Multidimensional analysis of cortical interneuron synaptic features reveals underlying synaptic heterogeneity

Patrick D. Dummer[1], Dylan I. Lee[2#], Sakib Hossain[1#], Runsheng Wang[1], Andre Evard[1], Casey M. Schneider-Mizell[3], Vilas Menon[2], Edmund Au[1,4*]

[1] Department of Pathology and Cell Biology, Columbia Irving Medical Center, New York NY, 10032.
[2] Department of Neurology, Columbia Irving Medical Center, New York NY, 10032.
[3] Allen Institute for Brain Science, Seattle WA, 98109.
[4]Columbia Translation Neuroscience Initiative Fellow, Columbia Irving Medical Center, New York NY, 10032.
[#] These authors contributed equally to the work.
[*] Correspondence to: ea2515@cumc.columbia.edu

Abstract:

Cortical interneurons are a diverse cell population, and it stands to reason that they deploy a diverse array of synapse types to fulfill their many roles in cortical function. However, little is known about interneuron synaptic diversity because we lack methods to extract features from synapse that might allow us to distinguish subgroups. Here, we develop an approach to aggregate image features from fluorescent confocal images of interneuron synapses and their post-synaptic targets, in order to characterize the heterogeneity of synapses at fine scale. We started by training a model that recognizes pre- and post-synaptic compartments and then determines the target of each interneuron synapse in the image. To accomplish this task, the model extracts hundreds of spatial and intensity features from each analyzed synapse, constructing a multidimensional data set, consisting of millions of synapses. We next used these image features to perform an unsupervised analysis on this dataset, and uncovered novel synaptic subgroups. These subgroups were spatially distributed in a highly structured manner that revealed the local underlying topology of the postsynaptic environment. Dendrite-targeting subgroups were clustered onto subdomains of the dendrite along the proximal to distal axis. Soma-targeting subgroups were enriched onto different postsynaptic cell types. Our analysis also discovered that the two main subclasses of interneurons, basket cells and somatostatin interneurons, utilize distinct strategies to enact inhibitory coverage. Thus, we present an image-based approach to study interneuron synapses that generates multidimensional data sets of synaptic

features. This approach helped uncover novel interneuron biology, and begins to establish a conceptual framework for studying interneuron synaptic diversity.

Introduction:

Cortical interneurons are a highly diverse cell population and our understanding of that diversity has accelerated as a result of high throughput transcriptomic-based studies in recent years [1-3]. A knowledge gap exists, however, for the diversity of cortical interneuron synapses. A growing body of evidence links cortical interneuron dysfunction to neuropsychiatric illness, a constellation of disorders typified by defects in synaptic function [4-8]. Furthermore, an enhanced understanding of interneuron synaptic diversity is key to linking transcriptomic cellular identity and the diverse functions of interneurons in the cortex. Thus, a framework for categorizing interneuron synaptic heterogeneity would be a powerful tool for revealing interneuron function in the healthy brain and how interneuron dysfunction contributes to disease etiology.

Our understanding of interneuron synaptic diversity is hampered by a lack of high throughput methodologies that can extract multidimensional data from individual synapses the way that single cell RNA-seq (scRNA-seq) data allows for the analysis cellular diversity. To address this, we reasoned that synapses are, at their core, a spatial apposition between 2 neurons. Thus, an image-based spatial analysis approach might allow us to extract biologically meaningful features from synapses. These features, in turn, could be analyzed at a population level to reveal underlying interneuron synaptic heterogeneity.

To develop our imaged-based approach, we made use of 2 well-studied principles in interneuron synaptic connectivity. First, interneurons form connections onto different cellular compartments of their postsynaptic partners – dendrite, soma and axon initial segment (AIS). Second, different interneuron subclasses exhibit strong target preferences for each compartment. Somatostatin-positive (Sst) interneurons prefer to target dendrite [9,10], basket cells (BCs) largely target soma and proximal dendrite [11,12], and chandelier cells (ChCs) have a strong preference for AIS [13,14]. The other important feature of interneuron synapses is that they form the same series of circuit motifs throughout the cortex in a repetitive fashion [15]. This repetitive connectivity lends itself to studying interneuron synapses at a population level where we can sample from a given area of the cortex and be reasonably confident that we will obtain a representative sample of all types of interneuron synapses.

Given these prior findings on interneuron synapse diversity, we therefore reasoned that representative imaged fields of interneuron synapses could serve as starting material for a computer model that was trained to perform a relatively simple task: assigning the target compartment for every synapse imaged. In the process of performing this supervised classification task, the model would extract multiple features from each synapse analyzed, which are then amenable to unsupervised analysis. An important requirement is that the features extracted by the model are themselves based solely on local image statistics, and not sub-selected or refined specifically for the classification task. Such a model would concurrently generate a rich, multidimensional dataset, which

we could leverage in unsupervised analyses to reveal underlying synaptic heterogeneity that is invisible to the human eye.

In this study, we start with standard 4-channel confocal images of interneuron synapses and their postsynaptic compartments. From this, we devised an automated, high-throughput method to extract spatial and intensity metrics between pre- and post-synaptic features. Using these features, we found that the trained model could assign targets for each synapse, verified by human scoring and at the scale of hundreds of thousands of synapses per analysis. Having generated this data set, we performed an unsupervised cluster analysis based on the features assigned to each synapse, which revealed multiple subgroups of dendrite- and soma-targeting synapses, We mapped these back onto the original raw images in order to glean biological insights and found that interneuron synaptic subgroups spatially segregate in a pattern that reveals the complex topology of the postsynaptic environment.

Results:

Interneurons target different compartments of the postsynaptic neuron – dendrite, soma and AIS. To establish an image-based system that can automatically assign a target for each synapse, we first validated an immunohistochemical strategy to label target compartments that would work in both cortical sections as well as in dissociated cortical culture (Figure 1A). To label soma and proximal dendrite, we used a combination of antibodies against voltage-gated potassium channels, $Kv_{2.1}$ and $Kv_{2.2}$ (collectively Kv2), combined in the same channel [16]. To label axon initial segment, we used an antibody against Ankyrin G [17]. Labeling dendrite was more challenging, and our attempts with MAP2 were unsuccessful because of uneven signal intensity and the fact that MAP2 only labels dendritic shaft (Supplemental Figure 1). While inhibitory input is primarily onto dendritic shaft, there remains a sizeable portion that innervates spine and neck [18,19]. As an alternative, we opted for a negative selection approach: we used an antibody against a pan-inhibitory postsynaptic scaffold protein, gephyrin and then relied on the absence of Kv2 and Ankyrin G to define the dendritic target compartment.

