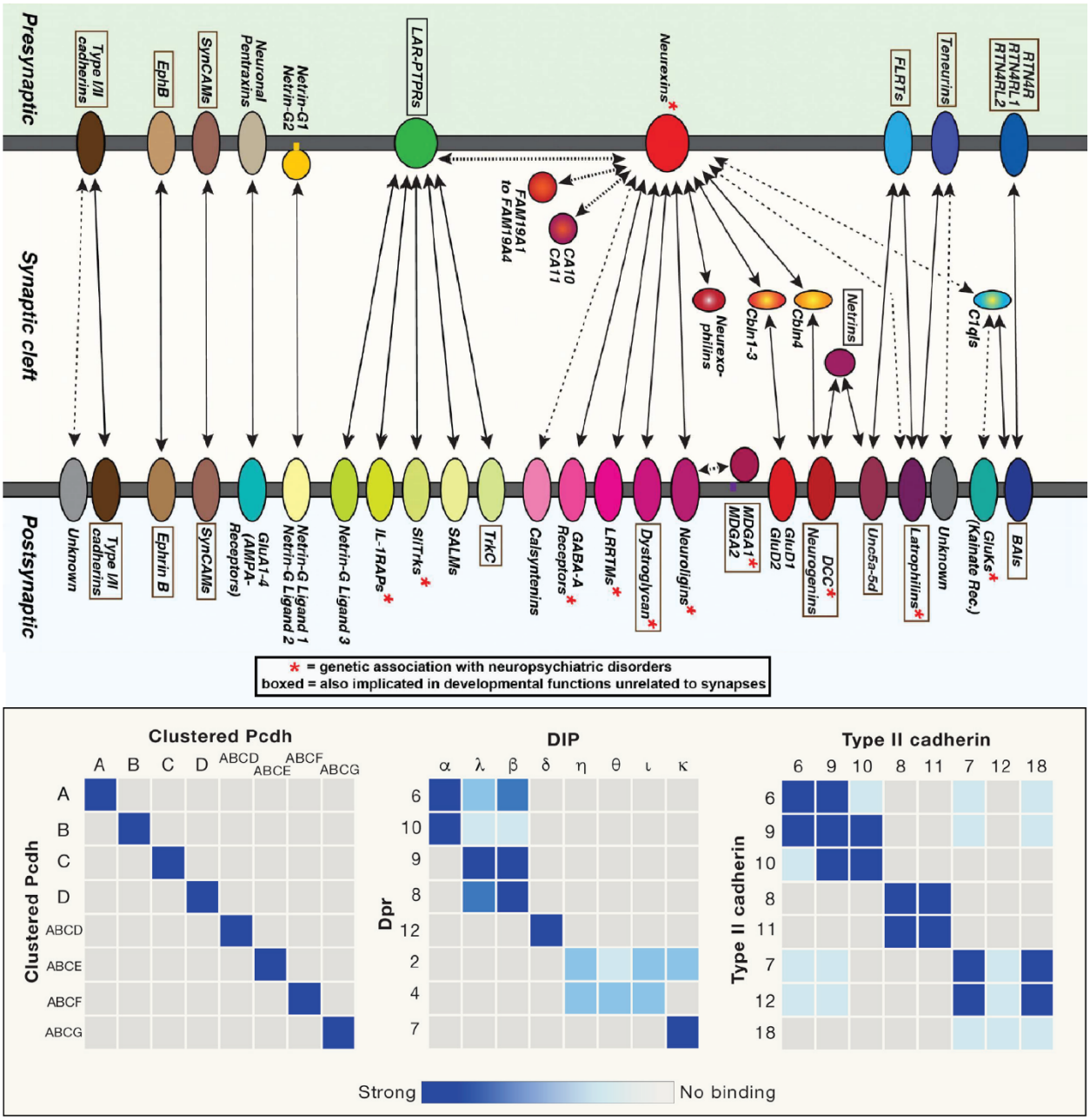

Figure 1.23 | Visual representation of the synaptic adhesion molecule code.
Upper: schematic diagram of candidate trans-synaptic synaptic adhesion molecule (SAM) complexes governing synapse assembly. Note that two families of presynaptic SAMs, neurexins and LAR-PTPRs, are hub molecules that interact with a series of postsynaptic SAM families and also

bind to each other in cis. Most candidate SAMs perform additional functions outside of synapses. Lines and arrows indicate interactions, with cis-interactions shown as dotted lines and less validated trans-interactions shown as dashed lines. *Lower*: Differential binding affinities of different adhesion molecule families. The clustered protocadherin (Pcdh) family displays highly specific homophilic interations. In contrast, two other SAM families display degenerate binding properties. DIP/Dpr interactions are largely heterophilic and often promiscuous. Type II cadherins are homophilic, but also exhibit heterophilic binding specificity. Abbreviations: LAR-PTPRs, LAR-type receptor phosphotyrosine phosphatases; DCC, deleted in colorectal cancer; EphB, Ephrin B; FLRT, fibronectin leucine-rich transmembrane; LRRTM, leucine-rich repeat transmembrane; Rec., receptor; RTN, reticulon; SALMs, synaptic adhesion-like molecules; SliTrks, Slit- and Trklike proteins; SynCAM, synaptic cell adhesion molecule; TrkC, tropomyosin receptor kinase C. Figures and legends adapted from Sudhof 2021 (upper) and Sanes and Zipursky, 2020 (lower).

and Semaphorins/Plexins are enriched in specific subcellular target compartments. This is not to say that none of the other families are involved in target recognition. Rather, although many of their members have been implicated in synaptic target recognition, accumulating evidence suggests that their utilization in this process is context dependent. Instead, most synaptic adhesion molecules are involved in shaping synaptic properties. We discuss specific roles for many of these proteins below. One thing to keep in mind, however, is that our methods for assessing functional changes are incredibly sensitive and nuanced, allowing us to easily detect changes. On the other hand, our methods to assess specific connections are comparatively immature, primarily consisting of counting synapse numbers using labeling techniques that often do not distinguish distinct connections. Thus, the lack of evidence of these gene families having a role in target recognition could arise from sensitivity issues in our detection methods, and generally unnuanced schema for categorizing synapses. A complete mapping of these interactions will require not only *in vitro* biochemical experiments using isolated proteins, but also *in vivo* validation through imaging and functional analysis.

In addition to these interactions being attractive or repulsive, *synaptic* target recognition implies that the intercellular interaction is synaptogenic, i.e., that the cell-cell contact is accompanied

by cellular cascades that promote synapse formation and/or stabilize the synapse. Indeed, this property distinguishes synaptic partner selection from other CAM-mediated cell recognition interactions like cellular migration and axon guidance. Generally speaking, it is still unclear which of the candidate synaptic adhesion molecules are primarily involved in cellular recognition, synapse assembly, shaping synaptic properties, or some combination of these functions. Although target recognition and synaptogenic signaling can be tackled by the same molecule(s), it is important to note that this is not necessarily the case. As discussed below, cell recognition and sustained contact can be established by one set of molecules, whereas the morphological and molecular changes associated with synapse formation can be driven by an entirely different set of molecules along a parallel path[26, 336]. Thus, although these processes can be linked, it is conceptually better not to conflate them. Synaptic target recognition occurs through transcellular intermolecular interactions between adhesion molecules expressed on the cell surface. It is thus primarily an extracellular process. In contrast, the assembly of synapses is mediated by cis interactions *on the membrane* or cytoplasmic signaling *inside the membrane*. Synaptic target recognition is a necessary precursor that potentiates synaptogenesis. Next, we will discuss synaptogenesis.

1.4.2.5. Synaptogenesis: building something together

Synaptogenesis – the stage of synapse assembly and plasticity – is initiated after the cell-cell contact of successful partner recognition (reviewed by [320, 337-339]). During synaptogenesis, synapses can follow a trajectory of growth, maintenance, or elimination. As long as the cell-cell contact is maintained, the pre- and post-synaptic cells engage in considerable transsynaptic cross-talk. This transsynaptic signaling occurs in both the anterograde and retrograde directions, and leads to reorganization of the pre- and post-synaptic machinery[318]. These changes can be activity-dependent (synaptic plasticity) or activity-independent (synaptic assembly). And in all cases, these changes

can lead to specific molecular and functional properties that define and distinguish one synapse from another. Importantly, this transsynaptic cross-talk can impact synaptic connectivity.

The initial stage of synaptogenesis is synapse assembly, whereupon a number of molecular programs are initiated via transsynaptic signaling in both the pre- and postsynaptic sides to stabilize (or destabilize) the synapse (reviewed in [268, 340]). Initial stages include recruitment of molecules, complexes, and organelles that will define the characteristics of the synapse. The specific molecules that are recruited are determined not only by the identities of the pre- and postsynaptic cells, but also by the cross-talk between them. On the presynaptic side this includes additional adhesion molecules, scaffolding molecules, ion channels, the presynaptic neurotransmitter release machinery, and organelles such as mitochondria, endoplasmic reticulum, proteasomes, and components for local translation[13][181, 268, 318, 320, 341-343]. These molecular changes are accompanied by morphological changes and subcellular compartmentalization in the form of presynaptic bouton formation. On the postsynaptic side, the transsynaptic cross-talk leads to the recruitment of neurotransmitter receptors, scaffolding proteins, organelles, and signal transduction complexes[344]. Of note, the postsynaptically recruited molecules are specifically matched to the presynaptic neurotransmitter, such as GABAergic receptors and the GABA/glycinergic scaffolding molecule gephyrin, to the postsynaptic site across from GABAergic presynapses. Moreover, these molecules are recruited even in the absence of synaptic transmission[318, 320], suggesting it can be mediated by synaptic adhesion molecules.

In order for synaptic transmission to function efficiently, reliably, and rapidly, neurotransmitter release sites on the presynaptic membrane must be precisely aligned with postsynaptic iono-

[13] Local translation in axons is controversial. The evidence for this process is reviewed by [157].

tropic receptors on the postsynaptic membrane. This has led to the hypothesis that synaptic adhesion molecules act as synaptic organizing molecules, whereby the trans-synaptic interaction initiates anterograde and retrograde localized signal transduction within the cytoplasm, thereby recruiting synaptic molecules to the contact site[339]. This hypothesis extends the synaptic adhesion molecule code model beyond simply cellular recognition, proposing in addition that specific transsynaptic interactions also assemble synapses and define specific synaptic properties[320]. Of note, transsynaptic interactions that mediate synaptic target recognition need not be involved in synaptic assembly or shaping synaptic properties, and vice versa. The intracellular signal transduction cascades that mediate synaptic assembly and shape synaptic properties not well understood.

1.4.2.5.1. Specification of presynaptic domains by retrograde transsynaptic signaling

Retrograde transsynaptic signaling is from the postsynaptic cell to the presynaptic cell. Retrograde signaling can be mediated by postsynaptic adhesion molecules or secreted molecules. The secreted factors, in turn, can be divided into molecules that diffuse freely within the synaptic cleft, and those that constitute the synaptic extracellular matrix and serve as extracellular scaffolding molecules. Retrograde signaling can be activity dependent, i.e., in response to presynaptic neurotransmitter release, or activity independent. During synapse assembly, this signaling leads to morphological and molecular changes at the presynaptic site. Here, the critical event in presynapse assembly is signaling across the presynaptic membrane, which in turn recruits neurotransmitter release machinery, presynaptic adhesion molecules, calcium channels, and receptors, and anchors the presynaptic scaffolding molecules to the membrane at the active zone.

On the presynaptic side, active zone scaffolding molecules mediate the vesicle release process, are required for functional synapses, and are a key indicator for the assembly of presynapses[341]. Interestingly, none of the principle active zone scaffolding molecules are transmembrane proteins,

indicating that recruitment of these molecules to the membrane opposite postsynaptic sites must be mediated by other molecules. No protein interaction or protein family has been found to be essential for membrane anchoring of principle active zone scaffolding proteins. Rather, heterologous expression assays have demonstrated that many synaptic adhesion molecules are sufficient to induce presynaptic assembly[345]. The most parsimonious model to explain these data is that recruitment and tethering of active zone scaffolding molecules is mediated by many redundant interactions[158]. These interactions could include direct interactions of scaffolding machinery with not only synaptic adhesion molecules, but also Ca2+ channels or membrane lipids. One intriguing model is liquid-liquid phase separation[341, 346, 347]: at sufficient concentrations, presynaptic active zone proteins RIM and RIM-BP undergo liquid-liquid phase separation, spontaneously forming membraneless subcellular compartments through low affinity interactions[348]. A similar condensate of synapsin and synaptophysin has been implicated in synaptic vesicle formation, packing, and motility[151, 213, 349], thereby allowing for the mesoscale domains of synaptic vesicles at synapses *in vivo*. We discuss condensates in greater depth below.

1.4.2.5.2. Specification of postsynaptic domains by anterograde transsynaptic signaling

Anterograde signaling is from the presynaptic cell to the postsynaptic cell. Anterograde synaptic signaling is mediated by presynaptic adhesion molecules and secreted molecules. The secreted factors, in turn, can be divided into molecules that diffuse freely within the synaptic cleft, and those that constitute the synaptic extracellular matrix and serve as scaffolding molecules at the synaptic cleft. Anterograde signaling can be activity dependent, via neurotransmitter release, or activity independent. During synapse assembly, this signaling leads to morphological and molecular changes at the postsynaptic site. Here, the critical event in postsynapse assembly is signaling across the postsynaptic membrane, which in turn recruits neurotransmitter receptors and anchors postsynaptic scaffolding molecules to the membrane at the signaling site. There are several models

for how this is achieved, including secreted molecules that diffuse across the synaptic cleft, and adhesion molecules that form transsynaptic bridges.

Presynaptically-secreted molecules can induce postsynaptic nucleation and compartmentalization. Perhaps the best characterized example of this model is Agrin-induced acetylcholine receptor (AChR) clustering at the neuromuscular junction (NMJ). Agrin is normally secreted by the presynaptic cell and interacts with its postsynaptic receptor Lrp4 and MuSK, leading to nucleation and AChR clustering at the postsynapse[350]. Importantly, this compartmentalization is independent of transsynaptic adhesion molecules: local, pipette-based application of Agrin to muscles induces nucleation and AChR recruitment. A similar role was proposed for pentraxins in the CNS, which cluster postsynaptic glutamate receptors and act as a transsynaptic bridge to presynaptic pentraxin receptor (reviewed in [350]). Similarly, growth factors[351, 352] and neurotransmitters[353, 354] have been shown to induce nucleation and receptor clustering independent of presynaptic specializations, indicating that these secreted molecules are sufficient for postsynapse assembly. It is important to note that, in a physiological context, these secreted molecules are typically found at concentrations high enough to induce such nucleation events only when secreted by presynaptic specializations or growth cones. Thus, although the physiological importance of these mechanisms outside of transsynaptic communication is unclear, these examples illustrate important adhesion molecule independent mechanisms of compartmentalization during synaptogenesis. Neurotransmitter receptors are also prime candidates in postsynapse assembly since these molecules interact directly with both scaffolding molecules and synaptic adhesion molecules. The cytoplasmic loop of several $GABA_A$ receptor subunits, for instance, binds to gephyrin, thus providing a molecular bridge between the cytoplasmic scaffolding molecule and the membrane-bound receptors[354]. This interaction is incredibly weak, however, and therefore has been discounted as a mechanism that

accounts for scaffold recruitment. Another role for neurotransmitter receptors in organizing synapses is via activity-dependent lipidation of postsynaptic components. Specifically, palmitoylation, a posttranslational addition of lipid, which is known to occur for gephyrin[355, 356] PSD95[356], and $GABA_A$ receptors[357, 358], has been shown to be an important mechanism for plasma membrane targeting and clustering. Lipid modification of gephyrin controls whether it is membrane-bound or soluble, and thus also its localization and clustering. Palmitoylation/depalmitoylation cycles are regulated by synaptic activity. Although interfering with palmitoylation decreased clustering, it did not abolish the process, suggesting redundant and independent mechanisms for membrane association and clustering, and that palmitoylation is associated with synaptic plasticity rather than assembly. In addition, synaptic connectivity appears to develop normally in the absence of neurotransmission, indicating that these mechanisms are not required for postsynapse assembly.

Synaptic adhesion molecules have also been proposed as central players in postsynapse assembly. The postsynaptic neuroligins are a prime candidate for mediating synaptic specific, especially given that their presynaptic counterparts, the neurexins, exhibit hundreds of isoforms due to alternative splicing[357], which in turn have cell-specific and circuit-specific expression profiles[273, 318, 359, 360]. Neuroligin-1 is exclusively found in glutamatergic synapses, whereas neuroligin-2 is exclusive to GABAergic synapses, and neuroligins -3 and -4 are found in both[358, 361-364]. Neuroligin-2 is the only transmembrane cell adhesion molecule identified that interacts with gephyrin[364]. It also interacts with collybistin[302, 365], another GABA-specific postsynaptic scaffolding molecule, and in cis with $GABA_A$ receptor subunits[324, 359]. These data all point to Neuroligin-2 as a principle actor in assembly of inhibitory synapses, but deletion studies show that it is not required for inhibitory synapse assembly, instead playing diverse roles in shaping synapses

across different inhibitory synapses[324]. The same is true for other neuroligins, namely that they are involved in shaping synapse properties, rather than orchestrating synapse assembly[268, 316, 318, 320]. Postsynaptic Adhesion GPCRs, namely the Latrophilin and BAI family members, are the only synaptic adhesion molecules consistently shown to be required for establishing synapses[332], as deletion of specific isoforms leads to severe decreases in specific subsets of subcellularly-targeted synapses. The BAI family consists of three poorly-characterized members, Bai1-3, with only Bai3 thus far having clear involvement in assembly of specific synapses. In the cerebellum, Bai3 is expressed postsynaptically by Purkinje cells and binds to C1ql1 expressed by climbing fiber cells. Bai3 deletion selectively blocks climbing fiber (but not parallel fiber) synapse formation[322]. Latrophilins also consist of three members, latrophilin1-3, and are better characterized. Latrophilin-1 is expressed postsynaptically in hippocampal neurons at both excitatory and inhibitory synapses throughout the cell, but deletion specifically decreased somatic inhibitory synapses[323], suggesting a role in the establishment of soma-targeting GABAergic synapses. Latrophilin-2 deletion in CA1 PyNs selectively decreases EC synapses on the distal dendritic tuft[331], whereas Latrophilin-3 deletion selectively decreases schaffer collateral inputs onto CA1 stratum oriens and stratum radiatum, corresponding to pyramidal cell basal and medial dendrites, respectively[345]. Latrophilin-mediated excitatory synapse assembly was initiated by forming a complex with presynaptic teneurin and FLRT counterparts. Of note, functional GPCR activity of latrophilin was required to rescue synapse formation in knockout animals[158], indicating that latrophilins activate synapse assembly through classic GPCR signaling. It is unclear whether the Bai family operates in a similar manner.

One attractive model that could encompass all the above findings is that subcompartmentalization of postsynaptic scaffold molecules may be driven by liquid-liquid phase separation[353, 366,

[367]. Such compartmentalization has been shown for both glutamatergic[333, 368] and GABAergic[369] postsynaptic scaffolding components. Of note, GABAergic and glutamatergic phase-transitioned scaffold condensates are mutually exclusive, form spontaneously past threshold concentrations, rely on multifarious weak interactions, and can have heterogenous molecular compositions. Indeed, a very recent study has linked postsynaptic latrophilins to recruitment of phase-transitioned postsynaptic scaffold condensates[144]. This mechanism would also explain the seemingly fluidic distribution of multiple separate irregular organizations of 2-dimensional $GABA_A$ receptor networks within a single synapse[343], as a rigid scaffold would lead to rigidly-organized receptors, which is not observed. As noted above, a similar mechanism has been proposed for presynaptic compartmentalization. Thus, rather than synaptic nucleation relying on one or a few strong interactions, this model proposes that the process could arise through many weak interactions. Importantly, this model suggests that multifarious transmembrane proteins all interacting with the same luminal scaffolds would only have to reach a specific localized threshold concentration before spontaneous condensate formation drives further clustering and stabilization. Parallelism in fluidic organization of the presynaptic active zone and the postsynaptic density would enable reciprocal interactions between stable condensates at the cytoplasmic face of both synaptic membranes, further stabilizing synaptic compartmentalization. Moreover, this configuration would enable rapid adjustments in alignment of vesicle release sites and postsynaptic receptors, thereby making neurotransmission more efficient. Such an organization strikes a balance between stability and flexibility, thereby enabling organizational and functional plasticity.

1.4.3. Post-contact models of synaptic specificity

The previously-discussed models of subcellular targeting specificity were all pre-contact methods for axon guidance, and thus deal with the axon navigating the environmental framework

to minimize the distance between itself and the target. But what happens when the axon reaches a potential target? The remaining models are all post-contact, and thus deal with target recognition, synaptogenesis, and synapse refinement. We begin with the domain model of subcellular targeting, which constitutes the most fundamental model for subcellular targeting specificity. We will then discuss regional regulation of synaptogenesis, followed by synapse elimination as a means for subcellular specificity. Again, we will be focusing on the three acting parties: the presynaptic cell, the target cell, and the environment. All of these models begin either at the point of contact between putative synaptic partners, or after contact has already been established. It is important to note that pre- and post-contact mechanisms are not mutually exclusive in determining the final, stereotyped connectivity. The developmental trajectory of a neural connection can harness both pre- and post-contact mechanisms to arrive at a specific, stereotyped connectivity. Moreover, the post-contact mechanisms detailed below are not mutually exclusive. Rather, different model connections will rely more or less heavily on different mechanisms for their ultimate specificity, thus providing fertile ground for the study of these different processes.

First, on the question of synaptic activity. The role of synaptic activity at different stages of neurodevelopment has been a topic of research for decades. Studies using knockout (KO) mice in which synaptic vesicle release is perturbed, and thus synaptic transmission in general is blocked, showed that the majority of synapses are formed and maintained even in the absence of neurotransmitter signaling[145, 320, 344, 351]. These studies indicate that axon guidance, synaptic targeting, and the initial stages of synaptogenesis are activity-independent. In other words, the initial form of synaptic connectivity appears to arise from 'hard wired' biological processes that are independent of individual experience[352]. This conclusion contrasts with studies showing that local appli-

cation of neurotransmitter can induce postsynaptic receptor clustering[214, 215], indicating that synaptic transmission is sufficient to induce the initial stages of synaptogenesis. Finally, electrical activity alone is sufficient to induce postsynaptic differentiation, as shown by DC electrical field stimulation[206, 370], indicating that clustering can occur through non-molecular stimuli. The physiological importance of the last two mechanisms is unclear, as neurotransmitter levels and sustained, focal electrical stimulation rarely reach sufficient levels to induce clustering outside of the synapse. Nevertheless, these studies reveal mechanisms whereby synaptic activity can play a role in synapse assembly, and thus may be important in shaping this initial connectivity. Regardless of its role in the initial connectivity, synaptic activity is important in refining connectivity after it is initially established, as discussed below.

Second, many of these molecules are recycled and reused in different developmental contexts. For instance, netrin-1, which we discussed as a diffusible mediator of axon guidance, also mediates neuronal migration[371], arborization[370, 372], and synaptogenesis[175, 223, 371]. The reuse of these molecules during distinct, often sequential, developmental stages confounds the effects of an intervention, often making it difficult to deconvolve the molecule's role at different developmental stages. Clever experimental designs have, in certain instances, sidestepped these issues and elucidated nuanced roles for molecules during different developmental stages. But most studies do not go to such lengths.

Third, even within the same developmental context, many molecules will perform distinct roles in different cellular contexts. As discussed above, Sema3A functions as a chemorepellent for pyramidal cell axons in the cerebral cortex[243], whereas in the cerebellum it acts as a chemoattractant for basket cell axons[174]. Thus, a molecule's function in a cellular process is often context-

dependent, relying not only on intrinsic properties, but also on its interacting partners and downstream signaling cascades.

We revisit the second and third points in the discussion.

1.4.3.1. Domain targeting of molecularly-enriched subcellular compartments

In the Cerebellum, cerebellar basket cells (CBCs) innervate Purkinje neurons (PkCs) soma and axon initial segment (AIS), forming a brush-like collection of synapses called a pinceau. This AIS targeting in particular is an example of synapses forming upon a subcellular gradient of locally-enriched adhesion molecules[373] (Figure 1.24). CBC targeting to the Purkinje cell AIS closely correlates with a subcellular gradient of the L1CAM family member neurofascin186 (NF186) adhesion molecule[174]. Notably, NF186 cytoplasmic domain contains Ankyrin-binding domains[374], and deletion of the AIS-localized scaffolding molecule Ankyrin G (AnkG) leads to NF186 diffusion into more apical regions of the PkC soma and more distal regions of the axon, with CBC boutons precisely following this more diffuse NF186 gradient. Further data showing that expression of a dominant-negative NF186 in PkCs decreased levels of the GABAergic presynaptic marker GAD65 at the AIS, suggesting that NF186 mediates CBC targeting of PkC

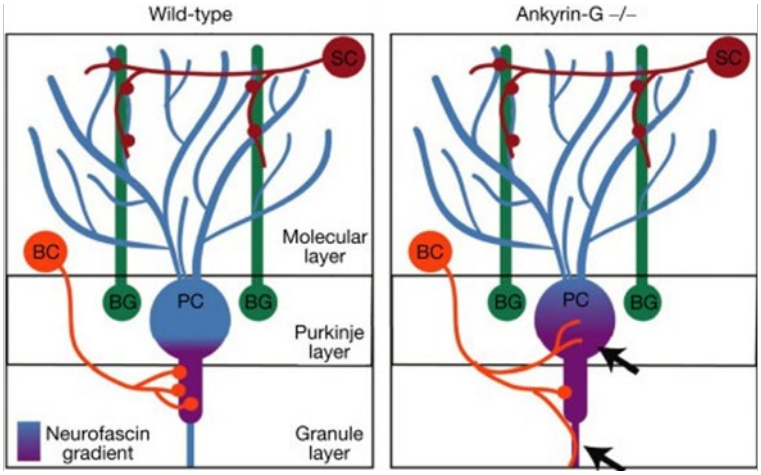

Figure 1.24 | Subcellular specificity of cerebellar basket cell inhibitory connections to Purkinje cells (PkCs). Schematics showing the inhibitory synaptic connectivity patterns of the wild type (*left*) and ankyrin-G deficient (*right*) mice. In wild type mice, a sharp gradient of neurofascin is present from AIS toward the soma. In ankyrin-G deficient mice, this gradient is no longer restricted to the AIS, which causes mistargeting of basket cell axons and reduced synapse formation. SC, stellate cell; BC, basket cell; PC, Purkinje cell; BG, Bergman glia. Figure and legend adapted from Williams and Ghosh, 2010.

AIS[374], though more recent data has challenged this conclusion[174, 175, 375]. These data led to a model whereby enrichment of attractive adhesion molecules within prespecified subcellular compartments of the target cell define the subcellular target region *a priori*, with the growth cone simply requiring guidance-mediated proximity to this region to initiate synapse formation[161, 172]. Given this model, the subcellular compartmentalization of postsynaptic adhesion molecules can occur cell intrinsically, as is the case with the AIS[16, 171, 200], or via inductive environmental influences, such as distal dendritic tuft in the cortex or SLM in the hippocampus[322]. This domain model has been shown in multiple settings, including the cortex[262], hippocampus[233, 266], retina[376], and spinal cord[232, 377].

Although the above model specifically proposes that attractive interactions drive domain-specific targeting, a more generalized model would include subcellularly-enriched repulsive cues that direct targeting away from specific domains, leading to synapse formation on domains that lack these cues. Indeed, this is precisely what occurs in the hippocampal circuit between dentate gyrus (DG) mossy fiber axons targeting CA3 pyramidal neurons (Figure 1.25). DG mossy fibers target CA3 pyramidal neurons (PyN) at proximal regions of the apical dendrite within the stratum lucosum (SL). Studies have shown that this targeting is based on interactions between semaphorins and their receptors, plexins. Unlike the previously-discussed semaphorins, Sema6A is a transmembrane protein expressed by the postsynaptic pyramidal neuron. PlexinA4, a Sema6A/Sema6B receptor, is expressed on the growth cone of mossy fiber afferents[232, 233, 235]. The PlexinA4-Sema6A interaction leads repulsion via collapse of the growth cone. Although Sema6A subcellular localization throughout the target CA3 pyramidal neuron, and therefore is present in all layers, its concentrations are highest within basal and apical dendrites, leaving a gradient expression pattern with increasing distance from the PyN soma. Given the expression pattern of this repulsive signal, one

might expect mossy fibers to target CA3 soma in the stratum pyramidalis (SP), but this is not the case. Instead, the postsynaptic CA3 PyNs express PlexinA2, which localizes to the PyN proximal apical dendrite within the SL and binds Sema6A in *cis*, thereby inhibiting its contact-mediated repulsive effects. This interaction decreases the effective Sema6A concentration within the SL to levels lower than the SP, thereby opening a window for mossy fibers to form connections at the stereotypical subcellular region, resulting in the characteristic innervation pattern of MFs on the SL. Deletion of PlexinA2 leads to the expected result, with mossy fibers targeting according to the Sema6A gradient and directing their axons to the SP. Of note, the PlexinA2 phenotype is rescued by Sema6A deletion, with the dual loss leading to a phenotype that is nearly indistinguishable from wild type[232, 233]. In contrast, deletion of presynaptic PlexinA4 leads to mossy fibers forming connections within all layers of CA3, as would be expected in the absence of a repulsive signal. This

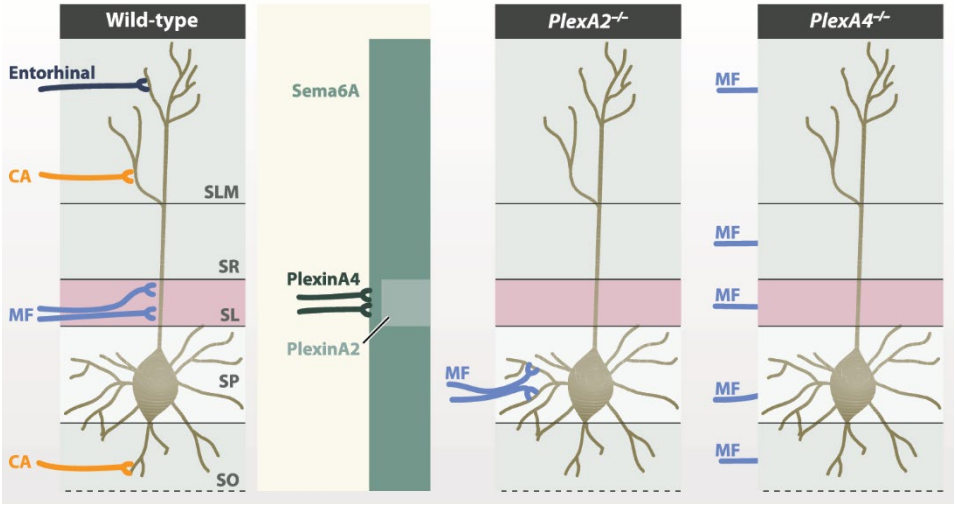

Figure 1.25 | Lamina-specific innervation of hippocampal CA3 neurons by mossy fibers (MF) is mediated by domain-restricted repulsive adhesion molecule interactions.
In CA3 axons from entorhinal cortex, commisural/associational afferents (CA), and dentate gyrus innervate distinct dendritic segments in the stratum lacunosum-moleculare (SLM), stratum radiatum (SR), stratum lucidum (SL), and stratum oriens (SO) as shown. Genetic studies suggest that it sends a repellent signal to MF axons that express its receptor, plexin-A4. Expression of another receptor, plexin-A2, in the SL attenuates this signal, however, allowing MFs to synapse on proximal dendrites. Figure and legend adapted from Sanes, 2009.

is an example whereby contact-mediated repulsion is coupled with domain-specific *cis* inhibition to drive subcellular targeting specificity on the target cell.

Multiple subcellular gradients can operate in tandem to yield the stereotyped subcellular targeting patterns. Indeed, this is illustrated by follow-up studies on the same DG-CA3 circuit. Given the above model, one might expect Sema6A deletion would lead to mossy fibers targeting all CA3 layers. This, however, is not the case. The persistence of 'appropriate' targeting in the Sema6A mutant suggested that an additional Plexin4A-mediated signal may still be present in these animals[233]. Indeed, in a follow-up study, Sema6B was found to have a similar repulsive effect as Sema6A, with these molecules functioning additively to repel MFs, thereby revealing functional redundancy[335]. These data illustrate how the organism can utilize molecular redundancy within the same system, thereby ensuring proper connectivity by bolstering the probability for correctly-operating molecular pathways. Unlike with Sema6A, Sema6B is not inhibited in cis by PlexinA2. Instead, the authors found that in addition to its cis inhibitory activity, PlexinA2 also acts an attractive signal, instructing MF innervation of the SL according to the original attractive domain model[334]. Thus, multiple subcellular gradients can operate in tandem, even utilizing overlapping molecules, to yield the stereotyped subcellular targeting patterns. In all of these instances, however, the important unifying principle of this model is that subcellular gradients of synaptic adhesion molecules on the target cell are directing domain-specific synaptic targeting.

1.4.3.2. Regional regulation of synaptogenesis: guidance- and targeting-independent regulation of subcellular specificity

The aforementioned models of subcellular specificity had to do with secreted factors or adhesion molecules acting as guidance cues to direct axons to specific subcellular regions on the postsynaptic cell, followed by synapse formation at the stereotyped location. These models suggest

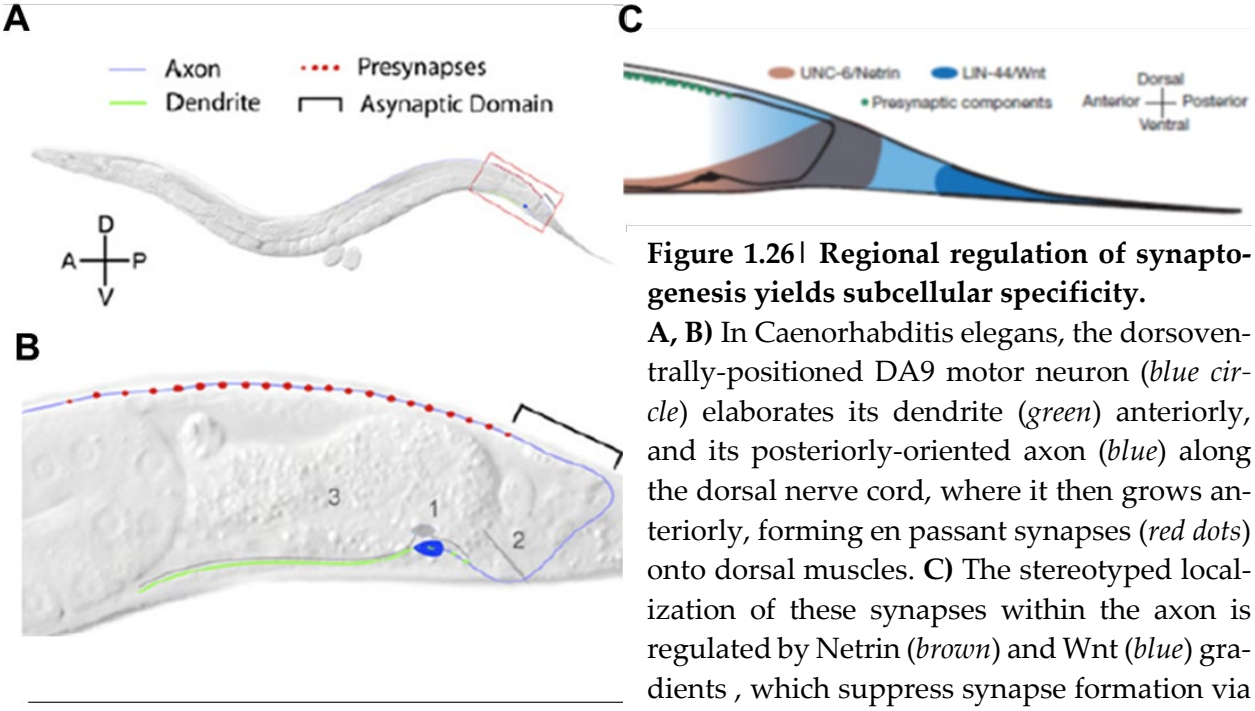

Figure 1.26| Regional regulation of synaptogenesis yields subcellular specificity.
A, B) In Caenorhabditis elegans, the dorsoventrally-positioned DA9 motor neuron (*blue circle*) elaborates its dendrite (*green*) anteriorly, and its posteriorly-oriented axon (*blue*) along the dorsal nerve cord, where it then grows anteriorly, forming en passant synapses (*red dots*) onto dorsal muscles. **C)** The stereotyped localization of these synapses within the axon is regulated by Netrin (*brown*) and Wnt (*blue*) gradients , which suppress synapse formation via DCC and frizzled receptors, respectively. Synapses (*green*) are only formed in low concentration regions far from these organizing centers. Removal of these gradients or their receptors in the DA9 neuron leads to ectopic synapse formation within otherwise asynaptic domains. Figures from Klassen & Shen, 2007 (A, B) and Poon et al., 2008 (C).

that proximity alone is sufficient to initiate synaptic targeting. There is evidence, however, that this is not always the case: sometimes synapse formation is a highly regulated process. In Caenorhabditis elegans, the dorsoventrally positioned DA9 motor neuron extends its axon posteriorly along the dorsal nerve cord, then anteriorly, where it forms en passant synapses onto dorsal muscles (Figure 1.26, A, B). The stereotypical location of these synapses is regulated by gradients of morphogens secreted from precisely patterned organizing centers (Figure 1.26, C)[320, 378]. Specifically, Netrin and Wnt suppress synapse formation via DCC and frizzled receptors on the axon, thereby restricting synapse formation to low concentration regions. Removal of these gradients or their receptors in the DA9 neuron leads to ectopic synapse formation within otherwise asynaptic

domains . Moreover, in DCC and frizzled knock-out animals, selective re-expression of these receptors in DA9 neurons rescues the phenotype. This mechanism, then, is distinct from the pre-contact axon pathfinding models discussed previously: axons are able to pass freely through those areas and are in contact with their target cells, but they are unable to initiate synapse formation. Here the environment is acting in an instructive manner to limit connections to specific regions by regulating synaptic assembly.

1.4.3.3. Synaptic plasticity: Activity-dependent changes to synapses

Synaptic plasticity is considered a distinct stage of synaptogenesis insofar as it encompasses activity-dependent changes to synapses, encompassing both synapse refinement and elimination. Synapses are not static entities in mature animals. Rather, they are incredibly dynamic, displaying activity-based expansion and diminution, with remarkably high rates of turnover. Such activity-dependent synapse refinement is a well-established model for how synaptic specificity can arise. Synaptic transmission can induce translocation of synapses to different subcellular regions, and competition between synapses can lead to specific elimination of promiscuous connections, and even in the establishment of distinct subcellular innervation domains[320], all of which result in stereotyped subcellular connectivity. We discuss the role of synaptic activity in shaping subcellular targeting below.

Synaptic plasticity is also primarily focused on mechanisms that shape synaptic properties. These properties can be established during synapse assembly, independent of activity, or evolve in response to synaptic transmission. This area of synaptic plasticity includes well-studied fields such as long-term potentiation and depression (LTP and LTD) and Hebbian learning, which are beyond the scope of this work. It is important to note, however, that most of the synaptic adhesion

molecules diagrammed above are involved in shaping synaptic properties rather than target recognition or synapse assembly[379]. We have already discussed the Neurexin/Neuroligin interaction above, and mention the Netrin-G/NGL interaction below, but offer these primarily as counterpoints to models of synaptic targeting and assembly. If the reader is so inclined, we recommend excellent reviews that discuss these molecules and their roles in shaping synaptic properties in greater depth[337, 380].

1.4.3.4. Activity-dependent refinement of promiscuous connections

In the cerebellar cortex, Purkinje neurons (PkCs) are the sole efferent and thus form a central node for information processing and transmission. Two glutamatergic afferents onto Purkinje neurons, namely granule cell parallel fibers (PFs) and climbing fibers (CFs) from CF cells in the inferior olive of the Medulla oblongata, illustrate how subcellular specificity can be defined during synaptogenesis via transsynaptic communication and competition (Figure 1.27)[338, 380-382]. Granule

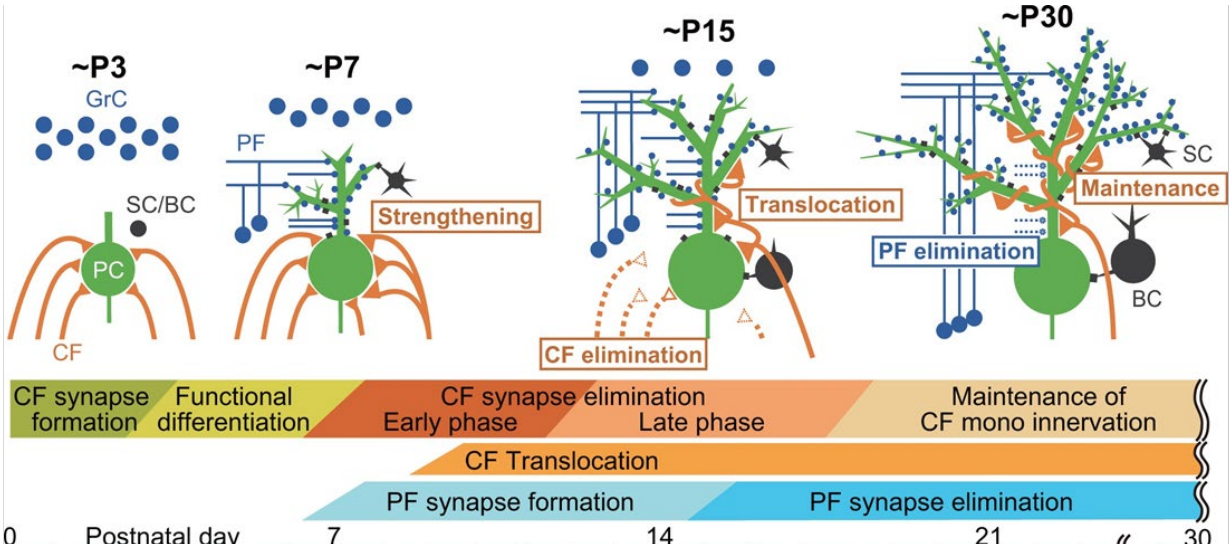

Figure 1.27 | Remodelling of CF and PF synaptic connections onto PCs during postnatal cerebellar development.
(*Upper panel*) Schematics depicting developmental changes in CF and PF synaptic connections to PCs at ~P3, ~P7, ~P15, and ~P30. (*Lower panel*) Key events related to postnatal development of CF-to-PC and PF-to-PC synapses from birth to ~P30. BC, basket cell; CF, climbing fiber; GrC, granule cell; PC, Purkinje cell; PF, parallel fiber; SC, stellate cell. Figure and legend from Kano & Watanabe, 2019.

cells form functionally weak but numerous (~100,000 PF synapses/PkC) synapses onto spines of distal dendrites from many PkCs. In contrast, in adult cerebellum, most PkCs are innervated by a single CF, forming functionally very strong but also very scarce excitatory input onto stubby spines of proximal dendrites. Notably, the stereotyped territory segregation of PF and CF innervation is activity-regulated and arise via synaptic competition. Since crosstalk between PF and CF afferents affect the final localization pattern of each, it is therefore prudent to provide a cursory discussion of the developmental trajectories of each individually, then to discuss both in tandem. In addition, the developmental trajectory of CF synapses introduces an additional model of subcellular targeting specificity, namely activity-dependent translocation.

CF innervation of PkCs is one of the best characterized models of activity-dependent refinement of synaptic circuitry in the mammalian brain. It is generally broken down into a series of partially overlapping developmental stages (Figure 1.27)[383]. During the *Synapse formation* stage (P0-P3), PkCs are initially innervated by multiple CF afferents. Immature PkCs do not have large primary dendrites, so CFs first terminate on perisomatic protrusions and thorns emerging from the PkCs somata. CFs continue to target soma even as PkC dendrite grows into the molecular layer at P6. During the *Functional differentiation* stage, (P3-P7), a single CF is selectively strengthened on the soma of each PkC. Next follow the overlapping processes of *CF elimination* (P7-P17), during which 'loser' CFs are selectively eliminated, and *CF translocation* (P9-P30), during which the winner CF translocates from PkC soma to the proximal (stem) dendrite. The final *maintenance* stage leads to the stereotypical CF mono-innervation and complete translocation of all CF synapses from the soma to the proximal region of the PkC stem dendrite.

CF elimination and translocation are both activity-mediated processes, with 'winner' and 'loser' being determined by synaptic competition. CF translocation appears to be mediated by homosynaptic competition between CFs. This was elegantly demonstrated using *in vivo* imaging, where selective elimination of the 'winner' CF by photoablation led to promotion and translocation of one of the 'loser' synapses into the 'winner' position[384]. This competition is mediated by CF induced calcium transients: disruption of postsynaptic calcium transients via $Ca_V2.1$ Ca^{2+} channel knockout led to strengthening and translocation of multiple CF inputs onto apical dendrites[385].

Whereas the translocation process is mediated by homosynaptic competition, the elimination process is mediated by heterosynaptic crosstalk. The early phase of CF elimination corresponds to the initial stages of CBC innervation at the PkC AIS and soma. As CBC innervation of the soma increases, the 'winner' CF translocates to more apical somatic regions and eventually the proximal region of the stem dendrite. Meanwhile, 'loser' CF, which do not translocate, are eliminated. This synchronicity raised the question of whether CBC inhibition and CF elimination are linked. Indeed, the bridge appears to be calcium signaling, which is positively regulated by CF activity and negatively regulated by CBC GABAergic inhibition. Genetic and pharmacological changes in GABAergic signaling altered elimination of 'loser' synapses from PkC soma, with decreased GABAergic activity leading to impaired CF elimination[386]. Of note, PkC calcium transients are significantly larger in mice with genetically decreased GABA transmission, which is reversed by increasing GABA sensitivity with diazepam. These studies indicate that heterosynaptic cross-talk between CBCs and CFs can lead to specific elimination of CF synapses. Moreover, this is an example of synaptic specificity arising from an initial wave of promiscuous targeting followed by selective elimination.

As noted above, in the adult cerebellum, parallel fiber (PF) afferents from granule cells innervate PkC distal dendrites. PF synapse formation occurs during the second postnatal week after CF innervation (P7-P14) (Figure 1.27). During this time PFs innervate the growing PkC dendritic stalk and to a much lesser extent apical regions of the PkC soma. Beginning at P15, PF synapses at PkC soma and proximal dendrites are selectively eliminated, contracting the PF innervation territory until it is confined to PkC distal dendrites.

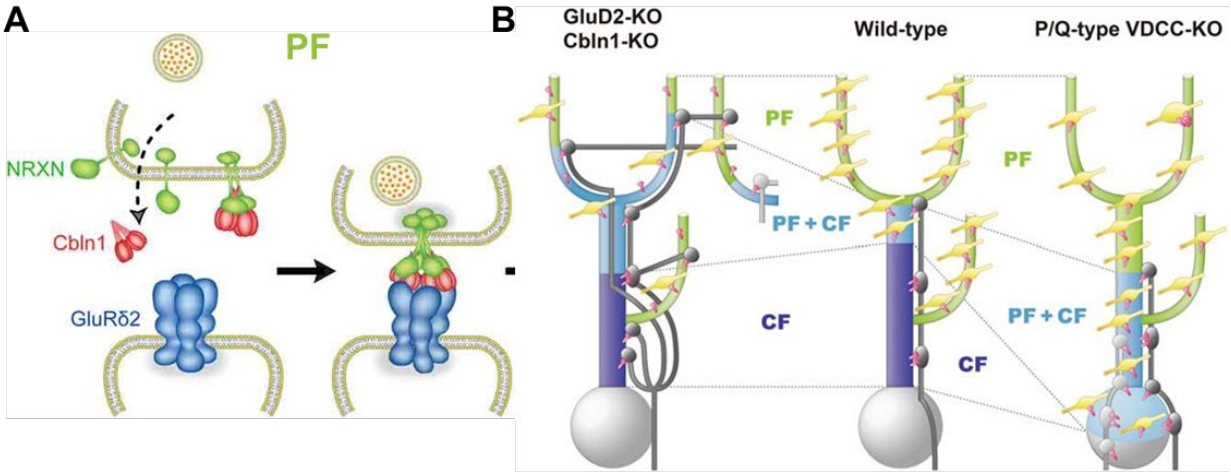

Figure 1.28 | Molecular basis of PF-PC synapse formation and impact of targeted disruptions on PF and CF innervation domains.
A) Molecular mechanism mediating PF-PC synapse formation. Cbln1 secreted from nascent PFs acts as a transsynaptic bridge between Nrxn at the PF presynapse and GlurD2 at the PC dendrite. When the contact between the PF terminal and PC spine takes place, GluRδ2 triggers Nrxn tetramerization and synapse formation. Note that Cbln1-induced dimerization of Nrxn is not sufficient to drive synapse formation. (Figure adapted from Mishina et al., 2012). **B)** Contrasting alterations in PC's excitatory synaptic wiring between GluD2/Cbln1 knockout and P/Q-type VDCC knockout mice. (*Middle*) In wild-type mice, CF and PF territories are sharply segregated, and CF mono-innervation is established. (*Left*) In GluD2/Cbln1 knockout mice, CF innervation territory expands, and multiple CF innervation occurs frequently by additional wiring at distal dendrites. (*Right*) In P/Q-type VDCC knockout mice, PF innervation territory expands, and multiple CF innervation occurs at basal dendrites and somata. Figure and legend from Kano & Watanabe, 2022. *Reproduced with permission from Springer Nature*.

The PF-PkC interaction is mediated by a transsynaptic interaction between the presynaptic adhesion molecule Neurexin 1, the presynaptically-expressed transsynaptic extracellular scaffolding protein Cbln1, and postsynaptic glutamate receptor GluRd2 (Figure 1.28, A)[338, 380, 387-389].

Ablation of Cbln1 from Granule cell PFs or GluD2 from the PkCs disrupts PF targeting, leading to decreased PF synapses on PkCs and an overall contraction of the PF innervation territory to more distal regions of the PkC dendrite (Figure 1.28, B)[258, 386]. Of note, these disruptions also lead to expansion of the CF territory to more distal dendritic regions. Since the Nrx1-Cbln1-GluD2 complex is required for GluD2 channel opening[383], both interventions lead to decreased PF-based PkC activation at distal dendrites. Similarly, blocking PF-based PkC activation via mGluR1-PKCg leads persistence of PF synapses on proximal dendrites, persistent CF somatic innervation, and multiple innervation of CFs due to issues with elimination of 'loser' CFs – disruption of homosynaptic competition required to identify 'loser' CFs for slated for elimination[336]. In contrast, PF innervation territory is defined by CF activity. PF innervation territory expands to proximal dendrites when CF projections were surgically lesioned, when cerebellar cortical activity was blocked by Na-channel blocker tetrodotoxin or AMPAR antagonist NBQX, or in the P/Q-type VDCC KO mouse[390]. The latter also leads to multiple CF innervation , and ectopic hyperspiny transformation at basal dendrites and somata. This is attributed to the lack of activity-dependent elimination of 'loser' CFs. In summary, then, both PF and CF afferents follow a process of promiscuous targeting followed by selective elimination, a model first proposed by Haydon and Drapeau[391].

Activity is also important for cortical interneuron subcellular targeting. Soma-targeting PV basket cells display an increasing number of boutons per soma in late prenatal development (P18-28). This innervation is based on a critical window of activity: intraocular injection of the neuronal activity blocker TTX during the third, but not fifth, postnatal week led to a decrease in the number

of perisomatic boutons in the contralateral (activity deprived) visual cortex[199]. Of note, the researchers did not assess whether non-somatic boutons were affected, indicating that this effect could apply to PVBC boutons generally, rather than their subcellular specificity.

1.4.3.5. Induction of a subcellular domain by afferents

The prior examples required subcellular region enrichment of specific molecules as a prerequisite for subcellular targeting, suggesting that the molecular gradients are always present prior to targeting. This is not always the case, however. Indeed, afferents can transsynaptically specify subcellular molecular gradients and domains. Perhaps the best characterized example of this process is synaptic spine formation. Since we are primarily interested in cortical GABAergic synapses, which do not induce spine formation, we instead focus on an example of domain induction that is more applicable to interneurons. Since a synaptic connection is already established by the time these mechanisms come into play, the following is, strictly speaking, not a model of subcellular targeting specificity. Instead, it falls within the domain of synaptogenesis and subcellular compartmentalization. We include it here, however, as a contrast with the prior domain model.

In the hippocampus and cortex, membrane-bound Netrin-G1 and -G2 are expressed on axons from different afferent pathways that terminate onto distinct laminae, where they interact with their postsynaptic receptors NGL-1 and -2, respectively (Figure 1.29)[199, 392]. Netrin-G1 and NGL1 localize to the SLM/OLM in hippocampus and L1 and L4 in cortex, whereas Netrin-G2 and NGL-2 localize to SR and SO in hippocampus. Deletion of presynaptic Netrin-G1 or -G2 does not affect axon pathfinding or laminar targeting, but does lead to selective disruption of the laminar localization of postsynaptic NGL-1 and -2, respectively[393]. Specifically, Netrin-G1 deletion leads to

NGL-1 diffusion along dendrites with no effect on NGL-2, and vice versa upon Netrin-G2 deletion. These results indicate that these transsynaptic interactions are not involved in synaptic targeting or synapse assembly, but rather that anterograde signaling from Netrin-Gs on the axon leads to recruitment of the corresponding NGLs to the innervated subregion of the dendrite[393, 394]. Subsequent studies showed that deletion of Netrin-Gs leads to distinct functional synaptic deficits[395] and fear- and anxiety-associated behavioral changes[236, 396], indicating that the Netrin-G/NGL interactions are important for shaping synaptic properties in these circuits.

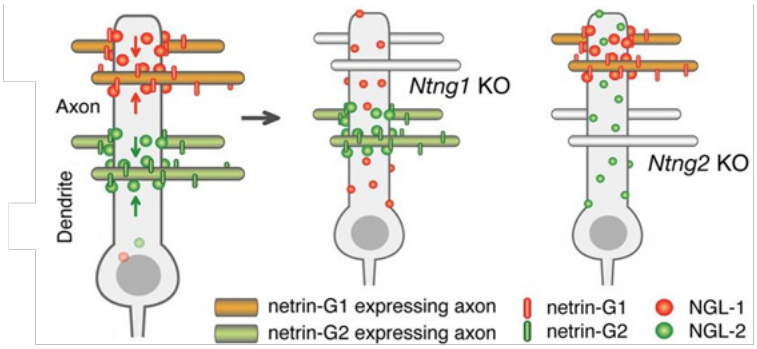

Figure 1.29 | Model for transneuronal regulation of receptor localization by axonal netrin-G proteins. Distinct axon populations, each expressing netrin-G1 or -G2, make contact with distinct segments of dendrites of target neurons. NGL-1 and -2, which are expressed in the target neurons, are anchored by axon-derived netrin-G1 and -G2, and are thereby precisely arranged onto the selected subdendritic segments. Without netrin-Gs as extrinsic cues, NGLs would be dispersed across multiple segments of dendrites. Figure and legend from Nishimura-Akiyoshi et al., 2007.

1.4.3.6. Reduce, Reuse, Recycle: Repurposing molecules leads to context-dependent functions

Although we have attributed many different biological functions to different biomolecules, these molecules do not have any intrinsic function per se. Many biomolecules are recycled and reused in different cellular contexts, where they can perform distinct roles. We discussed above how hippocampal neurons utilize presynaptic Netrin-G2 to induce NGL-2 clustering at the postsynapse. The Netrin-G2/NGL-2 interaction is also utilized in the retina by horizontal cells to drive lamina-specific targeting of rods. Here, however, the localizations are reversed: NGL-2 is

pre-synaptic and Netrin-G2 is postsynaptic. Generally such a change in localization would be unremarkable, but here it is accompanied by a profound functional shift. In this context, NGL-2 performs a more classic role in axon pathfinding and attractive domain targeting[4]. Deletion of NGL-2 from horizontal cells led to overgrowth of their axons and decreased synapse formation. This was rescued by NGL2 re-expression even after phenotype development. In contrast, overexpression of NGL2 in wild-type horizontal cells elevated synapse numbers above normal levels. Thus, NGL2 promotes the formation, maintenance, and restoration of synapses in the developing and mature retina by restricting axon growth and promoting synapse formation[5]. These data show how the same molecules can play different roles in different contexts. Indeed, this repurposing and context-dependent function is a general theme in biology. Many molecules are recycled and reused in different cellular and developmental contexts. We revisit this topic in the discussion.

1.5. Targeting specificity *ex vivo*

Dissociated culture of neurons is an invaluable tool for investigating general neuronal properties, but it is often assumed that specificity is lost following the dissociation and plating process. It is thus worthwhile to discuss what is known about synaptic specificity both *in vivo* and *ex vivo*. We will first define our expectations by discussing cortical interneuron specificity *in vivo*. We will then review literature findings for neuronal targeting *ex vivo*. This discussion will include findings from both glutamatergic and GABAergic cells of the cortex and hippocampus. Generally speaking, there are two different approaches for primary tissue culture, namely 1) organotypic slice cultures, where tissue is sliced and plated largely intact, and 2) dissociated culture, where the tissue is dissociated into a single cell suspension before plating. There are similarities and differences for both approaches, with the similarities largely associated with the dissection and isolation of tissue,

whereas the differences are due to the dissociation (or not) of the tissue prior to plating, and thus randomization of spatial organization and alteration of cellular density.

1.5.1. Specificity and exclusivity of cortical interneuron connectivity *in vivo*

Cortical interneurons form lamina-specific[6-9, 14], cell-type-specific[15, 16], and subcellular-compartment-specific[14, 17-20, 397] connections. It is important to note that in almost all these cases, the specificity observed in interneuron connectivity is expressed as preferential targeting rather than strict selectivity[398]. One model for interneuron cellular targeting proposed that these cells formed promiscuous synapses onto all proximal target cells, with little cell type specificity[10, 11]. Although individual case studies in the literature[5, 6, 12, 13] and a large-scale multi-patch paired recording study[14] suggest otherwise[6, 14], this non-selectivity model was only conclusively disproven recently via systematic assessment of cellular connectivity from painstaking reconstructions of electron microscopy data across an entire cortical column (a truly herculean effort!)[14]. This study clearly shows that morphologically-defined interneuron subtypes display preferential cellular targeting, including translaminar targeting that skips intermediate cellular layers. Despite the clarity of this conclusion, the actual circuit diagram is not nearly as clean as one might naively expect from existing models of cellular selectivity[17, 18]. Thus, data suggests interneuron cellular connectivity is primarily expressed as a preference rather than strict exclusivity.

In general, interneurons also display preferential (not exclusive) subcellular targeting[12, 19, 20]. For instance, somatostatin cells are generally considered dendrite-targeting, but careful quantitative assessment of these cells by immuno-EM showed that a small percent of Sst+ presynaptic boutons are found at the soma (5.5%) and AIS (1.2%)[55]. One model to explain these data is that

[14] For Jiang et al., 2015, see also the comment by Barth et al., 2016, also cited here, for important contributions and limitations of this study.

the Sst population is heterogeneous, with minor subpopulations displaying exclusivity in subcellular targeting. That is, assuming all Sst cells form the same number of connections (which they don't), 5.5% of Sst+ cells are a subpopulation(s) that exclusively targets soma and 1.2% by another(s) that exclusively target AIS. This is not the case. Although the Sst population itself *is* heterogeneous[14, 35, 37, 56-58, 399], individual Sst cells *also* display heterogeneous targeting[20, 177, 399]. Moreover, the relatively low percent of AIS-targeting synapses in the Sst population underrepresents the extent of non-chandelier cell synapses on AIS as a fraction of total AIS-targeting synapses. In fact, 41% of all AIS-targeting inhibitory synapses are from non-chandelier cells[177]. These data indicate that interneuron subcellular targeting specificity generally is preferential rather than exclusive. Among cortical interneurons, the one exception to this rule is chandelier cells, which target the axon initial segment at nearly 100%[399, 400], even displaying a preference for more distal regions of the AIS compared to non-ChCs[390].

1.5.2. Dissection: isolation and control

We will now discuss dissection and dissociation. *Ex vivo* plating and growth of intact (non-dissociated) tissue is called organotypic culture, whereas that of dissociated tissue is called dissociated culture. There are important similarities and differences between both techniques. Regarding the similarities, both are reductionist methods of assessing the tissue sample in isolation, that is, separated from other organ systems. This isolation derives from the dissection process, and therefore is a common aspect to both organotypic and dissociated culture systems. Tissue isolation is a powerful experimental method as it allows researchers to separate organ systems that may confound interpretations and/or results, thus yielding considerable power to define intrinsic processes from extrinsic ones. One example of tissue isolation was discussed above for compartmen-

talization, where plating of isolated neurons enables researchers to test whether subcellular compartments arise due to cell intrinsic or extrinsic programs. The same is true on a larger scale for organotypic slices, for instance, where culturing intact tissue explants enables investigation of whether a process is intrinsic to the tissue itself, or requires external signaling. A good example of this is interneuron subcellular targeting[401]. Using organotypic slices from the G42 mouse and GIN mouse lines, which preferentially label PV+ basket cells and L2/3 Martinotti cells, Di Cristo et al. found that subcellular targeting is recapitulated in organotypic slices *ex vivo*, indicating that subcellular targeting of these GABAergic neurons is independent of sensory or thalamic input. Remarkably, not only was PV BC soma targeting recapitulated in organotypic slices *ex vivo*, it also proceeds at the same maturation rate as *in vivo*[402]. In both contexts, this maturation is regulated by neuronal activity, as TTX treatment led to decreased number and size of perisomatic boutons both *in vivo* and *in vitro*. Thus, isolation from other neuronal tissue enables researchers to test whether the biological process under consideration is affected by afferent and/or efferent connections, the circulatory and immune systems, and the general milieu of secreted molecules signaling into the tissue itself from outside sources.

Perhaps most importantly, this isolated system provides researchers with incomparable power to manipulate the tissue environment by adding, removing, or changing individual components to test hypotheses. The breadth of manipulations varies widely, ranging from ions and small molecules to expressed genes and even addition/removal of organ systems. Using organotypic slice cultures from the G42 mouse and single cell transfection, Chattopadhyaya et al., showed that conditional knockdown of the GABA synthetic enzyme GAD67 in PV+ basket cells resulted in cell

autonomous defects in axon branching, perisomatic synapse formation, and complexity of the innervation field[403]. The same group showed in a later study that presynaptic expression of the synaptic adhesion molecule NCAM is involved in perisomatic innervation by PVBCs via Fyn[404].

The addition of different organ systems, tissues, or cell types is called co-culturing, and is an important means of distinguishing intrinsic and extrinsic programs. For instance, coculturing thalamus and cortex allowed researchers to assess whether thalamocortical targeting specificity is intrinsic to this system. These researchers found that although laminar specificity is preserved[135, 405], regional specificity is not[406]. Instead, thalamocortical afferents are rerouted by ephA/ephrinA gradients in the ventral telencephalon, an intermediate brain structure[223], a discovery made by co-culture of these three brain regions in their native orientations. Thus, the presynaptic and postsynaptic mechanisms underpinning laminar targeting specificity and laminar cell identity are largely intrinsic to the cortex, whereas those driving regional targeting and identity are generally extracortical and/or extrathalamic. A similar recapitulation of laminar specificity was witnessed in the hippocampus when investigating the laminar termination of entorhinal cortex afferents and commissural fibers[224], which, as discussed above, is an example of subcellular targeting specificity. Thus co-culture is a powerful tool to assess whether a process is intrinsically driven by the tissue (or cell) or induced by extrinsic environmental cues. Coculture systems can also be used to identify intrinsic signaling, as was elegantly shown for axodendritic polarization and neurite orientation in the cortex. Using cortical slices co-cultured with sparse, labeled dissociated cortical neurons, Polleux et al., found that a semaphorin gradient originating near the cortical surface acts attractively for dendrites and repulsively for axons, thereby orienting these cellular compartments radially[227, 407] (reviewed in [408]). Of note, the researchers identified the origin of the polarizing signal by positioning cortical slices at different orientations to one another (e.g. WM of one slice apposed

to MZ of another, or MZ of one slice oriented along the layers of another) and investigating the effect on axodendritic orientation of the dissociated labeled cells near the slice juxtaposition. These co-culture experiments illustrate the power and control afforded by an isolated, reductionist culture system in parsing intrinsic and extrinsic developmental programs. It is important to note that dissociated primary culture is itself a form of co-culture, wherein the heterogeneous cellular composition can be used to the advantage of the researcher.

1.5.3. Dissociation: a bit of chaos to disrupt intrinsic spatial organization

The major difference between the organotypic and dissociated cultures is the relative spatial organization of cells. Simply put, dissociation disrupts the spatial relationships between cells in two ways, 1) by randomizing their relative positioning, and 2) by altering the cellular density during plating.

By randomizing the cortex we are also disrupting the gradients and structures that were established during embryonic development and are utilized for axon pathfinding over long distances. Thus, the dissociation process not only disrupts these large-scale chemical gradients, but also randomizes cellular positioning upon plating, leading to randomized cellular environments. We discussed above how cellular scaffolds and extracellular gradients form a molecular framework that ensures the 'proper' presynapses form on the 'proper' targets. This molecular framework is disrupted by dissociation. The distance that any given gradient can reach is in turn confined to a hyper-local region. For secreted factors, this is small scale paracrine signaling: in a 2D culture system, the plating density is associated with the average distance between cells. As discussed above, the maximum effective distance for soluble molecules, such as neurotrophic factors or morphogens, is 25 cell diameters or approximately 250 um, with decreasing distances for less diffusible molecules such as ECM components[409]. For adhesion molecules, these gradients are abolished

entirely, reduced to molecular distances[15]. Thus, through dissociation, we are testing the extent to which interneuron subcellular synaptic specificity relies upon these long-distance signaling factors and cellular structures. Given the often-overlapping gradients in an intact cortex, this question would otherwise be impossible to test *in vivo* or in organotypic slices.

But dissociation is also testing the extent to which the local microenvironment determines cellular identity. Indeed, a major concern of dissociated culture is the effect that the randomization process has on cellular identity. A major question therefore arises: to what extent is the environment simply a permissive bystander to the unfolding of cell intrinsic processes, or is the local environment functioning in an instructive manner, guiding the neuron's development along a series of potential fate branch points. As discussed above, regional identity in the cortex is, at least initially, is driven by a complex network of transcriptional regulation and paracrine signaling that establishes polarized identity across the cortical sheet. Thus, regional identity is associated not only with the local cellular microenvironment, but in a knock-on manner, also more distal cellular microenvironments. Through this mechanism, regional identity of glutamatergic cells is intimately linked to cellular positioning along the A-P and L-M axes. Since dissociation randomizes cellular positioning across all axes, it is safe to assume that, in a dissociated culture context, regional identity is either retained but randomized, lost entirely, or reestablished at random. Either way, the generally uniform distribution of interneurons across different cortical regions discussed above suggests that matching targeting and target regional identity may not be a prerequisite for interneuron synaptic specificity.

[15] If axon pathfinding via adhesion molecule gradients are based on differential expression between neighbors, then cellular randomization will increase this differential massively, since it is highly unlikely that a neighboring cell will express similar levels of these molecules. This could be a confound.

The same does not appear to be true of cellular identity and positioning along the deep-to-superficial (radial) axis. Rather, at least for cortical glutamatergic neurons, cellular identity appears to be largely independent of radial cellular positioning. This principle is most clearly illustrated by stem cell differentiation experiments, where both glutamatergic[410-412] and GABAergic[412-414] cells retain molecular markers of and determinants for neuronal cell type identity in a dissociated context, with RNA expression and splicing profiles that generally correspond with their *in vivo* correlates[44, 412, 413, 415]. Of note, however, their synaptic specificity has not been assessed in this context, with researchers instead testing this property through transplantation studies. There is evidence to suggest that some identifying neuronal properties *may be* impacted by cellular radial positioning, however. The most significant effect was shown by Oishi et al., for L4 neuron identity, where *in vivo* repositioning of L4 neurons into L2/3 by shRNA interference led to a fate switch such that they resembled L2/3 pyramidal neurons. Of note, the fate switch was verified not only by molecular expression patterns, but also connectivity: the repositioned L4 cells also lacked their characteristic thalamocortical afferents[21]. Other research on radial positioning showed that multiple morphological properties of glutamatergic cells are altered in the reeler mouse[89, 130]. These *in vivo* experiments are more difficult to interpret, however, since these molecules are repurposed in many developmental processes, with the observed cellular identity changes and altered radial positioning being a parallel but causally unrelated phenotypes, rather than the radial positioning changes being causal predecessor to the identity changes. In contrast, connectivity is generally retained in the reeler cortex, as discussed above, suggesting these morphological changes are not sufficient to disrupt the cellular identity underlying afferent connectivity profiles. Moreover, the reeler cortex in reality is neither fully inverted nor randomized, as was shown by multi-probe in

situ gene expression profiling[150] and functional studies[78, 118, 131]. Regarding interneuron connectivity, however, attempting to parse the impact of spatial positioning, and thus local environmental cues, on synaptic specificity *in vivo* is not so easy. This is primarily due to the fact that interneurons migrating into the cortex tend to position themselves near their synaptic partners. This is true not only in the wild-type cortex, but also in the 'inverted' reeler cortex, where interneurons will also be positioned in a generally inverted manner[40, 416, 417]. Given this propensity for self-organization,the question of how local positioning impacts cellular identity and targeting is not easy to solve *in vivo*.

Dissociated culture offers an alternate approach to test these hypotheses. Although many studies around the turn of the century suggested synaptic specificity is retained in a culture dish[418], it is generally assumed, that specificity is lost following the dissociation and plating process. This is demonstrated by observations that Glutamatergic axons innervate the AIS and organize $GABA_A$ receptors in dissociated cultured hippocampus[419], which does not occur *in vivo*. Similarly, in dendrites, some gephyrin clusters will be juxtaposed with glutamatergic synapses[262], a non-physiological organization. In the extreme, isolated neurons on microislands, where they are forced to grow within a highly confined space, will target themselves and form autapses where glutamatergic presynapses will be in apposition to GABAergic postsynapses and vice versa[420]. Although these data suggest that targeting specificity is not retained after dissociation, it is important to recognize that all of these experiments used non-physiological cell densities, opening the possibility that the ectopic targeting is an 'any port in a storm' response, whereby low affinity molecular interactions that would otherwise be outcompeted are utilized because they're the strongest given the context. Indeed, using more physiological cell densities, Williams et al., recapitulated both the cellular and subcellular targeting specificity of hippocampal DG-to-CA mossy fiber synapses in

dissociated culture[421]. More recent studies have used stem cell-derived neurons to model hippocampal connectivity *in vitro*[391], showing that the stem-cell-derived CA3 PyNs are able to recapitulate their *in vivo* cellular targeting specificity in a dissociated context, an important step for disease modeling. Indeed, there has been a renewed interest recently in image-based profiling of synaptic connectivity using dissociated neuronal cultures[199].

1.5.4. Summary of *ex vivo* targeting specificity

Above we discussed the challenges and opportunities of investigating neuronal targeting specificity in dissociated cultures compared to *in vivo* conditions. *In vivo*, cortical interneurons tend to exhibit preferential targeting rather than strict selectivity, forming connections that are lamina-specific, cell-type-specific, and subcellular-compartment-specific. Whether this targeting is recapitulated *ex vivo* in a dissociated culture context is the primary question we address in this study.

Regarding the isolation and control of experimental conditions, both organotypic slice cultures and dissociated cultures allow researchers to separate tissue from confounding factors, enabling the study of intrinsic processes. Organotypic slices can replicate certain *in vivo* conditions, such as interneuron subcellular targeting, independently of external input. Co-culture experiments further enhance the understanding of intrinsic and extrinsic factors influencing cellular development and connectivity.

Dissociated cultures disrupt spatial organization, randomizing cellular positioning and altering cellular density. This disruption challenges the notion of strict targeting specificity observed *in vivo*. While some studies suggest that specificity is lost in dissociated cultures, recent research using more physiological cell densities has shown promising results in recapitulating cellular and

subcellular targeting specificity. Stem cell-derived neurons have also been utilized to model connectivity, indicating the potential for disease modeling and renewed interest in studying synaptic connectivity in dissociated cultures.

1.6. Summary of principles

To summarize, the developmental stage underlying synaptic specificity can be broken into those prior to synapse formation, synaptic targeting itself, and those following synapse formation. Prior to synapse formation, the process of selecting a specific synaptic target arises through an interaction between two distinct pairs of components, namely 1) the presynaptic cell interacting with its microenvironment, and 2) the postsynaptic target interacting with the target microenvironment. Here we are defining 'microenvironment' as anything other than the presynaptic cell and the postsynaptic target. There are also several prerequisites, namely 1) spatial proximity between presynaptic growth cone and postsynaptic target, 2) molecular complementarity between the two, and 3) a conducive microenvironment. The 'conducive' microenvironment is simply a cellular landscape within which these components acquire the necessary spatial proximity and express the 'appropriate' molecules at the appropriate subcellular regions. As such, this landscape can be functioning in either a permissive manner, acting as a passive bystander that allows cell intrinsic programs within the presynaptic cell and postsynaptic target to unfold, or in an instructive manner, guiding the neuron's development along a series of potential fate branch points. Synaptic target recognition is the physical interaction between the axon growth cone and the putative target, wherein the mutual expression of complementary adhesion molecules enables the presynaptic cell to select a specific postsynaptic target out of all possible partners. Here, the operative interaction is solely between the presynaptic cell and its putative target, and can be viewed as a binary decision: to form (or not to form) a synapse. Following successful synaptic targeting, the presynaptic

and postsynaptic components engage in reciprocal signaling to assemble the synapse. The synapse can now be considered a unit. In addition to the presynaptic and postsynaptic cells interacting with their respective microenvironments, this post-formation stage also includes direct and sustained transsynaptic cross-talk between the presynaptic and postsynaptic cells. This cross-talk includes both anterograde, presynaptic-to-postsynaptic, and retrograde, postsynaptic-to-presynaptic, transsynaptic signaling. Moreover, the effects of this signaling can be the local (isolated to proximal changes within the synapse itself) morphological and molecular changes discussed above, regional (extending beyond the immediate vicinity of the synapse), such as definition of molecularly distinct subcellular regions[385-387, 392, 422], and even signaling to the nucleus, such as survival signals[423, 424] and transsynaptic delivery of transcription factors[425]. Competition between nearby synapses on the postsynaptic cell can lead to specific elimination, resulting in stereotyped connectivity.

1.7. Conclusion and hypothesis

We set out to test whether cortical interneuron subcellular synaptic specificity requires the higher order cortical and extracortical organization present when these neurons begin to send out their axonal projections in their native context. To state the experiment formally, our null hypothesis postulates that cortical interneuron subcellular targeting is dependent on one or more of the following:

1. The interneuron's positioning within a specific cortical subregion or layer (and by extension, its migratory history).
2. The target cell's positioning within a specific cortical subregion or layer (and by extension, its migratory history).

3. Some pre-existing cellular structure(s) or molecular gradient(s) formed by intermediate (i.e. non-targeted) cells.
4. Indirect interaction(s) between the cortex and extracortical tissue. Examples include inputs from vasculature, extracortical neuronal projections into the cortex, or outward projections from cortical neurons to extracortical regions.

If subcellular targeting depends on any of the above principles, then disrupting these processes by excising the cortex, randomizing its contents, and growing it in isolation should also randomize interneuron targeting. Alternately, if subcellular targeting specificity *is* recapitulated in this randomized and isolated context, we can conclude that:

1. Direct contact between the pre- and postsynaptic adhesion molecules alone can account for interneuron subcellular targeting specificity. In other words, environmental cues, be they cellular structures, molecular gradients, or otherwise, are not required for these interneuron populations to form 'appropriate' subcellular connections.
2. Expression of the presynaptic molecules necessary for such targeting is intrinsically determined by the presynaptic cell, and that this designation occurs prior to interneuron migration into the cortex.
3. The molecular gradients responsible for localizing the target adhesion molecules to the 'appropriate' subcellular location are intrinsically determined in the target cell, and these gradients persist (or can be re-formed) after the target cell settles into its 'appropriate' layer.

This question is not simply one of basic biology, however. There is also translational potential here. Psychiatric diseases are due to abnormal synaptic circuitry. The interneuron subtypes we have chosen to study, soma-targeting basket cells (BCs), AIS-targeting chandelier cells (ChCs),

and dendrite-targeting somatostatin cells (SstCs) are all hypothesized to underlie a wide variety of psychiatric pathologies such as schizophrenia [425-429], autism [430-434], and bipolar disorder [435-438]. Interneuron dysfunction and resultant excitatory inhibitory imbalance is associated with seizure disorders [61, 439-441]. Synaptic circuitry is therefore a prime candidate for drug discovery. Recapitulation of these circuits in a culture dish would greatly facilitate these discoveries.

Bibliography to Chapter 1: Introduction

1. Brown, A., et al., *Topographic Mapping from the Retina to the Midbrain Is Controlled by Relative but Not Absolute Levels of EphA Receptor Signaling.* Cell, 2000. **102**(1): p. 77-88.
2. Feldheim, D.A., et al., *Genetic analysis of ephrin-A2 and ephrin-A5 shows their requirement in multiple aspects of retinocollicular mapping.* Neuron, 2000. **25**(3): p. 563-574.
3. Benson, D.L., D.R. Colman, and G.W. Huntley, *Molecules, maps and synapse specificity.* Nature Reviews Neuroscience, 2001. **2**(12): p. 899-909.
4. Muñoz, W., et al., *Layer-specific modulation of neocortical dendritic inhibition during active wakefulness.* Science, 2017. **355**(6328): p. 954-959.
5. Munoz, W., et al., *Layer-specific modulation of neocortical dendritic inhibition during active wakefulness.* Science, 2017. **355**(6328): p. 954-958.
6. Jiang, X., et al., *Principles of connectivity among morphologically defined cell types in adult neocortex.* Science, 2015. **350**(6264): p. aac9462.
7. Bortone, D.S., S.R. Olsen, and M. Scanziani, *Translaminar inhibitory cells recruited by layer 6 corticothalamic neurons suppress visual cortex.* Neuron, 2014. **82**(2): p. 474-85.
8. Jiang, X., et al., *The organization of two new cortical interneuronal circuits.* Nat Neurosci, 2013. **16**(2): p. 210-8.
9. Ma, Y., et al., *Distinct subtypes of somatostatin-containing neocortical interneurons revealed in transgenic mice.* J Neurosci, 2006. **26**(19): p. 5069-82.
10. Lu, J., et al., *Selective inhibitory control of pyramidal neuron ensembles and cortical subnetworks by chandelier cells.* Nat Neurosci, 2017. **20**(10): p. 1377-1383.
11. Hilscher, M.M., et al., *Chrna2-Martinotti Cells Synchronize Layer 5 Type A Pyramidal Cells via Rebound Excitation.* PLoS Biol, 2017. **15**(2): p. e2001392.
12. Wu, S.J., et al., *Cortical somatostatin interneuron subtypes form cell-type-specific circuits.* Neuron, 2023.
13. Barth, A.L., et al., *Comment on "Principles of connectivity among morphologically defined cell types in adult neocortex".* Science, 2016. **353**(6304): p. 1108-1108.
14. Schneider-Mizell, C.M., et al., *Cell-type-specific inhibitory circuitry from a connectomic census of mouse visual cortex.* bioRxiv, 2023: p. 2023.01.23.525290.
15. Kepecs, A. and G. Fishell, *Interneuron cell types are fit to function.* Nature, 2014. **505**(7483): p. 318-26.
16. Favuzzi, E., et al., *Distinct molecular programs regulate synapse specificity in cortical inhibitory circuits.* Science, 2019. **363**(6425): p. 413-417.
17. Kubota, Y., et al., *Neocortical inhibitory terminals innervate dendritic spines targeted by thalamocortical afferents.* J Neurosci, 2007. **27**(5): p. 1139-50.
18. Kubota, Y. and Y. Kawaguchi, *Dependence of GABAergic synaptic areas on the interneuron type and target size.* J Neurosci, 2000. **20**(1): p. 375-86.
19. Kubota, Y., et al., *Functional effects of distinct innervation styles of pyramidal cells by fast spiking cortical interneurons.* Elife, 2015. **4**.
20. Gonchar, Y., et al., *Axo-axonic synapses formed by somatostatin-expressing GABAergic neurons in rat and monkey visual cortex.* J Comp Neurol, 2002. **443**(1): p. 1-14.
21. Polleux, F., et al., *Pre- and Post-mitotic Events Contribute to the Progressive Acquisition of Area-specific Connectional Fate in the Neocortex.* Cerebral Cortex, 2001. **11**(11): p. 1027-1039.
22. Molyneaux, B.J., et al., *Neuronal subtype specification in the cerebral cortex.* Nat Rev Neurosci, 2007. **8**(6): p. 427-37.
23. Custo Greig, L.F., et al., *Molecular logic of neocortical projection neuron specification, development and diversity.* Nat Rev Neurosci, 2013. **14**(11): p. 755-69.
24. Wamsley, B. and G. Fishell, *Genetic and activity-dependent mechanisms underlying interneuron diversity.* Nat Rev Neurosci, 2017. **18**(5): p. 299-309.
25. Lim, L., et al., *Optimization of interneuron function by direct coupling of cell migration and axonal targeting.* Nat Neurosci, 2018. **21**(7): p. 920-931.
26. Waites, C.L., A.M. Craig, and C.C. Garner, *MECHANISMS OF VERTEBRATE SYNAPTOGENESIS.* Annual Review of Neuroscience, 2005. **28**(1): p. 251-274.
27. Sanes, J.R. and M. Yamagata, *Many Paths to Synaptic Specificity.* Annual Review of Cell and Developmental Biology, 2009. **25**(1): p. 161-195.
28. Yamagata, M., J.R. Sanes, and J.A. Weiner, *Synaptic adhesion molecules.* Curr Opin Cell Biol, 2003. **15**(5): p. 621-32.