A second necessary component for our image-based system is to label presynaptic interneuron boutons. For this, we made use of a Cre-dependent conditional reporter allele (Ai34 [20]) that expresses synaptophysin-tdTomato (syn-tdT). This allows us to label synapses with specificity through various Cre driver lines. To validate the Ai34 reporter, we crossed it to Nkx2-1^{Cre}, a pan MGE lineage interneuron Cre driver (Figure 1B) and Nex^{Cre}, a pan pyramidal cell Cre driver (Figure 1C) and counterlabeled with inhibitory

(Vgat and gephyrin) and excitatory (PSD95) synaptic markers. Consistent with our expectations, the vast majority of Nkx2-1^{Cre}; Ai34 labeled syn-TdT puncta were co-labeled with Vgat (97%) and apposed to gephyrin puncta (96%) while not showing a close spatial relationship with PSD95. In contrast, Nex^{Cre}; Ai34 labeled syn-tdT puncta were nearly all juxtaposed with PSD95 while not associating with gephyrin or Vgat.

Next, we crossed Ai34 to interneuron Cre driver lines and immunolabeled with postsynaptic markers for targets (Supplemental Figure 2A-C). As expected, syn-tdT puncta were enriched in Kv2-positive soma in Pv^{Cre} cortical sections while in Sst^{Cre} cortical sections, syn-tdT was enriched with gephyrin puncta in the absence of Kv2. We did not observe syn-tdT puncta present in layer 1 in Pv^{Cre}, but did so in Nkx^{Cre} and Sst^{Cre} cortical sections. This was as expected since Nkx^{Cre} and Sst^{Cre} label Sst interneurons that densely innervate distal pyramidal neuron dendrites.

Interneuron subclasses have strong target preferences – Sst interneurons for dendrite, BCs for soma and proximal dendrite and ChCs for AIS. We wished to label each subclass specifically in order to serve as a positive control for our approach. For Sst interneurons, we could simply use Sst^{Cre} but for BCs and ChCs, we needed an alternate approach since current Cre driver lines cannot distinguish between BCs and ChC (both of the parvalbumin- (Pv) positive lineage). Previously, it had been demonstrated that the Nkx2-1^{CreERT2} driver with a late gavage of tamoxifen (e17.5) labels ChCs with high enrichment in medial prefrontal cortex (pfc) [21]. By late embryogenesis, interneuron

production is primarily of the Pv lineage [22]. We therefore hypothesized that late gavage Nkx2-1^{CreERT2} labeled interneurons would be largely Pv-positive and while there was a reported enrichment for ChCs in medial pfc, there might be pockets of enrichment for BCs elsewhere in the cortex. To test this, we performed an e17.5 gavage on Nkx2-1^{CreERT2}; Ai9 mice and analyzed at P40 (n=8 mice). We immunolabeled for Sst and Pv and found that vast majority of fate mapped interneurons were Pv-positive (94%) (Supplemental Figure 3A,B) . We imaged multiple sections in the coronal plane throughout the rostral-caudal axis of cortex and applied an image analysis script that distinguishes ChC cartilage puncta from all other puncta (Supplemental Figure 3C). From this analysis we identified hotspots for labeled ChCs and areas where cartridges were absent (Supplemental Figure 3D, E). And since nearly all labeled neurons were Pv-positive, we surmised that these pockets were enriched for BCs. We immunolabeled Nkx2-1^{CreERT2}; Ai34 late gavage mice with postsynaptic targets and found that the apposition of syn-tdT puncta agreed with our identified hotspots enriched for BCs and ChCs (Supplemental Figure 2D, E).

Previous work on interneuron synaptic specificity suggest that interneuron synaptic targeting relies upon a strong intrinsic component where pre- and post-synaptic molecules confer specificity in a lock and key fashion [23-26]. Therefore, we surmised that interneurons may retain a degree of synaptic specificity even in dissociated culture where the cells are admixed and interneuron neurons must re-find their appropriate targets (Figure 1D). To further test assertion, we generated dissociated cortical cultures

from Sst^{Cre}; Ai34 and Nkx2-1^{CreERT2}; Ai34 pups where enriched BC and ChC cortical areas were isolated for separate cultures of labeled Sst, BCs and ChCs. Counterlabeled with postsynaptic target markers, we found that there appeared to be distinct differences in the target preferences for each of the 3 culture conditions that were consistent with subclass preferences for dendrite, soma and AIS, respectively (Figure 1E).

Having established labeling strategies for postsynaptic targets and a Cre-dependent labeling strategy for subclass-specific presynaptic puncta, we proceeded with training a model to assign targets to all syn-tdT puncta. To do so, we decided to take an iterative approach: first train the computer on images obtained from dissociated culture where targeting is more unambiguous. This first round of training would then provide a sound foundation for a second round of training in cortical sections where the environment is more complex.

However, we noted that even in dissociated culture, the raw immunohistochemical signal has nuanced features that that the human eye can recognize and account for but would prove difficult for a model to discern (Supplemental Figure 4). For instance, with soma targets, Kv2 signal is punctate and combined with presence of gephyrin puncta to define postsynaptic sites (Supplemental Figure 5A). With AIS, Ankyrin G signal is absent at the postsynapse and, due to the phospho-specific immunoreactivity of our gephyrin antibody [27], gephyrin is also absent at AIS synapses but Kv2 puncta are

present. Therefore, we defined AIS postsynapses as the localized absence Ankyrin G and the presence of a Kv2-positive puncta in high Ankryrin G+ AIS (Supplemental Figure 5B).

This type of subtlety necessitated that we train the model to define pre- and post-synaptic structures using features from multiple raw channels. To accomplish this, we utilized ilastik, a machine learning-based image analysis software program [28,29]. ilastik functions by defining images as probability maps. That is, each pixel in an image is defined as belonging to user-defined categories. Through user training (pixel training), ilastik learns to assign each pixel a probability score of belonging to each category and in so doing, a probability map is established that defines the most likely categorical representation of the image. Thus, through a series of probability maps (Figure 2A-D), we trained ilastik to recognize each postsynaptic target as an amalgamation of 7 probability channels. We also proceeded to define the presynaptic environment through 4 additional probability channels. Combined with the original 4-channel confocal image, ilastik works with 15 channels in total to define each image we capture.

The 4 probability channels defining the presynaptic environment were used to perform object segmentation using FIJI (Figure 2E, Supplemental Figure 6A). This defines each syn-tdT puncta as separate objects to be individually assessed downstream. Each object was used for the next round of training (object training) in which we train ilastik to recognize each syn-tdT puncta as targeting dendrite, soma, AIS, unknown or not a

synaptic punctum (Figure 2F,G; Supplemental Figure 6B-D). This portion of the training was greatly aided by starting with image datasets obtained from dissociated monolayer culture where targeting is clearer, synapse density is lower (by approximately 10-fold) and the pre- and post-synaptic structures are effectively 2-dimensional.

After pixel and object training in dissociated culture, we performed a second round of pixel and object training on images of cortical sections. Since the model at that time was already reasonably proficient at recognizing targets, our second round of training in cortical sections mainly focused on slight correction and refinement for both pixel and object. In total, our pixel training dataset consisted of 297 images (213 culture, 84 cortical sections) with ~180 000 training points. Our object training dataset consisted of 198 images (146 culture, 52 cortical section) with ~12 500 training points.

Our model provided pixel probability maps and target assignment for syn-tdT puncta that appeared visually appropriate for both dissociated culture and cortical sections (Supplemental Figure 7). Additionally, the large majority of puncta were definitively being assigned a target (> 98.5% of puncta were assigned a target). However, a true test of the classifier model's performance was to compare its results to manual human user scoring. We started with a new set of images outside of our training data set and randomly selected ~5400 syn-tdT puncta for user assessment (4673 from culture, 796 from cortical sections). A user was trained on how to navigate ilastik and access the pixel probability maps generated from the image so that the user and the computer had

access to the same information for assigning synapse targeting. The process of user assignment of all 5400 boutons required 75 hours, compared to several minutes for the model. Across all syn-tdT puncta, we found strong concordance (93.8% overall) between human and computer calls overall and that the accuracy was similar between boutons scored in culture (94.3%) versus cortical sections (90.3%) (Supplemental Figure 8).

Having developed a model in which we were confident in the target (dendrite, soma, AIS, unknown) calls made, we used it to assess synaptic targeting of interneuron subclasses in the intact cortical sections versus in dissociated culture. Prior work demonstrates that interneuron synaptic specificity is regulated by transsynaptic adhesion molecules and, thus, employs a strong intrinsic component [23-26]. Interneurons in dissociated culture lack multiple extrinsic cues for synaptic targeting: cortical structure, appropriate sensory input, network activity and paracrine factors produced in the cortical milieu. As such, cultured interneurons must principally rely on intrinsic mechanisms to form appropriate synaptic connections. Thus, we could examine the extent to which cell-intrinsic mechanisms dictate interneuron target specificity by comparing synaptic targeting in culture versus in cortical sections.

First, we compared targeting in cortical sections between 4 interneuron Cre driver lines: 1) Nkx2-1^{Cre}, which labels the entire MGE lineage (Sst, BC and ChC); 2) Pv^{Cre}, which labels the Pv lineage (BC and ChC); Sst^{Cre}, which labels the Sst lineage (Sst); and

Nkkx2-1^{CreERT2}, which labels primarily Pv-positive interneurons with BCs and ChCs enriched in different cortical regions (Figure 3A,B). We found significant differences in targeting between Nkx2-1 and Sst^{Cre} and between Sst^{Cre} and Pv^{Cre}. Regions enriched for BCs and ChCs in Nkx2-1^{CreERT2} also showed strong and significant preferences for soma and AIS, respectively. It is important to note that we also observed Sst and BCs forming targets onto targets other than dendrites and soma, which was expected in line with proportion of 'non-canonical' synapses reported in prior studies [9-11].

Second, we assessed targeting in dissociated cultures from Nkx2-1^{Cre}, Sst^{Cre} and Nkx2-1^{CreERT2} BC-enriched and ChC-enriched regions (Figure 3C,D). Here, we found a trend towards more dendrite targeting between Sst^{Cre} and Nkx2-1^{Cre} cultures and a significant increase in soma targeting between Nkx2-1^{Cre} and Nkx2-1^{CreERT2} cultures, consistent with known target preferences for interneuron subclass labeled in each line. A direct comparison between Sst, BC and ChC cultures showed that the target preferences between each subclass was significantly different from each other. While we observed an enrichment for AIS targeting in ChC regions, it was not close to the near 100% AIS target fidelity thought to be exhibited by ChCs [13,14]. We ascribe this discrepancy to the presence of non-ChC synapses in the same field and the relatively low number of synapses present in AIS-targeting cartridges versus other interneuron synapses, which served to dilute the AIS targeting numbers further.

Next, we compared targeting in cortical sections versus their counterpart in culture (Figure 3E). With Sst^{Cre}, overall targeting was similar, with an enrichment of dendrite targeting found in cortical sections. When we separated targeting across cortical layers we found that the elevated dendrite targeting was due to layer 1. This suggests that Sst interneuron targeting relies mainly on intrinsic mechanisms except in layer 1 which is lost when the cortex is dissociated. With cultured BCs target fidelity was not as robust in cultures compared with cortical sections and we observed a trend towards decreased in soma targeting compared to sections. Similarly, AIS targeting was decreased in ChC cultures.

Next, we examined targeting in cortical sections as a function of cortical depth (Figure 3F). We found that soma and dendrite target is not even across layers. As expected, we observed a strong enrichment of dendrite targeting in layer 1. This was present in both $Nkx2\text{-}1^{Cre}$ and Sst^{Cre} driver lines and reflects the fact that both label the Sst lineage that innervates pyramidal dendritic tufts. With Pv^{Cre}, we found an increase in soma targeting in layers 2/3 and a decrease in soma targeting in layer 6. This was also observed in the $Nkx2\text{-}1^{Cre}$ driver line but not as pronounced since the Cre driver labels both Pv and Sst lineages. The enrichment of soma targeting was absent in Sst^{Cre}. It seemed likely that layer variations in interneuron synaptic targeting reflected differences in the underlying postsynaptic environment. To test this, we crossed Sst^{Cre}; Ai34 to the *Reeler* mutant allele. In *Reeler* mutants, cortical structure is severely disrupted with deficits in inside-out radial migration and pyramidal cell polarity [30]. This was confirmed by

immunohistochemistry, which revealed a general inversion in cortical layering and disrupted pyramidal cell polarity (Supplemental Figure 9). Analysis of synaptic targeting as a function of cortical depth in Sst^{Cre}; RL/RL mice revealed gross differences in targeting across layers, including a loss of dendrite-exclusive targeting in superficial cortex and a loss of structure along the superficial to deep axis.

In order to assign a target for each synapse, ilastik takes into consideration measurements of the syn-tdT puncta itself and its spatial and intensity relationship to the 4 raw confocal channels and the 11 probability channels that define pre- and post-synaptic compartments. In total, ilastik assigns 696 metrics to each syn-tdT object, generating a highly multidimensional dataset from the millions of synapses analyzed; it is important to note that these features are not pre-selected based on their ability to discriminate the three targeting classes from each. As such, this allowed us to make use of this dataset to analyze cortical interneuron synapses at a population level in an unsupervised way. Our approach parallels those used in single cell RNA-seq (scRNA-seq) analysis, but instead of transcript identity and transcript abundance, our features are comprised of spatial and intensity metrics and their convolved values. We therefore re-purposed clustering approaches used in the analysis of scRNA-seq to identify underlying synaptic heterogeneity that would be invisible to the human eye.