29. Sanes, J.R. and S.L. Zipursky, *Synaptic Specificity, Recognition Molecules, and Assembly of Neural Circuits.* Cell, 2020. **181**(3): p. 536-556.
30. DeFelipe, J., et al., *New insights into the classification and nomenclature of cortical GABAergic interneurons.* Nat Rev Neurosci, 2013. **14**(3): p. 202-16.
31. Tremblay, R., S. Lee, and B. Rudy, *GABAergic Interneurons in the Neocortex: From Cellular Properties to Circuits.* Neuron, 2016. **91**(2): p. 260-92.
32. Poulin, J.F., et al., *Disentangling neural cell diversity using single-cell transcriptomics.* Nat Neurosci, 2016. **19**(9): p. 1131-41.
33. Tasic, B., et al., *Adult mouse cortical cell taxonomy revealed by single cell transcriptomics.* Nat Neurosci, 2016. **19**(2): p. 335-46.
34. Rosenberg, A.B., et al., *Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding.* Science, 2018. **360**(6385): p. 176-182.
35. Tasic, B., et al., *Shared and distinct transcriptomic cell types across neocortical areas.* Nature, 2018. **563**(7729): p. 72-78.
36. Gouwens, N.W., et al., *Classification of electrophysiological and morphological neuron types in the mouse visual cortex.* Nat Neurosci, 2019. **22**(7): p. 1182-1195.
37. Yao, Z., et al., *A taxonomy of transcriptomic cell types across the isocortex and hippocampal formation.* Cell, 2021. **184**(12): p. 3222-3241.e26.
38. Yao, S., et al., *A whole-brain monosynaptic input connectome to neuron classes in mouse visual cortex.* Nature Neuroscience, 2023. **26**(2): p. 350-364.
39. Anderson, S.A., et al., *Distinct origins of neocortical projection neurons and interneurons in vivo.* Cerebral cortex, 2002. **12**(7): p. 702-709.
40. Valiente, M. and O. Marín, *Neuronal migration mechanisms in development and disease.* Current Opinion in Neurobiology, 2010. **20**(1): p. 68-78.
41. Nery, S., G. Fishell, and J.G. Corbin, *The caudal ganglionic eminence is a source of distinct cortical and subcortical cell populations.* Nature neuroscience, 2002. **5**(12): p. 1279-1287.
42. Tyson, J.A. and S.A. Anderson, *GABAergic interneuron transplants to study development and treat disease.* Trends in neurosciences, 2014. **37**(3): p. 169-177.
43. Cooper, J.A., *A mechanism for inside-out lamination in the neocortex.* Trends in Neurosciences, 2008. **31**(3): p. 113-119.
44. Polleux, F., C. Dehay, and H. Kennedy, *Neurogenesis and commitment of corticospinal neurons in reeler.* J Neurosci, 1998. **18**(23): p. 9910-23.
45. Wichterle, H., et al., *Young neurons from medial ganglionic eminence disperse in adult and embryonic brain.* Nat Neurosci, 1999. **2**(5): p. 461-6.
46. Wichterle, H., et al., *In utero fate mapping reveals distinct migratory pathways and fates of neurons born in the mammalian basal forebrain.* Development, 2001. **128**(19): p. 3759-71.
47. Anderson, S.A., et al., *Distinct cortical migrations from the medial and lateral ganglionic eminences.* Development, 2001. **128**(3): p. 353-363.
48. Jovanovic, J.N. and A.M. Thomson, *Development of cortical GABAergic innervation.* Front Cell Neurosci, 2011. **5**: p. 14.
49. Wonders, C.P. and S.A. Anderson, *The origin and specification of cortical interneurons.* Nat Rev Neurosci, 2006. **7**(9): p. 687-96.
50. Miyoshi, G., et al., *Physiologically distinct temporal cohorts of cortical interneurons arise from telencephalic Olig2-expressing precursors.* J Neurosci, 2007. **27**(29): p. 7786-98.
51. Rymar, V.V. and A.F. Sadikot, *Laminar fate of cortical GABAergic interneurons is dependent on both birthdate and phenotype.* J Comp Neurol, 2007. **501**(3): p. 369-80.
52. Bartolini, G., G. Ciceri, and O. Marin, *Integration of GABAergic interneurons into cortical cell assemblies: lessons from embryos and adults.* Neuron, 2013. **79**(5): p. 849-64.
53. Lim, L., et al., *Development and Functional Diversification of Cortical Interneurons.* Neuron, 2018. **100**(2): p. 294-313.
54. Lodato, S., et al., *Excitatory Projection Neuron Subtypes Control the Distribution of Local Inhibitory Interneurons in the Cerebral Cortex.* Neuron, 2011. **69**(4): p. 763-779.
55. Tasic, B., et al., *Adult mouse cortical cell taxonomy revealed by single cell transcriptomics.* Nature Neuroscience, 2016. **19**(2): p. 335-346.
56. Zeisel, A., et al., *Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq.* Science, 2015. **347**(6226): p. 1138-1142.

57. Hodge, R.D., et al., *Conserved cell types with divergent features in human versus mouse cortex.* Nature, 2019. **573**(7772): p. 61-68.
58. Yao, Z., et al., *An integrated transcriptomic and epigenomic atlas of mouse primary motor cortex cell types.* Biorxiv, 2020: p. 2020.02. 29.970558.
59. Rudy, B., et al., *Three groups of interneurons account for nearly 100% of neocortical GABAergic neurons.* Dev Neurobiol, 2011. **71**(1): p. 45-61.
60. Lee, S., et al., *The largest group of superficial neocortical GABAergic interneurons expresses ionotropic serotonin receptors.* J Neurosci, 2010. **30**(50): p. 16796-808.
61. Taniguchi, H., J. Lu, and Z.J. Huang, *The spatial and temporal origin of chandelier cells in mouse neocortex.* Science, 2013. **339**(6115): p. 70-4.
62. Baude, A., et al., *Immunoreactivity for the GABAA receptor alpha1 subunit, somatostatin and Connexin36 distinguishes axoaxonic, basket, and bistratified interneurons of the rat hippocampus.* Cereb Cortex, 2007. **17**(9): p. 2094-107.
63. Llorca, A. and R. Deogracias, *Origin, Development, and Synaptogenesis of Cortical Interneurons.* Frontiers in Neuroscience, 2022. **16**.
64. Wang, X., et al., *Genetic Single Neuron Anatomy Reveals Fine Granularity of Cortical Axo-Axonic Cells.* Cell Reports, 2019. **26**(11): p. 3145-3159.e5.
65. Xu, H., et al., *Neocortical somatostatin-expressing GABAergic interneurons disinhibit the thalamorecipient layer 4.* Neuron, 2013. **77**(1): p. 155-67.
66. Yavorska, I. and M. Wehr, *Somatostatin-Expressing Inhibitory Interneurons in Cortical Circuits.* Front Neural Circuits, 2016. **10**: p. 76.
67. Urban-Ciecko, J. and A.L. Barth, *Somatostatin-expressing neurons in cortical networks.* Nat Rev Neurosci, 2016. **17**(7): p. 401-9.
68. Nigro, M.J., Y. Hashikawa-Yamasaki, and B. Rudy, *Diversity and Connectivity of Layer 5 Somatostatin-Expressing Interneurons in the Mouse Barrel Cortex.* J Neurosci, 2018. **38**(7): p. 1622-1633.
69. Fishell, G. and A. Kepecs, *Interneuron Types as Attractors and Controllers.* Annu Rev Neurosci, 2020. **43**: p. 1-30.
70. Mayer, C., et al., *Developmental diversification of cortical inhibitory interneurons.* Nature, 2018. **555**(7697): p. 457-462.
71. Waddington, C.H., *The strategy of the genes*. 1957: Routledge.
72. Slack, J.M., *Conrad Hal Waddington: the last renaissance biologist?* Nature Reviews Genetics, 2002. **3**(11): p. 889-895.
73. Quattrocolo, G., G. Fishell, and T.J. Petros, *Heterotopic Transplantations Reveal Environmental Influences on Interneuron Diversity and Maturation.* Cell Rep, 2017. **21**(3): p. 721-731.
74. Southwell, D.G., et al., *Intrinsically determined cell death of developing cortical interneurons.* Nature, 2012. **491**(7422): p. 109-13.
75. Wong, F.K., et al., *Pyramidal cell regulation of interneuron survival sculpts cortical networks.* Nature, 2018. **557**(7707): p. 668-673.
76. Ishino, Y., et al., *Regional Cellular Environment Shapes Phenotypic Variations of Hippocampal and Neocortical Chandelier Cells.* J Neurosci, 2017. **37**(41): p. 9901-9916.
77. Sultan, K.T., et al., *Progressive divisions of multipotent neural progenitors generate late-born chandelier cells in the neocortex.* Nat Commun, 2018. **9**(1): p. 4595.
78. Pla, R., et al., *Layer Acquisition by Cortical GABAergic Interneurons Is Independent of Reelin Signaling.* The Journal of Neuroscience, 2006. **26**(26): p. 6924-6934.
79. Hammond, V., et al., *Layer Positioning of Late-Born Cortical Interneurons Is Dependent on Reelin But Not p35 Signaling.* The Journal of Neuroscience, 2006. **26**(5): p. 1646-1655.
80. Förster, E., S. Zhao, and M. Frotscher, *Laminating the hippocampus.* Nature Reviews Neuroscience, 2006. **7**(4): p. 259-268.
81. Shen, Q., et al., *The timing of cortical neurogenesis is encoded within lineages of individual progenitor cells.* Nat Neurosci, 2006. **9**(6): p. 743-51.
82. Kohwi, M. and C.Q. Doe, *Temporal fate specification and neural progenitor competence during development.* Nat Rev Neurosci, 2013. **14**(12): p. 823-38.
83. Kolk, S.M. and P. Rakic, *Development of prefrontal cortex.* Neuropsychopharmacology, 2022. **47**(1): p. 41-57.
84. de Frutos, C.A., et al., *Reallocation of Olfactory Cajal-Retzius Cells Shapes Neocortex Architecture.* Neuron, 2016. **92**(2): p. 435-448.

85. Borrell, V. and O. Marín, *Meninges control tangential migration of hem-derived Cajal-Retzius cells via CXCL12/CXCR4 signaling.* Nature neuroscience, 2006. **9**(10): p. 1284-1293.
86. Kadoshima, T., et al., *Self-organization of axial polarity, inside-out layer pattern, and species-specific progenitor dynamics in human ES cell-derived neocortex.* Proc Natl Acad Sci U S A, 2013. **110**(50): p. 20284-9.
87. Hatanaka, Y., et al., *Semaphorin 6A–Plexin A2/A4 Interactions with Radial Glia Regulate Migration Termination of Superficial Layer Cortical Neurons.* iScience, 2019. **21**: p. 359-374.
88. Caviness Jr, V. *Reeler mutant mouse: a genetic experiment in developing mammalian cortex*. in *Soc Neurosci Symp*. 1977.
89. Dekimoto, H., T. Terashima, and Y. Katsuyama, *Dispersion of the neurons expressing layer specific markers in the reeler brain.* Development, Growth & Differentiation, 2010. **52**(2): p. 181-193.
90. Hevner, R.F., et al., *Postnatal shifts of interneuron position in the neocortex of normal and reeler mice: evidence for inward radial migration.* Neuroscience, 2004. **124**(3): p. 605-18.
91. Katsuyama, Y. and T. Terashima, *Developmental anatomy of reeler mutant mouse.* Dev Growth Differ, 2009. **51**(3): p. 271-86.
92. Di Donato, V., et al., *An Attractive Reelin Gradient Establishes Synaptic Lamination in the Vertebrate Visual System.* Neuron, 2018. **97**(5): p. 1049-1062 e6.
93. Chai, X., et al., *Reelin and cofilin cooperate during the migration of cortical neurons: a quantitative morphological analysis.* Development, 2016. **143**(6): p. 1029-40.
94. Rakic, P., *Specification of Cerebral Cortical Areas.* Science, 1988. **241**(4862): p. 170-176.
95. Vogt, C. and O. Vogt, *Allgemeine ergebnisse unserer hirnforschung*. Vol. 25. 1919: JA Barth.
96. Harris, J.A., et al., *Hierarchical organization of cortical and thalamic connectivity.* Nature, 2019. **575**(7781): p. 195-202.
97. Markram, H., et al., *Reconstruction and Simulation of Neocortical Microcircuitry.* Cell, 2015. **163**(2): p. 456-92.
98. Bastos, A.M., et al., *Canonical microcircuits for predictive coding.* Neuron, 2012. **76**(4): p. 695-711.
99. Beul, S.F. and C.C. Hilgetag, *Towards a "canonical" agranular cortical microcircuit.* Frontiers in Neuroanatomy, 2015. **8**.
100. Schuman, B., et al., *Four Unique Interneuron Populations Reside in Neocortical Layer 1.* J Neurosci, 2019. **39**(1): p. 125-139.
101. Staiger, J.F. and C.C.H. Petersen, *Neuronal Circuits in Barrel Cortex for Whisker Sensory Perception.* Physiological Reviews, 2021. **101**(1): p. 353-415.
102. Schuman, B., et al., *Neocortical Layer 1: An Elegant Solution to Top-Down and Bottom-Up Integration.* Annual Review of Neuroscience, 2021. **44**(Volume 44, 2021): p. 221-252.
103. Matho, K.S., et al., *Genetic dissection of the glutamatergic neuron system in cerebral cortex.* Nature, 2021. **598**(7879): p. 182-187.
104. Scala, F., et al., *Layer 4 of mouse neocortex differs in cell types and circuit organization between sensory areas.* Nature Communications, 2019. **10**(1): p. 4174.
105. Scala, F., et al., *Neocortical layer 4 in adult mouse differs in major cell types and circuit organization between primary sensory areas.* bioRxiv, 2019.
106. Briggs, F., *Organizing principles of cortical layer 6.* Frontiers in Neural Circuits, 2010. **4**.
107. Andjelic, S., et al., *Glutamatergic nonpyramidal neurons from neocortical layer VI and their comparison with pyramidal and spiny stellate neurons.* J Neurophysiol, 2009. **101**(2): p. 641-54.
108. Sóki, N., et al., *Investigation of synapses in the cortical white matter in human temporal lobe epilepsy.* Brain Research, 2022. **1779**: p. 147787.
109. Frazer, S., et al., *Transcriptomic and anatomic parcellation of 5-HT(3A)R expressing cortical interneuron subtypes revealed by single-cell RNA sequencing.* Nat Commun, 2017. **8**: p. 14219.
110. Banovac, I., et al., *The Distinct Characteristics of Somatostatin Neurons in the Human Brain.* Molecular Neurobiology, 2022. **59**(8): p. 4953-4965.
111. Tomioka, R., et al., *Demonstration of long-range GABAergic connections distributed throughout the mouse neocortex.* European Journal of Neuroscience, 2005. **21**(6): p. 1587-1600.
112. Jones, E.G. and T.P.S. Powell, *An electron microscopic study of the laminar pattern and mode of termination of afferent fibre pathways in the somatic sensory cortex of the cat.* Philosophical Transactions of the Royal Society of London. B, Biological Sciences, 1970. **257**(812): p. 45-62.
113. Vogt, B.A., *The role of layer I in cortical function*, in *Normal and altered states of function*. 1991, Springer. p. 49-80.

114. Cruikshank, S.J., et al., *Thalamic Control of Layer 1 Circuits in Prefrontal Cortex.* The Journal of Neuroscience, 2012. **32**(49): p. 17813-17823.
115. Roth, M.M., et al., *Thalamic nuclei convey diverse contextual information to layer 1 of visual cortex.* Nature neuroscience, 2016. **19**(2): p. 299-307.
116. Daigle, T.L., et al., *A Suite of Transgenic Driver and Reporter Mouse Lines with Enhanced Brain-Cell-Type Targeting and Functionality.* Cell, 2018. **174**(2): p. 465-480 e22.
117. Ciceri, G., et al., *Lineage-specific laminar organization of cortical GABAergic interneurons.* Nat Neurosci, 2013. **16**(9): p. 1199-210.
118. Yabut, O., et al., *Abnormal laminar position and dendrite development of interneurons in the reeler forebrain.* Brain Res, 2007. **1140**: p. 75-83.
119. Kirkcaldie, M.T.K., *Chapter 4 - Neocortex*, in *The Mouse Nervous System*, C. Watson, G. Paxinos, and L. Puelles, Editors. 2012, Academic Press: San Diego. p. 52-111.
120. Kang, H.J., et al., *Spatio-temporal transcriptome of the human brain.* Nature, 2011. **478**(7370): p. 483-489.
121. Cadwell, C.R., et al., *Development and Arealization of the Cerebral Cortex.* Neuron, 2019. **103**(6): p. 980-1004.
122. O'Leary, D.D., S.-J. Chou, and S. Sahara, *Area patterning of the mammalian cortex.* Neuron, 2007. **56**(2): p. 252-269.
123. Bishop, K.M., J.L.R. Rubenstein, and D.D.M. O'Leary, *Distinct Actions of Emx1, Emx2, andPax6 in Regulating the Specification of Areas in the Developing Neocortex.* The Journal of Neuroscience, 2002. **22**(17): p. 7627-7638.
124. Borello, U. and A. Pierani, *Patterning the cerebral cortex: traveling with morphogens.* Current Opinion in Genetics & Development, 2010. **20**(4): p. 408-415.
125. Nakagawa, Y., J.E. Johnson, and D.D. O'Leary, *Graded and areal expression patterns of regulatory genes and cadherins in embryonic neocortex independent of thalamocortical input.* Journal of Neuroscience, 1999. **19**(24): p. 10877-10885.
126. Bishop, K.M., G. Goudreau, and D.D.M. O'Leary, *Regulation of Area Identity in the Mammalian Neocortex by Emx2 and Pax6.* Science, 2000. **288**(5464): p. 344-349.
127. Sperry, R.W., *Chemoaffinity in the Orderly Growth of Nerve Fiber Patterns and Connections.* Proc Natl Acad Sci U S A, 1963. **50**(4): p. 703-10.
128. Zembrzycki, A., et al., *Genetic interplay between the transcription factors Sp8 and Emx2 in the patterning of the forebrain.* Neural development, 2007. **2**: p. 1-18.
129. Sahara, S., et al., *Sp8 exhibits reciprocal induction with Fgf8 but has an opposing effect on anterior-posterior cortical area patterning.* Neural development, 2007. **2**(1): p. 1-22.
130. Wagener, R.J., et al., *The Somatosensory Cortex of reeler Mutant Mice Shows Absent Layering But Intact Formation and Behavioral Activation of Columnar Somatotopic Maps.* The Journal of Neuroscience, 2010. **30**(46): p. 15700-15709.
131. Wagener, R.J., et al., *Thalamocortical Connections Drive Intracortical Activation of Functional Columns in the Mislaminated Reeler Somatosensory Cortex.* Cerebral Cortex, 2015. **26**(2): p. 820-837.
132. Fukuchi-Shimogori, T. and E.A. Grove, *Neocortex patterning by the secreted signaling molecule FGF8.* Science, 2001. **294**(5544): p. 1071-1074.
133. Schlaggar, B.L. and D.D.M. O'Leary, *Potential of Visual Cortex to Develop an Array of Functional Units Unique to Somatosensory Cortex.* Science, 1991. **252**(5012): p. 1556-1560.
134. Huffman, K.J., et al., *Formation of Cortical Fields on a Reduced Cortical Sheet.* The Journal of Neuroscience, 1999. **19**(22): p. 9939-9952.
135. Dufour, A., et al., *Area Specificity and Topography of Thalamocortical Projections Are Controlled by ephrin/Eph Genes.* Neuron, 2003. **39**(3): p. 453-465.
136. Bulfone, A., et al., *An olfactory sensory map develops in the absence of normal projection neurons or GABAergic interneurons.* Neuron, 1998. **21**(6): p. 1273-1282.
137. Heck, N., et al., *Activity-dependent regulation of neuronal apoptosis in neonatal mouse cerebral cortex.* Cerebral cortex, 2008. **18**(6): p. 1335-1349.
138. Verney, C., et al., *Independent Controls for Neocortical Neuron Production and Histogenetic Cell Death.* Developmental Neuroscience, 2000. **22**(1-2): p. 125-138.
139. Welker, E., et al., *Altered sensory processing in the somatosensory cortex of the mouse mutant barrelless.* Science, 1996. **271**(5257): p. 1864-7.
140. Ravary, A., et al., *Adenylate Cyclase 1 as a Key Actor in the Refinement of Retinal Projection Maps.* The Journal of Neuroscience, 2003. **23**(6): p. 2228-2238.

141. Gheorghita, F., et al., *Structural Basis for Map Formation in the Thalamocortical Pathway of the Barrelless Mouse.* The Journal of Neuroscience, 2006. **26**(39): p. 10057-10067.
142. Lu, H.-C., et al., *Role of Efficient Neurotransmitter Release in Barrel Map Development.* The Journal of Neuroscience, 2006. **26**(10): p. 2692-2703.
143. Dhande, O.S., et al., *Role of adenylate cyclase 1 in retinofugal map development.* Journal of Comparative Neurology, 2012. **520**(7): p. 1562-1583.
144. Verhage, M., et al., *Synaptic Assembly of the Brain in the Absence of Neurotransmitter Secretion.* Science, 2000. **287**(5454): p. 864-869.
145. Varoqueaux, F., et al., *Total arrest of spontaneous and evoked synaptic transmission but normal synaptogenesis in the absence of Munc13-mediated vesicle priming.* Proceedings of the National Academy of Sciences, 2002. **99**(13): p. 9037-9042.
146. Li, H., et al., *Laminar and Columnar Development of Barrel Cortex Relies on Thalamocortical Neurotransmission.* Neuron, 2013. **79**(5): p. 970-986.
147. Sharma, J., A. Angelucci, and M. Sur, *Induction of visual orientation modules in auditory cortex.* Nature, 2000. **404**(6780): p. 841-847.
148. Vue, T.Y., et al., *Thalamic Control of Neocortical Area Formation in Mice.* The Journal of Neuroscience, 2013. **33**(19): p. 8442-8453.
149. Guy, J. and J.F. Staiger, *The Functioning of a Cortex without Layers.* Front Neuroanat, 2017. **11**: p. 54.
150. Guy, J., et al., *Persistence of Functional Sensory Maps in the Absence of Cortical Layers in the Somsatosensory Cortex of Reeler Mice.* Cerebral Cortex, 2014. **25**(9): p. 2517-2528.
151. Dityatev, A., C.I. Seidenbecher, and M. Schachner, *Compartmentalization from the outside: the extracellular matrix and functional microdomains in the brain.* Trends in Neurosciences, 2010. **33**(11): p. 503-512.
152. Chia, P.H., P. Li, and K. Shen, *Cell biology in neuroscience: cellular and molecular mechanisms underlying presynapse formation.* J Cell Biol, 2013. **203**(1): p. 11-22.
153. Koch, C. and G. Laurent, *Complexity and the nervous system.* Science, 1999. **284**(5411): p. 96-98.
154. Weng, G., U.S. Bhalla, and R. Iyengar, *Complexity in biological signaling systems.* Science, 1999. **284**(5411): p. 92-96.
155. Lai, H.C. and L.Y. Jan, *The distribution and targeting of neuronal voltage-gated ion channels.* Nat Rev Neurosci, 2006. **7**(7): p. 548-62.
156. Trimmer, J.S., *Subcellular localization of K+ channels in mammalian brain neurons: remarkable precision in the midst of extraordinary complexity.* Neuron, 2015. **85**(2): p. 238-56.
157. Donato, A., et al., *Neuronal sub-compartmentalization: a strategy to optimize neuronal function.* Biol Rev Camb Philos Soc, 2019. **94**(3): p. 1023-1037.
158. Lyon, A.S., W.B. Peeples, and M.K. Rosen, *A framework for understanding the functions of biomolecular condensates across scales.* Nature Reviews Molecular Cell Biology, 2021. **22**(3): p. 215-235.
159. Diekmann, Y. and José B. Pereira-Leal, *Evolution of intracellular compartmentalization.* Biochemical Journal, 2012. **449**(2): p. 319-331.
160. Honigmann, A. and A. Pralle, *Compartmentalization of the Cell Membrane.* Journal of Molecular Biology, 2016. **428**(24, Part A): p. 4739-4748.
161. Gu, X., C. Jia, and J. Wang, *Advances in Understanding the Molecular Mechanisms of Neuronal Polarity.* Molecular Neurobiology, 2023. **60**(5): p. 2851-2870.
162. Rasband, M.N., *The axon initial segment and the maintenance of neuronal polarity.* Nat Rev Neurosci, 2010. **11**(8): p. 552-62.
163. Spruston, N., *Pyramidal neurons: dendritic structure and synaptic integration.* Nat Rev Neurosci, 2008. **9**(3): p. 206-21.
164. Tononi, G., O. Sporns, and G.M. Edelman, *A complexity measure for selective matching of signals by the brain.* Proceedings of the National Academy of Sciences, 1996. **93**(8): p. 3422-3427.
165. Edelman, G.M. and J.A. Gally, *Degeneracy and complexity in biological systems.* Proceedings of the National Academy of Sciences, 2001. **98**(24): p. 13763-13768.
166. Shepherd, G.M., *Neuron Doctrine: Historical Background*, in *Encyclopedia of Neuroscience*, L.R. Squire, Editor. 2009, Academic Press: Oxford. p. 691-695.
167. Nelson, A.D. and P.M. Jenkins, *Axonal Membranes and Their Domains: Assembly and Function of the Axon Initial Segment and Node of Ranvier.* Front Cell Neurosci, 2017. **11**: p. 136.
168. Gidon, A. and I. Segev, *Principles Governing the Operation of Synaptic Inhibition in Dendrites.* Neuron, 2012. **75**(2): p. 330-341.

169. Bloss, E.B., et al., *Structured Dendritic Inhibition Supports Branch-Selective Integration in CA1 Pyramidal Cells.* Neuron, 2016. **89**(5): p. 1016-30.
170. Kole, M.H. and G.J. Stuart, *Signal processing in the axon initial segment.* Neuron, 2012. **73**(2): p. 235-47.
171. Dzhashiashvili, Y., et al., *Nodes of Ranvier and axon initial segments are ankyrin G-dependent domains that assemble by distinct mechanisms.* J Cell Biol, 2007. **177**(5): p. 857-70.
172. van Beuningen, S.F.B., et al., *TRIM46 Controls Neuronal Polarity and Axon Specification by Driving the Formation of Parallel Microtubule Arrays.* Neuron, 2015. **88**(6): p. 1208-1226.
173. Hedstrom, K.L., et al., *Neurofascin assembles a specialized extracellular matrix at the axon initial segment.* J Cell Biol, 2007. **178**(5): p. 875-86.
174. Ango, F., et al., *Ankyrin-Based Subcellular Gradient of Neurofascin, an Immunoglobulin Family Protein, Directs GABAergic Innervation at Purkinje Axon Initial Segment.* Cell, 2004. **119**(2): p. 257-272.
175. Telley, L., et al., *Dual Function of NRP1 in Axon Guidance and Subcellular Target Recognition in Cerebellum.* Neuron, 2016. **91**(6): p. 1276-1291.
176. Lefebvre, J.L., J.R. Sanes, and J.N. Kay, *Development of Dendritic Form and Function.* Annual Review of Cell and Developmental Biology, 2015. **31**(1): p. 741-777.
177. Somogyi, P., et al., *Salient features of synaptic organisation in the cerebral cortex.* Brain Res Brain Res Rev, 1998. **26**(2-3): p. 113-35.
178. Riedemann, T., *Diversity and Function of Somatostatin-Expressing Interneurons in the Cerebral Cortex.* International Journal of Molecular Sciences, 2019. **20**(12): p. 2952.
179. Kirizs, T., et al., *Distinct axo-somato-dendritic distributions of three potassium channels in CA1 hippocampal pyramidal cells.* Eur J Neurosci, 2014. **39**(11): p. 1771-83.
180. Riquelme, D., et al., *Subcellular Localization and Activity of TRPM4 in Medial Prefrontal Cortex Layer 2/3.* Front Cell Neurosci, 2018. **12**: p. 12.
181. Virga, D.M., et al., *Activity-dependent subcellular compartmentalization of dendritic mitochondria structure in CA1 pyramidal neurons.* bioRxiv, 2023: p. 2023.03.25.534233.
182. Morales, M. and E. Fifkova, *Distribution of MAP 2 in dendritic spines and its colocalization with actin.* Cell and Tissue Research, 1989. **256**(3): p. 447-456.
183. Kosik, K.S. and E.A. Finch, *MAP2 and tau segregate into dendritic and axonal domains after the elaboration of morphologically distinct neurites: an immunocytochemical study of cultured rat cerebrum.* Journal of Neuroscience, 1987. **7**(10): p. 3142-3153.
184. Caceres, A., et al., *Immunocytochemical localization of actin and microtubule-associated protein MAP2 in dendritic spines.* Proceedings of the National Academy of Sciences, 1983. **80**(6): p. 1738-1742.
185. Worth, D.C., et al., *Drebrin contains a cryptic F-actin-bundling activity regulated by Cdk5 phosphorylation.* J Cell Biol, 2013. **202**(5): p. 793-806.
186. Aoki, C., et al., *Drebrin A is a postsynaptic protein that localizes in vivo to the submembranous surface of dendritic sites forming excitatory synapses.* J Comp Neurol, 2005. **483**(4): p. 383-402.
187. Ivanov, A., et al., *Drebrin A regulates dendritic spine plasticity and synaptic function in mature cultured hippocampal neurons.* J Cell Sci, 2009. **122**(Pt 4): p. 524-34.
188. Ketschek, A., et al., *Drebrin coordinates the actin and microtubule cytoskeleton during the initiation of axon collateral branches.* Dev Neurobiol, 2016. **76**(10): p. 1092-110.
189. Efimova, N., et al., *betaIII Spectrin Is Necessary for Formation of the Constricted Neck of Dendritic Spines and Regulation of Synaptic Activity in Neurons.* J Neurosci, 2017. **37**(27): p. 6442-6459.
190. Lim, S.T., et al., *A novel targeting signal for proximal clustering of the Kv2.1 K+ channel in hippocampal neurons.* Neuron, 2000. **25**(2): p. 385-97.
191. Baver, S.B. and K.M. O'Connell, *The C-terminus of neuronal Kv2.1 channels is required for channel localization and targeting but not for NMDA-receptor-mediated regulation of channel function.* Neuroscience, 2012. **217**: p. 56-66.
192. Sánchez-Ponce, D., et al., *Developmental expression of Kv potassium channels at the axon initial segment of cultured hippocampal neurons.* PLoS One, 2012. **7**(10): p. e48557.
193. Kihira, Y., T.O. Hermanstyne, and H. Misonou, *Formation of heteromeric Kv2 channels in mammalian brain neurons.* J Biol Chem, 2010. **285**(20): p. 15048-15055.
194. Jensen, C.S., et al., *Specific sorting and post-Golgi trafficking of dendritic potassium channels in living neurons.* J Biol Chem, 2014. **289**(15): p. 10566-81.
195. Du, J., et al., *The K+ channel, Kv2.1, is apposed to astrocytic processes and is associated with inhibitory postsynaptic membranes in hippocampal and cortical principal neurons and inhibitory interneurons.* Neuroscience, 1998. **84**(1): p. 37-48.

196. Bishop, H.I., et al., *Kv2 Ion Channels Determine the Expression and Localization of the Associated AMIGO-1 Cell Adhesion Molecule in Adult Brain Neurons.* Front Mol Neurosci, 2018. **11**: p. 1.
197. Bi, X.N., et al., *Polarized distribution of alpha 5 integrin in dendrites of hippocampal and cortical neurons.* Journal of Comparative Neurology, 2001. **435**(2): p. 184-193.
198. Tran, T.S., et al., *Secreted semaphorins control spine distribution and morphogenesis in the postnatal CNS.* Nature, 2009. **462**(7276): p. 1065-9.
199. Nishimura-Akiyoshi, S., et al., *Axonal netrin-Gs transneuronally determine lamina-specific subdendritic segments.* Proc Natl Acad Sci U S A, 2007. **104**(37): p. 14801-6.
200. Kupferman, J.V., et al., *Reelin signaling specifies the molecular identity of the pyramidal neuron distal dendritic compartment.* Cell, 2014. **158**(6): p. 1335-1347.
201. Meseke, M., et al., *Distal Dendritic Enrichment of HCN1 Channels in Hippocampal CA1 Is Promoted by Estrogen, but Does Not Require Reelin.* eNeuro, 2018. **5**(5).
202. Rivera, J.F., et al., *An evolutionarily conserved dileucine motif in Shal K+ channels mediates dendritic targeting.* Nat Neurosci, 2003. **6**(3): p. 243-50.
203. Jensen, C.S., et al., *Specific sorting and post-Golgi trafficking of dendritic potassium channels in living neurons.* J Biol Chem, 2014. **289**(15): p. 10566-10581.
204. Rhodes, K.J., et al., *KChIPs and Kv4 alpha subunits as integral components of A-type potassium channels in mammalian brain.* J Neurosci, 2004. **24**(36): p. 7903-15.
205. Chai, X., et al., *Reelin Stabilizes the Actin Cytoskeleton of Neuronal Processes by Inducing n-Cofilin Phosphorylation at Serine3.* The Journal of Neuroscience, 2009. **29**(1): p. 288-299.
206. Stanco, A., et al., *Netrin-1–α3β1 integrin interactions regulate the migration of interneurons through the cortical marginal zone.* Proceedings of the National Academy of Sciences, 2009. **106**(18): p. 7595-7600.
207. Munakata, H., et al., *Distribution and densitometry mapping of L1-CAM immunoreactivity in the adult mouse brain - light microscopic observation.* Bmc Neuroscience, 2003. **4**.
208. Kelly, E.A., et al., *Subcellular localization of intercellular adhesion molecule-5 (telencephalin) in the visual cortex is not developmentally regulated in the absence of matrix metalloproteinase-9.* Journal of Comparative Neurology, 2014. **522**(3): p. 676-688.
209. Horton, A.C., J.J. Yi, and M.D. Ehlers, *Cell type-specific dendritic polarity in the absence of spatially organized external cues.* Brain Cell Biol, 2006. **35**(1): p. 29-38.
210. Wu, Y.K., K. Fujishima, and M. Kengaku, *Differentiation of Apical and Basal Dendrites in Pyramidal Cells and Granule Cells in Dissociated Hippocampal Cultures.* PLOS ONE, 2015. **10**(2): p. e0118482.
211. Bruckner, G., et al., *Perineuronal nets provide a polyanionic, glia-associated form of microenvironment around certain neurons in many parts of the rat brain.* Glia, 1993. **8**(3): p. 183-200.
212. John, N., et al., *Brevican-containing perineuronal nets of extracellular matrix in dissociated hippocampal primary cultures.* Mol Cell Neurosci, 2006. **31**(4): p. 774-84.
213. Kim, N., et al., *Lrp4 is a receptor for Agrin and forms a complex with MuSK.* Cell, 2008. **135**(2): p. 334-42.
214. Zhang, H.L. and H.B. Peng, *Mechanism of acetylcholine receptor cluster formation induced by DC electric field.* PLoS One, 2011. **6**(10): p. e26805.
215. Orida, N. and M.-m. Poo, *Electrophoretic movement and localization of acetylcholine receptors in the embryonic muscle cell membrane.* Nature, 1978. **275**(5675): p. 31-35.
216. Chen, S.-Y. and H.-J. Cheng, *Functions of axon guidance molecules in synapse formation.* Current Opinion in Neurobiology, 2009. **19**(5): p. 471-478.
217. Fraser, S.E. and D.H. Perkel, *Competitive and positional cues in the patterning of nerve connections.* Journal of neurobiology, 1990. **21**(1): p. 51-72.
218. Gierer, A., *Model for the retino-tectal projection.* Proc R Soc Lond B Biol Sci, 1983. **218**(1210): p. 77-93.
219. Löschinger, J., F. Weth, and F. Bonhoeffer, *Reading of concentration gradients by axonal growth cones.* Philos Trans R Soc Lond B Biol Sci, 2000. **355**(1399): p. 971-82.
220. Kania, A. and R. Klein, *Mechanisms of ephrin–Eph signalling in development, physiology and disease.* Nature Reviews Molecular Cell Biology, 2016. **17**(4): p. 240-256.
221. Barallobre, M.J., et al., *Aberrant development of hippocampal circuits and altered neural activity in netrin 1-deficient mice.* Development, 2000. **127**(22): p. 4797-4810.
222. Chédotal, A., et al., *Semaphorins III and IV repel hippocampal axons via two distinct receptors.* Development, 1998. **125**(21): p. 4313-4323.
223. Polleux, F., et al., *Patterning of Cortical Efferent Projections by Semaphorin-Neuropilin Interactions.* Science, 1998. **282**(5395): p. 1904-1906.

224. Polleux, F., T. Morrow, and A. Ghosh, *Semaphorin 3A is a chemoattractant for cortical apical dendrites.* Nature, 2000. **404**(6778): p. 567-573.
225. Ba-Charvet, K.T.N., et al., *Slit2-mediated chemorepulsion and collapse of developing forebrain axons.* Neuron, 1999. **22**(3): p. 463-473.
226. Blockus, H. and A. Chedotal, *Slit-Robo signaling.* Development, 2016. **143**(17): p. 3037-44.
227. Francis, K. and B.O. Palsson, *Effective intercellular communication distances are determined by the relative time constants for cyto/chemokine secretion and diffusion.* Proceedings of the National Academy of Sciences, 1997. **94**(23): p. 12258-12262.
228. Yan, C., et al., *Automated and Accurate Detection of Soma Location and Surface Morphology in Large-Scale 3D Neuron Images.* PLOS ONE, 2013. **8**(4): p. e62579.
229. Boyer, N.P. and S.L. Gupton, *Revisiting Netrin-1: One Who Guides (Axons).* Frontiers in Cellular Neuroscience, 2018. **12**.
230. Blockus, H., et al., *Synaptogenic activity of the axon guidance molecule Robo2 underlies hippocampal circuit function.* Cell Reports, 2021. **37**(3): p. 109828.
231. Suto, F., et al., *Plexin-a4 mediates axon-repulsive activities of both secreted and transmembrane semaphorins and plays roles in nerve fiber guidance.* Journal of Neuroscience, 2005. **25**(14): p. 3628-3637.
232. Suto, F., et al., *Interactions between plexin-A2, plexin-A4, and semaphorin 6A control lamina-restricted projection of hippocampal mossy fibers.* Neuron, 2007. **53**(4): p. 535-47.
233. Tawarayama, H., et al., *Roles of Semaphorin-6B and Plexin-A2 in Lamina-Restricted Projection of Hippocampal Mossy Fibers.* The Journal of Neuroscience, 2010. **30**(20): p. 7049-7060.
234. Matsuoka, R.L., et al., *Transmembrane semaphorin signalling controls laminar stratification in the mammalian retina.* Nature, 2011. **470**(7333): p. 259-263.
235. Alto, L.T. and J.R. Terman, *Semaphorins and their Signaling Mechanisms*, in *Semaphorin Signaling: Methods and Protocols*, J.R. Terman, Editor. 2017, Springer New York: New York, NY. p. 1-25.
236. Soto, F., et al., *NGL-2 Regulates Pathway-Specific Neurite Growth and Lamination, Synapse Formation, and Signal Transmission in the Retina.* The Journal of Neuroscience, 2013. **33**(29): p. 11949-11959.
237. Xu, S., et al., *Interactions between the Ig-Superfamily Proteins DIP-α and Dpr6/10 Regulate Assembly of Neural Circuits.* Neuron, 2018. **100**(6): p. 1369-1384.e6.
238. Ashley, J., et al., *Transsynaptic interactions between IgSF proteins DIP-α and Dpr10 are required for motor neuron targeting specificity.* eLife, 2019. **8**: p. e42690.
239. Sergeeva, A.P., et al., *DIP/Dpr interactions and the evolutionary design of specificity in protein families.* Nature Communications, 2020. **11**(1): p. 2125.
240. Demyanenko, G.P. and P.F. Maness, *The L1 cell adhesion molecule is essential for topographic mapping of retinal axons.* Journal of Neuroscience, 2003. **23**(2): p. 530-538.
241. Buhusi, M., et al., *ALCAM regulates mediolateral retinotopic mapping in the superior colliculus.* Journal of Neuroscience, 2009. **29**(50): p. 15630-15641.
242. Heckman, E.L. and C.Q. Doe, *Establishment and Maintenance of Neural Circuit Architecture.* The Journal of Neuroscience, 2021. **41**(6): p. 1119-1129.
243. Williams, M.E., J. de Wit, and A. Ghosh, *Molecular mechanisms of synaptic specificity in developing neural circuits.* Neuron, 2010. **68**(1): p. 9-18.
244. Williams, S.E., et al., *Ephrin-B2 and EphB1 mediate retinal axon divergence at the optic chiasm.* Neuron, 2003. **39**(6): p. 919-935.
245. García-Frigola, C., et al., *Zic2 promotes axonal divergence at the optic chiasm midline by EphB1-dependent and-independent mechanisms.* 2008.
246. Herrera, E.s., et al., *Zic2 patterns binocular vision by specifying the uncrossed retinal projection.* Cell, 2003. **114**(5): p. 545-557.
247. Petros, T.J., B.R. Shrestha, and C. Mason, *Specificity and sufficiency of EphB1 in driving the ipsilateral retinal projection.* Journal of Neuroscience, 2009. **29**(11): p. 3463-3474.
248. Stein, B.E., et al., *The superior colliculus and visual thalamus*, in *Neuroscience in the 21st Century: From Basic to Clinical*. 2022, Springer. p. 1029-1054.
249. Hansen, M.J., G.E. Dallal, and J.G. Flanagan, *Retinal axon response to ephrin-as shows a graded, concentration-dependent transition from growth promotion to inhibition.* Neuron, 2004. **42**(5): p. 717-730.
250. Hindges, R., et al., *EphB forward signaling controls directional branch extension and arborization required for dorsal-ventral retinotopic mapping.* Neuron, 2002. **35**(3): p. 475-487.
251. Suetterlin, P. and U. Drescher, *Target-independent ephrina/EphA-mediated axon-axon repulsion as a novel element in retinocollicular mapping.* Neuron, 2014. **84**(4): p. 740-752.

252. Thakar, S., G. Chenaux, and M. Henkemeyer, *Critical roles for EphB and ephrin-B bidirectional signalling in retinocollicular mapping.* Nature communications, 2011. **2**(1): p. 431.
253. Hornberger, M.R., et al., *Modulation of EphA receptor function by coexpressed ephrinA ligands on retinal ganglion cell axons.* Neuron, 1999. **22**(4): p. 731-742.
254. Ango, F., et al., *Bergmann Glia and the Recognition Molecule CHL1 Organize GABAergic Axons and Direct Innervation of Purkinje Cell Dendrites.* PLOS Biology, 2008. **6**(4): p. e103.
255. Chao, D.L., L. Ma, and K. Shen, *Transient cell–cell interactions in neural circuit formation.* Nature Reviews Neuroscience, 2009. **10**(4): p. 262-271.
256. Rakic, P., *Neuron-glia relationship during granule cell migration in developing cerebellar cortex. A Golgi and electonmicroscopic study in Macacus rhesus.* Journal of Comparative Neurology, 1971. **141**(3): p. 283-312.
257. Nishiyama, H., *Dendritic translocation of climbing fibers: a new face of old phenomenon.* Cerebellum, 2015. **14**(1): p. 1-3.
258. Ichikawa, R., et al., *Territories of heterologous inputs onto Purkinje cell dendrites are segregated by mGluR1-dependent parallel fiber synapse elimination.* Proceedings of the National Academy of Sciences, 2016. **113**(8): p. 2282-2287.
259. Hioki, H., et al., *Cell type-specific inhibitory inputs to dendritic and somatic compartments of parvalbumin-expressing neocortical interneuron.* J Neurosci, 2013. **33**(2): p. 544-55.
260. Omiya, Y., et al., *VGluT3-expressing CCK-positive basket cells construct invaginating synapses enriched with endocannabinoid signaling proteins in particular cortical and cortex-like amygdaloid regions of mouse brains.* J Neurosci, 2015. **35**(10): p. 4215-28.
261. Bartos, M. and C. Elgueta, *Functional characteristics of parvalbumin- and cholecystokinin-expressing basket cells.* J Physiol, 2012. **590**(4): p. 669-81.
262. Williams, M.E., et al., *Cadherin-9 regulates synapse-specific differentiation in the developing hippocampus.* Neuron, 2011. **71**(4): p. 640-55.
263. Bodian, D. and N. Taylor, *Synapse Arising at Central Node of Ranvier, and Note on Fixation of the Central Nervous System.* Science, 1963. **139**(3552): p. 330-332.
264. Andres, K.H., *Über die Feinstruktur besonderer Einrichtungen in markhaltigen Nervenfasern des Kleinhirns der Ratte.* Zeitschrift für Zellforschung und Mikroskopische Anatomie, 1965. **65**(5): p. 701-712.
265. Khattab, F.I., *Synaptic Structures at Nodes of Ranvier in Spinal Cords of Mice.* Nature, 1967. **216**(5114): p. 496-497.
266. Betley, J.N., et al., *Stringent specificity in the construction of a GABAergic presynaptic inhibitory circuit.* Cell, 2009. **139**(1): p. 161-74.
267. Beatty, J.A. and C.L. Cox, *The multifaceted role of inhibitory interneurons in the dorsal lateral geniculate nucleus.* Visual Neuroscience, 2017. **34**: p. E017.
268. Südhof, T.C., *The cell biology of synapse formation.* Journal of Cell Biology, 2021. **220**(7).
269. Boyken, J., et al., *Molecular profiling of synaptic vesicle docking sites reveals novel proteins but few differences between glutamatergic and GABAergic synapses.* Neuron, 2013. **78**(2): p. 285-97.
270. O'Rourke, N.A., et al., *Deep molecular diversity of mammalian synapses: why it matters and how to measure it.* Nat Rev Neurosci, 2012. **13**(6): p. 365-79.
271. Hendry, S.H., et al., *Two classes of cortical GABA neurons defined by differential calcium binding protein immunoreactivities.* Exp Brain Res, 1989. **76**(2): p. 467-72.
272. Sohal, V.S., et al., *Parvalbumin neurons and gamma rhythms enhance cortical circuit performance.* Nature, 2009. **459**(7247): p. 698-702.
273. Varoqueaux, F., S. Jamain, and N. Brose, *Neuroligin 2 is exclusively localized to inhibitory synapses.* Eur J Cell Biol, 2004. **83**(9): p. 449-56.
274. Dresbach, T., et al., *Assembly of active zone precursor vesicles: obligatory trafficking of presynaptic cytomatrix proteins Bassoon and Piccolo via a trans-Golgi compartment.* J Biol Chem, 2006. **281**(9): p. 6038-47.
275. Terry-Lorenzo, R.T., et al., *Trio, a Rho Family GEF, Interacts with the Presynaptic Active Zone Proteins Piccolo and Bassoon.* PLoS One, 2016. **11**(12): p. e0167535.
276. Takács, V.T., et al., *Co-transmission of acetylcholine and GABA regulates hippocampal states.* Nature Communications, 2018. **9**(1): p. 2848.
277. Granger, A.J., et al., *Cortical ChAT+ neurons co-transmit acetylcholine and GABA in a target- and brain-region-specific manner.* eLife, 2020. **9**: p. e57749.

278. Root, D.H., et al., *Single rodent mesohabenular axons release glutamate and GABA.* Nature Neuroscience, 2014. **17**(11): p. 1543-1551.
279. Root, D.H., et al., *Selective Brain Distribution and Distinctive Synaptic Architecture of Dual Glutamatergic-GABAergic Neurons.* Cell Reports, 2018. **23**(12): p. 3465-3479.
280. Moore, N.J., et al., *Synaptic Connectivity between Renshaw Cells and Motoneurons in the Recurrent Inhibitory Circuit of the Spinal Cord.* The Journal of Neuroscience, 2015. **35**(40): p. 13673-13686.
281. Bailey, C.H., E.R. Kandel, and K.M. Harris, *Structural components of synaptic plasticity and memory consolidation.* Cold Spring Harbor perspectives in biology, 2015. **7**(7): p. a021758.
282. Bailey, C.H. and M. Chen, *Morphological basis of long-term habituation and sensitization in Aplysia.* Science, 1983. **220**(4592): p. 91-93.
283. Bourne, J.N. and K.M. Harris, *Coordination of size and number of excitatory and inhibitory synapses results in a balanced structural plasticity along mature hippocampal CA1 dendrites during LTP.* Hippocampus, 2011. **21**(4): p. 354-73.
284. Bailey, C.H. and M. Chen, *Morphological basis of short-term habituation in Aplysia.* Journal of Neuroscience, 1988. **8**(7): p. 2452-2459.
285. Craig, A.M. and H. Boudin, *Molecular heterogeneity of central synapses: afferent and target regulation.* Nat Neurosci, 2001. **4**(6): p. 569-78.
286. Blatow, M., et al., *A novel network of multipolar bursting interneurons generates theta frequency oscillations in neocortex.* Neuron, 2003. **38**(5): p. 805-17.
287. Pang, Z.P., et al., *Synaptotagmin-2 is essential for survival and contributes to Ca2+ triggering of neurotransmitter release in central and neuromuscular synapses.* J Neurosci, 2006. **26**(52): p. 13493-504.
288. Bouhours, B., et al., *Synaptotagmin2 (Syt2) Drives Fast Release Redundantly with Syt1 at the Output Synapses of Parvalbumin-Expressing Inhibitory Neurons.* J Neurosci, 2017. **37**(17): p. 4604-4617.
289. Chen, C., et al., *Synaptotagmin 2 Is the Fast Ca(2+) Sensor at a Central Inhibitory Synapse.* Cell Rep, 2017. **18**(3): p. 723-736.
290. Kerr, A.M., E. Reisinger, and P. Jonas, *Differential dependence of phasic transmitter release on synaptotagmin 1 at GABAergic and glutamatergic hippocampal synapses.* Proceedings of the National Academy of Sciences of the United States of America, 2008. **105**(40): p. 15581-15586.
291. Kochubey, O., N. Babai, and R. Schneggenburger, *A Synaptotagmin Isoform Switch during the Development of an Identified CNS Synapse.* Neuron, 2016. **90**(5): p. 984-99.
292. Martens, S., M.M. Kozlov, and H.T. McMahon, *How synaptotagmin promotes membrane fusion.* Science, 2007. **316**(5828): p. 1205-8.
293. Sommeijer, J.P. and C.N. Levelt, *Synaptotagmin-2 is a reliable marker for parvalbumin positive inhibitory boutons in the mouse visual cortex.* PLoS One, 2012. **7**(4): p. e35323.
294. Lee, S.-H., B. Schwaller, and E. Neher, *Kinetics of Ca2+ binding to parvalbumin in bovine chromaffin cells: implications for [Ca2+] transients of neuronal dendrites.* The Journal of Physiology, 2000. **525**(2): p. 419-432.
295. Uezu, A., et al., *Identification of an elaborate complex mediating postsynaptic inhibition.* Science, 2016. **353**(6304): p. 1123-9.
296. Paul, A., et al., *Transcriptional Architecture of Synaptic Communication Delineates GABAergic Neuron Identity.* Cell, 2017. **171**(3): p. 522-539 e20.
297. Volk, D.W., et al., *Reciprocal alterations in pre- and postsynaptic inhibitory markers at chandelier cell inputs to pyramidal neurons in schizophrenia.* Cerebral Cortex, 2002. **12**(10): p. 1063-1070.
298. Beneyto, M., et al., *Lamina-specific alterations in cortical GABA(A) receptor subunit expression in schizophrenia.* Cereb Cortex, 2011. **21**(5): p. 999-1011.
299. Rudolph, U. and H. Mohler, *GABAA receptor subtypes: Therapeutic potential in Down syndrome, affective disorders, schizophrenia, and autism.* Annu Rev Pharmacol Toxicol, 2014. **54**: p. 483-507.
300. Tan, K.R., U. Rudolph, and C. Luscher, *Hooked on benzodiazepines: GABAA receptor subtypes and addiction.* Trends Neurosci, 2011. **34**(4): p. 188-97.
301. Gao, Y. and S.A. Heldt, *Enrichment of GABAA Receptor alpha-Subunits on the Axonal Initial Segment Shows Regional Differences.* Front Cell Neurosci, 2016. **10**: p. 39.
302. Thomson, A.M. and J.N. Jovanovic, *Mechanisms underlying synapse-specific clustering of GABA(A) receptors.* Eur J Neurosci, 2010. **31**(12): p. 2193-203.
303. Ali, A.B. and A.M. Thomson, *Synaptic alpha 5 subunit-containing GABAA receptors mediate IPSPs elicited by dendrite-preferring cells in rat neocortex.* Cereb Cortex, 2008. **18**(6): p. 1260-71.

304. Serwanski, D.R., et al., *Synaptic and nonsynaptic localization of GABAA receptors containing the alpha5 subunit in the rat brain.* J Comp Neurol, 2006. **499**(3): p. 458-70.
305. Honig, B. and L. Shapiro, *Adhesion Protein Structure, Molecular Affinities, and Principles of Cell-Cell Recognition.* Cell, 2020. **181**(3): p. 520-535.
306. Chen, W.V. and T. Maniatis, *Clustered protocadherins.* Development, 2013. **140**(16): p. 3297-302.
307. Duan, X., et al., *Type II cadherins guide assembly of a direction-selective retinal circuit.* Cell, 2014. **158**(4): p. 793-807.
308. de Wit, J., et al., *Unbiased discovery of glypican as a receptor for LRRTM4 in regulating excitatory synapse development.* Neuron, 2013. **79**(4): p. 696-711.
309. Lanoue, V., et al., *The adhesion-GPCR BAI3, a gene linked to psychiatric disorders, regulates dendrite morphogenesis in neurons.* Mol Psychiatry, 2013. **18**(8): p. 943-50.
310. Krishnaswamy, A., et al., *Sidekick 2 directs formation of a retinal circuit that detects differential motion.* Nature, 2015. **524**(7566): p. 466-470.
311. Schreiner, D. and J.A. Weiner, *Combinatorial homophilic interaction between gamma-protocadherin multimers greatly expands the molecular diversity of cell adhesion.* Proc Natl Acad Sci U S A, 2010. **107**(33): p. 14893-8.
312. Dionne, G., et al., *Mechanotransduction by PCDH15 Relies on a Novel cis-Dimeric Architecture.* Neuron, 2018. **99**(3): p. 480-492.e5.
313. Brasch, J., et al., *Visualization of clustered protocadherin neuronal self-recognition complexes.* Nature, 2019. **569**(7755): p. 280-283.
314. Sigoillot, S.M., et al., *The Secreted Protein C1QL1 and Its Receptor BAI3 Control the Synaptic Connectivity of Excitatory Inputs Converging on Cerebellar Purkinje Cells.* Cell Rep, 2015. **10**(5): p. 820-832.
315. Takeo, Y.H., et al., *GluD2- and Cbln1-mediated competitive interactions shape the dendritic arbors of cerebellar Purkinje cells.* Neuron, 2021. **109**(4): p. 629-644.e8.
316. Kakegawa, W., et al., *Anterograde C1ql1 signaling is required in order to determine and maintain a single-winner climbing fiber in the mouse cerebellum.* Neuron, 2015. **85**(2): p. 316-29.
317. Tasic, B., et al., *Shared and distinct transcriptomic cell types across neocortical areas.* bioRxiv, 2017.
318. Südhof, T.C., *Synaptic Neurexin Complexes: A Molecular Code for the Logic of Neural Circuits.* Cell, 2017. **171**(4): p. 745-769.
319. Shapiro, L. and D.R. Colman, *The diversity of cadherins and implications for a synaptic adhesive code in the CNS.* Neuron, 1999. **23**(3): p. 427-430.
320. Südhof, T.C., *Towards an Understanding of Synapse Formation.* Neuron, 2018. **100**(2): p. 276-293.
321. O'Sullivan, M.L., et al., *FLRT proteins are endogenous latrophilin ligands and regulate excitatory synapse development.* Neuron, 2012. **73**(5): p. 903-10.
322. Anderson, G.R., et al., *Postsynaptic adhesion GPCR latrophilin-2 mediates target recognition in entorhinal-hippocampal synapse assembly.* Journal of Cell Biology, 2017. **216**(11): p. 3831-3846.
323. Sando, R., X. Jiang, and T.C. Sudhof, *Latrophilin GPCRs direct synapse specificity by coincident binding of FLRTs and teneurins.* Science, 2019. **363**(6429).
324. Chen, L.Y., et al., *Conditional Deletion of All Neurexins Defines Diversity of Essential Synaptic Organizer Functions for Neurexins.* Neuron, 2017. **94**(3): p. 611-625 e4.
325. Sclip, A. and T.C. Südhof, *LAR receptor phospho-tyrosine phosphatases regulate NMDA-receptor responses.* eLife, 2020. **9**: p. e53406.
326. Emperador-Melero, J., G. de Nola, and P.S. Kaeser, *Intact synapse structure and function after combined knockout of PTPδ, PTPσ, and LAR.* eLife, 2021. **10**: p. e66638.
327. Basu, R., M.R. Taylor, and M.E. Williams, *The classic cadherins in synaptic specificity.* Cell Adh Migr, 2015. **9**(3): p. 193-201.
328. Yamagata, M., J.A. Weiner, and J.R. Sanes, *Sidekicks: Synaptic adhesion molecules that promote lamina-specific connectivity in the retina.* Cell, 2002. **110**(5): p. 649-660.
329. Maness, P.F. and M. Schachner, *Neural recognition molecules of the immunoglobulin superfamily: signaling transducers of axon guidance and neuronal migration.* Nat Neurosci, 2007. **10**(1): p. 19-26.
330. Li, J., et al., *Alternative splicing controls teneurin-latrophilin interaction and synapse specificity by a shape-shifting mechanism.* Nature Communications, 2020. **11**(1): p. 2140.
331. Sando, R. and T.C. Südhof, *Latrophilin GPCR signaling mediates synapse formation.* eLife, 2021. **10**: p. e65717.
332. Matúš, D., et al., *The Essential Role of Latrophilin-1 Adhesion GPCR Nanoclusters in Inhibitory Synapses.* bioRxiv, 2023: p. 2023.10.08.561368.

333. Wang, S., et al., *Alternative splicing of latrophilin-3 controls synapse formation.* Nature, 2024. **626**(7997): p. 128-135.
334. Klassen, M.P. and K. Shen, *Wnt Signaling Positions Neuromuscular Connectivity by Inhibiting Synapse Formation in C. elegans.* Cell, 2007. **130**(4): p. 704-716.
335. Poon, V.Y., M.P. Klassen, and K. Shen, *UNC-6/netrin and its receptor UNC-5 locally exclude presynaptic components from dendrites.* Nature, 2008. **455**(7213): p. 669-73.
336. Haydon, P.G. and P. Drapeau, *From contact to connection: early events during synaptogenesis.* Trends in Neurosciences, 1995. **18**(4): p. 196-201.
337. Kano, M. and M. Watanabe, *Synaptogenesis and Synapse Elimination*, in *Handbook of the Cerebellum and Cerebellar Disorders*, M. Manto, et al., Editors. 2013, Springer Netherlands: Dordrecht. p. 281-299.
338. Kano, M. and M. Watanabe, *Synaptogenesis and Synapse Elimination*, in *Handbook of the Cerebellum and Cerebellar Disorders*, M.U. Manto, et al., Editors. 2022, Springer International Publishing: Cham. p. 309-332.
339. Regalado, M.P., et al., *Transsynaptic Signaling by Postsynaptic Synapse-Associated Protein 97.* The Journal of Neuroscience, 2006. **26**(8): p. 2343-2357.
340. Chanaday, N.L. and E.T. Kavalali, *Role of the endoplasmic reticulum in synaptic transmission.* Current Opinion in Neurobiology, 2022. **73**: p. 102538.
341. Emperador-Melero, J. and P.S. Kaeser, *Assembly of the presynaptic active zone.* Current Opinion in Neurobiology, 2020. **63**: p. 95-103.
342. Bresler, T., et al., *Postsynaptic Density Assembly Is Fundamentally Different from Presynaptic Active Zone Assembly.* The Journal of Neuroscience, 2004. **24**(6): p. 1507-1520.
343. Sando, R., et al., *Assembly of Excitatory Synapses in the Absence of Glutamatergic Neurotransmission.* Neuron, 2017. **94**(2): p. 312-321.e3.
344. Sigler, A., et al., *Formation and Maintenance of Functional Spines in the Absence of Presynaptic Glutamate Release.* Neuron, 2017. **94**(2): p. 304-311.e4.
345. Chen, X., et al., *Phase separation at the synapse.* Nature neuroscience, 2020. **23**(3): p. 301-310.
346. Wu, X., et al., *RIM and RIM-BP Form Presynaptic Active-Zone-like Condensates via Phase Separation.* Molecular Cell, 2019. **73**(5): p. 971-984.e5.
347. Milovanovic, D., et al., *A liquid phase of synapsin and lipid vesicles.* Science, 2018. **361**(6402): p. 604-607.
348. Park, D., et al., *Cooperative function of synaptophysin and synapsin in the generation of synaptic vesicle-like clusters in non-neuronal cells.* Nature Communications, 2021. **12**(1): p. 263.
349. Hoffmann, C., et al., *Synapsin condensation controls synaptic vesicle sequestering and dynamics.* Nature Communications, 2023. **14**(1): p. 6730.
350. Madhavan, R. and H.B. Peng, *HGF induction of postsynaptic specializations at the neuromuscular junction.* Journal of neurobiology, 2006. **66**(2): p. 134-147.
351. Kwon, H.-B. and B.L. Sabatini, *Glutamate induces de novo growth of functional spines in developing cortex.* Nature, 2011. **474**(7349): p. 100-104.
352. Oh, W.C., et al., *De novo synaptogenesis induced by GABA in the developing mouse cortex.* Science, 2016. **353**(6303): p. 1037-1040.
353. Bai, G. and M. Zhang, *Inhibitory postsynaptic density from the lens of phase separation.* Oxford Open Neuroscience, 2022. **1**.
354. Dejanovic, B., et al., *Palmitoylation of Gephyrin Controls Receptor Clustering and Plasticity of GABAergic Synapses.* PLOS Biology, 2014. **12**(7): p. e1001908.
355. Fukata, Y., et al., *Local palmitoylation cycles define activity-regulated postsynaptic subdomains.* Journal of Cell Biology, 2013. **202**(1): p. 145-161.
356. Fang, C., et al., *GODZ-mediated palmitoylation of GABAA receptors is required for normal assembly and function of GABAergic inhibitory synapses.* Journal of Neuroscience, 2006. **26**(49): p. 12758-12768.
357. Treutlein, B., et al., *Cartography of neurexin alternative splicing mapped by single-molecule long-read mRNA sequencing.* Proc Natl Acad Sci U S A, 2014. **111**(13): p. E1291-9.
358. Sudhof, T.C., *Neuroligins and neurexins link synaptic function to cognitive disease.* Nature, 2008. **455**(7215): p. 903-11.
359. Trotter, J.H., et al., *A combinatorial code of neurexin-3 alternative splicing controls inhibitory synapses via a trans-synaptic dystroglycan signaling loop.* Nature Communications, 2023. **14**(1): p. 1771.
360. Song, J.Y., et al., *Neuroligin 1 is a postsynaptic cell-adhesion molecule of excitatory synapses.* Proceedings of the National Academy of Sciences of the United States of America, 1999. **96**(3): p. 1100-1105.

361. Budreck, E.C. and P. Scheiffele, *Neuroligin-3 is a neuronal adhesion protein at GABAergic and glutamatergic synapses.* Eur J Neurosci, 2007. **26**(7): p. 1738-48.
362. Hoon, M., et al., *Neuroligin-4 is localized to glycinergic postsynapses and regulates inhibition in the retina.* Proc Natl Acad Sci U S A, 2011. **108**(7): p. 3053-8.
363. Takacs, V.T., T.F. Freund, and G. Nyiri, *Neuroligin 2 is expressed in synapses established by cholinergic cells in the mouse brain.* PLoS One, 2013. **8**(9): p. e72450.
364. Poulopoulos, A., et al., *Neuroligin 2 drives postsynaptic assembly at perisomatic inhibitory synapses through gephyrin and collybistin.* Neuron, 2009. **63**(5): p. 628-42.
365. Panzanelli, P., et al., *Distinct mechanisms regulate GABAA receptor and gephyrin clustering at perisomatic and axo-axonic synapses on CA1 pyramidal cells.* J Physiol, 2011. **589**(Pt 20): p. 4959-80.
366. Zeng, M., et al., *Phase Transition in Postsynaptic Densities Underlies Formation of Synaptic Complexes and Synaptic Plasticity.* Cell, 2016. **166**(5): p. 1163-1175.e12.
367. Zeng, M., et al., *Reconstituted Postsynaptic Density as a Molecular Platform for Understanding Synapse Formation and Plasticity.* Cell, 2018. **174**(5): p. 1172-1187.e16.
368. Bai, G., Y. Wang, and M. Zhang, *Gephyrin-mediated formation of inhibitory postsynaptic density sheet via phase separation.* Cell Res, 2021. **31**(3): p. 312-325.
369. Liu, Y.-T., et al., *Mesophasic organization of GABAA receptors in hippocampal inhibitory synapses.* Nature Neuroscience, 2020. **23**(12): p. 1589-1596.
370. Manitt, C., et al., *Netrin participates in the development of retinotectal synaptic connectivity by modulating axon arborization and synapse formation in the developing brain.* Journal of Neuroscience, 2009. **29**(36): p. 11065-11077.
371. Goldman, J.S., et al., *Netrin-1 promotes excitatory synaptogenesis between cortical neurons by initiating synapse assembly.* Journal of Neuroscience, 2013. **33**(44): p. 17278-17289.
372. Zheng, M., et al., *Netrin-1 promotes synaptic formation and axonal regeneration via JNK1/c-Jun pathway after the middle cerebral artery occlusion.* Frontiers in Cellular Neuroscience, 2018. **12**: p. 13.
373. Davis, J.Q., S. Lambert, and V. Bennett, *Molecular composition of the node of Ranvier: identification of ankyrin-binding cell adhesion molecules neurofascin (mucin+/third FNIII domain-) and NrCAM at nodal axon segments.* Journal of Cell Biology, 1996. **135**(5): p. 1355-1367.
374. Buttermore, E.D., et al., *Pinceau Organization in the Cerebellum Requires Distinct Functions of Neurofascin in Purkinje and Basket Neurons during Postnatal Development.* The Journal of Neuroscience, 2012. **32**(14): p. 4724-4742.
375. Huang, Z.J., *Subcellular organization of GABAergic synapses: role of ankyrins and L1 cell adhesion molecules.* Nat Neurosci, 2006. **9**(2): p. 163-6.
376. Ashrafi, S., et al., *Neuronal Ig/Caspr recognition promotes the formation of axoaxonic synapses in mouse spinal cord.* Neuron, 2014. **81**(1): p. 120-129.
377. Meerabux, J.M., et al., *Human netrin-G1 isoforms show evidence of differential expression.* Genomics, 2005. **86**(1): p. 112-116.
378. Yasuda, M., et al., *An activity-dependent determinant of synapse elimination in the mammalian brain.* Neuron, 2021. **109**(8): p. 1333-1349.e6.
379. Chamma, I. and O. Thoumine, *Dynamics, nanoscale organization, and function of synaptic adhesion molecules.* Molecular and Cellular Neuroscience, 2018. **91**: p. 95-107.
380. Kano, M. and T. Watanabe, *Developmental synapse remodeling in the cerebellum and visual thalamus.* F1000Res, 2019. **8**.
381. Carrillo, J., N. Nishiyama, and H. Nishiyama, *Dendritic Translocation Establishes the Winner in Cerebellar Climbing Fiber Synapse Elimination.* The Journal of Neuroscience, 2013. **33**(18): p. 7641-7653.
382. Leto, K., et al., *Consensus Paper: Cerebellar Development.* Cerebellum, 2016. **15**(6): p. 789-828.
383. Hashimoto, K., et al., *Postsynaptic P/Q-type Ca²⁺ channel in Purkinje cell mediates synaptic competition and elimination in developing cerebellum.* Proceedings of the National Academy of Sciences, 2011. **108**(24): p. 9987-9992.
384. Nakayama, H., et al., *GABAergic Inhibition Regulates Developmental Synapse Elimination in the Cerebellum.* Neuron, 2012. **74**(2): p. 384-396.
385. Kurihara, H., et al., *Impaired parallel fiber→ Purkinje cell synapse stabilization during cerebellar development of mutant mice lacking the glutamate receptor δ2 subunit.* Journal of Neuroscience, 1997. **17**(24): p. 9613-9623.
386. Carrillo, E., et al., *Delta glutamate receptors are functional glycine- and D-serine–gated cation channels in situ.* Science Advances, 2021. **7**(52): p. eabk2200.

387. Miyazaki, T., et al., *Ablation of glutamate receptor GluRδ2 in adult Purkinje cells causes multiple innervation of climbing fibers by inducing aberrant invasion to parallel fiber innervation territory.* Journal of Neuroscience, 2010. **30**(45): p. 15196-15209.
388. Ito-Ishida, A., et al., *Presynaptically released Cbln1 induces dynamic axonal structural changes by interacting with GluD2 during cerebellar synapse formation.* Neuron, 2012. **76**(3): p. 549-64.
389. Mishina, M., et al., *Molecular mechanism of parallel fiber-Purkinje cell synapse formation.* Frontiers in neural circuits, 2012. **6**: p. 90.
390. Chattopadhyaya, B., et al., *Experience and activity-dependent maturation of perisomatic GABAergic innervation in primary visual cortex during a postnatal critical period.* J Neurosci, 2004. **24**(43): p. 9598-611.
391. Nakashiba, T., et al., *Complementary expression and neurite outgrowth activity of netrin-G subfamily members.* Mech Dev, 2002. **111**(1-2): p. 47-60.
392. Nakashiba, T., et al., *Netrin-G1: a novel glycosyl phosphatidylinositol-linked mammalian netrin that is functionally divergent from classical netrins.* Journal of Neuroscience, 2000. **20**(17): p. 6540-6550.
393. Matsukawa, H., et al., *Netrin-G/NGL complexes encode functional synaptic diversification.* J Neurosci, 2014. **34**(47): p. 15779-92.
394. Zhang, Q., et al., *Diversification of behavior and postsynaptic properties by netrin-G presynaptic adhesion family proteins.* Mol Brain, 2016. **9**: p. 6.
395. Zhang, Q., et al., *Netrin-G1 regulates fear-like and anxiety-like behaviors in dissociable neural circuits.* Sci Rep, 2016. **6**: p. 28750.
396. Soto, F., L. Zhao, and D. Kerschensteiner, *Synapse maintenance and restoration in the retina by NGL2.* eLife, 2018. **7**: p. e30388.
397. Fino, E. and R. Yuste, *Dense inhibitory connectivity in neocortex.* Neuron, 2011. **69**(6): p. 1188-203.
398. Karnani, M.M., M. Agetsuma, and R. Yuste, *A blanket of inhibition: functional inferences from dense inhibitory connectivity.* Current opinion in Neurobiology, 2014. **26**: p. 96-102.
399. Schneider-Mizell, C.M., et al., *Chandelier cell anatomy and function reveal a variably distributed but common signal.* bioRxiv, 2020.
400. Di Cristo, G., et al., *Subcellular domain-restricted GABAergic innervation in primary visual cortex in the absence of sensory and thalamic inputs.* Nat Neurosci, 2004. **7**(11): p. 1184-6.
401. Chattopadhyaya, B., et al., *GAD67-mediated GABA synthesis and signaling regulate inhibitory synaptic innervation in the visual cortex.* Neuron, 2007. **54**(6): p. 889-903.
402. Chattopadhyaya, B., et al., *Neural Cell Adhesion Molecule-Mediated Fyn Activation Promotes GABAergic Synapse Maturation in Postnatal Mouse Cortex.* The Journal of Neuroscience, 2013. **33**(14): p. 5957-5968.
403. Yamamoto, N., T. Kurotani, and K. Toyama, *Neural Connections Between the Lateral Geniculate Nucleus and Visual Cortex in Vitro.* Science, 1989. **245**(4914): p. 192-194.
404. Bolz, J., et al., *Formation of target-specific neuronal projections in organotypic slice cultures from rat visual cortex.* Nature, 1990. **346**(6282): p. 359-362.
405. Molnár, Z. and C. Blakemore, *Lack of regional specificity for connections formed between thalamus and cortex in coculture.* Nature, 1991. **351**(6326): p. 475-477.
406. Frotscher, M. and B. Heimrich, *Formation of layer-specific fiber projections to the hippocampus in vitro.* Proceedings of the National Academy of Sciences, 1993. **90**(21): p. 10400-10403.
407. Barnes, A.P. and F. Polleux, *Establishment of Axon-Dendrite Polarity in Developing Neurons.* Annual Review of Neuroscience, 2009. **32**(1): p. 347-381.
408. Shen, Q., et al., *The timing of cortical neurogenesis is encoded within lineages of individual progenitor cells.* Nature Neuroscience, 2006. **9**(6): p. 743-751.
409. Eiraku, M., et al., *Self-organized formation of polarized cortical tissues from ESCs and its active manipulation by extrinsic signals.* Cell stem cell, 2008. **3**(5): p. 519-532.
410. Gaspard, N., et al., *An intrinsic mechanism of corticogenesis from embryonic stem cells.* Nature, 2008. **455**(7211): p. 351-357.
411. Maroof, A.M., et al., *Prospective isolation of cortical interneuron precursors from mouse embryonic stem cells.* J Neurosci, 2010. **30**(13): p. 4667-75.
412. Maroof, A.M., et al., *Directed differentiation and functional maturation of cortical interneurons from human embryonic stem cells.* Cell Stem Cell, 2013. **12**(5): p. 559-72.
413. Colasante, G., et al., *Rapid Conversion of Fibroblasts into Functional Forebrain GABAergic Interneurons by Direct Genetic Reprogramming.* Cell Stem Cell, 2015. **17**(6): p. 719-734.

414. Weyn-Vanhentenryck, S.M., et al., *Precise temporal regulation of alternative splicing during neural development.* Nature Communications, 2018. **9**(1): p. 2189.
415. Oishi, K., et al., *Identity of neocortical layer 4 neurons is specified through correct positioning into the cortex.* Elife, 2016. **5**.
416. Muller, T.H., D. Swandulla, and H.U. Zeilhofer, *Synaptic connectivity in cultured hypothalamic neuronal networks.* J Neurophysiol, 1997. **77**(6): p. 3218-25.
417. Christie, S.B. and A.L. De Blas, *GABAergic and glutamatergic axons innervate the axon initial segment and organize GABA(A) receptor clusters of cultured hippocampal pyramidal cells.* J Comp Neurol, 2003. **456**(4): p. 361-74.
418. Christie, S.B., et al., *Synaptic and extrasynaptic GABA(A) receptor and gephyrin clusters.* Changing Views of Cajal's Neuron, 2002. **136**: p. 157-180.
419. Rao, A., E.M. Cha, and A.M. Craig, *Mismatched appositions of presynaptic and postsynaptic components in isolated hippocampal neurons.* Journal of Neuroscience, 2000. **20**(22): p. 8344-8353.
420. Sarkar, A., et al., *Efficient Generation of CA3 Neurons from Human Pluripotent Stem Cells Enables Modeling of Hippocampal Connectivity In Vitro.* Cell Stem Cell, 2018. **22**(5): p. 684-697 e9.
421. Verstraelen, P., et al., *Image-Based Profiling of Synaptic Connectivity in Primary Neuronal Cell Culture.* Frontiers in Neuroscience, 2018. **12**.
422. Castrén, E., *Trophic Factors: Neurotrophic Factors*, in *Neuroscience in the 21st Century: From Basic to Clinical*, D.W. Pfaff, Editor. 2013, Springer New York: New York, NY. p. 1555-1589.
423. Sugiyama, S., et al., *Experience-dependent transfer of Otx2 homeoprotein into the visual cortex activates postnatal plasticity.* Cell, 2008. **134**(3): p. 508-520.
424. Lewis, D.A., *Chandelier cells: shedding light on altered cortical circuitry in schizophrenia.* Molecular Psychiatry, 1998. **3**(6): p. 468-471.
425. Lewis, D.A., et al., *Cortical parvalbumin interneurons and cognitive dysfunction in schizophrenia.* Trends Neurosci, 2012. **35**(1): p. 57-67.
426. Lewis, D.A. and R.A. Sweet, *Schizophrenia from a neural circuitry perspective: advancing toward rational pharmacological therapies.* J Clin Invest, 2009. **119**(4): p. 706-16.
427. Fazzari, P., et al., *Control of cortical GABA circuitry development by Nrg1 and ErbB4 signalling.* Nature, 2010. **464**(7293): p. 1376-80.
428. Ariza, J., et al., *The Number of Chandelier and Basket Cells Are Differentially Decreased in Prefrontal Cortex in Autism.* Cereb Cortex, 2018. **28**(2): p. 411-420.
429. He, L.J., et al., *Conditional deletion of Mecp2 in parvalbumin-expressing GABAergic cells results in the absence of critical period plasticity.* Nat Commun, 2014. **5**: p. 5036.
430. Ito-Ishida, A., et al., *Loss of MeCP2 in Parvalbumin-and Somatostatin-Expressing Neurons in Mice Leads to Distinct Rett Syndrome-like Phenotypes.* Neuron, 2015. **88**(4): p. 651-8.
431. Lu, H., et al., *Loss and Gain of MeCP2 Cause Similar Hippocampal Circuit Dysfunction that Is Rescued by Deep Brain Stimulation in a Rett Syndrome Mouse Model.* Neuron, 2016. **91**(4): p. 739-747.
432. Ure, K., et al., *Restoration of Mecp2 expression in GABAergic neurons is sufficient to rescue multiple disease features in a mouse model of Rett syndrome.* Elife, 2016. **5**.
433. Zhu, S., et al., *Genetic disruption of ankyrin-G in adult mouse forebrain causes cortical synapse alteration and behavior reminiscent of bipolar disorder.* Proc Natl Acad Sci U S A, 2017. **114**(39): p. 10479-10484.
434. Martinowich, K., R.J. Schloesser, and H.K. Manji, *Bipolar disorder: from genes to behavior pathways.* J Clin Invest, 2009. **119**(4): p. 726-36.
435. Lopez, A.Y., et al., *Ankyrin-G isoform imbalance and interneuronopathy link epilepsy and bipolar disorder.* Mol Psychiatry, 2017. **22**(10): p. 1464-1472.
436. Hughes, T., et al., *Elevated expression of a minor isoform of ANK3 is a risk factor for bipolar disorder.* Transl Psychiatry, 2018. **8**(1): p. 210.
437. Schwaller, B., et al., *Parvalbumin deficiency affects network properties resulting in increased susceptibility to epileptic seizures.* Mol Cell Neurosci, 2004. **25**(4): p. 650-63.
438. Dutton, S.B., et al., *Preferential inactivation of Scn1a in parvalbumin interneurons increases seizure susceptibility.* Neurobiol Dis, 2013. **49**: p. 211-20.
439. Tai, C., et al., *Impaired excitability of somatostatin- and parvalbumin-expressing cortical interneurons in a mouse model of Dravet syndrome.* Proc Natl Acad Sci U S A, 2014. **111**(30): p. E3139-48.
440. Malik, R., et al., *Tsc1 represses parvalbumin expression and fast-spiking properties in somatostatin lineage cortical interneurons.* Nat Commun, 2019. **10**(1): p. 4994.

441. Raza, S.A., et al., *HIPP neurons in the dentate gyrus mediate the cholinergic modulation of background context memory salience.* Nature Communications, 2017. **8**(1): p. 189.

Chapter 2: Materials and methods

2.1. Sample preparation and image acquisition

2.1.1. Animal husbandry

All animal handling and maintenance were performed according to the regulations of the Institutional Animal Care and Use Committee of Columbia University Medical center. Mouse lines used in this study are detailed in Table 1. Whenever possible, mice were maintained as homozygotes for breeding purposes. The Nkx2.1^{Cre} allele is a BAC transgenic, so homozygous breeders were identified by the proportion of Cre positivity of the offspring. Nkx2.1^{CreER} homozygotes are embryonic lethal, and thus were maintained as heterozygotes. Since the Sst^{CreER} allele is a knock-in-knock-out, all animals used in this study were Sst^{CreER} heterozygotes. For all experiments involving the Reln$^{RL/RL}$ Reeler mutant, Reln$^{RL/wt}$ or Reln$^{wt/wt}$ littermates were used as controls. Reln$^{RL/RL}$ animals are infertile, so all crosses to generate these animals were from heterozygous parents. All animal information, including ID, cross information, and genotype were stored in a database for reference. All embryonic time points were counted from discovery of vaginal plug (E0.5).

2.1.2. Genotyping

DNA was sourced from tails and digested in proteinase K (proK) in (50 mM Tris, 1 mM EDTA, 0.5% Tween, 1x Proteinase K, pH 8.8) at 65°C for at least one hour, and proK was heat inactivated at 95°C for 10 minutes prior to use for genotyping reactions. Animals were genotyped using published or custom primers as outlined in Table 2. Genotyping protocols used either 2x Hotstart PCR Master mix (NEB) or 2x DreamTaq Green Master mix (Thermo Scientific K1081) according to manufacturer's protocols. Target-specific thermal cycling protocols are available

upon request. PCR products were separated by electrophoresis using 1.5% to 2% agarose gels in TAE buffer and visualized with ethidium bromide or Gel Red [Biotium]. Animal cross information and genotyping results including gel images were stored in a database for cross-reference. Heat-inactivated DNA was systematically stored should animals require re-genotyping.

2.1.3. Tamoxifen Induction

For all experiments involving CreER induction with tamoxifen, dosages were calculated on a ug/g basis, with animals weighed at time of treatment. Table 3 shows the specific dosage calculations for each condition. Experiments involving Nkx2.1^{CreER} mice were performed according to Taniguchi et al., 2013[1] with minor modifications. In brief, Nkx2.1^{CreER};Ai34 studs were crossed with Swiss Webster females and monitored for timed plugs. Pregnant females were orally gavaged with tamoxifen at embryonic day 17.5 (E17.5) to induce Cre recombination. Since tamoxifen treatment can complicate pregnancy, in cases where emergency cesarian sections were performed, pups were fostered with the closest age-matched litter, at oldest P2. Experiments involving Sst^{CreER} mice were performed in adult mice using a tamoxifen ramping protocol developed by Raza et al., 2017[2]. Briefly, animals were treated with increasing doses of tamoxifen over the course of three days as follows: Day1, 80 µg/g; Day2 160 µg/g, Day3, 320 µg/g. We found this protocol to have much higher rates of survival than the initial protocol developed by Taniguchi et al., 2011[3].

2.1.4. Perfusions and tissue preparation

Mice were deeply anesthetized by intraperitoneal injection of ketamine/xylazine (87.5 mg/kg and 12.5 mg/kg, respectively) and confirmed by toe pinch prior to non-survival surgery. Buffer formulations for perfusions and associated products are listed in the Perfusion and Slice Preparation Buffers subsection of Appendix A. For samples used in PV- and Sst-based cellular localization experiments (see Immunostaining below), tissue was prepared according to Gallerani & Au 2020[4]

via trans-cardiac perfusion with Phosphate Buffered Saline (PBS) solution, then 4% PFA in PBS. Brains were extracted and post-fixed overnight at 4°C in the same fixative. Samples were sliced into 50 µm sections on a Leica VT1000S vibratome and stored in propylene glycol:glycerol:PBS solution (3:3:4) antifreeze solution at -30°C until use in immunostaining.

For experiments involving immunostaining of synaptic molecules, samples were prepared according to Notter et al., 2014[5]. Briefly, mice were intracardially perfused with 15-20 mL ice cold, oxygenated ACSF [125 mM NaCl, 26 mM $NaHCO_3$, 25 mM glucose, 2.5 mM KCl, 2.5 mM $CaCl_2$, 2 mM $MgCl_2$, 1.25 mM NaH_2PO_4], pH 7.4,). After decapitation, the brain was extracted from the skull and sliced into 2 mm sections using a CytoVista™ Coronal Mouse Brain Slicer [Thermo Scientific V11320]. Tissue slices were then immersion fixed in 4% PFA in PBS for one hour, washed extensively with PBS, cryoprotected in 30% sucrose in PBS overnight at 4°C, then embedded in OTC [Sakura Finetek 4583] on dry ice and stored at -80°C. Just prior to immunostaining, tissue was sliced in to 40 µm sections at -20°C on a cryostat and transferred to PBS.

2.1.5. Cortical dissection and dissociation

Dissection and dissociation were performed as outlined in Nunnelly et al., 2022[3] using buffers and products listed in Appendix A, Dissection and Dissociation media. Briefly, cortices from P3-P5 pups were micro-dissected into chilled Hibernate A (HibA) media (Hibernate A[Gibco], 1% Pen/Strep [Gibco]). Briefly, the brain was excised from skull and meninges were carefully removed. The cortex was separated from subcortical regions and hippocampus and cortices were stored intact on ice until dissociation. The Dissociation media formulation and associated products is listed in Appendix A. Briefly, Papain [Worthington Biochemical LK003178] was resuspended to 20 U/mL in 5 mL HibA media and activated by incubating for ~20 minutes at 37°C, mixed with DNAseI (200 U/mL final, [Worthington Biochemical LK003172]), then added to minced cortices.

Cortical tissue was dissociated at 37 °C for 15 min on an end-over-end rotator. Papain was removed after centrifugation at 200 × *g* for 5 min, and tissue was resuspended in HibA. Tissue was dissociated by trituration in HibA media and filtered through a 40 µm filter to ensure single cell suspension.

2.1.6. Tissue culture

Concentration and viability of dissociated cells was determined using a Countess II FL cell counting system [Life Technologies] with trypan blue exclusion assay to determine viability. Survival above 80% was deemed acceptable for culturing. Cells were then centrifuged and re-suspended in plating media to a concentration of 6×10^7 viable cells per mL, and 15-25 µL cell suspensions were spotted into the center of a well on poly-D lysine pre-coated 24-well plates. Cells were allowed to adhere for 30-45 minutes in a 37°C, 5.0% CO_2 tissue culture incubator, then pre-warmed plating media was added to a final volume of 500 µL/well. The following day (DIV1), media was fully changed to growth media, and cultures were maintained by 50% media changes performed every other day. Total well volume was increased to 750 µL on DIV7 and 1 mL on DIV14. Unless otherwise noted, cultures were fixed on DIV21 in 4% PFA [Electron Microscopy Sciences] in PBS for 10 minutes, washed 4x in PBS, and 1 mL PBS was added for storage. Fixed culture plates were wrapped in Parafilm and stored in a 4°C refrigerator until processed for immunostaining. All tissue culture media formulations are listed in the Tissue Culture Reagents and Media subsection of Appendix A.