To begin, we performed an unsupervised analysis on synapses from $Nkx2\text{-}1^{Cre}$ and Sst^{Cre} dissociated culture datasets. Synapses in culture are largely 2-dimensional and

as a result more uniform in their immunolabeling, making the dataset inherently less noisy. Second, we reasoned that downstream analysis would be easier in culture since the relationships between synapses and targets would be easier to interpret in 2D monolayer culture rather than the more complex environment of a cortical section.

We used a standard workflow for unsupervised analysis, which consisted of dimensionality reduction followed by clustering. Briefly, we took the object feature metrics computed by ilastik as input, z-scored each feature, performed dimensionality reduction using a standard autoencoder, and ran the Louvain community detection algorithm on the Euclidean distances in autoencoder space to cluster objects putative classes of boutons (more details in Materials and Methods). We tested the robustness of the autoencoder using different seeds. We calculated reconstruction loss with mean squared error and observed similar losses across multiple seeds using different machines (3.7% $\pm$ 0.13 across 15 seeds). Having established that the autoencoder approach was robust for dimensionality reduction, we then generated Uniform Manifold Approximation and Projection (UMAP) plots, which employ a non-linear transformation to embed objects into a 2-dimensional space for visualization. Overlaying the supervised classification predictions on the UMAP plot revealed that the different target classes (dendrite, soma and AIS) group together, indicating that the 2-D embedding for visualization can distinguish between syn-tdT puncta with different targets (Figure 4A). For our bouton clustering, we used two assessments to optimize over parameters: 1) comparison of intercluster and intracluster dispersion maximizing the Calinski-Harabasz

score and minimizing the Davies-Bouldin score, and 2) segregation of annotated target classes, with each individual cluster enriched for a single target (Supplemental Figure 10). Through this process, we arrived at our static model grouping synapses in dissociated culture, in which we identified 9 clusters: 6 dendrite-targeting subgroups, 2 soma-targeting subgroups, and 1 AIS group (Figure 4B). Based on these subgroups, we identified the image-based features that most clearly delineated each subgroup (Figure 4C) and the top 20 features with the greatest variation between subgroups (Supplemental Figure 11).

Although we used multiple metrics and some prior information to ensure that our bouton clusters were robust, the question remains as to whether these groupings have biological significance beyond the major targeting classes. To tackle this question, we examined the spatial organization of our subclusters, with the goal of assessing whether there was non-random distribution. This was possible by tracking boutons from each synaptic subgroups back onto their source images via their x, y, z coordinates within the source tile.

We first ran a spatial compactness analysis, whereby we assessed the extent to which boutons from a subcluster were more likely to be spatially proximal to each other, as opposed to boutons from a different cluster. Visual examination revealed that synapses belonging to the same subgroup were likely localized in space. We then performed a systematic compactness analysis, whereby we identified the 10 nearest neighbors for

each bouton (using a modified Louvain approach), and assessed whether these neighbors were enriched for boutons of the same cluster, as compared to the null distribution of all bouton clusters (Figure 5A). By organizing the mean local enrichment scores for each bouton class versus all others, we made 2 observations: First, the spatial distribution of synaptic subgroups is significantly non-random, in that boutons of the same cluster tend to group together in space. This is not surprising for the AIS cluster, given the spatial compartments its boutons target as compared to the other clusters. By contrast, it was not a given that the two soma-targeting clusters and the dendrite-targeting clusters would also show spatial segregation with respect to their counterparts. Second, it was clear that some subgroups were spatially correlated or spatially anti-correlated with other subgroups (Supplemental Figure 12A), suggesting that individual clusters might show higher-order spatial organization.

Given that these relationships are founded on subcellular targeting of distinct compartments, we next asked whether there is a natural, optimized sequence for these bouton classes. To do so, we reasoned that in an optimized sequence, adjacent subgroups would be optimally spatially correlated, that is, the optimal sequence would have a maximal sum of off-diagonal values from the enrichment matrix (Supplemental Figure 12). We calculated the sums of the off-diagonal values for every permutation of the 9 synaptic subgroups to empirically determine the optimal sequence, which we found to be 8, 4, 5, 0, 1, 3, 7, 2, 6 (Figure 5B, Supplemental Figure 12B). By re-ordering the enrichment matrix according to this optimized sequence, we find that the AIS cluster

is first, then the 2 soma subgroups, then the 6 dendrite subgroups (Figure 5C). A closer examination of the dendrite subgroups revealed additional substructure. We found 3 pairs of subgroups that showed high spatial correlation with each other and, with each successive pair, an increasingly greater anti-correlation with soma and AIS subgroups. That is, our analysis indicated that dendritic subgroups are spatially arranged along the proximal to distal axis of the dendrite (in relation to the soma). Indeed, when we re-color the dendritic subgroups as red-orange-white heatmap and re-map them onto the raw image, the proximal-to-distal arrangement is clear (Figure 5D).

Our findings suggest that arrangement of synaptic subgroups reflect a topographic map of the postsynaptic environment. To further test the robustness of this finding, we performed an independent clustering of the boutons, using a different dimensionality reduction technique (principal component analysis) and a different clustering approach (k-means clustering on Euclidean distance in PC space). The clusters we identified using this alternative approach were largely similar to those obtained from the autoencoder+Louvain approach, and when the optimal consensus sequence for PCA+K-means was mapped onto raw images, we observed similar soma-enriched and dendrite-enriched clusters along the proximal to distal axis (Supplemental Figure 13). Satisfied with the robustness of our clustering approach, we next tested whether dendrite-targeting subgroups were similarly arranged on the dendrites of pyramidal neurons in cortical sections.

In the cortex, pyramidal neurons are highly polarized. They bear an apical dendrite that projects superficially, often ramifying in layer 1, and an axon and basal dendrite that project deep. We wondered whether the same proximal-to-distal arrangement of dendrite-targeting subgroups might also be present in cortical sections. We therefore performed a similar synaptic cluster analysis on a Sst^{Cre}; Ai34 cortical column (Figure 5F). We noted that there were approximately twice as many synaptic subgroups identified compared to cultures. We mapped the synaptic subgroups back onto the column and found 6 dendrite-targeting subgroups that were highly enriched in layer 1, corresponding to the distal dendrite of pyramidal neurons. We also found subgroups that were enriched in superficial and deep layers, which may correspond to proximal and mid dendrite. Therefore, the dendritic subgroups we identified in culture and those identified in cortical sections both localize to dendrites in a proximal-to-distal manner. Taken together, our analysis revealed parallels in dendrite-targeting subgroups along the proximal-to-distal axis in dissociated culture and in cortical sections (Figure 5E, F).