For experiments involving IL-13, the peptide [Biolegend 575904] was resuspended in growth media to working concentrations just prior to addition. Treatment was initiated at DIV13 and samples were fixed 24 hours later. This earlier time point was chosen to match the P5-P15 window of synaptogenesis identified by Barron et al., 2023[2]. All tissue culture was performed using 24-well

optical plastic tissue culture plates [Ibidi 82426]. As outlined in Appendix A, plates were pre-treated with 20 µg/mL poly-D lysine [Sigma Aldrich P6407] in borate buffer (10 mM Sodium Borate, pH 8.5 (sodium tetraborate and boric acid [Sigma S9640 and B6768, respectively]) overnight at room temperature, then rinsed with sterile water and allowed to dry before plating. Tissue culture time points were counted from the day of plating (DIV0).

2.1.7. Immunostaining

Cryopreserved tissue slices were processed as free-floating sections according to previous work[6] with slight variation. In brief, free-floating tissue slices were blocked and permeabilized in 10% Donkey serum, 0.01% Triton-X100 in PBS for one hour, then endogenous mouse IgG was blocked by overnight incubation at 4°C with an anti-mouse Fab fragment [Jackson immunoresearch 715-007-003]. Samples were washed 4x in TBS, then incubated with primary antibodies for 2-3 days, washed, and incubated with secondary antibodies overnight. All primary and secondary antibodies used in this study are listed in Table 4 and Table 5, respectively. Following the initial blocking and permeabilization step, all incubation steps were performed using 1% donkey serum in TBS at 4°C, and all wash steps were performed using 4x washes with TBS. TBS was used instead of PBS because the anti-gephyrin antibody is phospho-specific. Slices were mounted by paintbrush on SuperFrost Plus charged slides [Fisher Scientific 12-550-15] in TrueBlack mounting media [Biotium 23017], coverslipped, cured overnight at RT, then stored in the dark at 4°C prior to imaging. Since the tissue cultures were grown on plates specifically designed for microscopy, all immunostaining was performed in well. Cultures were blocked and permeabilized as above, with exclusion of Fab incubation. Primary and secondary incubations were also performed as outlined above, but incubations were only overnight. Following the final wash after secondary incubation, a drop of Aqua-Poly/Mount [Polysciences 18606] or TrueBlack mounting

media was added directly to the growth region and the well was covered with 15 mm round coverslip. Coverslipped samples were stored in the dark at 4°C prior to imaging. PV and SST immunostaining of PFA-perfused tissue was performed according to Gallerani and Au 2020[7].

2.1.8. Fluorescence microscopy

PV- and Sst-stained samples were imaged on Zeiss epifluourescent tiling microscope at 200x. For cultures, we chose representative regions for bulk cultures and for single cells, areas where individual cells were isolated and in the center of the growth region. For slices, we chose coronal sections between Bregma 0.61 and -1.31 mm and imaged from anterior cingulate to S1 of cortex. Synapses were imaged on a Zeiss LSM 800 confocal microscope using a 40x oil objective, 0.5x zoom, and a pinhole size of 37 µm to gather 1024x1024 voxel Z stacks of variable depth. These imaging conditions yielded an X,Y,Z resolution of 0.312, 0.312, 0.500 µm/voxel. Following a low-resolution preview scan, multi-tile regions of interest were designated for imaging at the aforementioned resolution. Initial focal points were established by the addition of support points and verified by autofocus prior to imaging. The designated ROIs were imaged overnight via automation inherent to the Zen Blue Software [Zeiss].

2.2. Sample Nomenclature

This topic is discussed at length in the Nomenclature section of Chapter 1. Briefly, all images were systematically assigned unique names as outlined in Appendix B based on the biological sample of origin and imaging process, according to the following pattern: *[Experiment ID]_[Sample ID]_[Biological conditions]_[Staining conditions]_[Scene ID]*_m*[Tile ID]*_[*variable stem*]. The *Experiment ID* consisted of the mouse birthdate for slices or the seeding date for culture. The

Sample ID was the mouse ear tag for slices or the well identifier (e.g. A6) for culture. The *biological conditions* included the driver line and reporter, as well as any relevant mutations or treatment conditions. The *staining conditions* included the antibodies used in staining indexed by channel number. The *Scene ID* (i.e. tile region) was assigned in Zen Blue [Zeiss] either during the imaging setup or post-hoc during scene splitting. The *Tile* IDs were designated in FIJI based on the image metadata. Each of these name components was kept constant throughout downstream steps, thus providing an internal structure for image alignment and data referencing during image processing. Moreover, since image/sample nomenclature was rooted in the biological ID, we were able to use this identifier to cross-reference animal husbandry and genotyping information. The *variable stem* was assigned algorithmically based on data type and/or processing stage. This final tag ensured that different data from the source file would not overwrite each other. **Image processing**

Image processing is discussed in-depth below in Custom code and in Chapter 1. Briefly, multi-scene images were initially split into individual scenes in Zen Blue software and exported as .czi files. All images were systematically named prior to processing such that each image had a unique name with human-readable identifiers (experiment number, sample name, genotype, staining, scene number). Image metadata was read and extracted using the Bio-formats[8] plugin in FIJI[9], and source metadata was stored in a designated metadata folder for future reference. Scenes were initially split into individual tiles and stored as .tif files. All downstream processing occurred on a tile-by-tile basis, with stitched scenes only used for visualization purposes. Images were processed in ImageJ/FIJI using freely available plugins or custom code written in the ImageJ macro language or Jython. Images were converted to the HDF5 format using the ilastik ImageJ plugin prior to supervised learning in Ilastik software[10]. Ilastik was trained on two pixel classifiers: one for presynaptic and one for target, which were then exported as probability maps. Ilastik generated

probability maps were imported into FIJI and concatenated with the source image file. 3D object segmentation of presynaptic boutons was performed in FIJI using the 3D spot segmentation algorithm based on seeds from a 3D local maxima filter, both of which are from the 3D ImageJ Suite (mcib3d plugin)[6]. Unsupervised classification results were mapped onto boutons using custom code written in FIJI/ImageJ using the CLIJ2 plugin[11].

2.4. Custom Code

At most stages in the pipeline, custom code was written to facilitate image batch processing and analysis. This code was used for the following functions:

2.4.1. Metadata extraction

This code is discussed in greater depth in Chapter 1. Briefly, it extracts useful image information from the image metadata, such as the original image dimensions (number of pixels in X and Y, plus the number of Z slices, channels, and timepoints), acquisition parameters (the X, Y, and Z voxel sizes in µm, time interval in ms, channel names, excitation information, pinhole size, detector gain), scene information (e.g. the number of tiles in each scene, the order of acquisition, and their X, Y, and Z locations on the microscope stage (in µm) as well as their their relative X, Y, and Z positions within the scene/tile region (in pixels/voxels)), and microscope information (microscope name, objective used, nominal magnification, and detectors). The acquisition date, a timestamp to the millisecond, is also extracted and, since it is consistent throughout all scenes acquired during the imaging session, is used as a unique unifying string for grouping. These data are extracted primarily from the Bio-formats metadata, a library that standardizes image metadata from many different microscope filetypes[11]. When the required information is unavailable through Bio-formats, it is extracted from the original image metadata. These aggregated data are

automatically stored in tabular form as a .csv file at a standardized location within a central metadata repository. Also stored are the Bio-formats metadata (as an XML file), the original image metadata (as a list in TXT format), and any post-hoc changes to the image that occurred prior to entry into the pipeline (so long as those changes are documented in the original metadata).

2.4.2. Stitching

This code is discussed in greater depth in the Stitching section of Chapter 1. Briefly, for multi-tile scenes, the metadata extraction code automatically generates a generalized tile configurations map that designates the X, Y, and Z coordinates of each tile within its respective scene. The generalized tile configuration maps are automatically stored as TXT files at a standardized location within the metadata directory, and are thus accessible at any stage in the pipeline. We wrote custom code that harnesses the *Grid/Collection Stitching* plugin in FIJI[10, 12] to automatically stitch tiles at any stage in the pipeline into their respective multi-tile scenes using this generalized tile configurations map.

2.4.3. File Grouping

This code is discussed in greater depth in File Grouping section of Chapter 1. We wrote custom code that utilizes the file naming conventions and the extracted metadata to automatically group related files based on multiple parameters. In particular, individual tiles are automatically grouped by sample format (i.e. slice vs culture), experiment, sample (i.e. well for culture, mouse ID for slice), slide, and scene. We use the acquisition date in the metadata (a date-time string with millisecond resolution) to group images acquired during the same imaging session independent of scene name.

2.4.4. Image standardization

We built in multiple checkpoints for image standardization to ensure that the inputs meet expectations for the given pipeline. We discuss them in greater depth in Chapter 1. We utilized the Bio-formats plugin to convert source images from their various microscope-specific image formats into the universal TIF image format and split multi-tile scenes into individual tiles, which serve as the unit of processing throughout the remainder of the pipeline. We utilized the extracted metadata for quality control (QC) stages to ensure images are structured as required and expected. These QC steps assess the following: image dimensions, pixel/voxel size, bitdepth, channel order, and dimension order. Whenever possible, images are automatically adjusted to meet pre-designated requirements (e.g. dimension order or pixel size adjustments), and when not possible, these inputs are rejected (e.g. missing channels, 2D images for a 3D pipeline).

2.4.5. Image Correction

We implemented background and shading correction in the X/Y dimensions by adapting the previously-published BaSiC plugin[13] to run efficiently for large, multi-channel tiled Z-stacks. This code is discussed in depth in the XY correction section of chapter 1. The code utilizes the metadata extraction and grouping algorithms mentioned above, as well as a rapid extended depth of focus algorithm implemented on the GPU via CLIJ2[8]. The approach collapses the Z stack while maintaining shading aberrations, then calculates corrections based on all tiles of the designated group using BaSiC. Corrections are stored in the metadata repository 1) to avoid recalculation of the same files, and 2) to generate a generalized/universal shading correction for rapid correction without calculation.

We also implemented a batch correction step to correct for systematic error in staining intensity and quality between independent imaging experiments. This code is discussed in depth in the

Batch correction section of chapter 1. This correction step utilizes the metadata extraction and grouping algorithms above to group tiles based on one or more pre-defined categories. The algorithm extracts tile-based channel statistics (min, mean, max, SD, variance, quartiles, automatic threshold values) and histograms, both *in toto* and masked. These values are saved within the metadata folder, and then used to calculate the group-wide channel statistics and histograms. If establishing a standard for normalization, the same statistics are calculated across all groups (i.e. globally) and then saved as a basis for normalizing subsequent images sent through the pipeline. Each group is then adjusted such that the group-wide normalization value matches the global standard. The current study utilized sample format (e.g. slice vs culture) and acquisition date (i.e. one or more scenes imaged without interruption) as the basis for grouping and thus two separate global standards were calculated for slice and culture. We used the non-zero 75th quantile (Q75NZ) as the basis for normalization based on Graf et al., 2021[9] for this study, though we wrote the algorithm to be extensible. Tile, group, and global data are stored in the metadata folder, and therefore each group requires only one round of calculation.

2.4.6. Pixel classification

Corrected and uncorrected images were converted into HDF5 format using the Ilastik4imageJ plugin[10] prior to supervised training or batch application of the classifier model. Custom code was written to standardize this process and apply it in batch. As stated above, we used the Ilastik software to train two pixel classifiers: one for presynaptic signal and one for target signal. The pixel classifiers were used to generate probability maps, where the pixel intensity in each channel correlates with the probability that the pixel belongs to the designated class. Probability maps were exported in the HDF5 format.

2.4.7. Image concatenation

After pixel classification, the input image and both probability maps were imported into FIJI. Probability maps were standardized to be on a 0-100 intensity scale and converted to 8-bit if required. The three images were then concatenated into a 15 channel image and exported as HDF5 files.

2.4.8. Object segmentation

To segment objects of the desired class, we first generated a 0/1 binary mask from the corresponding channel in the probability map. A 0.20 probability threshold (i.e. 20% chance the pixel belongs to the desired class) was chosen as a cutoff for class inclusion. The mask was then applied to the corresponding channel in the input image, thereby removing pixels with low probability of belonging to the desired class. Starting seeds for object identification were generated by finding 3D local maxima on a 3D Gaussian filtered version of the masked image. The masked input was then segmented into objects using the 3D spot segmentation algorithm from the 3D ImageJ Suite (mcib3d plugin)[14]. A 3D size cutoff was applied to ensure objects fit within a designated volumetric range. Finally, the objects were filtered via 2D size filtering in the X, Y, and Z dimensions on our 3D objects to remove objects that fit the volumetric cutoff but were particularly elongated in any one dimension. Segmentation results were saved as binary masks in the TIF format. It is described in greater detail in the Object segmentation section of Chapter 1.

2.4.9. Class mapping

We wrote custom code to map the tabular classification results from the unsupervised model back onto the original binary image used for bouton segmentation. In order to avoid indexing issues, we based the mapping on object coordinates rather than index. The algorithm is rapid and

GPU-based via CLIJ2[15], and could be adapted to map any numerical values in a csv onto a corresponding binarized mask. It is described in greater detail in the Object class mapping section of Chapter 1.

2.5. Object classification

2.5.1. Supervised object classification

Supervised object classification was performed in Ilastik using the concatenated image from image concatenation above as 'Raw' input and the binary segmentation mask from Object segmentation above as inputs. Classified objects were exported as .TIF format, with tabular metrics and object classification exported in the .CSV format. No custom code was used for supervised object classification.

2.5.2. Unsupervised training by autoencoder embedding

Ilastik object metrics were filtered to remove columns containing redundant metrics (e.g. columns with covariance metrics for the same two image channels). Columns related to pixel intensity metrics (e.g. mean intensity, skewness of intensity) and object morphometrics (e.g. object size, convex hull area), but not location (e.g. object centers, bounding boxes), were used as object features for further analysis. Objects comprised of two or fewer pixels, as assessed in the 'Size in pixels' metric from Ilastik, were removed from the dataset. NA values in columns related to kurtosis or skewness of intensities were set to 0. Columns with zero variance, as assessed by *VarianceThreshold* in the scikit-learn library[16], were removed from the dataset.

The resulting dataset was randomly partitioned into three groups: 80% of objects were assigned to a 'training' partition, to determine scaling and train the autoencoder (see below); 10%

were assigned to a 'validation' partition, to assess the training progress of the autoencoder; and the remaining 10% was assigned to a 'test' partition, to assess accuracy across runs of the pipeline.

Ilastik features in the training partition were scaled using the *MinMaxScaler* of scikit-learn, translating and scaling each feature to range between zero and one. The transformation determined from the training partition was applied to the validation and test partitions.

Dimensionality reduction of scaled Ilastik object features was performed using a four-layer feedforward autoencoder[17] implemented in the Keras library in Python[18]. The autoencoder accepts scaled ilastik features as input, and attempts to output the same features. The autoencoder consisted of a two-layer encoder comprised of 64- and 16-node linear layers; and a two-layer decoder comprised of 64- and 432-node linear layers. All layers except the last were followed by rectified linear unit (ReLU) activation functions. The autoencoder was trained on the training partition for 150 epochs on an Nvidia RTX 3080 graphics card. After each epoch, reconstruction accuracy was assessed on the validation partition using mean squared error. The epoch with the lowest validation error was saved, and the resulting encoder was used to reduce the training, validation, and test objects from a 432-dimensional Ilastik feature space into a 16-dimensional reduced embedding space. Autoencoder embeddings were batch corrected with the Harmony library in Python[19].

2.5.3. Unsupervised clustering

Unsupervised object clustering was performed in Python using the Scanpy library[20], using GPU acceleration provided by the RAPIDS library . 16-dimensional autoencoder embeddings were loaded as features of an Anndata object[21], on which subsequent analyses were performed. A k-nearest neighbors network was computed with the *scanpy.pp.neighbors* function (n_neighbors

= 30), and the neighbors network was used the calculated a UMAP embedding[22] of objects with the *scanpy.tl.umap* function.

Louvain clustering[23] was applied to the neighborhood network (*scanpy.tl.louvain*) at a variety of resolutions ranging 0.1 to 3.0. Clustering resolution and network neighborhood size were selected according to the following criteria: maximization of Calinski-Harabasz[24] and minimizing Davies-Bouldin[25] indicies, as assessed by *metrics.calinski_harabasz_score* and *metrics.davies_bouldin_score* in scikit-learn respectively; clusters targeting distinct cellular compartments, based on Ilastik predicted targets; and visual coherence of clusters in UMAP space.

2.6. Statistical analysis

2.6.1. Spatial clustering analysis

We hypothesized that synapses belonging to the same subgroup were likely to be clustered together through Louvain clustering. To test this hypothesis, we generated Spatial Compactness Analysis (SPA), a non-parametric clustering-algorithm-agnostic iterative procedure that evaluates both the compactness of the Louvain clusters and their inter-cluster dependencies (relationships) in the physical (biological) space. To measure cluster compactness, we compute nearest neighbors using the object center coordinates. Specifically, for each cluster, we compute the 10-NN of each sample in the physical space and record their cluster labels. We aggregate these labels over the entire cluster and perform a specific normalization process with respect to a 'null' neighbor distribution. This normalization step is necessary because clusters with more samples would inherently have a higher probability of having observations that would be 10-NNs of observations in other clusters. To obtain the null distribution for a given cluster, we find the distribution of 10-NN neighbors counts proportional to the percentage of total samples in each cluster. The null neighbor computation represents the distribution of neighbors when there is no biological relationship in the

Louvain clustering, and we sample based purely on cluster size. Therefore, dividing the coordinate-based neighbor count by the null neighbor count measures a deviation from randomness. Repeating this process for each cluster enables us to construct a spatial correlation matrix. For any pair of clusters, a ratio greater than 1 indicates spatial correlation, while a ratio less than 1 indicates spatial anti-correlation. A cluster's ratio value with respect to itself indicates the cluster's own spatial compactness.

2.6.2. Statistical comparisons

For most statistical comparisons we utilized built-in functions within Prism. Proportional comparisons between conditions, such as the supervised classifier comparisons in Chapter 4, break some of the fundamental assumptions of ANOVA. In these cases we utilized a generalized linear model (GLM). Relative cell type proportions are naturally distributed in the closed interval [0 - 1]. Therefore, we used a logistic regression with a logit link to model response variables in that interval. However, compositional data tends to be enriched with zeroes, depleted of ones and left-skewed. We modeled this with a quasibinomial distribution that uses an extra parameter to account for overdispersion. We also accounted for heteroskedastic residuals using robust standard errors calculated with White's original method using the R package sandwich[26]. We assessed the model assumptions and the influence of outliers using the R package performance[27]. We computed the 95% confidence intervals based on these robust standard errors using the R package broom. Finally, for each model, we ran multiple iterations randomizing the relative cell type proportions across subjects to assess model calibration and the rate of false positives.

Bibliography to Chapter 2: Materials and methods

1. Taniguchi, H., et al., *A resource of Cre driver lines for genetic targeting of GABAergic neurons in cerebral cortex.* Neuron, 2011. **71**(6): p. 995-1013.
2. Gallerani, N. and E. Au, *Loss of Clustered Protocadherin Diversity Alters the Spatial Distribution of Cortical Interneurons in Mice.* Cereb Cortex Commun, 2020. **1**(1): p. tgaa089.
3. Notter, T., et al., *A protocol for concurrent high-quality immunohistochemical and biochemical analyses in adult mouse central nervous system.* Eur J Neurosci, 2014. **39**(2): p. 165-75.
4. Nunnelly, L.F., et al., *St18 specifies globus pallidus projection neuron identity in MGE lineage.* Nat Commun, 2022. **13**(1): p. 7735.
5. Barron, J.J., et al., *Group 2 innate lymphoid cells promote inhibitory synapse development and social behavior.* bioRxiv, 2023.
6. Linkert, M., et al., *Metadata matters: access to image data in the real world.* J Cell Biol, 2010. **189**(5): p. 777-82.
7. Schindelin, J., et al., *Fiji: an open-source platform for biological-image analysis.* Nature Methods, 2012. **9**(7): p. 676-682.
8. Berg, S., et al., *ilastik: interactive machine learning for (bio)image analysis.* Nature Methods, 2019. **16**(12): p. 1226-1232.
9. Ollion, J., et al., *TANGO: a generic tool for high-throughput 3D image analysis for studying nuclear organization.* Bioinformatics, 2013. **29**(14): p. 1840-1841.
10. Haase, R., et al., *CLIJ: GPU-accelerated image processing for everyone.* Nature Methods, 2020. **17**(1): p. 5-6.
11. Preibisch, S., S. Saalfeld, and P. Tomancak, *Globally optimal stitching of tiled 3D microscopic image acquisitions.* Bioinformatics, 2009. **25**(11): p. 1463-5.
12. Peng, T., et al., *A BaSiC tool for background and shading correction of optical microscopy images.* Nature Communications, 2017. **8**(1): p. 14836.
13. Graf, J., et al., *FLINO: a new method for immunofluorescence bioimage normalization.* Bioinformatics, 2021. **38**(2): p. 520-526.
14. Pedregosa, F., et al., *Scikit-learn: Machine Learning in Python.* ArXiv, 2011. **abs/1201.0490**.
15. Kramer, M.A., *Nonlinear principal component analysis using autoassociative neural networks.* AIChE Journal, 1991. **37**(2): p. 233-243.
16. Chollet, F.a.o. 2015; Available from: https://keras.io.
17. Korsunsky, I., et al., *Fast, sensitive and accurate integration of single-cell data with Harmony.* Nature Methods, 2019. **16**(12): p. 1289-1296.
18. Wolf, F.A., P. Angerer, and F.J. Theis, *SCANPY: large-scale single-cell gene expression data analysis.* Genome Biology, 2018. **19**(1): p. 15.
19. Virshup, I., et al., *anndata: Annotated data.* bioRxiv, 2021: p. 2021.12.16.473007.
20. McInnes, L. and J. Healy, *UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction.* ArXiv, 2018. **abs/1802.03426**.
21. Blondel, V.D., et al., *Fast unfolding of communities in large networks.* Journal of Statistical Mechanics: Theory and Experiment, 2008. **2008**: p. 10008.
22. Caliński, T. and J. Harabasz, *A dendrite method for cluster analysis.* Communications in Statistics, 1974. **3**(1): p. 1-27.
23. Davies, D.L. and D.W. Bouldin, *A Cluster Separation Measure.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 1979. **PAMI-1**(2): p. 224-227.
24. Zeileis, A., *Object-oriented Computation of Sandwich Estimators.* Journal of Statistical Software, 2006. **16**(9): p. 1 - 16.
25. Lüdecke, D., et al., *performance: An R package for assessment, comparison and testing of statistical models.* Journal of Open Source Software, 2021. **6**(60).
26. Zhu, F., et al., *Architecture of the Mouse Brain Synaptome.* Neuron, 2018. **99**(4): p. 781-799 e10.
27. Zhang, X. and V. Bennett, *Restriction of 480/270-kD ankyrin G to axon proximal segments requires multiple ankyrin G-specific domains.* J Cell Biol, 1998. **142**(6): p. 1571-81.

Chapter 3: Image-based multidimensional analysis of synapses

3.1. Introduction

Traditionally, synapses are studied using electrophysiology, neuroanatomical reconstructions, or biochemistry. Although these methods offer considerable benefits, they are not without their limitations. Electrophysiology and neuroanatomical reconstruction are both high resolution techniques that enable investigations of synapses on a cellular and subcellular level *in situ*. Both techniques suffer from throughput limitations, however. Researchers are only able to study individual cells one at a time and, based on the amount of work required to completely process a single cell, datasets rarely exceed a few dozen samples. In contrast, biochemical techniques are high throughput, enabling researchers to study thousands or hundreds of thousands of synapses at a time. These techniques, however, require removal of the synapse from its native context and yield an average measure for the whole population. Synapses are incredibly diverse, however, with differential enrichment in different brain regions[1]. Given the limitations of current techniques, we sought to develop a high-throughput, high-resolution fluorescent image processing and analysis platform to study individual synapses on a population level *in situ*. Specifically, we wanted to be able to study these populations synapses-by-synapse using a wide array of metrics. Moreover, we wanted to be able to use the platform for both hypothesis testing and discovery.

In recent years several genetic tools have become available to identify and manipulate specific interneuron populations. Mouse lines harboring specific Cre drivers and Cre-dependent reporters facilitate irreversible fluorescent labeling and genetic manipulation of specific cortical interneuron populations. Cre-dependent reporter lines enable us to interrogate different properties of these interneurons. The Ai34 reporter labels Cre-positive cells with synaptophysin::tdTomato (Syp-tdT), enabling genetically-targeted visualization of axon termini and presynaptic boutons. This genetic

toolbox therefore enables unprecedented access to specific neuronal subtypes, allowing researchers to study synaptic connectivity at a population level in normal and disease contexts. In addition, advances in machine learning, AI-based statistical modeling, and computing power have enabled researchers to analyze imaging data at unprecedented depth. We combined these technologies to develop an image processing and analysis platform that enables us to study neuronal synapses at a population-wide level *in* situ using hundreds of metrics per synaptic bouton. Coupling these data with a supervised classification model allowed us to robustly classify hundreds of thousands of synapses per run based on canonical subcellular targeting class, showing its utility for hypothesis testing. Further, given the depth of data associated with each bouton, we built in the ability to analyze the dataset using an unsupervised learning model, which in turn revealed an unprecedented amount of diversity within these canonical targeting classes.

The platform can be broken down into five stages (Figure 3.1): 1) input addition and standardization, 2) image correction,[16] 3) pixel classification, 4) object segmentation, and 5) object classification. The first step reads the biological images in their various forms, identifies their intrinsic structure, and breaks them into standardized, individual units for downstream processing. The second step corrects for common issues with fluorescent bio-images, such as optical aberrations and batch-to-batch variability. The third step utilizes machine learning algorithms to classify the fluorescent signals according to biologically meaningful criteria. The fourth step extracts objects of interest from the resultant images, then extracts hundreds of metrics per object. The fifth

[16] Image correction includes both XY-correction of optical artefacts and batch correction to standardize sample-to-sample fluorescence intensity values.

step uses these data to classify the objects based on supervised learning models for hypothesis testing or

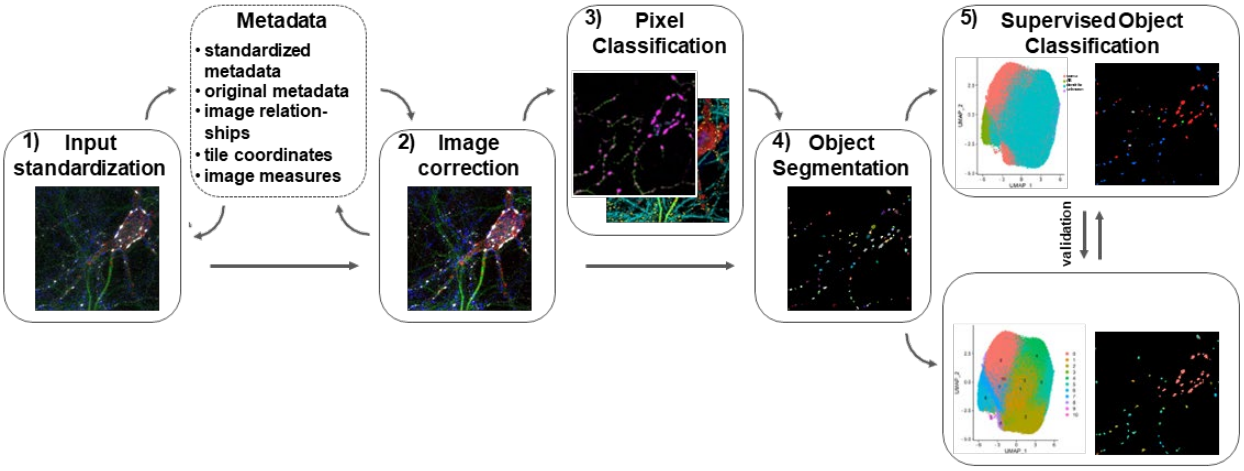

Figure 3.1| A schematic displaying the five stages of the platform.
The platform generally consists of five stages of image processing and analysis:
1) Image standardization: Inputs are added to the pipeline and standardized for downstream processing. This stage identifies the intrinsic structure of the image files, assigns each image in the series a unique name based on the image metadata and nomenclature rules, then converts the images and metadata into standardized formats. The code stores original, standardized, and custom metadata in a central repository. It then uses these metadata to assess whether the current file structure accords with pre-determined experimental parameters. **Note:** different metadata are extracted and utilized at various stages of processing. We display this component of the platform with a dashed outline because, although it is not an image processing stage per se, it is utilized at multiple stages for data storage, quality control, file organization, image correction, and image stitching purposes. It therefore serves as an organizing principle for the various processing stages.
2) Image intensity correction: Images are corrected for issues common to fluorescent bio-images, such as optical aberrations and batch-to-batch intensity variability. This stage utilizes the hierarchical and categorical relationships identified in the first stage to group images for processing. Measures from individual tiles, image series, and experimental conditions are stored in the metadata folder. The conditions-wide metrics are initially derived retrospectively (i.e., from the images themselves) and set as a standard for prospective correction of subsequent image series.
3) Pixel classification, which utilizes machine learning models to classify the raw fluorescent signals according to biologically meaningful criteria.
4) Object segmentation: this stage identifies objects within the pixel-classified and raw images. The segmentation algorithm is modular and therefore malleable to different object shapes as required for experiment requirements. Also generated is the measurand image, which serves as the image from which intensity-based object metrics are derived during the object classification stage.
5) Object classification: This stage measures object statistics from the segmentation and measurand images, then uses these data to classify objects. Object classification can use supervised learning models (***5a***) for hypothesis testing, or unsupervised models (***5b***) for discovery. The former fits

the objects to user-designated classes based on a user-trained model, whereas the latter classifies objects based on the internal structure of the dataset itself. Of note, these approaches are not mutually exclusive, and can be utilized in parallel to validate each other.

unsupervised models for discovery. Different metadata are extracted and utilized at various stages of processing. Overall, these stages allow for multiple machine learning-based steps to process and analyze fluorescence microscopy images, yielding hundreds of spatial and morphological metrics per pre-synaptic bouton. This approach and the depth of data obtained enabled us to robustly classify synapse subtypes based not only on pre-determined classes for hypothesis testing, but also using a hypothesis-blind approach for discovery of novel bouton subtypes.

3.2. Experiment design

3.2.1. Initial question and impetus to create the pipeline

Our experimental question was initially based on a biological question of synaptic targeting. Briefly, different subtypes of cortical GABAergic interneurons synapse onto distinct subcellular regions of the target cell (Figure 3.2, A). The three subclasses of interest are dendrite-targeting somatostatin (Sst) cells, the soma- and proximal dendrite-targeting basket cells, and axon initial segment (AIS)-targeting chandelier cells. Presented practically, we wanted to know whether these cortical interneuron subclasses recapitulated subcellular specificity in a culture dish. The results from these experiments are discussed at length in chapter 2, but we realized very early on in our investigations that quantifying our results by manual annotation would be, from a practical standpoint, impossible. In culture, a single interneuron can form over 3000 synapses; we were labeling thousands of cells in a single well. In slice, a single 320 μm x 320 μm x 20 μm 3D tile image[17]

[17] Volumetrically, this is approximately 2 nL.

easily contained tens of thousands of labeled interneuron synapses; we would have to image hundreds of tiles to avoid region-to-region or sample-to-sample bias. And GABAergic interneurons are a *minority* population in the cortex! In order to make any statistical argument for or against our hypothesis, we would have to automate the classification and counting process. Doing this required considerable trial and error, but we believe that the platform we present here, though still in its infancy and not without its faults, not only solves our synaptic bouton classification problem, but also sheds new light on synaptic bouton classification altogether.

3.2.2. Labeling the presynaptic compartment: reporter selection

In order to test our hypothesis, we had to label two aspects of our sample: 1) the presynaptic boutons of the different interneuron classes, and 2) the different subcellular target compartments. For the former, we utilized mouse genetic tools based on the Cre/loxP driver/reporter system to differentially label distinct interneuron cell types. As diagrammed in Figure 3.2, B, by using a genetically engineered 'driver' line, Cre is selectively expressed based on endogenous expression patterns of the 'driver' gene, here Sst^{Cre}. Once expressed, Cre translocates to the nucleus, where it targets *loxP* sites in another genetically engineered 'reporter' allele for recombination. Here, the *loxP* sites flank a stop cassette that blocks downstream transcription. Cre-mediated recombination excises the stop cassette, leading to expression of a downstream reporter protein. Since a) the reporter is under control of a constitutively active promoter (here *CAG*), and b) recombination is an irreversible genetic modification, Cre recombination leads to a) permanent and b) heritable reporter expression. The specific interneuron driver lines we chose and our reasoning for doing so are discussed in chapter 2. This reasoning is summarized in Figure 3.2, C. Detailed information on mouse lines used in this study can be found in Table 1. Note that, unless otherwise stated, all

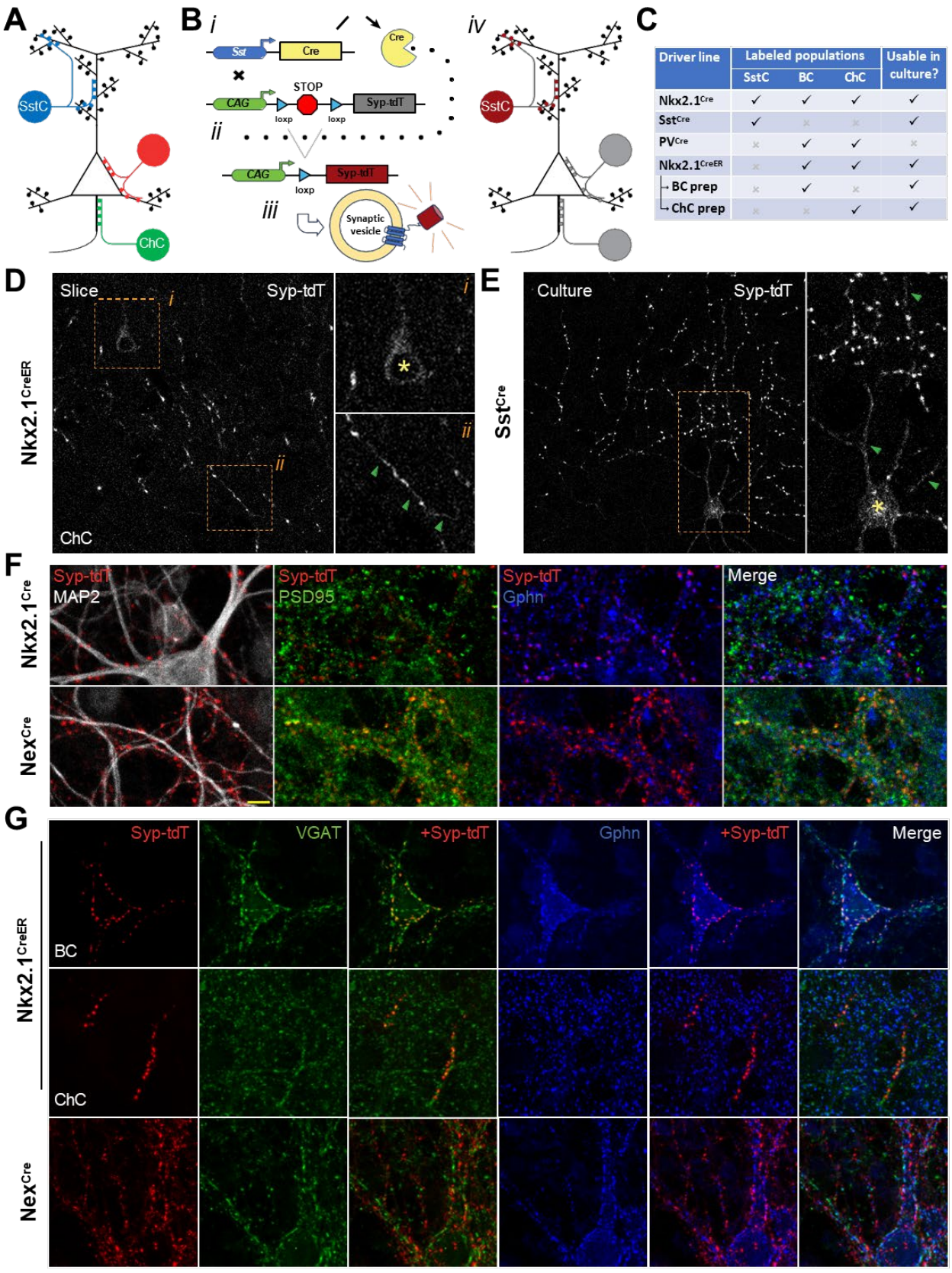
A
SstC
ChC
B
i
Sst
Cre
Cre
STOP
CAG
loxp
loxp
Syp-tdT
ii
CAG
loxp
Syp-tdT
iii
Synaptic vesicle
iv
SstC
C
Driver line
Labeled populations
SstC
BC
ChC
Usable in culture?
Nkx2.1Cre ✓ ✓ ✓ ✓
SstCre ✓ × × ✓
PVCre × ✓ ✓ ×
Nkx2.1CreER × ✓ ✓ ✓
→ BC prep × ✓ × ✓
→ ChC prep × × ✓ ✓
D
Nkx2.1CreER
Slice
Syp-tdT
i
ii
ChC
E
SstCre
Culture
Syp-tdT
F
Nkx2.1Cre
NexCre
Syp-tdT
MAP2
Syp-tdT
PSD95
Syp-tdT
Gphn
Merge
G
Nkx2.1CreER
NexCre
Syp-tdT
VGAT
+Syp-tdT
Gphn
+Syp-tdT
Merge
BC
ChC

Figure 3.2 | Presynaptic cell labeling strategy
A) Schematic of the three interneuron classes and their subcellular targets.
B) Schematic of the genetic labeling strategy. ***i***. An interneuron class-specific Cre driver (here Sst-^{Cre}) expresses cre upon promoter activation. ***ii***. Cre excises the a stop codon flanked by loxP sites (LSL sequence) in a reporter gene (here Ai34). The LSL sequence is positioned between a constitutively active promoter (CAG) and a fluorescent reporter gene (Syp-tdT). ***iii***. Cre-mediated LSL excision leads to permanent and heritable expression of Syp-tdT, which localizes to synaptic vesicles. ***iv***. This enables presynaptic bouton visualization in cells that have activated the Cre driver.
C) Table of Cre driver lines, labeled populations, and utility in culture. Nex^{Cre}, which labels cortical glutamatergic neurons, is not included here. This driver was only used for validation experiments presented here.
D-E) examples of 'off-target' labeling in the Ai34 reporter line in slice and culture. D) Syp-tdT expression profile from a ChC-enriched region of the Nkx2.1^{CreER};Ai34 mouse in slice. Insets show magnified images of the indicated regions. Labled soma is indicated by asterisk and neurites by green arrows. E) Syp-tdT expression profile from a Sst^{Cre};Ai34 mouse in culture. Inset shows a magnified image of the indicated region, with annotation as in D.
F) Syp-tdT labels presynaptic boutons that target the expected postsynaptic densities (PSDs). Nkx2.1^{Cre};Ai34 (upper) labels all MGE-derived GABAergic interneuron populations; Nex^{Cre};Ai34 (lower) labels cortical glutamatergic neurons. PSD95 is a glutamatergic postsynaptic marker; Gphn is a GABAergic postsynaptic marker.
G) Syp-tdT labels presynaptic boutons of the expected neurotransmitter type. Nkx2.1^{CreER};Ai34 samples from BC or ChC-enriched regions labels GABAergic interneurons and Nex^{Cre};Ai34 (lower) labels glutamatergic cells. VGAT is a GABAergic presynaptic marker; Gphn is a GABAergic postsynaptic marker.

further references to a driver line imply the presence of the Ai34 reporter allele. For example, *Nkx2.1^{CreER}* should be understood as *Nkx2.1^{CreER};Ai34*.

Since we used a fluorescent reporter line and are here discussing the image processing and analysis platform, a discussion of the reporter is fitting. We used the Ai34 reporter line (Figure 3.2, B), which expresses the synaptophysin::tdTomato (Syp-tdT) chimera protein following Cre-mediated excision of a stop cassette. Synaptophysin localizes to neuronal presynapses, whereas tdTomato is a bright red fluorescent protein, so the Syp-tdT chimera localizes tdTomato to presynaptic boutons. In sum, coupling a Cre driver line with the Ai34 reporter allowed us to selectively

label presynaptic boutons of different interneuron populations with a red fluorescent protein. Example images of the Syp-tdT signal in slice and culture are shown in Figure 3.2, D and E, respectively. It is important to emphasize here that we used the *same reporter* to label *different populations*. This methodological parity in labeling allows us to make direct comparisons between different synaptic populations. We know, for instance, that among different interneuron populations labeled using the same Ai34 reporter, any changes in Syp-tdT-labeled bouton morphology or intensity are due to biological differences between these populations, and not technical differences in labeling. Fewer uncontrolled variables yield a better experimental design and a clearer interpretation.

Some variables are beyond our control, however. We chose to use the Ai34 reporter line for our studies because of its specificity for labeling presynaptic boutons. During validation experiments, however, we witnessed low level Syp-tdT signal in neurites and soma of labeled cells (Figure 3.2, D and E, arrowheads and asterisks, respectively). Syp-tdT localization to somata was witnessed in both tissue slices and cultures (arrowheads). Although the off-target labeling of neurites was evident in slice (Figure 3.2, D, arrowheads), it was more prominent in cultures (Figure 3.2, E, arrowheads), with more immature cultures in particular displaying greater Syp-tdT localization to neurites (data not shown). This suggests that synaptophysin becomes increasingly more specific to presynaptic boutons with progressive neuronal maturation. These conclusions were further supported by counterstaining with pre- and post-synaptic markers, as discussed below. Although this off-target labeling was initially a nuisance, we not only subverted, but even exploited it to extract more information than initially anticipated. We did this by incorporating a *post-hoc* pixel classification step, which we discuss in greater detail in the Pixel Classification section below. Despite our ability to account for, correct, and at times even exploit such specificity issues, the need to

incorporate this step is illustrative of the principle that decreased labeling specificity requires increased *post-hoc* compensation during image processing, and vice versa. Generally speaking, less *post-hoc* image processing implies a more robust experimental system.

We next sought to validate this reporter for interneuron synaptic targeting. To do so, we stained cortical Sst^{Cre}, $Nkx2.1^{CreER}$, $Nkx2.1^{Cre}$, and Nex^{Cre} cultures for VGAT, gephyrin, and PSD95 (Figure 3.2, F-G). Nex^{Cre} labels cortical glutamatergic (but not GABAergic) neurons, and therefore labels glutamatergic presynaptic boutons. It served as a positive control for glutamatergic markers and negative control for GABAergic ones. We used VGAT as a marker of GABAergic presynapes, gephyrin for GABAergic postsynaptic densities (PSDs), and PSD95 for glutamatergic PSDs. Shown are images from Sst^{Cre} and Nex^{Cre} cultures. As expected, Syp-tdT boutons from the $Nkx2.1^{Cre}$ mouse were juxtaposed with gephyrin but not PSD95, and vice versa for Nex^{Cre} (Figure 3.2, F). Sst^{Cre} and $Nkx2.1^{CreER}$ displayed similar results as $Nkx2.1^{Cre}$ (data not shown). Also as expected, Syp-tdT boutons from $Nkx2.1^{CreER}$ BC and ChC preps were VGAT-positive and showed high apposition to gephyrin but not PSD95. In contrast, Syp-tdT boutons from the Nex^{Cre} culture were VGAT-negative and showed high apposition to PSD95 but not gephyrin. Sst^{Cre} and $Nkx2.1^{Cre}$ displayed similar patterns of juxtaposition as $Nkx2.1^{CreER}$ (not shown). These results help to validate our genetic approach to label GABAergic presynapses using different interneuron-specific drivers and the Ai34 reporter. In addition, Syp-tdT signal previously identified as 'neurite' was neither VGAT-positive nor juxtaposed with gephyrin, further supporting our conclusions of 'off-target' labeling. These experiments display an orthogonal approach to validate tools and verify conclusions.

3.2.3. Labeling the target compartments: antibody selection

We next turned to visualizing the subcellular target compartments. A schematic of the compartments and the antibodies used to visualize them is shown is Figure 3.3, A. We used Ankyrin G (AnkG) as a marker of the AIS, Gephyrin as a marker of GABAergic PSDs, and combined staining of Kv2.1 and Kv2.2 (Kv2) as a general marker of the soma and proximal dendrites. Examples of the staining patterns for these antibodies are shown for culture and slice (Figure 3.3, A). Since we are here discussing principles for choosing labels, it is worthwhile to delve into our reasoning and preliminary experiments for choosing this particular antibody panel to test our hypothesis.

We used AnkG as a marker of AIS. We based our decision here largely on wide body of literature associated with this protein at the AIS[2-19]. Simply put, AnkG is the canonical marker of this subcellular compartment, antibodies directed at AnkG are well-validated, and these antibodies yielded the expected results in our culture system and in slice (Figure 3.3, A).

We used a combination of Kv2.1 and Kv2.2 monoclonal antibodies in the same channel to label soma and proximal dendrite[20-25]. We based this choice on literature findings: individually these proteins have distinct cell- and layer-specific expression patterns in pyramidal cells[26]. We reasoned that visualizing both Kv2 antibodies in the same channel would limit the potential for false negative classification of soma targeting boutons. This combination yielded the expected results in both culture and slice (Figure 3.3, A). Henceforth we refer to the combined labeling of Kv2.1 and Kv2.2 simply as Kv2.

Labeling the dendritic target compartment was more challenging. MAP2 is a canonical marker of dendritic shafts[27, 28], and we initially tested antibodies to MAP2 as a direct label of dendrites. We encountered several issues, however. First, when counterstained with drebrin and gephyrin,

markers of dendritic spines and inhibitory PSDs, respectively, we witnessed a number of GABAergic boutons that were not juxtaposed with MAP2 (Figure 3.3, B). This result was in part expected.

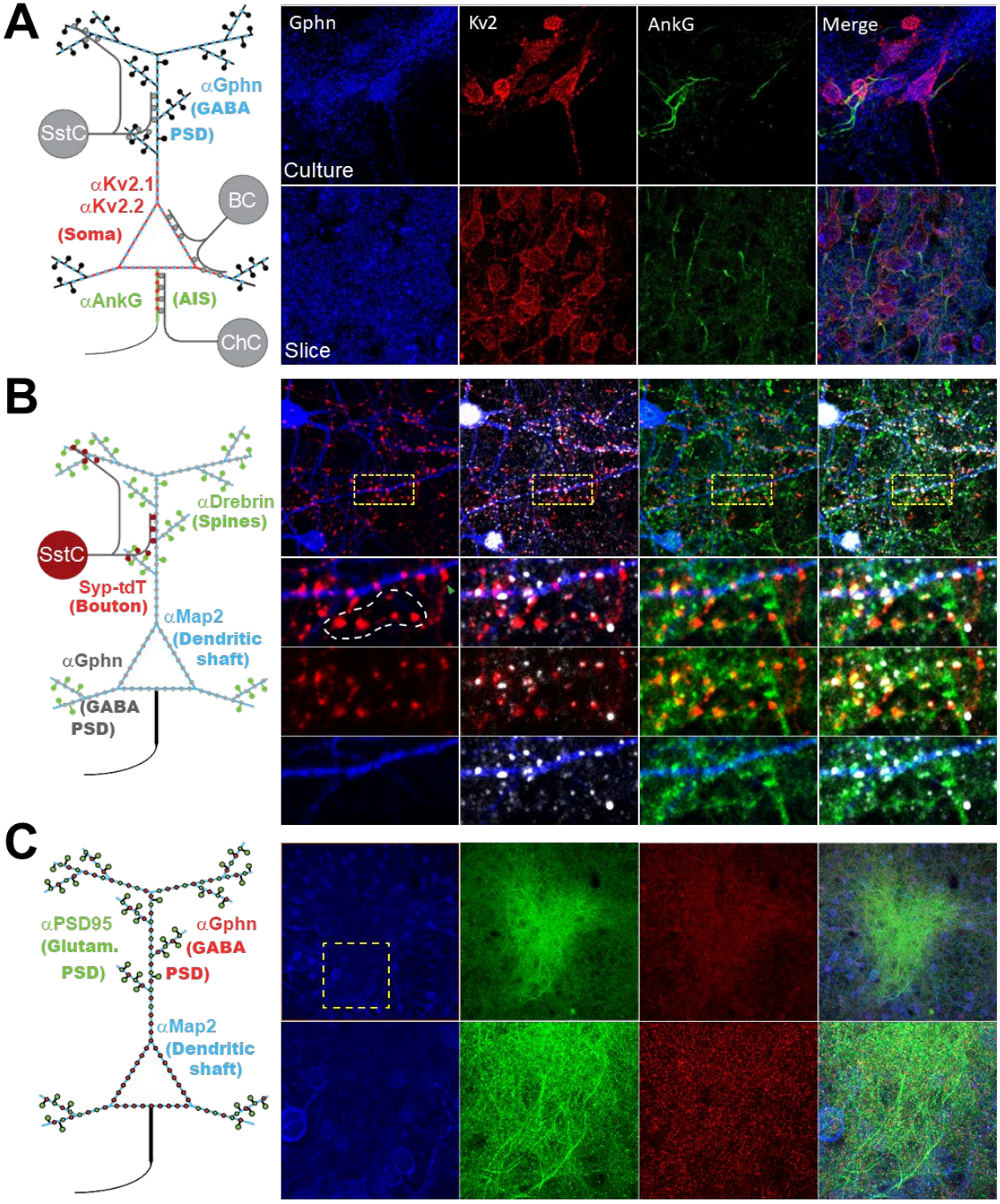

Figure 3.3 | Target Compartment labeling strategy.
A) Final panel used to visualize the target compartments. *Left*: A schematic of the cortical interneuron populations and their subcellular targets. Displayed are the antibodies used and the most prominent subcellular localization of their targeted proteins. *Right*: Example staining images in culture (*upper*) and slice (*lower*). Signals are colored to match the schamatic.
B) MAP2 does not label dendritic spines targeted by SstCre cells. *Left*: Schematic of the antibody labeling strategy to assess dendritic spine targeting by SstCre interneurons. *Right*: Example an Sst^{Cre};Ai34 dissociated culture stained as in the schematic. The region of interest is magnified and shown in the lower panels. Note that the boutons circled in white are juxtaposed with drebrin and gephyrin, but not MAP2. Unlike the boutons on the dendritic shaft above, these boutons target dendritic spines.
C) MAP2 staining intensity decreases in high synaptic density regions. *Left*: Schematic of the antibody labeling strategy to assess synaptic density in low intensiity MAP2 regions witnessed in culture. *Right*: Example a high density growth region in dissociated culture, stained as in the schematic. The region of interest is magnified and shown in the lower panels. Note that the region with high PSD95 and Gphn staining has low MAP2 intensity.

Although inhibitory dendritic input primarily targets dendritic shaft, a sizeable subset of GABAergic boutons target spine necks and heads[29, 30], which are not labeled by MAP2[26, 29-33]. Second, we also noted an unexpected decrease in MAP2 intensity in high density growth regions in culture (Figure 3.3, C). These low intensity regions were evident with three different MAP2 antibodies (data not shown), indicating this staining pattern is independent of the antibody used. These low-signal regions were inversely correlated with PSD95 and gephyrin signal, indicating that this was not a technical issue arising from decreased antibody penetrance, but rather a biological phenomenon due to lower levels of the MAP2 epitope within these regions. Given the inverse correlation with PSD95 staining, we suspect the decreased MAP2 signal may be activity-dependent, as has been shown previously in slice[34]. Based on these issues with MAP2 we decided it was not a reliable indicator of the dendritic subcellular compartment for our purposes.

We therefore decided to use gephyrin, a general marker of inhibitory post synaptic densities (PSDs) instead. Unlike MAP2, gephyrin signal was not affected by cellular density (Figure 3.3, C) and labeled spines targeted by GABAergic synapses (Figure 3.3, B). In addition, our extensive

use of gephyrin in validating our genetic approach to labeling GABAergic presynapses (Figure 3.2, F-G) provided further confidence in decision to move forward using this antibody. We therefore settled on our final antibody panel of Ankyrin G, Kv2, and gephyrin to label the three subcellular compartments targeted by ChCs, BCs, and SstCs, respectively (Figure 3.3, A). Since gephyrin PSDs are present at all target compartments, we defined any bouton with positive gephyrin signal but negative for both Kv2 and AnkG as dendrite-targeting. We defined any bouton with no gephyrin, Ankyrin G, or Kv2 signal as unknown-targeting. As discussed below, the unknown-targeting class comprised less than 1.5% of all boutons, validating this staining approach as largely comprehensive. Although we found this combination generally sufficed, as outlined in greater depth in the Pixel Classificiation section below, we did encounter some specificity and sensitivity issues. We therefore trained a second supervised pixel classifier (Figure 3.1, #3) to compensate for and/or exploit these issues, thereby providing a more comprehensive and representative picture than was possible with the target staining alone.

3.3. File organization

3.3.1. Nomenclature[18]

For the pipeline to be functional and its outputs to be meaningful, we needed to be able to track an image and its associated data across different stages and in various formats. A single source image added to the pipeline can yield dozens of associated data files in various formats. Systematically naming these files in a consistent and predictable manner was key to organizing and associating data from the same sample. In other words, our structured system of nomenclature served as an organizing principle upon which we built the pipeline. Moreover, the input and output

[18] The nomenclature, including specific definitions for each field, is outlined in greater depth in Appendix B. Here we discuss the rationale behind it and how it was utilized to streamline processing.

data are organized both hierarchically and categorically, which is reflected in our naming system. Understanding this organization is critical for understanding the text that follows, and therefore also the platform itself. It is for this reason that nomenclature is a critical step and requires discussion.

We took a structured approach to sample nomenclature that would ensure the filenames

1. are human-readable and meaningful,
2. reflect the hierarchical nature and/or the experimental categories of the samples, and
3. are unique at each processing stage, but
4. contain elements that persist throughout all stages of processing.

These points are best illustrated by the nomenclature structure itself, and how it plays out in a concrete example. We systematically named each image or its associated data according to the following pattern:

[Experiment ID]_[Sample ID]_[Biological conditions]_[Staining conditions]_[Scene]_*m*[Tile]_[variable stem ID]***

For example, the following slice sample from mouse #7441:

20201120_7441_SstCre-Ai34_AnkG-Kv2-Gphn _TR1_m03_BatchQ75NZ.tif

As per point #1 above, critical experiment parameters are immediately recognizable within this filename: it includes a date (2020-11-20, here the mouse's birthdate), the mouse ID (7441), and its genotype (Sst^{Cre};Ai34), as well as the staining conditions for the sample (it is a tissue slice stained for AnkG, Kv2, and Gphn). As per point #2, it also includes important grouping/hierarchical information: the specific source image is the third tile (m03) from the first tile region/scene (TR1). These portions, which have been emphasized in bold, comprise the image tile's ***base name***, which is kept constant at all stages of the pipeline and is present in every output (point #4). Importantly, this base name is assigned at the first stage in the pipeline, when multi-tile scenes (or tile regions) are split into individual tiles (the processing unit of our pipeline). This specific file's

position in the pipeline is designated by the *[variable stem ID]* portion, shown above in italics: it is the output image from batch-correction using the 75th (non-zero) quantile (*BatchQ75NZ*.tif). This designator comprises the ***stem*** of the filename, which is varied in a predictable manner and used to distinguish data from different stages of the pipeline that are associated with the same source tile. Appending the variable (but non-unique) stem to the constant (but unique) base yields a unique name for each file at every stage of the pipeline (point #3).

Beyond simply containing meaningful information, this systematic nomenclature also has considerable utility. Most importantly, the constant base plus variable stem structure is the basis for automatic file/data handling throughout the pipeline. Each stage of the pipeline was designed to batch process/analyze any number of tiles, i.e., the user is able to designate any number of tiles as inputs, and the pipeline will process them sequentially without any further user input. Input lists for a single run, therefore, can easily reach lengths of hundreds or thousands of files. To complicate this picture further, as outlined in the pipeline schematic (Figure 3.1), various stages of the pipeline can require multiple input files to process any given tile. Each of these inputs is handled in a distinct way during processing: improper designation will yield errors during processing or erroneous outputs. So it is critical that the distinct types be properly designated at the outset. Therefore, two forms of file alignment must occur here to process any given tile correctly:

1. *tile alignment*: distinct inputs from the *same tile* (e.g. 7441_[…]_TR1_m03_*) must be aligned so they are processed together, and
2. *type alignment*: inputs of a *specific type* (e.g. input image vs presynaptic output vs target output) must be identified as *that* input type, such that each is handled properly during processing.

Our nomenclature system solves both alignment problems. First, since the base/tile name identifies the tile and is consistent throughout the pipeline, it is easy to identify which files are associated with which tile at any stage simply by matching bases. Second, since distinct data types are predictably designated with distinct stems, it is easy to identify an input as *that* type because it has *that* stem. For instance, [*tilename_mX*]_Presynaptic.h5 and [*tilename_mX*]_target.h5 are easily identified by user and computer as two different forms of data from the same tile. Thus, *tile alignment* is reduced to *base/tile name matching*, whereas *type alignment* is reduced to *stem matching*.

Even with such a system, manually aligning inputs is inefficient and error prone, especially if dealing with hundreds or thousands of tiles in a single run. Since the file nomenclature follows a designated set of rules, we can utilize it to perform these alignment steps automatically. Each stage of our pipeline performs the *tile alignment* step automatically in seconds, thereby avoiding issues from data misalignment: tiles missing required inputs are skipped automatically, as are tiles with multiple inputs of the same type. We use *type alignment* checkpoints to ensure that each input matches the expected type: files with unexpected stems are rejected, thereby avoiding pipeline errors and erroneous results.

Beyond file/data handling, this filename pattern also facilitates automatic grouping of related (i.e., hierarchically-associated) files. Another important consequence from our nomenclature system is that it facilitates integration of data gathered within the pipeline with data that exists outside of it. We discuss our utilization of the nomenclature for automated tile grouping in greater detail in Appendix B. In summary, our nomenclature system provides each image with a unique identifier, similar to a barcode, that contains structural information about sample organization. This 'barcode' allows us to store each object and its associated data within a database, thereby facilitating downstream analysis of the entire dataset.

3.3.2. Adding an input: standardization of metadata, files, and file structures

We sought to create a pipeline that could be adapted for any number of biological problems, with as few limitations as possible on the imaging equipment or the image structure used. Biological images, however, are frequently saved in proprietary, microscope-specific file formats with non-standardized internal file structures. Moreover, the diversity in biological questions implies that acquisition parameters themselves can be highly variable, e.g., different questions may require different magnifications, image dimensions, file structure, etc. The experimenter must be able to designate those criteria at the outset, (and so the pipeline itself must be able to accommodate such changes). Further, once the experimental design has been determined, there must be checks in place to ensure that the input data adheres to the required criteria. Simply put, the input data must fit within pre-defined set of criteria: non-standard inputs will lead to processing errors, or worse, erroneous results.

Critical for image standardization is knowing how the image file is structured and how it was acquired[35-38]. This information is contained in the image metadata, which details how the image was acquired and how the file is internally structured. Examples of such data include the *image dimensions* (the size, in pixels or voxels, of the image in the X, Y, Z, C (channels), and T (time) dimensions), *calibration values* (the X, Y, and Z voxel sizes in µm, and the time interval in ms), and *scene information* (the total number of tiles in the scene, as well as their names, order of acquisition, and X, Y, and Z locations in µm on the microscope stage). These metadata are contained as key-value lists within the image file. Problematically, manufacturers do not follow a standardized practice for *what* they list or *how* they list it[39]. For example, the width of 512x512 pixel image may be stored as *width=512* by one manufacturer and *sizeX=512* by another. Thus, the very information needed to standardize input images is itself unstandardized.

The very first problem we encounter, then, is standardization. The above discussion boils down to three stages of standardization, all of which are performed in the first step of the pipeline:

1. Standardization of metadata: the pipeline should be able to identify the file's internal structure and acquisition parameters (i.e. metadata) and save this information in a standardized form at a standard location.
2. Standardization of file formats: the pipeline should be able to read non-standardized file types and convert them into a universal format.
3. Standardization of file structure: for any inputs that do not meet predetermined, standardized structure, the pipeline should be able to convert them to that form if possible, or reject them if not.

We addressed the first two standardization problems by positioning the Bio-formats software at the interface between microscope output and pipeline input. Bio-formats utilizes an expansive library to read biological imaging data from over 150 proprietary and non-proprietary image formats, and to write these data to standardized, open file formats[40]. It is therefore specifically designed to solve the first two standardization problems.

The first step in our pipeline addresses the first problem listed above, i.e., it extracts metadata from the source image and stores it in a standardized form and in a standardized location within a central repository (Figure 3.4, A). We provide a comprehensive list of the extracted metadata fields in table 6. Some examples of these data include the *image dimensions* (the size, in pixels or voxels, of the image in the X, Y, Z, C (channels), and T (time) dimensions), *calibration values* (the X, Y, and Z voxel sizes in μm, and the time interval in ms), and *scene information* (the total number of

tiles in the scene, as well as their names[19], order of acquisition, and X, Y, Z locations in μm on the microscope stage). Importantly, the specific set of stored metadata values could be adapted to researcher requirements. These data are extracted primarily from the standardized Bio-formats metadata, but when unavailable there, are either pulled from the original image metadata or derived from metadata that is present. These aggregated data are automatically appended to a *source metadata table* (a CSV file) located at a standardized location within a central metadata repository. Also stored in the repository are the *Bio-formats/OME metadata* (as an XML file), the *original image metadata* (as a list in TXT format), and if present in the original metadata, an *audit trail* that documents post-hoc changes to the source image that occurred prior to entry into the pipeline (as a TXT file). Finally, we use the extracted metadata to automatically derive information required for downstream steps of the pipeline, such as the *tile configurations map*, which designates how the tiles of a multi-tile image are physically arranged relative to each other, or the *tile groupings*, which automatically classifies tiles into specific groups based on metadata and nomenclature (both discussed later). This information is then structured as needed and stored at predetermined locations within the metadata directory for ease of access downstream.

The second step in the pipeline addresses the second problem above, namely standardization of file format. We use Bio-formats to import the user-designated, microscope-specific source images into FIJI, where we can then process it as if it were any other image (Figure 3.4, B). We use the above metadata to designate commands for *how* to open the file, and what (if any) processing steps are needed to make the file conform to the predesignated standards. For example, multi-tile scenes are split automatically into individual tiles (the unit of processing throughout the remainder of the pipeline), and renamed according to their pre-designated tile name. This ensures that, from

[19] Note that the *scene* and *tile* names are derived directly from the image metadata, at the very first step in the pipeline.

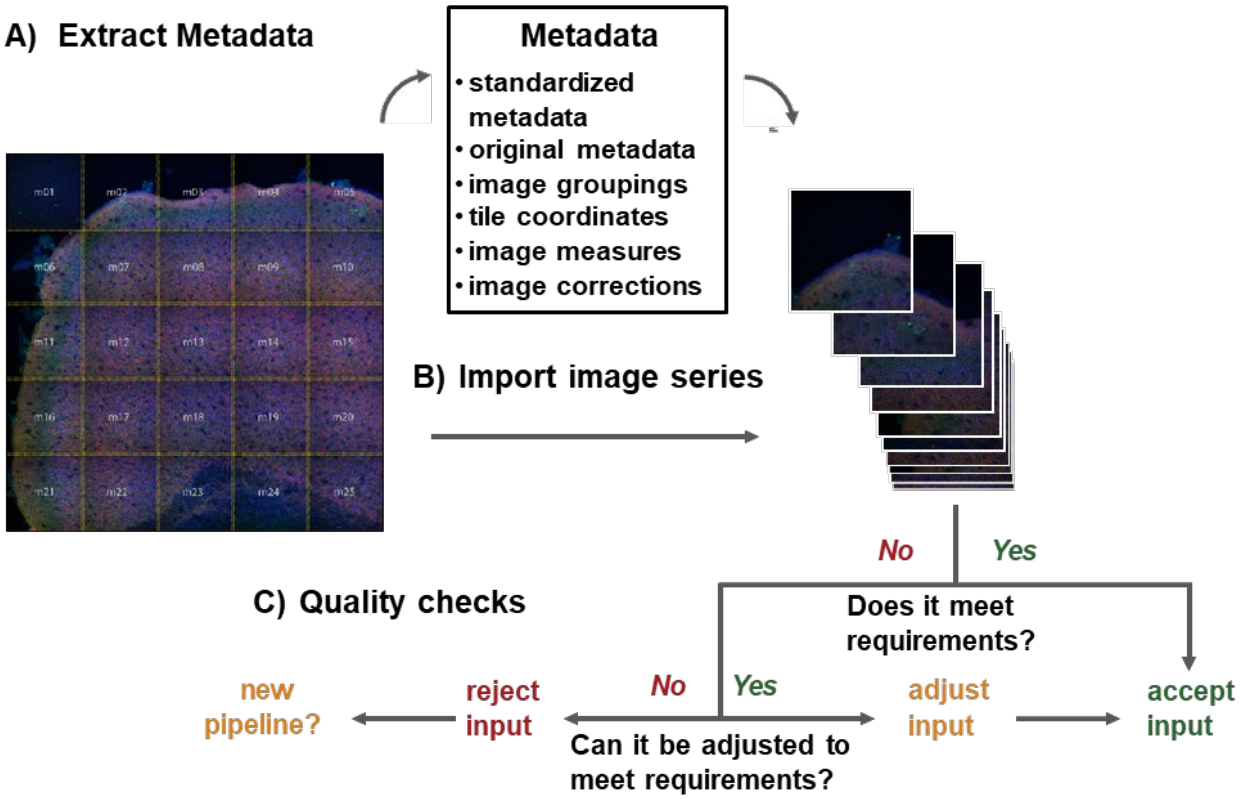

Figure 3.4 | Input standardization and quality control checks.
A flowchart of the initial stage of the platform, where inputs are imported and standardized. Different bioimaging experiment require different experimental parameters and image structures, with file formats largely dependent on the microscope used for acquisition. This means that the metadata, file, image, and sample structure are all variable. This stage standardizes the image metadata and file structure, derives the sample structure, then performs quality control checks to ensure that the inputs accord with a pre-designated image structure.
A) Metadata extraction. The platform extracts the original image metadata, Bio-Formats (i.e., standardized) metadata, and custom metadata from the input images. It identifies hierarchical and categorical relationships between tiles within the acquisition series and between other image series already in the pipeline, and generates a generalized tile configuration file to automate image stitching. These data are stored in a central repository in standardized forms.
B) Image import. The images in the acquisition series are imported and split into individual tiles (the processing unit of the pipeline), then algorithmically assigned unique names using a standardized nomenclature based on the image metadata.
C) Quality checks. The image structure, such as its dimension numbers, sizes, and order, are compared to pre-designated standards. If the image does not meet the requirements, it is adjusted to meet those standards if possible or rejected from further analysis if not. Tiles are saved in a standardized and generally accessible file format at a standard location relative to the input directory.

the outset, the inputs into the pipeline conform to the nomenclature expectations. Ultimately, the

standardized tile is saved in the generic and universally accessible TIF format within a standard directory for ease of access.

The third step of the pipeline runs the input through a series of quality control (QC) stages to ensure the input images conform to the predesignated criteria (Problem 3) (Figure 3.4, C). These steps assess the following fundamental aspects of the file structure: image dimensions, pixel/voxel size, channel order, and dimension order. Images are automatically adjusted to meet requirements when possible (e.g. images are scaled to match required pixel sizes, channels are rearranged if misaligned), and rejected when not possible (e.g. missing channels, 2D images for a 3D pipeline, static images in a time-lapse pipeline). These QC and standardization steps not only ensure that inputs can be handled during downstream processing, but also that data from independent experiments are comparable if run through the same pipeline.

3.3.3. File Grouping and Data Organization

As described above, the first stage of our pipeline splits multi-tile scenes into individual tile images. This splitting process greatly facilitates downstream processing, but it also ignores the hierarchical structure of the data itself. There are times, however, when the tiles must be re-grouped according to various nodes in this hierarchy: sample-to-sample intensity correction and stitching of individual tiles into a scene both depend on the hierarchical relationships within the data. This splitting function also ignores categorical relationships within the data: samples with the same staining conditions or sample format (e.g., slice vs culture) have the same experimental conditions, and therefore belong to the same group. Our pipeline utilizes the image metadata and filename structure to identify these hierarchical and categorical relationships automatically.

We have written code to automatically parse the filename and group tiles both hierarchically and categorically based on nomenclature patterns and metadata values. We discuss this code and

provide examples for its utilization in Appendix B: Sample nomenclature. We utilize different tile groupings at almost every stage of the pipeline, from pre-processing to analysis. Simply put, these automated groupings facilitate automated decision-making throughout the pipeline. As discussed in the Batch correction section below, we use categorical groupings to shuttle different samples into different processing steps when required and hierarchical groupings to rapidly and efficiently align thousands of files from numerous inputs. The more sophisticated the pipeline becomes, the more important it will be to automatically identify and carve out different groupings.

We have, in effect, created an ecosystem for images, their metadata, and the resultant image data. Each object is provided a unique identifier and automatically grouped categorically and hierarchically. This allows images and their associated data to be tracked throughout the processing pipeline. Data from one stage can be easily accessed and utilized for subsequent stages. This organization should greatly facilitate imaging experiments.

3.3.3.1. Stitching

A clear example of the utility of this ecosystem is image stitching. Stitching takes individual tile inputs and organizes them based on their acquisition coordinates into a single, fused scene (Figure 3.5). The stitching process itself is automated, and ultimately uses a native stitching function within FIJI to stitch individual tiles into a scene. To perform this operation, the stitching function needs to know the name and X, Y, and Z coordinates of each tile within the scene. The *tile configuration map* serves as this stitching blueprint; it is a TXT file that structures these data in a standardized way to make it readable for the stitching function. Simply put, the tile configuration map designates which files the stitching algorithm needs to open and where each image goes within the scene. We encountered two problems for the stitching algorithm: 1) the tile configuration map is not automatically generated when scenes are split, i.e. we had to write code to build one for each

scene, and 2) Since the filenames in the tile configuration map must match the tile filenames in the directory exactly, each (functional) tile configuration map only works for the specific files it designates, i.e., we would need different tile configuration files to stitch the same scene at different stages of the pipeline.

We solved the first problem using the image metadata. Embedded within the source metadata is the X, Y, and Z position in μm of the objective on the stage at onset of imaging. As part of the metadata extraction algorithm, we convert these positions to pixels/voxels, then set them relative to the scene origin (the format required by the stitching algorithm). We solved the second problem using our nomenclature system. Simply put, for each scene we create a *tile configuration map* with generalized *tilenames*, and is used as a blueprint to stitch tiles into their corresponding scene at any stage within the pipeline. Storing these data at a standardized location within the central metadata repository enables algorithmic access to these data at any stage of the pipeline. We have

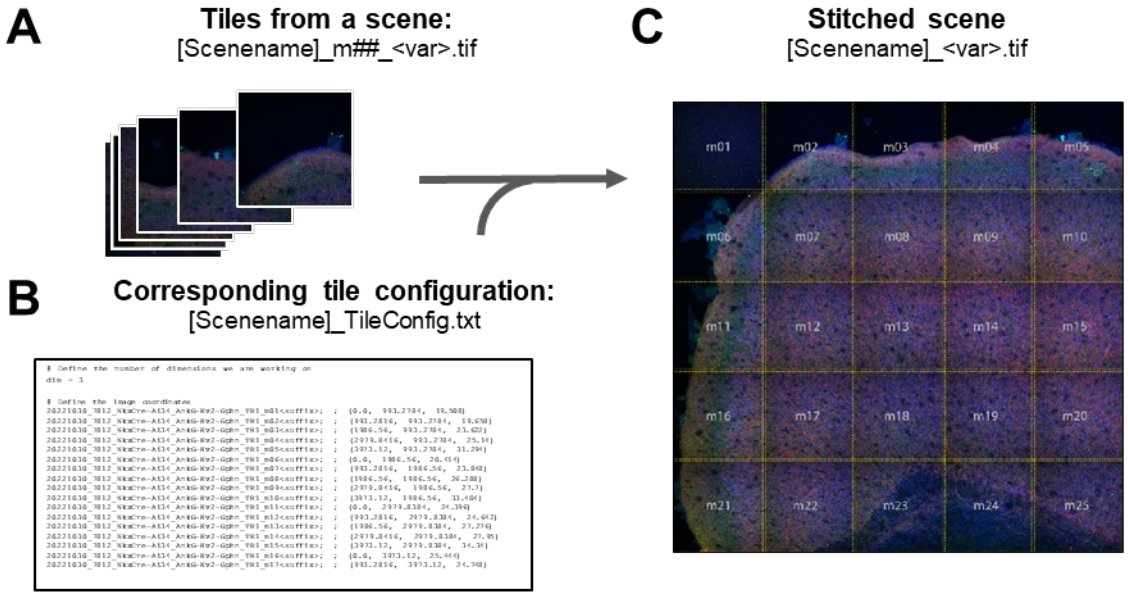

Figure 3.5 | Generalized scene stitching
The stitching algorithm identifies the scene name based on the tile nomenclature (***A***), then finds the corresponding generalized tile configuration file in the metadata directory and adjusts the tile configurations contents to match the files in the input directory (***B***). Finally, stitches the tiles in the scene that are present within the input directory (***C***).

written code that, given an input directory and tile, automatically identifies the corresponding generalized tile configuration map, creates a specific map for all images in the scene within the input directory, and stitches those images according to an extensible set of user options. The code automatically removes missing files from the output tile configuration map. It has built-in options for 3D images, such as ignoring the Z position or flattening the image via various Z projection algorithms (minimum, maximum, mean, standard deviation, and median intensity, as well as extended depth of focus). These projection algorithms utilize the GPU and operate on a tile-by-tile basis, decreasing RAM requirements and processing times. Since the acquisition-based tile positions may not actually be aligned (e.g. due to an uncalibrated microscope), we also include an option to calculate the optimal fit. This solution is permanently saved as a generalized tile configuration map in a separate location of the metadata directory, and if present, the user can opt to use this version rather than the original, metadata-based one. Since all the required input data is extracted automatically and all these steps are executed algorithmically, stitching is performed almost instantaneously with far less room for user error.

In combination, this provides an example for our structured approach to file nomenclature, metadata extraction, and file organization greatly facilitates algorithmic access to and standardized manipulation of images.

3.4. Image correction

Our pipeline relies heavily on multi-variable (big data) models and machine learning to classify pixels and objects. Humans are very good at spotting systematic differences and, when we recognize those differences are irrelevant, we can often look past them to look for more relevant differences. Computers, unless specifically trained to do so, are not able to look past such differences. An isolated machine learning model accepts the input data as is, basing its decisions on

statistical differences within the dataset and how it is trained to respond to such differences. If the data contains systematic errors that the model has not been trained to take into account, it will view those errors as relevant divergences in the dataset and make decisions accordingly. In other words, instead of recognizing relevant differences, the system will pick up on the systematic errors. Simply put: junk in, junk out.

Common sources of systematic error in bio-imaging data include intensity aberrations due microscope optics and sample-to-sample variability in staining quality ('batch effect'). We can correct for such systematic errors with additional machine learning models, but oftentimes it is easier and more robust to take a rational approach. We therefore implemented two image correction steps to correct for known systematic errors in image intensity: 1) X/Y correction and 2) sample-to-sample (or 'batch') correction.

3.4.1. X/Y correction

Optical microscopy data often exhibit a radial decrease in intensity from the center of the image, an effect known as vignetting[39, 41]. This intensity drop-off is an optical effect of the objective lens itself, varies based on aperture and focal length settings, and is present to some degree in all lenses, but especially so in fluorescence images. The intensity effects due to vignetting is easily seen in stitched images, as shown in Figure 3.6 A. We corrected for vignetting in our pipeline by adapting the previously-published BaSiC plugin[42] to run efficiently for large, multi-channel, multi-tile Z-stacks (Figure 3.6, B-E). We chose BaSiC because it takes a 'retrospective' approach to correction (i.e. it relies on the actual image data from an image series to calculate the correction), is robust against typical image artefacts, works for diverse imaging conditions, and requires relatively few input images.

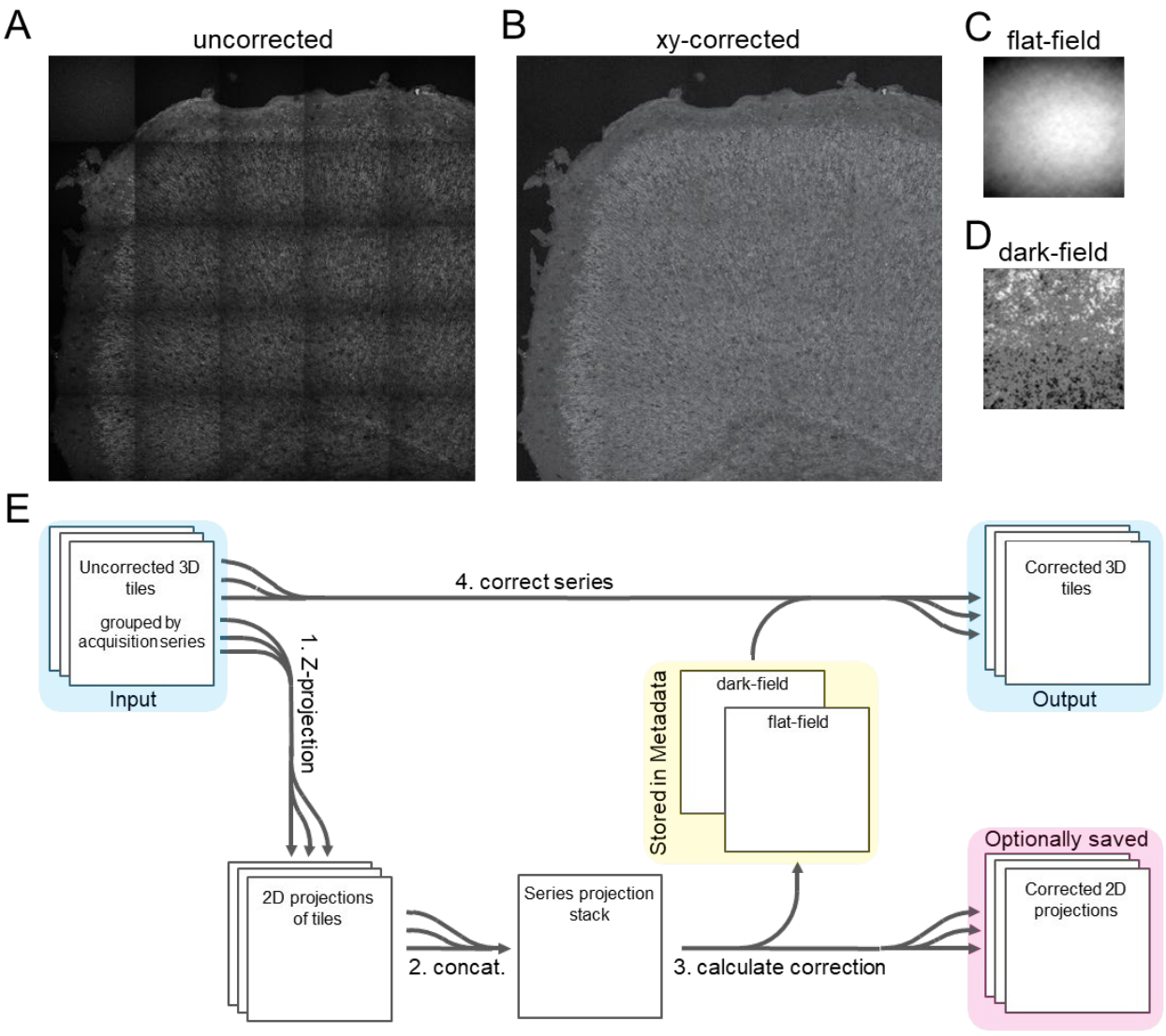

Figure 3.6 | Fast XY correction of tiled Z stacks.
A) An example of the effect of vignetting on a tiled image series. **B)** Correction of vignetting using the BaSiC algorithm. **C)** The flat-field calculation for the displayed series. **D)** The dark-field calculation for the displayed series. **E)** Schematic for fast correction of tiled Z stacks. Step 1 collapses each tiled Z stack into a 2D image that maintains the original shading aberrations (step 1). The resulting 2D EDF tile images in the group were then combined into an image stack (step 2), which is in turn used to calculate the group-wide flat-field and dark-field XY corrections in BaSiC (step 3). BaSiC outputs the flat-field and dark-field calculations as images, which we store in the metadata repository. The other direct output from the calculation stage is a stack of XY-corrected EDF images. As an option, these images can be split back into individual tiles and saved. We then used the resultant flat-field and dark-field calculations as a basis for correcting each tile as a full Z-stack (step 4).

We assumed that, since all tiles in the same acquisition series were acquired at the same time

using the same acquisition parameters, these images would also exhibit similar vignetting. We therefore decided to process these images as a unit. The original BaSiC algorithm was designed for use with multi-tile 2D images or single-tile 2D time series images (i.e. both 3D images), rather than multi-tile Z-stacks (i.e. 4D images). Given BaSiC's input structure, to process an acquisition series as a unit we must concatenate all associated tiles into a single image stack. Preliminary testing with a 4-channel , 62-tile, 40-Z-slice (i.e. 2480 images per channel) acquisition series required us to load all 12GB into memory, and BaSiC processing took over an hour per channel. For comparison, the largest similarly-stained acquisition series currently in our dataset is 242 tiles and 35 Z-slices (i.e. 8470 images per channel), would require 50 GB RAM just to load the file, and assuming linear scaling for processing time, would take ~3.4 hours per channel. Since such processing times and memory requirements are not tenable for a high-throughput system, we decided to break down this correction step to be handled more efficiently Figure 3.6, E. The BaSiC algorithm operates in two stages: the first, a 'calculation' stage, calculates the flat-field and dark-field aberrations across the image stack (step 3 in the schematic). The second, a 'correction' stage, corrects each image of the stack by removing the flat-field and dark-field aberrations (step 4). The BaSiC plugin allows users to input pre-calculated flat-field and dark-field images in order to perform the correction(Figure 3.6, C and D). We exploited this feature to increase efficiency (Figure 3.6, E). We noted that the XY aberration within any given tile will generally be consistent across the Z stack, but there is no guarantee that all images within the Z stack will be in focus. We therefore utilized a rapid extended depth of focus (EDF) algorithm implemented on the GPU via CLIJ2[15] to collapse each tiled Z stack into a 2D image that maintains the original shading aberrations (step 1). The resulting 2D EDF tile images in the group were then combined into an image stack (step 2), which was in turn used to calculate the group-wide flat-field and dark-field XY

corrections in BaSiC (step 3). BaSiC outputs the flat-field and dark-field calculations as images, which we store in the metadata repository. The other direct output from the calculation stage is a stack XY-corrected EDF images, which can be split back into individual tiles and saved if desired. We then used the resultant flat-field and dark-field calculations as a basis for correcting each tile as a full Z-stack (step 4).

Despite the increase in total steps, the efficiency gains are quite remarkable. Since the correction calculation operates on 2D images, this approach decreases total RAM requirements and calculation times by a factor equal to the size of the Z stack (in our case, 35-80 fold). Moreover, splitting the correction phase into tile-by-tile processing similarly decreases total RAM requirements by a factor equivalent to the number of tiles in the acquisition series (~240-fold for our largest). Importantly, this decrease in RAM requirements enables parallel processing of multiple groups, further increasing efficiency. Using the same system, calculating the correction takes less than a minute per channel, whereas applying the correction to each tile is accomplished in a matter of seconds. Since the flat-field and dark-field calculations are permanently stored in the metadata directory, we avoid the need to recalculate should we have to rerun the same files. Moreover, since we store the detectors and objective used for acquisition within metadata, we can calculate 'universal' flat-field and dark-field profiles for those conditions and save it as a specific profile for prospective correction, e.g. of small acquisition series that are underpowered for their own retrospective calculation, or very large acquisition series (e.g. thousands of tiles) where the calculation stage would require considerable time. This again highlights the strengths of this ecosystem at different experimental scales.