Next, we examined the 2 soma-targeting subgroups (4 and 5) (Figure 4B). We wondered if they might localize to different portions of the soma like we found with dendrite subgroups (Supplemental Figure 14). However, when we map subgroups 4 and 5 back onto raw images, their distribution is clumped. Further, by examining an ilastik-based metric that is a proxy for proximity to AIS, we found no differences between subgroups 4 and 5, again suggesting that subgroup 4 and 5 do not map to different subdomains of the soma. We therefore examined the spatial distribution of

subgroups 4 and 5 in relation to the raw confocal channel for Kv2, which labels soma and proximal dendrite. This revealed that subgroups 4 and 5 contact different postsynaptic soma (Figure 6A). We quantified this feature by soma and found the distribution to be tri-modal (Figure 6B): soma that contained >70% subgroup 4 (type 4 soma), soma that contained >70% subgroup 5 (type 5 soma), and soma that contained ~50% mixture of subgroups 4 and 5 (type 4/5 soma). scRNA-seq studies have identified multiple subtypes of BCs and it was possible that different BC subtypes might form synapse that correspond to soma subgroups 4 and 5 or might specifically contact soma type 4, 4/5 and 5. To test this, we cultured Ai34-labeled BCs at low density with unlabeled cortical neurons, which allowed us to examine the axonal arbors of individual BCs. In all individual BCs examined (n=15), we found both subgroup 4 and 5 synapses formed by the same cell and that the same BC formed synapses onto soma that were type 4, 4/5, and 5 (Supplemental Figure 15).

This finding suggested to us that types 4, 5, and 4/5 soma represented distinct postsynaptic populations that dictated the nature of soma subgroup synapses that formed onto it. As such, we wondered if Sst interneurons, which also form soma-targeting synapses but to a lesser degree, also targeted soma in the same fashion. A comparison of Nkx2-1^{Cre} and Sst^{Cre} cultures shows that types 4, 5, and 4/5 soma can be found in the same proportions for genotypes, consistent with the notion of distinct postsynaptic cell types, rather than differences in targeting as a result of interneuron subclass (Figure 6C). If the postsynaptic populations are in fact distinct, we would

expect that types 4, 5 and 4/5 cells to exhibit distinct features. To test this, we quantified the number of synapses formed onto types 4, 5 and 4/5 and found that type 4 soma contained significantly more synapses than type 4/5, with type 5 containing the fewest number (Figure 6D). The increase in synapse number for type 4 was only observed in Nkx2-1^{Cre} cultures, not Sst^{Cre} cultures, suggesting that BCs labeled in Nkx2-1^{Cre} cultures were likely responsible for the elevated soma targeting in type 4 (Figure 6E). Next, we examined the size and Kv2 intensity of type 4, 4/5 and 5 soma. We found that type 4/5 was significantly larger and more had more intense Kv2 signal than types 4 and 5. In our dataset, we found a sizeable number of type 4 and type 5 cells that exclusively contained subgroups 4 and 5. We found that these cells contained significantly fewer synapses, had the smallest soma size and had the lowest intensity of Kv2 (Supplemental Figure 16). Taken together, our data suggest that subgroup 4 and 5 soma-targeting distinguishes postsynaptic neuron types (4, 4/5 and 5), which exhibit distinct features and likely represent distinct postsynaptic populations.

We examined individual Sst, BC and ChCs to test whether interneuron subclasses preferentially employed synaptic subgroups (Supplemental Figure 17). Unsurprisingly, we found that BCs made use of subgroups 4 and 5 more than Sst and ChC and that ChCs preferentially used subgroup 8 (AIS-targeting). With Sst interneurons, there was a significant increase in the use of subgroup 6, which corresponds to the subgroup that targets the most distal portion of the dendrite. But overall, Ssts, BCs and ChCs in culture make use of all 9 subgroups of synapses to varying degrees. This prompted us

to examine the axonal branches of individual interneurons to test whether the sequence of synapse formed by subclasses might be distinct.

We mapped subgroups back onto the raw images of individual BCs and Sst and when we re-colored them as red-orange-white heatmap based on the optimal sequence we determined earlier (Figure 5), there were noticeable differences (Figure 7A, B). BC axonal branches bore synapse subgroups that were consistent with the optimal synaptic sequence. In contrast, Sst axonal branches bore synaptic subgroups that were often non-sequential. To quantify this, we developed a synaptic 'sequentiality' score that accounts for the subgroup identity of each synapse along an axonal branch and counts the average number of 'jumps' from the optimal sequence for each successive synapse (Supplemental Figure 18). For example, if a subgroup 3 synapse is followed by a subgroup 6 synapse, this would constitute a jump of 3 (the optimal sequence is 8, 4/5, 0, 1, 3, 7, 2, 6,) (Figure 7C). We quantified the average sequentiality score for axonal branches from individual BCs and Ssts (n=10 and n=12 cells) and found that BCs axonal branches were significantly more sequential than Sst axonal branches (Figure 7D, E). A potential confound, however, was that BCs are far more likely to form multiple successive soma-targeting synapses, which would result in successive jumps of 0, depressing their sequentiality score. We therefore performed a secondary analysis in which we only scored sequentiality for contiguous stretches of axon containing dendrite-targeting subgroups (Figure 7F). Even with this more stringent measure, BC axonal branches are still significantly more sequential than Sst. We interpreted these data to

suggest that BCs form synapses along the processes of postsynaptic partners, thereby resulting in synapses that follow their topology and are, thus, more sequential. In contrast, Sst axons form a small number of synapses onto each postsynaptic partner and then 'skip' to the next neuron resulting in synapses that are more randomized.

To test this assertion further, we analyzed single cell reconstructions from a large electron microscopy volume of mouse visual cortex [31]. A collection of 58 perisomatic targeting cells (putative BCs) and 54 dendrite-targeting cells (putative SSTs) along a column have been extensively proofread [32], with hundreds to thousands of synaptic outputs for each cell, focusing only on synaptic outputs targeting excitatory neurons (n=292,673, see Materials and Methods). Note that no chandelier cells were in these reconstructions. To quantify the relationship between synaptic targeting and axonal morphology, for each cell we identified regions of axon in which synaptic outputs target the same postsynaptic cell, which we called contiguous targeting domains (Figure 7G). Each contiguous targeting domain was characterized by its number of synapses (Figure 7H), which ranged from 1–17 synapses. Consistent with our analysis of axon subgroup sequentiality, we found that BCs had significantly larger targeting domains compared to dendrite targeting cells ($p<10^{-300}$, Mann-Whitney U test), owing to a long tail of large components (Figure 7I). Taken together, our analysis of synaptic subgroup sequentiality aligns well with observed connectivity properties observed in the intact cortex.