3.4.2. Batch correction

A major source of systematic error is the sample-to-samples variability in staining intensity and quality between different biological samples and independent imaging experiments. We therefore included a stage to correct for this source of variance. When considering this stage of the pipeline, we had several requirements that needed to be met:

1. Works for 2D and 3D multi-channel images.
2. Includes a masking step, so that only relevant pixels/voxels are included in the correction calculations.
3. Correction is series-by-series rather than tile-by-tile. That is, the series is the unit for correction; every tile within the series is adjusted using the same coefficients/correction.
4. The series data is calculated from individual tiles (not a concatenated stack).
5. The correction standard initially is *set retrospectively* (i.e. from the actual data), but for all subsequent runs is *applied prospectively*. That is, we did not want the basis for correction to change from one run to the next.
6. Different correction standards can be set for different conditions. For example, slice and culture samples needed to have distinct correction standards for the same markers.

We tested several off-the-shelf solutions but were unable to find one that satisfied all our needs. We therefore decided to build our own solution. The correction step utilizes the grouping algorithm above to group tiles based on one or more user-designated categories.

Figure 3.7, A shows a schematic of the batch correction algorithm. In the first phase, the algorithm cycles through every tile in the group and (1a) measures channel intensity statistics (min,

mean, max, SD, variance, automatic threshold values), (1b) extracts histograms, and (1c) calculates the non-zero quartile values from the histogram. If the user supplies a mask, then it will

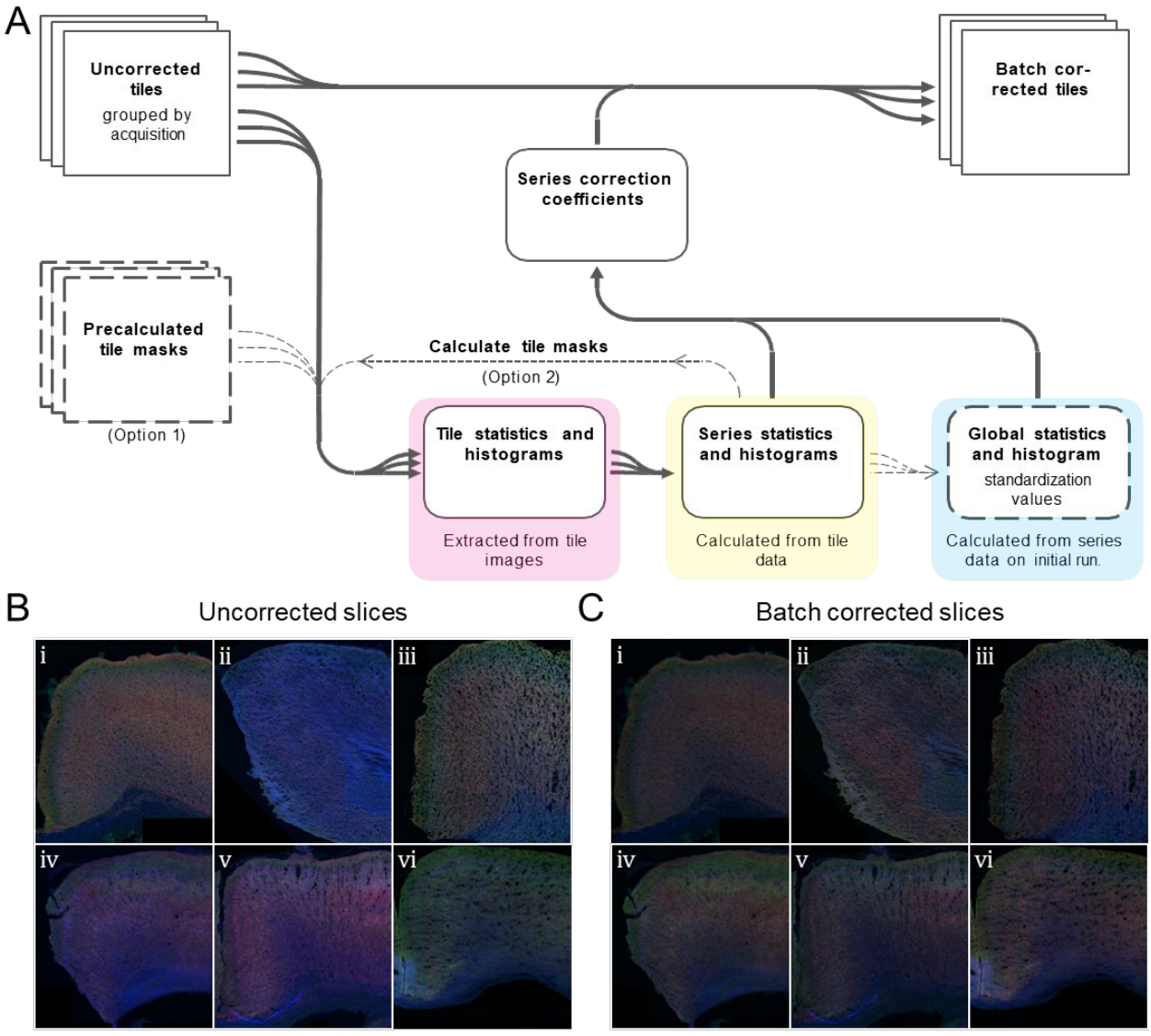

Figure 3.7 | Sample-to-sample (batch) intensity correction.
A) A schematic of the batch correction algorithm. The first stage (*magenta*) measures channel intensity statistics across the whole tile. These data are permanently stored in the metadata directory (i.e. they are measured only once). Measurements are from both masked and unmasked tiles. The mask can be supplied by the user (option 1) or derived from a series-wide threshold value calculated in the second stage (option 2). The second stage (*yellow*) calculates series-wide metrics from the associated tile measures in the metadata directory. These values serve as the basis for series-wide batch correction. If a mask is not supplied (option 2), a series threshold value is calculated from the series histogram, then used to mask each tile. This option requires re-opening each tile to extract a mask and recalculate masked values. Stage three (*cyan*) calculates the global

values from all available series values, which in turn serve as a basis for standardization of intensity of all subsequent runs. These values are calculated by default on the initial run and stored permanently in the metadata directory. Global values can be recalculated from all available series data later if desired, in which case any current global values are archived.
B-C) Uncorrected (***B***) and batch-corrected (***C***) slices. Slices were batch corrected using the 75th non-zero quantile matching method. Data are presented without display adjustments.

measure, extract, and calculate the same data for the masked pixels/voxels only. The second phase utilizes the tile metrics and tile histograms to calculate (2a) channel intensity statistics and automatic threshold values (2b) a cumulative histogram, and (2c) quartile values for the entire group. If the user has not provided a mask, they are prompted at the outset to designate an autothreshold method for masking. The algorithm will then cycle through the tiles a second time, this time masking the images based on the designated threshold value calculated for the group, then measures, extracts, and calculates the masked statistics and masked histograms for each tile, followed by calculations of these values for the group. The third phase sets the global standards to which each group is adjusted. It does this essentially by repeating the second phase, but this time uses the group-wide data as inputs to calculate (3a) global statistics, (3b) histograms, and (3c) quartiles. This global calculation is done only on the first run or upon user override. Finally, in the fourth phase, the algorithm (4a) calculates the correction coefficients for each channel from the group values and global standards, then (4b) adjusts each tile based on a user-designated correction method. A set of uncorrected and corrected slices is shown in Figures 7 B and C, respectively.

We included several features to make this batch correction efficient, robust, and malleable:

First, the tile, group, and global metrics and histograms, as well as the correction calculations (i.e., the outputs from phases 1-4a) are saved to a designated location in the metadata directory, with measurement values saved in list form using pre-defined keys, and histograms saved as 1D

images using algorithmically-derived filenames. Masked and unmasked values are saved using distinct keys, so these data do not overwrite each other. Permanently storing these data not only ensures consistency from run to run, but also provides a major efficiency gain: the measurement and extraction step (i.e. phase 1) is by far the most time-consuming part of phases 1-4a because it requires reading tiles from disk. Since the tile data are permanently stored in the metadata folder, tiles with existing values do not have to be re-measured on subsequent runs. This means that processing times for this stage do not increase as more images are added to the pipeline, an important feature when datasets can have thousands or even tens of thousands of tiles.

Second, we utilized the hierarchical structure of our sample nomenclature to align thousands of files from numerous inputs rapidly and efficiently. During this stage, multiple types of input data must be aligned for each user-designated input tile. Using a brute-force, iterative search approach, the total number of potential file combinations scales exponentially with a linear increase in the number of inputs. We used the hierarchical experiment-sample-scene-tile nomenclature structure to perform this alignment recursively from experiment down to tile, with higher order mismatches skipping lower order searches. This alignment approach linearizes the relationship between inputs and potential combinations, and led to a ~50-fold reduction in processing times at this stage compared to a brute-force approach.

Third, since the group calculations are based on tile measures that are readily available in the metadata directory, our approach avoids high RAM requirements and long read, write, and calculation times for large groups. Instead, calculating the group and global values takes a milliseconds. Since the resulting values are stored in the metadata folder using distinct key identifiers, they are readily accessible for any downstream runs.

Fourth, tiles can be grouped in many ways. Since these values are stored in the metadata directory, recalculating group values for a different grouping method does not require re-measuring tile values. Instead, the new grouping values are simply recalculated from re-ordered existing tile data. Group data for each grouping method are stored using unique keys, meaning that selecting a new grouping method does not overwrite data from other grouping methods. The current study utilized acquisition series (i.e. AcquisitionDate timestamp) as the grouping method because it ensured that all images within the set 1) share the same acquisition parameters, 2) are from the same slide/plate, independent of changes in scene or even sample.

Fifth, the global standards are calculated and set by default based on the input data from the first run (i.e. *retrospectively*), but not recalculated on subsequent runs. For subsequent runs, the correction coefficients are calculated based on the pre-existing global standard (i.e. *prospectively*), meaning that images from all runs utilize the same correction standard, and are thus comparable. Since data from the first run may not be representative, we also built in the option to re-calculated and re-set the global standards from all available data later if desired. Upon re-calculation, the pre-existing global standards are automatically archived, thereby giving users the option to return to any prior standards if desired.

Sixth, we built in the ability to have different global standards for different experimental conditions. For instance, for the current study we designated different standards for different sample formats (i.e., slice vs culture). Similarly, one could designate different global standards based on any change in experiment format, such as different staining conditions or biological conditions. Importantly, each input is handled automatically based on its corresponding groupings from the auto-generated *groupings table* stored in the metadata directory. This automatic file handling and

decision-making step ensures that inputs are processed as expected, with little need for human input and little room for human error.

Seventh, we included several correction methods, including non-zero quantile matching for the 50th and 75th quantiles (Q50NZ and Q75NZ, respectively) with and without log transform. The current study used the linear Q75NZ method for correction, a standard for slide-to-slide correction of bioimaging data from multiplexed tissue arrays[43]. The code used for any given correction method is modular, so adding other correction methods in the future is relatively simple, making this step of the pipeline malleable and extensible.

In summary, this stage makes an initial pass at correcting batch-to-batch variation. The correction methods built into this step are admittedly rudimentary. That said, the structured storage of hierarchically-organized metrics that, ultimately, reflect experimental desing creates data ecosystem that could be utilized for more complex batch correction methods. Thus, we have created a framework that is extensible and malleable. For example, image acquisition parameters, which we extract in the first step of the pipeline and store in metadata folder, could be utilized to guide batch correction. Similarly, AI-based methods could be substituted for our naïve arithmetic transformations. Additionally, other image correction steps, such as denoising or deconvolution could be incorporated at this stage. Separating signal from noise is critical. Here, we provide a framework to do that.

3.5. Pixel Classification

Biologists often use antibody staining as means of marking biological substructures, but it is important to keep in mind that the resultant staining pattern is not *actually* the substructures. Rarely does the staining pattern line up perfectly with the biological substructure or cell type for which it is a marker. Some antibodies bind off-target proteins, others don't bind the target protein in certain

biological contexts (e.g. due to post-translational modifications), and sometimes the target protein isn't throughout the entirety of the biological substructure for which it is a marker. Biologists understand these exceptions and use context clues and knowledge of the biological system to interpret the staining data in a way that makes sense given the experimental question and biological context.

An example may help to illustrate this point in more concrete terms. Our experimental question investigates GABAergic interneuron subcellular targeting. To investigate this question experimentally, we used antibodies to stain proteins that canonically localize to the subcellular compartments of interest as markers of these subcellular compartments: Ankyrin G for the axon initial segment (AIS), Kv2.1 + Kv2.2 (Kv2) for the soma and proximal dendrite (soma), and Gephyrin for GABAergic post-synaptic densities (PSDs) generally. Figure 3.8, A-D shows the staining patterns for each of these antibodies at low resolution. Figure 3.8, E shows synaptophysin-tdTomato (syp-tdT) labeling of presynaptic boutons from an AIS-targeting GABAergic interneuron labeled by Nkx2.1^{CreER} driver; our indicator of GABAergic synapses at the AIS. Although we are using these antibodies as markers for the subcellular compartment, these antibodies are not staining the subcellular compartments *per se*. Each of these antibodies presents subtle but important divergences from the 'marker' paradigm. Their staining patterns at the AIS illustrate this point well Figure 3.8, A'-E'. Ankyrin G is a canonical marker of the AIS. Although the Ankyrin G antibody we used displayed significant enrichment at the AIS, it also stained neurites in general at a lower intensity Figure 3.8. A, B, meaning that the Ankyrin G staining is not exclusive to the AIS. Moreover, it is not localized throughout the entire AIS: it is specifically excluded from regions with high Kv2 clustering Figure 3.8, A'-C', which has previously been documented[15]. We are using

Kv2.1 and Kv2.2 as markers for neuronal soma and proximal dendrite Figure 3.8, C', but as mentioned already Figure 3.8, C', E'' and described previously, both proteins also localize to the AIS[15, 44], where they cluster to form potassium channels adjacent to GABAergic PSDs[15]. Of note, these Kv2-enriched PSD subdomains of the AIS do not contain Ankyrin G. Gephyrin is a canonical marker of GABAergic PSDs, but not all Gephyrin antibodies recognize all GABAergic

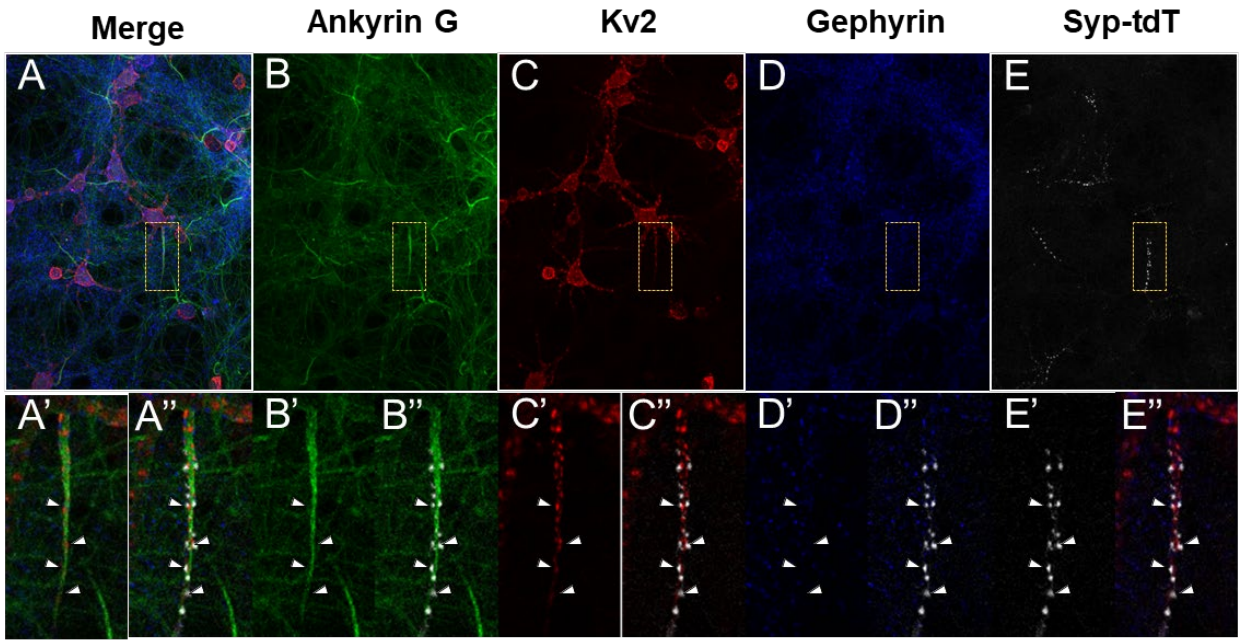

Figure 3.8| Staining pattern of subcellular markers generally and at the AIS.
The top row shows the general staining pattern for the markers, and the bottom row shows the staining pattern at the axon initial segment (AIS). The displayed AIS is boxed. In the second row, the marker's staining pattern is indicated by X', and its relation to presynaptic boutons from an AIS targeting GABAergic interneuron is indicated in X''.
A) Merged image of Ankyrin G, Kv2, and Gephyrin staining. **B)** Ankyrin G signal at the AIS. Ankyrin G is a canonical marker of the AIS. Note that this antibody also stains neurites in general at a lower intensity, and is excluded from Kv2-rich regions of the AIS. **C)** Kv2.1 and Kv2.2 combined (Kv2) signal. Kv2 is a marker of neuronal soma and proximal dendrites. Note that it also localizes to the AIS, where it is found in apposition to AIS-targeting GABAergic synapses. **D)** Gephyrin staining. Gephyrin clustering is a canonical marker of GABAergic post-synaptic densities (PSDs). This gephyrin antibody (mAb7a) recognizes gephyrin that is phosphorylated at S270. Gephyrin localized to PSDs at the AIS is not phosphorylated at S270, as indicated by the lack of signal in apposition to the Syp-tdT signal. **E)** Syp-tdT signal from an Nkx2.1^{CreER};Ai34 AIS-targeting GABAergic interneuron. The Syp-tdT signal here serves as a positive indicator of GABAergic synapses at the AIS.

PSDs. Specifically, the gephyrin antibody we used in this study (mAb7a, the industry standard) recognizes a phosphorylated form of gephyrin (S270), which clusters and forms a scaffold for $GABA_A$ receptors. Figure 3.8, D, D', D'', E'' illustrates this staining pattern within our system. Importantly, gephyrin PSDs at the AIS are *not* phosphorylated at S270, meaning that this Gephyrin antibody does not recognize GABAergic PSDs at the AIS[45], which we witness: few of the Syp-tdT-labeled boutons at the AIS are in apposition to gephyrin staining. Of note for our system in particular, the regions of Kv2 channel clustering at the AIS are the same regions where Ankyrin G is excluded and Gephyrin lacks S270 phosphorylation[46] Figure 3.8, C, and is adjacent to the cisternal organelle complex, an AIS-specific endoplasmic reticulum specialization where ChC (but not non-ChC inhibitory) boutons typically cluster[47]. As biologists, we use this knowledge of the biological context to reinterpret the raw staining patterns: the regions where Kv2 is present and AnkG is absent are not only part of the AIS, but (despite the absence of Gephyrin staining) are the part of the AIS where GABAergic PSDs associated with ChC synapses are enriched. Thus, there is a level of processing that occurs between viewing the raw staining data and inferring biological meaning from those data. Biologists don't just view the staining patterns; they also interpret them.

We wanted to train a computer to interpret the raw staining data as a biologist does. We therefore took a supervised learning approach to train machine learning models to classify pixels of our confocal images based on biologically meaningful criteria. We trained two pixel classifiers, one for presynaptic data (Figure 3.9) and another for target data (Figure 3.10). We chose to train two classifiers because each classifier performs a mutually exclusive classification, i.e., an increased probability to be categorized as one class requires decreasing probability it is categorized some other class. We are, however, interested in synaptic targeting, and thus do not want the presynaptic

and target signals to be in any way dependent on each other, let alone mutually exclusive. In contrast, we do want the possible classes within the presynaptic classifier to be mutually exclusive, and similarly for the classes within the target classifier. Simply put: using a single classifier would lead to competition between presynaptic and target classes, whereas we are interested in their intersection. In all cases, we chose to use probability maps as the outputs from the pixel classifiers. These probability maps are multichannel images, with a total number of channels corresponding to the total number of classes for the classifier. The intensity values for each channel are on a 0-100 scale and reflect the percent probability that the pixel falls into that class. Given that the classes are mutually exclusive, the sum of all channels for any given pixel/voxel is 100.

3.5.1. Presynaptic pixel classification

The presynaptic classifier is shown in Figure 3.9. This classifier was trained exclusively using the synaptophysin::tdTomato (Syp-tdT) reporter channel. We chose four classes for the presynaptic classifier: 1) background (i.e. unlabeled), 2) labeled soma, 3) labeled neurite, and 4) presynaptic bouton. Although our study is primarily concerned with synaptic targeting (and so also the presynaptic bouton class), we chose to include the labeled soma and labeled neurite classes because together these three classes encompass the different subcellular compartments within which we observed real (i.e. non-background) labeling from cells expressing the Ai34/Syp-tdT reporter. That is, these are the categorical distinctions a biologist would make when viewing the signal in this channel, and we wanted to train our classifier to make the same distinctions. In addition, including

these classes opens up further avenues for downstream investigation, such as comparisons of labeled cell numbers or morphological analysis of labeled cells.

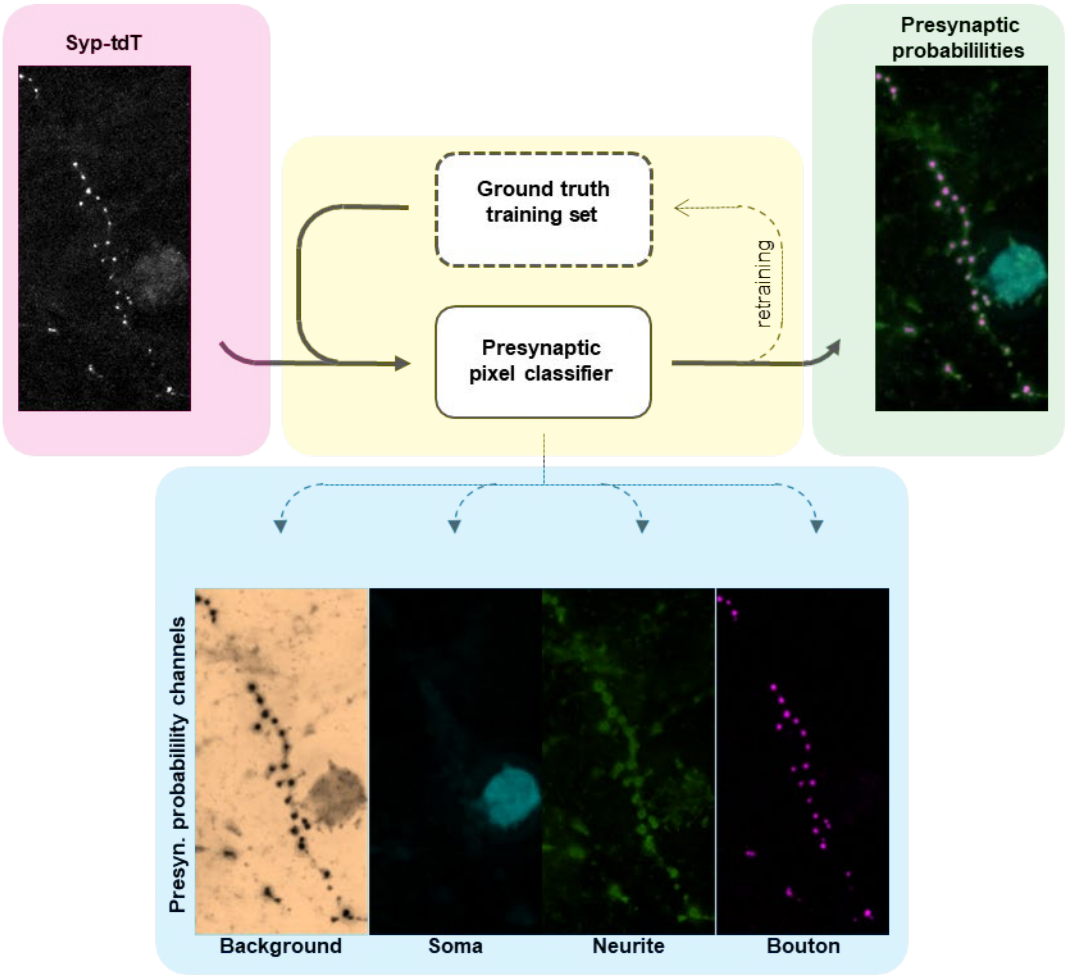

Figure 3.9 | Presynaptic pixel classifier.
Only Syp-tdT (magenta), the presynaptic marker channel, is used to train the presynaptic pixel classifier (*yellow*). Probability maps of the four classes modeled by the presynaptic classifier are shown separately (*cyan*): 1) background (i.e. unlabeled), 2) labeled soma, 3) labeled neurite, and 4) presynaptic bouton. A merged image showing the cumulative probabilities for all positive signals is shown (*green*). Images are from a sparsely-labeled Sst^{Cre} primary cortical culture.

3.5.2. Target pixel classification

The target classifier is shown in Figure 3.10. This classifier was trained exclusively using the Ankyrin G, Kv2, and Gephyrin channels. We chose seven classes for the target classifier: 1) background, 2) AIS, 3) soma & proximal dendrite, 4) gephyrin PSDs, 5) non-AIS Ankyrin G (i.e. target

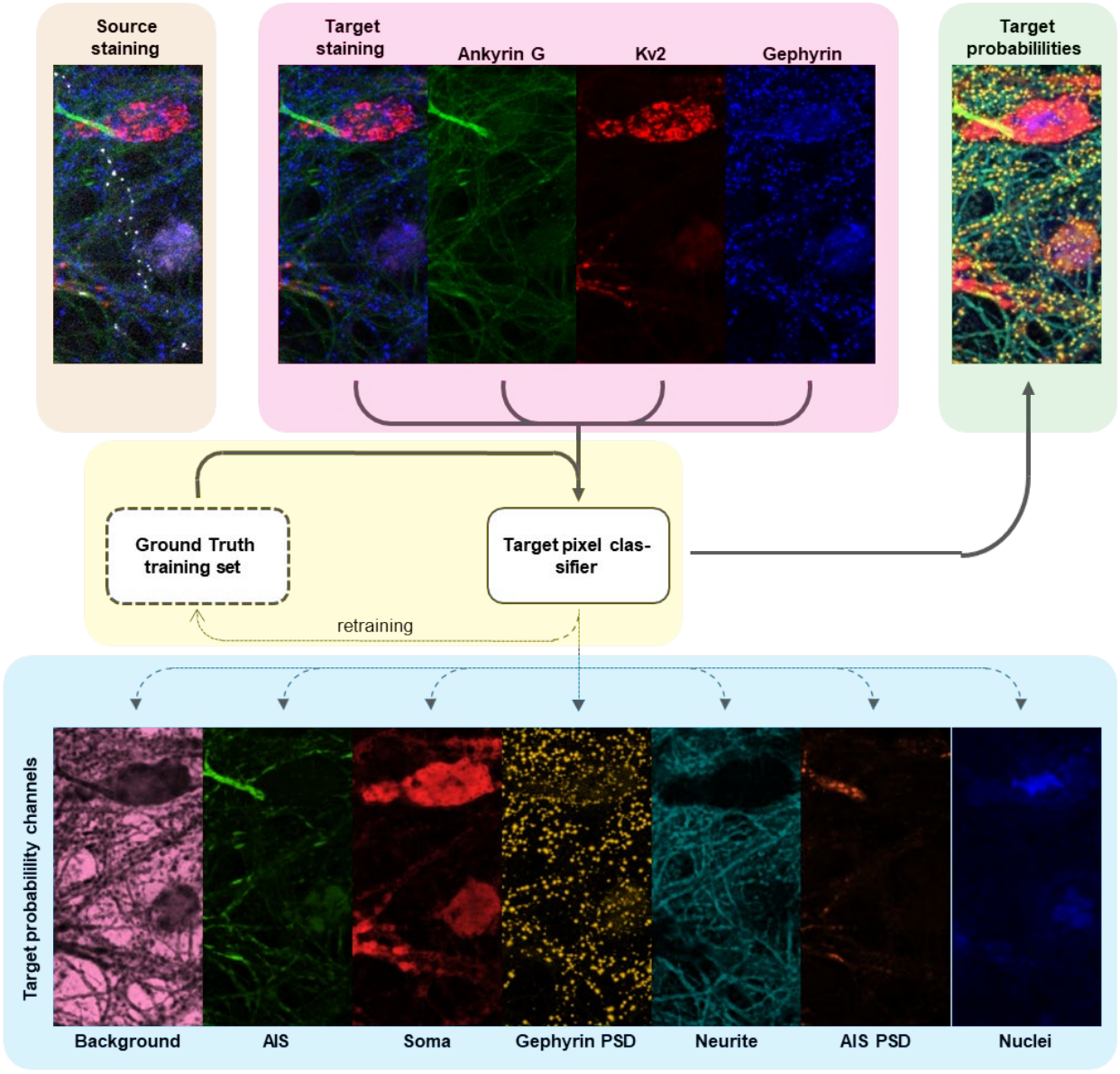

Figure 3.10 | Target pixel classifier.
The target classifier was trained exclusively using the Ankyrin G, Kv2, and Gephyrin channels (*magenta*) of the source image (*orange*). Probability maps for the seven classes modeled by this classifier are shown (*cyan*): 1) background, 2) AIS, 3) soma & proximal dendrite, 4) gephyrin PSDs, 5) non-AIS Ankyrin G (i.e. target neurite), 6) KV2 at the AIS (i.e. AIS PSDs), and 7) nuclei. An overlay of all probability maps less background is shown (*green*). The ROI shown here is the same as in Figure 9.

neurite), 6) KV2 at the AIS (i.e. AIS PSDs), and 7) nuclei (Figure 3.10). Again, although our study is primarily concerned with targeting at the AIS, soma/proximal dendrite, and distal dendrite, we decided to include the other (non-background) classes because they are a more accurate reflection

of the subcellular compartments identifiable from this combinatorial staining. As discussed above, we chose to distinguish AIS PSDs from the AIS in general because a) this classification more accurately represents our biological knowledge of the structure itself, and b) it provided us with a surrogate PSD at the AIS, which given the Gephyrin antibody specificity otherwise would be missing. We included the nuclei class because this compartment was readily identifiable from the Kv2 signal, and its inclusion would enable us to quantify total cell numbers.

3.6. Object classification

The next stage of our pipeline moves from analyzing images as a disconnected series of pixels/voxels to identification, measurement, and classification of discrete objects. As outlined in Figure 3.11, this stage consists of four steps: 1) designating/constructing the image base data from which object measurements are extracted (the object measurand image), 2) synthesizing pixels/voxels into discrete, extended objects (the segmented objects), 3) extracting the measurements for each object from the measurand image, and 4) classifying the objects based on these measurements and the classification model. We took two approaches to object classification, a) a supervised learning approach and b) an unsupervised (hypothesis-blind) approach. In our current pipeline, object metrics are extracted as part of the supervised learning step and saved in tabular form as part of the outputs from this stage. However, object measurement could be a separate step. We use these same metrics as inputs for unsupervised classification, meaning that both classification steps are ultimately based on the same data. We discuss these steps here.

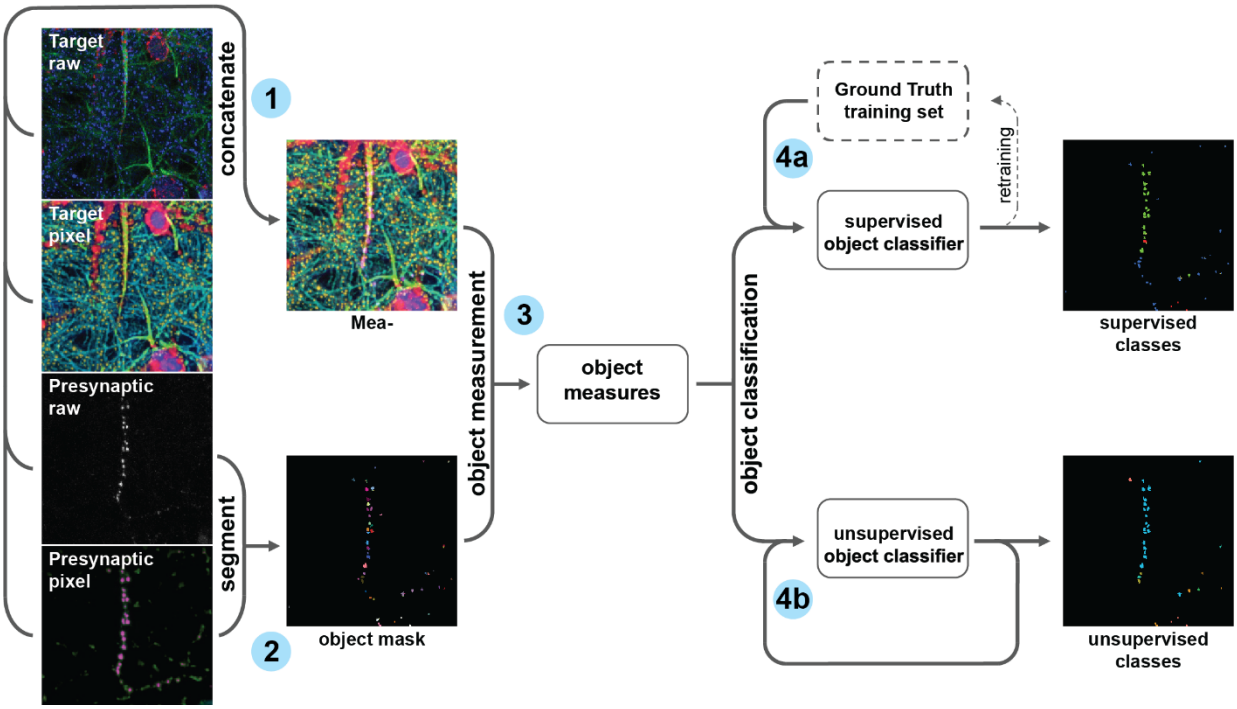

Figure 3.11 | Object classification schematic.
The object classification moves from analyzing images as a disconnected series of pixels/voxels to identification, measurement, and classification of discrete objects. This stage consists of four steps: **1)** Measurand generation. This image is the the base data from which object measurements are extracted. It is concatened from the pixel classifier inputs and outputs. Here the measurand is a 15-channel image. **2)** Object segmentation. This step synthesizes pixels/voxels into discrete, extended objects, then assigns each object a unique identifier. **3)** Object measurement. In our pipeline, a total of 431 unique metrics are extracted for each object and its local neighborhood. These include morphometrics, intensity-based photometrics, and cross-channel intensity covariants. **4)** Object classification. There are two approaches to object classification, 4a) a supervised learning approach and 4b) an unsupervised (hypothesis-blind) approach. Both approaches use the same object measurements as input data, and are therefore complimentary modes of analysis. Images are from a ChC-enriched primary cortical culture where Nkx.21^{CreER} labeled the cells. The ROI shown here is the same as in Figure 8.

3.6.1. Object Measurand Concatenation

This step designates the measurand for object-based measurements, i.e., the image from which object measurements are extracted. This step essentially takes any number of distinct input images for a given tile and combines (i.e., concatenates) channels into a multi-channel image stack with a

pre-designated channel order. One could, presumably, use any number of input channels at this stage, with the only downsides being increased calculation times and (potentially) decreased interpretability. We wanted the inputs for our object classification stage to include not only the outputs from both pixel classifiers (i.e., the biological interpretations), but also the images used as inputs for the pixel classifiers (the input channels). We therefore concatenate the pixel classifier input channels (i.e. Syp-tdT, Ankyrin G, Kv2, and Gephyrin) and the and target pixel classifier output channels into a single 15-channel stack Figure 3.11, #1. The pixel classifier images are in the form of probability maps. This 15-channel image serves as the first input for the supervised object classifier, i.e., the measurand for presynaptic bouton measurements.

3.6.2. Object segmentation

This stage utilizes the pixel classifier probability maps and the corresponding input channel(s) to derive objects via segmentation. We use the probability maps to designate a probability cutoff to remove low probability pixels from the input channel, which is then fed into a segmentation algorithm to extract objects of a designated class. This stage is malleable and can be used to extract objects of any class from the input images, so long as there is a corresponding probability map. We extracted presynaptic bouton objects using the Syp-tdT input channel and presynaptic bouton probability channel Figure 3.11, #2. Moreover, the specific segmentation function utilized is modular, and therefore can be modified to accommodate objects of different shapes. We utilized the 3D spot segmentation algorithm from the 3D ImageJ Suite (mcib3d plugin[48]), which uses input seeds (calculated via 3D local maxima) as a starting point for object identification, then designates pixels as components of that object via a process of 3D expansion. We tested other 3D segmentation algorithms within the 3D Suite (e.g. simple segmentation, iterative thresholding, and hysteresis thresholding), and found that the Spot segmentation algorithm did the best job segmenting and

splitting the (largely spherical) presynaptic boutons, was relatively simple, consistent, and easily automated. We save the segmentation output as a binary image, which in turn serves as the second input for the supervised object classifier (i.e., it designates the 3D extended objects themselves).

3.6.3. Object measurement

The next step uses the object segmentation mask and measurand image to extract measurements for each object Figure 3.11, #3. Although the object measurement step is built into our supervised classification stage, these steps are functionally and conceptually distinct. It is therefore more fitting to discuss them separately. Object measurement begins by transforming the binary segmentation image into an object label map via a connected components labeling algorithm. Simply put, this algorithm considers adjacent, non-zero pixels/voxels to be part of the same object. Every resultant object is then designated a unique, non-zero identifier value (its index, label, or ID), with background/non-object designated as zero. (Practically speaking, each set of connected pixels/voxels is given a unique intensity). These objects are then used as the basis for measuring summary statistics across the object. The measurements include location-based metrics, morphometrics, and photometrics. Location-based metrics include information about the object's relative coordinates within the image. We did not include any location-based metrics as inputs for object classification. Morphometrics include size- and shape-based measurements, such as volume, sphericity, and min/mean/max radii. Photometrics are intensity-based measurements such as min/mean/max intensity for each channel, as well as cross-channel metrics like intensity covariance. These intensity-based statistics can be measured within the object itself or from the object's neighborhood (i.e., pixels/voxels within a given radius, excluding the object itself). In total, we extracted 431 non-redundant metrics per bouton across the 15 channels. These metrics serve as

inputs for both the supervised Figure 3.11, #4a and unsupervised Figure 3.11, #4b object classification classifiers.

3.6.4. Supervised object classification

The supervised object classification step fits the objects to a pre-defined classification model based on the object's metrics. As discussed, it uses the measurand image and object mask as inputs Figure 3.12, A. We chose five object classes Figure 3.12, B: 1) Non-bouton, 2) Dendrite-targeting, 3) Soma-targeting, 4) AIS-targeting, and 5) Unknown-targeting. We then established classificatioin criteria for each object class. 1) The Non-bouton class encompasses any objects that are not boutons, i.e., objects that either a) do not fit the expected shape and size (e.g. single voxel, small, non-spherical/flat, or very large objects), or b) do not have significant overlap with contiguous high Presynaptic Bouton probability. Non-boutons are generally technical artefacts from the segmentation stage, and we excluded these objects from downstream analysis. 2) The Dendrite-targeting class is any bouton with clear overlap of high probability Gephyrin PSD, but does not overlap with Soma, AIS, AIS PSD. 3) The Soma-targeting class is any bouton that overlaps with Soma but does not overlap with AIS or AIS PSD. 4) the AIS-targeting class is any bouton that overlaps with AIS or AIS-PSD. Finally, 5) the Unknown-targeting class is any object that fits the bouton criteria but does not fit any of the other three bouton classes, i.e., it does not overlap with Gephyrin PSD, Soma, AIS, or AIS PSD. The Unknown-targeting class comprised approximately 0.99% of all objects and 1.4% of all boutons. The model itself was trained by expert observers using a subset of the input data (Figure 3.10, C). Object classification is then assigned based on maximum probability across all classes (Figure 3.10, D). We excluded any Non-bouton classified objects for downstream analysis, yielding bouton targeting.

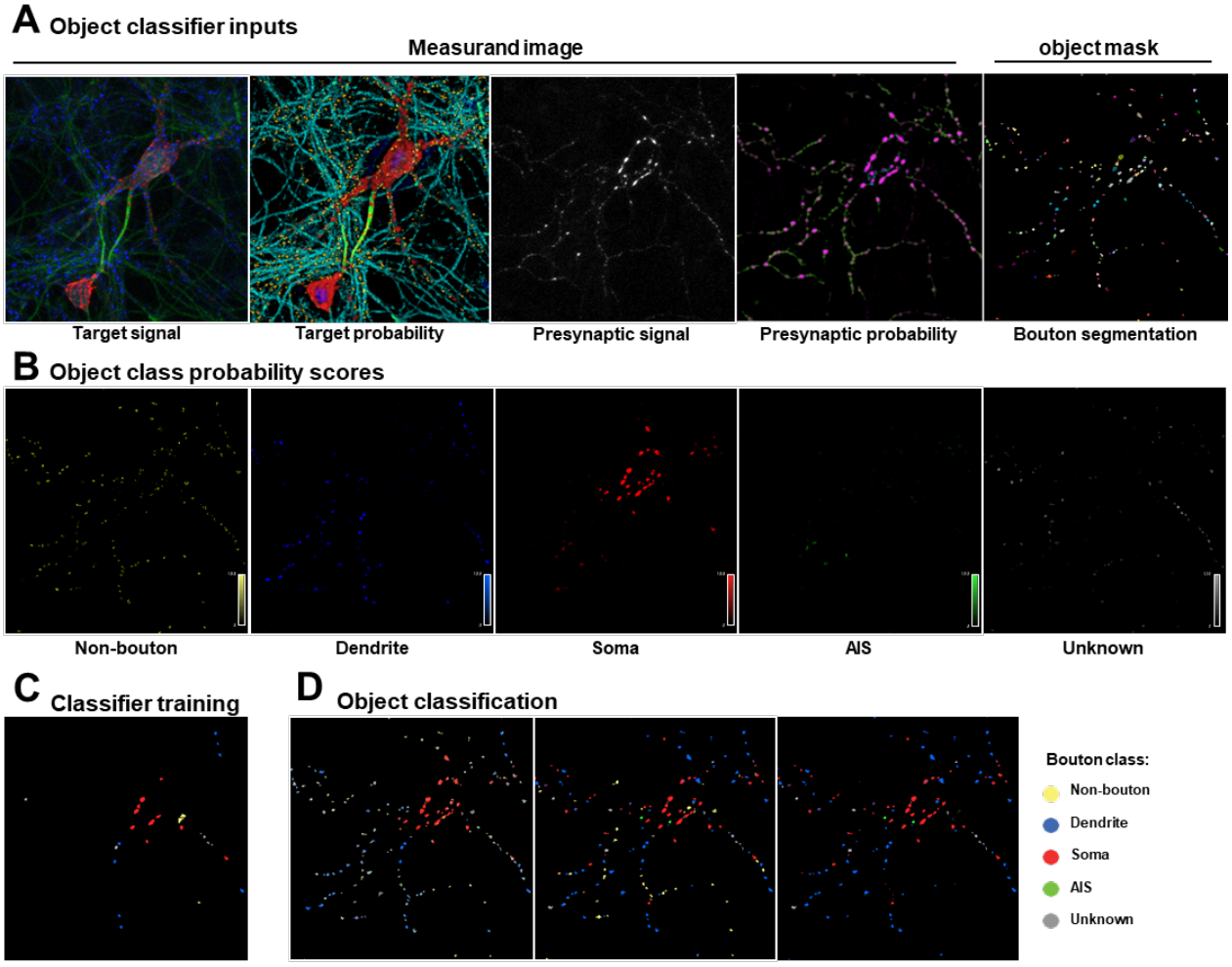

Figure 3.12 | Object classification inputs and outputs.
A) The object classification inputs consist of the measurement basis image and an object mask. Our basis image was concatenated from the raw target signal and target pixel classifier probability maps, as well as the presynaptic signal and presynaptic probability maps. The object mask results from segmentation of the raw presynaptic signal using the bouton probability score for thresholding, followed by a 3D spot segmentation algorithm that uses local maxima as seeds. **B)** Object probability scores for the five object classes: Non-bouton and Dendrite-, Soma-, AIS-, and Unknown-targeting. Class definitions and classification criteria are provided in the text. **C)** Ground truth input for the current tile. The probability scores are based on minimal user training. **D)** Object classification is assigned based on maximum probability across all classes. Excluding the Non-bouton class yields the bouton targeting. Images are from a BC-enriched primary cortical culture where Nkx.21^{CreER} labeled the cells.

3.6.5. Supervised classifier validation

We validated our model using a ground-truth validation set of ~5400 independently classified objects that had been excluded from training. The validation set consisted of 30 tiles from 10 independent experiments, with 20-300 randomly-selected objects per tile. It included boutons from both slice and culture and from all driver lines. Figure 3.13, A shows an example tile subregion. We compared the model predictions Figure 3.13, B to the validation set Figure 3.13, C, noting mismatches between ground truth and prediction for each class Figure 3.13, D. We then generated a confusion matrix to calculate true positive, true negative, false positive, and false negative scores to assess model performance. Overall model performance was high, with an accuracy[20] of 94.2% across all objects and 95.3% across all boutons, with similar accuracy for each bouton class individually (Figure 3.13, E). Specificity[21] and negative predictive value[22] had similarly high scores across all categories. Precision[23] and sensitivity[24] were similarly high for all objects and all boutons, though we did witness decreased performance for both metrics among specific object classes. To assess how our model's predictive performance deviated from ground truth for specific classes, we generated confusion matrices to assess precision and false discovery rate (FDR)[25] (Figure 3.13,

[20] Informally, *accuracy* is the fraction of predictions that the model got correct. Formally, $accuracy = (TP + TN)/(TP + TN + FP + FN)$, where TP = true positives, TN = true negatives, FP = false positives, and FN = false negatives.

[21] *specificity* is also known as the *true negative rate* (TNR). Informally, it is the probability that an actual negative is predicted to be negative. Formally, $specificity = TN/(TN + FP)$.

[22] *Negative predictive value* (NPV) or negative predictive power informally tells you how often a model is correct when it makes a negative prediction, i.e., how often a negative result represents a true negative. Formally, $NPV = TN/(TN + FN)$.

[23] *precision* is also known as the *positive predictive value* (PPV). Informally, precision measures how often a model is correct when it makes a positive prediction, i.e., the proportion of positive predictions out all correct predictions. Formally, $precision = TP/(TP + TN)$.

[24] *sensitivity* is also referred to as *recall* or *true positive rate* (TPR). Informally, sensitivity describes how often the model predicts the class when it actually *is* that class. In other words, it is a measure of how well a model can identify true positives, or the proportion of actual positives that it identified correctly. Formally, $sensitivity = TP/(TP + FP)$.

[25] False discovery rate (FDR) informally is how often a model is incorrect when it makes a positive prediction. Formally, $FDR = FP/(TP + FN) = 1 - precision$.

F) and sensitivity and false negative rate (FNR)[26] (Figure 3.13, G). The model performed worst

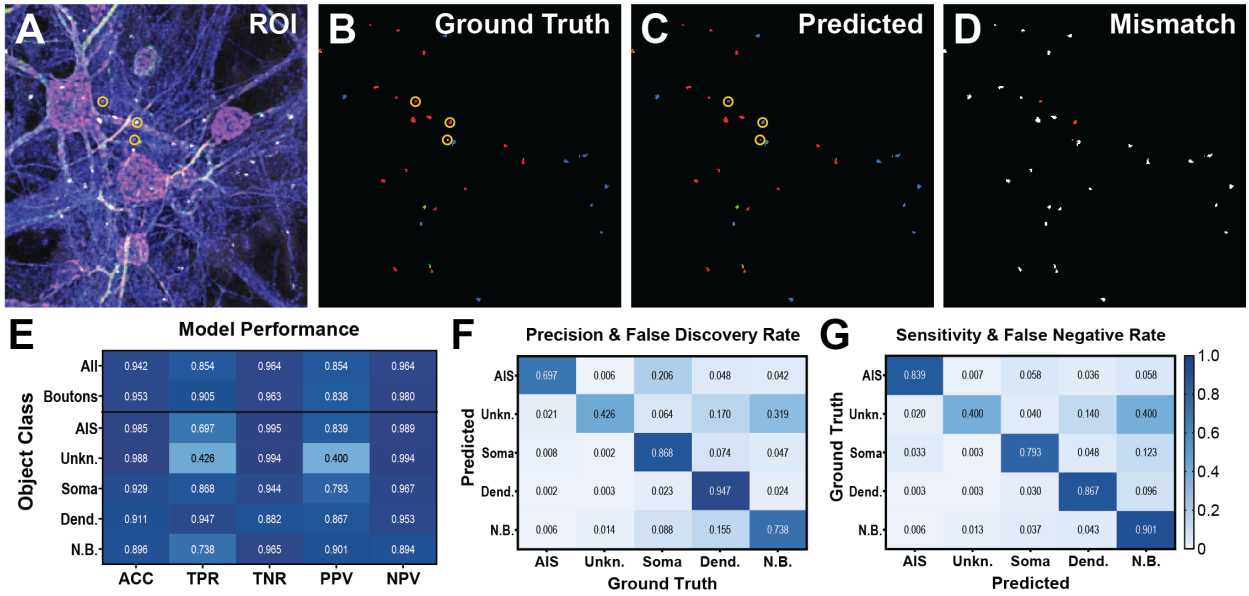

Figure 3.13 | Supervised object classifier performance.

A) An example region of interest (ROI) used in the validation sample. The validation set consisted of 4742 randomly-selected objects from 30 tiles and 10 independent experiments. The validation set was classified without model feedback. The objects circled in orange are mismatches between ground truth and prediction. **B)** Ground truth classification for the ROI. Mismatches are circled as in ***A***. **C)** Model prediction for the ROI. Mismatches are circled as in ***A***. **D)** Matches are shown in white, mismatches in red. **E)** Model performance metrics across all classes, all bouton classes, and each object class individually. ACC: accuracy; TPR: sensitivity, recall, or true positive rate; TNR: true negative rate; PPV: precision or positive predictive value; NPV: negative predictive value. Please refer to the text for formal and informal definitions of these performance indicators. **F)** Precision and false discovery rate of the object classifier. Shown is the confusion matrix where each value has been normalized by the total predicted counts for that class (TP+FN). The diagonal shows precision, or how often a model is correct when it makes a positive prediction, (*precision = TP/(FP+FN)*). The off-diagonals show false discovery rate, or how often a model is incorrect when it makes a positive prediction, (*FDR = FP/(TP+FN) = 1 - precision*). Note that the axes of F are flipped compared to G to emphasize the precedence of the predicted class for this metric.

G) Sensitivity (i.e., recall or true positive rate, TPR) and false negative rate of the object classifier. Shown is the confusion matrix where each value has been normalized by the total ground truth counts for that class (TP+FP). The diagonal shows sensitivity, how well a model can identify true positives, or the proportion of actual positives that it identified correctly, (*sensitivity = TP/(TP+FP)*). The off-diagonals show the false negative rate, or the proportion of actual positives that the model identified incorrectly, (*FNR = FN/(TP+FP) = 1- sensitivity*) for all class combinations. Note that the axes of G are flipped compared to F to emphasize the precedence of ground truth and prediction for these metrics.

for the Unknown-targeting class, with this class having 40% precision and 40% sensitivity. The confusion matrices revealed these issues were primarily due to false prediction of Non-bouton as Unknown-targeting (FNR=0.400) and misclassification of this class as Non-bouton (FDR=0.318), respectively. In total, 67% of all Unknown-targeting false positives were actually Non-bouton, and 56% of all misclassified Unknown-targeting were improperly classified as Non-bouton. Also of note for this class were the moderately high FDR (0.170) and FNR (0.140) for the dendrite class, indicating a secondary node of misclassification and misprediction for this bouton type. As stated above, however, this class is negatively defined and constitutes less than 1.5% of all boutons. Also present but far less prominent performance issue was a decreased precision for the AIS-targeting class (0.697) primarily due to their misclassifications as Soma-targeting boutons (FDR=0.206), which accounted for 68% of the misclassified AIS-targeting boutons and 20% of all AIS-targeting boutons. Of note, the AIS-targeting class is the second smallest class, comprising 3.5% of all objects and 5.0% of all boutons. Given the small proportion of this class overall, increased training with AIS-enriched images could rectify these performance issues. Alternately, the metrics we are currently using may be suboptimal for positively predicting this class. Overall, these data show that our supervised bouton classification model scores well on all performance metrics.

3.6.6. Ability to detect changes in subcellular targeting

We initially set out to build a platform to screen interneuron connectivity. There are two general requirements for such a platform: 1) scalability and 2) ability to detect changes. The latter is in part covered by the performance metrics analysis above. This analysis, however, does not take into account the biological sample's ability to change due to intervention. We therefore set out to

[26] False negative rate (FNR) informally is the proportion of actual positives that the model identified incorrectly. Formally, $FNR = FN/(TP + FP) = 1 - sensitivity$.

run a small scale pharmacological screen to assess whether our culture system was amenable to changes. To do this we partnered with Anna Molofsky's group at The University of California San Francisco, who recently found that IL13 transiently increases cortical interneuron synapse numbers *in vivo* and *ex vivo* slices[49]. In a personal communication, they noted that their preliminary data suggested the increase may be subtype-specific. We therefore used our culture system to test this hypothesis. We used Nkx2.1^{Cre};Ai34 mice to label presynaptic boutons of all MGE-derived interneurons. We then generated dissociated cortical cultures, which we grew to DIV14. We selected this time point because we found from live imaging that it corresponded to peak synaptic growth with a high level of plasticity. We then treated with increasing concentrations of IL13 for 24 hours and assessed subcellular targeting using our supervised classification system described above. This small-scale screen consisted of six different treatments with two biological replicates each. It was repeated on three independent occasions. Unexpectedly, we did not witness a change in total synapse numbers, indicating that the biological response to IL13 in culture and *in vivo* are distinct. We found, however, a profound shift in subcellular targeting at higher IL13 concentrations ($p<0.0001$ for interaction between IL13 concentration and subcellular target, by 2-way ANOVA). Specifically, we witnessed shifts of ~10% and ~20% at 100 and 200 ug/mL (both $P<0.0001$) from dendrite to soma targeting (Figure 3.14, A and B). AIS targeting was unaffected, suggesting the effect is cell type specific. Although small in scale, this experiment is proof of concept that our platform is scalable and that a dissociated culture system can be used to detect pharmacologically-induced changes in interneuron subcellular targeting. As illustrated by the unexpected lack of

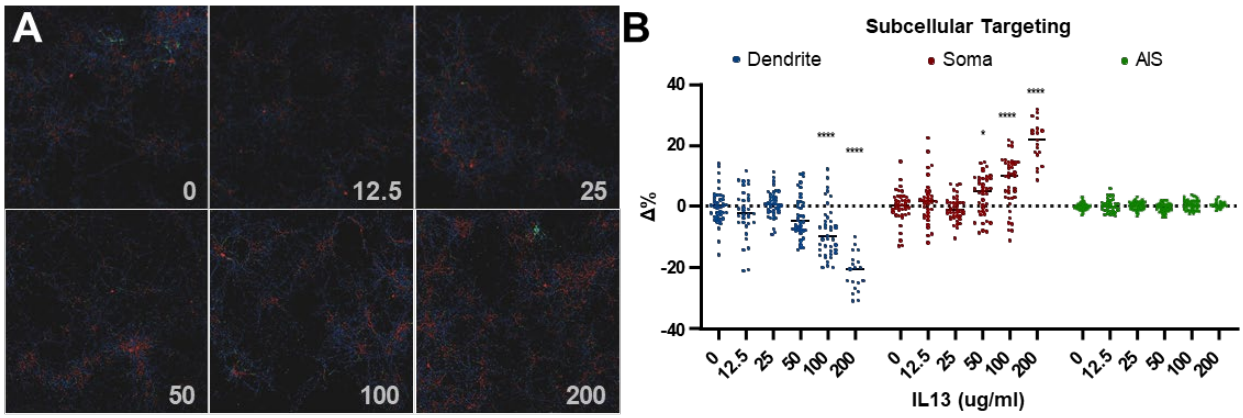

Figure 3.14 | IL13 shifts interneuron subcellular targeting preference from dendrite to soma. **A)** Examples of bouton subcellular targeting from an Nkx2.1Cre;Ai34 culture after treatment with IL13. Cultures were treated on DIV14 with vehicle or the designated concentrations of IL13 in ug/mL and analyzed 24 hours later. Boutons are colored by target. blue: dendrite; red: soma; green: AIS. **B)** Change in targeting distribution compared to vehicle. Each point is calculated from all boutons in a single tile. Data are from two biological replicates per condition, with 25 tiles per sample. Experiment was repeated three times. Representative results are shown. 2-Way ANOVA with Šídák's multiple comparisons test. * *p<0.05;* **** *p<0.0001.*

change in synapse numbers in culture, it will be important to identify the limits of recapitulation in this system and to better define *in vivo* correlates in development.

3.6.7. Unsupervised object classification

We next asked whether an unsupervised analysis of the objects would a) arrive at the same classes, and b) reveal finer classifications of boutons in an unbiased manner. Our pipeline for unsupervised clustering was adapted from single cell RNA sequencing techniques. Briefly, we took the object feature metrics computed by Ilastik as input, performed basic filtering and rescaling, reduced the dimensionality of the dataset using an autoencoder, and clustered objects embedded in this reduced space into putative classes of boutons.

Preprocessing applied standard machine learning techniques: we kept Ilastik features related only to object morphology (e.g. object size and perimeter concavities) or intensity (e.g. mean intensity of an image channel, covariance of two channel intensities), but not location (e.g. center or

bounding box of the object) or object classification (e.g. assigned class, probability of class); uninformative features with zero variance were removed; and object metrics were rescaled to the 0-1 range to ensure that autoencoder accuracy and training is not unduly influenced by individual features.

We applied dimensionality reduction to counteract the curse of dimensionality, the phenomenon that when analyzing high-dimensional datasets, data points become sparse, making it difficult to determine the underlying structure of the data. Dimensionality reduction attempts to map high-dimensional data to a lower dimension while attempting to preserve the relative structure of the data. We utilized an autoencoder[50], a neural network which is trained to reconstruct its input as output. The network features a 'bottlenecked' middle layer of lower dimension than the inputs, forcing the network to learn a reduced representation of the input data during training. Outputs from this layer serve as the embedding—the reduced-dimensionality coordinates. We selected on a four-layer network, with two layers before the embedding layer and two layers after to reconstruct the inputs, and a 16-dimensional embedding, which maximized the reconstruction accuracy on a pilot dataset while minimizing the embedding dimensionality. Training was performed on 80% of the objects, validation for each training epoch was performed on a distinct 10% of objects, and model comparison was performed on the remaining 10% of objects. The 16-dimensional embedding was subsequently used for visualization of objects in UMAP space, and for clustering.

Figure 3.15 shows all objects from a preliminary run rendered using Uniform Manifold Approximation and Projection (UMAP), a non-linear dimension reduction technique to visualize the relationships between objects[51]. The algorithm creates a network of nearest neighbors between points, then arranges these points in two dimensions, optimizing to maintain neighbors in the network. As a consequence, points nearby in UMAP space tend to be closely related. Overlaying the

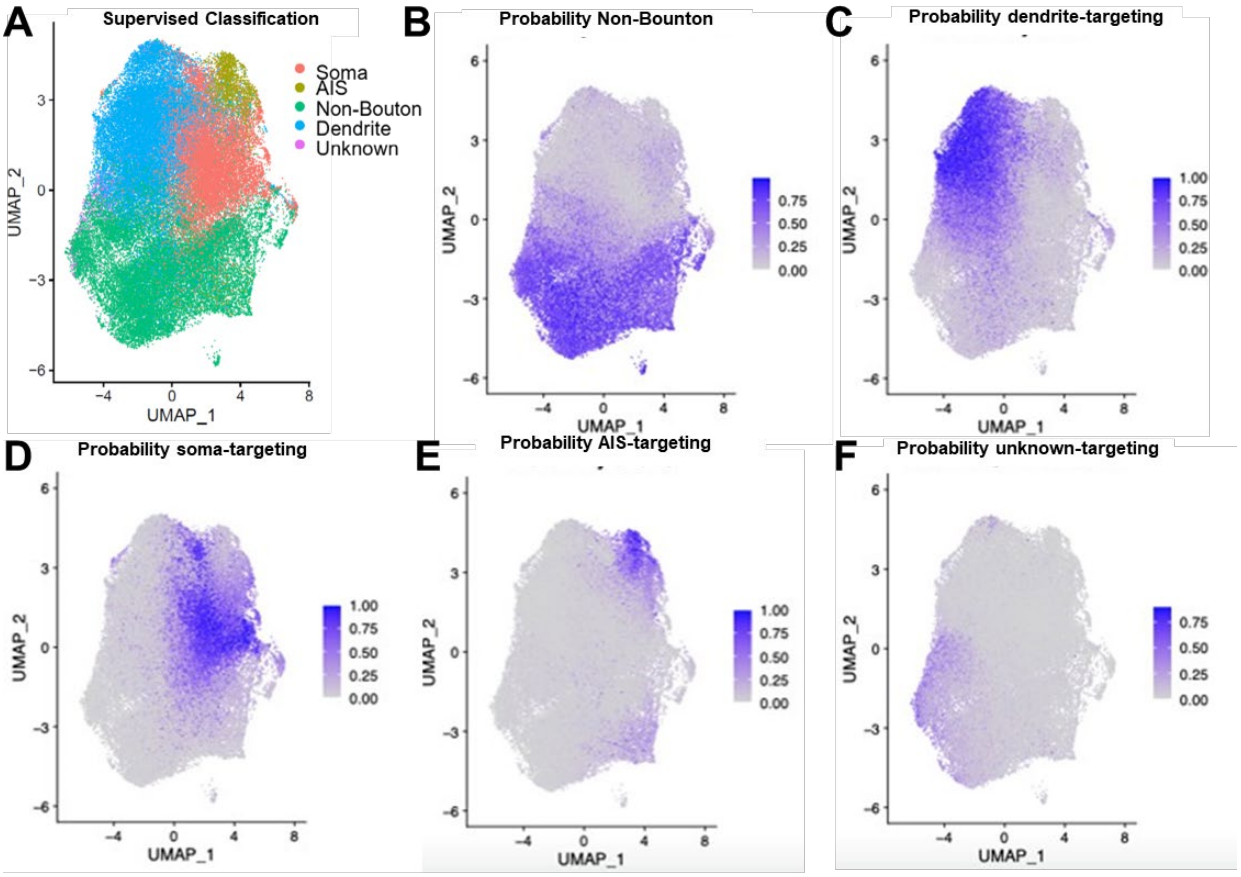

Figure 3.16 | Supervised classification calls as autoencoded parameters rendered into UMAP space.

A) The supervised object classification results were visualized on UMAP space based on unsupervised learning parameters. **B-F)** Probability of supervised classes modeled on UMAP. The non-bouton class (***B***) is largely segregated from the bouton classes. Based on these results, we excluded the non-bouton class from downstream analysis. The Dendrite-targeting (***C***), soma-targeting (***D***), AIS-targeting (***E***), and Unknown-targeting (***F***) classes all show distinct regional enrichment patterns, indicating the autoencoder recognized distinct signatures for these bouton classes.

supervised classification predictions on the UMAP plot revealed that the different classes largely cluster together, separately from other classes (Figure 3.15 A), indicating that the autoencoder embedding distinguishes between boutons of different classes. We next mapped the probability scores onto the UMAP space (Figure 3.15, B-F). These data revealed that the non-bouton class was largely segregated from the bouton classes (Figure 3.15, B). We therefore excluded these objects from further analysis. Dendrite-targeting (Figure 3.15, C), soma-targeting (Figure 3.15, D),

and AIS-targeting (Figure 3.15, E) boutons were largely restricted to their own regions within the UMAP, further validating the unsupervised approach. Objects with high probability for unknown-targeting (Figure 3.15, F) displayed considerable mixing with the non-bouton and dendrite classes, but little intermixing with soma- or AIS-targeting classes, further confirming that these are borderline cases as suggested by our ground truth validation. Overall, these data further validate our supervised model and indicated that the autoencoder had considerable potential for unsupervised learning.

We next used our autoencoder embedding as a basis for clustering. Simply put, we wanted to test whether the unsupervised learning model based on the same data would arrive at similar classes as the supervised classifier and/or identify new classes. We used the Louvain method of clustering, which detects communities of similar objects based on a weighted neighborhood network[52]. We selected this method for its robust detection of arbitrarily-structured communities, minimal parameter space, and ease of implementation. This algorithm takes two parameters—resolution, which controls the number of clusters detected, and 'k', the number of nearest neighbors to include when constructing the neighborhood network. These parameters were selected to (1) optimize two metrics of comparing intercluster and intracluster dispersion: maximizing the Calinski-Harabasz score[41] and minimizing the Davies-Bouldin score[53], under the assumption that clusters should comprise similar objects, and (2) reflect the segregation of target classes, with each cluster enriched for a single target. We discuss these parameters in greater depth in Chapter 4. This preliminary clustering analysis yielded 10 distinct classes that generally accorded with the supervised classifications: separate clusters encompassing soma-, dendrite-, and AIS- targeting classes (Figure 3.17, A and B). Moreover, the clustering analysis revealed considerable heterogeneity within the dendrite- and soma-targeting bouton classes. The heterogeneity within the soma-

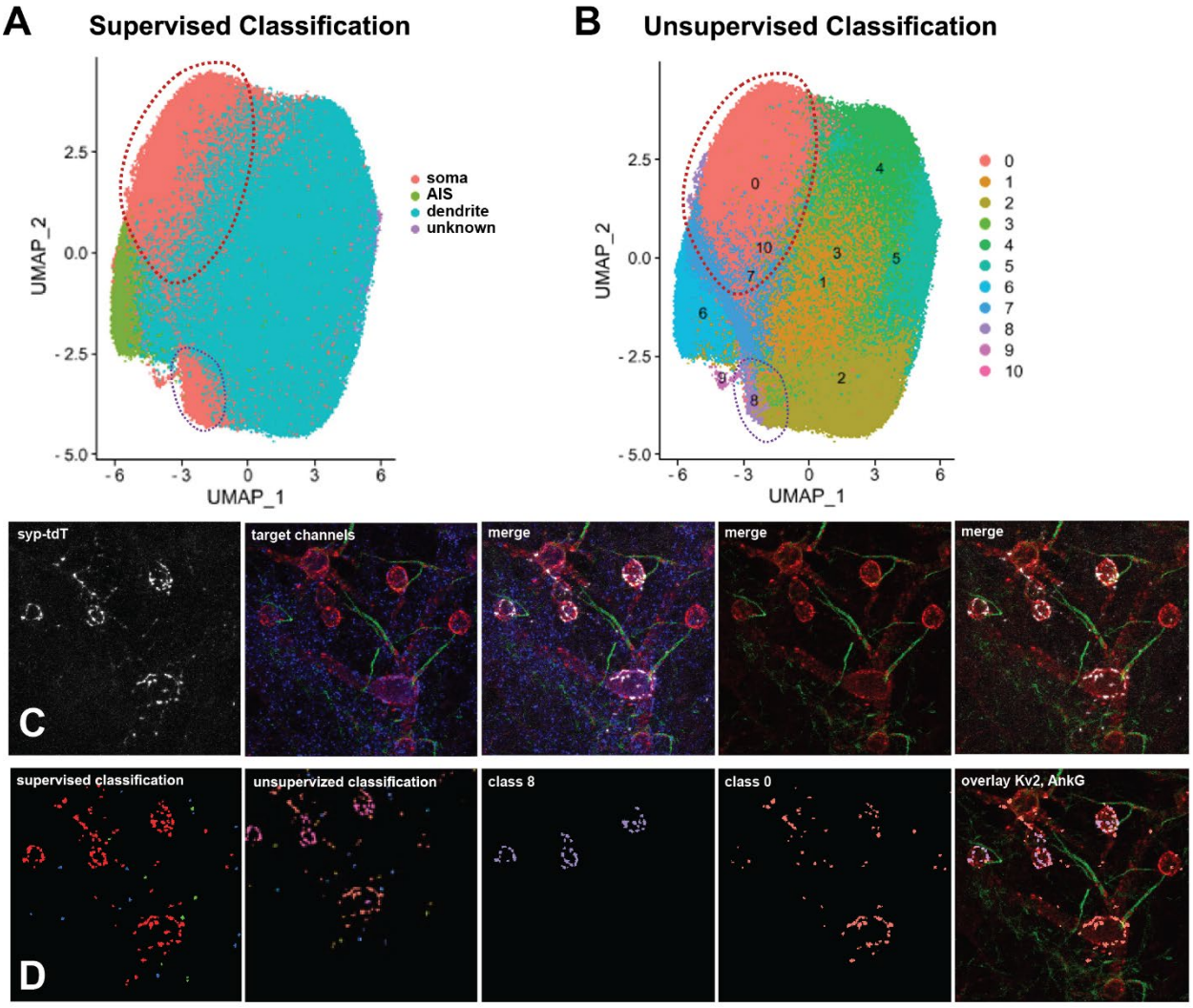

Figure 3.17 | Unsupervised classification reveals novel bouton subtypes of canonical targeting classes.

A) Supervised classification rendered onto unsupervised model and displayed as a UMAP. The soma-targeting class is split into multiple distinct regions, which are circled. **B)** Classification from a clustering algorithm based on the autoencoder parameters that comprise the unsupervised model. The unsupervised model suggests 10 distinct bouton classes that largely fit as subgroups within the targeting classes predicted by the supervised model. Two distinct soma-targeting classes (0 and 8) are circled. **C)** Original staining data for a predominantly soma-targeting interneuron. The Syp-tdT signal is from a single neuron labeled using the Nkx2.1^{CreER} driver and Ai34 reporter. This driver line labels soma- and AIS-targeting GABAergic interneurons. **D)** Object classification results mapped onto the images from ***C***. The soma-targeting boutons from the supervised classifier are split into two distinct classes by the unsupervised classifier: class 8 targets soma of small, round neurons, whereas class 0 targets large pyramidal neurons.

Note that the clusters presented here are from preliminary experiments and therefore differ from those in Chapter 4. The splitting of the soma-targeting class into two distinct clusters was observed in both, as discussed in Chapter 4.

targeting boutons was particularly enticing since the unsupervised classification showed at least two

separate regions of enrichment within the UMAP (Figure 3.17, A), which were designated as separate clusters by the clustering algorithm (classes 0 and 8, Figure 3.17, B).

3.6.8. Object class mapping

One of the major strengths of this approach is that we can visualize the classification results *in situ*. That is, we can see the classes identified by our unsupervised classifier within their original biological context by mapping them back onto the original images. To do this, we wrote custom code that takes the tabular classification results from the autoencoder and maps them onto the original object segmentation image. Of note, the algorithm is rapid and GPU-based via CLIJ2[54], and could easily be adapted to map any tabular numerical values onto a corresponding binarized mask. The code also outputs summary statistics for each tile, including counts and percentages of each class and, if supplied, overlap with the supervised classifier calls for all objects within the tile. The output is a single channel Z stack, with classes mapped by intensity value. To visualize each class independently, we also wrote code to split these classes into separate channels. Figure 3.17, C shows the staining data from an isolated interneuron expressing Syp-tdT via the Nkx2.1^{CreER} driver and Ai34 reporter, which primarily labels Soma- and AIS-targeting GABAergic interneurons. The boutons are primarily Soma targeting based on visual inspection (Figure 3.17, C) and supervised classification (Figure 3.17, D). This cell, however, displays both class 0 and class 8 Soma-targeting boutons, which show strikingly distinct characteristics: class 0 targets large, pyramidal neurons, whereas class 8 targets neurons with small, round soma (Figure 3.17, D). In other words, our unsupervised classification approach revealed novel, biologically-relevant bouton subtypes within the canonical soma targeting class. These preliminary results demonstrate the power of our approach to investigate synaptic diversity.

3.7. Conclusions

In this chapter we present an image-based tool to classify objects based on a multidimensional analysis of underlying metrics. We built this system to serve as a foundation upon which other, more complex image-based questions can be asked. In particular, the inclusion of a centralized repository for metadata allowed us to create a data ecosystem that streamlines processing and greatly increases extensibility. We used this workflow to assess cortical interneuron subcellular targeting at scale. Given the relative ease of training and its adaptability, this model could be extended to other tissue formats, such as organoids or different brain regions. In addition, the classification schema could be modified for different target staining to assess other forms of synaptic targeting, such as glutamatergic neuron targeting, or even cellular or regional specificity. Finally, the workflow could be adapted to ask biological questions outside of neuroscience. We hope that this serves as an important steppingstone. We provide a more in-depth discussion of this work and our findings from the next chapter in the Discussion section below.

Bibliography to Chapter 3: Image-based multidimensional analysis of synapses

1. Ango, F., et al., *Ankyrin-Based Subcellular Gradient of Neurofascin, an Immunoglobulin Family Protein, Directs GABAergic Innervation at Purkinje Axon Initial Segment.* Cell, 2004. **119**(2): p. 257-272.
2. Huang, Z.J., *Subcellular organization of GABAergic synapses: role of ankyrins and L1 cell adhesion molecules.* Nat Neurosci, 2006. **9**(2): p. 163-6.
3. Boiko, T., et al., *Ankyrin-Dependent and -Independent Mechanisms Orchestrate Axonal Compartmentalization of L1 Family Members Neurofascin and L1/Neuron–Glia Cell Adhesion Molecule.* The Journal of Neuroscience, 2007. **27**(3): p. 590-603.
4. Dzhashiashvili, Y., et al., *Nodes of Ranvier and axon initial segments are ankyrin G-dependent domains that assemble by distinct mechanisms.* J Cell Biol, 2007. **177**(5): p. 857-70.
5. Gasser, A., et al., *An ankyrinG-binding motif is necessary and sufficient for targeting Nav1.6 sodium channels to axon initial segments and nodes of Ranvier.* J Neurosci, 2012. **32**(21): p. 7232-43.
6. Smith, K.R., et al., *Psychiatric risk factor ANK3/ankyrin-G nanodomains regulate the structure and function of glutamatergic synapses.* Neuron, 2014. **84**(2): p. 399-415.
7. Jenkins, P.M., et al., *Giant ankyrin-G: a critical innovation in vertebrate evolution of fast and integrated neuronal signaling.* Proc Natl Acad Sci U S A, 2015. **112**(4): p. 957-64.
8. Tseng, W.C., et al., *Giant ankyrin-G stabilizes somatodendritic GABAergic synapses through opposing endocytosis of GABAA receptors.* Proc Natl Acad Sci U S A, 2015. **112**(4): p. 1214-9.
9. Lopez, A.Y., et al., *Ankyrin-G isoform imbalance and interneuronopathy link epilepsy and bipolar disorder.* Mol Psychiatry, 2017. **22**(10): p. 1464-1472.
10. Zhu, S., et al., *Genetic disruption of ankyrin-G in adult mouse forebrain causes cortical synapse alteration and behavior reminiscent of bipolar disorder.* Proc Natl Acad Sci U S A, 2017. **114**(39): p. 10479-10484.
11. Tai, Y., et al., *Axo-axonic Innervation of Neocortical Pyramidal Neurons by GABAergic Chandelier Cells Requires AnkyrinG-Associated L1CAM.* Neuron, 2019. **102**(2): p. 358-372 e9.
12. Grubb, M.S. and J. Burrone, *Activity-dependent relocation of the axon initial segment fine-tunes neuronal excitability.* Nature, 2010. **465**(7301): p. 1070-4.
13. Rasband, M.N., *The axon initial segment and the maintenance of neuronal polarity.* Nat Rev Neurosci, 2010. **11**(8): p. 552-62.
14. Kole, M.H. and G.J. Stuart, *Signal processing in the axon initial segment.* Neuron, 2012. **73**(2): p. 235-47.
15. King, A.N., C.F. Manning, and J.S. Trimmer, *A unique ion channel clustering domain on the axon initial segment of mammalian neurons.* J Comp Neurol, 2014. **522**(11): p. 2594-608.
16. Evans, M.D., et al., *Rapid Modulation of Axon Initial Segment Length Influences Repetitive Spike Firing.* Cell Rep, 2015. **13**(6): p. 1233-1245.
17. Wefelmeyer, W., D. Cattaert, and J. Burrone, *Activity-dependent mismatch between axo-axonic synapses and the axon initial segment controls neuronal output.* Proc Natl Acad Sci U S A, 2015. **112**(31): p. 9757-62.
18. Nelson, A.D. and P.M. Jenkins, *Axonal Membranes and Their Domains: Assembly and Function of the Axon Initial Segment and Node of Ranvier.* Front Cell Neurosci, 2017. **11**: p. 136.
19. Lim, S.T., et al., *A novel targeting signal for proximal clustering of the Kv2.1 K+ channel in hippocampal neurons.* Neuron, 2000. **25**(2): p. 385-97.
20. Mandikian, D., et al., *Cell type-specific spatial and functional coupling between mammalian brain Kv2.1 K+ channels and ryanodine receptors.* J Comp Neurol, 2014. **522**(15): p. 3555-74.
21. Bishop, H.I., et al., *Distinct Cell- and Layer-Specific Expression Patterns and Independent Regulation of Kv2 Channel Subtypes in Cortical Pyramidal Neurons.* J Neurosci, 2015. **35**(44): p. 14922-42.
22. Trimmer, J.S., *Subcellular localization of K+ channels in mammalian brain neurons: remarkable precision in the midst of extraordinary complexity.* Neuron, 2015. **85**(2): p. 238-56.
23. Jensen, C.S., et al., *Trafficking of Kv2.1 Channels to the Axon Initial Segment by a Novel Nonconventional Secretory Pathway.* J Neurosci, 2017. **37**(48): p. 11523-11536.
24. Bishop, H.I., et al., *Kv2 Ion Channels Determine the Expression and Localization of the Associated AMIGO-1 Cell Adhesion Molecule in Adult Brain Neurons.* Front Mol Neurosci, 2018. **11**: p. 1.
25. Kaech, S. and G. Banker, *Culturing hippocampal neurons.* Nat Protoc, 2006. **1**(5): p. 2406-15.
26. Kosik, K.S. and E.A. Finch, *MAP2 and tau segregate into dendritic and axonal domains after the elaboration of morphologically distinct neurites: an immunocytochemical study of cultured rat cerebrum.* Journal of Neuroscience, 1987. **7**(10): p. 3142-3153.
27. Gonchar, Y., et al., *Axo-axonic synapses formed by somatostatin-expressing GABAergic neurons in rat and monkey visual cortex.* J Comp Neurol, 2002. **443**(1): p. 1-14.

28. Somogyi, P., et al., *Salient features of synaptic organisation in the cerebral cortex.* Brain Res Brain Res Rev, 1998. **26**(2-3): p. 113-35.
29. Iascone, D.M., et al., *Whole-Neuron Synaptic Mapping Reveals Spatially Precise Excitatory/Inhibitory Balance Limiting Dendritic and Somatic Spiking.* Neuron, 2020. **106**(4): p. 566-578.e8.
30. Karimi, A., et al., *Cell-type specific innervation of cortical pyramidal cells at their apical dendrites.* eLife, 2020. **9**: p. e46876.
31. Hayashi, K. and T. Shirao, *Change in the shape of dendritic spines caused by overexpression of drebrin in cultured cortical neurons.* J Neurosci, 1999. **19**(10): p. 3918-25.
32. Steward, O. and S. Halpain, *Lamina-Specific Synaptic Activation Causes Domain-Specific Alterations in Dendritic Immunostaining for MAP2 and CAM Kinase II.* The Journal of Neuroscience, 1999. **19**(18): p. 7834-7845.
33. Ropelewski, A.J., et al., *Standard metadata for 3D microscopy.* Scientific Data, 2022. **9**(1): p. 449.
34. Hammer, M., et al., *Towards community-driven metadata standards for light microscopy: tiered specifications extending the OME model.* Nature Methods, 2021. **18**(12): p. 1427-1440.
35. Sarkans, U., et al., *REMBI: Recommended Metadata for Biological Images—enabling reuse of microscopy data in biology.* Nature Methods, 2021. **18**(12): p. 1418-1422.
36. Goldberg, I.G., et al., *The Open Microscopy Environment (OME) Data Model and XML file: open tools for informatics and quantitative analysis in biological imaging.* Genome Biology, 2005. **6**(5): p. R47.
37. Linkert, M., et al., *Metadata matters: access to image data in the real world.* Journal of Cell Biology, 2010. **189**(5): p. 777-782.
38. Linkert, M., et al., *Metadata matters: access to image data in the real world.* J Cell Biol, 2010. **189**(5): p. 777-82.
39. Peng, T., et al., *A BaSiC tool for background and shading correction of optical microscopy images.* Nature Communications, 2017. **8**(1): p. 14836.
40. Goldman, D.B., *Vignette and Exposure Calibration and Compensation.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010. **32**(12): p. 2276-2288.
41. Haase, R., et al., *CLIJ: GPU-accelerated image processing for everyone.* Nature Methods, 2020. **17**(1): p. 5-6.
42. Graf, J., et al., *FLINO: a new method for immunofluorescence bioimage normalization.* Bioinformatics, 2021. **38**(2): p. 520-526.
43. Sánchez-Ponce, D., et al., *Developmental expression of Kv potassium channels at the axon initial segment of cultured hippocampal neurons.* PLoS One, 2012. **7**(10): p. e48557.
44. Tyagarajan, S.K. and J.-M. Fritschy, *Gephyrin: a master regulator of neuronal function?* Nature Reviews Neuroscience, 2014. **15**(3): p. 141-156.
45. Schneider-Mizell, C.M., et al., 2020.
46. Ollion, J., et al., *TANGO: a generic tool for high-throughput 3D image analysis for studying nuclear organization.* Bioinformatics, 2013. **29**(14): p. 1840-1841.
47. Barron, J.J., et al., *Group 2 innate lymphoid cells promote inhibitory synapse development and social behavior.* bioRxiv, 2023.
48. Kramer, M.A., *Nonlinear principal component analysis using autoassociative neural networks.* AIChE Journal, 1991. **37**(2): p. 233-243.
49. McInnes, L. and J. Healy, *UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction.* ArXiv, 2018. **abs/1802.03426**.
50. Blondel, V.D., et al., *Fast unfolding of communities in large networks.* Journal of Statistical Mechanics: Theory and Experiment, 2008. **2008**: p. 10008.
51. Caliński, T. and J. Harabasz, *A dendrite method for cluster analysis.* Communications in Statistics, 1974. **3**(1): p. 1-27.
52. Davies, D.L. and D.W. Bouldin, *A Cluster Separation Measure.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 1979. **PAMI-1**(2): p. 224-227.
53. Taniguchi, H., J. Lu, and Z.J. Huang, *The spatial and temporal origin of chandelier cells in mouse neocortex.* Science, 2013. **339**(6115): p. 70-4.
54. Lim, L., et al., *Optimization of interneuron function by direct coupling of cell migration and axonal targeting.* Nat Neurosci, 2018. **21**(7): p. 920-931.

Chapter 4: Principles of cortical interneuron synaptic organization

4.1. Introduction

In vivo, distinct GABAergic cortical interneuron populations target specific subcellular domains of the target cell (Figure 4.1, A). Somatostatin cells (Sst, blue) target distal dendrites, basket cells (BC, red) target soma and proximal dendrites, and chandelier cells (ChC, green) target the axon initial segment. We initially set out test whether these cortical interneurons require a precisely patterned cortical environment to form the proper subcellular connections. We hypothesized that interneuron subcellular targeting specificity is largely independent of the broader cortical environment (laminar, positioning, afferent connections, etc.), and as such would be recapitulated in a dissociated culture setting (Figure 4.1, B). Note that recapitulation implies not only that targeting specificity is intrinsically determined by the presynaptic cell, but also that compartmentalization of the target molecules in the correct subcellular regions is intrinsically determined by the target cell.