Discussion:

In this study, we developed an image-based workflow to assign interneuron synaptic targeting at scale and with high concordance with human user scoring. As expected, this approach detected significant differences in targeting across interneuron Cre driver lines both in cortical sections as well as in dissociated culture. This afforded us the opportunity to directly compare target fidelity between cortical interneurons in the intact brain and potential drop-off in fidelity for interneurons in dissociated culture. Overall, we found that interneuron subclasses in culture showed similar target preference to interneurons in culture: Sst interneurons mainly targeted dendrite, BCs were more likely to target soma and ChCs targeted AIS more than the other subclasses. However, their relative synaptic target fidelity compared to their counterparts in cortical sections varied by subclass.

Sst interneurons in culture displayed remarkably similar targeting to Sst interneurons in cortical sections. The only exception was layer 1, which is almost exclusively dendrite-targeting. This was to be expected since cortical layering and polarity are lost in dissociated culture, both of which are required for layer 1 to form. BC interneurons, on the other hand, were not as high fidelity and had reduced soma targeting. One possibility is that robust soma innervation may require additional cues such as cortical structure and/or extrinsic signaling provided by the intact cortex. Another possibility is that BCs fail to attain full maturity in culture. Consistent with this notion, EM connectomic analysis of basket cell synapses through early postnatal time points found

a successive refinement synapses onto soma over time [33]. Thus, BCs in culture might represent an arrested state of maturity, which accounts for the decreased soma targeting. This is also consistent with our findings on Sst neurons, since the same study showed early and persistent dendrite targeting from Sst cells, potentially accounting for the higher target fidelity of Sst interneurons in culture. In addition to culture maturity, parvalbumin expression is not robust in vitro (which is why we did not use Pv^{Cre}; Ai34 for our culture experiments). Parvalbumin expression is sensitive to neuronal activity levels [34] and it is therefore also possible that activity levels in culture may not be insufficient or inappropriate for Pv lineage cells to achieve maturation. Taken together, it suggests that refinements in culture conditions could help to bring in vitro and in vivo synapse targeting into closer alignment.

Similarly, ChCs (the other Pv lineage subclass) also lagged behind their in vivo counterparts in target fidelity. We observed diminished AIS targeting and in fact targeting of ChCs to non-AIS targets. ChCs are thought to have near-absolute target preference for AIS [13,14,35], which may suggest that ChCs, like BCs, also fail to fully mature in culture. In addition, pyramidal cell polarity in the intact cortex allows for organized presentation of AIS targets, a configuration that could be required for high fidelity targeting of AIS. Nevertheless, we observe clear chandelier cell-like cartridges in our dissociated culture, which we only ever observed in Nkx2-1^{CreERT2} ChC-rich cultures. It is, to the best of our knowledge, the first described instance of ChC cartridge

formation in vitro and suggests that AIS targeting is at least partially under intrinsic regulation.

To establish our image-based system, it was necessary to clearly define the target compartments as well as presynaptic puncta for the model. To do so, further information was provided to the model via ilastik probability channels, which had the effect of adding complexity to the features measured. Thus, in the process of determining synaptic targeting, the model concurrently generated a multidimensional dataset with hundreds of metrics assigned to each synaptic punctum. Given the parallels between this dataset and that of scRNA-seq, we re-purposed approaches used for analyzing single cell transciptomics to help us in identifying underlying synaptic heterogeneity beyond the major targeting classes. By analyzing the spatial arrangement of these newly-identified synaptic subgroups, we found that subgroups are spatially arranged in a manner that reflects the heterogeneity of the postsynaptic environment (Summarized in Supplemental Figure 19).

With dendrite-targeting synaptic subgroups, we identified 3 pairs of subgroups that were organized in a proximal to distal arrangement along the dendrite. By performing a similar analysis in cortical sections, we found dendrite-targeting subgroups that specifically mapped onto layer 1 or were enriched in superficial and deep layers, consistent with the notion that a proximal-to-distal arrangement is also present in the intact cortex. That cortical interneurons target specific compartments of their

postsynaptic partner is a well-studied organizing principle of interneuron connectivity. Our data, however, reveals further granularity in specificity whereby distinct synaptic subgroups segregate onto subdomains of the dendrite.

A wealth of literature describes the compartmentalized nature of electrical signals in dendrite subdomains [36-38]. It is thought that these local electrical compartments integrate their signals to encode for complex information [39,40]. The role of inhibitory input onto these compartments can regulate or even override signals [41] and our data suggest that the nature of those inhibitory inputs is distinct across dendritic subcompartments, prompting one to speculate that they might also display different functional properties. Indeed, inhibitory synapses targeting distal dendrite exert a stronger than anticipated inhibitory effect on dendritic activity that is distinct from the effects of proximal dendrite-targeting synapses [42].

We also identified further heterogeneity within the soma-targeting group, which segregated into two 2 soma-targeting subgroups. Unlike dendrite-targeting subgroups, we did not find that their spatial distribution was onto subdomains of the soma and proximal dendrite. Rather, we found that the distribution of soma-targeting subgroups mapped onto 3 postsynaptic populations based on the relative abundance of each subgroup. We also noted that the 3 postsynaptic populations exhibited significant differences in size, Kv2 intensity, and the number of synapses/soma.

Our interpretation of the spatial distribution of both soma- and dendrite-targeting subgroups is that the postsynaptic neuron plays an instructive role as to the nature of the inhibitory synapse formed onto it - proximal to distal compartments of the dendrite dictate which subgroup of synapse can form while the somata of different neuron populations exhibit a preference for relative proportions of soma-targeting subgroups innervating it. Additionally, we find that Sst, BC and ChC subclasses can form the same subgroups of synapses, albeit at different proportions. Taken together, our findings suggest that interneurons are generalists in the types of synapses they can form, dictated by which compartment or cell soma type they encounter. This is not to say that interneurons play a passive role. To be sure, interneuron subclass (and subtype) identity drives strong target preferences for the subcompartments or the choice of postsynaptic partner [43,44]. Only once it does form a synapse, our data indicates that the nature of that synapse is determined by the target. Notably, a rich complement of proteins is found in the inhibitory postsynaptic compartment [45], and therefore the postsynapse well-positioned to establish a wide selection of synapse types.

Finally, we found that BCs and Sst interneurons enact inhibitory coverage using distinct strategies. We noted that synapses along BC axons were more likely to follow the topology of the postsynaptic neuron (greater sequentiality). We hypothesized that BCs formed contiguous synapses along the processes of their postsynaptic partners. In contrast, Sst synapses were significantly less sequential along their axons, which we interpreted as Sst axons being more likely to cut across multiple targets forming a

limited number of synapses onto each target cell. To test this, we analyzed interneuron synaptic connections from a connectome database derived from serial EM reconstructions, comparing dendrite-targeting versus basket cells. This analysis confirmed our interpretation on axon sequentiality as soma-targeting cells formed significantly more contiguous synapses onto the same postsynaptic neuron than dendrite-targeting cells.

Glutamate photo-uncaging experiments have revealed that both BCs [46] and Sst interneurons [47] promiscuously form connections and that the likelihood of those connections is directly correlated with distance from target. Indeed, a follow up study indicates that the axon project patterns of BCs and Ssts are highly predictive of connectivity, thus obeying Peter's rule [48]. Our findings lend additional anatomical insight. Here, we show that individual BC axons form dense and contiguous connections along the same postsynaptic cell, enabling strong and responsive inhibition [49], whereas Sst axons form fewer contiguous synapses, enabling them to establish a diffuse, lattice-like coverage with many pyramidal neurons [50,51]. These different strategies for enacting inhibitory coverage suggest to us that there may be subclass-intrinsic differences in the developmental programs regulating axon pathfinding and synaptogenesis, which may warrant further investigation.

The key innovation of our approach is the use of multidimensional imaging data to characterize individual synapses, and to assess large numbers of synapses at a

population level at the resolution of cellular subcompartments. Coupled with posthoc analysis of the spatial distribution of synaptic subgroups, it represents a novel and powerful method to study synapses going forwards, one that can be modified to suit experimental needs. In this study, we used the spatial and intensity relationships between interneuron synapses and their target compartments as the basis of our analysis, but other strategies could also be employed. For instance, presynaptic puncta could by immunolabeled for synaptic proteins to confer a greater degree of molecular specificity while using the same image-based analysis system we developed. Another strategy would be to genetically label postsynaptic cells to assess interneuron cell type specific targeting. Indeed, interneuron subtypes have been shown to exhibit cell-type specificity and distinct patterns of dendritic innervation onto subpopulations of pyramidal neurons [43,44] and interneurons [52,53]. Another approach would be to incorporate live imaging to reveal the dynamics of synaptogenesis or synaptic plasticity among subgroups. Here, fluor-conjugated gephyrin [54] accompanied by live labeling of other target compartments would be required.

There are some caveats that should be considered with our findings. First, our analysis was performed at a population level, which has the downside of blurring individual differences. A good example of this is our use of an optimal sequence of synaptic subgroups, which is very likely a simplified view. This is because the optimal sequence was derived from an overall analysis subgroup spatial distribution, resulting in a ‘best fit’

sequence. In actuality, sequentiality is almost certainly more nuanced, perhaps with parallel sequence types occurring depending on postsynaptic cell type or even across dendritic branches. In our current analysis, we are blind to these types of distinctions and it will require further work in future to carefully resolve. Second, it is possible that interneuron synaptic diversity manifests along multiple dimensions. Our analysis only speaks to the diversity in spatial and intensity relationships between pre- and post-synaptic compartments and there might be parallel or even orthogonal forms of synaptic diversity that we are not detecting in our analysis. By changing the starting parameters and/or inputs of the analysis, different forms of diversity might be uncovered.

It is notable the extent to which synaptic target fidelity and synaptic features are retained in dissociated culture. For one, we found significant target preferences across Sst, BCs and ChCs in culture. But perhaps more remarkably, analysis of synaptic subgroups in culture and their spatial distribution was predictive and helped inform our follow up analysis in cortical sections (Figure 5E and F; Figure 7). The parallels between neurons in culture and in the intact brain argue in favor of a prominent role for intrinsically determined cellular programs in the establishment of subdomains in dendrites and interneuron synaptic connections. It also lends credence to the use of in vitro approaches as a simple and tractable model going forwards to address some of the outstanding questions pertaining to interneuron synaptic diversity.

One of the key questions prompted by our findings is whether there is a molecular delineation to the synaptic subgroups we identified. Certainly, we hypothesize that differential spatial and intensity metrics across synaptic subgroups might serve as a proxy for differences in molecular composition. Indeed recent work finds high molecular diversity in inhibitory synapses [55]. Here, an in vitro model (in conjunction with analysis in the cortex) could prove to be a useful starting point for interrogating molecular composition and regulation of synaptic subgroup identity. Similarly, a culture-based model might indirectly serve as a platform for studying the process of dendritic subcompartmentalization since interneuron dendrite-targeting groups spatially segregate. Finally, there is growing interest in employing culture-based disease models for elucidating disease etiology and for the development of novel treatments. A multidimensional representation of synapses in culture would provide a highly comprehensive read-out for the healthy and diseased state, one that would be useful for hypothesis generation and could be adapted to high throughput drug screening.

In recent years, a number of groups have generated serial EM connectomic databases of mouse neocortex, for example [32,56,57,58]. Object registration of features belonging to every pre- and post-synaptic neuron and their synaptic connections has offered an unprecedented glimpse into the organization of cortical synaptic connectivity. Our image-based, multidimensional approach serves as a complementary and parallel strategy for studying synapses. Its benefits lie in its broad accessibility (requiring only basic confocal microscopy) and its versatility (compatibility with mouse genetics and

immunohistochemistry). These features make it a flexible vehicle for hypothesis testing through mouse genetics (as we did with the *Reeler* mutant), molecular perturbation, or the interrogation of other experimental variables. Findings from our approach can then be cross-referenced with the ground truth provided by EM connectomics. Indeed, this was the strategy we employed to confirm differences in inhibitory coverage between Sst and BCs.

Our image-based approach also has parallels to previous work on excitatory synaptic diversity [59]. By analyzing the presence and intensity relationships between 2 postsynaptic fluorescently-labeled synaptic proteins, the authors found synaptic heterogeneity in excitatory neurons throughout the brain. In contrast, our work focuses on inhibitory synapses in the cortex, investigates their association with cellular compartments, and incorporates both pre- and post-synaptic features in our analysis. Zhu and colleagues found that excitatory synapse types were spatially enriched across brain structures or within layers of the cortex and hippocampus [59]. In this study, we find a different organizing principle. With interneuron synapses, we observe synaptic subgroup distribution locally in subcellular compartments of the postsynaptic cell and across postsynaptic populations. This contrast makes biological sense given the differences between excitatory and inhibitory synaptic connectivity. Cortical interneurons project locally through repetitive synaptic motifs, so it stands to reason that their complement of synapse types would be patterned locally and reflect local postsynaptic topography (subcompartments of dendrite and different species of postsynaptic soma).