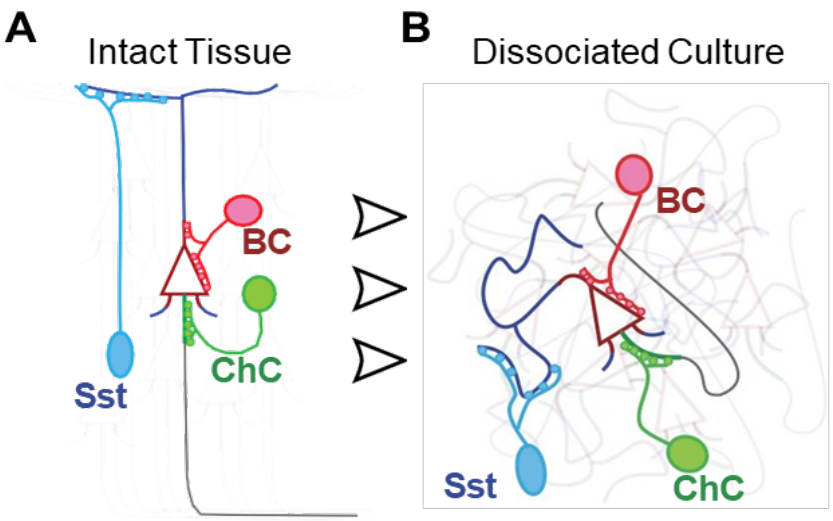

Figure 4.1 | GABAergic interneuron subcellular targeting specificity in vivo and model for interneuron specificty in culture.
A) In vivo, distinct GABAergic cortical interneuron populations target specific subcellular domains of the target cell. Somatostatin cells (Sst, blue) target distal dendrites, basket cells (BC, red) target soma and proximal dendrites, and chandelier cells (ChC, green) target the axon initial segment. **B)** We hypothesized that interneuron subcellular targeting specificity is largely independent of the broader cortical environment (laminar,

positioning, afferent connections, etc.), and as such would be recapitulated in a dissociated culture setting. Note that recapitulation implies not only that targeting specificity is intrinsically determined by the presynaptic cell, but also that compartmentalization of the target molecules in the correct subcellular regions is intrinsically determined by the target cell.

To state the experiment formally, our null hypothesis postulates that cortical interneuron subcellular targeting is dependent on one or more of the following:

1. The interneuron's positioning within a specific cortical subregion or layer.
2. Some pre-existing cellular structure(s) or molecular gradient(s) formed by intermediate (i.e. non-targeted) cells and utilized by the presynaptic and/or postsynaptic cell(s).
3. Indirect interaction(s) between the cortex and extracortical tissue. Examples include signaling from the choroid plexus, vasculature, and afferent or efferent (extracortical) projections.

In other words, if subcellular targeting depends on any of the above principles, then disrupting these processes by excising the cortex, randomizing its contents, and growing it in isolation should also randomize interneuron targeting. Alternately, if subcellular targeting specificity *is* recapitulated in this randomized and isolated context, we can conclude that:

1. Expression of the presynaptic molecules necessary for such targeting is intrinsically determined by the presynaptic cell. Specifically, interneuron identity is designated s prior to settling and axodendritic polarization and is stable in a randomized cellular context.
2. The molecular gradients responsible for localizing the target adhesion molecules to the 'appropriate' subcellular location are intrinsically determined in the target cell. Specifically, the molecular identity of the target subcellular compartments is determined prior to settling. These gradients form during or after axodendritic polarization, and are stable in a randomized cellular context.

3. Direct contact between the pre- and postsynaptic adhesion molecules alone is *sufficient* for 'appropriate' interneuron subcellular targeting specificity.

4.2. Experiment design

Testing this model required differentially labeling three different interneuron classes. To do this we harnessed a number of well-characterized genetic tools based on the Cre-lox driver-reporter system (see Chapter 1, Figure 4.2). Specifically, we used the following driver lines: Nkx2.1^{Cre} to label all MGE-derived interneurons (Figure 4.2, A), Sst^{Cre} to label the Sst+ population of dendrite-targeting interneurons (Figure 4.2, B), PV^{Cre} to label all PV-expressing cells in cortical slices (Figure 4.2, C); and the inducible Nkx2.1^{CreER} line with late embryonic (E17.5) tamoxifen induction to label both chandelier and basket cells[27] (Figure 4.2, D and E, respectively). We used Nkx2.1^{CreER} for the culture system instead of the canonical PV^{Cre} line because cortical interneurons do not express PV until late adolescence or early adulthood *in vivo*, and expression does not occur at all in our culture conditions (data not shown). We also included the Nex^{Cre} driver line, which labels cortical glutamatergic excitatory neurons, as a negative control for GABAergic synapses and a positive control for glutamatergic synapses during validation experiments, as discussed in greater depth in Chapter 1, Labeling the presynaptic compartments: reporter selection.

[27] Late tamoxifen induction of the Nkx2.1^{CreER} mouse labels late-born cortical interneurons. These are largely ChCs and BCs. See the Cortical Development section of the Introduction for the molecular and developmental basis for differential labeling in this mouse compared to the constitutively active Nkx2.1^{Cre} mouse.

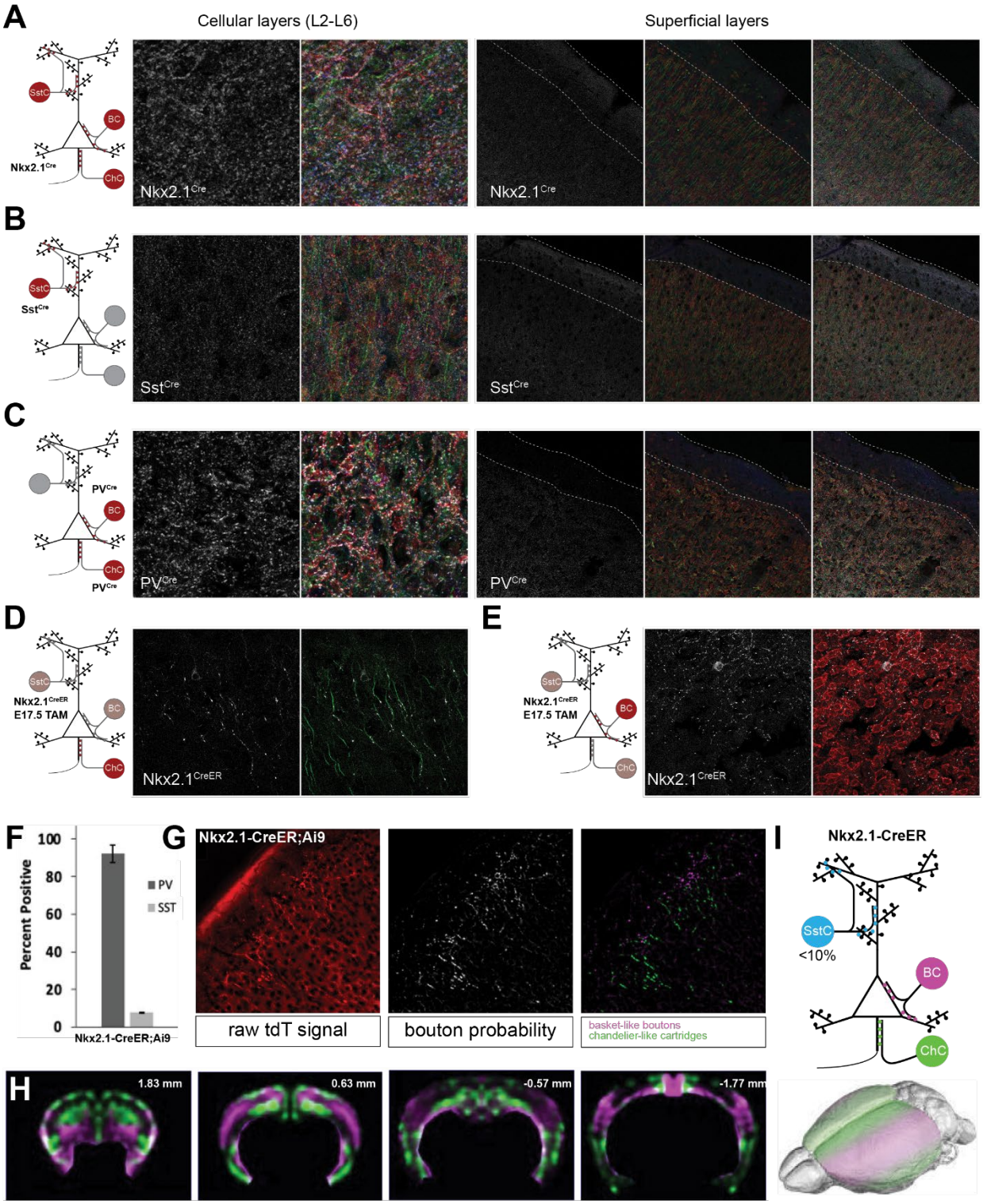

Figure 4.2 | Presynaptic labeling of distinct interneuron populations using mouse driver lines and the Ai34 reporter.

A-C) The labeling strategies used for the specific cortical interneuron populations. In each case, a schematic of the Cre driver and the labeled interneuron population is shown at left. A high resolution example image showing the driver line's targeting specificity within the cellular (L2-L6)

layers of the cortex is shown in middle. A low resolution image of the superficial layers of the cortex at right shows the distal dendrite (L1) targeting pattern of each labeled population. Note that PV is not expressed in our culture conditions, precluding use of the PV^{Cre} line for our tissue culture experiments. We opted to use the $Nkx2.1^{CreER}$ line for these experiments instead.
D-E) Isolated examples of chandelier (***D***) and basket (***E***) cells labeled after late embryonic (E17.5) tamoxifen (TAM) induction of the inducible CreERT2 (CreER) construct in the $Nkx2.1^{CreER}$ mouse. Note that $Nkx2.1^{Cre}$ (***A***) labels all MGE-derived neurons, whereas E17.5 TAM induction of the $Nkx2.1^{CreER}$ mouse (***D-E***) selectively labels late-born MGE interneurons. Potentially labeled interneuron populations not shown in the image are displayed in muted colors.
F) Late tamoxifen induction of the $Nkx2.1^{CreER}$ mouse labels almost exclusively the PV+ population of cortical interneurons, with minimal labeling of Sst+ cells. Since most interneurons labeled by the $Nkx2.1^{CreER}$ mouse are PV+ and Sst-, we concluded that the majority of labeled cells are basket or chandelier cells.
G) Since there are no positive markers to differentiate chandelier cells and basket cells, we used bouton morphology as a means to differentiate regions enriched for these populations. Shown are examples of morphology-based bouton classification to distinguish chandelier cell cartridges from non-cartridge boutons in cortical slices. Cartridges are shown in green, whereas non-cartridge boutons are shown in magenta. Based on the data from ***F***, we concluded that the non-cartridge boutons are from soma-targeting basket cells.
H) A heat map of the of cartridge and non-cartridge bouton distribution within the cortex of the $Nkx2.1^{CreER}$ mouse at different Bregma coordinates along the anterior-posterior axis. Sections were processed as free-floating, and we thus lost information about left-right orientation for these slices. We have therefore assumed bilateral symmetry.
I) A schematic summarizing the data from ***D-H*** characterizing the populations labeled by E17.5 TAM induction of the $Nkx2.1^{CreER}$ mouse. Note that the labeled populations of AIS-targeting chandelier cells and soma-targeting basket cells are differentially enriched across the cortex in this mouse.

Since all MGE-derived interneurons come from Nkx2.1-expressing progenitors, we were concerned about the specificity of our strategy of using the $Nkx2.1^{CreER}$ driver line to label BCs and

ChCs. We therefore sought to validate this mouse to label these populations. We first wanted to assess the extent to which late tamoxifen induction of the $Nkx2.1^{CreER}$ mouse labeled 1) Sst+ interneurons, 2) PV+ interneurons, and 3) Sst-/PV- interneurons. Since Sst and PV localize to the cell body, we used the Ai14 reporter, which like Ai34 is a Cre-dependent reporter, but unlike Ai34 leads to expression of bright tdTomato that localizes throughout the cell. We then stained a series of coronal sections sampled from across the cortex for Sst and PV. We trained different pixel

classifiers to identify Sst+, PV+ soma, and tdT+ soma and boutons (data not shown), then segmented the labeled soma and assessed overlap to derive all staining combinations across all cortical sections. In this way we were able to classify tdT+ soma as being Sst cells (Sst+/PV-), PV cells (Sst-/PV+), dual-postive (Sst+/PV+) or neither (Sst-/PV-). Approximately 90% tdT-labeled neurons were PV cells, and fewer than 10% of all were Sst cells (Figure 4.2, F). We witnessed minimal labeling of Sst+/PV+ cells generally, and none that were also tdT+. These data ruled out the concern for any significant labeling of Sst interneurons.

Since the Nkx2.1^{CreER} mouse labels both BCs and ChCs, we wanted to know whether we could target specific regions of this mouse during dissection to generate cultures enriched for these cell types. We therefore asked whether there were cortical regions of this mouse that were differentially enriched for labeled BCs and ChCs. BCs and ChCs do not have markers that categorically distinguish these cell types. BCs and ChCs do, however, have morphologically distinct boutons: ChCs form cartridges, whereas BCs do not. We therefore used this feature to assess regional enrichment. To do this, we used the tdT pixel classifier mentioned above, and extracted only the tdT+ boutons. We were unable to train a pixel classifier using the default feature sets in Ilastik to differentiate cartridge from non-cartridge tdT+ boutons. We therefore opted for a different approach. ChC cartridges have distinct, elongated, vessel-like morphology compared to dispersed, individual boutons or perisomatic baskets. Moreover, cartridges tend to be oriented radially relative to the cortical surface. We therefore used native feature extraction code in FIJI to assess vesselness and orientation within the bouton probability channel. We then built a scale that took these features into account to classify pixels as either cartridge or non-cartridge (Figure 4.2, G). We then created statistical maps of local average probability normalized to overall bouton probability to map local enrichment of cartridge and non-cartridge boutons. Finally, we registered these enrichment maps to

the Allen Mouse Common Coordinate Framework (AllenCCF), a 3D brain reference atlas that can be used register brain slices from different animals to a common coordinate system. By registering all slices to the same common reference atlas, were able to generate a 3D reconstruction of all labeled cells throughout the brain from which we derived a statistical model for Nkx2.1^{CreER} cellular labeling. The results for the slice series are shown in Figure 4.2, H. Based on these data we confirmed a previously observed medial enrichment of ChCs as well as a decrease in cartridges along the rostral->caudal axis[1]. Additionally, we demonstrated for the first time a ventrolateral enrichment of cartridges in the piriform cortex and a dorsolateral enrichment BCs from S1 to the end of the neocortex. These results are summarized in Figure 4.2, I. Based on these data we concluded that the ChC and BC populations labeled by late embryionic (E17.5) tamoxifen induction of the Nkx2.1^{CreER} mouse are sufficiently distinct and could be targeted during dissection for enrichment.

Having validated the Nkx2.1^{CreER} line for labeling BCs and ChCs, and armed with a dissection strategy, we next crossed our Cre driver lines with the Ai34 reporter line, which permanently labels cre-expressing cells with presynaptically-localized synaptophysin::tdTomato (Syp-tdT) (Figure 4.2, A-E). We discuss our validation of this presynaptic reporter in Chapter 3, Labeling the presynaptic compartments: reporter selection.

We next turned to visualizing the subcellular target compartments of the three interneuron types of interest (Figure 4.3, A). As discussed in Labeling the target compartments: antibody selection of Chapter 3, we eventually settled on using Ankyrin G to label AIS, a combination of Kv2.1 and Kv2.2 (Kv2) to label soma and proximal dendrite, and Gephyrin to label GABAergic PSDs in general. A schematic of the compartments and the antibodies used to visualize them is shown in Figure 4.3, B.

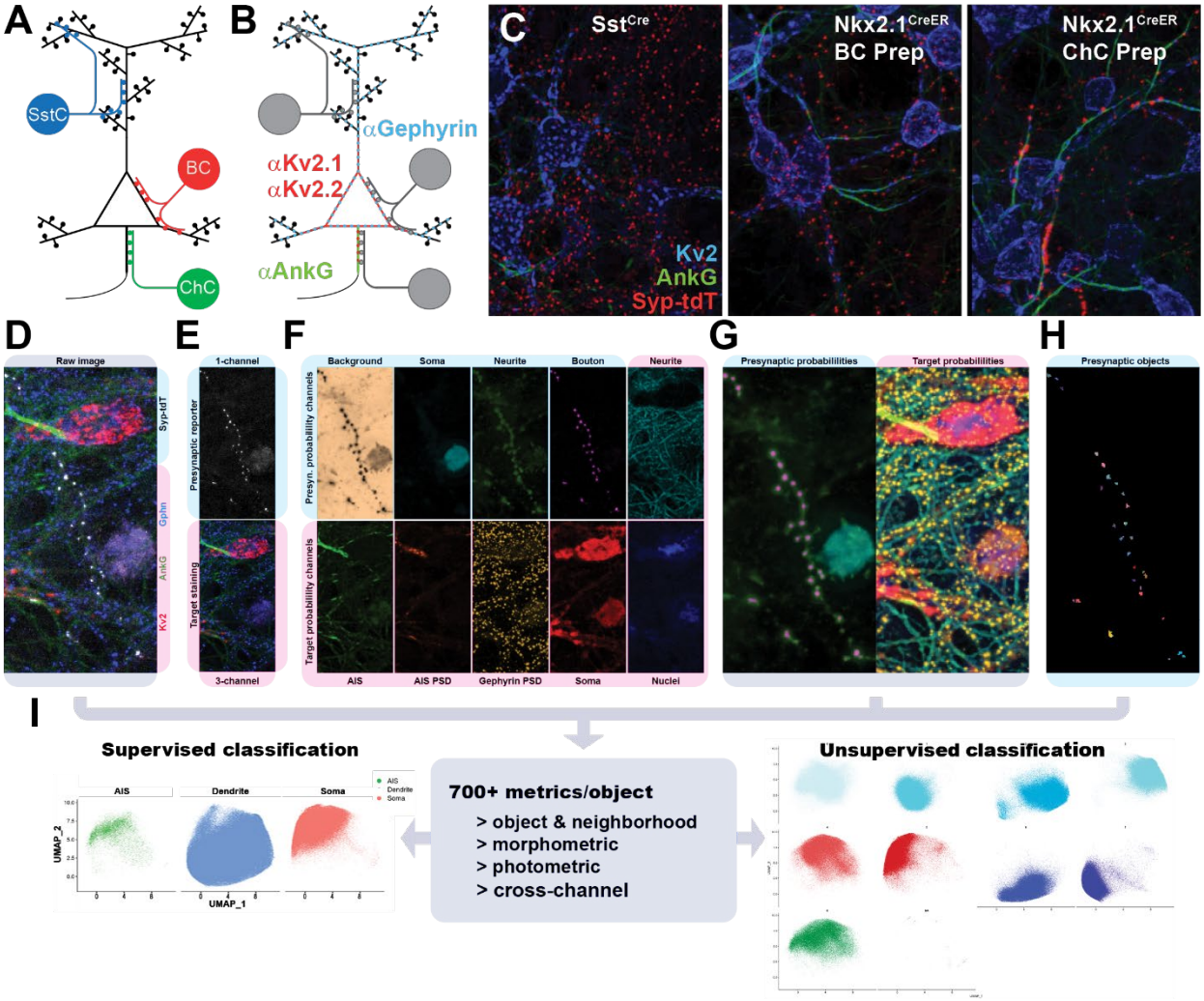

Figure 4.3 | Target compartment labeling and bouton targeting classification strategies.
A) Schematic of the three interneuron classes and their subcellular targets. **B)** Schematic of the antibody staining approach to visualize the subcellular target compartments. **C)** Example images of subcellular targeting in dissociated culture from three different sample preparations. Antibody labeling is colored as in B, but gephyrin staining is excluded. **D)** Example raw 4-channel input image. **E)** Inputs for each machine learning (ML)-based pixel classification. Presynaptic inputs are in blue, target in magenta. **F)** ML-based pixel classification outputs. Presynaptic classifier bins presynaptic reporter (Syp-tdT)-positive pixels into four distinct subclasses (background, bouton, neurite, soma). Target classifier bins target channels into seven target subcompartments (background, Soma/Proximal dendrite, AIS, Gephyrin PSDs, Neurite, Kv2 at the AIS, and nuclei). **G)** Pixel classifications are exported as probability maps, with one channel per class. **H)** The raw Syp-tdT channel and presynaptic probability maps are used for bouton identification and segmentation. **I)** Metrics are extracted for each object from the fifteen channels (raw, presynaptic, target) and used for supervised (*left*) and/or unsupervised (*right*) object classification.

Given this experimental framework, we were now ready to test our model. To restate the hypothesis briefly, we sought to test whether a) cortical interneuron subcellular synaptic specificity (Figure 4.3, A) requires some higher order cortical organization present when these neurons begin to send out their axonal projections, or b) if the cell intrinsic programming and direct cell-cell interactions that remain after excision and randomization are sufficient for 'proper' subcellular targeting. Randomized subcellular targeting in the randomized context would support the former, whereas recapitulation would support the latter.

Preliminary experiments using bulk cultures from Sst^{Cre} mice and BC and ChC preps from $Nkx2.1^{CreER}$ mice (Figure 4.3, C) suggested a differential bias in Sst^{Cre} cultures towards dendrite-targeting boutons and $Nkx2.1^{CreER}$ towards soma- and AIS-targeting. The strongest indicator was the presence of cartridge-like structures in the $Nkx2.1^{CreER}$ ChC culture, which were completely absent from the Sst^{Cre} culture.

These promising qualitative results prompted us to quantify the data. The large number of labeled boutons meant that manual counting and classification would be impossible from a practical standpoint. We therefore developed a machine learning based platform to classify and quantify boutons by subcellular target at scale (Figure 4.3, D-I). We discuss this tool at length in Chapter 3. Armed with this tool, we revisited our question from a quantitative and statistical perspective.

4.3. Supervised classification of cortical interneuron boutons

4.3.1. GABAergic interneuron subcellular targeting is recapitulated in dissociated culture

We first wanted to establish a gold standard to which to compare subcellular targeting in culture. We established these benchmarks by evaluating subcellular targeting of these different driver lines in cortical slice (Figure 4.4, A-B). $Nkx2.1^{Cre}$ labels all of the MGE-derived interneuron clas-

ses. We therefore used the targeting distribution of this mouse as a baseline when comparing between driver lines. We expected Sst^{Cre} to display increased dendrite- and decreased soma- and AIS-targeting. We expected PV^{Cre} to show the opposite trend: more soma and AIS targetin, and less dendrite targeting. In aggregate, we expected $Nkx2.1^{CreER}$ to show higher proportions of AIS-targeting, but were agnostic to the amount of soma and dendrite targeting. Moreover, we expected the ChC-enriched regions of this mouse display more AIS-targeting and BC-enriched regions more soma-targeting. As shown in Figure 4.4, B, this is precisely what we witnessed. The $Nkx2.1^{Cre}$ slices displayed comparable levels of soma and dendrite targeting; PV^{Cre} displayed predominantly soma targeting; Sst^{Cre} displayed a strong preference for dendrite-targeting AIS targeting. $Nkx2.1^{CreER}$ displayed regional variability: overall, the targeting proportions were similar to $Nkx2.1^{Cre}$, but with slight enrichment of AIS targeting; in BC-enriched regions this AIS targeting was largely absent; in ChC-enriched regions it was more prominent.

Given this benchmark, we were finally prepared to ask whether subcellular targeting is recapitulated in dissociated culture (Figure 4.4, C). We generated dissociated cultures from $Nkx2.1^{Cre}$, Sst^{Cre}, and $Nkx2.1^{CreER}$ mice. For the latter, we included both bulk cultures and BC and ChC-targeted preps (Figure 4.4, D-E). As discussed in Chapter 1, PV is not expressed in dissociated cultures (data not shown), so we excluded the PV^{Cre} driver line from this analysis. Again, we used $Nkx2.1^{Cre}$ as a baseline, expecting similar trends as in slice. If subcellular targeting is not recapitulated, we should expect no difference between the three driver lines. If it is, then the Sst^{Cre} culture should display higher levels of dendrite-targeting and lower levels of soma- and AIS-targeting than $Nkx2.1^{Cre}$, whereas the $Nkx2.1^{CreER}$ culture should have lower proportion of dendrite-targeting boutons and a higher level of soma- and AIS-targeting. As shown in Figure 4.4, E, this is indeed

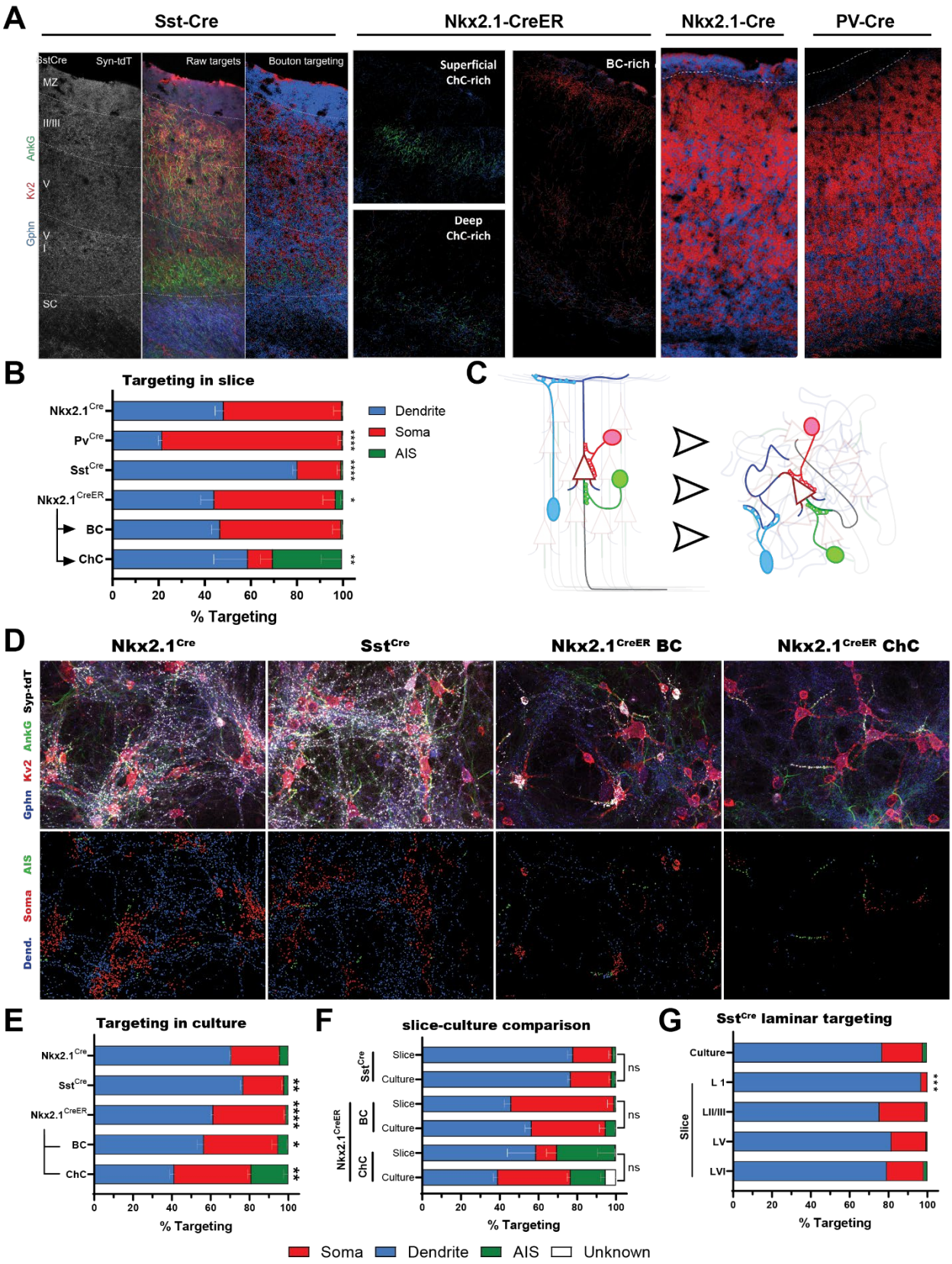
A
Sst-Cre
Nkx2.1-CreER
Nkx2.1-Cre
PV-Cre
SstCre
Syn-tdT
Raw targets
Bouton targeting
MZ
II/III
V
V
I
SC
AnkG
Kv2
Gphn
Superficial ChC-rich
Deep ChC-rich
BC-rich
B
Targeting in slice
Nkx2.1Cre
PvCre
SstCre
Nkx2.1CreER
BC
ChC
Dendrite
Soma
AIS
% Targeting
C
D
Nkx2.1Cre
SstCre
Nkx2.1CreER BC
Nkx2.1CreER ChC
Syp-tdT
AnkG
Kv2
Gphn
AIS
Soma
Dend.
E
Targeting in culture
% Targeting
F
slice-culture comparison
SstCre
Nkx2.1CreER
BC
ChC
Slice
Culture
ns
% Targeting
G
SstCre laminar targeting
Culture
L 1
LII/III
LV
LVI
Slice
% Targeting
Soma
Dendrite
AIS
Unknown

Figure 4.4 | Recapitulation of interneuron subcellular targeting in a dissociated culture system. **A)** Example images of bouton targeting in slice. Syp-tdT and target signals are shown for Sst^{Cre} only; all others show maximum projection images of bouton targeting. Approximate laminar boundaries are denoted in white. For $Nkx2.1^{CreER}$, targeting is shown for superficial and deep ChC-enriched regions and BC-enriched regions. **B)** Quantification of targeting by Cre driver line in tissue slices. Statistical comparisons use the $Nkx2.1^{Cre}$ sample as the expected targeting distribution. The $Nkx2.1^{CreER}$ samples do not distinguish between enriched regions. **C)** Schematic of the subcellular targeting of cortical interneurons in intact tissue (*left*) and their hypothesized targeting in dissociated culture (*right*). Dendrite-targeting Sst cells are in blue, soma- and proximal dendrite-targeting basket cells (BCs) are in red, and AIS-targeting chandelier cells (ChCs) are in green. **D)** Example images of bouton targeting in dissociated culture for $Nkx2.1^{Cre}$, Sst^{Cre}, and $Nkx2.1^{CreER}$ BC and ChC preps. Staining is above, and targeting below. **E)** Quantification of targeting distributions in culture. Statistical comparisons use the $Nkx2.1^{Cre}$ sample as the expected targeting distribution. **F)** Comparison of targeting in tissue slices and dissociated culture for Sst^{Cre} and $Nkx2.1^{CreER}$. The latter compares BC-rich slice regions with BC-enriched cultures or ChC-rich slice regions with ChC-enriched cultures. There was no significant difference between slice and culture conditions. **G)** Comparison of SstCre targeting in dissociated culture and to different layers in slice. ** P<0.5, ** P<0.1, *** P<0.01, **** P<0.001.*

what we witnessed. The $Nkx2.1^{Cre}$ culture displayed an intermediate phenotype between the primarily dendrite-targeting Sst^{Cre} culture and the enriched soma-targeting of $Nkx2.1^{CreER}$ bulk cultures or BC prep. AIS-targeting was significantly enhanced in the $Nkx2.1^{CreER}$ ChC prep. These data strongly suggested retention of identity and recapitulation of targeting in dissociated culture. Indeed, direct comparison between slice and culture revealed no statistical difference between these conditions (Figure 4.4, F). These data suggest that GABAergic interneuron subcellular targeting is primarily intrinsically determined.

Given the current model for Martinotti cell targeting of L1[2] and the absence of distal dendritic tuft formation in dissociated culture[3], we were surprised not to see a difference between slice and culture for Sst^{Cre}. We noted that in slice, L1-targeted boutons are a minority of all boutons labeled by Sst^{Cre} (Figure 4.4, A, G). We therefore hypothesized that Sst^{Cre} targeting in culture more closely resembled the cellular layers (L2-L6) of the cortex. Comparing Sst^{Cre} culture targeting to targeting in different cortical layers revealed that the heavily-enriched dendrite targeting in L1 was indeed

an outlier (Figure 4.4, G). Thus, the culture targeting profile of Sst^{Cre} closely matched targeting of the cellular layers (L2-L6) of the cortex. In summary, these data strongly support a model whereby MGE interneuron subcellular targeting specificity and identity are recapitulated in dissociated culture.

4.3.2. GABAergic interneuron boutons are highly organized within the cortex

Adapting our platform to slice data allowed us for the first-time to assess cortical interneuron targeting in its native context on a population-wide scale. Given that the cortex is highly structured, we next asked how subcellular targeting of different interneuron populations differs across this structured environment.

For these studies we decided to swap the $Nkx2.1^{CreER}$ mouse for PV^{Cre}. We did this for several reasons. First, since we were now working in an *in vivo* context, we were no longer limited by PV expression issues in culture, allowing us the option to use the PV^{Cre} driver to label these interneuron populations. Second, from a practical standpoint, for the $Nkx2.1^{CreER}$ mouse, the need to manually induce Cre recombination through tamoxifen induction within a narrow embryonic window leads to considerable litter-to-litter variability in labeled populations. PV^{Cre}, on the other hand, labels cells automatically based on intrinsic PV expression and so does not have these consistency issues. PV^{Cre} is therefore a more robust labelling method. Third, the PV^{Cre} mouse labels a larger population of interneurons than the $Nkx2.1^{CreER}$ mouse. Specifically, within the molecularly defined taxonomy of cortical interneurons[4], it labels the entire *Pvalb* subclass of the MGE lineage and therefore provides a more comprehensive picture of this subclass of interneurons. Finally, as an extension of the last point on interneuron taxonomy, the combination of $Nkx2.1^{Cre}$, Sst^{Cre}, and PV^{Cre} offers a taxonomically nested set: $Nkx2.1^{Cre}$ labels the entire *MGE* neighborhood of cortical

GABAergic interneuron presynaptic boutons, which in turn comprises three subclasses. PV^{Cre} labels the *Pvalb* subclass, whereas Sst^{Cre} labels the *Sst* and *Sst-Chodl* subclasses. This driver combination, then allows us to assess subcellular targeting specificity of distinct taxonomically-defined MGE interneuron populations, as well as the MGE interneuron population as a whole.

As outlined in the introduction, the cortex is conventionally divided into six distinct layers (L1-L6), each of which has a distinct cellular composition. We anticipated the greatest differences in targeting enrichment to track with acellular (L1) and cellular (L2-L6) layers of the cortex. L1 is largely acellular, being comprised mainly of complex network of axonal and dendritic connections. The vast majority of GABAergic inputs within this layer from MGE-derived interneurons come from L2/3 and L5 Sst Martinotti cells (MCs), which synapse onto pyramidal cell distal dendritic tuft. Very few PV interneurons project into this layer, and very little soma- and AIS-targeting occurs here. In contrast, L2-L6 are all cellular layers, with distinct glutamatergic neurons inhabiting (and thereby defining) each layer. Compared to L1, L2-L6 are heavily enriched for soma- and AIS-targeting, though high levels of (non-proximal) dendrite targeting also occurs. Although the axo-axonic connections more prevalent in the cellular layers of the cortex, they comprise a small minority population overall compared to soma or dendrite targeting synapses.

Given this organization, we expected different cortical interneuron populations to display differential enrichment of bouton subtypes between L1 and L2-L6. Specifically, we had the following expectations:

1. L1 should be almost exclusively dendrite-targeting boutons. This targeting should be evident in $Nkx2.1^{Cre}$ and Sst^{Cre}, but not PV^{Cre} mice.
2. Compared to L1, L2-L6 should be much more heavily enriched for soma- and AIS targeting, and less for dendrite-targeting.

3. Within L2-L6, soma- and AIS-targeting should be as follows: $PV^{Cre} > Nkx2.1^{Cre} > Sst^{Cre}$.
4. Within L2-L6, dendrite-targeting should be as follows: $Sst^{Cre} > Nkx2.1^{Cre} > PV^{Cre}$.
5. Within L2-L6 we expected PV^{Cre} to label primarily (but not exclusively) soma-targeting boutons, and the Sst^{Cre} labeled population to be enriched for (but not exclusively) dendrite-targeting boutons.
6. We expected the AIS-targeting population to comprise a minority of all interneuron boutons, add to be enriched as follows: $PV^{Cre} > Nkx2.1^{Cre} > Sst^{Cre}$.
7. Finally, since $Nkx2.1^{Cre}$ is comprised of both Sst and PV interneurons, we expected the boutons labeled by $Nkx2.1^{Cre}$ to be a composite of the PV^{Cre} and Sst^{Cre} populations.

As shown in Figure 4.4, A, and in greater detail below, all but one of these expectations were met. Specifically, we were unable to satisfactorily confirm #6 due to an overall lack of AIS targeting across all mice (quantified in Figure 4.4, B). Although we did witness AIS-targeting in all three driver lines, we saw considerably less than expected, especially for PV^{Cre}. We were therefore not confident to make conclusions about this population within these driver lines. We suspect that this underrepresentation arises in part from specificity issues of our supervised classifier (see Chapter 3.6.5 discussion of supervised classifier performance metrics) and too lenient inclusion of non-bouton objects during segmentation, which are in turn misclassified as soma- or dendrite-targeting (covered in greater depth in the Discussion section below). We therefore present the AIS targeting data alongside the soma- and dendrite-targeting data. Given our low confidence in these results, however, we caution against over interpretation.

Mapping the supervised bouton classification back onto the original images not only confirmed these expectations, but also revealed unexpected enrichment patterns of bouton populations across cortical architecture (Figure 4.5, A and B). As discussed above, we had anticipated major

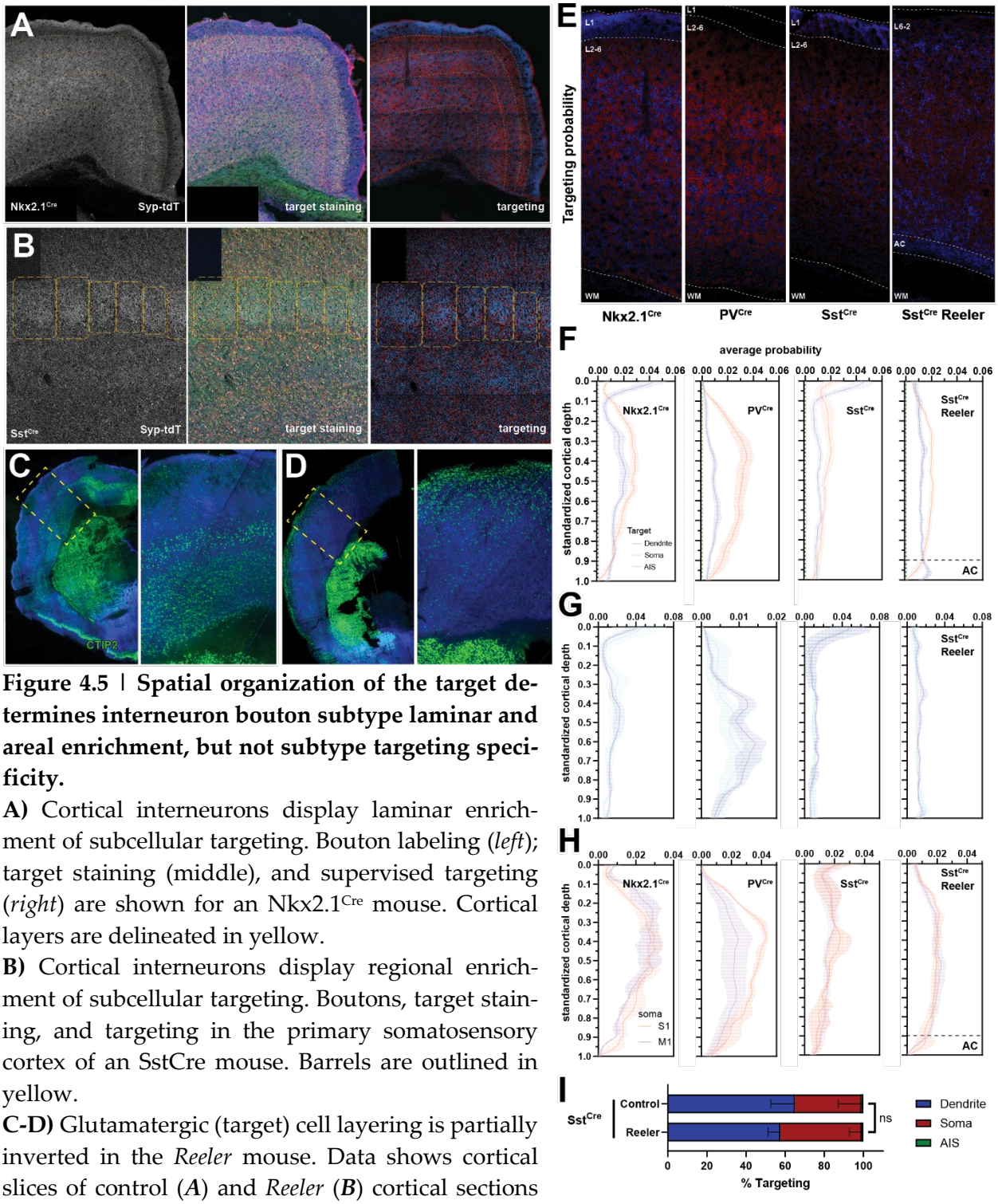

Figure 4.5 | Spatial organization of the target determines interneuron bouton subtype laminar and areal enrichment, but not subtype targeting specificity.

A) Cortical interneurons display laminar enrichment of subcellular targeting. Bouton labeling (*left*); target staining (middle), and supervised targeting (*right*) are shown for an Nkx2.1^{Cre} mouse. Cortical layers are delineated in yellow.

B) Cortical interneurons display regional enrichment of subcellular targeting. Boutons, target staining, and targeting in the primary somatosensory cortex of an SstCre mouse. Barrels are outlined in yellow.

C-D) Glutamatergic (target) cell layering is partially inverted in the *Reeler* mouse. Data shows cortical slices of control (***A***) and *Reeler* (***B***) cortical sections stained for *CTIP2* staining (green), a marker of early-born glutamatergic cells that reside in deep layers. Background autofluorescence is shown in blue as a reference. Note the partially inverted and/or randomized positioning of these cells in the *Reeler* mouse compared to control.

E) Examples of subcellular targeting probabilities for different interneuron populations across a cortical column. The spatial distribution of Sst interneuron targeting is disrupted in the Reeler mouse. WM, white matter. AC, acellular region.
F) Distinct interneuron bouton populations display distinct targeting enrichment profiles across a cortical column. Data shows the average probability of encountering a bouton of the designated targeting class between the cortical surface (0) and start of WM (1). Note that the Nkx2.1^{Cre} mouse, which labels both the PV+ and Sst+ interneuron populations, displays a targeting profile that resembles a combination of the PV^{Cre} and Sst^{Cre} profiles. Note also that the enrichment profile of Sst^{Cre} is partially inverted in the *Reeler* mouse, with dendrite-targeting enrichment occurring at the cortical base rather than surface.
G-H) Laminar enrichment profiles of bouton targeting differ by cortical region. As in ***F***, but with dendrite (***G***) and soma (***H***) targeting profiles by cortical region (M1 and S1BF). Note that the two dendrite targeting peaks in the PV^{Cre} mouse are more superficial in S1 compared to M1. This mouse also displayed enriched soma-targeting in S1.
I) Cortical spatial disorganization in Reeler mouse does not alter overall subcellular targeting specificity of Sst interneurons.

differences between L1 and L2-L6. Much to our surprise, we also witnessed distinct bouton class enrichment patterns within L2-L6. These patterns were evident both across cortical layers and between cortical regions. Specifically, we witnessed layer-like strata throughout the cortex (Figure 4.5, A) and barrels in the somatosensory cortex (S1) (Figure 4.5, B). Of note, the differential bouton enrichment in L1 and within barrels was evident from the Syp-tdT channel alone, and did not require a counterstain to visualize (Figure 4.5, A and B). We therefore sought not only to characterize bouton spatial organization, but also to understand its basis.

Our culture data strongly suggests that interneuron subcellular targeting specificity is largely intrinsically determined. In contrast, these distinct laminar and regional enrichment profiles strongly suggest that these bouton classes are highly organized in a way that reflects the underlying cortical structure. Our culture experiments disrupted not only cortical structure, but also the cortical milieu, afferents, and efferents. We therefore wanted to test the extent to which cortical laminar structure alone is important for interneuron subcellular targeting. We hypothesized that disrupting

glutamatergic cell spatial organization would also disrupt cortical interneuron bouton organization, but that overall subcellular targeting proportions would be minimally affected.

As described in the introduction, the laminar organization of the cortex arises from the inside-out birth and migration pattern of glutamatergic neurons, where newly-born glutamatergic neurons migrate up the radial glial fiber, past earlier born neurons, to the cortical surface. This migration process is based on a reelin gradient established by Cajal-Retzius cells near the cortical surface: removing the reelin gradient disrupts glutamatergic neuron migration, leading to disrupted laminar organization[3, 5-15]. Specifically, upon reelin disruption, cortical layering generally follows a reversed order, with specific regions displaying high levels of intermixing and ectopic positioning[13, 16, 17]. We therefore used the well-characterized Reeler mouse, which has a recessive loss of function mutation in the *reln* gene, to disrupt cortical lamellar organization *in vivo*, while still maintaining an intact cortical milieu, afferents, and efferents[9]. We confirmed that cortical structure is indeed disrupted in the Reeler mouse via Ctip2 staining, a marker of early-born cortical L5-6 glutamatergic neurons. Figure 4.5, C and D show the CTIP2 staining pattern in control and Reeler mutant mice respectively, which shows that the glutamatergic cells of the cortex are generally inverted in the Reeler mouse.

We decided to use the Sst^{Cre} population to test this hypothesis. We did this for several reasons. First, within our three driver lines, Sst^{Cre} displayed the highest level of bouton spatial organization across the cortex. Second, Sst Martinotti cells have been shown to leave a trailing axon in the nascent L1 as they transition from tangential migration near the cortical surface to radial migration into deeper layers[8]. This process is hypothesized to result in the characteristic L1 targeting of this population. Cortical laminar organization plays a central role in these processes. Third, although laminar positioning of most MGE interneurons is independent of reelin signaling[11, 18], L2/3 Sst-

positive MCs do appear to utilize reelin for proper positioning[19]. We therefore reasoned that this population would also be the most susceptible to disruptions in cortical structure, and thus was the best candidate to assess the interplay between subcellular targeting specificity and cortical organization. We used *Sst^{Cre};reln$^{Rl/Rl}$* (Sst^{Cre} Reeler) mice as the experimental group, and used *Sst^{Cre};reln$^{wt/*}$* littermates, where ' * ' indicates any allele, as controls (Sst^{Cre} Control or simply Sst^{Cre}).

To assess these spatial differences quantitatively, we collapsed our 3D classification images into 2D projection images that show class regional average (Figure 4.5, E). We then used these averaged maps to assess class enrichment along the superficial to deep axis. We did this by measuring signal intensity along perpendiculars to both the L1 surface and L6/WM border. We normalized each measure by its total length and averaged all normalized measures to yield a statistical profile of absolute class enrichment between the surface (0) to base (1) of the cortex. Data are presented as smoothed, weighted averages from at least 50 measures per cortical region from at least 5 mice per condition (Figure 4.5, F-H). To assess differences by layer and by region simultaneously, we performed this profiling in both the primary motor (M1) and primary somatosensory barrel field (S1BF) regions of the cortex for dendrite- and soma-targeting (Figure 4.5, G and H, respectively).

As is clear from these profiles, all three driver lines showed distinct regional peaks and troughs in in targeting specificity across this superficial to deep axis in both M1 and S1BF, reflecting layer-specific enrichment of these different bouton populations. Moreover, the populations labeled by the three driver lines generally reflected taxonomic expectations, with Sst^{Cre} and PV^{Cre} displaying mutually exclusive enrichment profiles, that were both present in Nkx2.1^{Cre}. Since the cortical layers of the Reeler mouse approximate inversion, we wondered whether there would be enrich-

ment of dendrite targeting boutons at the base of the cortex instead of surface. Indeed, this disruption was reflected in the spatial positioning of dendrite-targeting boutons. Figure 4.6, E-H show the spatial organization and targeting probability in the Sst^{Cre} control and Reeler mice. As is clear from these profiles, the strong enrichment of dendrite-targeting boutons in L1 at the cortical surface of Sst^{Cre} control mice absent in the Reeler mutant. In contrast, dendrite targeting is enriched within an acellular zone (AC) near the L6/WM border. Quantification across multiple mice revealed that this dendrite targeting enrichment was a characteristic feature of Reeler mutants, confirming a partial reversal of the trends we saw in control (Figure 4.5, E-F). Interestingly, the dendrite and soma targeting classes did display distinct profiles within the cellular 'L2-6' of the Reeler cortex. Apart from the increased targeting of acellular region near the base of the cortex, however, the Reeler mouse showed very little evidence of layering *per se* along the superficial to deep axis. Instead, we witnessed a 'mound' of soma targeting boutons that gradually increased from the cortical surface that plateaued at 25% of cortical depth until tapering again at 80% depth. Thus, most laminar Sst targeting enrichment patterns are disrupted when the target cells are disorganized.

We also saw differences between cortical regions. PV^{Cre}, for example, displayed significantly more soma-targeting boutons in S1BF than in M1, with the enrichment peaking at ~35% of cortical depth, the approximate location of L4 barrels. Surprisingly, dendrite targeting boutons also showed subtlety different enrichment patterns in M1 and S1BF. Specifically, this bouton class displayed two distinct peaks at 35-45% and 60-70% of cortical depth. Remarkably, these peaks occurred approximately 10% closer to the cortical surface in S1BF than in M1, further underscoring regional differences in interneuron bouton organization. We also witnessed enriched dendrite-targeting for the Sst^{Cre} population within barrels (Figure 4.5, B). Any differential enrichment across layers and

regions witnessed in Sst^{Cre} control mice was absent in the Reeler mutant, as shown by the comparisons of M1 and S1 (Figure 4.5, G and H). Finally, we found no statistical difference between control and Reeler conditions in Sst^{Cre} interneuron subcellular targeting (Figure 4.5, E)[28], further supporting our conclusions from culture experiments that interneuron subcellular targeting specificity is largely intrinsically determined. In aggregate, these data show that cortical interneuron boutons are highly organized within the cortex, displaying differential laminar enrichment across the superficial to deep axis as well as distinct region-specific enrichment. In addition, the Reeler data confirm that cortical interneuron bouton spatial organization arises in response to the underlying lamellar organization of glutamatergic target neurons.

4.4. Unsupervised clustering of cortical interneuron boutons

Our supervised bouton classification model utilizes hundreds of measures per object to classify boutons by canonical targeting classes. We wondered, however, whether the canonical classification system was telling the whole picture. We therefore used an unsupervised clustering approach to approach bouton classification from a discovery standpoint. As with our supervised classifier, we first developed this approach using our culture system. The unsupervised machine learning algorithm and clustering approaches, as well as parameters and batch correction methods are discussed in Chapter 1 and revisited in the Discussion. We here discuss how we developed a stable, unsupervised model of bouton subcellular classification from our complete culture dataset, and the biological insights we gained from it.

[28] Note that the distribution of Sst^{Cre} subcellular targeting in slice shown here differs from that shown in Figure 4. In each case, the compared samples run were in parallel through the same iteration of classifiers. These experiments, however, were performed independently and used distinct iterations of the supervised classifiers for quantification. We therefore caution against comparing between these results directly. For now, we attribute the differences to noise (Gremlins…).

4.4.1. Developing a stable, unsupervised model of interneuron subcellular targeting in culture

Our final dataset of boutons from culture was derived from 10 independent experiments (plates) and 30 different samples (wells). We imaged these samples over 27 imaging sessions, yielding a total of 1380 tiles split across 163 scenes. In aggregate, these data yielded just under 3.1 million boutons, each of which had 431 non-redundant metrics. Given this dataset, our first task was to generate a stable, unsupervised model for bouton classification based on subcellular targeting. We used the four-layer network with 16-dimensional autoencoder embedding discussed in Chapter 1 to train the network and reduce the dimensionality of our dataset. We found that our dataset had considerable sample-to-sample variability. Figure 4.6, A shows a Uniform Manifold Approximation and Projection (UMAP) of select boutons from different samples overlaid with their sample ID. The same data is shown sample-by-sample overlaid with the supervised targeting classification in Figure 4.6, B. We were concerned that unsupervised clustering on these data as is would lead to clustering by sample, rather than by subcellular target. We therefore used Harmony[20] as a means of batch correction. This approach led to significantly higher intermixing, as is evident in Figure 4.6, C and D, which show the same samples after batch correction.

We next wanted to assess divisions within this population by clustering. As discussed in Chapter 1, we used the Louvain method of clustering, which detects communities of similar objects based on a weighted neighborhood network[21]. We selected this method for its robust detection of arbitrarily-structured communities, minimal parameter space, and ease of implementation. The Louvain algorithm takes two parameters: resolution, which sets the 'granularity' of the downstream clustering, and 'k', the number of nearest neighbors included when constructing the neighborhood network. Using a low resolution yields more basic class types, i.e., a few large clusters, whereas high granularity implies more refined classification, i.e., many smaller clusters. A higher k value

leads to smoother bounds between classes, but also increases classification error. Figure 4.6, E shows the relationship between these parameters and the number of resultant clusters for k-neighbor values of 5, 10, and 15, and resolution values of 0.1, 0.2, 0.4, 0.6, 1, 1.3, 1.6, and 2. Since some of the resultant clusters can be small and thus bias downstream analysis, we set a retention threshold of 0.5%, 3-fold lower than the 1.5% of our 'Unknown' targeting class, ignoring any populations below this cutoff. Figure 4.6, F shows the number of retained clusters as a function of varying k-neighbors and resolution. As expected, using a lower resolution led to fewer classes, whereas using a higher k value led to greater retention, as the very distinct boundaries defining very small populations became blurred, and they were in turn absorbed in larger cluster.

Before settling on any specific parameters, we wanted to know what would be driving the clustering in the first place. Specifically, we wanted to confirm that the clustering would be based on systematic differences in bouton targeting, not sample batches. We did this by assessing the basis for population divisions as we generated finer and finer clusters. We therefore assessed whether increasingly granular clustering would lead to increased batch identity or increased targeting identity. To do this, we utilized the LISI score, which measures the effective number of objects of known class represented in the local neighborhood of each object in the identified cluster. Simply put, the LISI score is a statistical representation of the number of objects that need to be sampled in the neighborhood before two are drawn from the same class. As such, the value ranges from 1 to N, the number of known classes, with 1 indicating perfect purity and N indicating perfect intermixing. We used this metric to test cluster intermixing of two parameters, namely sample identity (which would allow us to assess the extent to which batch effect was still present after correction), and supervised targeting class (which would assess the extent to which the clusters were segregating along known synaptic targeting classification).

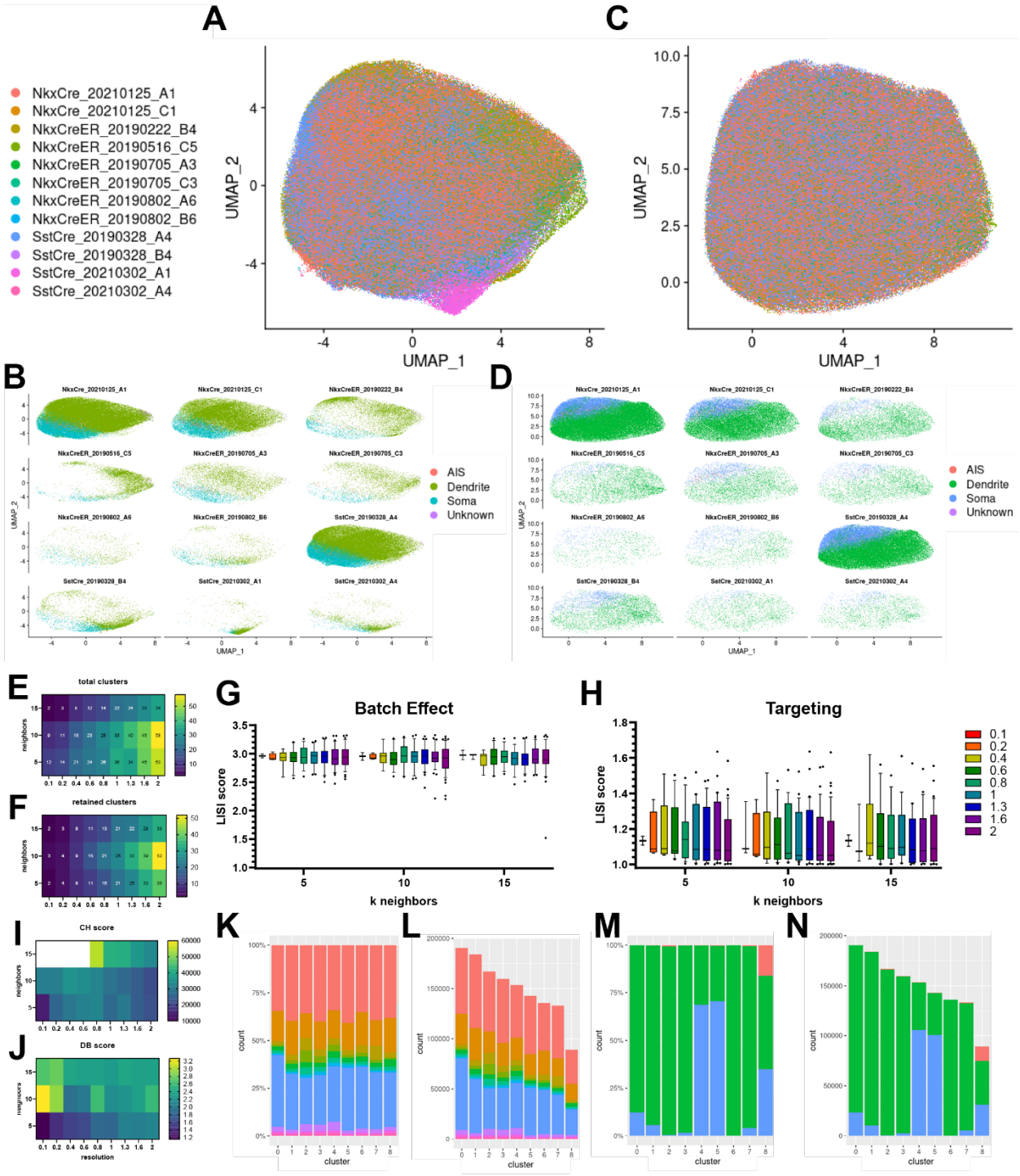

Figure 4.6 | Generation of a model for interneuron bouton subcellular targeting in dissociated culture by unsupervised classification.

A-B) UMAP showing autoencoder embedding of interneuron boutons without batch correction. In ***A***, boutons are shown in aggregate, colored by sample. In ***B***, the same UMAP is shown, but split to show positions of individual samples on the UMAP embedding. Boutons are colored by supervised class.

C-D) As in ***A-B***, but after batch correction using Harmony.

E-F) Heatmaps showing the relationship between the k-neighbors and resolution hyperparameters used in Louvain clustering. The total number of clusters (***E***) identified ithin the batch-corrected dataset using different combinations of these parameters.Note the increased number of identified clusters with higher resolution. ***F)*** Clusters with population sizes below 0.5% of the total were subsequently removed. ***F*** shows the total number of retained clusters. Note that increasing the number of k neighbors used to define the neighborhood also leads to increased retention, as distinct boundaries between populations are blurred, leading to fewer sharply-defined small populations that would otherwise be removed

G-H) LISI scores were used to assess intermixing of independent classification for the same boutons. Sample intermixing (***G***) was used to assess batch effect, whereas supervised targeting class was used to asses the impact of synaptic targeting (***H***) on the cluster delineations at each parameter combination tested for the Louvain clustering. LISI scores range from 1 (absolute purity) to N (absolute intermixing), where N is the number of possible categories. Data show combined scores for all retained clusters identified using the given parameter combination, and are presented as box and whiskers plots, with whiskers indicating the 10th-90th quantiles.

I-J) Heatmaps showing (***I***) the Calinski-Harabasz (CH) and (***J***) the Davies-Bouldin (DB) scores for all parameter combinations. A higher CH score indicates that the identified clusters are dense and well-separated, and so we wanted to maximize this metric. The DB score is used to signify 'similarity' between clusters, so we wanted to minimize this score. Importantly, we wanted to select scores that provide yielded a peak (CH) or valley (DB) relative surrounding scores when plotted: continuously ascending or descending lines provide no reason to choose one score over another, being instead indicative of issues in parameter selection or the structure of the data itself. The combination of k=10 neighbors and resolution of 0.4 was chosen based on this analysis, and resulted in nine clusters that crosswed the 0.5% cutoff.

Note that for the CH score, values for k=15 neighbors with resolution at or below 0.6 are presented as out of range. The CH scores for these conditions were extremely high compared to other conditions, thereby masking changes among the lower scores. Since these parameter combinations displayed a descending trend, and so were uninformative from a parameter selection standpoint, we have masked them to provide a clearer picture of the trends within more informative but lower-scoring conditions.

K-N) Distribution of boutons from different samples (***K, L***) and the targeting distribution (***M-N***) for the retained clusters, presented as a proportion of total (***K, M***) or absolute values (***L, N***). Note the high level of sample intermixing (***K*** and ***L***) compared to targeting enrichment (***M*** and ***N***), indicating that the clustering partitions are primarily driven by synaptic targeting rather than sample-to-sample differences.

We first wanted to rule out that clustering would be driven by batch effect. A high amount of sample intermixing would result in a high LISI score, we therefore wanted this value to be as high as possible, and to show little to no change with increasingly granular clustering. Figure 4.6, G shows the LISI scores for all retained clusters across the different conditions. The data are presented as box and whiskers plots, with whiskers indicating the 10^{th}-90^{th} quantiles. In general, we witnessed little to no change in sample-specific enrichment across these conditions within the retained clusters. In contrast, when all clusters were included, we saw decreased LISI scores (i.e., greater batch effect) for conditions with k=5 neighbors, especially with decreasing resolution values (data not shown). Although this result is at first counterintuitive (decreased resolution should lead to larger populations and so more intermixing), the most highly-affected conditions were also those with the highest proportion of clusters excluded due to our size cutoff (Figure 4.6, E and F). This result indicates that sample-based effects are likely still present, but are only a concern in very small populations. Since this effect was noticeably absent after excluding small clusters, we concluded that sample-specific signatures were of little concern within the range of conditions tested. This analysis alleviated concerns that batch effect was driving clustering.

Next we wanted to ensure that cluster segregation was occurring along the lines of synaptic specificity. We therefore used the supervised targeting data as a benchmark, expecting that ever more granular clustering would lead to further enrichment for a single subcellular target. Since we wanted to see increasing levels of population purity, we wanted the LISI scores to decrease with increasing resolution. As shown in Figure 4.6, H, this is exactly what we saw. Although the effect size is relatively small within this resolution range, there is a clear decrease in overall LISI score with increasing resolutions at all k values, indicating that clustering tended to occur in accordance with subcellular targeting.

Next, we wanted to approached parameter selection from an objective, statistical standpoint. We assumed that (1) the identified bouton clusters should comprise similar objects, and (2) that the segregation from unsupervised clustering should also segregate along the lines of supervised targeting, with each identified cluster enriched for a single subcellular target. We therefore selected resolution and k-neighbors values based on two separate analyses. For (1) we wanted to optimize two metrics that compare inter-cluster and intra-cluster dispersion, namely the Calinski-Harabasz (CH) and the Davies-Bouldin (DB) scores. The CH score describes how similar an object is to its own cluster (intra-cluster dispersion, or cohesion) compared to other clusters (inter-cluster dispersion , or separation)[22]. A higher CH score indicates that the identified clusters are dense and well-separated, and so we wanted to maximize this metric. The DB score compares the distance between clusters and the size of the clusters themselves[23]. The DB score is used to signify 'similarity' between clusters, so we wanted to minimize this score. Importantly, we wanted to select scores that yielded a peak (CH) or valley (DB) relative surrounding scores when plotted: continuously ascending or descending lines provide no reason to choose one score over another, being instead indicative of issues in parameter selection or the structure of the data itself. Figures 6, I and J show the for the CH and DB scores, respectively. As can be clearly seen, we witnessed gradual peaks in the CH score using both k=5 and k=10 nearest neighbors, with apexes at k=5/res=0.8 (CH=32700) and k=10/res=0.4 (CH=35600). Note that the values for k=15 neighbors with resolution at or below 0.6 are presented as out of range. The CH scores for these conditions were extremely high compared to other conditions, having values between 160,000 (res=0.1) and 60,000 (res=0.6) that masked more subtle changes among lower scores. Since these conditions displayed a descending trend, they were uninformative from a parameter selection standpoint. We therefore excluded these conditions from candidate selection and have hidden their values to provide a clearer picture of the

trends within more informative conditions. The DB score, in contrast, displayed very different trends. Although conditions using k=5 neighbors generally displaying very low scores, which is what we wanted, the trendline was generally increasing and therefore uninformative. Conditions using k=10 neighbors in contrast displayed a sharp drop that reached a nadir at res=0.4, then rapidly increased again by 0.8, and then was followed by another, more gradual valley between 0.8 and 2. Of note, the nadir in the DB score at k=10,res=0.4 corresponded with the peak in CH score using the same condition, indicating that these conditions minimized intra-cluster and maximized inter-cluster differences. In other words, these conditions provided us with the most compact and separated clusters among all conditions tested.

These conditions resulted in nine clusters that passed the size cutoff (we excluded 9 others, most of which had fewer than 10 members). The counts for the included and excluded clusters is presented in Table 7. Figure 4.6, K and L, respectively show the sample distribution within the nine retained clusters as a percent of total population or in absolute values. These data show strikingly similar proportions of all samples across all clusters, further alleviating concerns that the identified clusters would be heavily affected by batch effect. In contrast, we saw considerable enrichment of specific targeting classes within these clusters, as shown by the targeting distribution as a percent of cluster population and in absolute values (Figure 4.6, L and N, respectively). Of note, we witnessed target-specific enrichment patterns within distinct clusters, further confirming that our chosen parameters were distinguishing by subcellular target.

4.4.2. Unsupervised classification of cortical interneuron boutons identifies novel subclasses

Our prior analysis left us reasonably assured that our unsupervised clustering was detecting real differences in subcellular synaptic targeting specificity. We therefore wanted to investigate the differences between the identified clusters. Since we found the targeting-associated LISI scores

to be low for clusters 0-7 (Figure 4.7, A), we were confident in designating each as a distinct subclass of the supervised targeting classes. To do this, we assessed the distribution of supervised targeting classes within these clusters (Fig. 6, M and N), and assigned each as dendrite or soma targeting based on the greatest enrichment of supervised targeting class. This analysis revealed that the greatest partitioning occurred within the dendrite-targeting class, which was split into six distinct clusters, namely 0-3, 6, and 7. In contrast, the soma-targeting class was split into two subclasses, namely clusters 4 and 5. To show this partitioning, we recolored these clusters to match our standard color palette for dendrite and soma targeting used in the supervised clustering (Figure 4.7, B). This class assignment was further supported by the location of these clusters on UMAP space relative to the supervised targeting classes. Figure 4.7, C shows the supervised classes overlaid onto the UMAP as a topological map, showing distinct regions occupied by each of these targeting classes. As is clear when viewing each of the clusters in isolation (Figure 4.7, D), these subclasses generally fall within the UMAP space occupied by the assigned targeting class, with different subclasses occupying distinct regions of these territories. As noted, only cluster 8 displayed considerable intermixing of the classes designated by our supervised model. We have colored this cluster green since it is the only one that displayed significant enrichment for the AIS-targeting class (Figure 4.6, M-N) and generally occupied the same territory as this class on UMAP space (Figure 4.7, C, D). We also note that (1) our supervised classifier displayed a comparatively low true positive rate for AIS-targeting (0.697) (see Chapter 1 discussion of supervised classifier performance metrics), opening the possibility that the unsupervised model could outperform our supervised model for some AIS-targeting boutons, and (2) even utilizing different clustering parameters (i.e., tuning values for k neighbors and resolution), we were unable to find a combination that led to more than 20% enrichment for the AIS-targeting class, instead splitting it into different

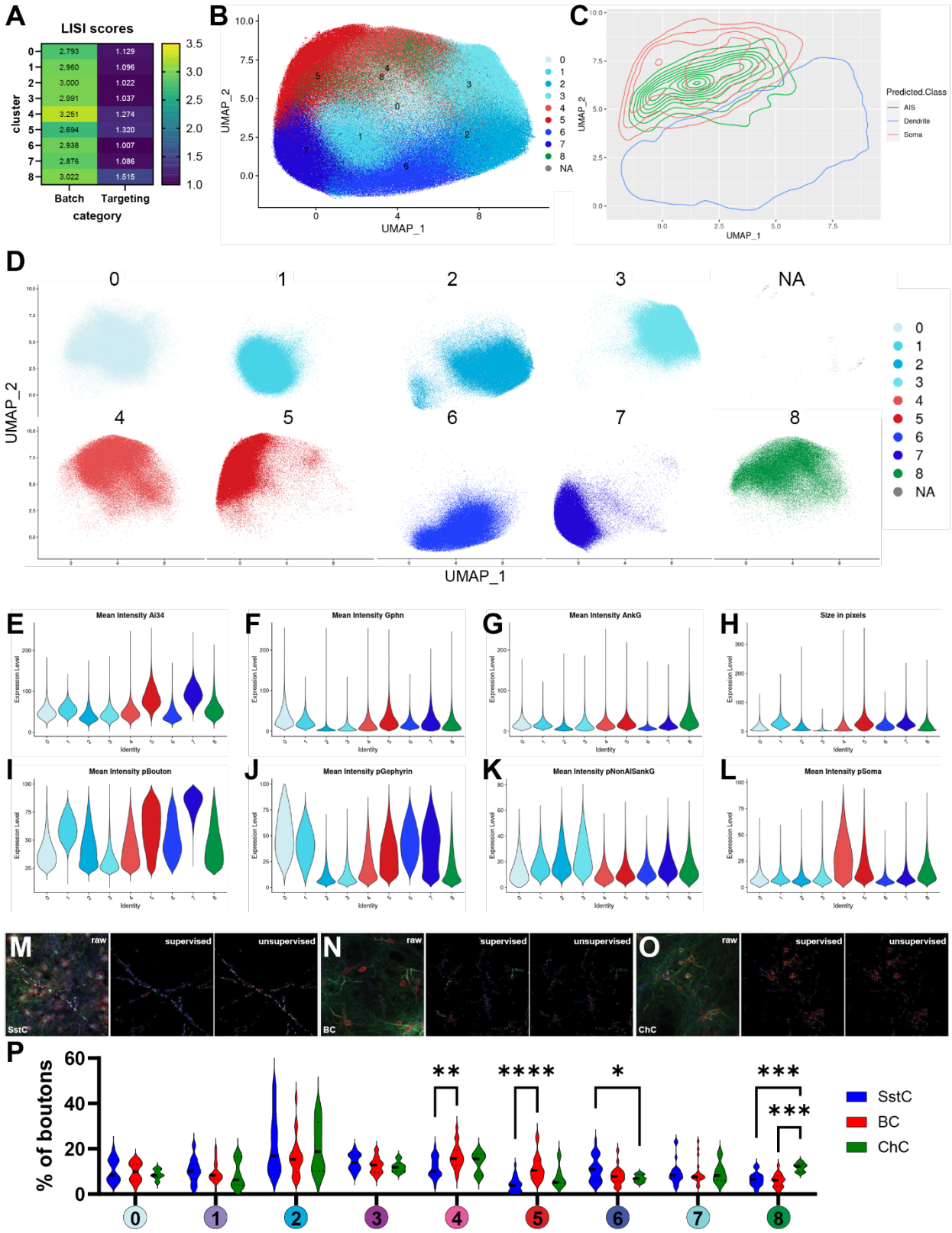

Figure 4.7 | Unsupervised classification of cortical interneuron boutons in culture identifies canonical interneuron bouton classes and reveals novel bouton subclasses.

A) LISI scores associated with batch effect (*left*) and predicted targeting (*right*) for the retained clusters identified by the selected unsupervised model. Values closer to 1 indicate less population mixing. For Batch scores, a value of 1 would indicate boutons from that cluster are derived exclusively from a single sample (i.e., presence of batch effect). For targeting, 1 would indicate all boutons in the cluster target a single subcellular region defined by the supervised classifier (i.e., dendrite, soma, AIS, or unknown). **B)** UMAP of unspervised bouton classification in culture. The retained clusters are designated numerical identifiers 0-8 based on cluster size. Cluster colors were assigned as discussed in the text. Clusters excluded due to small size were combined and are shown here as NA. **C**) Topological representation of the supervised bouton classification rendered onto UMAP. Note the two peaks within the soma-targeting class (red). **D)** Unsupervised clusters shown in isolation. Cluster colors and identifiers are presented as in B. **E-L)** Selected feature expression in the identified bouton clusters. Features have been ordered by type. On first row: mean raw signal intensity of (***E***) Syp-tdT, (***F***) gephyrin, and (***G***) Ankyrin G. At right: Morphological features, namely (***H***) object size in voxels (volume). On second row: from pixel classification, mean probablity scores of (***I***) bouton, (***J***) gephyrin psd, (***K***) non-AIS ankyrin G (i.e., neurite), and (***L***) soma. Displayed are eight of the 431 features used for clustering. **M-O)** Mapping of the unsupervised clasificationonto isolated cells in culture. In each case, images are as follows: Syp-tdT overlaid with targets (*left*), supervised classifier prediction (*middle*) and unsupervised cluster identity (*right*). Images show boutons from an isolated (***O***) SstCre interneuron, or Nkx2.1^{CreER} interneurons from (***P***) ChC-rich or (***Q***) BC-rich culture preparations. **P)** Specific unsupervised bouton subtypes are enriched within distinct cell populations in culture. Data are compiled from isolated interneurons as in ***M-O***. ** P<0.5, ** P<0.1, *** P<0.01, **** P<0.001.*

clusters. Given these caveats, we were comfortable designating cluster 8 as AIS-targeting within the clusters identified by this model.

We were surprised by the amount of subdivision within the dendrite and soma targeting classes and were curious about what differentiated these populations. We therefore looked at the differential expression of features used to train the autoencoder (Figure 4.7, E-L). We found differences within these clusters not only in raw signal intensity from the antibody staining (Figure 4.7, E-G), but also morphological features such as object size (Figure 4.7, H), and probability-based features from our pixel classifiers (Figure 4.7, I-L). Indeed, features associated with the pixel classifiers tended to show the most variability between clusters, indicating the added power of this

approach. Within the dendrite-targeting clusters, some of the major differentially-expressed features were bouton, gephyrin PSD, and Ankyrin G neurite probability (Figure 4.7, I-K, respectively). Within the identified soma-targeting subclasses, we saw differences in bouton size and bouton, gephyrin PSD, and soma probability (Figure 4.7, H-J, L, respectively) among others. We found that cluster 8, the tentatively designated AIS-targeting cluster, was the cluster most strongly associated with Ankyrin G raw signal (Figure 4.7, G), as well as AIS and AIS PSD probability (not shown), further validating our assignment of cluster 8 as the AIS targeting class. We next mapped the unsupervised clustering back onto the original images. As shown in Figure 4.7, M-O, for isolated cells, our class assignment for these clusters largely accords with the supervised targeting data and targeting assignment by visual inspection. As with supervised classification, the different clusters did not strictly map onto distinct isolated cells, though we did see distinct enrichment profiles for different cells, suggesting biological validity to the clusters. Finally, we compared cluster distribution across isolated SstCs, BCs, and ChCs, witnessing significantly different distribution profiles for each (Figures 7, P). Specifically, we witnessed enriched soma targeting (4 & 5) in BCs, AIS targeting (8) in ChCs, and dendrite targeting (6) in SstCs. This further verified that the unsupervised classification is detecting differences in subcellular targeting and validated our subclass assignments.

4.4.3. Dendrite-targeting bouton subclasses partition based on proximity to soma

Since the soma and AIS targeting bouton classes are defined by targets that are spatially defined within spatially restricted subcompartments, we expected the classes identified by the Louvain clustering method to also be physically clustered. In other words, since boutons of the same class should target the same subcellular compartment, they should also be close to one another in

physical space. We therefore sought to use physical proximity as an independent means of validating the results from our unsupervised clustering. To do this, we performed a Louvain analysis based on the x, y, z coordinates for each Autoencoder identified object in Sst^{Cre} and Nkx2-1^{Cre} bulk cultures to assess whether classes are physically clustered. We performed the Louvain analysis using k=10 nearest neighbors, then included an additional step to control for differences in relative abundance for each cluster, as discussed below. Since this analysis may be unfamiliar to some, and there are some nuances involved in how to interpret the data, we first present the results from our analysis, then use it as a heuristic to describe the analysis itself and how to read them, and finally go into depth on the results themselves.

The results from this analysis are shown in the correlation matrix in Figure 4.8, A. Here, the correlation plot is organized by class ID, which is assigned automatically based on cluster size (0 = largest), with the color designated as previously based on identified subcellular targeting. The numerical value in each cell is the spatial correlation score, which, intuitively, can be thought of as a measure of compactness/closeness of a cluster. This 'compactness' score is computed by considering the identity distribution of k-nearest neighbors in the physical space surrounding each bouton assigned to that class, and then comparing this distribution to what would be expected by random chance. This process is iterated for each bouton, arriving at a statistical score for the compactness (or spatial clustering) of the class as a whole. For row i, column j on the correlation matrix, a larger number indicates more of class i's k-nearest neighbors are coming from class j; a smaller number indicates fewer of class i's k-neighbors are coming from class j.

There are four important things to note here. First, the matrix displays a *directional* relationship and therefore must be read in a specific manner, namely row to column. Starting with row i and reading the value in column j asks how many of class i's k-nearest neighbors are coming from

class j. Doing the inverse, i.e., starting from column j and reading the value in row i, does *not* provide any information about the inverse relationship. Rather, the inverse relationship is depicted in the cell at row j, column i.

Second, as a logical extension of the last point, the diagonals along the matrix show intra-group or homotypic, relationships, i.e., whether the population is spatially clustered or self-avoidant, whereas the off-diagonals show inter-group, or heterotypic, relationships. These off-diagonally can be thought of as a 'correlation coefficient' that measures the strength of association between observations between two different bouton classes.

Third, the compactness score is a log2 transform of the ratio between the observed and expected (i.e., null) interaction. Therefore, a value of 0 indicates that the true identities of the k-nearest neighbors are no different from any k-random neighbors. Both positive and negative deviations from 0 are meaningful, indicating non-random spatial relationships within and/or between populations: values less than 0 indicate the objects are spatially anti-correlated, with these populations inhabiting distinct spatial regions; values greater than 0 indicate non-random spatial clustering. The greater the magnitude of deviation from 0, the less random the observation. Moreover, since the compactness/closeness score is a logarithmic transform, the scale should not be interpreted linearly: a value of ±1, for instance, is 2-fold deviation from expected, whereas ±2 is a 4-fold deviation. The value of -3.64, the greatest magnitude in our dataset, is 12.5-fold more anti-clustered than expected by random chance.

Fourth, we are calculating the ratio of observed to expected as a means of controlling for differences in class sizes. Inherently, it would be more likely for any given bouton's k-nearest-neighbors to contain boutons from more populous classes than from smaller ones. Therefore, we are using the probability of a random encounter to control for this effect. This difference in class

sizes explains why the matrix is not symmetrical. An important mathematical consequence of this normalization method is that it can overrepresent clustering in smaller classes. Simply put, with progressively smaller classes, the probability that they would have a neighbor belonging to another small class becomes exceedingly small. As we are normalizing by an ever-smaller denominator, the effect of such an encounter on the compactness score becomes extremely high. We have filtered out exceedingly small populations to avoid such overrepresentation, presenting the counts for the included classes in Table 7. Even so, the mathematical nuances of this score are important to keep in mind while interpreting the results.

As described above, we wanted to use physical clustering as an independent method to validate our statistical clustering method. Specifically, we anticipated that boutons of the same type will be spatially clustered. Such homotypic spatial clustering is identifiable based on intra-group compactness score along the diagonal of the correlation matrix (Figure 4.8, A). Specifically, we expected the diagonals to be greater than 0. This is exactly what we saw (Figure 4.8, A). As indicated by the values along the diagonal, all identified bouton clusters displayed within-group correlation scores greater than 0, indicating they are non-randomly distributed and spatially clustered within the culture dish, as expected. Specifically, the AIS-targeting (cluster 8) displayed the highest intra-group score of +2.75 (6.74 more clustered than random) as might be expected given the highly restricted area that defines the AIS. Although this was the smallest class that we included, at 6.6% of the total population it is unlikely that the score is artefactually inflated to this value due to population size. The next highest intra-group correlations were from the two soma-targeting classes (4 and 5), with class 5 having a score of +2.28 (4.07-fold) displaying slightly higher clustering levels than class 4 with +1.77 (3.41-fold). These data confirmed our suspicions that soma-

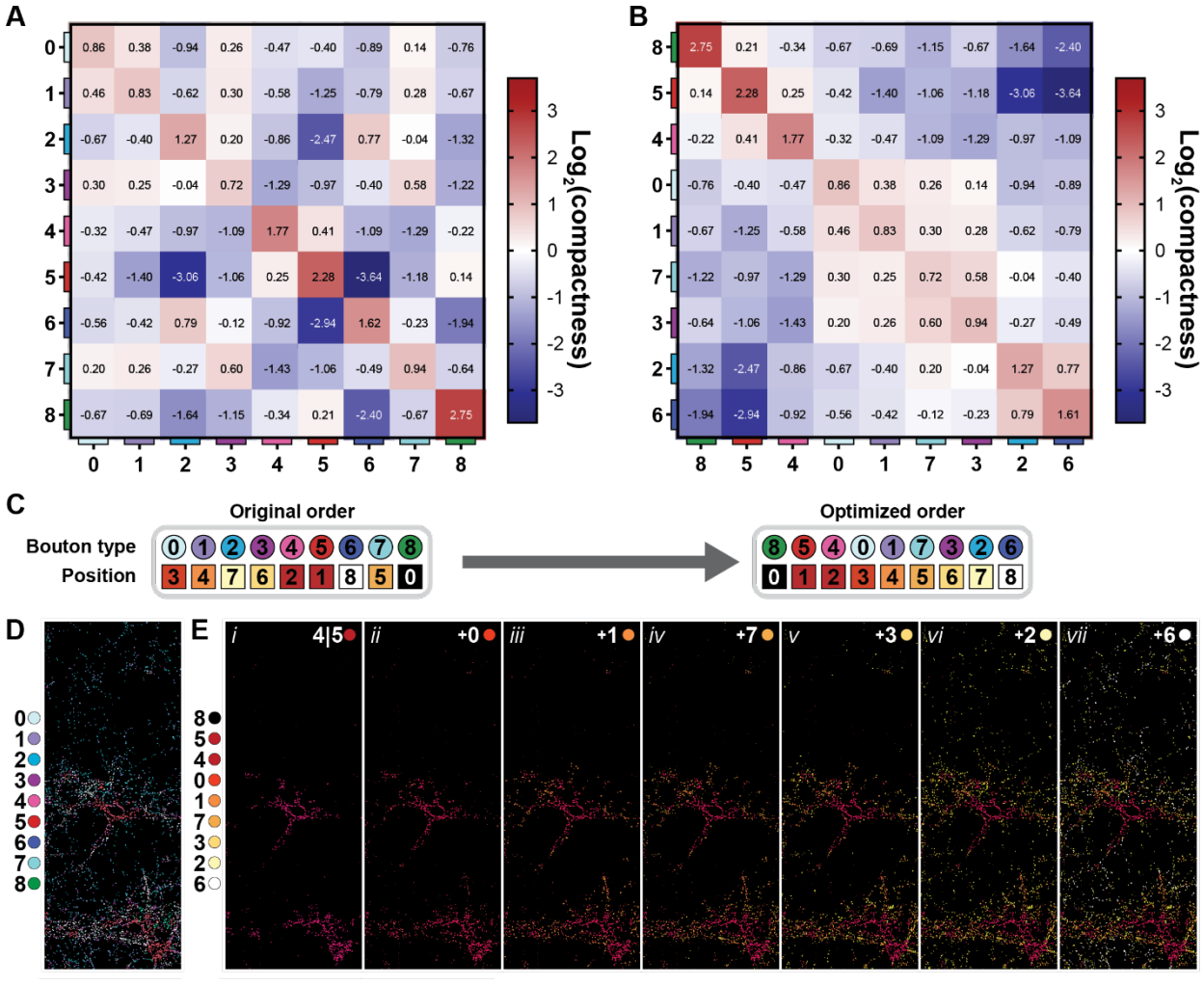

Figure 4.8 | Unsupervised dendrite-targeting bouton subtypes in culture follow a proximal-to-distal spatial organization relative to the target soma.

A) Log_2 Compactness score matrix based on spatial nearest neighbor analysis. '0' indicates spatial relationship expected by random chance. Positive values indicate higher spatial clustering, whereas negative values indicate spatial avoidance. **B)** As in *A*, but reordered to maximize sub- and super-diagonal values, with soma-targeting subpopulations (types 4 and 5) set as an anchor. The relationships suggest the identified dendrite-targeting subpopulations are ordered sequentially at increasing distances from the soma. **C)** *Left*: Original order (as in *A*) and *Right*: compactness-optimized order (as in *B*) of bouton subtypes. The position in the sequence is coded as a heatmap. **D)** Example ROI overlaid with unsupervised bouton subpopulations with original color palette. **E)** ROI from *D*, but with boutons colored by sequence position as in *C*. (*i*) shows soma-targeting boutons only, with no distinction between types 4 and 5. (*ii-vii*) show sequential addition of dendritic subtypes based on the sequence in *B*. The boutons are recolored as a heatmap to emphasize sequentiality. Note the increasingly proximal-to-distal spatial organiza-

tion of the dendrite-targeting bouton subtypes. **F)** UMAP of unsupervised clustering with subtypes outlined by subcellular target compartment suggested by *B* and *E*. **G)** Schematic showing the target cell subcompartments revealed through the unsupervised analysis.

targeting boutons would be spatially clustered. Unexpectedly, we also witnessed high scores for two dendrite-targeting classes, namely classes 6 with +1.61 (3.07-fold) and class 2 with +1.27 (2.41-fold), indicating that these dendrite-targeting bouton classes highly spatially clustered. All other dendrite-targeting bouton classes displayed non-random intra-group clustering, with the lowest scores being for class 1 with 0.83 (1.78-fold) and class 3 with 0.72 (1.65-fold). Although we expected the dendrite-targeting classes to be less clustered, the extent of physical clustering observed was particularly surprising given our labeling method for this compartment: unlike the AIS or soma, which are large, extended subcellular compartments, gephyrin PSDs are themselves highly distributed, punctate structures, so evidence that these are themselves clustered indicates a level of self-organization within the culture dish that was quite unexpected.

Encouraged by these findings, we became intrigued by the non-random signatures from the inter-class (heterotypic) relationships on the off-diagonals. We witnessed both correlated [e.g., 2→6, +0.77 (1.71-fold); 6→2, +0.79 (1.73-fold)] anticorrelated [e.g., 5→2, -3.06 (8.33-fold); 8→6, -2.40 (5.28-fold)] classes. Indeed, the strongest relationship in the entire dataset was mutual avoidance between class 5 soma-targeting and class 6 dendrite-targeting boutons. The 5→6 score of -3.64, for instance, indicates a 12.5-fold enrichment over random chance. Given that these relationships are founded on subcellular targeting of distinct compartments, we next asked whether there is a natural, optimized order for these bouton classes. To do this, we set the AIS-targeting class as an anchor, and iteratively identified the next class in the sequence by maximizing the sum

of the sub- and super-diagonals for every remaining permutation of inter-class relationship[29]. Using this iterative process, we arrived at 8, 5, 4, 0, 1, 7, 3, 2, 6 (Figure 4.8, C) as the most representative permutation for the sequential, spatial ordering of the identified bouton subclasses. Of note, this consensus sequence generally followed the pattern of AIS→Soma→Dendrite.

Having arrived at this sequence, we re-ordered our correlation matrix hoping this re-organization would reveal additional intrinsic spatial relationships between the identified bouton classes (Figure 4.8 B). This revealed three distinct, spatially correlated nodes within the dendrite-targeting class consisting of 0 & 1, 7 & 3, and 2 & 6, which displayed progressively stronger anti-correlation with the AIS- and soma- targeting classes. This relationship suggests that this sequence reflects an increasing distance from soma. We verified this visually by recoloring our original bouton classes (Figure 4.8, C and D) as a heat map based on this sequence (Figure 4.8, C), and mapping the re-colored heatmap back on to Sst^{Cre}; Ai34 bulk culture (Figure 4.8, E). This mapping shows that the identified dendrite-targeting classes are arranged at increasing distances from target cell soma. In addition, we noted that the class most enriched in individually labeled dendrite-targeting Sst^{Cre} cells was type 6, the most distal of the dendrite-targeting subclasses (Figure 4.7, P). This spatial analysis led us to conclude that unsupervised classification identified distinct dendritic inhibitory boutons arranged in a proximal-to-distal spatial orientation in relation to the target cell soma (Figure 4.8, F and G). This spatial organization is partially reflected in the UMAP, with the most distal node (2 & 6) occupying a region at the opposite pole from the soma (5 & 4), and the remaining dendrite-targeting classes occupying more medial positions (Figure 4.8, F). A schematic for the

[29] The sub- and super-diagonals are the cells below and above two adjacent diagonals. Visually, the sub- and super-diagonals are encompassed by drawing a square around two adjacent diagonals. For example, the sub-diagonal of cell 0,0 is cell 1,0 (C=+0.46); the super-diagonal of cell 1,1 is cell 0,1 (C=+0.38). Maximizing the sum of these values also maximizes inter-group spatial clustering. That is, it ensures that adjacent diagonals are also maximally spatially correlated. Maximizing these values across the correlation matrix yields a matrix with an optimized sequence of spatially correlated bouton classes.

proposed positioning of the dendrite-targeting boutons is shown in Figure 4.8, G. We conclude that the spatial relationships between the identified subclasses represents a principle of spatial self-organization of interneuron boutons that arises within a culture dish.

4.4.4. Unsupervised soma-targeting boutons subtypes are associated with distinct target cell profiles

We next wanted to investigate the partitioning of the soma-targeting subtypes (classes 4 & 5). PV basket cells are a morphologically and molecularly heterogenous population[2, 24]. We therefore asked whether this partitioning reflected distinct presynaptic basket cell types by comparing the distribution of classes 4 and 5 boutons from individual basket cells in culture (Figure 4.9, A). Although we witnessed basket cells that preferred one type or the other, no individual basket cell displayed exclusively one bouton type, suggesting the partitioning between the soma-targeting bouton classes is not a marker of presynaptic cell identity. We therefore looked to the postsynaptic cell. Our spatial analysis indicated that both soma-targeting classes displayed strong homotypic compactness scores, with weaker, but still positive, heterotypic relationships. We postulated two models to explain these data (Figure 4.9, B): 1) a soma polarization model, where distinct subtypes are associated with distinct subcellular compartments, e.g., one subtype closer to AIS and the other closer to dendrite, and 2) a somatic heterogeneity model, where distinct soma-targeting bouton subtypes are enriched on distinct soma, with some overlap. If the data reflected the soma polarization model, we expected to see differences between the two subtypes in the mean

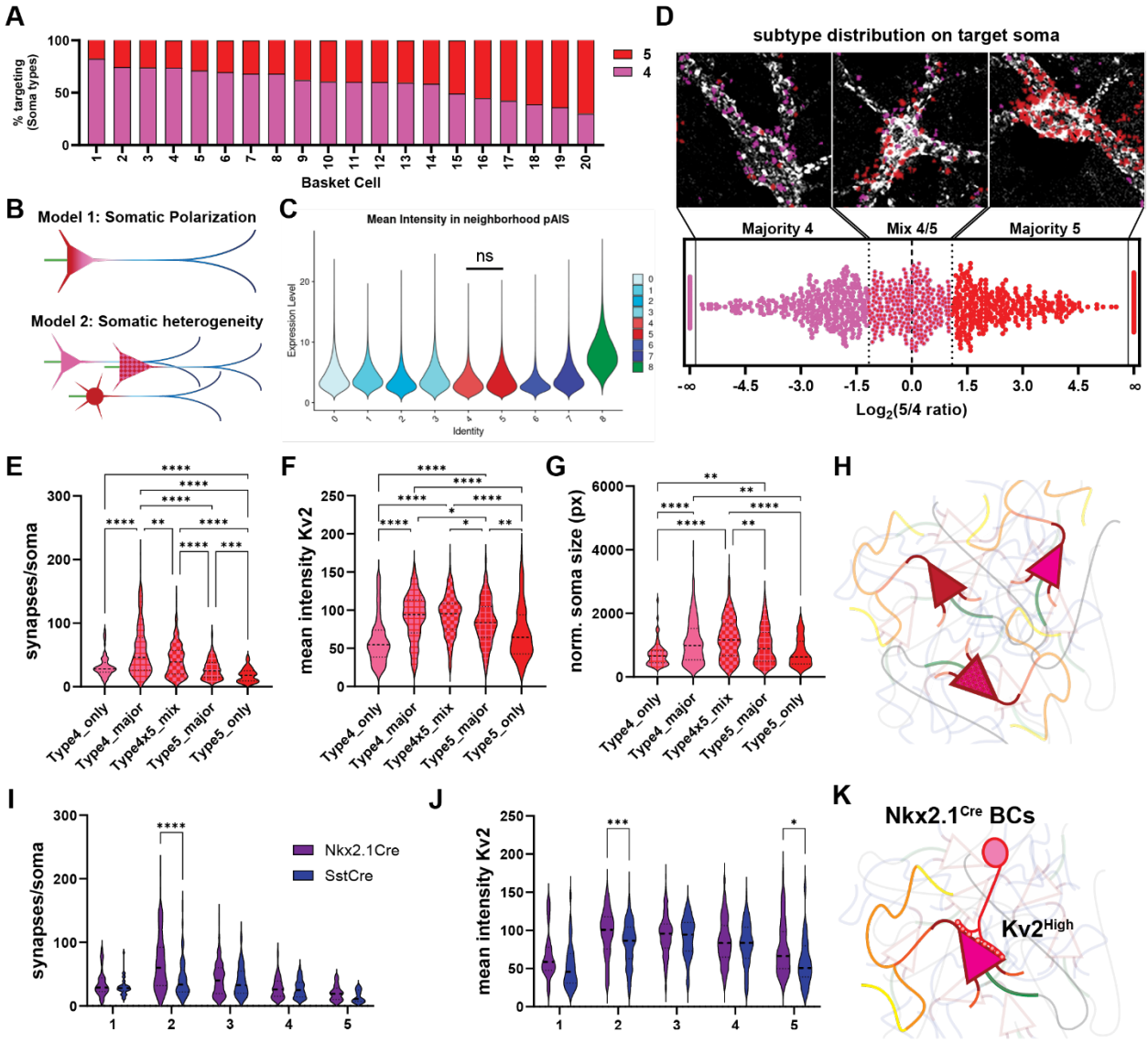

Figure 4.9 | Unsupervised soma-targeting bouton subtypes are associated with distinct target cell profiles: evidence of preferential cellular targeting in culture.

A) Distribution of class 4 and 5 soma targeting boutons across 20 individual basket cells from Nkx2.1^{CreER} cultures showing class preference but not exclusivity in some cells. **B)** schematics for somatic polarization and somatic heterogeneity models to that could explain the inter- and intra-type compactness scores for the two soma-targeting bouton subtypes. See text for details. **C)** Mean probability of AIS within a 10 voxel radial neighborhood of the bouton. Contrary to the the polarization model, there was no statistical difference between types 4 and 5. Other bouton subtypes are shown for reference. **D)** *Upper*: example images showing target cell soma with majority subtype 4, a mix of 4 and 5, or majority 5. Kv2 staining is shown to demarcate the target cell soma. *Lower*: Distribution of soma-targeting subtypes on target soma. Data shows Log2 ratio of sybtype 5 to subtype 4 boutons per soma, with data pooloed from Sst^{Cre} and Nkx2.1^{Cre} bulk cultures. Soma were categorized as shown for downstream analysis. Soma at ±∞ have exclusively type 4 or 5

boutons. Dotted lines at ±1.2 indicate the thresholds for majority designation, an OR of 2.3, or appoximately 70% one type over the other. **E-G)** Soma categorized as in D were assessed for total synapses/soma (***E***), mean Kv2 intensity/soma (***F***) and soma size (***G***). **H)** Schematic of the target regions and soma types identified through this analysis. ***I-J)*** as in ***E*** and ***F***, respectively, but with comparison between SstCre and Nkx2.1^{Cre} driver lines. Note the consistent difference in the majority type 4 population between SstCre and Nkx2.1^{Cre}. The Nkx2.1^{Cre} line labels both SstCs and BCs, suggesting the difference is due to the BC population not labeled by Sst^{Cre}. **K)** Schematic of the inferred BC population forming a high number of type 4 boutons onto target cells with high Kv2 levels. * *P<0.5,* ** *P<0.1,* *** *P<0.01,* **** *P<0.001.*

intensity of AIS in the neighborhood. As shown (Figure 4.9, C), there was no statistical difference between soma-targeting subtypes for this metric. We also found no difference in mean intensity of nuclei in the neighborhood (data not shown), suggesting the distinction is not based on soma vs proximal dendrite targeting. We therefore rejected the polarization model. Next we tested the somatic heterogeneity model. Given this model, we expected to see different levels of enrichment for types 4 and 5 on different target cell soma. We therefore manually segmented ~1000 soma based on their Kv2 staining and assessed subtype enrichment on a soma-by-soma basis by calculating the subtype enrichment score via a log2 transform of the 5/4 ratio (Figure 4.9, D). To ensure that our analysis was sufficiently powered for each individual soma, we set a minimum inclusion threshold of 5 boutons/soma. The mean number of boutons per soma was 39.5. Graphing the enrichment scores of all soma revealed five peaks corresponding to five different soma categories: exclusively 4 (-∞), majority 4 (-6 – -1.2), mix 4/5 (-1.2 – +1.2), majority 5 (+1.2 – +6), and exclusively 5 (+∞). Based on these data, and the observation that individual basket cells display both types 4 and 5 soma-targeting boutons (Figure 4.9, A), we posited that the distinction between somatic subtypes 4 and 5 reflects, in part, a property of the postsynaptic cell. We therefore assessed the mean Kv2 intensity, soma size, and number of boutons/soma across these populations, and found significant differences across all of these parameters (Figures 9, E-G), indicating that these soma populations are indeed distinct and lending further support the somatic heterogeneity model

(Figure 4.9, H). Finally, since Nx2.1^{Cre} and Sst^{Cre} label distinct interneuron populations, we asked whether the soma targeted by these driver lines were also distinct. We found that the majority 4 soma class displayed increased number of synapses per soma (Figure 4.9, I) and increased mean Kv2 intensity (Figure 4.9, J) in Nx2.1^{Cre} culture compared to Sst^{Cre}. Since Sst^{Cre} labels only dendrite-targeting somatostatin cells, whereas Nkx2.1^{Cre} labels both somatostatin cells and soma-targeting basket cells, we concluded that these differences must be accounted for by basket cells displaying preferential cellular targeting in dissociated culture (Figure 4.9, K).

4.4.5. Basket and somatostatin cells display distinct axonal growth strategies in dissociated culture

We noticed from our individually-labeled cultures that BCs and SstCs have distinct morphologies, with BCs forming highly convoluted, tight networks of connections, and SstCs forming straight neurites with more dispersed boutons over larger areas (Figure 4.10, A). Given these morphological differences, we hypothesized that these cell types utilize different axonal growth strategies to achieve distinct forms of inhibitory coverage, and that these differences would be reflected by the bouton subtypes along an axonal branch. Specifically, we reasoned that we could use the bouton sequence from our spatial compactness analysis to assess how an axon utilizes the cellular environment during growth and synapse formation: along an axon, smaller changes in sequence from bouton to bouton would reflect a stronger adherence to the underlying cellular topography, whereas large changes would reflect more abrupt jumps from one subcompartment to another. We formalized this analysis as the *sequentiality score*, which is displayed with a schematic for calculation in Figure 4.10, B. Given the spatial sequence uncovered from our compactness analysis, we can directly transform a bouton of class β into its corresponding position in the sequence, σ. This

transformation in turn allows us to calculate the sequentiality score for two adjacent boutons on an axonal branch as follows:

$$S = 1 - \frac{|\sigma_i - \sigma_{i+1}|}{\sigma_{max} - \sigma_{min}}$$

where σ_i is the position in the sequence corresponding to a bouton of class β at index i on the axon, and σ_{i+1} is that of the next bouton. The absolute value of the difference, $|\Delta\sigma|$, then, provides a metric for the magnitude of jump along the sequence. σ_{max} and σ_{min} are, respectively, the maximum and minimum sequence values within the full set of boutons analyzed. We subtract from 1 so that smaller jumps in the sequence yield higher sequentiality scores. The average sequentiality score per branch, $\bar{S}_{branch}$, is then calculated as:

$$\bar{S}_{branch} = 1 - \frac{\sum_{i=0}^{n} |\sigma_i - \sigma_{i+1}|}{(n-1)(\sigma_{max} - \sigma_{min})}$$

Where n is the total number of boutons on the branch. S and $\bar{S}_{branch}$ are unitless metrics that can be compared across different bouton sets and range from 1 (no change in identity across the entire branch) to 0 (the greatest possible jumps from bouton to bouton). A value of 0.5 would correspond to random order given an equal distribution of all bouton types. We next wanted to establish biologically-relevant standards for comparing our observed $\bar{S}_{branch}$ values. Since $\bar{S}_{branch}$ is a function of identities and order of boutons on an axonal branch, we calculated the expected value for each branch, $\bar{S}_{expected}$, by randomizing the bouton order across 10,000 iterations. $\bar{S}_{expected}$ serves as an internal control for the actual value of each branch, $\bar{S}_{actual}$, by defining random sequentiality: values greater than $\bar{S}_{expected}$ indicate stronger adherence to the underlying cellular topography than expected by random chance, whereas smaller values indicate more sub-compartment jumps.

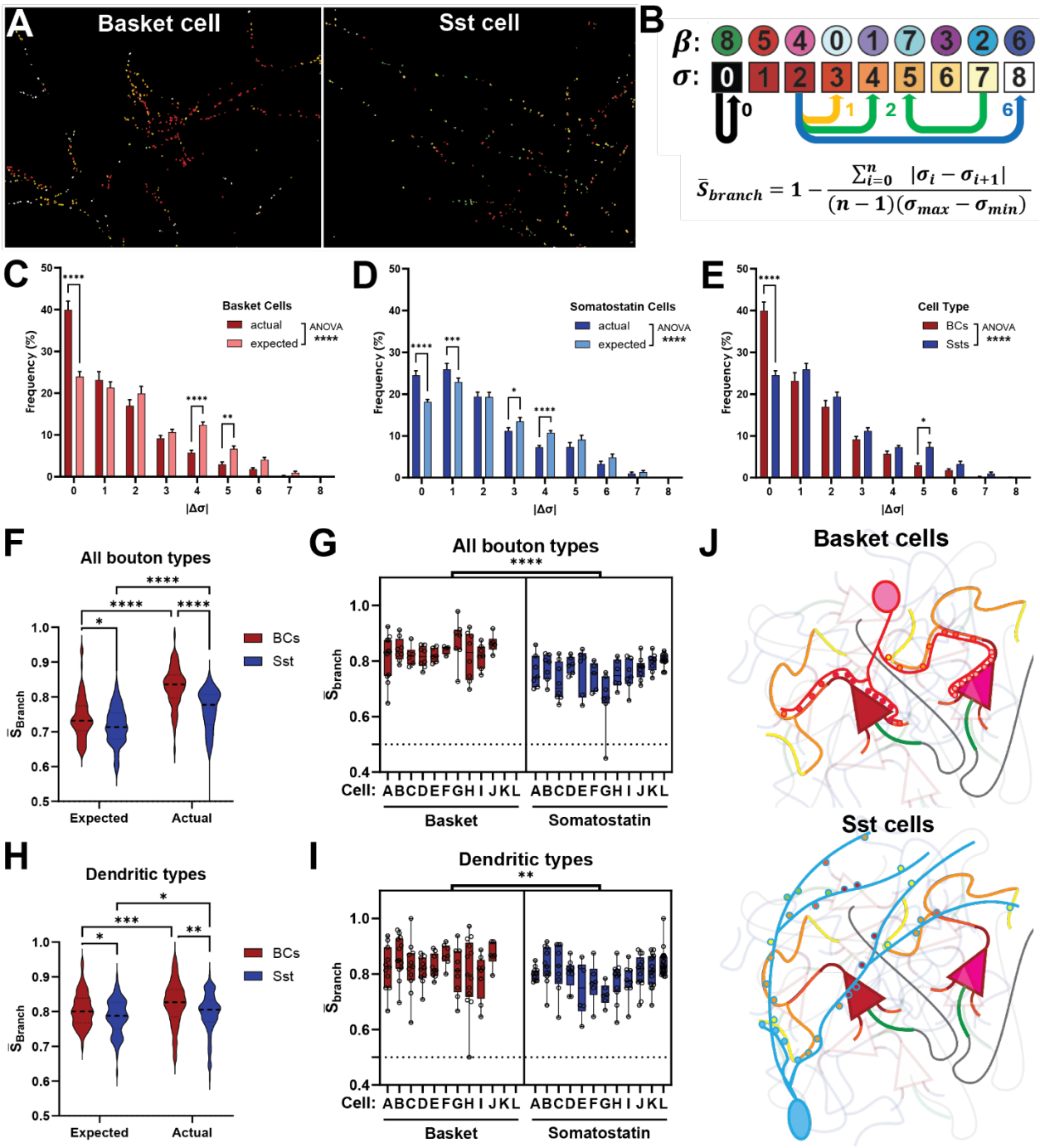

Figure 4.10 | Bouton subtype sequentiality reveals distinct targeting strategies for Basket and Sst cells in culture.

A) Representative morphologies and bouton sequence values for an individual BC (*left*) and SstC (*right*) in dissociated culture. **B)** Sequentiality measurement schematic. **C-E)** Frequency distribution of sequentiality scores for all boutons from isolated BCs (***C, E***) and SstCs (***D, E***) in culture. Actual distributions are compared to expected for that cell type (***C, D***), or between cell types (***E***). **F-I)** average sequentiality scores per axonal branch for all bouton types (***F, G***) or for dendritic types only (***H, I***). Data are shown cumulatively (***F, H***) or on a cell-by-cell basis (***G, I***) Cumulative

data is also compared to the expected cumulative distribution. **J)** Schematic for axonal growth and synaptic targeting of basket cells (*top*) and Sst cells (*bottom*) in culture based on sequentiality data. Basket cells form many synapses along the target compartment of a single cell. Sst cells form synapses across many target cells. * *P<0.5, ** P<0.1, *** P<0.01, **** P<0.001*.

We first compared the actual distribution of sequentiality scores, $|\Delta\boldsymbol{\sigma}|$**,** to the expected distribution from our randomized simulations for individual BCs (n=10) from Nkx2.1^{CreER} BC preps and SstCs (n=12) from Sst^{Cre} cells diluted in unlabeled cortical preps. Both cell types displayed distributions with significantly smaller jumps than expected (Figure 4.10, C and D, respectively), indicating the targeting sequence is non-random. Moreover, the distribution for BCs was shifted towards smaller jumps than SstCs (Figure 4.10, E). In line with these data, we found that the actual sequentiality per branch ($\bar{S}_{actual}$) for both BCs and SstCs was significantly higher than expected ($\bar{S}_{expected}$), with BCs further displaying higher $\bar{S}_{actual}$ values than SstCs (Figure 4.10, F). This trend was witnessed across multiple individual cells, with relatively consistent values between branches (Figure 4.10, G). We wanted to ensure that this difference between and SstCs was not driven solely by soma-targeting boutons, and therefore ran the same analyses on branches comprised only of dendritic subtypes. We witnessed the same non-random sequentiality for BCs and SstCs, with BCs again displaying higher values than SstCs (Figure 4.10, H and I). Based on this analysis, we conclude that both BCs and SstCs display non-random axonal growth patterns in culture, and further, that BCs adhere more closely to underlying cellular topography than SstCs. Axon guidance can proceed through two distinct molecular mechanisms: 1) via surface-bound (haptotactic) cues, such as subcellular gradients of adhesion molecules, or 2) via soluble (chemotactic) cues, such as secreted molecules that form gradients by diffusing throughout the environment. These sequentiality data indicate that Basket cells axons tend to follow subcellular gradients more closely than Somatostatin cells, suggesting the former are more reliant on haptotactic cues,

and the latter chemotactic cues, for axon guidance. We summarize these results schematically in Figure 4.10, J.

4.4.6. Unsupervised classification in slice identifies bouton classes with distinct laminar distributions

Given the findings from our unsupervised classification in culture, and the laminar enrichment differences we witnessed in cortical columns using our supervised classifier, we next wondered whether an unsupervised approach might differentially identify bouton subclasses based on spatial distribution in cortical slices.

We ran into major roadblocks in data processing given the number of objects assessed. To put this into concrete terms, our complete culture dataset contained just under 3.1 million boutons from 1435 tiles, our complete slice dataset from 40 slices was comprised of just over 800 tiles and contained 42.1 million boutons. To put this into perspective, with the tabular input bouton data consisting of approximately 1 GB per tile, to run all of these boutons through our unsupervised classifier, which currently processes the data in memory, would require over 800 GB of RAM, resources we simply do not have. At the time of writing, we are in the process of adapting our unsupervised pipeline to store data to disk, thereby decreasing RAM requirements at the expense of longer calculation times. As such, the following data was not generated using a stable model, and we stress that these results are preliminary. Nevertheless, the experiments presented here are not only a natural extension of the other experiments within this chapter, but the results from this preliminary study also accord with findings discussed earlier in this chapter. We therefore include them with the disclaimer that they are preliminary and may not hold given a more comprehensive and stable model.

Our preliminary test set included two cortical columns from two different cortical slices of a single mouse. The four two scenes were comprised of 6 tiles total, and yielded 413,500 boutons (average 69,000 boutons per tile). Our first question was whether the unsupervised autoencoder-based learning method would be able to differentiate between the canonical targeting classes from our supervised model. As previously, we tested this by assessing the distribution of probability scores and predicted targeting class on a UMAP rendering of the embedded autoencoder data. As expected, targeting probability scores (Figure 4.11, A-C) and classification (Figure 4.11, D) were by in large relegated to distinct regions of the UMAP space, indicating our autoencoder model was able to identify distinct differences between these populations even in this new sample format. We then performed unsupervised clustering, empirically selecting parameters that distinguished the three canonical bouton classes. This approach led to the identification of 16 distinct subclasses. Soma-targeting boutons were almost exclusively contained within a single population (cluster 4), with the remainder distributed within clusters 1, 2, and 8. The AIS-targeting population was found exclusively in a single cluster (cluster 2), but as with our unsupervised culture experiments, the identified population also contained a high proportion non-AIS targeting (mainly dendrite, some soma), indicating the unsupervised classifier may be encountering similar issues for this targeting class in both contexts. The remaining clusters subdivided the dendrite-targeting class.

Given how compositionally distinct L1 is from other cortical layers, we wondered whether the unsupervised clustering would identify dendrite-targeting boutons in this layer as distinct from other dendrite-targeting boutons in the cortex. We therefore mapped the bouton's depth (i.e., its distance in pixels from the cortical surface) onto the UMAP. Remarkably, this analysis cleanly divided the UMAP into three distinct populations (Figure 4.11, F), which we confirmed to be L1, L2-6, and subcortical white matter (WM) by mapping the bouton cluster identity back onto the

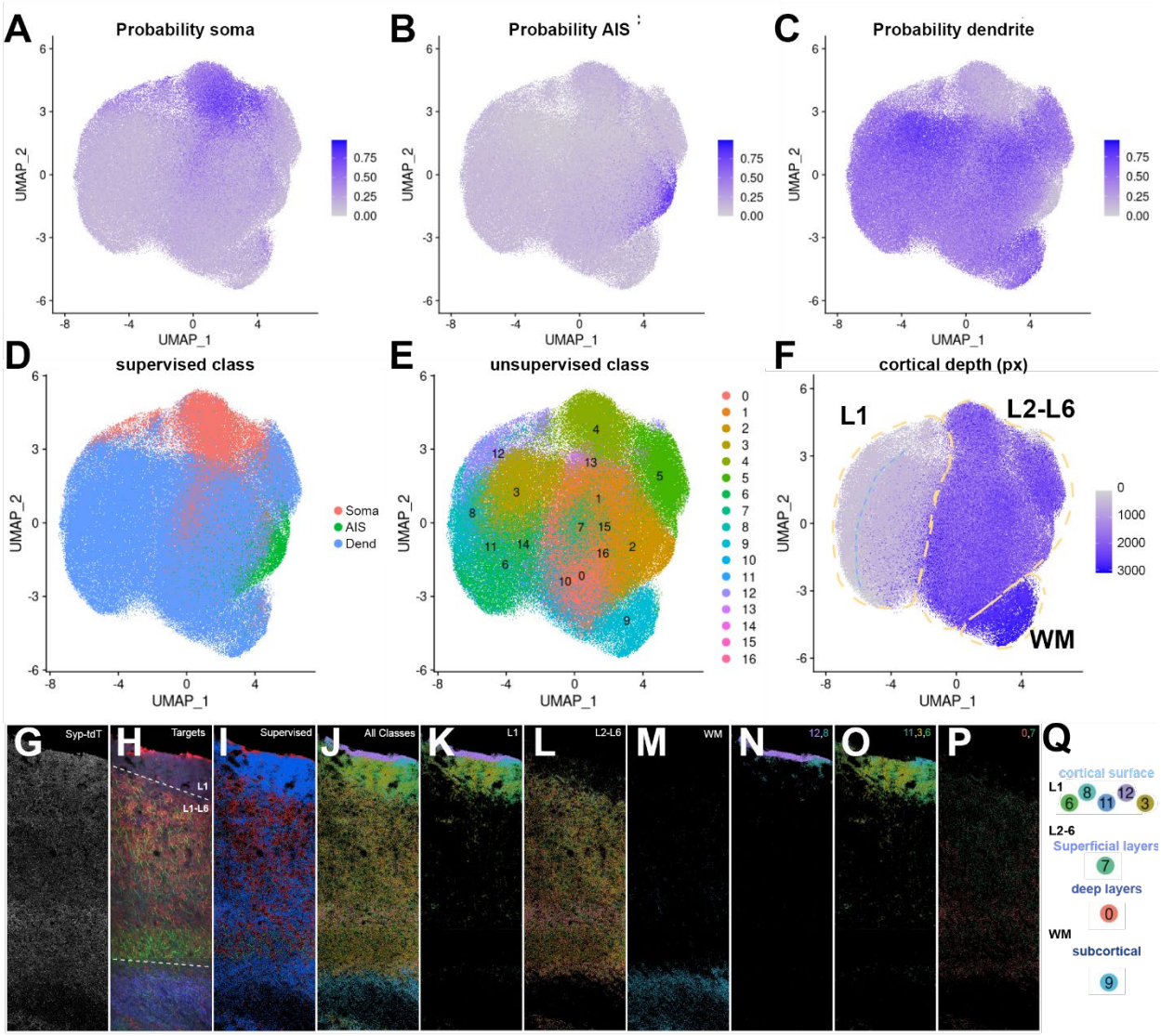

Figure 4.11 | Unsupervised classification of cortical interneuron boutons in slice identifies bouton subclasses with distinct laminar distributions.

Data show preliminary results from an application of our unsupervised learning algorithm to a cortical column from an Sst-Cre mouse.

A-F) Autoencoder embedding of slice bouton data rendered onto UMAP space, overlaid with the bouton's supervised classifier soma-targeting probability score (***A***), AIS-targeting probability score (***B***), dendrite-targeting probability score (***C***), supervised class (***D***), unsupervised class (***E***), and cortical depth (i.e., distance from cortical surface) in pixels (***F***). Approximate boundaries between populations in L1, L2-L6, and subcortical white matter (WM) are outlined in yellow, and the boundary between identified, "superficial" (8 & 12) and "deep" L1 classes is indicated by purple and aqua dashed line.

G-P) Images of an Sst-Cre cortical column run through the unsupervised classifier. The Syp-tdT (***G***) and target (***H***) input signals a representative cortical column. For the targets, Kv2 is in red, Ankyrin G in green, and gephyrin in blue. Approximate boundaries between L1, L2-L6 and WM

are shown on the target image for reference. Mapping of supervised (***I***), and unsupervised (***J***) class for all labeled boutons in the column. Unsupervised classes shown in K-P were selected based on asymmetric laminar distributions, specifically those enriched in L1 (***K***), L2-6 (***L***), WM (***M***), superficial L1 (***N***), deep L1 (***O***), and superficial or deep L2-L6 (***P***).
Q. A schematic summarizing the bouton identified classes with distinct laminar distributions.

original image (Figure 4.11, G-M). This mapping not only confirmed that specific clusters were located in L1 (3, 6, 8, 11, 12, and 14), L2-6 (1, 2, 4, 5, 7, 10, 13, 15, 16), or WM (9) (Figure 4.11, K, L, and M, respectively), but also revealed the L1 population itself was divided into 'superficial' (8, 12) and 'deep' (3, 6, 11, 14) subpopulations (Figure 4.11, N and O, respectively), and that within L2-6, cluster 7 tended to be in superficial, and cluster 0 in deep layers (Figure 4.11, P). These data are summarized schematically in Figure 4.11, Q. We added a line to the UMAP in Figure 4.11, F at the border between the 'superficial' and 'Deep' L1 populations to further emphasize the spatial relationships indirectly encoded by the autoencoder and detected by the clustering algorithm. These results are particularly remarkable since we specifically excluded any spatial positioning information as inputs for the autoencoder, indicating that the unsupervised model was able to differentiate these populations indirectly from morphological and intensity metrics alone. Moreover, we saw depth-based gradations in both dimensions of the UMAP, suggesting that a high degree of the differences identified between these bouton subclasses are associated with their positioning along the superficial to deep axis.

4.5. Conclusions

In this chapter we present our biological findings on cortical interneuron synapses using an image-based multidimensional analysis of their synapses. Specifically, we show how we used the tool presented in Chapter 1 to test outstanding questions in cortical interneuron synaptic biology. To summarize, we first showed that cortical interneuron subcellular targeting can be recapitulated

in dissociated culture. Next, we showed that synaptic boutons are highly organized within the cortex, and that this spatial organization is based on the positioning of their target cells. We then utilized our unsupervised classifier to reveal previously unrecognized heterogeneity within the canonical bouton targeting classes. Next, we found that these newly identified classes are spatially organized within the culture dish. Then, we utilized this spatial organization to discover that SstCs and BCs enact distinct forms of inhibitory coverage. Finally, we applied the unsupervised clustering technique to a cortical column, which in turn revealed a remarkable level of bouton spatial organization within the cortex. We provide a more in-depth discussion of the methodology and biological findings in the Discussion section below. In addition, we discuss how these results fit within a larger body of work that challenges long-standing paradigms about neuronal connectivity. We end by presenting a conceptual framework that reframes the narrative of synaptic connectivity. Our hope is that this framework will serve as a basis for predictive models of synaptic connectivity.

Bibliography to Chapter 4: Principles of cortical interneuron synaptic organization

1. Kupferman, J.V., et al., *Reelin signaling specifies the molecular identity of the pyramidal neuron distal dendritic compartment.* Cell, 2014. **158**(6): p. 1335-1347.
2. Yao, Z., et al., *A taxonomy of transcriptomic cell types across the isocortex and hippocampal formation.* Cell, 2021. **184**(12): p. 3222-3241.e26.
3. Cooper, J.A., *A mechanism for inside-out lamination in the neocortex.* Trends in Neurosciences, 2008. **31**(3): p. 113-119.
4. Caviness Jr, V. *Reeler mutant mouse: a genetic experiment in developing mammalian cortex*. in *Soc Neurosci Symp*. 1977.
5. Lord, M.C.P. and V.S. Caviness Jr., *Determinants of cell shape and orientation: A comparative Golgi analysis of cell-axon interrelationships in the developing neocortex of normal and reeler mice.* Journal of Comparative Neurology, 1979. **187**(1): p. 49-69.
6. Naus, C.C.G. and F.E. Bloom, *Immunohistochemical analysis of the development of somatostatin in the reeler neocortex.* Developmental Brain Research, 1988. **43**(1): p. 61-68.
7. Hevner, R.F., et al., *Postnatal shifts of interneuron position in the neocortex of normal and reeler mice: evidence for inward radial migration.* Neuroscience, 2004. **124**(3): p. 605-18.
8. Hammond, V., et al., *Layer Positioning of Late-Born Cortical Interneurons Is Dependent on Reelin But Not p35 Signaling.* The Journal of Neuroscience, 2006. **26**(5): p. 1646-1655.
9. Pla, R., et al., *Layer Acquisition by Cortical GABAergic Interneurons Is Independent of Reelin Signaling.* The Journal of Neuroscience, 2006. **26**(26): p. 6924-6934.
10. Rymar, V.V. and A.F. Sadikot, *Laminar fate of cortical GABAergic interneurons is dependent on both birthdate and phenotype.* J Comp Neurol, 2007. **501**(3): p. 369-80.
11. Yabut, O., et al., *Abnormal laminar position and dendrite development of interneurons in the reeler forebrain.* Brain Res, 2007. **1140**: p. 75-83.
12. Katsuyama, Y. and T. Terashima, *Developmental anatomy of reeler mutant mouse.* Dev Growth Differ, 2009. **51**(3): p. 271-86.
13. Steindler, D.A. and S.A. Colwell, *Reeler mutant mouse: maintenance of appropriate and reciprocal connections in the cerebral cortex and thalamus.* Brain Research, 1976. **113**(2): p. 386-393.
14. Boyle, M.P., et al., *Cell-type-specific consequences of reelin deficiency in the mouse neocortex, hippocampus, and amygdala.* Journal of Comparative Neurology, 2011. **519**(11): p. 2061-2089.
15. Dekimoto, H., T. Terashima, and Y. Katsuyama, *Dispersion of the neurons expressing layer specific markers in the reeler brain.* Development, Growth & Differentiation, 2010. **52**(2): p. 181-193.
16. Wagener, R.J., et al., *The Somatosensory Cortex of reeler Mutant Mice Shows Absent Layering But Intact Formation and Behavioral Activation of Columnar Somatotopic Maps.* The Journal of Neuroscience, 2010. **30**(46): p. 15700-15709.
17. Lim, L., et al., *Optimization of interneuron function by direct coupling of cell migration and axonal targeting.* Nat Neurosci, 2018. **21**(7): p. 920-931.
18. Korsunsky, I., et al., *Fast, sensitive and accurate integration of single-cell data with Harmony.* Nature Methods, 2019. **16**(12): p. 1289-1296.
19. Blondel, V.D., et al., *Fast unfolding of communities in large networks.* Journal of Statistical Mechanics: Theory and Experiment, 2008. **2008**: p. 10008.
20. Caliński, T. and J. Harabasz, *A dendrite method for cluster analysis.* Communications in Statistics, 1974. **3**(1): p. 1-27.
21. Davies, D.L. and D.W. Bouldin, *A Cluster Separation Measure.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 1979. **PAMI-1**(2): p. 224-227.
22. Karube, F., Y. Kubota, and Y. Kawaguchi, *Axon Branching and Synaptic Bouton Phenotypes in GABAergic Nonpyramidal Cell Subtypes.* The Journal of Neuroscience, 2004. **24**(12): p. 2853-2865.
23. Wang, Y., et al., *Anatomical, physiological, molecular and circuit properties of nest basket cells in the developing somatosensory cortex.* Cereb Cortex, 2002. **12**(4): p. 395-410.
24. Waddington, C.H., *The strategy of the genes*. 1957: Routledge.

Chapter 5: Discussion

5.1. Summary of methodological approach

5.1.1. A multidimensional imaging platform to assess synaptic connectivity

We present here an image-based tool to score synapse target specificity of interneurons at scale. The workflow is effective in both tissue sections and dissociated culture. Given the relative ease of training and its adaptability, the model could be extended to other tissue formats, such as organoids or different brain regions. In addition, the classification schema could be modified for different target staining to assess other forms of synaptic targeting, such as glutamatergic neuron targeting, or even cellular or regional specificity. The workflow utilizes two approaches to presynaptic bouton classification: supervised for hypothesis testing, and unsupervised for discovery. We show examples of both classification approaches in this study to address unanswered questions in interneuron biology. We discuss these findings in greater detail below. Here we will discuss the workflow itself, compare it to existing approaches, and outline places for improvement.

The workflow consists of five general steps: 1) input standardization, 2) image correction, 3) pixel classification, 4) object segmentation and measurement extraction, and 5) object classification. Different metadata are extracted and in a centralized repository, thereby facilitating bookkeeping, image processing, and analysis. These stages are largely modular, making it easy to add, remove, or modify intermediate processing steps if required. As a principle, we wanted to minimize user interaction at any given stage of processing. Ideally, the user would have to do no more than designate the input directory and output directory, confirm a standard set of parameters, and press ‘Run’, with the code handling all subsequent steps for any number of input images. We also sought to reduce human error by minimizing file handling: a) any stage that requires more than one input type for the same image aligns these inputs automatically from the available files,

and b) the outputs for each stage serve as inputs for the next. We have structured our file- and object-naming system that enables users and software to track various forms of the same sample through all stages the pipeline, facilitating integration of these data into a database. The platform does not require any special techniques or equipment: our data is based on immunofluorescence microscopy images acquired using a laser-scanning confocal microscope, both standards of the field for decades. Moreover, the pipeline is easily generalizable: it can be adapted for use in almost any biological question that utilizes fluorescence microscopy simply by changing the stains, imaging parameters, classifiers, and segmentation steps.

The key innovation of our approach is the harnessing of multidimensional imaging data for parallel supervised and unsupervised classification approaches to characterize synaptic connectivity. Our inclusion of a pixel classification step generated biologically-relevant quantitative annotations to fluorescence microscopy images. In addition, it acts as a force multiplier that enabled us to extract over 400 metrics per object from standard four-channel images. Coupled with the large-scale, population-wide dataset, this depth and breadth of data provided sufficient power to identify distinct synaptic populations. Our supervised classifier has high precision, accuracy, and sensitivity, indicating overall high levels of concordance with manual human scoring. Unsupervised classification of the same objects revealed previously unrecognized diversity within these canonical classes. Moreover, parallel classification approaches served as a means of cross-validation. Since the underlying metrics used for classification are location-independent, we were able to utilize spatial patterning as an orthogonal means of validation and discovery. This approach allowed us to identify an unprecedented level of self-organization in both dissociated culture and intact tissue slices.

5.1.2. Comparison to other studies

Our approach to characterizing presynaptic boutons is similar to other synapse imaging platforms developed independently by Seth Grant's lab at the University of Edinburgh[1-4] and Mark Bathe's lab at MIT[5-7]. The study from Zhu and colleagues[1] is the best representation for comparing our platform to that from the Grant lab. Here, Zhu et al., used two fluorescently-labeled postsynaptic scaffolding molecule, PSD95 and SAP102, expressed from their native genetic loci to characterize excitatory synapses throughout the brain. The group extracted multiple features and performed an unsupervised multidimensional analysis to identify synaptic heterogeneity across the brain. There are many strengths to the approach used in this and subsequent studies from the same group[2, 4]. First, the all-genetic labeling approach side-steps a lot of batch-to-batch variability inherent to immunostaining. Second, the far higher resolution used by these authors yields a more detailed picture of synaptic structures. Third, the scale of imaging (whole brain) provides a much more comprehensive picture of the synaptome. Finally, mapping the synapses onto a common coordinate framework both contextualizes and generalizes the results. Despite being limited to only two channels and purely post-synaptic labeling, the authors found a remarkable level of synaptic diversity throughout the brain, with specific enrichment within different brain structures and cortical/hippocampal layers. Our approach coupled immunolabeling of subcellular target compartments with a conditional transgenic labeling approach to visualize the presynaptic boutons of specific cortical inhibitory interneuron populations. Our approach represents multiple steps forward. First, the use of targeted population labeling not only provided us with an overall picture of the MGE interneurons synaptome (via Nkx2.1^{Cre}), it also allowed us to deconvolve this picture into distinct subpopulations (Sst ^{Cre} and PV ^{Cre}). This approach, then, provides an inroad to the investi-

gation of cell-type specific synaptomes. Second, visualizing postsynaptic proteins and target compartments in tandem with the presynaptic label provides a more comprehensive picture of the synapse. Not only did this provide important information on synaptic targeting and specificity, it also allowed us to make inferences about transsynaptic cross-talk. In other words, visualizing both the presynaptic and postsynaptic components in tandem provides information about neuronal communication. Finally, our spatial analysis of the unsupervised classification represents an orthogonal approach to validate the statistically-defined synapse classes. Due to technical limitations mentioned in Chapter 4, we were unable to perform the same unsupervised classification at scale on cortical slices. It will be interesting to see whether the spatial organization hinted at from our preliminary analysis holds up across multiple samples and labeling approaches.

The Bathe lab has also developed a similar imaging approach to profile and classify synapses[7]. Specifically, the approach utilizes multiplexed fluorescence microscopy of over a dozen antibody-linked nucleic acid probes to image synaptic proteins in a single sample. Guo et al coupled this labeling approach with super-resolution microscopy to visualize the organization of these probes at the nanometer scale. In a testament to the extensibility of this approach, subsequent publications have utilized this platform to compare the molecular compositions of excitatory and inhibitory synapses[5] and to assess the effects of disease-associated genetic interventions[6]. Although there is no published use of the platform on intact tissue slices, it has been utilized in complex 3D cultures of stem cell derived organoids[8]. There are multiple similarities between this approach and ours, though it is also worthwhile to point to some subtle differences.

First, we will discuss sample preparation and acquisition parameters. Since the Bathe lab has published multiple studies using their platform, we will primarily focus on that from Danielson

and colleagues[5]. Like our platform, the Bathe lab's approach visualizes both pre- and post-synaptic labels, providing a comprehensive picture of the synapse. Despite this similarity, the focus was quite different. The studies from the Bathe lab assessed synapse diversity primarily from a molecular and morphological standpoint. We on the other hand were primarily interested in synaptic diversity as it pertains to targeting specificity. Thus, the Bathe lab labeled general synaptic proteins, whereas our labels were chosen to identify target compartments. This speaks to the malleability and power of these approaches to investigate diverse biological questions. Despite this similarity, a major difference is the type and number of biological targets visualized. Compared to the Grant lab, which uses a 2-channel all-genetic approach, and our study, which uses a 4-channel hybrid genetic and immunostaining approach, the Bathe lab developed an (ultimately) all-immuno approach that is capable of imaging upwards of 12 biological targets within the same sample. The expansion of biological targets is a major innovation, which we discuss in greater detail alongside other multiplexed imaging approaches below. We will now discuss the all-immuno approach. There are both positives and negatives to this approach. A) it is a generalized approach to labeling synapses, where all synapses expressing the targeted biomolecule are labeled. This is a top-down approach that relies on post-hoc deconvolution and marker expression to identify different synaptic populations. We took a bottom-up approach of labeling specific neuronal populations and asking how the synapses differed between those populations. This bottom-up approach embraces cellular diversity as a basis for synaptic diversity. The former is more equipped for discovery, whereas the latter is better equipped for hypothesis testing. Both approaches are valid. B) antibodies enable a huge level of control and versatility at the expense of batch-to-batch systematic error. Antibody panels consist of modular units that can be swapped based on the question at hand. As demonstrated by subtle changes across studies[5-8], this approach is far more malleable than a fully genetic

approach. This malleability comes at a cost, however. Antibodies are notoriously finicky. Sample-to-sample and batch-to-batch variability with antibody staining is a major issue. Each antibody-antigen interaction acts independently, multiplexed imaging experiments rarely have established internal controls, and the number of targets (in the dozens) is often insufficiently powered for robust statistics-only based correction of systematic differences. To be fair, this is also a limitation of our own study. We discuss this problem in greater detail below. Finally, another major difference in imaging parameters is resolution. Compared to our study, Danielson et al. used far higher resolution. The higher resolution provides a higher level of confidence in characterization, since the point spread function at low resolution yields greater signal overlap. The use of a far higher resolution limits the amount of sample imaged, however. This is a trade-off that ultimately depends upon the biological question at hand. Danielson et al. were primarily interested in the molecular composition of synapses, which required high resolution imaging. We were interested in synaptic connectivity, and chose our imaging parameters based on the size of the target compartments. This compromise is evident in the size of the dataset: Danielson et al. had a dataset of 180,000 synapses, whereas ours was over 1.6 million. That is, we sacrificed resolution for the sake of scale. Advances in microscopes and imaging speeds at high resolution will enable high-resolution imaging at scale.

Second is image processing and analysis. Like our incorporation of the presynaptic pixel classifier, this platform utilizes an AI/ML approach for synapse detection based on a probability map (DoGNet)[9]. DoGNet is architecturally more sophisticated than our pixel classification approach: it utilizes a convolutional neural network to calculate synapse probability, whereas we are using a random forest approach. On the other hand, our approach presupposes (and accounts for) a higher degree of biological sophistication: whereas DoGNet differentiates between only two classes (synapse vs non-synapse), we differentiate the presynaptic channel into four mutually exclusive classes

(background, bouton, labeled soma, neurite). As discussed above, synaptic proteins are often found in non-synaptic contexts, and 'off-target' staining is common in biological imaging. Multi-classing can help to segregate these signals, thereby integrating biologically-relevant context that would otherwise be lost from a more simplistic approach to feature extraction. We incorporate this concept into our feature measurements by extracting features not only from the raw staining image, but also from the presynaptic and target probability maps. The differences in these approaches is evident in the number of features used for classification. Despite imaging 12 biological targets, Danielson et al., extracted only 108 features. Our four-channel imaging, on the other hand, yielded over 400 non-overlapping features. This added depth provides additional power for discovery. This group also utilized a multidimensional, unsupervised statistic approach to classify synapses[5-7]. Although similar at face value, there are some subtle differences between our approaches. A) Unlike these studies, we also incorporated a supervised branch of object classification. The supervised and unsupervised approaches work synergistically to cross-validate each other, thereby providing more confidence in the ultimate models. B) We also incorporated spatial analysis as an orthogonal mode of validation. Such an analysis makes sense given our focus on subcellular targeting, but may not translate as a validation method for all questions. That said, independent validation of unsupervised classification is critical.

5.1.3. Current and future imaging approaches to synaptic diversity

Highly multiplexed and multidimensional imaging platforms provide a new avenue to discover diversity and characterize biological objects. One clear conclusion from the studies mentioned above is that synapses are not only remarkably diverse, but also that there is an inherent organization to this diversity. This makes biological sense given the hierarchical organization of neuronal diversity[10-15] and the statistical approaches used for cluster identification[16]. At the end

of the day the observed diversity is a function of the extracted features. This presents a potential issue in synapse categorization. Since the classification is ultimately based on the feature set, and different experiments will utilize different features (e.g. they visualize different biomolecules), the classification from one experiment will not translate to those from another. Much like the growing pains of single cell RNA-sequencing (scRNAseq)[17], however, technical biases within current pipelines, subtle differences between feature sets, and variability in analytical methods may undermine reproducibility and transference of previously identified classes to new datasets. Across different scRNAseq datasets, large sets of variably expressed genes can be used to cut through systematic experimental error and identify replicable cell types with high accuracy[17]. Such statistical approaches are not possible when the number of visualized targets is low. This picture is further complicated when biomolecules are independently visualized using targeted molecular probes.

Ultimately, a comprehensive analysis of synaptic diversity, and how it is structured, may require a standardized feature set from panel-less, genome-wide imaging data. Toward this end, considerable advances have been made recently in multiplexed imaging and high dimensional histology techniques to overcome the spectral limits of conventional fluorescence microscopy and increase the number of biological molecules imaged in a single sample. These approaches include including spectral unmixing[18, 19], cyclic staining[20-23], and DNA barcoding[24, 25] (reviewed in [26-29]). As described above, such multiplexed approaches have already been used to profile synaptic diversity in tissue culture[5, 7, 8, 30] and neuronal connectivity in intact brain[31]. Orthogonal labeling of the same structure will offer redundancy, ultimately yielding a more robust platform. Combining these multiplexed imaging techniques with our approach to pixel classification and feature extraction would enable significantly higher dimensional analysis and finer parsing of synaptic

subclasses and phenotypes. It may also offer an avenue for batch correction and classification transference.

This is not to say that limited classification systems, like those discussed above, are invalid. The validity of any given classification system must be based on its utility. Hypothesis-blind approaches are, at the end of the day, tools for hypothesis generation. These hypotheses, in turn, require orthogonal approaches to validate the discoveries. In other words, a classification system is validated by its ability to provide biological insights beyond simply diversity. This brings up a good question: how are we to interpret the features used for classification? Unlike scRNAseq datasets, where features are targetable genetic reads, the features in multidimensional imaging datasets are morphological and photometric readouts of biomolecules extended in space. Whereas differences in gene expression profiles often offer a clear pathway forward, the biological significance of differentially expressed features from multidimensional imaging is less clear. Abstract mathematical entities, like variance or kurtosis of a biomolecule within a segmented region of space, are not readily targetable for intervention. The similarity between imaging and single cell RNAseq datasets breaks down here. Multidimensional imaging features offer phenotypic descriptors, the biological relevance of which may not always be immediately apparent. Instead of peering deeper into the class-defining features, the spatially-extended nature of imaging data offers something that scRNAseq cannot: the ability to assess spatial relationships *between* identified object classes. As discussed in greater depth below, spatial analysis of our unsupervised bouton classes ultimately revealed fundamental differences in how basket and Sst cells distribute inhibitory coverage onto target cells. We therefore offer an example for how data from these platforms can be utilized. We hope this pipeline serves as a useful guide for designing and implementing biological image processing and analysis in neuroscience and other fields.

Synapses are highly polar objects. We approached object identification by segmenting the presynaptic bouton only, then measuring signal intensity or probability using scalar (i.e., non-directional) metrics within the object bounds or its immediate surroundings. Previous studies have assessed similar questions of synaptic diversity from a post-synaptic perspective[1-3]. A true assessment of the synaptome would combine both approaches, with segmentation of the full (pre- and post-) synapse. Such an analysis yield a more comprehensive survey of these subcellular structures by enabling the extraction of more information: 1) the same scalar object metrics extracted for the presynaptic bouton here could be extracted also from a) the postsynaptic component, b) the synaptic cleft (pre-post overlap), and c) the synapse as a whole (combined object), and 2) identification of pre- and post-synaptic components would enable vector-based (i.e., directional) measurements, which would more robustly reflect synaptic targeting and provide more nuanced metrics of synapse structure and biology. Finally, an all image-based method that abandons user-designated object-associated features altogether may offer a more objective and sensitive approach to synapse classification. Given the highly polar structure of synapses, such an approach will require a prior step of object alignment, which is crucial given the largely random orientation of synapses in tissue. Such an approach would offer the added advantage of unsupervised feature selection, which could in turn be helpful for discovering new synaptic phenotypes.

5.1.4. A hurdle to progress: batch-to-batch variation

We used two stages of batch correction in this study, 1) naïve image intensity correction, and 2) Harmony[32] for batch correction of data for the unsupervised classifier. Our current method for batch correction of images stems from correction methods proposed for multiplexed tissue array imaging[33], whereby channel signal intensities for an entire acquisition series are matched via the 75th non-zero quantile to a reference intensity. We are using Harmony, a state-of-the-art method

for batch correction of single cell RNA sequencing (scRNAseq) data to correct for batch-to-batch variation in object (bouton) metrics for the unsupervised object classification. Both batch correction methods have several drawbacks, however. First, a major issue arises from the fundamental assumption operating between both batch correction methods, namely that in aggregate, signal intensities and/or metrics should be consistent from sample to sample. Assuming consistent intensities and adjusting intensities accordingly may obscure true biological variance between samples by over-correcting. This is particularly important when comparing interventions or conditions that are expected to alter the specific signals under investigation. A good example here is the need to use distinct benchmarks for culture and slice data during the intensity correction stage. Second, overall signal intensity is a very blunt metric that does not take into account the wide variety of qualitative changes that can occur between staining sessions. For instance, the dynamic range between two distinct structures stained by the same antibody (e.g. intensity of Kv2 at soma vs proximal dendrite or AIS) can vary considerably from one experiment to another. Our approach does not take into account such changes.

There are several methods to counteract these issues: 1) inclusion of a standardized control for batch-to-batch comparison, 2) incorporation of sample knowledge, such as image metadata or acquisition parameters[34], as an orthogonal set of metrics to guide standardization, 3) the use of machine learning methods, such as the use of generative adversarial networks[35, 36], specifically for batch correction and/or model generalization. The method of quantile-matching used in multiplexed tissue array imaging should include an internal control, such as a standard set of cells or tissue included on each slide, to be used as a standard of comparison between staining sessions. Unfortunately, including a standard only gets one so far: it can control for differences in staining and acquisition parameters from one experiment to the next, but cannot control for differences in

tissue preparation (such as perfusion, post-fixation, pre-staining storage, etc.,) that may impact the quality of staining.

Experimental error is classically broken into two types: random error, which can be dealt with by statistical analysis and sufficiently powering experiments, and systematic error, which is associated with systematic differences in sample preparation and measurement. The latter are commonly referred to as batch effects, and can be identified and corrected post-hoc. Systematic error is a major issue in bioimaging, where minor changes in sample preparation or acquisition can lead to systematic differences between samples. The issue here is that the biologically relevant differences between sample conditions are not separable *a priori* from systematic error, and sometimes it is not readily apparent which category the systematic differences fall into. Over-correcting for systematic can mask biologically relevant changes.

The typical approach to identify and correct for systematic error is replication. Biological data is often organized hierarchically. For example, biologists distinguish between technical replicates, which are replicated measures of the same sample, and biological replicates, which are replications of different samples. Large scale imaging experiments are more complex. At the lowest level, they consist of individual pixels within a single image (a tile). Multiple individual tiles comprise a sample region. Multiple sample regions can be used to image the same tissue specimen. Multiple tissue specimens will comprise a single biological sample. In the case of mouse experiments, multiple mice will comprise a litter. Multiple litters a genotype. And so on, and so on. This hierarchy of replication is important for experimental design. And it comes with a number of assumptions: we expect higher degrees of variance between two different levels than within a single level. But incorporating such a hierarchy into our statistical approaches is not straightforward. Often we assume, as we did in this study, that measures should generally be consistent across experiments. As

stated above, this can lend itself to over-correction and potentially mask biologically relevant differences. We suggest that future approaches to batch correction in bioimaging should be developed with this hierarchy in mind. We did not take such an approach in this study. But we have created a pipeline where these data are automatically collected and systematically stored. This is the first step. Such standardized data structures can be incorporated into models to parse systematic error from biologically meaningful differences. Future work will be needed to tackle this problem.

5.2. Summary of biological findings

5.2.1. Insights into interneuron development from dissociated culture experiments

In this study, we labeled three distinct cortical interneuron populations that target different subcellular regions of the postsynaptic cell. We then utilized the image processing and analysis pipeline described above to assess subcellular targeting of these populations in intact tissue slices and dissociated culture. This experimental design allowed us to compare directly interneuron targeting fidelity within these two contexts, and to test any potential deviations from expected in dissociated culture. We found that, overall, all three interneuron populations generally recapitulated their *in vivo* targeting profiles: somatostatin cells (SstCs) preferentially targeted dendrite, basket cells (BCs) targeted soma and proximal dendrite, and chandelier (ChCs) cells targeted the axon initial segment (AIS). That is, we found no statistical difference in targeting proportions between intact slice and dissociated culture.

At face value, this finding suggests that subcellular targeting is recapitulated in dissociated culture. We cautiously accept this conclusion, but suggest a more nuanced interpretation. We discuss the nuanced interpretation in greater depth below. To explain the caution, we'll use ChCs as an example. Electron microscopy studies of individual ChCs in slice have shown that these cells display near 100% AIS targeting[37, 38]. Our data appears to conflict with this observation: none of

the ChC regions analyzed displayed over 50% AIS targeting. There are three possibilities for this disagreement: 1) the ChCs actually form off-target synapses 2) there is contamination of non-ChC synapses, and 3) our measure is noisy, i.e., false positive classification of 'non-bouton' objects within the 'bouton' class. We hold the EM data as a gold standard, and therefore conclude that the deviation in slice must be due to the latter two possibilities. First, contamination. Unlike the EM studies, which investigate targeting of individual ChCs, we are assessing targeting across entire ChC-enriched cortical regions. Based on this approach, we do not anticipate a 'pure' population of ChC synapses. Our 'ChC' culture samples are similarly prone to contamination. These cells derive from targeted perinatal dissections of the same ChC-enriched regions. (We did this to ensure parity in comparison between conditions.) Thus, contamination explains one component of the difference. Regarding noise: we built our classifiers to be inclusive of interneuron boutons with a broad range of sizes and intensities. The initial formulation to discern 'bouton' signal from non-bouton signal utilized an orthogonal method of verification: counterstaining with the GABAergic presynaptic marker VGAT. Due to microscope channel limitations, we could not include this counterstain in our subsequent targeting studies. Thus, in the absence of this counterstain, we may be including objects that by eye resemble boutons but are not actually presynapses. Classifying such objects as 'boutons' would subsequently result in an increase in 'off-target' categorization. Such noise is prone to batch-to-batch variability, as discussed above, and therefore is difficult to control. It would also be present in both slice and culture samples. We believe these false positives likely explain another component of this deviation. Additional experiments with an orthogonal means of verifying synapses and the ability to trace each synapse back to its source cell would bolster our confidence in this conclusion. Assuming that contamination and noise are statistically consistent

between conditions and stabilized through replication, we conclude that the amount of off-target synapses quantified using the imaging workflow is similar between slice and culture.

Our observation that subcellular targeting is recapitulated in a dissociated culture context sheds light on two important aspects of interneuron development. Specifically, it eliminates the following theoretic possibilities:

1. cortical interneuron identity (subcellular targeting profile) relies upon the local cortical environment after settling (cellular positioning, etc.)
2. subcellular targeting relies upon a precisely structured cortical environment to guide axons to their targets and/or refine promiscuous connectivity to its ultimate specificity.

Instead, these data suggest these interneuron classes have little reliance on their terminal environment in determining subcellular targeting preferences. We discuss the implications of this finding on our understanding of interneuron identity and mechanisms of subcellular targeting in turn.

5.2.2. Persistence of interneuron identity

The possibility for a context-dependent change in cellular identity is best encapsulated by two competing models of cellular specification: the progenitor specification model and the progressive specification model[39]. The progenitor specification model, in its most extreme form, posits that interneuron identity is established at neuronal birth through environmental cues that bestow a genetic blueprint for differentiation and maturation. The progressive specification model, on the other hand, posits that the epigenetic information necessary for sculpting interneuron identity is acquired later in development, through interactions with the cortical environment. From a reductionist standpoint, the former posits that the cortical environment simply potentiates intrinsically-

determined fate, whereas the latter proposes that the cortical environment acts instructively on a more plastic identity to determine the interneurons ultimate fate. *In vivo* attempts to alter cortical spatial organization by interfering with glutamatergic neuron migration have shown that interneurons adopt their laminar position in response to cues from the pyramidal cells they will ultimately target[40-43]. That is, despite abnormal positioning of the target cells, interneurons will position themselves as 'normal' near those cells, thereby ensuring a normal cellular environment and paracrine signaling. This difficulty repositioning interneurons into an abnormal cellular environment *in vivo* has confounded attempts to test these models. Dissociation and plating, on the other hand, decouples interneurons from their target cells and fully randomizes the cellular environment. It therefore offers a platform to sidestep this confound and more completely test these models.

Our results suggest that an interneuron's subcellular targeting profile is genetically and irreversibly specified prior to axodendritic polarization. This is not to say that the entirety of molecular expression profiles, firing patterns, etc., are predetermined at birth. Heterotopic transplantation studies have shown that specific molecular markers, cellular morphology, and physiology of interneurons are impacted by the local environment[44, 45]. Assuming that these molecular and morphological changes are due to selective differentiation rather than selective survival, the most parsimonious explanation is that broad brushstrokes of interneuron cellular identity are largely specified before they reach their final settling place in the brain, in line with a progenitor specification model, whereas finer details of interneuron phenotype are progressively specified at different developmental stages, as different genetic programs are irreversibly selected in response to the local environment. In other words, specific limits are set at birth for the ensuing plasticity.

What aspects of cellular identity abide by which model? Our data indicates that the primary designator of an interneuron's identity, its targeting preference, is established prior to axodendritic

polarization. That is, it does not arise in response to the local cellular environment after settling. This aligns with results from molecular expression profiling of interneurons during development, which showed that distinct, cardinal populations of interneurons were identifiable during migration[46-48] and even at different progenitor regions within the ganglionic eminences[49, 50]. In contrast, the final cellular environment is critical in determining selective survival[51, 52] and fine-tuning terminal aspects of identity, such as cellular morphology and activity-dependent expression profiles[44, 45]. The morphological aspects in particular are supported by our data, which show that interneuron synapse positioning is ultimately determined by the positioning of the target. The fact that we see recapitulation of subcellular targeting in a dissociated culture means that we can use this method to test more conclusively whether interneuron identity is specified at birth or results from progressive stages of specification over development. In-depth molecular, physiological, and morphological profiling of interneurons in dissociated culture vs intact tissue would provide more information on the degree of recapitulation in this randomized context.

5.2.3. Models for cortical interneuron subcellular targeting

We outlined multiple models of synaptic specificity in the introduction, and generally divided them into 'pre-contact' and 'post-contact' models. We also differentiated between the presynaptic and postsynaptic cells, which are ultimately involved in the synaptic connection, from third party 'environmental' mediators, that are not. Our data conclusively shows that the spatial organization of the cortical environment is not a prerequisite for 'appropriate' interneuron subcellular targeting. This finding allows us to more precisely pinpoint the degree to which these interneuron populations rely on these different models for 'proper' subcellular targeting. Generally speaking, we can conclude that cortical interneurons have little reliance on third party mediators to establish this

connectivity. Rather, their targeting specificity is heavily reliant on intrinsically determined molecular expression and localization profiles within the presynaptic and postsynaptic cells. Regarding the presynaptic cells, as discussed above, this means that the operative receptors on the growth cone are determined prior to axodendritic polarization. For the target cells, this means that not only are the postsynaptic recognition molecules and guidance cues similarly pre-determined, but also that they generally localize to the correct subcompartments via cell intrinsic mechanisms. This also means that cortical interneuron targeting specificity is heavily biased towards hyper-local mechanisms[30], such as target-directed axon guidance cues and/or domain-based target recognition models. In other words, they have little reliance on 'pre-contact' third-party mediators of axon guidance, including large scale haptotactic and chemotactic guidance cues as well as cells that act as guideposts or scaffolding. This contrasts sharply with what is known of long-distance glutamatergic connections, which rely heavily on these mechanisms for regional, laminar, and cellular specificity[54-61]. This makes biological sense given the differences between projection neurons and interneurons: projection neurons project to distal regions, and thus are highly reliant on the intermediate environment to guide their axons to their ultimate destination; Interneurons, on the other hand, project locally, with the same cell type forming repetitive motifs in many different brain regions. To be sure, our results do not exclude all pre-contact mechanisms outlined in the introduction. Interneurons may place a heavy emphasis on local chemotactic cues for axon guidance, provided these are expressed by the target cells themselves. Such hyper-local chemotactic guidance cues have been shown for interneuron populations in other brain regions, such as cerebellar

[30] Here we define 'hyper-local' based on the effective intercellular communication range defined by Francis & Palsson, 1997 [53].

basket cells[61], and it is reasonable that cortical interneurons would also utilize similar mechanisms.

At face value, our findings on somatostatin cells appear to contradict the previously-proposed mechanism for how they target pyramidal cell distal dendritic tuft in L1[46]. Lim and colleagues showed that translaminar Sst cells target this region by leaving a trailing process during radial migration out of the superficial layers. Since the distal dendritic tuft does not form in dissociated cultures[62], we are unable to test whether translaminar somatostatin cells preferentially target distal dendritic tuft when uncoupled from migration. That said, it is noteworthy that, according to our spatial organization data, somatostatin cells were specifically enriched for the most distal of the dendritic subtypes (type 6), though it is unclear whether these boutons in culture bear any molecular resemblance to those that target distal dendritic tuft of intact tissue. Based on these differences, we propose that somatostatin cells utilize different mechanisms for translaminar and intralaminar targeting of dendritic compartments in intact tissue. This makes sense since the translaminar trailing process is distinct from the intralaminar axons it sends out after settling. In dissociated culture, Somatostatin cell targeting is likely more akin to intralaminar targeting, as supported by the proportional enrichment of subcellular targeting, which matched that of the cellular layers (L2-6). Whether translaminar Somatostatin cells utilize the same local guidance cues and target recognition molecules for both processes is unclear, however. These questions could be addressed by coculture experiments, where dissociated MGE from animals with genetically label Martinotti cells are seeded onto organotypic slices with and without L1 or the cellular layers.

Another model for synaptic specificity is promiscuous targeting followed by specific elimination. Microglia have been shown to mediate large scale synaptic pruning in an activity-dependent manner[63, 64], with specialized microglia that remodel GABAergic synapses[65]. Our culture

conditions, however, do not support growth and survival of microglia[66], meaning this mechanism for targeted synapse elimination is absent. This does not exclude specific elimination altogether, but instead limits the potential mechanism(s) in our culture model to processes intrinsic to the pre- and postsynaptic cells. Live imaging, conditions that support microglia[66], and pharmacological manipulations[67] could all be utilized to test this possibility in our culture model.

In the parlance of 'necessary' and 'sufficient,' our data suggests that third party mediators of axon guidance and synapse refinement are not necessary for proper interneuron subcellular targeting, and that hyper-local molecular cues from the postsynaptic cell are sufficient to explain this recapitulation. We note, however, that the models for subcellular targeting presented in the introduction are not mutually exclusive, and that these cells may utilize additional pre- and post-contact mechanisms to optimize targeting efficiency and specificity *in vivo*. In the following sections we propose a framework for understanding and interpreting synaptic targeting that provides a more nuanced view of these findings.

We also note that we did not investigate cellular targeting specificity. Chandelier, basket, and somatostatin cells all display high levels of laminar specificity in intact tissue[68-71], but it is unclear what underlies this phenomenon. Spatial analysis of chandelier connections from electron microscopy data, for instance, showed no evidence of target selectivity beyond spatial proximity [38]. This finding accords with primary reliance on hyper-local targeting models outlined above, and suggests that cellular specificity may be based on nothing more than settling position. Alternately, the general confinement of boutons to a single layer for intralaminar targeting could be due to repulsive guidance cues in adjacent layers. In line with this latter model, full reconstructions of individual axo-axonic cells (AACs) revealed that some AACs displayed translaminar targeting that skipped intermediate layers (e.g. targeting of AISs in L2/3 and L5, but not L4)[70]. This finding

suggests an interplay between environment-mediated axon guidance and cellular recognition in target AIS selection. To be sure, this axon guidance could operate through classical haptotactic/chemotactic molecular cues, as occurs with projection neurons[54, 57, 58, 72-75], or by coupling migration with axonal targeting, as has been shown for translaminar somatostatin cells[46]. Either way, these data suggest that the spatial organization of the environment may be important for interneuron cellular specificity. We did not assess cellular specificity in this work because 1) we lacked the means of labeling these subtypes and their targets, 2) large-scale cellular connectivity of interneurons has only recently been mapped, and 3) channel limitations forced us to prioritize subcellular specificity over cellular specificity. New methods to label specific subtypes of interneurons[14, 76, 77], recent large-scale work to map their cellular specificity[69, 71, 78], and multiplexed imaging approaches (reviewed in [26-29]) all help to sidestep these limitations. In combination with the platform introduced here, tackling this question is now much more feasible.

5.2.4. Redefining synaptic targeting classes

As noted above, one of the major advantages of our approach is parallel supervised and unsupervised classification. We used the same data discussed above for unsupervised clustering to identify bouton subclasses using a hypothesis-blind method. We utilized metrics such as the CH, DB, and LISI scores to objectively choose parameters for clustering. This approach yielded nine distinct bouton types, each of which displayed high levels of enrichment for one of the three primary targeting classes from our supervised approach: six dendrite-targeting, two soma-targeting, and one AIS-targeting bouton types. We found that individual basket, chandelier, and somatostatin cells were specifically enriched for some of these subtypes over others, which accords with our expectations and helps to validate this classification model as biologically meaningful.

The identified subtypes were not randomly distributed, but instead displayed a remarkable level of spatial organization within a culture dish: the two soma-targeting subtypes were associated with distinct soma profiles, whereas different dendrite-targeting subtypes were positioned at increasing distances from the soma. Regarding the two soma-targeting subtypes, we found individual soma exclusively associated with or heavily enriched for one soma-targeting subtype or another, suggesting the subtype is determined in part by properties of the postsynaptic cell. We also witnessed individual basket cells displaying enrichment (but not exclusivity) for one subtype or another, suggesting preferential cellular targeting in dissociated culture. These findings are in line with findings from hippocampal Dentate Granule cells, which have morphologically and functionally distinct synaptic terminals on the same axon depending on their postsynaptic target[79, 80]. They also bolster the notion that retrograde transsynaptic signaling is a major determinant of presynaptic properties. Further characterization of target cells (e.g., via soma morphology or cell type marker expression) and comparison to overall GABAergic bouton somatic targeting (e.g., via parallel staining with VGAT) would help to bolster this observation.

Regarding the six dendrite-targeting subtypes, our spatial clustering analysis found that these bouton types are located at increasing distances from soma. Specifically, the unsupervised classifier was able to identify boutons at distinct proximal, medial, and distal dendritic regions relative to soma. We note that this proximal-to-distal sequential ordering of the dendrite-targeting classes reflects distance from *any* soma, not distance from the *target cell's* soma along the dendrite. Although the latter is what neuroscientists typically mean when discussing the proximal, medial, and distal dendritic compartments, the randomized positioning and orientation of cells, combined with the lack of markers for these compartments in dissociated culture complicates such an analysis. Future experiments will be required to address whether the ordering identified here corresponds

with positioning *along* the dendrite. That said, this finding suggests a previously unrecognized level of spontaneous self-organization of interneurons in dissociated culture. In line with this observation, the most distal dendrite-targeting subtype (type 6) was enriched in individual somatostatin cells, which target more distal regions of the dendritic tree than basket cells.

In the cortex, a defining feature of Martinotti cells, a subset of somatostatin cells, is translaminar targeting of distal dendritic tuft in L1. It is unclear whether the most (spatially) distal dendritic classes identified in culture correspond to those targeting distal dendritic tuft in intact tissue. Although a possibility, other studies suggest this is probably not the case. Martinotti cells couple axonal targeting of dendritic tuft to their radial migration out of the marginal zone[46], which cannot occur in this dissociated context where these cells do not migrate. In addition, the distal dendritic compartment is not formed via cell intrinsic mechanisms in culture[62], suggesting the synaptic recognition molecules typically enriched at this subcellular region in tissue are distributed throughout the dendritic tree in dissociated culture. Thus, the increased targeting of more distal dendrite in somatostatin cultures is likely due to a previously unrecognized targeting mechanism that is not shared by basket cells, such as the different axon growth strategies discussed below. Moreover, somatostatin cells likely utilize this targeting mechanism within the cellular layers of the cortex, after settling. Although it has not been a primary focus of research on somatostatin cells, all Sst subtypes form intralaminar synapses[76]. It will be important to characterize these synapses further, both with regard to their molecular composition and their subcellular positioning on the target cell. Profiling of intralaminar and translaminar somatostatin synapses, and how they compare to the synapses formed in a dissociated culture, will be helpful for defining the biological relevance of the synapse classes identified by unsupervised clustering.

This spatial analysis also revealed that BCs and SstCs display different axon growth patterns. Our sequentiality analysis indicates that both BCs and SstCs display non-random axonal growth patterns in culture, and further, that BCs adhere more closely to underlying cellular topography than SstCs. Since we are unaware of terms to differentiate such neurite growth strategies, we will coin the term 'vineal' (of or pertaining to vines, < Latin *vīnea*, vine[81]) to describe neurite growth along cellular structures, and 'avian' (bird-like) to describe the high propensity to jump across these structures. Axon guidance can proceed through two distinct molecular mechanisms: 1) via surface-bound (haptotactic) cues, such as subcellular gradients of adhesion molecules, or 2) via soluble (chemotactic) cues, such as secreted molecules that form gradients by diffusing throughout the environment. These sequentiality data indicate that Basket cell axons tend to follow subcellular gradients more closely than Somatostatin cells, suggesting the former are more reliant on hapto-tactic cues, and the latter chemotactic cues, for axon guidance. These growth patterns reflect different strategies for achieving inhibitory coverage that align closely with our understanding of the role these cells play in regulating networks. Indeed, a wide body of literature suggests the spatial positioning of inhibitory synapses relative to the soma leads to different strength and properties of inhibitory control[82-89]. Distal dendritic targeting by SstCs act as a rheostat to modulate target cell responsiveness, whereas somatic and proximal dendritic inhibition by BCs is far stronger, acting as a binary switch that alternately shuts down or enables target cell firing. Our sequentiality data suggests that the axonal growth strategies of BCs and SstCs are an orthogonal mechanism to achieve the same ends: assuming a constant number of boutons per presynaptic cell and an equal synaptic strength per synapse, the more vineal strategy of BCs will lead to stronger and more robust inhibition of fewer cells, whereas the more avian strategy of SstCs will lead to more distributed, but ultimately weaker coverage of individual cells.

5.2.5. Cortical interneuron subcellular targeting in intact tissue

What can we conclude about supervised targeting in slice? Proportionally, our supervised targeting data accorded with expectations: PV ^{Cre} boutons were primarily soma and proximal dendrite targeting, Sst ^{Cre} boutons were primarily dendrite targeting, and Nkx2.1^{Cre}, which labels both PV+ and Sst+ interneurons, displayed an intermediate phenotype. Even more striking was the level of spatial organization of these bouton types. Our population level analysis revealed that cortical interneuron boutons are highly organized. We witnessed both laminar and regional differences in interneuron subcellular targeting.

Regarding laminar enrichment, we found that the spatial patterning generally tracked with the standard divisions of the cortical layers. These population-level synaptic results largely accord with prior work showing the axonal trees of individual interneuron subtypes are often confined to specific laminae[15, 69]. Consistent with what is known about cortical interneuron development and connectivity[14, 90-92], we found that the subcellular targeting profile of the entire MGE lineage is largely accounted for by the Sst and PV populations, each of which displays a distinct targeting profile. Specifically, we witnessed a greater density of dendrite-targeting synapses in deeper layers in PV cells, whereas dendrite-targeting of Sst cells was denser in more superficial layers.

As expected, most boutons in the acellular layer (L1) of the cortex were classified as dendrite targeting. PV boutons were almost entirely absent from this layer, indicating that, as expected, the majority of MGE interneuron boutons found here are from Sst Martinotti cells. In addition, our unsupervised classification approach identified distinct populations of dendrite-targeting boutons within L1, suggesting sublamination. This result aligns with a wide body of literature on sublamination of L1 based on differential molecular expression[93-98] and connectivity[93, 99]. It is unclear what underlies the distinction identified by our unsupervised classifier in slice. These different

bouton classes may derive from different Sst subtypes. According to this model, the class distinction and differential positioning would be accounted for by distinct Sst subtypes preferentially targeting different subdendritic regions of the same postsynaptic cells. For example, two subpopulations of Martinotti cells, fanning-out SST-Calb2 and T-shaped SST-Myh8 cells, have distinct subcellular and sublaminar innervation patterns: the former target distal dendritic tuft and apical branches in lower L1, and the latter almost exclusively tuft in upper L1[76, 99]. Alternately, the class distinction could reflect distinct targets. In line with this possibility, different pyramidal cell types have distinct distal dendritic tuft morphologies and branching patterns within L1, and show different dendritic densities in upper and lower L1[15]. In this case, rather than distinguishing boutons from different presynaptic Sst types, the classifier would be making distinctions based on target cell type or targeted subcellular region. These possibilities are neither mutually exclusive nor exhaustive, of course. Experiments that label different Sst interneuron subtypes and/or target pyramidal cells could address these questions.

We also witnessed labeled boutons in the cortical white matter (WM). Synapses in the WM have been reported in human[100]. We witnessed WM boutons in both Nkx2.1^{CreER} and Sst^{Cre} mice. The WM boutons in the Nkx2.1^{CreER} mouse are likely from the Meis2 interneuron population[13, 101]. These cells reside in the WM and project locally within WM, deep L6, and superficial striatum[101]. Somatostatin cells have been observed in WM of both human[102] and mouse[76, 103]. In mouse, these cells also express NPY and nNos, generally correspond to the Sst Chodl population, and have a projection neuron phenotype[10, 14]. It is unclear whether they form synapses in WM, however. More work will have to be done to characterize these connections and to identify the source cells.

We also observed differences in synaptic target enrichment between cortical regions. In particular, we witnessed increased dendrite-targeting density for the Sst^{Cre}-labeled population within barrels of the somatosensory cortex. The increased concentration of dendrite-targeting boutons in barrels of the Sst^{Cre} mouse is likely due to recently characterized Sst-Hpse and Sst-Crh cells, both of which are non-Martinotti cells that preferentially target barrels[76, 104, 105]. It will be important to take a systematic approach to characterizing synapses of different interneuron subtypes.

Of note, our data is a statistical representation using data combined from multiple mice. These patterns were far more pronounced within individual mice, and even more so within a single cortical region. We suspect the aggregate data masks many of these patterns because we are standardizing the surface to base cortical distance using a linear transformation that assumes consistent spacing of layers across subjects and across regions. We did this because we did not include an independent measure for cortical layer, and using the observed bouton layering pattern for this purpose would be circular. We know, however, that layer spacing differs along the rostral-caudal axis and between cortical regions. Registering the slices to a common coordinate framework, such as that provided by the AllenCCF mouse brain reference atlas[106], would allow us to perform non-linear standardization based on slice reference coordinates. This in turn would provide a more representative illustration of the spatial patterning of these bouton populations that preserves subtle laminar differences.

This study represents an important step forward in generating a fine-mapping of interneuron synaptic identity within the mouse brain. We show on a population-wide level that inhibitory synapses display high degrees of laminar and regional patterning in the mouse brain. Important work has already been done on this front from the perspective of excitatory postsynapses[1]. It will be

interesting to compare the spatial distribution of inhibitory and excitatory synapses within the cortex.

5.2.6. Interneuron targeting in the reeler cortex

What can we conclude from the reeler experiments? We found little quantitative difference between control and reeler conditions in Sst^{Cre} interneuron subcellular targeting. These data further support our conclusions that interneuron subcellular targeting specificity is largely independent of cortical organization and primarily intrinsically determined by the presynaptic and postsynaptic cells. The striking laminar enrichment patterns of dendrite- and soma-targeting boutons witnessed in all other mice were largely absent in the reeler mutant, with little to no differential enrichment across cortical laminae or regions. These data strongly support a model wherein cortical interneuron bouton spatial organization arises in response to the underlying laminar organization of glutamatergic target neurons. The conclusion that interneuron bouton patterning arises from the spatial localization of the targets is in line with prior work on interneuron migration[41] as well as our culture results from this study. Our findings are also in line with studies showing cortico-cortical, cortico-thalamic and thalamo-cortical circuits are largely intact in the reeler mouse[107].

reeler mutants displayed a partial reversal in the spatial distribution of dendrite- and soma-targeting boutons compared to control. The partial reversal is interesting, and elucidates the molecular mechanisms underlying Sst Martinotti cell targeting of L1. L1 targeting by Martinotti cells has been shown to unfold in a developmental process that links tangential and radial migration with axon pathfinding and synaptic specificity. In brief, Martinotti cells form synapses on pyramidal cell distal dendritic tufts in layer 1 during their inward radial migration from the cortical surface. Remarkably, we see a partial reversal of Sst dendrite targeting enrichment in the reeler mouse: dendritic targeting is decreased near the cortical surface, but increased in the acellular zone

near the cortical base. These data suggest that MCs may synapse onto the appropriate targets even in the partially reversed cortical structure. Assuming MCs utilize the same migration/targeting model established by Lim et al.[46] within this altered structural context, our data suggests that MCs respond to the partially inverted cortical environment during tangential migration by following a deeper migratory path near the subventricular zone. They would then leave a trailing axon in this subventricular region during an outward radial migration towards more superficial region, ending this process in the 'appropriate' location near their inappropriately positioned synaptic targets.

This model is partially supported by previously published studies. Inverted pyramidal neurons are a common feature of the reeler cerebral cortex[108, 109], with apical dendrites often pointing towards the subventricular zone. Independent studies have shown that the laminar positioning of cortical interneurons generally correlates with the allocation of projection neurons, even in cases where cortical layering is inverted via disrupted Reelin signaling[41, 110]. Such 'appropriate' positioning was shown for L5/6 Calbindin-positive (CB+) GABAergic interneurons[110], a hallmark expression and positioning profile of T-shaped Martinotti cells[76, 105]. These findings suggest that these MCs adopt a laminar position in response to Reelin-independent cues. In other words, 'appropriate' positioning is likely driven by molecular gradients established by the target cells themselves, and only indirectly related to the reelin gradient insofar as it is responsible for 'appropriate' positioning of the target cells at the cortical surface. This is not the case for all MCs, however. Studies have shown that laminar positioning of L2/3 prototypical MCs[40] and perhaps calretinin-positive L5 Fanning-out MCs[110] is Reelin-dependent. Specifically, these cells display positional randomization and redistributed across all layers, as would be expected for Reelin-dependent migration. Indeed, these MC subpopulations may account for the ectopic (i.e., not inverted) Sst im-

munoreactive fibers enriched in superficial regions of the reeler mouse cortex[111]. Since MC targeting of L1 is closely tied to migration, these data suggest MC-subtype specific effects of the reeler mutation on cortical circuitry. Laminar positioning is not the same as synaptic targeting, however. These models will have to be tested by labeling specific MC subpopulations[76, 112, 113] or migratory streams[46] and investigating targeting in the context of Reelin disruption.

5.2.7. Reframing synaptic specificity from studies in dissociated culture

We found that three populations of cortical interneurons displayed targeting preferences in dissociated cultures that were indistinguishable from those witnessed *in vivo*. Our results help to validate dissociated culture as a reductionist model for interrogating synaptic connectivity. These findings are in opposition to the consensus view that neurons in dissociated culture will predominantly target promiscuously, and synaptic specificity is *not recapitulated*[114]. This view is based on early culture studies of the autonomic nervous system that showed examples of synapse formation in culture for which there are no *in vivo* counterparts and a significant shift in the proportions of physiological connections, such that small minority populations *in vivo* became an overwhelming minority *in vitro*, and vice versa (reviewed in [115]). The belief that dissociated neuronal cultures express synaptic specificity was subsequently bolstered by advances in cellular labeling techniques and subsequent experiments that not only refuted stringent specificity in culture[116], but also expanded the list of non-physiological connections in dissociated cultured neurons[117-119]. In low-density hippocampal cultures, for instance, glutamatergic cells will form non-physiological synapses with mismatched postsynaptic $GABA_AR$ clusters onto the AIS of target cells[117], a region never innervated by hippocampal glutamatergic neurons *in vivo*[118]. At the furthest extreme, isolated neurons on microislands will target themselves, forming autapses with mismatched presynaptic and postsynaptic components[119]. These studies, as well as other supportive data, have spread

the belief that much of the observed synaptic specificity in mature animals derives from promiscuous synaptogenesis followed by specific elimination[120].

Our results on cortical interneuron subcellular targeting push back on this interpretation, suggesting a 'specificity first' model does occur in culture. We are not the first to show recapitulation of synaptic targeting in a dissociated context, however. Indeed, studies have shown that multiple CNS systems recapitulated physiological synaptic targeting in a dissociated context. Williams et al., focusing on the hippocampus, observed that dentate gyrus (DG) neurons in dissociated culture behaved similar to their *in vivo* counterparts, preferentially forming synapses onto CA3 rather than CA1 neurons[114]. Remarkably, the resultant synapses had distinctive phenotypes reminiscent of DG→CA3 Mossy fibers, including size, subcellular location, and physiological profile. Also in hippocampal cultures, GABAergic presynapses were preferentially found on target cell somata, whereas glutamatergic presynapses preferred dendrites, in accord with data from hippocampal slices[121]. These hippocampal innervation patterns were activity-independent, suggesting initial domain specificity without refinement. Similarly, dendrites of dissociated cerebellar granule cells correctly recognize their target axons in vitro[122], and stem cell derived Renshaw cells of the spinal cord displayed correct targeting preferences for V1 interneuron subtypes when co-cultured[123]. Thus multiple CNS systems have been shown to recapitulate both cellular and subcellular targeting specificity in a reductionist, dissociated context.

To further complicate this picture, these affirmative findings tend to be inconsistent in different labs. Linda Van Aelst's group has replicated our findings on cortical interneuron subcellular specificity using our culture protocol (personal communication, Linda Van Aelst). On the other hand, Beatriz Rico's group had previously attempted the same experiments, but instead of witnessing robust subcellular specificity reported here, they observed promiscuous targeting (personal

communication, Beatriz Rico). The same appears to be true of other CNS systems. Different groups have had difficulty replicating the findings from Williams et al. showing preferential DG→CA3 cellular specificity (personal communication, Kristin Baldwin). Despite utilizing the same cells to probe the same connections in dissociated culture, one group observed robust targeting specificity whereas another observed promiscuity. Taken as is, one group's findings suggest a 'specificity first' model whereas the other suggests a 'promiscuity with selective elimination' model.

Let's accept all these results at face value. How do we reconcile these studies? In conversation with Beatriz Rico, we noted that the primary methodological difference between our approaches was culture density. Our protocol uses seeding densities far higher than theirs, and indeed, higher than typically used for dissociated cultures. In line with their results, we found that subcellular targeting was more promiscuous with lower neuron densities or poor culture health (data not shown). Practically speaking, then, we can write off many of these contradictory findings as differences in neuronal populations investigated and culture conditions utilized (more on this below). More philosophically, these experiments reveal that mechanisms utilized during circuit formation in a normal physiological context can be overruled or discarded, and non-physiological mechanisms may be adopted, depending on the local environment. Specifically, they point towards a role for activity-independent molecular competition, both at and outside of the contact site, in regulating target recognition and synaptic specificity[117, 121]. We witness robust recapitulation of physiological interneuron targeting in more physiological, high-density cultures, and less robust recapitulation in less physiological, low-density ones. This heightened cell density likely enables greater levels of molecular competition, which in turn should contribute to greater selectivity in synaptic

connections. At non-physiological cell densities, neurons either die or adapt. Those that adapt embrace an 'any port in a storm' strategy of connectivity that overrules the more physiological subcellular targeting specificity: chandelier cells innervate soma and dendrites (our observations) and glutamatergic cells innervate AIS[117]. In the most non-physiological environment, neuronal isolation, these cells abandon the well-characterized clustered protocadherin-based mechanism for self-recognition and self-avoidance[124] and form non-physiological synapses with mismatched pre- and postsynaptic components[119].

Any attempt at reconciliation must be in line with widely accepted molecular models of synaptic connectivity and cellular differentiation[48, 125, 126]. Sperry's Chemoaffinity Hypothesis proposes that genetically-encoded molecular interactions underlie the specific patterning of neuronal connections[127]. As discussed below, this hypothesis has been refined into the synaptic adhesion molecule code model, where synaptic target recognition is mediated by a combinatorial code of complementary synaptic adhesion molecules[128-131]. Molecular expression profiles are not static, molecular affinities are neither binary nor exclusive, and synaptic molecules do not exist in isolation, however. Differential binding affinities and expression profiles work both cooperatively and competitively at different stages of circuit assembly[128, 131-133]: instead of acting in mutual exclusion, these mechanisms ultimately synergize to yield preferential outcomes *based on the immediate environment*. The critical distinction here is the level of emphasis placed on the immediate environment as an actor in determining synaptic specificity. Disregarding a role for the local environment does not leave a place for mechanisms to explain relative cellular competition and relative affinities. This suggests that the standard lock-and-key model for synaptic specificity should be modified to include not only the presynaptic and postsynaptic adhesion molecules at the contact site, but also the microenvironment within which they operate. This 'microenvironment' includes

the extracellular space surrounding the contact site, as well as intracellular and membrane-bound presynaptic and postsynaptic molecules within the local neighborhood. We discuss the implications of this principle and propose modifications to this conceptual framework for synapse formation below. Suffice it to say, one major implication is that connectivity is not rigidly determined by cellular expression profile, but malleable and probabilistic based on the microenvironment: whether physiological or not, the resulting connectivity is 'appropriate' to the cellular environment and neuronal plasticity.

This reframing places the burden of validity for a tissue culture model of connectivity on phenotyping the resultant synapses in both the physiological and reductionist contexts. Ultimately, defining the limitations of such a model system will require not only higher resolution definitions of synapse types and deeper phenotyping of connectivity (both in the physiological context and in the reductionist system), but also more precise characterization of overall culture conditions and the local microenvironment around a synapse. We noted that our culture system utilizes a high seeding density that is more in line with physiological cell densities than typical 2D culture models. Starting conditions are distinct from end-point conditions, however. Despite using a consistent seeding density, we witnessed wide variations in adherence and survival from one preparation to the next (data not shown). We suggest that studies on neuronal connectivity that utilize culture models should include orthogonal metrics of end-point culture health, such as cellular density and diversity, to ensure reproducibility. Further, the cellular topography within a dissociated culture is not homogonous: local neighborhood metrics should be included at multiple scales to help contextualize findings and normalize results. It will be interesting to see whether targeting specificity scales continuously with competition, or if there is a threshold density below which it collapses.

Further optimization and standardization of culture conditions will be required to scale this approach for more systematic screening assays.

In line with this concept, we note that the total number of connections formed by individual cells was severely decreased in our high-density dissociated cultures compared to intact slice (data not shown)[38, 134-136]. Thus, although we did not witness a change in targeting identity or proportional distribution, the overall *efficiency per cell* is severely reduced with randomization and 2D plating. That is, instead of forming promiscuous connections, the interneurons formed fewer connections. We suspect that this decrease in total synapses results from fewer high affinity targets being within reach, and that dissociated 3D cultures with densities similar to the cortex would more closely parallel the physiological number of connections per cell. These also data suggest that the spatial organization of the cortex is not a prerequisite for proper interneuron targeting specificity *per se*. Rather, cortical organization ensures that more targets are within the limited reach of the axonal arbor, enabling more proper connections. Thus, cortical lamination likely evolved in part as a means of utilizing spatial proximity to optimize circuit assembly, both as a means of limiting detrimental connections and increasing the probability (and total number) of beneficial ones.

5.2.8. A screening platform for neuronal connectivity

One of the most exciting prospects given our findings is the ability to use our platform to screen cortical connectivity at scale. We show a small scale screen of cortical interneuron connectivity following treatment with the cytokine IL13. We found that increased doses of IL13 over a 24-hour treatment period led to significant shifts from dendritic to somatic subcellular targeting. This shows that our dissociated culture system is sensitive enough to detect changes in synaptic connectivity. Moreover, our performance metrics indicate that our supervised classifier is on par

with the human eye. There are two pressing questions in the field of neuroscience: 1) the molecular basis of synaptic connectivity, and 2) the molecular basis of psychiatric diseases, and the desperate need for treatments. Our discoveries open the door to using our approach to address these questions at scale.

It has been proposed that a synaptic adhesion molecule code underlies the specific connectivity of the brain. We discussed this hypothesis at length in the introduction, and we propose an alternate model below. Our platform offers the ability to test this hypothesis at scale. In a landmark study, Favuzzi and colleagues showed that the cortical interneuron subtypes studied here utilize distinct molecular mechanisms for subcellular targeting[137]. The researchers utilized scRNAseq across development to define a list of putative adhesion molecules in each of these different interneuron populations. They ultimately arrived at a list of 26 presynaptic adhesion molecules and 92 putative postsynaptic partners across these three cell types. The authors verified the involvement of one presynaptic molecule for each of the three interneuron populations by painstakingly testing them *in vivo*. Dissociated culture would facilitate more extensive testing of the list of 100+ pre- and post-synaptic molecules. We propose that our platform could be integrated into this workflow to identify and test synaptic adhesion molecules in target recognition and synaptic specificity. Specifically, the workflow would involve: a) verification of targeting recapitulation in culture, b) *in vivo* identification of putative target molecules by scRNAseq, c) *in vitro* target screening by loss and gain of function interventions, and d) *in vivo* confirmation of hits using tissue slices. This workflow would involve building a supervised classifier to assess synaptic targeting both in culture and in slice. This workflow is not limited to cortical interneuron subcellular targeting. It could be utilized for any form of synaptic connectivity, so long as the targeting is recapitulated in culture. In addition, it does not necessitate the use of dissociated culture: organotypic culture using intact

slices would also work. This workflow, then offers an unprecedented ability to test the molecular basis of neuronal connectivity.

A similar approach could be utilized to discover new therapies for psychiatric diseases. Psychiatric diseases are associated with aberrant information flow. This altered flow of information ultimately arises due to abnormal synaptic connectivity or dysregulated synaptic communication. Interneuron dysfunction is a common theme in many psychiatric disorders[138, 139] and epilepsy[140, 141]. Postmortem studies of brains form individuals with schizophrenia (SCZ) have revealed decreased BC and ChC cell and synapse numbers and altered GABAergic receptor subtype expression[139, 142-147]. Autism spectrum disorder (ASD) has been associated with BC, ChC and SstCs[148-150] and excitatory/inhibitory imbalance[151]. Both parvalbumin and somatostatin cells have been implicated in various genetic forms of epilepsy[140, 141]. These associations are backed by genetic mouse models, such as ErbB4 in SCZ[152-154], MECP2 in ASD[149, 150], and TSC1[141] or Scn1a[140] in epilepsy, all of which have shown interneuron-specific phenotypes. If disease-associated phenotypes can be recapitulated in dissociated culture, it would offer a drug discovery platform for these diseases. Specifically, showing that human stem cell derived interneurons recapitulate synaptic connectivity in culture would offer an inroad to personalized medicine in psychiatric diseases. Differentiation protocols exist for GABAergic cortical interneurons and their glutamatergic targets. This, then, offers a putative path forward for using our approach and our discoveries to develop a platform for psychiatric drug discovery. Development in psychiatric drugs has largely stagnated. This is in part because the current methodology for discovery is observing and exploiting side effects when these populations are given drugs not intended for treatment. Our findings offer an alternate path. Interneuron synaptic connectivity could be utilized as a means of testing molecular libraries. Moreover, if changes in connectivity are witnessed, we would also anticipate

functional changes to the networks that arise in culture. This offers parallel approach to screening that couples imaging with a functional readout. Thus, our findings provide a potential foundation for psychiatric drug development.

What would be required to turn this into a screening platform? Simply put, it is scaling. We show a small-scaled screen of interneuron connectivity following 24-hour treatment with IL13. This experiment utilized 24 wells and, given our microscope setup, took over 40 hours to image. Screening molecular libraries, which can number in the tens of thousands of compounds, will require faster imaging, faster image processing, and a larger supply of cells. It will also require internal controls and greater standardization, as well as an approach for batch correction that is able to remove systematic error while retaining biologically-relevant systematic differences. Experimental scaling will also require hardware scaling. We have already encountered issues in processing based on the sheer size of the dataset. The 62 million bouton number quoted above is not even from the complete set of images. It is from a subset of tiles that we identified and used to analyze targeting profiles in cortical columns. The full dataset at the time of writing is likely four times as large, and probably needs to be padded out with more biological replicates. The hardware requirements are staggering: using our current code, which stores all of the data in working memory, would require terabytes (yes, plural) of RAM. We are in the process of adapting our code to store the calculations on disk, but this comes at a major efficiency cost. Back of the envelope calculations put processing times for the pilot set of cortical columns at a week using this method. Many of these issues could be surmounted by moving our platform to cloud-based computing systems. Such a move, however, requires expertise, is a lot of work, and can be very costly. We also note that scaling will require considerable advances in processing power and efficient code. Optimizing these conditions will take considerable time, effort, and expense; but the payoff is

potentially huge. We hope that our efforts here provide the theoretical and practical groundwork to make this a reality.

5.3. The synaptic potentiality landscape

Typically, synaptic biology is described through anecdotes: descriptions of model systems used as examples to convey generalized concepts of synaptic targeting, diversity, structure, plasticity, function, etc. What follows is an attempt to synthesize current descriptive models and findings from disparate studies on synaptic target recognition and synaptogenesis into a conceptual framework that could be used to formulate a predictive model for these post-contact processes. It is by no means a complete or comprehensive account, and primarily serves to restate, extend, and constrain current ideas on these processes based on our findings in culture. Thus, one of the primary goals here is to define relevant variables and to describe how they relate to one another. We end by suggesting what we believe would be needed to formalize such a predictive model mathematically.

The impetus for this framework initially arose from an attempt to reconcile seemingly contradictory findings from *in vitro* and *in vivo* studies. As discussed above, our ability to recapitulate cortical interneuron *in vivo* synaptic targeting specificity *in vitro* contrasts sharply with prior attempts showing that the same cell types formed promiscuous synapses *in vitro* in dissociated culture. As eloquently stated by Richard Bunge,

> *If cultured neurons are found speaking a language we cannot at first comprehend, we may be well advised not to dismiss these signals as nonsense but rather to consider them as clues which we must learn to understand if we are to gain added insight into the mechanisms*

of development and functioning of neural systems. (Bunge, 1973)[115]

We work from the premise that observations of non-physiological synaptic phenomena *in vitro* abide by the same natural laws as their more physiological counterparts *in vitro* and *in vivo*, and therefore that these apparently contradictory findings can be explained by the same set of fundamental principles. We hope that attempting to reconcile these non-physiological results will shed light on aspects of the phenomena that are less pronounced in more physiological contexts, and therefore have been underrepresented in prior descriptive models of synaptic targeting and synaptogenesis.

5.3.1. Expanding the conceptual framework for synapse formation, diversity, and function

We concluded earlier that localized molecular expression patterns, molecular affinities, and the immediate environment work in concert, ultimately yielding preferential synaptic states. In other words, synaptic states are determined probabilistically from 'bottom up' biochemical interactions. Accepting this principle reframes the narrative of synaptic targeting, formation and function. From this perspective, neuronal connectivity exists in a dynamic, multidimensional space that ultimately synergizes into a potential energy landscape. That is, neuronal connectivity should be understood through a heuristic similar to Waddington's epigenetic landscape (Figure 5.1)[48, 125, 126, 155]. According to this framework, synaptic states are stabilized or

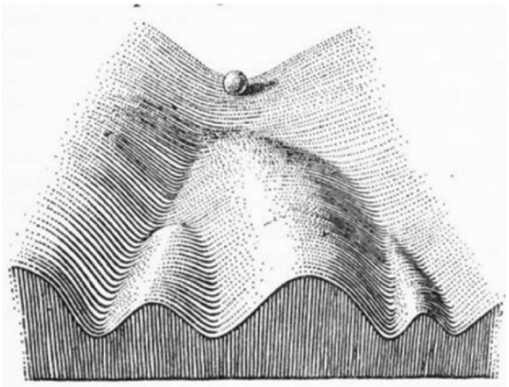

Figure 5.1 | Waddington's visualization of the potential energy landscape. The metaphorical epigenetic landscape proposed by Waddington is provides a conceptual framework for cell type specification during cellular differentiation. A similar framework is helpful for conceptualizing synaptic connectivity, differentiation, and diversity. From Waddington, C.H., *The strategy of the genes*. 1957: Routledge.

destabilized as energetic maxima and minima are created or eliminated. 'Synaptic state' here encompasses the full breadth of synaptic properties: existence, morphology, molecular composition, polarized/depolarized. The local minima represent localized states of stability: different minima are associated with different stable states of synaptic properties. The local maxima represent energy barriers between these ground states. They act as a buffer from environmental noise up to a certain threshold, but above this threshold, the synapse rapidly transitions into another stable state. Thus, the landscape can be understood as a probability function for each synaptic state. Unlike Waddington's epigenetic landscape, which is ultimately a descriptor of cell state, this synaptic potentiality landscape is extended in space and evolves over time in response to local conditions. Synaptic properties are embraced or discarded as the energy landscape moves from one ground state to another. The resultant energy barriers and spatiotemporal segregation synergize, allowing each synapse to operate semi-independently: a partially compartmentalized structure within a larger cellular network.

This framework adds a probabilistic, combinatorial, and spatially-distributed spin to Sperry's Chemoaffinity Hypothesis[127] specifically for the post-contact processes of synaptic development. Such a framework has been widely embraced for pre-contact processes, as exemplified by orthogonal chemical gradients acting sequentially[54, 156] and combinatorially[157-167] at relative (not absolute) levels[156] to form a molecular coordinate system for axonal and dendritic guidance (discussed in the introduction and reviewed in[60, 157, 168]). The lock-and-key model applies Sperry's Chemoaffinity Hypothesis to synaptic target recognition, proposing that synaptic specificity is achieved through molecular recognition between proteins on the surface of neurons[128, 169]. Specifically, Sperry concluded that:

> *"...cells and fibers of the brain and cord must carry some kind of identification tags, presumably cytochemical in nature, by which they are distinguished one from another almost, in many regions, to the level of a single neuron..." (Sperry, 1963)*[127]

According to this model, Sperry's 'identification tags' are conceptualized as cell surface proteins that act like "locks" and "keys," where the key (a specific protein or molecule) on the axon terminal of one neuron fits into the lock (a complementary receptor or molecule) on the dendrite of the target neuron[127, 128, 169-171]. The initial formulation of Sperry's Chemoaffinity Hypothesis proposed that each neuron expressed unique, genetically-encoded chemical tags that mediate target recognition. Based on the number of neurons within vertebrates being orders of magnitude greater than the size of the genome, this hypothesis has been refined into the adhesion molecule code (AMC) model, where synaptic target recognition is mediated by a combinatorial code of complementary synaptic adhesion molecules[128-131]. Similar to a zip-code, this model suggests that distinct, combinatorially-expressed adhesion molecules will dictate distinct aspects of regional, laminar, cellular, and subcellular connectivity using distinct lock-and-key combinations for ever-finer target specification[128, 169]. According to this model, the number of molecules required to specify a synaptic pair is likely large[128, 172], and the code may exhibit multiple levels of degeneracy[133, 173].

Although the AMC model has considerable experimental support[128, 171] and predictive power[137, 174, 175], it also has its limitations. By focusing only on presynaptic and postsynaptic adhesion molecules at the contact site, it does not include a role for the microenvironment around the putative synapse in regulating synaptic specificity, formation, and function. Many studies have shown, however, that the immediate microenvironment around the contact site plays a critical role

in regulating target recognition and synapse formation both *in vivo* and *in vitro*[62, 117, 119, 176-195]. That is, target selection is directed by a spataially-distributed network of adhesive and non-adhesive interactions in the neighborhood of the contact site[176, 177]. This microenvironment includes extrasynaptic regions of the presynaptic[178-181] and postsynaptic[182-184] membranes (i.e., those outside the contact region), as well as the intracellular[185, 186] and extracellular microenvironments around the contact site. The extracellular conditions include not only biophysical conditions that impact protein folding like pH, temperature, and ionic strength, but also diffuse molecules in the extracellular space[187, 188] and components of the extracellular matrix[62, 189-192] either directly in indirectly involved in target recognition, and factors that impact competition like cellular density[119, 193, 194] and diversity[117, 182, 195].

Contrary to the lock-and-key model, which suggests that binary expression or absolute concentrations of complementary molecular pairs regulate synaptic target recognition, combinatorial molecular mechanisms[124, 196, 197] and relative expression levels[133, 173, 176, 177, 198] seem to be the rule in target recognition[128]. Since target recognition and synaptogenesis are localized to a small contact site, the relative preference must be established through dynamic, spatially-distributed interaction networks on both the presynaptic and postsynaptic cells. This requires cross-talk within a growth cone or neurite simultaneously sensing multiple regions of the environment and 'choosing' the most stable interaction[176, 177, 185]. Studies also show that a localized cell-cell contact can prompt activity-independent molecular competition within the neighborhood of the contact site, initiating a distributed network effect that can alter the surrounding synapses. This effect has been shown postsynaptically in sparse dissociated cultures, where non-physiological, weak mismatched glutamatergic synapses on target cell AIS and dendrite were eliminated after contact by a GABAergic growth cone[117, 182]. In dendrites innervated only by glutamatergic cells, contact by a

GABAergic cell not only eliminated weak mismatched glutamatergic synapses, but also led to molecular and morphological changes in proximal, more physiological matched glutamatergic synapses that remained. Relative recognition and competitive specificity both suggest that molecules beyond the immediate contact site play a role in target recognition and synaptogenesis.

Finally, synaptic target recognition is just the first step within a dynamic and continuous procession of post-contact developmental processes. The lock-and-key model does not consider downstream anterograde and retrograde signaling processes critical in initiating synaptogenesis and establishing synaptic properties. The synaptic potentiality landscape proposed here tweaks this lock-and-key model by presenting target recognition as a dynamic, combinatorial, distributed, and probabilistic process involving not only the contact site, but also the immediate microenvironment around it. It proposes that the strength of the summed binding affinities at a contact site alone is not sufficient to explain the phenomena. Rather, this affinity is compared to other proximal contacts within the neighborhood through a distributed, dynamic network of molecular interactions. These distributed networks function on both the pre- and postsynaptic side, and through a process of activity-independent cis and trans molecular competition, determine whether to initiate synaptogenesis. Critically, it is this distributed initiation process that distinguishing synaptic target recognition from cellular recognition in general. It positions target recognition as the initial stage in a broader framework of post-contact synapse development, all of which can impact the wiring diagram[199] and physiological properties of the synapse[129, 130].

5.3.2. Topography of the synaptic potentiality landscape: reconciling paradoxes in neuronal connectivity

What determines the topography of this synaptic potentiality landscape? Like Waddington's epigenetic landscape (Figure 5.2), the synaptic potentiality landscape is ultimately based on physical constraints, molecular affinities, and chemical concentrations. Each molecular interaction is a binary decision between molecular states (e.g. to bind or not to bind) with an associated probability function based on intrinsic properties and environmental conditions. These individual probabilities sum to form an overall probability function for every possible state. This overall probability function is the potentiality landscape. Embracing such a framework challenges standing 'top-down' rules of synapse formation[31], but it also reconciles seemingly contradictory findings. To illustrate this point, let's reframe target recognition and synaptogenesis according to this framework.

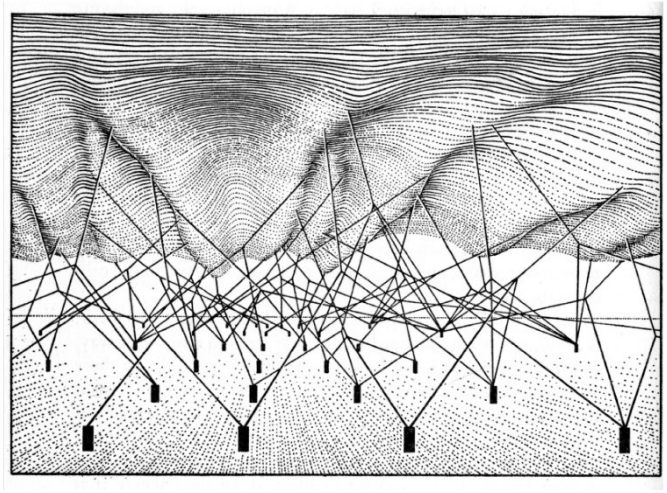

Figure 5.2 | The complex system of interactions underlying the potential energy landscape. Waddington postulated that the shape of the slope in the epigenetic landscape is determined by the complex interaction between several molecules, here represented by the guy-ropes, the strength of which is ultimately controlled by genes, represented by the pegs in the ground anchoring the guy-ropes. In the synaptic potentiality landscape, the shape of the slope is determined by a distributed network of molecular interactions within the neighborhood of the contact site. From Waddington, C.H., *The strategy of the genes*. 1957.

[31] More specifically, it reframes these 'top-down' rules as probabilistic patterns rather than hard and fast natural laws. Therefore, this reframing does not challenge the observations or conclusions, it cautions against their absolute applicability. In this respect, the 'top-down' rules mentioned here should be understood more as guidelines.

First, let's imagine a single neuron in isolation. What does the synaptic potentiality landscape of this neuron look like? Since synaptic target recognition requires physical contact between two cellular surfaces, the probability for synapse formation drops to zero beyond the membrane. Practically speaking, then, the synaptic potentiality landscape for any given neuron is spatially constrained to its cell surface. Since the neuron is localized in space, this landscape is also constrained to that space. Lucky for us this is a thought experiment, so we aren't limited by such pesky material realities. If we know the molecular composition of the cell and the environment, we could use that cell to measure the potentiality landscape across the environment. This would effectively map the synaptic potentiality landscape for that cell were it in contact with that part of the environment[32]. Thus, the synaptic potentiality landscape is practically constrained to the cell surface, but potentially extends throughout the environment.

What does that landscape look like on the cell surface? The specific potentiality function at any given location on the cell membrane is determined by the local chemical compositions and their intrinsic molecular affinities. Let's imagine that this neuron is growing on an extended surface that replicates the ECM. In such a case, the neuron will engage standard cell intrinsic processes of axodendritic polarization and compartmentalize as discussed in the introduction. Thus, the synaptic potentiality landscape is not homogenous: axonal regions will have a higher probability to form presynapses and dendritic regions will have a higher probability to form postsynapses.

Will it form a synapse? We know from studies discussed previously that the answer depends on the local environment. In general, there is a low probability a neuron will spontaneously form a synapse without direct cell-cell contact. Special conditions, such as electric field stimulation[200,

[32] Before brushing this off as a mere intellectual play, it is important to note that we *can* do something very similar via post-hoc image processing. For example, the presynaptic channels from one image could be registered onto the target channels of a different image. Thus, non-rigid image transformation could be used to generate a synthetic static map of this landscape.

[201] or high concentrations of diffuse synaptogenic molecules[202-205] can shift this probability, leading to the formation of contact-independent hemisynaptic structures. In such cases, this framework suggests that the specific hemisynaptic structures that do form, i.e., presynapse or postsynapse and their specific molecular composition, will be determined by the local molecular composition of the targeted compartment. That is, the topography of this landscape at any given point in space and time is a function of the local molecular composition of the cellular compartment and its immediate environment.

Studies have shown that neurons tend to avoid themselves based on repulsive protocadherin code of self-recognition[124, 128, 131, 206, 207]. Traditionally, this is viewed as a top-down rule of neuronal connectivity. Our model reframes this as just another energy barrier: neurons tend to extend their neurites away from themselves, thereby maximizing surface area, in part because this is a lower energy process than interacting with itself. We discussed earlier cases where neurons are forced to interact with themselves, thereby overcoming this energy barrier (and breaking the above rule). In such cases, these neurons will form autapses with mismatched presynaptic and postsynaptic components[119]. This finding breaks another top-down rule of neuronal connectivity. Specifically, neurons form presynaptic and postsynaptic components are matched based on the neurotransmitter: glutamatergic boutons with glutamatergic PSDs, GABAergic boutons with GABAergic PSDs[129, 130]. Our model reconciles this result: given the environmental context, mismatched autapses are a low energy state that would be outcompeted in a more physiological context. Specifically, these autapses are higher energy states than self-avoidance, but provided forced self-interaction, forming a mismatched autapse yields a more stable state than *not* utilizing the synaptic molecules for this purpose. From a biophysical perspective, the aggregate synaptogenic *trans* interactions at the cell surface simply outcompete the aggregate non-synaptogenic *cis* interactions,

leading to adhesion molecule clustering and recruitment of an otherwise unutilized reserve pool of 'mismatched' cytoplasmic presynaptic and postsynaptic molecules to the contact site. The clustered, presynaptic and postsynaptic molecules are held in place by a few weak interactions. In aggregate, this mismatched state is energetically more stable than maintaining these molecules in a diffuse state outside of a synaptic assembly. Again, given a different context, these mismatched autapses would be outcompeted by more energetically favorable states. Each point on the cell membrane, then, has a degree of probabilities associated with different states, with the ultimate state established by the molecular composition, immediate environment, and system noise.

Let's go back to our first case, where the neuron is not forced to interact with itself. Now let's add another neuron, and let's say that these neurons come into contact. What does the synaptic potentiality landscape for each of these neurons look like? The Neuron Doctrine postulates that, as a rule[33], information flow is polarized, with dendrites acting as inputs and axons acting as outputs[208]. In accordance with this rule, dendrites will generally have a high probability to form PSDs and a low probability for presynaptic boutons, and vice versa for axons. The synaptic potentiality landscape will thus also be generally polarized. Unlike the 'top down' postulates of the Neuron Doctrine, however, this synaptic potentiality landscape allows for exceptions. Dendro-dendritic synapses in the olfactory bulb[209, 210] and axo-axonic synapses such as AIS-targeting ChCs in the cortex or presynaptic bouton-targeting GABApre interneurons in the spinal cord[211, 212] break this rule of the Neuron Doctrine[208]. They are explained here by rare (but evolutionarily conserved) energetically favorable states that result from the specific molecular composition and compartmentalization within these cells. Given our two neurons, they *could* form an axo-axonic, dendro-dendritic, or axo-dendritic synapse provided favorable conditions. Less stable states will tend to

[33] The 'law of dynamic polarization'

be outcompeted by more stable ones, and thus occur at lower frequencies. Given that there are only two neurons in this interaction, the low level of competition will enable low affinity molecular interactions during target recognition, and so also a higher number of low stability states. More specialized (i.e., molecularly distinct) synapses will tend to display greater stringency in targeting specificity than less specialized ones. Given this framework, it is perhaps unsurprising that such highly specialized synapses show stringent specificity[37, 211], whereas less specialized ones display a hierarchy of preferences *in vivo*[213-215]. The alternate synapses that such highly specialized cells would form are simply outcompeted by neurons in their surroundings. Thus, the dichotomy between stringent specificity and hierarchy of preference can be reframed as degrees of stability in the context of competition: a hierarchy of preference still exists for neuron types that display 'stringent specificity,' these less-stable connections are simply outcompeted by surrounding cells within a physiological context.

Let's focus on the more common and better understood case of cells that preferentially form axo-dendritic synapses in a physiological context, such as a glutamatergic pyramidal cell of the cortex. And let's say that it comes into contact with the target cell. Given that we have only two cells in our thought experiment, molecular competition is very low, and most adhesion molecules on the presynaptic and postsynaptic cell's membranes (i.e. those outside of the contact site) will have defaulted to low affinity cis interactions (or no interactions at all). During the stage of synaptic recognition at the contact site, mutually attractive adhesion molecules will lower the energy and mutually repulsive ones will increase it. These affinities act cooperatively and competitively to determine the ultimate strength of the trans interactions. That is, a combinatorial code of recognition is at play. But unlike the standard AMC model that focuses on the contact point alone, we must also take into account other *cis* and *trans* interactions in the neighborhood. Our contact site

will be competing for a limited pool of molecules to stabilize the interaction, and so is operating in a distributed, competitive manner across both membranes. The contact site is competing not only with proximal *cis* interactions (an unengaged reserve pool), but also with *trans* axo-dendritic, dendro-dendritic, and axo-axonic connections with partially overlapping codes. Ultimately, these molecular interactions will default to the lowest energy state. If the initial energy barriers to break existing connections can be overcome, and the *trans* interactions at the contact site are lower energy than the competing *cis* and *trans* ones, the interactions at our contact site will stabilize. Given that our presynaptic cell preferentially targets dendrites, we known that these *trans* interactions will be much stronger at the target cell's dendrites than at other subcellular regions. But this does not preclude weaker *trans* interactions at other subcellular regions of the target cell, such as the soma or AIS, from outcompeting even weaker *cis* or *trans* interactions in that neighborhood. Indeed, studies have shown that glutamatergic cells from the hippocampus will form non-physiological synapses onto soma and AIS of target cells in sparse cultures[117, 182]. Moreover, these weak glutamatergic connections will be outcompeted (and ultimately eliminated) by GABAergic cells forming more physiological connections. Note that although these GABAergic connections onto the AIS are *more* physiological, it is highly unlikely that they were formed exclusively by AIS-targeting ChCs. Rather, as with our last example, the hierarchy of preference for target recognition is on display. In this case, the low level of competition is revealing unstable states arising from low affinity interactions that otherwise would be outcompeted in a more physiological context. Thus, the probability function for any synaptic state is modulated in part by activity-independent cellular competition distributed across the membrane.

Let's say that we got lucky and our two cells have not only come into contact, but also formed a stable *trans* interaction. We've determined that they *can* form a synapse, but *will* they? Simply

put, if the energy landscape at the contact site favors synapse formation, a synapse will form spontaneously. The current view suggests that synapse initiation is determined by retrograde and anterograde signaling transsynaptic signaling, which transforms the contact site into a nucleation site for synaptogenesis[129]. Specifically, this nucleation is thought to occur via signaling from synaptic organizers at the membrane (typically thought to be adhesion molecules) into the cytoplasm, thereby recruiting synaptic components that establish a stable, functional synaptic structure. Importantly, it makes no difference whether nucleation progresses through a few strong interactions or many weak ones, the result is the same either way: a dynamic, evolutionarily conserved process of molecular recruitment and assembly that yields an even more stable state, a synapse. We agree such signaling from the membrane and molecular recruitment is critical for initiating synaptogenesis, but suggest that it is only part of the picture. To explain this position, let's go back to pre-nucleation, where we have a stable *trans* interaction between our two cells of interest. Whether or not the energy landscape at the point of contact favors nucleation depends not only on the molecular interactions at that location, but also those in the surrounding neighborhood. As studies have shown, initiation of synaptogenesis is determined not simply by transsynaptic communication, but also by a distributed network of molecular interactions within the proximity of the contact site[117, 176, 177, 182, 187, 188]. These interactions include the distributed *cis* and *trans* interactions on the membrane described before, as well as intracellular and extracellular interactions. For example, pre- and postsynaptic components spontaneously cluster (form condensates) in a concentration-dependent manner even in the absence of cell-cell contacts, suggesting that these molecules exist in equilibrium between condensed and diffuse states[131, 216-229]. Adhesion molecules interacting in *trans* will stabilize clusters of pre- and postsynaptic components, shifting the balance toward the condensate state[228]. The total quantity of these nucleated components is limited, however. Thus,

higher reserves outside of the contact site will increase the probability of recruitment, whereas lower reserves will decrease it. That is, nucleation is a competitive process: synaptogeneisis is more likely to occur at low affinity contact sites when there is less competition for the requisite molecules in the neighborhood. As with target recognition, increased competition will increase specificity. Indeed it is difficult to conceptualize how his synaptic target recognition and synapse formation could operate via relative (not absolute) expression levels [133, 173, 176, 177, 198] unless they function through competitive network interactions distributed far beyond the contact site. Thus, initiation of synaptogenesis is determined not only by the strength of signaling from the contact site into the cytoplasm, but also by the synaptogenic potential at the contact site (a function of the neighborhood).

These conclusions suggest that distinct regions on a neuron will have distinct synaptogenic potentials. Moreover, since distinct synaptic states have distinct requisite molecules, each state will have a distinct synaptogenic potential determined by the neighborhood concentrations of those requisite molecules. The same two neurons, with ostensibly the same presynaptic and postsynaptic components, can therefore form an array of molecularly distinct synapses as a result of differential synaptogenic potential and varied levels of molecular competition in the neighborhood. The discussion of component clustering above shows how activity-independent molecular interactions in the cytoplasm and membranes in the neighborhood can impact the synaptogenic potential. Synaptogenic potential can also be modulated by extracellular molecules, including diffusible secreted factors acting over long distances[187, 188, 230] or locally secreted extracellular scaffolding proteins from neuronal and non-neuronal cells[189-192, 230-236] or via network proximal activity-dependent processes associated with synaptic plasticity (e.g. homo- and hetero-synaptic crosstalk)[178-181, 237-239]. Defining the molecular mechanisms that determine this synaptogenic potential will be critical

not only for understanding synapse formation, but also uncovering the hierarchies within the combinatorial code of synaptic recognition. Importantly, based on the above discussion, 'synaptogenic potential' should not be understood as a simple measure applicable to all synapses. Rather, similar to the hierarchy of preference, it is a multidimensional metric defining the probabilities to form different synaptic states at that location. At a cell-cell contact, each membrane will have distinct synaptogenic potentials that will synergistically determine the probabilities for all possible synaptic states at the contact site. Critically, stable contacts can exist without acting as nucleation sites, thereby establishing a reserve pool of contact sites for activity-dependent synaptic plasticity[178-181, 237-239]. Thus synapse formation and elimination are accompanied by nucleation and denucleation, respectively, unstable transitionary states wherein the neighborhood is dynamically adjusting the reserve pool of synaptogenic molecules and so also the synaptogenic potential within it.

Let's say that our luck still hasn't run out, and a synapse forms at our contact site. We stated earlier that the synapse is a low energy state between unstable transitionary states. Why is the synapse more stable? The heightened stability is due to evolutionarily conserved, self-reinforcing interactions that dynamically maintain low energy states. The synaptic scaffolding condensates discussed in the introduction are good examples of such energetically favorable states[216, 217, 219, 221-223, 225]. Here, a concentration-dependent fluid-fluid phase transition leads to the formation of a subcellular compartment consisting of high concentrations of highly specialized pre- or postsynaptic components. The stability of these compartments is evident from studies where the application of an environmental stimulus leads to the formation of 'hemisynaptic' assemblies (pre- or postsynaptic structures without their post- or presynaptic counterparts) independent of cell-cell contacts[200-203]. These structures persist for hours after the initial stimulus is removed, a testament

to the self-reinforcing nature of the reciprocally attractive molecular interactions within (and repulsive interactions at the surface of) the condensate. When such condensates form at a cell-cell junction, the result is two self-sustaining subcompartments bridged by a network of stable adhesion molecules. Stabilization on one side leads to stabilization on the other, a virtuous cycle that that is not only energetically favorable, but also self-reinforcing.

So we have our synapse. Compositionally, what does it look like? A synapse consists of two highly compartmentalized networks of molecular interactions on either side of the synaptic cleft, connected by a network of transsynaptic adhesion molecules. In total, a single synapse will be comprised of hundreds of distinct, highly specialized molecules[240]. These components can be broken down categorically by associated functions. All presynapses contain exocytic and endocytic components, proton loading mechanisms vesicle release machinery, active zone proteins. All postsynapses contain neurotransmitter receptors, scaffolding molecules, and signal transduction mechanisms. Proteomics studies suggest all synapses share certain generic components, such as synaptic vesicle vesicular ATPases (vATPases) and proteins mediating synaptic vesicle endocytosis[240]. Such ubiquitous synaptic molecules appear to be the exception rather than the rule, however. Research has revealed a remarkable diversity of chemical synapses within the CNS[241-243]. Given that over a thousand synaptic molecules have been identified, the compositional diversity of these molecular networks is potentially staggering. Perhaps unsurprisingly, synaptic proteomes from more closely related neuronal populations displayed greater levels of compositional overlap. That is, much of this diversity can be explained by the molecular identities (expression profiles) of the presynaptic and postsynaptic cells. Two distinct types of presynaptic neurons will form morphologically, molecularly, and functionally distinct synapses on the same postsynaptic cell. Similarly, the same presynaptic cell can form distinct synapses onto two (or more) different types

of postsynaptic cells (Our data here, see also[79, 80, 195, 242]). Molecular expression profiles constrain composition, but they do not dictate it entirely. A single presynaptic cell can form multiple different types of synapses onto the same postsynaptic cell (Our data here, but see also [117, 182]), indicating that the specific composition of the pre- and postsynaptic structures ultimately depends on the molecular composition of (and competition within) the neighborhood and local environment around the contact site.

Even more critical than stability or composition is functionality. Subcellular compartmentalization adjusts local molecular concentrations, altering the energetics of reactions. Pre- and postsynaptic assemblies create highly specialized microenvironments that increase efficiency of certain reactions and decrease the efficiency of others. That is, the same molecules will establish different network connections and function differently in a condensed synaptic assembly than in a diffuse cytoplasmic state[219, 229]. Moreover, the spatial segregation of these molecular networks enables them to function as partially isolated entities that are nevertheless integrated within a larger whole. The relationship between functional segregation and integration within a network has been formalized mathematically as '*complexity'*, and serves as a metric for the degree of independence between small and large subsets of a system[244-250]. Such complexity exists not only within the molecular networks comprising the synapse itself, but also between those synaptic networks and its local neighborhood. We discussed activity-independent homo- and heterosynaptic cross-talk above. We know also know from research on research on synaptic plasticity that activity-dependent events at one synapse will affect other synapses in the neighborhood[178, 179, 181, 238, 239, 251, 252]. Given the compositional differences between synaptic types and the spatially distributed nature of synaptic network effects, it is perhaps unsurprising that molecularly similar synapses (homosynaptic) ones will abide by distinct coordinated plasticity rules than dissimilar (heterosynaptic) ones[180,

[237, 238], (reviewed in [178, 179, 181, 239]), or that molecularly distinct subcellular compartments (e.g., proximal vs distal dendrites) abide by distinct plasticity rules[238, 251-254]. Indeed, the proximity relationships defined by these plasticity studies may help to elucidate the spatial extent of these distributed network effects impacting events at a given contact site. Clarifying the causal boundaries of this neighborhood will be critical for formulating a mathematical model to predict synaptic state transitions.

5.3.3. Defining synaptic state: conceptual and practical considerations

We mentioned earlier that 'Synaptic state' encompasses the full breadth of synaptic properties: existence, morphology, molecular composition, polarized/depolarized, physiological profile, etc. We propose that any specific combination of these properties represents a distinct synaptic state. Thus, 'synaptic state' is a multidimensional, quantitative descriptor associated with any point on the cell (or potentially any point in space). From this perspective, every synapse is unique, and synaptic diversity is the rule, not the exception. Conceptually, the complete set of all possible combinations encompasses the complete set of possible synaptic states.

In reality, we know that certain properties are mutually exclusive. That is, due to fundamental constraints, the observed synaptic states represent a small subset of all potential states. For example, gephyrin condensates exclude glutamatergic postsynaptic molecules and PSD95 condensates exclude GABAergic postsynaptic molecules[216, 217, 219, 221-223, 225] (c.f. introduction). This mutual exclusivity extends far beyond presynaptic neurotransmitter enzymes and postsynaptic scaffolding molecules. Using fluorescence activated synapse sorting (FASS) of synaptosomes coupled with mass spectrometry (FASS/MS), Van Oostrum and colleagues found very little overlap between the proteomes of GABAergic and glutamatergic synapses (~1% of the 1134 synapse-enriched pro-

teins)[240]. This segregation is not simply due to transcriptional differences, but also post-translational compartmentalization. When both glutamatergic and GABAergic neurotransmitter synthesis enzymes are expressed in the same cell, these components segregate into distinct presynaptic structures with distinct functional properties[195, 255, 256]. Pre- and Postsynaptic structures with high concentrations of both glutamatergic- and GABAergic-specific molecules are high energy, unstable states. Such states either spontaneously collapse or never form in the first place. Specific combinations that we *do* witness represent low energy states: distinct stable combinations of molecular, morphological, and functional properties. Provided favorable conditions, specific synaptic structures with distinct properties will form spontaneously. Ultimately, the specific state that arises is determined by the molecules available, their intrinsic binding affinities, the local conditions, and the amount of noise in the system. In other words, the result depends on the expression profiles of the presynaptic and postsynaptic cells and the immediate environment around the synapse.

We also know that certain properties tend to occur together. For example, co-expression of specific calcium-mediated vesicle-binding (Syt2) and calcium-buffering (PV) proteins enable rapid synaptic communication in fast-spiking interneurons[257-265], as discussed in the introduction. Indeed, this observation can be extended to entire molecular networks of co-enriched molecules[240, 266]. Van Oostrum and colleagues, for instance, identified multiple modules of highly co-enriched proteins within both GABAergic and glutamatergic synapse populations[240]. It is important to emphasize here that biological systems operate as degenerate networks[249, 267-271]: many distinct molecular compositions *can* lead to phenotypically similar synapses. This is most clearly depicted by clustering on a UMAP. Within this space, every possible state is represented, but only specific

combinations are realized. Although each individual synapse is unique, there are specific properties that are more likely to co-occur than others. According to this formulation, then, the total number of unstable synaptic states far exceeds the number of stable ones.

Finally, from a practical standpoint, we do not have the ability to measure the full breadth of properties simultaneously for any given synapse. Our insights into these properties are fundamentally limited by the tools we are using. An incomplete dataset can still be useful, however. We show in this work how visualization of four molecules enabled us to extract over 430 unique metrics. These 'properties' comprise a partial view of the multidimensional landscape these synapses inhabit. Dimension reduction and clustering analysis suggested that each of these synapses exists as one of nine classes. This clustering analysis provides us with a picture of the network's degeneracy: the ability of structurally different elements to yield the same output[249].

5.3.4. Biophysical assumptions and their implications for neuronal connectivity

This framework makes several assumptions based on fundamental biophysical and genetic principles. To be sure, none of the following assumptions are controversial, and most biologists will be familiar with each of them. Since they form the basis of this framework, however, it is important to state them clearly. We spell them out here and discuss their implications regarding neuronal connectivity.

First, this framework implicitly accepts that molecules will be repurposed in different contexts if it is energetically favorable to do so. Fundamentally, molecular repurposing implies *multifunctionality*, the ability of the same element to perform multiple functions or achieve multiple outcomes within different contexts. That is, a molecule's function is context-dependent, defined by its interactions within the molecular network and local conditions rather than some intrinsic and unchanging property of the molecule itself. This multifunctionality allows for genetic efficiency:

this versatility allows the system to dynamically adjust to changing conditions and perform a wide range of functions without requiring specialized components for each task. The same gene product can perform different functions in the same cell at different developmental stages[272-275]. The same molecule can also perform different functions at the same developmental stage in different cells[57, 61]. We discussed examples of both in the introduction. Context-dependent functionality also implies that removing such components can have context-dependent, heterogeneous effects. We discussed an example of this in the introduction, where deletion of Netrin-G2/NGL-2 had distinct effects depending on their pre- or postsynaptic expression[276-281]. These represent distinct synaptic states; deletion of the same component in these different contexts differentially destabilized each of these states in different ways.

Second, it implicitly embraces molecular functional redundancy[34]. Whenever functionally redundant molecules are available and their involvement is energetically favorable, they'll be incorporated. This functional redundancy adds robustness to the system at the expense of efficiency. But it also conceptually reframes the underlying molecular interactions. They do not operate hierarchically, as previously proposed[129, 130, 208]. Rather, they function as degenerate networks, where structurally different elements can yield the same output[249, 267-271, 282, 283]. The effect of altering these elements (e.g., by removing a component) or the environment depends on how central it is to the stability of the synaptic state(s). Taking out a central node could lead to complete collapse of most, if not all synapses. Functionally redundant molecules buffer these effects. Removing a redundant component could lead to instability, but strictly speaking, no phenotypic change. We

[34] We are here using the more familiar phrase 'functional redundancy' to evoke the more precisely defined concept of '*degeneracy.*' Strictly speaking, *redundancy* occurs when the same function is performed by *identical* elements, whereas *degeneracy* is the ability for *structurally different* elements to perform the same function in a context-dependent fashion[249, 269]. Note the formal relationship between *degeneracy*, defined here, and *multifunctionality*, the ability of *structurally identical* elements to perform *different* functions within different contexts. Of note, in physics degeneracy refers to systems that exhibit multiple discrete or distinct energy values or states[269].

discussed examples of this in the introduction, with Semaphorins/Plexins in the hippocampus[284-287]. Gene products do not have to be homologous to be functionally redundant. This is exemplified by the roles of the evolutionarily unrelated adhesion molecule *fasciclin* and the cytoplasmic tyrosine kinase *abelson* in drosophila axon pathfinding[288]. Deleting either gene individually causes no significant abnormalities, but the absence of both leads to major defects in growth cone guidance, rerouting commissural axons and formation of ectopic connections. We also witness this phenomenon with disease mutations. The genetic terms 'penetrance' and 'expressivity' ultimately describe probabilistic relationships between disease mutations and disease states. In synaptic diseases, incomplete penetrance and/or expressivity are based on functional redundancy and synaptic state stability. Moreover, destabilization of one synaptic state may result in the adoption of another. This is witnessed with the complete removal of all neurexins, which has little effect on gross synapse numbers and instead leads to diverse alterations in synaptic functional and morphological properties[129, 289]. When an intervention prompts large-scale state changes across multiple synaptic populations, this is known as "synaptome reprogramming"[1, 241, 290]. Combined with the heterogeneous molecular composition of synapses, state degeneracy explains the observation that partial phenotypes are more common than complete loss in knockout models[129, 130]. We discuss this in greater depth below.

Third, the potentiality landscape is spatially extended. The synaptogenic potential at a contact site is determined through a distributed network of molecular interactions that extends far beyond the site of contact. This blurs the typically sharp boundaries for what we call a synapse: a comprehensive understanding of a synapse requires not only definition of the molecules within the presynaptic and postsynaptic compartments and transsynaptic interactions, but also a comprehensive characterization of the neighborhood within which it exists. This assumption implicitly relates to

subcellular compartmentalization. Many molecules will spontaneously segregate into spatially distinct regions of the cell after reaching specific threshold concentrations. Compartmentalization is a fundamental principle of cellular biology: heightened local concentrations alter the energetics of reactions, thereby increasing efficiency. Moreover, these spatially segregated (but nevertheless connected) molecular networks function as partially isolated entities. As noted above, this relationship between interconnectivity and independence of nodes within a network has been formalized mathematically as '*complexity*'[244-250]. It accords with data showing that individual synapses can operate independently, while still being integrated into the wider cellular network. This interconnectivity and semi-independence of molecular networks extends beyond simply synapses. A fundamental distinction can be made here between molecular networks within the cell (cell intrinsic) and those in the environment (cell extrinsic). Clearly, these networks are themselves connected, and so the (ultimately false) dichotomy between 'cell' and 'environment' is based on degree of network independence rather than two fully isolated entities. Cell intrinsic processes shape the environment, and vice versa. Truly independent entities are fully outside each other's sphere of influence. Defining these spatial limits of biological causality *in vivo*, as done by Francis & Palsson for diffuse molecules *in vitro* [53], will help to constrain this complexity and identify the operative 'environment' for any given cell.

Fourth, this formulation is inherently dynamic. The dynamic nature of synapse formation implies that the relative impact of any individual component on the potentiality landscape is not static. Rather, an adhesion molecule important for establishing that initial pseudo-stable nucleation state during recognition may be inconsequential once the synapse is established, as has been witnessed experimentally[137, 291] (reviewed in [292]). This fits with well-established models of 'critical periods' during development.

Fifth, it also allows for nested state changes, where one aspect of the system can change dramatically, while leaving other aspects relatively unperturbed. An illustration of this concept from the potentiality landscape is a deep valley with a small hill in the center. The local minima on either side of the hill are distinct states. An example is polarization and depolarization. The synapse is easily able to jump between these two states, which in turn prompts distinct molecular interactions (Ca2+ influx, vesicle binding, neurotransmitter release, etc.). But the overall structure and stability of the synapse itself (represented by the deep valley) is not changed.

As noted above, none of these principles should be controversial or unfamiliar. Biologists call upon functional repurposing whenever they discover a new function for a biomolecule; they invoke functional redundancy when a loss of function intervention does not yield a loss of phenotype. In other words, these concepts are used retrospectively. As such, they are currently prospective only insofar as they predict that biological systems can behave unpredictably. We propose that the synaptic potentiality landscape provides a framework to utilize these concepts prospectively: it provides the theoretical basis for a mathematical formalization to predict stable synaptic states based on cellular composition and context.

5.3.5. Nothing is necessary, everything is sufficient: embracing nomenclature for a probabilistic, degenerate framework

At its heart, this framework is probabilistic, not determinist. Truly embracing this concept forces us to evolve our nomenclature. It suggests probabilistic and quantitative descriptions of phenomena rather than the strict binaries of *necessary* and *sufficient* that are intrinsic to deterministic models. To explore this concept further, it's helpful to understand how the current deterministic language is insufficient to describe the phenomena. This will reveal why it is necessary to abandon it in favor of a more accurate, probabilistic nomenclature.

Loss-of-function and gain-of-function experiments are staples of modern biology. The common interpretation of loss-of-function experiments is that they reveal molecules to be *necessary*, whereas gain-of-function experiments show them to be *sufficient*. Descriptions of specific molecules being *necessary* and *essential* for specific connections are common in the literature. A careful assessment of the data, however, will reveal that most of these loss-of-function interventions result in partial phenotypes: *decreased* synapse numbers on the target instead of full elimination[61, 137, 293-300]. Strictly speaking, a partial phenotype indicates that the removed molecule is *not necessary*; rather, its loss *decreases the probability* of forming the stereotyped connections. The synaptic potentiality landscape naturally lends itself to such descriptions. It also offers a framework for understanding those 'robust' synapses that remain. They can be explained by molecular redundancy, repurposing of the lost molecule, synapses outside of a critical window, or simply stochastic noise. It is also not at odds with truly deterministic phenomena. It simply reframes them as specific limit cases of the probability function, where the probability is expressed in binarized 1's and 0's. In this way, it allows us to reconcile the data while providing accurate language to describe partial phenotypes.

We can also look at this from the other side, namely gain-of-function experiments. Artificial synapse assays, the standard gain-of-function experiment in the connectivity field, suggest it is relatively easy to create a synapse. Indeed, many synaptic adhesion molecules induce synapse assembly[114, 130, 132, 291, 301-306], including molecules where knockout animals show little to no change in synapse numbers[303, 305, 307]. Viewed from a deterministic lens, these results are puzzling[129, 130]. The problem becomes even more paradoxical with studies that show synapse formation can be induced independent of complementary adhesion molecules. Examples include postsynapse formation by local application of diffusely secreted molecules[202, 203] and electrical

field stimulation[200, 201], or presynapse formation by interaction with polylysine coated beads[204, 205]. Of note, the resultant 'hemisynaptic' assemblies are morphologically and molecularly identical to standard pre- or postsynaptic structures. Regarding a role for adhesion molecules in synapse formation, ultimately these experiments suggest that *none are necessary* and *many are sufficient*.

Given a deterministic framework, these findings are paradoxical[129, 130]. A probabilistic, degenerate framework makes sense of these results. Such 'sufficient' conditions simply result in *an energetic state that favors synapse formation*. In other words, the nucleation step proposed to be driven by synaptic organizers[129, 130] can be achieved through any number of favorable conditions. Most of these nucleation events are driven by the neuron encountering complementary synaptic adhesion molecules on a stable surface. Some of them are completely independent of such molecules. Either way, the result is the same: increased probability to form a synaptic assembly.

At first glance, this reformulation may appear to be in opposition to the adhesion molecule code model of synaptic recognition. It is not. Rather, it generalizes this model and makes it dynamic. It proposes that cells (and even subcellular regions) will have optimized codes depending on their current state and molecular composition. That is, cells do not adhere to a singular, static synaptic adhesion molecule code. The interaction probabilities and their relative importance are dynamic and spatially extended; they adjust in accordance with the broader cellular and molecular context[156, 173, 187, 188, 295]. On the spatial front, this accords with studies showing that relative, not absolute, adhesion molecule concentrations control targeting preference[54, 128, 131, 133, 156] or that gradients of diffusely-secreted molecules can impact the probability of synapse formation[187, 188]. On the temporal front, this accords with well-established concept of developmental critical periods for synapse formation and maintenance[137, 291] (reviewed in [292]). Molecules important to stabilize cell-cell interactions during synaptic target recognition may be less important after synaptogenesis,

when molecular crosstalk and network effects have stabilized the synapse[137, 291]. The probability function associated with any individual code will change depending on the amount of cellular competition, for instance in dissociated culture versus intact tissue.

5.3.6. Turning a heuristic into a predictive model

We provide a conceptual framework here to understand synapse formation. The framework reconciles a large number of fringe cases of synaptic connectivity that are unexplained by current models. At its heart it is probabilistic, and arises from molecular interactions between the presynaptic cell, the postsynaptic cell, and the local environment. It was inspired by Waddington's epigenetic landscape but ultimately is based on a wide body of literature on neuronal connectivity, biochemistry, physical state modeling, and network modeling. It suggests that synapses operate as incompletely isolated degenerate networks[249, 267-269].

As it stands, however, this is a heuristic. Ultimately the utility of a model is based on its predictive power. In order for this framework to be formulated as a predictive model, it will have to be formalized mathematically. There are four requirements here:

1. a comprehensive definition of stable (degenerate) synaptic states,
2. the underlying variables that determine them,
3. an experimental system to test them, and
4. a mathematical theory to tie these together.

Regarding the first two requirements, a more comprehensive molecular and functional characterization of individual synapses, and the conditions within which they occur, will help to define the model's constraints and enable comparison, transference, and testing. This is not a one lab job. It will require standardization and centralization of definitions and datasets from many different

sources and experimental designs. NCBI/NLM has established a navigable repository for gene expression datasets. The same needs to happen for multidimensional imaging data in neuroscience. These data must incorporate a systematic accounting of intrinsic (the molecules probed and the measures utilized) and extrinsic (environmental) variables. The latter includes not only treatments and mutations, but also contextual information, such as the tissue source (e.g., age and region dissociated for primary culture) and end stage sample data (e.g., culture age, fixation protocol, etc.). Once established, this archive should be integrated with ever-growing databases of neuronal connectivity, cellular expression profiles, molecular interaction networks, and molecular affinities. As discussed above, it is not enough just to know *what* is expressed, we also have to know *where* those molecules localize. A more complete understanding of subcellular localization profiles would greatly help here. This in turn will require more comprehensive annotation of subcellular compartments beyond the synapse. Such an ecosystem will enable computational researchers to assess spatial correlations and formalize synapse identities.

At the end of the day, molecular profiles of synapses and their immediate environments will be required. A multidimensional imaging approach like ours, with a more comprehensive panel of synaptic, subcellular, and cellular markers will be critical to achieve this goal. This will require advances in multiplexed imaging of proteins. We discussed progress on this front, as well as perceived barriers, earlier in the discussion. It is clear, however, that we are still years away from an imaging approach that parallels singe cell RNAseq at the protein level. In the short term, however, we *can* make probabilistic statements about synaptic states based on incomplete knowledge of molecular composition. We show here, for instance, how the distribution of synaptic targeting shifts with increasing concentrations of IL-13. We can formulate this relationship as a probabilistic statement between presynapse, postsynapse, environment, and noise. In this case, ‘synaptic state’

is simply the bouton targeting class, and the stability of each is reflected in their relative proportions and classification probabilities. The 'underlying variables' are admittedly incomplete (and indirect) measures of molecular composition: antibody staining and fluorescent protein expression. Nevertheless, we are able to conclude that changes in the environment (IL-13 concentrations) impacts synaptic state stability. A similar conclusion can be reached from the different Cre driver lines, where we categorically change the molecular composition of the presynapse by altering cellular identity. Unsupervised classification yielded insights into unexpected differences in synaptic states based on presynaptic cell identity. Clearly more comprehensive knowledge of presynaptic and postsynaptic molecular composition will enable finer distinctions in synaptic state classification. Nevertheless, this approach helped us to uncover previously unrecognized differences in axon growth strategies of BCs and SstCs. That is, we are able to make significant advances even with an incomplete picture. The unmeasured properties, then, constitute noise: aspects of the system to which we are blind, and thus a source of error in any predictive system. Incorporating this unmeasured 'noise' into a formalized model as a variable will be critical to quantifying *how predictive* it is.

This last point brings up the third requirement: an experimental system for testing. We propose that tissue culture offers a means to massively increase throughput and therefore also test various models and conditions. We discussed at length the potential to use our system for adhesion molecule screening and drug screening. These are examples of varying synaptic molecular composition and environmental conditions, respectively. In both cases, the primary outcome is a probabilistic statement about synaptic states.

Finally, regarding the mathematics. We show in this study how multidimensional analysis can provide information on stable synaptic states. Measuring synaptic properties and defining stable

states is only the first step to formulating a predictive model, however. A predictive model will require an understanding of the underlying network's degeneracy, i.e., how distinct combinations of synaptic properties and environmental conditions can yield similar synaptic phenotypes[249, 267, 270, 271, 283]. There is a rich mathematical theory for modeling degenerate networks that could be applicable here[269, 308-310]. Neurons and the networks they create are complex systems: they consist of multiple incompletely isolated networks that operate with semi-independence[246, 247, 249, 250, 308-311]. Incorporating this notion further complicates the picture. Complex systems[246, 247, 250] and their relation to degeneracy[249, 269] have been modeled mathematically for anatomically-connected brain regions. It remains to be seen whether these concepts can be reformulated for molecular interactions in 3D space, where structural connectivity between the operative entities is fundamentally distinct. Finally, we discussed how both activity-independent (molecular) and activity-dependent (plasticity) competition can affect synaptic connectivity. This suggests that there are at least two distinct but related forms of synaptic competition that must be incorporated into the model. Assuming *a priori* winner and loser identities can simulate processes occurring downstream of this identification, but it cannot identify how synapses become winners or losers in the first place[312]. Mathematical frameworks to detect (rather than assert) winners and losers have been formulated for *cellular* competition (survival vs death)[312], and could be a guide to a mathematical framework for *synaptic* competition. That said, cellular and synaptic competition are fundamentally distinct processes. For cellular competition, homotypic conditions can be used to establish a baseline of survival free from competition, which is in turn compared to heterotypic (competitive) conditions that forces winners and losers. Synapse formation differs insofar as heterotypic conditions not only result in the loss of specific synaptic states, but are oftentimes a prerequisite for the *emergence* of others. Moreover, current formulations for cellular competition are limited to just two populations.

There are likely hundreds of distinct, stable synaptic states, however, which are in part determined through distributed network effects within the microenvironment. Thus, reformulating the current framework would require not only extension beyond two competing identities, but also incorporation of context-dependent spatially extended conditions as parameters. Although difficult, such a framework would help us to determine which synaptic states actually compete with one another, while also identifying the factors that drive synaptic competition and synaptic state acquisition. In other words, it would help to map the synaptic potentiality landscape itself.

Ultimately, we show a static picture of synaptic states. Synaptic development is a dynamic process, however. Synapse formation may be initiated by direct contact between a presynaptic and postsynaptic cell, but the events before and after have a bearing on the synaptic phenotype that ultimately arises. Even after maturation there is tremendous plasticity and turnover. Understanding how these states evolve over time is of great interest. But it is also a difficult problem. There is a rich mathematical theory of dynamical systems that could be useful here. Indeed, such approaches have been proposed and used to model cellular differentiation[48, 125, 155]. Whether these approaches are applicable to a spatially-extended 3D environment is unclear, however. It is also unclear whether such approaches can be synthesized with the previously mentioned methods to model degeneracy and complexity. It will be exciting to see whether these mathematical frameworks can be wrangled into a formalized model with real predictive power. At the end of the day, however, such a formalized model will have to be tested and validated using empirical data. We hope that the system we presented here provides the groundwork for these advances. [35]

[35] Thank you for reading! I genuinely appreciate your dedication and attention.

Bibliography to Chapter 5: Discussion

1. Zhu, F., et al., *Architecture of the Mouse Brain Synaptome.* Neuron, 2018. **99**(4): p. 781-799 e10.
2. Cizeron, M., et al., *A brainwide atlas of synapses across the mouse life span.* Science, 2020. **369**(6501): p. 270-275.
3. Curran, O.E., et al., *A single-synapse resolution survey of PSD95-positive synapses in twenty human brain regions.* European Journal of Neuroscience, 2021. **54**(8): p. 6864-6881.
4. Bulovaite, E., et al., *A brain atlas of synapse protein lifetime across the mouse lifespan.* Neuron, 2022. **110**(24): p. 4057-4073.e8.
5. Danielson, E., et al., *Molecular Diversity of Glutamatergic and GABAergic Synapses from Multiplexed Fluorescence Imaging.* eneuro, 2021. **8**(1): p. ENEURO.0286-20.2020.
6. Falkovich, R., et al., *A synaptic molecular dependency network in knockdown of autism- and schizophrenia-associated genes revealed by multiplexed imaging.* Cell Rep, 2023. **42**(5): p. 112430.
7. Guo, S.-M., et al., *Multiplexed and high-throughput neuronal fluorescence imaging with diffusible probes.* Nature Communications, 2019. **10**(1): p. 4377.
8. Tomov, M.L., et al., *Resolving cell state in iPSC-derived human neural samples with multiplexed fluorescence imaging.* Communications Biology, 2021. **4**(1): p. 786.
9. Kulikov, V., et al., *DoGNet: A deep architecture for synapse detection in multiplexed fluorescence images.* PLoS computational biology, 2019. **15**(5): p. e1007012.
10. Tasic, B., et al., *Adult mouse cortical cell taxonomy revealed by single cell transcriptomics.* Nat Neurosci, 2016. **19**(2): p. 335-46.
11. Gray, L.T., et al., *Layer-specific chromatin accessibility landscapes reveal regulatory networks in adult mouse visual cortex.* Elife, 2017. **6**.
12. Rosenberg, A.B., et al., *Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding.* Science, 2018. **360**(6385): p. 176-182.
13. Tasic, B., et al., *Shared and distinct transcriptomic cell types across neocortical areas.* Nature, 2018. **563**(7729): p. 72-78.
14. Yao, Z., et al., *A taxonomy of transcriptomic cell types across the isocortex and hippocampal formation.* Cell, 2021. **184**(12): p. 3222-3241.e26.
15. Gouwens, N.W., et al., *Classification of electrophysiological and morphological neuron types in the mouse visual cortex.* Nat Neurosci, 2019. **22**(7): p. 1182-1195.
16. Blondel, V.D., et al., *Fast unfolding of communities in large networks.* Journal of Statistical Mechanics: Theory and Experiment, 2008. **2008**(10): p. P10008.
17. Crow, M., et al., *Characterizing the replicability of cell types defined by single cell RNA-sequencing data using MetaNeighbor.* Nature Communications, 2018. **9**(1): p. 884.
18. Shi, L., et al., *Highly-multiplexed volumetric mapping with Raman dye imaging and tissue clearing.* Nature Biotechnology, 2022. **40**(3): p. 364-373.
19. Seo, J., et al., *PICASSO allows ultra-multiplexed fluorescence imaging of spatially overlapping proteins without reference spectra measurements.* Nature Communications, 2022. **13**(1): p. 2475.
20. Lin, J.-R., et al., *Highly multiplexed immunofluorescence imaging of human tissues and tumors using t-CyCIF and conventional optical microscopes.* eLife, 2018. **7**: p. e31657.
21. Gerdes, M.J., et al., *Highly multiplexed single-cell analysis of formalin-fixed, paraffin-embedded cancer tissue.* Proceedings of the National Academy of Sciences, 2013. **110**(29): p. 11982-11987.
22. Radtke, A.J., et al., *IBEX: an iterative immunolabeling and chemical bleaching method for high-content imaging of diverse tissues.* Nature Protocols, 2022. **17**(2): p. 378-401.
23. Radtke, A.J., et al., *IBEX: A versatile multiplex optical imaging approach for deep phenotyping and spatial analysis of cells in complex tissues.* Proceedings of the National Academy of Sciences, 2020. **117**(52): p. 33455-33465.
24. Black, S., et al., *CODEX multiplexed tissue imaging with DNA-conjugated antibodies.* Nature Protocols, 2021. **16**(8): p. 3802-3835.
25. Saka, S.K., et al., *Immuno-SABER enables highly multiplexed and amplified protein imaging in tissues.* Nature Biotechnology, 2019. **37**(9): p. 1080-1090.
26. Moffitt, J.R., E. Lundberg, and H. Heyn, *The emerging landscape of spatial profiling technologies.* Nature Reviews Genetics, 2022. **23**(12): p. 741-759.
27. Hickey, J.W., et al., *Spatial mapping of protein composition and tissue organization: a primer for multiplexed antibody-based imaging.* Nature Methods, 2022. **19**(3): p. 284-295.

28. Lewis, S.M., et al., *Spatial omics and multiplexed imaging to explore cancer biology.* Nature Methods, 2021. **18**(9): p. 997-1012.
29. Andreou, C., R. Weissleder, and M.F. Kircher, *Multiplexed imaging in oncology.* Nature Biomedical Engineering, 2022. **6**(5): p. 527-540.
30. Falkovich, R., et al., *A synaptic molecular dependency network in knockdown of autism- and schizophrenia-associated genes revealed by multiplexed imaging.* Cell Reports, 2023. **42**(5): p. 112430.
31. Shen, F.Y., et al., *Light microscopy based approach for mapping connectivity with molecular specificity.* Nature Communications, 2020. **11**(1): p. 4632.
32. Korsunsky, I., et al., *Fast, sensitive and accurate integration of single-cell data with Harmony.* Nature Methods, 2019. **16**(12): p. 1289-1296.
33. Graf, J., et al., *FLINO: a new method for immunofluorescence bioimage normalization.* Bioinformatics, 2021. **38**(2): p. 520-526.
34. Lin, A. and A.X. Lu, *Incorporating knowledge of plates in batch normalization improves generalization of deep learning for microscopy images.* bioRxiv, 2022: p. 2022.10.14.512286.
35. Sandfort, V., et al., *Data augmentation using generative adversarial networks (CycleGAN) to improve generalizability in CT segmentation tasks.* Scientific Reports, 2019. **9**(1): p. 16884.
36. Qian, W.W., et al., *Batch equalization with a generative adversarial network.* Bioinformatics, 2020. **36 Supplement_2**: p. i875-i883.
37. Somogyi, P., et al., *Salient features of synaptic organisation in the cerebral cortex.* Brain Res Brain Res Rev, 1998. **26**(2-3): p. 113-35.
38. Schneider-Mizell, C.M., et al., *Chandelier cell anatomy and function reveal a variably distributed but common signal.* bioRxiv, 2020.
39. Wamsley, B. and G. Fishell, *Genetic and activity-dependent mechanisms underlying interneuron diversity.* Nat Rev Neurosci, 2017. **18**(5): p. 299-309.
40. Yabut, O., et al., *Abnormal laminar position and dendrite development of interneurons in the reeler forebrain.* Brain Res, 2007. **1140**: p. 75-83.
41. Pla, R., et al., *Layer Acquisition by Cortical GABAergic Interneurons Is Independent of Reelin Signaling.* The Journal of Neuroscience, 2006. **26**(26): p. 6924-6934.
42. Hevner, R.F., et al., *Postnatal shifts of interneuron position in the neocortex of normal and reeler mice: evidence for inward radial migration.* Neuroscience, 2004. **124**(3): p. 605-18.
43. Lodato, S., et al., *Excitatory Projection Neuron Subtypes Control the Distribution of Local Inhibitory Interneurons in the Cerebral Cortex.* Neuron, 2011. **69**(4): p. 763-779.
44. Quattrocolo, G., G. Fishell, and T.J. Petros, *Heterotopic Transplantations Reveal Environmental Influences on Interneuron Diversity and Maturation.* Cell Rep, 2017. **21**(3): p. 721-731.
45. Ishino, Y., et al., *Regional Cellular Environment Shapes Phenotypic Variations of Hippocampal and Neocortical Chandelier Cells.* J Neurosci, 2017. **37**(41): p. 9901-9916.
46. Lim, L., et al., *Optimization of interneuron function by direct coupling of cell migration and axonal targeting.* Nat Neurosci, 2018. **21**(7): p. 920-931.
47. Mayer, C., et al., *Developmental diversification of cortical inhibitory interneurons.* Nature, 2018. **555**(7697): p. 457-462.
48. Fishell, G. and A. Kepecs, *Interneuron Types as Attractors and Controllers.* Annu Rev Neurosci, 2020. **43**: p. 1-30.
49. Mi, D., et al., *Early emergence of cortical interneuron diversity in the mouse embryo.* Science, 2018. **360**(6384): p. 81-85.
50. McKenzie, M.G., et al., *Non-canonical Wnt Signaling through Ryk Regulates the Generation of Somatostatin- and Parvalbumin-Expressing Cortical Interneurons.* Neuron, 2019. **103**(5): p. 853-864 e4.
51. Southwell, D.G., et al., *Intrinsically determined cell death of developing cortical interneurons.* Nature, 2012. **491**(7422): p. 109-13.
52. Wong, F.K., et al., *Pyramidal cell regulation of interneuron survival sculpts cortical networks.* Nature, 2018. **557**(7707): p. 668-673.
53. Francis, K. and B.O. Palsson, *Effective intercellular communication distances are determined by the relative time constants for cyto/chemokine secretion and diffusion.* Proceedings of the National Academy of Sciences, 1997. **94**(23): p. 12258-12262.
54. Dufour, A., et al., *Area Specificity and Topography of Thalamocortical Projections Are Controlled by ephrin/Eph Genes.* Neuron, 2003. **39**(3): p. 453-465.

55. Barallobre, M.J., et al., *Aberrant development of hippocampal circuits and altered neural activity in netrin 1-deficient mice.* Development, 2000. **127**(22): p. 4797-4810.
56. Chédotal, A., et al., *Semaphorins III and IV repel hippocampal axons via two distinct receptors.* Development, 1998. **125**(21): p. 4313-4323.
57. Polleux, F., et al., *Patterning of Cortical Efferent Projections by Semaphorin-Neuropilin Interactions.* Science, 1998. **282**(5395): p. 1904-1906.
58. Polleux, F., T. Morrow, and A. Ghosh, *Semaphorin 3A is a chemoattractant for cortical apical dendrites.* Nature, 2000. **404**(6778): p. 567-573.
59. Ba-Charvet, K.T.N., et al., *Slit2-mediated chemorepulsion and collapse of developing forebrain axons.* Neuron, 1999. **22**(3): p. 463-473.
60. Blockus, H. and A. Chedotal, *Slit-Robo signaling.* Development, 2016. **143**(17): p. 3037-44.
61. Telley, L., et al., *Dual Function of NRP1 in Axon Guidance and Subcellular Target Recognition in Cerebellum.* Neuron, 2016. **91**(6): p. 1276-1291.
62. Kupferman, J.V., et al., *Reelin signaling specifies the molecular identity of the pyramidal neuron distal dendritic compartment.* Cell, 2014. **158**(6): p. 1335-1347.
63. Schafer, D.P., et al., *Microglia sculpt postnatal neural circuits in an activity and complement-dependent manner.* Neuron, 2012. **74**(4): p. 691-705.
64. Lehrman, E.K., et al., *CD47 protects synapses from excess microglia-mediated pruning during development.* Neuron, 2018. **100**(1): p. 120-134. e6.
65. Favuzzi, E., et al., *GABA-receptive microglia selectively sculpt developing inhibitory circuits.* Cell, 2021. **184**(15): p. 4048-4063. e32.
66. Goshi, N., et al., *A primary neural cell culture model to study neuron, astrocyte, and microglia interactions in neuroinflammation.* Journal of Neuroinflammation, 2020. **17**(1): p. 155.
67. Yasuda, M., et al., *An activity-dependent determinant of synapse elimination in the mammalian brain.* Neuron, 2021. **109**(8): p. 1333-1349.e6.
68. Lu, J., et al., *Selective inhibitory control of pyramidal neuron ensembles and cortical subnetworks by chandelier cells.* Nat Neurosci, 2017. **20**(10): p. 1377-1383.
69. Schneider-Mizell, C.M., et al., *Cell-type-specific inhibitory circuitry from a connectomic census of mouse visual cortex.* bioRxiv, 2023: p. 2023.01.23.525290.
70. Wang, X., et al., *Genetic Single Neuron Anatomy Reveals Fine Granularity of Cortical Axo-Axonic Cells.* Cell Reports, 2019. **26**(11): p. 3145-3159.e5.
71. Jiang, X., et al., *Principles of connectivity among morphologically defined cell types in adult neocortex.* Science, 2015. **350**(6264): p. aac9462.
72. Polleux, F., et al., *Pre- and Post-mitotic Events Contribute to the Progressive Acquisition of Area-specific Connectional Fate in the Neocortex.* Cerebral Cortex, 2001. **11**(11): p. 1027-1039.
73. Fraser, S.E. and D.H. Perkel, *Competitive and positional cues in the patterning of nerve connections.* Journal of neurobiology, 1990. **21**(1): p. 51-72.
74. Gierer, A., *Model for the retino-tectal projection.* Proc R Soc Lond B Biol Sci, 1983. **218**(1210): p. 77-93.
75. Löschinger, J., F. Weth, and F. Bonhoeffer, *Reading of concentration gradients by axonal growth cones.* Philos Trans R Soc Lond B Biol Sci, 2000. **355**(1399): p. 971-82.
76. Wu, S.J., et al., *Cortical somatostatin interneuron subtypes form cell-type-specific circuits.* Neuron, 2023.
77. Daigle, T.L., et al., *A Suite of Transgenic Driver and Reporter Mouse Lines with Enhanced Brain-Cell-Type Targeting and Functionality.* Cell, 2018. **174**(2): p. 465-480 e22.
78. Gamlin, C.R., et al., *Integrating EM and Patch-seq data: Synaptic connectivity and target specificity of predicted Sst transcriptomic types.* bioRxiv, 2023: p. 2023.03.22.533857.
79. Maccaferri, G., K. Tóth, and C.J. McBain, *Target-specific expression of presynaptic mossy fiber plasticity.* Science, 1998. **279**(5355): p. 1368-1371.
80. Acsády, L., et al., *GABAergic Cells Are the Major Postsynaptic Targets of Mossy Fibers in the Rat Hippocampus.* The Journal of Neuroscience, 1998. **18**(9): p. 3386-3403.
81. Dictionary, O.E., *vineal, adj., Etymology*, in *Oxford English Dictionary*. 2023, Oxford University Press.
82. Gidon, A. and I. Segev, *Principles Governing the Operation of Synaptic Inhibition in Dendrites.* Neuron, 2012. **75**(2): p. 330-341.
83. Higley, M.J., *Localized GABAergic inhibition of dendritic Ca2+ signalling.* Nature Reviews Neuroscience, 2014. **15**(9): p. 567-572.
84. Kubota, Y., et al., *Functional effects of distinct innervation styles of pyramidal cells by fast spiking cortical interneurons.* Elife, 2015. **4**.

85. Markram, H., et al., *Reconstruction and Simulation of Neocortical Microcircuitry.* Cell, 2015. **163**(2): p. 456-92.
86. Bloss, E.B., et al., *Structured Dendritic Inhibition Supports Branch-Selective Integration in CA1 Pyramidal Cells.* Neuron, 2016. **89**(5): p. 1016-30.
87. Tremblay, R., S. Lee, and B. Rudy, *GABAergic Interneurons in the Neocortex: From Cellular Properties to Circuits.* Neuron, 2016. **91**(2): p. 260-92.
88. Boivin, J.R. and E. Nedivi, *Functional implications of inhibitory synapse placement on signal processing in pyramidal neuron dendrites.* Curr Opin Neurobiol, 2018. **51**: p. 16-22.
89. Haufler, D., et al., *Simulations of cortical networks using spatially extended conductance-based neuronal models.* The Journal of Physiology, 2023. **601**(15): p. 3123-3139.
90. Rudy, B., et al., *Three groups of interneurons account for nearly 100% of neocortical GABAergic neurons.* Dev Neurobiol, 2011. **71**(1): p. 45-61.
91. Butt, S.J., et al., *The temporal and spatial origins of cortical interneurons predict their physiological subtype.* Neuron, 2005. **48**(4): p. 591-604.
92. Butt, S.J., et al., *The requirement of Nkx2-1 in the temporal specification of cortical interneuron subtypes.* Neuron, 2008. **59**(5): p. 722-32.
93. Schuman, B., et al., *Four Unique Interneuron Populations Reside in Neocortical Layer 1.* J Neurosci, 2019. **39**(1): p. 125-139.
94. Jones, E.G. and T.P.S. Powell, *An electron microscopic study of the laminar pattern and mode of termination of afferent fibre pathways in the somatic sensory cortex of the cat.* Philosophical Transactions of the Royal Society of London. B, Biological Sciences, 1970. **257**(812): p. 45-62.
95. Vogt, B.A., *The role of layer I in cortical function*, in *Normal and altered states of function*. 1991, Springer. p. 49-80.
96. Cruikshank, S.J., et al., *Thalamic Control of Layer 1 Circuits in Prefrontal Cortex.* The Journal of Neuroscience, 2012. **32**(49): p. 17813-17823.
97. Roth, M.M., et al., *Thalamic nuclei convey diverse contextual information to layer 1 of visual cortex.* Nature neuroscience, 2016. **19**(2): p. 299-307.
98. Vogt, C. and O. Vogt, *Allgemeine ergebnisse unserer hirnforschung*. Vol. 25. 1919: JA Barth.
99. Munoz, W., et al., *Layer-specific modulation of neocortical dendritic inhibition during active wakefulness.* Science, 2017. **355**(6328): p. 954-958.
100. Sóki, N., et al., *Investigation of synapses in the cortical white matter in human temporal lobe epilepsy.* Brain Research, 2022. **1779**: p. 147787.
101. Frazer, S., et al., *Transcriptomic and anatomic parcellation of 5-HT(3A)R expressing cortical interneuron subtypes revealed by single-cell RNA sequencing.* Nat Commun, 2017. **8**: p. 14219.
102. Banovac, I., et al., *The Distinct Characteristics of Somatostatin Neurons in the Human Brain.* Molecular Neurobiology, 2022. **59**(8): p. 4953-4965.
103. Tomioka, R., et al., *Demonstration of long-range GABAergic connections distributed throughout the mouse neocortex.* European Journal of Neuroscience, 2005. **21**(6): p. 1587-1600.
104. Naka, A., et al., *Complementary networks of cortical somatostatin interneurons enforce layer specific control.* Elife, 2019. **8**.
105. Nigro, M.J., Y. Hashikawa-Yamasaki, and B. Rudy, *Diversity and Connectivity of Layer 5 Somatostatin-Expressing Interneurons in the Mouse Barrel Cortex.* J Neurosci, 2018. **38**(7): p. 1622-1633.
106. Wang, Q., et al., *The Allen Mouse Brain Common Coordinate Framework: A 3D Reference Atlas.* Cell, 2020. **181**(4): p. 936-953 e20.
107. Steindler, D.A. and S.A. Colwell, *Reeler mutant mouse: maintenance of appropriate and reciprocal connections in the cerebral cortex and thalamus.* Brain Research, 1976. **113**(2): p. 386-393.
108. Lutz, D., et al., *Proteolytic cleavage of transmembrane cell adhesion molecule L1 by extracellular matrix molecule Reelin is important for mouse brain development.* Scientific Reports, 2017. **7**(1): p. 15268.
109. Landrieu, P. and A. Goffinet, *Inverted pyramidal neurons and their axons in the neocortex of reeler mutant mice.* Cell and tissue research, 1981. **218**: p. 293-301.
110. Hammond, V., et al., *Layer Positioning of Late-Born Cortical Interneurons Is Dependent on Reelin But Not p35 Signaling.* The Journal of Neuroscience, 2006. **26**(5): p. 1646-1655.
111. Naus, C.C.G. and F.E. Bloom, *Immunohistochemical analysis of the development of somatostatin in the reeler neocortex.* Developmental Brain Research, 1988. **43**(1): p. 61-68.
112. Ma, Y., et al., *Distinct subtypes of somatostatin-containing neocortical interneurons revealed in transgenic mice.* J Neurosci, 2006. **26**(19): p. 5069-82.

113. Oliva, A.A., et al., *Novel hippocampal interneuronal subtypes identified using transgenic mice that express green fluorescent protein in GABAergic interneurons.* Journal of Neuroscience, 2000. **20**(9): p. 3354-3368.
114. Williams, M.E., et al., *Cadherin-9 regulates synapse-specific differentiation in the developing hippocampus.* Neuron, 2011. **71**(4): p. 640-55.
115. Barondes, S., *Neuronal recognition*. 2013: Springer Science & Business Media.
116. Sanes, D.H. and M.-m. Poot, *In vitro analysis of position- and lineage-dependent selectivity in the formation of neuromuscular synapses.* Neuron, 1989. **2**(3): p. 1237-1244.
117. Christie, S.B. and A.L. De Blas, *GABAergic and glutamatergic axons innervate the axon initial segment and organize GABA(A) receptor clusters of cultured hippocampal pyramidal cells.* J Comp Neurol, 2003. **456**(4): p. 361-74.
118. Somogyi, P., et al., *Glutamate decarboxylase immunoreactivity in the hippocampus of the cat: distribution of immunoreactive synaptic terminals with special reference to the axon initial segment of pyramidal neurons.* The Journal of Neuroscience, 1983. **3**(7): p. 1450.
119. Rao, A., E.M. Cha, and A.M. Craig, *Mismatched appositions of presynaptic and postsynaptic components in isolated hippocampal neurons.* Journal of Neuroscience, 2000. **20**(22): p. 8344-8353.
120. Haydon, P.G. and P. Drapeau, *From contact to connection: early events during synaptogenesis.* Trends in Neurosciences, 1995. **18**(4): p. 196-201.
121. Benson, D.L. and P.A. Cohen, *Activity-independent segregation of excitatory and inhibitory synaptic terminals in cultured hippocampal neurons.* Journal of Neuroscience, 1996. **16**(20): p. 6424-6432.
122. Ito, S. and M. Takeichi, *Dendrites of cerebellar granule cells correctly recognize their target axons for synaptogenesis in vitro.* Proceedings of the National Academy of Sciences, 2009. **106**(31): p. 12782-12787.
123. Hoang, P.T., et al., *Subtype Diversification and Synaptic Specificity of Stem Cell-Derived Spinal Interneurons.* Neuron, 2018. **100**(1): p. 135-149.e7.
124. Brasch, J., et al., *Visualization of clustered protocadherin neuronal self-recognition complexes.* Nature, 2019. **569**(7755): p. 280-283.
125. Slack, J.M., *Conrad Hal Waddington: the last renaissance biologist?* Nature Reviews Genetics, 2002. **3**(11): p. 889-895.
126. Waddington, C.H., *The strategy of the genes*. 1957: Routledge.
127. Sperry, R.W., *Chemoaffinity in the Orderly Growth of Nerve Fiber Patterns and Connections.* Proc Natl Acad Sci U S A, 1963. **50**(4): p. 703-10.
128. Sanes, J.R. and S.L. Zipursky, *Synaptic Specificity, Recognition Molecules, and Assembly of Neural Circuits.* Cell, 2020. **181**(3): p. 536-556.
129. Südhof, T.C., *Towards an Understanding of Synapse Formation.* Neuron, 2018. **100**(2): p. 276-293.
130. Südhof, T.C., *The cell biology of synapse formation.* Journal of Cell Biology, 2021. **220**(7).
131. Honig, B. and L. Shapiro, *Adhesion Protein Structure, Molecular Affinities, and Principles of Cell-Cell Recognition.* Cell, 2020. **181**(3): p. 520-535.
132. Yamagata, M., X. Duan, and J.R. Sanes, *Cadherins Interact With Synaptic Organizers to Promote Synaptic Differentiation.* Front Mol Neurosci, 2018. **11**: p. 142.
133. Xu, S., et al., *Interactions between the Ig-Superfamily Proteins DIP-α and Dpr6/10 Regulate Assembly of Neural Circuits.* Neuron, 2018. **100**(6): p. 1369-1384.e6.
134. Blazquez-Llorca, L., et al., *Spatial distribution of neurons innervated by chandelier cells.* Brain Structure and Function, 2015. **220**(5): p. 2817-2834.
135. Turner, N.L., et al., *Multiscale and multimodal reconstruction of cortical structure and function.* bioRxiv, 2020.
136. *Development, morphology and topography of chandelier cells in the auditory codex of the cat.* 1985.
137. Favuzzi, E., et al., *Distinct molecular programs regulate synapse specificity in cortical inhibitory circuits.* Science, 2019. **363**(6425): p. 413-417.
138. Marin, O., *Interneuron dysfunction in psychiatric disorders.* Nat Rev Neurosci, 2012. **13**(2): p. 107-20.
139. Lewis, D.A., et al., *Cortical parvalbumin interneurons and cognitive dysfunction in schizophrenia.* Trends Neurosci, 2012. **35**(1): p. 57-67.
140. Dutton, S.B., et al., *Preferential inactivation of Scn1a in parvalbumin interneurons increases seizure susceptibility.* Neurobiol Dis, 2013. **49**: p. 211-20.
141. Malik, R., et al., *Tsc1 represses parvalbumin expression and fast-spiking properties in somatostatin lineage cortical interneurons.* Nat Commun, 2019. **10**(1): p. 4994.
142. Lewis, D.A., *Chandelier cells: shedding light on altered cortical circuitry in schizophrenia.* Mol Psychiatry, 1998. **3**(6): p. 468-71, 466-7.

143. Volk, D.W., et al., *Reciprocal alterations in pre- and postsynaptic inhibitory markers at chandelier cell inputs to pyramidal neurons in schizophrenia.* Cerebral Cortex, 2002. **12**(10): p. 1063-1070.
144. Cruz, D.A., et al., *Selective alterations in postsynaptic markers of chandelier cell inputs to cortical pyramidal neurons in subjects with schizophrenia.* Neuropsychopharmacology, 2009. **34**(9): p. 2112-24.
145. Lewis, D.A. and R.A. Sweet, *Schizophrenia from a neural circuitry perspective: advancing toward rational pharmacological therapies.* J Clin Invest, 2009. **119**(4): p. 706-16.
146. Beneyto, M., et al., *Lamina-specific alterations in cortical GABA(A) receptor subunit expression in schizophrenia.* Cereb Cortex, 2011. **21**(5): p. 999-1011.
147. Glausier, J.R., K.N. Fish, and D.A. Lewis, *Altered parvalbumin basket cell inputs in the dorsolateral prefrontal cortex of schizophrenia subjects.* Mol Psychiatry, 2014. **19**(1): p. 30-6.
148. Ariza, J., et al., *The Number of Chandelier and Basket Cells Are Differentially Decreased in Prefrontal Cortex in Autism.* Cereb Cortex, 2018. **28**(2): p. 411-420.
149. Ito-Ishida, A., et al., *Loss of MeCP2 in Parvalbumin-and Somatostatin-Expressing Neurons in Mice Leads to Distinct Rett Syndrome-like Phenotypes.* Neuron, 2015. **88**(4): p. 651-8.
150. Ure, K., et al., *Restoration of Mecp2 expression in GABAergic neurons is sufficient to rescue multiple disease features in a mouse model of Rett syndrome.* Elife, 2016. **5**.
151. Levitt, P. and D.B. Campbell, *The genetic and neurobiologic compass points toward common signaling dysfunctions in autism spectrum disorders.* J Clin Invest, 2009. **119**(4): p. 747-54.
152. Fazzari, P., et al., *Control of cortical GABA circuitry development by Nrg1 and ErbB4 signalling.* Nature, 2010. **464**(7293): p. 1376-80.
153. Del Pino, I., et al., *Erbb4 deletion from fast-spiking interneurons causes schizophrenia-like phenotypes.* Neuron, 2013. **79**(6): p. 1152-68.
154. Del Pino, I., et al., *Abnormal wiring of CCK(+) basket cells disrupts spatial information coding.* Nat Neurosci, 2017. **20**(6): p. 784-792.
155. Schiebinger, G., et al., *Optimal-transport analysis of single-cell gene expression identifies developmental trajectories in reprogramming.* Cell, 2019. **176**(4): p. 928-943. e22.
156. Brown, A., et al., *Topographic Mapping from the Retina to the Midbrain Is Controlled by Relative but Not Absolute Levels of EphA Receptor Signaling.* Cell, 2000. **102**(1): p. 77-88.
157. Kania, A. and R. Klein, *Mechanisms of ephrin–Eph signalling in development, physiology and disease.* Nature Reviews Molecular Cell Biology, 2016. **17**(4): p. 240-256.
158. Williams, S.E., et al., *Ephrin-B2 and EphB1 mediate retinal axon divergence at the optic chiasm.* Neuron, 2003. **39**(6): p. 919-935.
159. García-Frigola, C., et al., *Zic2 promotes axonal divergence at the optic chiasm midline by EphB1-dependent and-independent mechanisms.* 2008.
160. Herrera, E.s., et al., *Zic2 patterns binocular vision by specifying the uncrossed retinal projection.* Cell, 2003. **114**(5): p. 545-557.
161. Petros, T.J., B.R. Shrestha, and C. Mason, *Specificity and sufficiency of EphB1 in driving the ipsilateral retinal projection.* Journal of Neuroscience, 2009. **29**(11): p. 3463-3474.
162. Stein, B.E., et al., *The superior colliculus and visual thalamus*, in *Neuroscience in the 21st Century: From Basic to Clinical*. 2022, Springer. p. 1029-1054.
163. Hansen, M.J., G.E. Dallal, and J.G. Flanagan, *Retinal axon response to ephrin-as shows a graded, concentration-dependent transition from growth promotion to inhibition.* Neuron, 2004. **42**(5): p. 717-730.
164. Hindges, R., et al., *EphB forward signaling controls directional branch extension and arborization required for dorsal-ventral retinotopic mapping.* Neuron, 2002. **35**(3): p. 475-487.
165. Suetterlin, P. and U. Drescher, *Target-independent ephrina/EphA-mediated axon-axon repulsion as a novel element in retinocollicular mapping.* Neuron, 2014. **84**(4): p. 740-752.
166. Thakar, S., G. Chenaux, and M. Henkemeyer, *Critical roles for EphB and ephrin-B bidirectional signalling in retinocollicular mapping.* Nature communications, 2011. **2**(1): p. 431.
167. Hornberger, M.R., et al., *Modulation of EphA receptor function by coexpressed ephrinA ligands on retinal ganglion cell axons.* Neuron, 1999. **22**(4): p. 731-742.
168. Chen, S.-Y. and H.-J. Cheng, *Functions of axon guidance molecules in synapse formation.* Current Opinion in Neurobiology, 2009. **19**(5): p. 471-478.
169. Kandel, E.R., et al., *Formation and Elimination of Synapses*, in *Principles of Neural Science, 6e*. 2021, McGraw Hill: New York, NY.
170. Shapiro, L. and D.R. Colman, *The diversity of cadherins and implications for a synaptic adhesive code in the CNS.* Neuron, 1999. **23**(3): p. 427-430.

171. Sanes, J.R. and M. Yamagata, *Many Paths to Synaptic Specificity.* Annual Review of Cell and Developmental Biology, 2009. **25**(1): p. 161-195.
172. Duan, X., et al., *Type II Cadherins Guide Assembly of a Direction-Selective Retinal Circuit.* Cell, 2014. **158**(4): p. 793-807.
173. Xu, C., et al., *Control of Synaptic Specificity by Establishing a Relative Preference for Synaptic Partners.* Neuron, 2019. **103**(5): p. 865-877.e7.
174. Sun, Y.C., et al., *Integrating barcoded neuroanatomy with spatial transcriptional profiling enables identification of gene correlates of projections.* Nat Neurosci, 2021. **24**(6): p. 873-885.
175. Harris, K.D., et al., *Classes and continua of hippocampal CA1 inhibitory neurons revealed by single-cell transcriptomics.* PLoS biology, 2018. **16**(6): p. e2006387.
176. Schwabe, T., H. Neuert, and Thomas R. Clandinin, *A Network of Cadherin-Mediated Interactions Polarizes Growth Cones to Determine Targeting Specificity.* Cell, 2013. **154**(2): p. 351-364.
177. Schwabe, T., et al., *Differential Adhesion Determines the Organization of Synaptic Fascicles in the Drosophila Visual System.* Current Biology, 2014. **24**(12): p. 1304-1313.
178. Chistiakova, M., et al., *Heterosynaptic plasticity: multiple mechanisms and multiple roles.* The Neuroscientist, 2014. **20**(5): p. 483-498.
179. Chistiakova, M., et al., *Homeostatic role of heterosynaptic plasticity: models and experiments.* Frontiers in computational neuroscience, 2015. **9**: p. 89.
180. Jungenitz, T., et al., *Structural homo-and heterosynaptic plasticity in mature and adult newborn rat hippocampal granule cells.* Proceedings of the National Academy of Sciences, 2018. **115**(20): p. E4670-E4679.
181. Chater, T.E. and Y. Goda, *My neighbour hetero—deconstructing the mechanisms underlying heterosynaptic plasticity.* Current Opinion in Neurobiology, 2021. **67**: p. 106-114.
182. Christie, S.B., et al., *Synaptic and extrasynaptic GABAA receptor and gephyrin clusters.* Prog Brain Res, 2002. **136**: p. 157-80.
183. Brickley, S.G. and I. Mody, *Extrasynaptic GABA(A) receptors: their function in the CNS and implications for disease.* Neuron, 2012. **73**(1): p. 23-34.
184. Kaufman, A.M., et al., *Opposing Roles of Synaptic and Extrasynaptic NMDA Receptor Signaling in Cocultured Striatal and Cortical Neurons.* The Journal of Neuroscience, 2012. **32**(12): p. 3992-4003.
185. Sabo, S.L., R.A. Gomes, and A.K. McAllister, *Formation of Presynaptic Terminals at Predefined Sites along Axons.* The Journal of Neuroscience, 2006. **26**(42): p. 10813-10825.
186. Wu, Y.E., et al., *The balance between capture and dissociation of presynaptic proteins controls the spatial distribution of synapses.* Neuron, 2013. **78**(6): p. 994-1011.
187. Poon, V.Y., M.P. Klassen, and K. Shen, *UNC-6/netrin and its receptor UNC-5 locally exclude presynaptic components from dendrites.* Nature, 2008. **455**(7213): p. 669-73.
188. Klassen, M.P. and K. Shen, *Wnt Signaling Positions Neuromuscular Connectivity by Inhibiting Synapse Formation in C. elegans.* Cell, 2007. **130**(4): p. 704-716.
189. John, N., et al., *Brevican-containing perineuronal nets of extracellular matrix in dissociated hippocampal primary cultures.* Mol Cell Neurosci, 2006. **31**(4): p. 774-84.
190. Bruckner, G., et al., *Perineuronal nets provide a polyanionic, glia-associated form of microenvironment around certain neurons in many parts of the rat brain.* Glia, 1993. **8**(3): p. 183-200.
191. Hedstrom, K.L., et al., *Neurofascin assembles a specialized extracellular matrix at the axon initial segment.* J Cell Biol, 2007. **178**(5): p. 875-86.
192. Dityatev, A., C.I. Seidenbecher, and M. Schachner, *Compartmentalization from the outside: the extracellular matrix and functional microdomains in the brain.* Trends in Neurosciences, 2010. **33**(11): p. 503-512.
193. Bekkers, J. and C. Stevens, *Excitatory and inhibitory autaptic currents in isolated hippocampal neurons maintained in cell culture.* Proceedings of the National Academy of Sciences, 1991. **88**(17): p. 7834-7838.
194. Segal, M.M., *Epileptiform activity in microcultures containing one excitatory hippocampal neuron.* Journal of Neurophysiology, 1991. **65**(4): p. 761-770.
195. Burlingham, S.R., et al., *Induction of synapse formation by de novo neurotransmitter synthesis.* Nature Communications, 2022. **13**(1): p. 3060.
196. Schreiner, D. and J.A. Weiner, *Combinatorial homophilic interaction between gamma-protocadherin multimers greatly expands the molecular diversity of cell adhesion.* Proc Natl Acad Sci U S A, 2010. **107**(33): p. 14893-8.
197. Togashi, H., et al., *Interneurite affinity is regulated by heterophilic nectin interactions in concert with the cadherin machinery.* J Cell Biol, 2006. **174**(1): p. 141-51.

198. Courgeon, M. and C. Desplan, *Coordination between stochastic and deterministic specification in the Drosophila visual system.* Science, 2019. **366**(6463): p. eaay6727.
199. Valdes-Aleman, J., et al., *Comparative Connectomics Reveals How Partner Identity, Location, and Activity Specify Synaptic Connectivity in Drosophila.* Neuron, 2021. **109**(1): p. 105-122.e7.
200. Zhang, H.L. and H.B. Peng, *Mechanism of acetylcholine receptor cluster formation induced by DC electric field.* PLoS One, 2011. **6**(10): p. e26805.
201. Orida, N. and M.-m. Poo, *Electrophoretic movement and localization of acetylcholine receptors in the embryonic muscle cell membrane.* Nature, 1978. **275**(5675): p. 31-35.
202. Kwon, H.-B. and B.L. Sabatini, *Glutamate induces de novo growth of functional spines in developing cortex.* Nature, 2011. **474**(7349): p. 100-104.
203. Oh, W.C., et al., *De novo synaptogenesis induced by GABA in the developing mouse cortex.* Science, 2016. **353**(6303): p. 1037-1040.
204. Burry, R.W., *Development of apparent presynaptic elements formed in response to polylysine coated surfaces.* Brain Research, 1982. **247**(1): p. 1-16.
205. Burry, R.W. and D.M. Hayes, *Development and elimination of presynaptic elements on polylysine-coated beads implanted in neonatal rat cerebellum.* Journal of Neuroscience Research, 1986. **15**(1): p. 67-78.
206. Chen, W.V. and T. Maniatis, *Clustered protocadherins.* Development, 2013. **140**(16): p. 3297-302.
207. Chen, W.V., et al., *Pcdhalphac2 is required for axonal tiling and assembly of serotonergic circuitries in mice.* Science, 2017. **356**(6336): p. 406-411.
208. Shepherd, G.M., *Neuron Doctrine: Historical Background*, in *Encyclopedia of Neuroscience*, L.R. Squire, Editor. 2009, Academic Press: Oxford. p. 691-695.
209. Lagier, S., et al., *GABAergic inhibition at dendrodendritic synapses tunes γ oscillations in the olfactory bulb.* Proceedings of the National Academy of Sciences, 2007. **104**(17): p. 7259-7264.
210. Rall, W., et al., *Dendrodendritic synaptic pathway for inhibition in the olfactory bulb.* Experimental neurology, 1966. **14**(1): p. 44-56.
211. Betley, J.N., et al., *Stringent specificity in the construction of a GABAergic presynaptic inhibitory circuit.* Cell, 2009. **139**(1): p. 161-74.
212. Ashrafi, S., et al., *Neuronal Ig/Caspr recognition promotes the formation of axoaxonic synapses in mouse spinal cord.* Neuron, 2014. **81**(1): p. 120-129.
213. Shen, K., *Molecular mechanisms of target specificity during synapse formation.* Current Opinion in Neurobiology, 2004. **14**(1): p. 83-88.
214. Shen, K., R.D. Fetter, and C.I. Bargmann, *Synaptic Specificity Is Generated by the Synaptic Guidepost Protein SYG-2 and Its Receptor, SYG-1.* Cell, 2004. **116**(6): p. 869-881.
215. Cash, S., A. Chiba, and H. Keshishian, *Alternate neuromuscular target selection following the loss of single muscle fibers in Drosophila.* J Neurosci, 1992. **12**(6): p. 2051-64.
216. Milovanovic, D., et al., *A liquid phase of synapsin and lipid vesicles.* Science, 2018. **361**(6402): p. 604-607.
217. Zeng, M., et al., *Reconstituted Postsynaptic Density as a Molecular Platform for Understanding Synapse Formation and Plasticity.* Cell, 2018. **174**(5): p. 1172-1187.e16.
218. Wu, X., et al., *RIM and RIM-BP Form Presynaptic Active-Zone-like Condensates via Phase Separation.* Molecular Cell, 2019. **73**(5): p. 971-984.e5.
219. Chen, X., et al., *Phase separation at the synapse.* Nature neuroscience, 2020. **23**(3): p. 301-310.
220. Emperador-Melero, J. and P.S. Kaeser, *Assembly of the presynaptic active zone.* Current Opinion in Neurobiology, 2020. **63**: p. 95-103.
221. Bai, G., Y. Wang, and M. Zhang, *Gephyrin-mediated formation of inhibitory postsynaptic density sheet via phase separation.* Cell Res, 2021. **31**(3): p. 312-325.
222. Hoffmann, C. and D. Milovanovic, *Gephyrin: a scaffold that builds a phase at the inhibitory postsynapses.* Cell Research, 2021. **31**(3): p. 245-246.
223. Lyon, A.S., W.B. Peeples, and M.K. Rosen, *A framework for understanding the functions of biomolecular condensates across scales.* Nature Reviews Molecular Cell Biology, 2021. **22**(3): p. 215-235.
224. Park, D., et al., *Cooperative function of synaptophysin and synapsin in the generation of synaptic vesicle-like clusters in non-neuronal cells.* Nature Communications, 2021. **12**(1): p. 263.
225. Bai, G. and M. Zhang, *Inhibitory postsynaptic density from the lens of phase separation.* Oxford Open Neuroscience, 2022. **1**.
226. Qi, C., et al., *Molecular mechanisms of synaptogenesis.* Frontiers in Synaptic Neuroscience, 2022. **14**.
227. Hoffmann, C., et al., *Synapsin condensation controls synaptic vesicle sequestering and dynamics.* Nature Communications, 2023. **14**(1): p. 6730.

228. Wang, S., et al., *Alternative splicing of latrophilin-3 controls synapse formation.* Nature, 2024. **626**(7997): p. 128-135.
229. McDonald, N.A., R.D. Fetter, and K. Shen, *Assembly of synaptic active zones requires phase separation of scaffold molecules.* Nature, 2020. **588**(7838): p. 454-458.
230. Yuzaki, M., *Two Classes of Secreted Synaptic Organizers in the Central Nervous System.* Annual Review of Physiology, 2018. **80**(1): p. 243-262.
231. McRae, P.A., et al., *Sensory deprivation alters aggrecan and perineuronal net expression in the mouse barrel cortex.* J Neurosci, 2007. **27**(20): p. 5405-13.
232. Giamanco, K.A. and R.T. Matthews, *Deconstructing the perineuronal net: cellular contributions and molecular composition of the neuronal extracellular matrix.* Neuroscience, 2012. **218**: p. 367-84.
233. Flores, C.E. and P. Méndez, *Shaping inhibition: activity dependent structural plasticity of GABAergic synapses.* Frontiers in cellular neuroscience, 2014. **8**: p. 327.
234. de Winter, F., et al., *The Chemorepulsive Protein Semaphorin 3A and Perineuronal Net-Mediated Plasticity.* Neural Plast, 2016. **2016**: p. 3679545.
235. Favuzzi, E., et al., *Activity-dependent gating of parvalbumin interneuron function by the perineuronal net protein brevican.* Neuron, 2017. **95**(3): p. 639-655. e10.
236. Testa, D., A. Prochiantz, and A.A. Di Nardo, *Perineuronal nets in brain physiology and disease.* Semin Cell Dev Biol, 2019. **89**: p. 125-135.
237. Ravasenga, T., et al., *Spatial regulation of coordinated excitatory and inhibitory synaptic plasticity at dendritic synapses.* Cell reports, 2022. **38**(6).
238. Weber, J.P., et al., *Location-dependent synaptic plasticity rules by dendritic spine cooperativity.* Nature communications, 2016. **7**(1): p. 11380.
239. Chistiakova, M. and M. Volgushev, *Heterosynaptic plasticity in the neocortex.* Experimental brain research, 2009. **199**(3): p. 377-390.
240. van Oostrum, M., et al., *The proteomic landscape of synaptic diversity across brain regions and cell types.* Cell, 2023. **186**(24): p. 5411-5427. e23.
241. Grant, S.G.N. and E. Fransén, *The Synapse Diversity Dilemma: Molecular Heterogeneity Confounds Studies of Synapse Function.* Frontiers in Synaptic Neuroscience, 2020. **12**.
242. Nusser, Z., *Creating diverse synapses from the same molecules.* Current Opinion in Neurobiology, 2018. **51**: p. 8-15.
243. O'Rourke, N.A., et al., *Deep molecular diversity of mammalian synapses: why it matters and how to measure it.* Nature Reviews Neuroscience, 2012. **13**(6): p. 365-379.
244. Tononi, G., O. Sporns, and G.M. Edelman, *A measure for brain complexity: relating functional segregation and integration in the nervous system.* Proceedings of the National Academy of Sciences, 1994. **91**(11): p. 5033-5037.
245. Tononi, G., O. Sporns, and G.M. Edelman, *A complexity measure for selective matching of signals by the brain.* Proceedings of the National Academy of Sciences, 1996. **93**(8): p. 3422-3427.
246. Koch, C. and G. Laurent, *Complexity and the nervous system.* Science, 1999. **284**(5411): p. 96-98.
247. Weng, G., U.S. Bhalla, and R. Iyengar, *Complexity in biological signaling systems.* Science, 1999. **284**(5411): p. 92-96.
248. Sporns, O., G. Tononi, and G.M. Edelman, *Connectivity and complexity: the relationship between neuroanatomy and brain dynamics.* Neural networks, 2000. **13**(8-9): p. 909-922.
249. Edelman, G.M. and J.A. Gally, *Degeneracy and complexity in biological systems.* Proceedings of the National Academy of Sciences, 2001. **98**(24): p. 13763-13768.
250. Székely, G., *An approach to the complexity of the brain.* Brain research bulletin, 2001. **55**(1): p. 11-28.
251. Cichon, J. and W.-B. Gan, *Branch-specific dendritic Ca2+ spikes cause persistent synaptic plasticity.* Nature, 2015. **520**(7546): p. 180-185.
252. Magó, Á., et al., *Synaptic Plasticity Depends on the Fine-Scale Input Pattern in Thin Dendrites of CA1 Pyramidal Neurons.* J Neurosci, 2020. **40**(13): p. 2593-2605.
253. Harvey, C.D. and K. Svoboda, *Locally dynamic synaptic learning rules in pyramidal neuron dendrites.* Nature, 2007. **450**(7173): p. 1195-1200.
254. Govindarajan, A., et al., *The dendritic branch is the preferred integrative unit for protein synthesis-dependent LTP.* Neuron, 2011. **69**(1): p. 132-146.
255. Root, D.H., et al., *Single rodent mesohabenular axons release glutamate and GABA.* Nature neuroscience, 2014. **17**(11): p. 1543-1551.

256. Saunders, A., et al., *A direct GABAergic output from the basal ganglia to frontal cortex.* Nature, 2015. **521**(7550): p. 85-89.
257. Blatow, M., et al., *A novel network of multipolar bursting interneurons generates theta frequency oscillations in neocortex.* Neuron, 2003. **38**(5): p. 805-17.
258. Sohal, V.S., et al., *Parvalbumin neurons and gamma rhythms enhance cortical circuit performance.* Nature, 2009. **459**(7247): p. 698-702.
259. Pang, Z.P., et al., *Synaptotagmin-2 is essential for survival and contributes to Ca2+ triggering of neurotransmitter release in central and neuromuscular synapses.* J Neurosci, 2006. **26**(52): p. 13493-504.
260. Bouhours, B., et al., *Synaptotagmin2 (Syt2) Drives Fast Release Redundantly with Syt1 at the Output Synapses of Parvalbumin-Expressing Inhibitory Neurons.* J Neurosci, 2017. **37**(17): p. 4604-4617.
261. Chen, C., et al., *Synaptotagmin 2 Is the Fast Ca(2+) Sensor at a Central Inhibitory Synapse.* Cell Rep, 2017. **18**(3): p. 723-736.
262. Kerr, A.M., E. Reisinger, and P. Jonas, *Differential dependence of phasic transmitter release on synaptotagmin 1 at GABAergic and glutamatergic hippocampal synapses.* Proceedings of the National Academy of Sciences of the United States of America, 2008. **105**(40): p. 15581-15586.
263. Kochubey, O., N. Babai, and R. Schneggenburger, *A Synaptotagmin Isoform Switch during the Development of an Identified CNS Synapse.* Neuron, 2016. **90**(5): p. 984-99.
264. Martens, S., M.M. Kozlov, and H.T. McMahon, *How synaptotagmin promotes membrane fusion.* Science, 2007. **316**(5828): p. 1205-8.
265. Sommeijer, J.P. and C.N. Levelt, *Synaptotagmin-2 is a reliable marker for parvalbumin positive inhibitory boutons in the mouse visual cortex.* PLoS One, 2012. **7**(4): p. e35323.
266. Uezu, A., et al., *Identification of an elaborate complex mediating postsynaptic inhibition.* Science, 2016. **353**(6304): p. 1123-9.
267. Goaillard, J.-M. and E. Marder, *Ion channel degeneracy, variability, and covariation in neuron and circuit resilience.* Annual review of neuroscience, 2021. **44**: p. 335-357.
268. Whitacre, J. and A. Bender, *Degeneracy: a design principle for achieving robustness and evolvability.* Journal of theoretical biology, 2010. **263**(1): p. 143-153.
269. Tononi, G., O. Sporns, and G.M. Edelman, *Measures of degeneracy and redundancy in biological networks.* Proceedings of the National Academy of Sciences, 1999. **96**(6): p. 3257-3262.
270. Haddad, S.A. and E. Marder, *Circuit Robustness to Temperature Perturbation Is Altered by Neuromodulators.* Neuron, 2018. **100**(3): p. 609-623.e3.
271. Marder, E., et al., *Neuromodulation of Circuits with Variable Parameters: Single Neurons and Small Circuits Reveal Principles of State-Dependent and Robust Neuromodulation.* Annual Review of Neuroscience, 2014. **37**(Volume 37, 2014): p. 329-346.
272. Stanco, A., et al., *Netrin-1–α3β1 integrin interactions regulate the migration of interneurons through the cortical marginal zone.* Proceedings of the National Academy of Sciences, 2009. **106**(18): p. 7595-7600.
273. Manitt, C., et al., *Netrin participates in the development of retinotectal synaptic connectivity by modulating axon arborization and synapse formation in the developing brain.* Journal of Neuroscience, 2009. **29**(36): p. 11065-11077.
274. Goldman, J.S., et al., *Netrin-1 promotes excitatory synaptogenesis between cortical neurons by initiating synapse assembly.* Journal of Neuroscience, 2013. **33**(44): p. 17278-17289.
275. Zheng, M., et al., *Netrin-1 promotes synaptic formation and axonal regeneration via JNK1/c-Jun pathway after the middle cerebral artery occlusion.* Frontiers in Cellular Neuroscience, 2018. **12**: p. 13.
276. Soto, F., L. Zhao, and D. Kerschensteiner, *Synapse maintenance and restoration in the retina by NGL2.* eLife, 2018. **7**: p. e30388.
277. Soto, F., et al., *NGL-2 Regulates Pathway-Specific Neurite Growth and Lamination, Synapse Formation, and Signal Transmission in the Retina.* The Journal of Neuroscience, 2013. **33**(29): p. 11949-11959.
278. Nishimura-Akiyoshi, S., et al., *Axonal netrin-Gs transneuronally determine lamina-specific subdendritic segments.* Proc Natl Acad Sci U S A, 2007. **104**(37): p. 14801-6.
279. Matsukawa, H., et al., *Netrin-G/NGL complexes encode functional synaptic diversification.* J Neurosci, 2014. **34**(47): p. 15779-92.
280. Zhang, Q., et al., *Diversification of behavior and postsynaptic properties by netrin-G presynaptic adhesion family proteins.* Mol Brain, 2016. **9**: p. 6.
281. Zhang, Q., et al., *Netrin-G1 regulates fear-like and anxiety-like behaviors in dissociable neural circuits.* Sci Rep, 2016. **6**: p. 28750.

282. Llorca, A. and R. Deogracias, *Origin, Development, and Synaptogenesis of Cortical Interneurons.* Frontiers in Neuroscience, 2022. **16**.
283. Ratliff, J., E. Marder, and T. O'Leary, *Neural circuit robustness to acute, global physiological perturbations.* bioRxiv, 2018: p. 480830.
284. Meerabux, J.M., et al., *Human netrin-G1 isoforms show evidence of differential expression.* Genomics, 2005. **86**(1): p. 112-116.
285. Suto, F., et al., *Interactions between plexin-A2, plexin-A4, and semaphorin 6A control lamina-restricted projection of hippocampal mossy fibers.* Neuron, 2007. **53**(4): p. 535-47.
286. Tawarayama, H., et al., *Roles of Semaphorin-6B and Plexin-A2 in Lamina-Restricted Projection of Hippocampal Mossy Fibers.* The Journal of Neuroscience, 2010. **30**(20): p. 7049-7060.
287. Alto, L.T. and J.R. Terman, *Semaphorins and their Signaling Mechanisms*, in *Semaphorin Signaling: Methods and Protocols*, J.R. Terman, Editor. 2017, Springer New York: New York, NY. p. 1-25.
288. Elkins, T., et al., *Genetic analysis of a Drosophila neural cell adhesion molecule: interaction of fasciclin I and Abelson tyrosine kinase mutations.* Cell, 1990. **60**(4): p. 565-575.
289. Chen, L.Y., et al., *Conditional Deletion of All Neurexins Defines Diversity of Essential Synaptic Organizer Functions for Neurexins.* Neuron, 2017. **94**(3): p. 611-625 e4.
290. Grant, S.G.N., *Synapse diversity and synaptome architecture in human genetic disorders.* Human Molecular Genetics, 2019. **28**(R2): p. R219-R225.
291. Li, J., et al., *Alternative splicing controls teneurin-latrophilin interaction and synapse specificity by a shape-shifting mechanism.* Nature Communications, 2020. **11**(1): p. 2140.
292. Heckman, E.L. and C.Q. Doe, *Establishment and Maintenance of Neural Circuit Architecture.* The Journal of Neuroscience, 2021. **41**(6): p. 1119-1129.
293. Hayano, Y., et al., *IgSF11 homophilic adhesion proteins promote layer-specific synaptic assembly of the cortical interneuron subtype.* Science Advances, 2021. **7**(29): p. eabf1600.
294. Kurusu, M., et al., *A screen of cell-surface molecules identifies leucine-rich repeat proteins as key mediators of synaptic target selection.* Neuron, 2008. **59**(6): p. 972-85.
295. Ango, F., et al., *Ankyrin-Based Subcellular Gradient of Neurofascin, an Immunoglobulin Family Protein, Directs GABAergic Innervation at Purkinje Axon Initial Segment.* Cell, 2004. **119**(2): p. 257-272.
296. Ango, F., et al., *Bergmann Glia and the Recognition Molecule CHL1 Organize GABAergic Axons and Direct Innervation of Purkinje Cell Dendrites.* PLOS Biology, 2008. **6**(4): p. e103.
297. Duan, X., et al., *Type II cadherins guide assembly of a direction-selective retinal circuit.* Cell, 2014. **158**(4): p. 793-807.
298. Krishnaswamy, A., et al., *Sidekick 2 directs formation of a retinal circuit that detects differential motion.* Nature, 2015. **524**(7566): p. 466-470.
299. Tai, Y., et al., *Regulation of chandelier cell cartridge and bouton development via DOCK7-mediated ErbB4 activation.* Cell Rep, 2014. **6**(2): p. 254-63.
300. Tai, Y., et al., *Axo-axonic Innervation of Neocortical Pyramidal Neurons by GABAergic Chandelier Cells Requires AnkyrinG-Associated L1CAM.* Neuron, 2019. **102**(2): p. 358-372 e9.
301. Sando, R., X. Jiang, and T.C. Sudhof, *Latrophilin GPCRs direct synapse specificity by coincident binding of FLRTs and teneurins.* Science, 2019. **363**(6429).
302. Luo, B., et al., *ErbB4 promotes inhibitory synapse formation by cell adhesion, independent of its kinase activity.* Translational Psychiatry, 2021. **11**(1): p. 361.
303. Scheiffele, P., et al., *Neuroligin Expressed in Nonneuronal Cells Triggers Presynaptic Development in Contacting Axons.* Cell, 2000. **101**(6): p. 657-669.
304. Biederer, T. and P. Scheiffele, *Mixed-culture assays for analyzing neuronal synapse formation.* Nat Protoc, 2007. **2**(3): p. 670-6.
305. Budreck, E.C. and P. Scheiffele, *Neuroligin-3 is a neuronal adhesion protein at GABAergic and glutamatergic synapses.* Eur J Neurosci, 2007. **26**(7): p. 1738-48.
306. Bomkamp, C., et al., *Mechanisms of PTPσ-Mediated Presynaptic Differentiation.* Frontiers in Synaptic Neuroscience, 2019. **11**.
307. Tsetsenis, T., et al., *Direct visualization of trans-synaptic neurexin-neuroligin interactions during synapse formation.* J Neurosci, 2014. **34**(45): p. 15083-96.
308. Barabási, D.L., et al., *Neuroscience needs network science.* Journal of Neuroscience, 2023. **43**(34): p. 5989-5995.
309. Seguin, C., O. Sporns, and A. Zalesky, *Brain network communication: concepts, models and applications.* Nature Reviews Neuroscience, 2023. **24**(9): p. 557-574.

310. Bullmore, E. and O. Sporns, *Complex brain networks: graph theoretical analysis of structural and functional systems.* Nature Reviews Neuroscience, 2009. **10**(3): p. 186-198.
311. Verney, C., et al., *Independent Controls for Neocortical Neuron Production and Histogenetic Cell Death.* Developmental Neuroscience, 2000. **22**(1-2): p. 125-138.
312. Pak, T.F., J. Pitt-Francis, and R.E. Baker, *A mathematical framework for the emergence of winners and losers in cell competition.* Journal of Theoretical Biology, 2024. **577**: p. 111666.

Tables

Table 1: Mouse lines used in this study

Type	Internal Name	Alternate Name	Strain	Strain #	Source
Driver Lines	Nkx2.1^{Cre}	BAC-Nkx2.1^{Cre}	C57BL/6J-Tg(Nkx2-1-cre)2Sand/J	008661	Jax
	Nkx2.1^{CreER}	Nkx2.1^{CreER} Nkx2.1^{CreERT2}	STOCK Nkx2-1$^{tm1.1(cre/ERT2)Zjh}$/J	014552	Jax
	PV^{Cre}	PV^{Cre}, Pvalb-IRES-Cre	B6;129P2-*Pvalb$^{tm1(cre)Arbr}$*/J	008069	Jax
	Sst^{Cre}	Sst-IRES-Cre	STOCK Sst$^{tm2.1(cre)Zjh}$/J	13044	Jax
	Sst^{CreER}	SOM-CreER, SOM-CreERT2, Sst-^{CreER}, Sst^{CreERT2}	B6(Cg)-Sst$^{tm1(cre/ERT2)Zjh}$/J	010708	Jax
	Nex^{Cre}	Math2/Cre, Neu-rod6-cre, Neu-rod6/Cre, Nex^{Cre}, NEX-Cre	Neurod6$^{tm1(cre)Kan}$	MGI:2668659	Schwab, et al., 2000
Reporter Lines	Ai34	Ai34D Ai34(RCL-Syp/tdT)-D	B6;129S-Gt(ROSA)26Sor$^{tm34.1(CAG-Syp/tdTomato)Hze}$/J	012570	Jax
	Ai14	Ai14 , Ai14D Ai14(RCL-tdT)-D	B6;129S6-Gt(ROSA)26Sor$^{tm14(CAG-tdTomato)Hze}$/J	007914	Jax
Mutant Lines	Reeler	reeler, Reln^{RL}	B6C3Fe *a/a-Reln^{rl}*/J	000235	Eric Olson
Other	SW	Swiss Webster	Tac:SW		Taconic

Table 2: Genotyping primers

reaction/target	primer name	alternate name(s)	sequence
250bp Internal Control	Myog F	*Myog F, 250bp Internal Control Fwd*	TTACGTCCATCGTGGACAGC
250bp Internal Control	Myog R	*Myog R, 250bp Internal Control Rev*	TGGGCTGGGTGTTAGTCTTA
265bp Internal Control	isl1-Flox F1	*isl1-Flox F1, 265bp Internal Control Fwd*	TCT CTC ATT GGG GAA GCA AC
265bp Internal Control	isl1-Flox R1	*isl1-Flox R1, 265bp Internal Control Rev*	TCC AGT GAA GGC CTT CC AGT
Ai34/Ai9	oPDD001	oPDD001, ROSA_Ai--_tgLoxP_F	AGCTGATCCGGAACCCTTAATA
Ai34/Ai9	oPDD002	oPDD002, ROSA_Ai34_hSyn_R1	TTCAGCCGAGGAGGAGTAGT
Ai34/Ai9	oPDD004	oPDD004, ROSA_Ai34_hSyn_F2| Ai34_ea1_F	CCCACCTCCTTCTCCAATCA
Ai34/Ai9	oPDD005	oPDD005, ROSA_Ai--_tdTom_R2, Ai34_ea1_R	CCT TGG AGC CGT ACA TGA AC
Ai34/Ai9	oPDD007	oPDD007, ROSA_Ai--_tdTom_R3, Ai34_ea2_R	TGG AGC CGT ACA TGA ACT GG
Ai34D	12209	*12209, Ai34D Mutant, hSyp*	GGA GTG TGC CAA CAA GAC GGA GA
Ai34D	12210	*12210, Ai34D Mutant, hSyp*	CCA GCC TGT CTC CTT GAA CAC GA
Ai9	oIMR9105	*oIMR9105, Ai9 Mut Fwd, tdTomato*	CTGTTCCTGTACGGCATGG
CreER	CreERT2 F	CreERT2 F	GTGCCTGGCTAGAGATCCTG
CreER	CreERT2 R	CreERT2 R	AGAGACTTCAGGGTGCTGGA
GFP	EGFP-353For	*EGFP-353For , GFP Fwd*	ACACCCTGGTGAACCGCA TCGAG
GFP	EGFP-648Rev	*EGFP-648Rev , GFP Rev*	GCGCTTCTCGTTGGGGTCTTTGC
NexCre	NexCre #4	*NexCre #4, NexCre Forward*	GAGTCCTGGAATCAGTCTTTTTC
NexCre	NexCre #5	*NexCre #5, NexCre WT Reverse*	AGAATGTGGAGTAGGGTGAC
NexCre	NexCre #6	*NexCre #6, NexCre Cre Reverse*	CCGCATAACCAGTGAAACAG
Nkx2.1Cre	14314	*14314, Nkx2.1Cre Transgene Reverse*	CGG TTA TTC AAC TTG CAC CA
Nkx2.1Cre	17400	*17400, Nkx2.1Cre Transgene Forward*	CTC TGG TGG CTG CCT AAA AC
Nkx2.1CreER	14015	*14015, Nkx2.1CreER Common F*	GCC TCC ACT CAA GCC AAT TA
Nkx2.1CreER	14016	*14016, Nkx2.1CreER Wild type Reverse*	CCT GGC CCT GTC TGT ACG
Nkx2.1CreER	oIMR9377	*oIMR9377, Nkx2.1CreER Mutant Reverse*	ATG TTT AGC TGG CCC AAA TG
QIA-CRE	CRE-QIAGEN-FW	*CRE-QIAGEN-FW, QIA-CRE Forward*	GAA CCT GAT GGA CAT GTT CAG G
QIA-CRE	CRE-QIAGEN-REV	*CRE-QIAGEN-REV, QIA-CRE Reverse*	AGT GCG TTC GAA CGC TAG AGC CTG T
Reeler	34153	34153, Reeler, Common, Rxn A,B	CAC CCT TGA CTG AAA CAG GT
Reeler	34154	34154, Reeler, WT, Rxn A	GAA TCT GCT ACA CAG TTG ACA
Reeler	34156	34156, Reeler, Mut, Rxn B	AAG TAT TTT CTA ATT TTA TCA ACT TGC
Reeler O	JBXFwd1	Reeler O Cmn F, JBXFwd1	gtcctcactctgcccttt
Reeler O	JBXRev1(wt)	Reeler O WT R, JBXRev1(wt)	caggaatgaagcagactctc

Reeler O	JBXRev2(mut)	Reeler O Mt R, JBXRev2(mut)	tcacggacaaactgctct
ROSA mut	oIMR9103	*oIMR9103, ROSA mut Rev, WPRE*	GGC ATT AAA GCA GCG TAT CC
ROSA WT	oIMR9020	*oIMR9020, ROSA WT Fwd*	AAG GGA GCT GCA GTG GAG TA
ROSA WT	oIMR9021	*oIMR9021, ROSA WT Rev*	CCG AAA ATC TGT GGG AAG TC
Sex	SRY F	*SRY F, Sex Fwd*	TCATGAGACTGCCAACCACAG
Sex	SRY R	*SRY R, Sex Rev*	CATGACCACCACCACCACCAA
Sst>CreER	9167	*9167, Sst>CreER Mut F - RxnB*	TGT ACG GTC AGT AAA TTG GAC
Sst>CreER	9176	*9176, Sst>CreER Mut R - RxnB*	CAT TGA CAG GTA CCC AAC TG
Sst>CreER	9177	*9177, Sst>CreER WT F - RxnA*	TGC TGA AAC GCA ATG TTT GT
Sst>CreER	9178	*9178, Sst>CreER WT R - RxnA*	TTC CTG GAT AGG GCA GTG AG
SstCre	9989	*9989, SstCre Mutant Forward*	TGG TTT GTC CAA ACT CAT CAA
SstCre	11224	*11224, SstCre Common*	GGG CCA GGA GTT AAG GAA GA
SstCre	11225	*11225, SstCre Wild type Forward*	TCT GAA AGA CTT GCG TTT GG
PV Cre A	17283	*17283, PV Cre A Mutant Forward*	AAA TGC TTC TGT CCG TTT GC
PV Cre A	oIMR8290	*oIMR8290, PV Cre A Wild type Forward*	CAG AGC AGG CAT GGT GAC TA
PV Cre A	oIMR8291	*oIMR8291, PV Cre A Wild type Reverse*	AGT ACC AAG CAG GCA GGA GA
PV Cre A	oIMR9377	*oIMR9377, PV Cre A Mutant Reverse*	ATG TTT AGC TGG CCC AAA TG
PV Cre B	16210	*16210, PV Cre B Wild Type Forward*	CCA TTC CCT CAT CCA CAG AG
PV Cre B	27606	*27606, PV Cre B Mutant Forward*	GAC GCA ATC GTC CGA TCC
PV Cre B	49521	*49521, PV Cre B Common*	CAA ACA AAC TGA ACA GAA ACT CAG G

Table 3: Tamoxifen induction of CreER driver lines

Tamoxifen induction of CreER:			
Tamoxifen dosage	**Final Conc. (µg/g)**	**TAM Stock (mg/mL)**	**Vol. Stock (µL/g)**
Nkx2.1>CreER		*via gavage of E17.5 pregnant dam*	
High	200	20	10
Low	20	2	10
Single Cell	5	1	5
Sst>CreER		*via gavage or IP injection of adult mice*	
Day 1 (2mg)	80	20	4
Day 2 (4mg)	160	20	8
Day 3 (8 mg)	320	20	16

Table 4: Primary antibodies

Target	Host	mono/ Poly	clone	Subtype	Source	Catalog	Dilution
Ankyrin G	Rb	pAb	-	IgG	Synaptic Systems	386 003	1:1000
Gephyrin (pS270, Gp)	Gp	mAb	GpmAb7	IgG1k	Synaptic Systems	147-318	1:200
Kv2.1	Ms	mAb	K89/34	IgG1k	Neuromab	75-014	1:1000
Kv2.2	Ms	mAb	N372B/1	IgG1k	Neuromab	75-369	1:500
MAP2	Chk	pAb	-	IgY	Synaptic Systems	188 006	1:1000
PV	Gp	-	-	-	-		-
PV	Chk	-	-	-	-	-	-
Sst	Rt	mAb	YC7	IgG2b	Chemicon	MAB354	1:300
Synaptophysin	Ms	mAb	SVP-38	IgG1	Sigma	S5768	1:200
VGAT	Rb	pAb	-	IgG	Synaptic Systems	131 002	1:500

Table 5: Secondary antibodies

Host Species	Target Species	Conjugate	Ex (optimal)	Em	Source	Catalog	Dilution
Dk	Chk	DL405	405	452	Jackson Immunoresearch	703-475-155	1:500
Dk	Ms	DL405	405	452	Jackson Immunoresearch	715-475-150	1:500
Gt	Rb	CF405M	405	452	Biotium	20373-500uL	1:1000
Dk	Rb	DL405	405	452	Jackson Immunoresearch	711-475-152	1:500
Dk	Gp	CF488A	488	515	Biotium	20169	1:1000
Dk	Gp	AF488	488	515	Jackson Immunoresearch	706-545-148	1:500
Dk	Ms	AF488	488	515	Jackson Immunoresearch	715-545-150	1:500
Dk	Rb	CF488A	488	515	Biotium	20015	1:1000
Dk	Rt	AF488	488	515	Jackson Immunoresearch	712-545-153	1:500
Dk	Gp	AF647	633	640	Jackson Immunoresearch	706-605-148	1:500
Dk	Rt	AF647	633	640	Jackson Immunoresearch	712-605-153	1:500

Table 6: Saved source metadata fields by category, with data type and source

Below is a list of the metadata fields stored within the metadata table. Columns are as follows:

- 'Category' designates the metadata's source of origin. 'Uninitialized' metadata is independent of the file itself, and designates metaparameters for extraction and file storage. All other metadata is considered 'Initialized,' i.e., derived via Bioformats either directly or indirectly from the image metadata.
- 'Name' designates the field name.
- 'Type' designates the data type. I = integer, S = string, B = Boolean, F = float. [X] indicates the data are stored as a [*SeriesDelimiter*]-delimited array. [*SeriesDelimiter*] is stored as a field in the metadata.
- 'Source' specifies whether the metadata is designated (i.e., is a metaparameter), directly extracted from the metadata, or derived from extracted metadata.

Category	Name	Type	Source
Uninitialized	Index	I	Designated
Uninitialized	SourceDirectory	S	Derived
Uninitialized	SourceFile	S	Derived
Uninitialized	SceneName	S	Derived
Uninitialized	SceneNameRegex	S	Derived
Uninitialized	Format	S	Extracted
Uninitialized	BF-version	S	Extracted
Uninitialized	BF-supported	B	Extracted
Uninitialized	FileGroupOption	S	Extracted
Initialized	BF-Complete	B	Extracted
Initialized	OMEmetadataPath	S	Designated
Initialized	UsedFileCount	I	Extracted
Initialized	UsedFileDelimiter	S	Designated

Initialized	UsedFileNames	S	Extracted
Initialized	IsMultiSeries	B	Derived
Initialized	SeriesCount	I	Extracted
Scene	SeriesIsSequential	B	Derived
Scene	SeriesIsConsistent	B	Derived
Scene	SeriesNumMin	I	Derived
Scene	SeriesNumMax	I	Derived
Scene	TileNamePrefix	S	Designated
Scene	TileNumMin	I	Derived
Scene	TileNumMax	I	Derived
Scene	TileNamePattern	S	Derived
Scene	TileNameRegex	S	Derived
Scene	SeriesDelimiter	S	Designated
Scene	SubSeriesDelimiter	S	Designated
Instrument	MicroscopeType	S	Extracted
Instrument	nDetectors	I	Derived
Instrument	DetectorIDs	[S]	Extracted
Instrument	DetectorModel	S	Extracted
Instrument	Offset	[F]	Extracted
Instrument	DetectorType	[S]	Extracted
Instrument	DetectorZoom	[F]	Extracted
Instrument	ObjectiveID	S	Extracted
Instrument	ObjectiveModel	S	Extracted
Instrument	Immersion	S	Extracted
Instrument	LensNA	F	Extracted
Instrument	NominalMagnification	F	Extracted
Instrument	WorkingDistance	F	Extracted
Instrument	WorkingDistanceUnit	S	Extracted
Series Data	SeriesIndex	[S]	Designated
Series Data	SeriesNames	[S]	Extracted
Series Data	TileNumbers	[S]	Derived
Series Data	TileNames	[S]	Derived
Acquisition	AcquisitionDate	S	Extracted
Objective	ObjectiveSettingsID	S	Extracted
Objective	Medium	S	Extracted
Objective	RefractiveIndex	F	Extracted
Pixel	BigEndian	B	Extracted
Pixel	DimensionOrder	S	Extracted
Pixel	Interleaved	B	Extracted
Pixel	PhysicalSizeX	F	Extracted
Pixel	PhysicalSizeXUnit	S	Extracted

Pixel	PhysicalSizeY	F	Extracted
Pixel	PhysicalSizeYUnit	S	Extracted
Pixel	PhysicalSizeZ	F	Extracted
Pixel	PhysicalSizeZUnit	S	Extracted
Pixel	TimeIncrement	F	Extracted
Pixel	TimeIncrementUnit	S	Extracted
Pixel	SignificantBits	I	Extracted
Pixel	PixelType	S	Extracted
Pixel	SizeX	I	Extracted
Pixel	SizeY	I	Extracted
Pixel	SizeZ	I	Extracted
Pixel	SizeT	I	Extracted
Pixel	SizeC	I	Extracted
Channel	ChannelNumbers	[S]	Derived
Channel	AcquisitionMode	S	Extracted
Channel	ChannelColors	[I]	Extracted
Channel	EmissionWavelengths	[F]	Extracted
Channel	EmissionWavelengthU-nits	[S]	Extracted
Channel	ExcitationWavelengths	[F]	Extracted
Channel	ExcitationWavelengthU-nits	[S]	Extracted
Channel	Fluors	[S]	Extracted
Channel	ChannelIDs	[S]	Extracted
Channel	IlluminationTypes	[S]	Extracted
Channel	ChannelNames	[S]	Extracted
Channel	PinholeSizes	[F]	Extracted
Channel	PinholeSizeUnits	S	Extracted
Channel	SamplesPerPixels	I	Extracted
Channel	ChannelDetectorIDs	[S]	Extracted
Stage position	StagePositionX	[F]	Extracted
Stage position	StageXUnit	[S]	Extracted
Stage position	StagePositionY	[F]	Extracted
Stage position	StageYUnit	[S]	Extracted
Stage position	StagePositionZ	[F]	Extracted
Stage position	StageZUnit	[S]	Extracted
Stitching	SpatialDimensions	I	Derived
Stitching	StagePositionsOverlap	F	Extracted
Stitching	StagePositionsOverlap-Source	S	Designated
Stitching	StitchingPositionsOverlap	F	Designated

Stitching	StitchingPositionsOverlapForced	B	Designated
Stitching	StitchingPositionsStartX	[F]	Derived
Stitching	StitchingPositionsStartY	[F]	Derived
Stitching	StitchingPositionsStartZ	[F]	Derived
Stitching	StitchingPositionsUnit	S	Designated
Stitching	StitchingPositionsFileExists	B	Designated
Stitching	StitchingPositionsDir	S	Designated
Stitching	StitchingPositionsFile	S	Designated
Stitching	StitchingGeneralizedSuffix	S	Designated

Table 7: Sizes of unsupervised bouton classes

Size is the total bouton counts in the designated cluster. % is the percentage of all boutons. Clusters in red did not reach minimum size requirements and were excluded from downstream analysis. Related to Chapter2, Figures 6, 7, and 8.

cluster	size	%
0	190491	14.05%
1	183882	13.56%
2	166980	12.31%
3	159702	11.78%
4	153496	11.32%
5	142826	10.53%
6	135733	10.01%
7	132908	9.80%
8	89933	6.63%
9	130	0.01%
10	80	0.01%
11	21	0.00%
12	11	0.00%
13	2	0.00%
14	2	0.00%
15	2	0.00%
16	2	0.00%
17	2	0.00%

Table 8: Permissions statements for reproduced figures.

Figure	Figure title	Permission statement
1.4	Cortical development stages prior to circuit integration.	Reprinted from Tyson, J.A. and S.A. Anderson, GABAergic interneuron transplants to study development and treat disease. Trends in neurosciences. 2014 37(3): p. 169-177, Copyright © 2014 Elsevier, with permission from Elsevier.
1.5	The MGE lineage of cortical interneurons	Reprinted from Yao, Z., et al., A taxonomy of transcriptomic cell types across the isocortex and hippocampal formation. Cell. 2021, 184(12): p. 3222-3241.e26, Copyright © 2021 Elsevier, with permission from Elsevier.
1.6	Schematic summarizing the major MGE-derived cortical GABAergic interneurons based on their connectivity patterns	Reprinted from Llorca, A. and R. Deogracias, Origin, Development, and Synaptogenesis of Cortical Interneurons. Frontiers in Neuroscience, 2022. 16. Copyright © 2022 Llorca & Deogracias, under Creative Commons Attribution License (CC BY)
1.7	The developmental landscape reflects attractor dynamics	Used with permission of Annual Reviews, from Fishell, G. and A. Kepecs, Interneuron Types as Attractors and Controllers. Annu Rev Neurosci, 2020. 43: p. 1-30. Copyright © 2020 by Annual Reviews; permission conveyed through Copyright Clearance Center, Inc
1.8	The evolution of corticogenesis: inside-out birth and migration patterns yield ontogenicaly-related cortical columns	Reprinted from Kolk, S.M. and P. Rakic, Development of prefrontal cortex. Neuropsychopharmacology, 2022. 47(1): p. 41-57. Copyright © 2022 Kolk & Rakic, under Creative Commons Attribution License (CC BY)
1.9	Lamination of the cerebral cortex	Vogt, C. and O. Vogt, Allgemeine ergebnisse unserer hirnforschung. Vol. 25. 1919: JA Barth. Reproduced with permission from Springer Nature.
1.10	Regional organization of the mouse cortex	Reprinted from Kirkcaldie, M.T.K., Chapter 4 - Neocortex, in The Mouse Nervous System, C. Watson, G. Paxinos, and L. Puelles, Editors. 2012, Academic Press: San Diego. p. 52-111., Copyright © 2012 Elsevier, with permission from Elsevier.
1.11	Spatial patterning of the cortex	Reprinted from Borello, U. and A. Pierani, Patterning the cerebral cortex: traveling with morphogens. Current Opinion in Genetics & Development, 2010. 20(4): p. 408-415., Copyright © 2010 Elsevier, with permission from Elsevier.
1.12	Examples of topographic maps: the pathways transmitting tactile information from the muzzle to the cortex	Staiger, J.F. and C.C.H. Petersen, Neuronal Circuits in Barrel Cortex for Whisker Sensory Perception. Physiological Reviews, 2021. 101(1): p. 353-415. Copyright © 2021 the American Physiological Society. Reproduced with permission from the American Physiological Society.

1.13, A-D	Independence of laminar and regional iden-tity: evidence from the reeler cortex	Reprinted from Wagener, R.J., et al., Thalamocortical Connections Drive Intracortical Activation of Functional Columns in the Mis-laminated Reeler Somatosensory Cortex. Cerebral Cortex, 2015. 26(2): p. 820-837. Copyright © 2015 Wagener, R.J., et al., under Creative Commons Attribution License (CC BY)
1.13, E-F	Independence of laminar and regional iden-tity: evidence from the reeler cortex	Guy, J. and J.F. Staiger, The Functioning of a Cortex without Lay-ers. Front Neuroanat, 2017. 11: p. 54. by permission of Oxford Uni-versity Press.
1.14	General localization of voltage-gated ion channels in a model neuron	Lai, H.C. and L.Y. Jan, The distribution and targeting of neuronal voltage-gated ion channels. Nat Rev Neurosci, 2006. 7(7): p. 548-62., reproduced with permission from SNCSC'.
1.15	Branch and sub-branch organi-zation of morphologically de-fined pyramidal neuron den-dritic domains	Reprinted from Bloss, E.B., et al., Structured Dendritic Inhibition Supports Branch-Selective Integration in CA1 Pyramidal Cells. Neuron, 2016. 89(5): p. 1016-30., Copyright © 2016 Elsevier, with permission from Elsevier.
1.16	Glutamatergic cell in the cortex have layered subcellular com-partments	© 2021 Springer Science+Business Media, LLC, part of Springer Nature from Araya, R., Dendritic morphology and function. Neuro-science in the 21st Century: From Basic to Clinical, 2022: p. 571-606, reproduced with permission from SNCSC'.
1.17	Mechanisms of axon and den-drite guidance	Reprinted from Chen, S.-Y. and H.-J. Cheng, Functions of axon guidance molecules in synapse formation. Current Opinion in Neu-robiology, 2009. 19(5): p. 471-478, Copyright © 2020 Elsevier, with permission from Elsevier.
1.18	Lamina-specific innervation of hippocampal neurons by dis-tinct inputs	Used with permission of Annual Reviews, from Sanes, J.R. and M. Yamagata, Many Paths to Synaptic Specificity. Annual Review of Cell and Developmental Biology, 2009. 25(1): p. 161-195; Copy-right © 2009 by Annual Reviews, permission conveyed through Copyright Clearance Center, Inc.
1.19, A-C	Differential use of ephrin/Eph adhesion molecules in axon guidance during visual system development	Reprinted from Kania, A. and R. Klein, Mechanisms of ephrin–Eph signalling in development, physiology and disease. Nature Reviews Molecular Cell Biology, 2016. 17(4): p. 240-256.Copyright © 2016 Macmillan Publishers Limited. Reproduced with permission from SNCSC.
1.19, D-E	Differential use of ephrin/Eph adhesion molecules in axon guidance during visual system development	Reprinted from Dufour, A., et al., Area Specificity and Topography of Thalamocortical Projections Are Controlled by ephrin/Eph Genes. Neuron, 2003. 39(3): p. 453-465., Copyright © 2003 by Cell Press, with permission from Elsevier.

1.20	Schematic Representation of the Bergmann Glial Fiber Scaffold and CHL1 in Directing Stellate Axons to Innervate Purkinje Dendrites	Reprinted from Ango, F., et al., Bergmann Glia and the Recognition Molecule CHL1 Organize GABAergic Axons and Direct Innervation of Purkinje Cell Dendrites. PLOS Biology, 2008. 6(4): p. e103. Copyright: © 2008 Ango et al. under Creative Commons Attribution License (CC BY)
1.22	Scale-free graph of the iPSD proteome	Reprinted from Uezu, A., et al., Identification of an elaborate complex mediating postsynaptic inhibition. Science, 2016. 353(6304): p. 1123-9. Copyright © 2016, American Association for the Advancement of Science. Reprinted with permission from AAAS.
1.23, upper	Visual representation of the synaptic adhesion molecule code	Reprinted from Llorca, A. and R. Deogracias, Origin, Development, and Synaptogenesis of Cortical Interneurons. Frontiers in Neuroscience, 2022. 16. under Creative Commons Attribution License (CC BY)
1.23, lower	Visual representation of the synaptic adhesion molecule code	Reprinted from Sanes, J.R. and S.L. Zipursky, Synaptic Specificity, Recognition Molecules, and Assembly of Neural Circuits. Cell, 2020. 181(3): p. 536-556., Copyright © 2020 Elsevier, with permission from Elsevier.
1.24	Subcellular specificity of cerebellar basket cell inhibitory connections to Purkinje cells (PkCs)	Reprinted from Williams, M.E., J. de Wit, and A. Ghosh, Molecular mechanisms of synaptic specificity in developing neural circuits. Neuron, 2010. 68(1): p. 9-18, Copyright © 2010, with permission from Elsevier.
1.25	Lamina-specific innervation of hippocampal CA3 neurons by mossy fibers (MF) is mediated by domain-restricted repulsive adhesion molecule interactions	Used with permission of Annual Reviews, from Sanes, J.R. and M. Yamagata, Many Paths to Synaptic Specificity. Annual Review of Cell and Developmental Biology, 2009. 25(1): p. 161-195. Copyright © 2009; permission conveyed through Copyright Clearance Center, Inc.
1.26, A-B	Regional regulation of synaptogenesis yields subcellular specificity	Reprinted from Klassen, M.P. and K. Shen, Wnt Signaling Positions Neuromuscular Connectivity by Inhibiting Synapse Formation in C. elegans. Cell, 2007. 130(4): p. 704-716., Copyright © 207 Elsevier, with permission from Elsevier.
1.26, C	Regional regulation of synaptogenesis yields subcellular specificity	Used with permission ofSpringer Nature BV, from Poon, V.Y., M.P. Klassen, and K. Shen, UNC-6/netrin and its receptor UNC-5 locally exclude presynaptic components from dendrites. Nature, 2008. 455(7213): p. 669-73. Copyright: © 2008; permission conveyed through Copyright Clearance Center, Inc.
1.27	Remodelling of CF and PF synaptic connections onto PCs during postnatal cerebellar development	Reprinted from Kano, M. and T. Watanabe, Developmental synapse remodeling in the cerebellum and visual thalamus. F1000Res, 2019. 8. Copyright: © 2019 Kano M and Watanabe T, under Creative Commons Attribution License (CC BY)

1.28, A	Molecular basis of PF-PC synapse formation and impact of targeted disruptions on PF and CF innervation domains	Reprinted from Mishina, M., et al., Molecular mechanism of parallel fiber-Purkinje cell synapse formation. Frontiers in neural circuits, 2012. 6: p. 90. Copyright © 2012 Mishina, Uemura, Yasumura and Yoshida, under Creative Commons Attribution License (CC BY)
1.28, B	Molecular basis of PF-PC synapse formation and impact of targeted disruptions on PF and CF innervation domains	Kano, M. and M. Watanabe, Synaptogenesis and Synapse Elimination, in Handbook of the Cerebellum and Cerebellar Disorders, M.U. Manto, et al., Editors. 2022, Springer International Publishing: Cham. p. 309-332. Copyright © Springer Nature Switzerland AG 2022. Reproduced with permission from Springer Nature.
1.29	Model for transneuronal regulation of receptor localization by axonal netrin-G proteins	Used with permission ofThe National Academy of Sciences of the USA, fromNishimura-Akiyoshi, S., et al., Axonal netrin-Gs transneuronally determine lamina-specific subdendritic segments. Proc Natl Acad Sci U S A, 2007. 104(37): p. 14801-6. Copyright © 1957; permission conveyed through Copyright Clearance Center, Inc.
5.1	Waddington's visualization of the potential energy landscape.	Used with permission of George Allen & Unwin Ltd., from Waddington, C.H., The strategy of the genes. 1957: Routledge. Copyright © 1957; permission conveyed through Copyright Clearance Center, Inc.
5.2	The complex system of interactions underlying the potential energy landscape.	Used with permission of George Allen & Unwin Ltd., from Waddington, C.H., The strategy of the genes. 1957: Routledge. Copyright © 1957; permission conveyed through Copyright Clearance Center, Inc.

Appendices

Appendix A: Buffer and media formulations:

Dissection and Dissociation media:

Dissection media: HibA

Working solution for dissection and dissociation of mouse brain, and basal solution for papain digestion. Aliquot into 50 mL tubes and store at 4°C.

- 500 mL Hibernate A [Thermo Scientific A1247501]
- 5 mL Pen/Strep

Dissociation media

Working solution for dissociation of cortical tissue. Resuspend one vial of papain in 4.5 mL HibA media (20 U/mL final). Separately, resuspend one vial of DNAseI in 0.5 mL HibA (2000 U/mL working). Activate by incubating for ~20 minutes at 37°C. Add DNAse I (200 U/mL final) after papain activation Resuspended papain and DNAse I solutions can be stored at 4°C for up to one week. Activated papain should be used within 24 hours.

- 5 mL HibA media
- 1 vial papain [Worthington Biochemical LK003178]
- 1 vial DNAseI [Worthington Biochemical LK003172]

Tissue culture media formulations:

Basal media: NBA

Base recipe for plating and growth media. Aliquot into 50 mL tubes and store at 4˚C. Good until most recent expiration date, at least 6 months.

- 500 mL neurobasal-A medium (Life technologies 10888-022)
- 5 mL GlutaMaX [Life technologies 35050-061]
- 5 mL P/S
- 5 ml HEPES

Growth media: NBA/B27 PLUS

make one 50ml tube of media at a time. Good for ~2 weeks at 4°C. NB: Aliquot B27 PLUS into 1ml aliquots upon first thaw and store at -20˚°C.

- 50ml NBA basal media

- 1 mL B27 PLUS [life technologies A3582801]

Plating media: NBA/B27 Plus/10%FBS/Laminin

Thaw laminin at RT or (preferably) 4˚C. Do not refreeze. Add laminin just before use. Good for ~2 weeks if laminin is not added. Discard any excess plating media after laminin addition. NB: Aliquot Laminin into 20 µl aliquots upon first thaw and store at -20˚°C.

- 9 mL Growth media (NBA/B27 PLUS)
- 1 ml FBS
- 20 µL (final 2 µg/L) Laminin [Sigma Aldrich L2020-1MG]

Tissue culture dish and plate preparation:

Tissue Dishes and plates:

- 24-well Imaging plates: µ-Plate 24 Well Black ID 14 mm, Ibidi Cat.No: 82426
- 96-well Imaging plates: µ-Plate 96 Well Black, Ibidi Cat.No: 89626

10X Borate buffer: (0.01M, pH8.5)

Makes 500 mL 10x Borate Buffer. Store at 4°C.

- 1.24g boric acid (Sigma B6768, H3BO3, FW 61.83)
- 1.9g borax (sodium tetraborate, Sigma S9640, Na2B4O7.10H2O, FW 381.4)
- 500 mL dH2O
- pH to 8.4, filter

Poly-D lysine coating solution: PDL

Makes a 20 µg/mL Poly-D lysine solution. Used to coat dishes prior to cell culture. Add final PDL solution to culture dish and incubate at room temperature overnight in a tissue culture hood. Wash 4x with cell culture grade water. NB: Resuspend poly-D lysine hydrobromide (PDL), Mol weight 70-150 kDa [Sigma Aldrich PD0899-50MG] to 500 µg/mL in cell culture grade water, aliquot into 1mL tubes, and store at -20°C.

- 1 mL 10x Borate buffer.
- 1 mL 500 µg/mL PDL
- 18 mL tissue culture grade water.

Perfusion and slice preparation buffers:

Arificial cerebro-spinal fluid (ACSF):

Used to perfuse animals for immunostaining of synapses. Before use, gas with 5% CO_2 / 95% O_2 for 10-15 minutes, then add $CaCl_2$ to 2.5 mM final concentration.

- 125 mM NaCl
- 26 mM $NaHCO_3$
- 25 mM glucose
- 2.5 mM KCl
- 2.5 mM $CaCl_2$*
- 2 mM MgCl
- 1.25 mM NaH_2PO_4
- pH 7.4

* $CaCl_2$ is prone to precipitation. Add it after bubbling and just before use.

30% sucrose buffer:

Used for cryopreservation of tissue. Add tissue to sucrose solution and incubate overnight at 4°C. Tissue will sink to the bottom of the tube, indicating endogenous water has been displaced and the tissue is ready for freezing.

- 70 mL PBS
- 30 g sucrose

Appendix B: Sample nomenclature

General nomenclature structure for sample/image names:

User input at scope:

- **[Experiment ID]**_**[Sample ID(modifier)]**_[Biological conditions(modif.)]_[Staining conditions(modif.)]_

Example output image:

- **[Experiment ID]**_**[Sample ID(modifier)]**_[Biological conditions(modif.)]_[Staining conditions(modif.)]_*[Scene]*_*[Tile]*

Nomenclature Rules:

- **Bold** fields (**[Experiment ID]** and **[Sample ID]**) are required inputs from the user. The [Experiment ID]_[Sample ID] combination *must* constitute a unique string for sample reference. No two samples should have the same (combined) value. Note that sample ID *modifiers are ignored* for uniqueness tests.
- *Italicized* fields e.g. *[Scene]* and *[Tile]* are *automatically generated* and should not be entered by the user.
- Underscores (i.e. _) indicate breaks *between* canonical, ordered information, e.g. between the [Experiment ID] field and the [Sample ID] field.
- The name inputted at the microscope *must* end with an underscore.
- Hyphens (i.e. -) indicate breaks *within* canonical, ordered information. Examples include different genes within the [Genotype] field (e.g. SstCre-Ai34 designating driver-reporter genotypes) or different identifiers within the [Sample ID] field (e.g. 7769-17 designating ear-toe tags of the mouse ID).
- Parentheticals indicate modifiers to the above hyphenated fields, e.g. the (Age&Sex) modifier to [Sample ID], or the (alleles) modifier to [Genotype]. Note that parentheticals are optional (but informative) inputs.

- Only one of the non-bold, non-italicized, i.e. [Biological] and [Staining] conditions fields is required. If this information is not included in full, it must be substituted with a string that references these conditions.

Field definitions:

[Experiment ID]

Type: non-hyphenated date, [year][month][day] **Description**: Designator for the experiment.

- For **cultures**, this value should be the date of seeding.
- For **tissue samples** , this value should be the date of birth or sample acquisition, and should correspond to the animal DOB (or acquision) in the Mouse Book. If these data are missing, use the date of perfusion as a last resort.

[Sample ID(modifier)]

Type: string **Description**: Sample designator

- For **cultures**:
    - This value should be the well, as designated by Row and Column (e.g. B4). If more than one plates were generated on this date, a plate prefix is to be added, e.g. Plate1-B4.
    - If the source sample differs between wells, e.g. B4 is from mouse 7441 and C4 from 7559, these data should be stored as a modifier to this field, e.g. B4(7441).
- For **tissue samples**:

- This value should be the mouse ID as recorded in the Mouse Book. The [ear tag]-[toe tag] ID should be prioritized, but if unavailable, the [cross tag]-[toe tag] ID should be used.

- Sample fixation/perfusion information, such as Days *in vitro* (DIV) or Embryonic(E)/Postnatal(P) age and sex (M|F), should be stored as a modifier, e.g. B4(P4F-DIV21) for culture or 7769-17(P56M) for slides.

[Biological conditions(modifiers)]

Type: string. **Description**: This value stores experimentally relevant biological conditions

- Genotypes and alleles:

  - If the relevant alleles are dominant (and non-deleterious), no gene modifier is required. For example, $Sst^{Cre/Cre}$ and $Sst^{Cre/wt}$ animals show no (known) phenotypic differences, so SstCre is enough information. Recessive alleles should always have both alleles in parentheses, e.g. $RELN^{RL/wt}$ $RELN^{RL/RL}$ should be designated RELN(RL-wt) and RELN(RL-RL), respectively.
  - If the homozygous allele combination is deleterious, modifiers should be used. For example, Sst^{CreER} is a knock-in/knock-out, so $Sst^{CreER/wt}$ acts as a dominant allele and should be labeled SstCreER, whereas $Sst^{CreER/CreER}$ acts as a Sst knock-out and should be labeled Sst(CreER-CreER).

- Treatments and protocol/dosages:

  - Any non-genetic intervention should be included here. If the animal was treated with drug, infected with virus, or grown under specific experimental conditions.

For example, viral transduction of Sst^{Cre};Ai34+ sample with AAV expressing Gephyrin-GFP and lentivirus expressing shErbB4 would be indicated as Sst-Cre-Ai34-AAV(Gphn-GFP)-LV(shErbB4).

- Note that if the treatment refers to a specific genotype (e.g. tamoxifen induction of CreER+ animals), it should be included as a modifier for the genotype, e.g. an $Nkx2.1^{CreER+}$ mouse gavaged with 200ug/g tamoxifen at E17.5 would read as NkxCreER(gE17-200).

[Staining conditions(modifier)]

Type: string. **Description**: This value indicates the stain(s) utilized to visualize the sample.

Required if the sample has multiple distinct staining conditions, e.g. slices labeled with more than one antibody panel.

- **[Primary antibody (secondary dye)]** or **[Fluorescent protein (microscope excitation)]** – suggested for user information. Required if the sample is stained with more than one antibody/dye panel.
- **Slide number** – required if there is more than one slide per staining condition. Ensures scene values are unique within the slide.

[Scene]

Type: string, [string][digit] **Description**: Scene designator, based on microscope setup.

Used interchangeably with 'tile region'. The scene is automatically generated from image metadata and is designated by the user during the microscope imaging setup. The scene (or

tile region) is a set or group of adjacent tiles that can be stitched together into a continuous image. A single imaging session can yield one or more scenes / tile regions, and each scene / tile region can be comprised of one or more tiles.

[Tile]

Type: string, m[digit] **Description**: Tile designator, based on microscope setup.

Automatically generated from image metadata and designated by the user during microscope imaging setup. This is the unit for processing in our pipeline, and can be comprised of multiple component (sub)images. For example, a single tile with (XYCZT) dimensions of 1024 x 1024 x 4 x 40 x 1 consists of 160 single-channel images (4 channels x 40 Z slices), but is treated as a unit within the pipeline.

Nomenclature for slides:

Structure:

[Date of birth]_[Animal ID(Age&Sex)]_[Genotype(alleles)]_ [antibodies(dyes)]_*[Scene]*_*[Tile]*

Example of slide labeling:

1. **[Date of birth]**_[Animal ID(Age&Sex)]
2. [Genotype(alleles)]
3. **[Staining conditions(modifier)]**
4. **[Staining conditions(modifier) cont.]**

The information from the slide labeling should be used to name the images during acquisition. In other words, use a regularized version of the slide label during microscope setup and auto-save.

Example 1: Animal ID is from ear tag. Age&Sex are undocumented. Biological conditions includes only the shorthand Genotype because all alleles are dominant (i.e. no allele modifier is needed). Nkx^{CreER} was induced by gavage with 200ug/g tamoxifen at E17.5, designated as the (gE17-200) modifier. Staining conditions designate primary antibody and (secondary dye). The *[Scene]* value is generated from the user's custom scene-name (*Scene-3*) during microscope setup and an autolabeled value from the image metadata (*TR5*). The *[tile]* values are automatically generated in accordance with the microscope setup.

User input at scope:

- 20221001_7769-17 _NkxCreER(gE17-200)-Ai34_ AnkG(405)-Kv2(488)-SypTdT(561)-Gphn(647)_

Example output image:

- 20221001_7769-17 _NkxCreER(gE17-200)-Ai34_ AnkG(405)-Kv2(488)-SypTdT(561)-Gphn(647)_ *Scene-3-TR5_m4*

Example 2: Animal ID is from cross and toe tags. Age&Sex are not documented. Biological conditions includes Genotype (shorthand) for dominant & alleles for recessive alleles. The stainining conditions are stored in a metadata file (generated independently and stored separately) and designated by the value S1. The *[Scene]* and *[tile]* values are automatically generated from the microscope setup.

User input at scope:

- 20221001_7559x7577-32 _SstCre-Ai34-RELN(RL-RL)_S1_

Example output image name:

- 20221001_7559x7577-32 _SstCre-Ai34-RELN(RL-RL)_S1_*TR3_m4*

Automated Tile Grouping

The automated tile grouping code in our pipeline relies upon the structured outlined above to identify categorical and hierarchical relationships between images. In order to understand how we identify these relationships and how we use them, we first have to go back into the nomenclature and the metadata to understand the utility and limitations of each, and how we exploited both to increase efficiency and to facilitate processing stages. Presented below is the nomenclature structure, as well as two real examples of its use from our dataset.

Nomenclature structure:

[Experiment ID]_[Sample ID]_[Biological conditions]_[Staining conditions]_[Scene]_m[Tile]_[variable stem ID]

example 1: 20201120_7441_SstCre-Ai34_AnkG-Kv2-Gphn _TR1_m03_BatchQ75NZ.tif

example 2: 20210302_A1_SstCre-Ai34_AnkG-Kv2-Gphn _TR1_m04_xyCorr-EDF.tif

Based on naming patterns alone, we can automatically classify/group tile data based on their experiment, sample, scene, genotype, and staining conditions. We know, for instance, that the tiles above are from different experiments. And since the experiment-sample-scene-tile relationship is hierarchical, we also know that no matter what follows the experiment field, these data are from different samples, scenes, and tiles. In contrast, the *[Biological conditions]* and *[Staining conditions]* fields are categorical and cut across these hierarchies. For example, these samples utilized the same driver and reporter alleles (Sst^{Cre} and Ai34, respectively) and were stained using the same antibodies (AnkG, Kv2, and Gphn), and are thus grouped by these shared properties. Another less apparent category embedded in the *[Sample ID]* field is the *sample format*, i.e. whether data derives from slice or from culture. As mentioned above, slice and culture samples follow different naming conventions, with slice samples containing a 4-digit number (the mouse ear tag ID#) and culture samples following a [A-D][1-6] pattern (the sample's well ID on the tissue culture plate). Based on these patterns, we know that example1 is from slice, whereas example2 is from culture. Importantly, these embedded categories are extensible: adding a new sample format simply requires implementing an additional naming convention, e.g., oA3 for the organoids in well A3 or C5-sl2515 for organotypic slice culture from mouse 2515 in C5.

There are instances, however, when nomenclature alone is insufficient to infer that tiles belong to the same group. To provide a real life example, we set up a series of imaging sessions to image culture samples in a 24-well plate. Each acquisition session acquired images from multiple wells (i.e. multiple samples), with multiple scenes per well. Since any given imaging session outputs only a single image file, we were unable to designate the full sample nomenclature at the scope.

Instead, we used the tile region name to document the source well during acquisition, then post-hoc split the images by scene and renamed the files to accord with our nomenclature rules. In other words, the images from these wells had different scene names and sample names but were all part of the same acquisition series. To complicate this picture further, despite using similar acquisition parameters, slight differences in the amount of oil on the objective led to systematic differences in image intensity between one acquisition series and another. That is, neither the experiment grouping, nor the sample grouping reflected the true relationship between the tiles acquired within the same session (and so neither could be used as a grouping parameter to delineate the systematic error). The image metadata, however, provides a solution: every tile from the same acquisition series shares the same time stamp within the AcquisitionDate field. This time stamp is to the millisecond, and so, for all intents and purposes, is a unique barcode for any given acquisition series. We are therefore able to use this value to group all tiles acquired in the same imaging session, regardless of potential naming differences. Moreover, since it is not possible to change acquisition parameters or slides/plates in the middle of an acquisition series, all images within a series must share the acquisition parameters and be from the same slide/plate. We utilize this acquisition series grouping during the Image correction stages of the pipeline discussed below.