Excitatory neurons on the other hand are largely long-range neurons whose afferent synapses project to distal points; Here, synaptic diversity reflecting different functional areas of the brain might be expected. In future, employing a uniform image-based analysis of both inhibitory and excitatory systems could prove highly informative, allowing us to study the spatial distribution of excitatory and inhibitory synaptic subgroups simultaneously and infer how the two systems functionally interact.

Figure Legends:

Figure 1. ***Labeling target compartments and presynaptic puncta for image-based analysis system.*** (A) Immmunohistochemical labeling of postsynaptic target compartments in (top row) cortical sections and (bottom row) dissociated cortical culture. Target compartments labeled: soma and proximal dendrite ($Kv_{2.1}$ and $Kv_{2.2}$ combined, Kv2), axon initial segment (Ankyrin G), dendrite (gephyrin). (B, C) Validation of genetic labeling of presynaptic synaptophysin-tdTomato+ puncta. (B) Pan-MGE lineage interneuron Cre driver, Nkx2-1^{Cre}; Ai34 presynaptic puncta counter labeled with inhibitory synaptic marker (gephyrin, blue) and excitatory synaptic marker (PSD95, green). (C) Pan-pyramidal cell Cre driver, Nex^{Cre}; Ai34 presynaptic puncta counterlabeled with gephyrin (blue) and PSD95 (green). Yellow arrowheads indicate overlap/close apposition between syn-tdT puncta and markers. (D) Schematic showing canonical target preference for 3 interneuron subclasses: Sst interneuron, basket cell (BC) and chandelier cell (ChC). Right panel shows targeting in cortex, left panel shows target recapitulation in dissociated culture. (E) Dissociated cultures in which Sst, BC and ChCs were genetically-labeled with Ai34 reporter. Counterlabeled with Kv2 (blue) and Ankyrin G (green).

Figure 2. ***Overview of image-based model for assigning synaptic target to individual interneuron synaptic puncta.*** (A-D) Automated assignment of pre- (blue) and post- (pink) synaptic compartments to 11 probability channels. (A) Raw 4-channel confocal image for Ai34-label presynaptic puncta (white), Kv2 (red), Ankyrin G (green) and gephyrin (blue). (B) Raw channels split into pre- (top, blue) and post- (bottom, pink) synaptic compartments. (C) Raw pre- and post-synaptic channels split into probability channels. Presynaptic probability channels used to perform (D) object segmentation of syn-tdT presynaptic puncta. (E) Overlay of probability channels showing example images that were used for object training to define dendrite-, soma- and AIS-targeting synapses. (F) Target probability score assigned to each synapse, color coded according to heatmap (right) for each target compartment.

Figure 3. ***Quantification and comparison of interneuron targeting in cortical sections and in dissociated culture.*** (A) Example cortical columns showing target assignment of Ai34-labeled synaptic puncta for Sst^{Cre}, Nkx2-1^{Cre}, Pv^{Cre} and Nkx2-1^{CreERT2} BC- and ChC-rich regions. (B) Stacked plots depicting interneuron synaptic target preferences for the 4 Cre driver lines in (A). (C) Example images of dissociated cultures in which Sst, BC and ChC are labeled by Ai34 (white) and counterlabeled with target compartments – Kv2 (red), Ankyrin G (green), gephyrin (blue). (D) Stacked plots depicting interneuron target preferences for dissociated cultures using various Cre driver lines (Nkx2-1 for entire MGE lineage, Sst^{Cre} for Sst interneurons, Nkx2-1^{CreERT2} bulk for primarily Pv lineages, BC cultures from Nkx2-1^{CreERT2} BC-rich region, ChC cultures from Nkx2-1^{CreERT2} ChC-rich region. (E) Direct comparison of stacked plots of target preference of cortical sections versus dissociated culture counterpart. (F) Interneuron targeting of soma (red) and dendrite (blue) as a function of cortical depth in

cortical columns of Nkx2-1^{Cre}, Pv^{Cre}, Sst^{Cre}, and Sst^{Cre}; *Reeler* mice. Grey boxes denote approximate depths of cortical layers. (A) Nkx2-1^{Cre} n=6, 2 cortical columns/animal; Sst^{Cre} n=7, 2 cortical columns/animal; Pv^{Cre} n=8, 2 cortical columns/animal; Nkx2-1^{CreERT2} BC- and ChC-rich n=3, 2 cortical regions/animal. (B) Nkx2-1^{Cre} n=4 cultures, 20-40 tiled regions/culture; Sst^{Cre} n=4 cultures, 20-40 tiled regions/culture; Nkx2-1^{CreERT2} bulk culture n=3, 20-35 tiled regions/culture; Nkx2-1^{CreERT2} BC-rich n=4; Nkx2-1^{CreERT2} ChC-rich cultures n=4. (A, B) 2-way ANOVA, post hoc Tukey's test. * $p<0.05$, ** $p<0.01$, ***$p<0.001$, ****$p<0.0001$. (F) Nkx2-1^{Cre} n=6; Pv^{Cre} n=8; Sst^{Cre} n=7; Sst^{Cre}; Rl/Rl n=5.

Figure 4. *Unsupervised analysis of Ilastik-derived synaptic features reveals the presence of synaptic subgroups.* (A) UMAP visualization of synapses from the in vitro culture data set, generated from autoencoder-based intermediate layer nodes as the reduced dimensional space. Points are colored by Ilastik-derived (supervised learning) target calls. Co-segregation of soma-, dendrite- and AIS-targeting synapses is apparent in this visual representation. (B) Same plot as in A, but with points colored by cluster identity following Louvain community detection to identify clusters, showing subgroups enriched in dendrite- and soma-targeting synapses. (C) Violin plots showing the distribution of ilastik-derived feature values, with features selected based on maximal discriminability between each cluster and all the others shown in (B). Features are ordered to highlight that clusters are distinguished by combinations of features with continuous values rather than sharply-defined breaks. Feature Legend: A) mean intensity of gephyrin probability; B) variance of intensity in gephyrin probability; C) Covariance of channel intensity of gephyrin probability; D) variance of intensity in neurite probability; E) Variance of intensity in bouton probability; F) covariance of channel intensity soma probability versus gephyrin probability; G) covariance of channel intensity gephyrin probability versus non-AIS Ankyrin G probability; H) minimum intensity in non-AIS Ankyrin G probability; I) mean intensity in postsynaptic target background; J) kurtosis of intensity in presynaptic background; K) minimum intensity in bouton probability; L) variance of intensity in Ai34 raw channel; M) covariance of channel intensity in Kv2 versus presynaptic background probability; N) variance of intensity in soma probability; O) mean intensity in neurite probability; P) total intensity in RFP soma probability; Q) covariance of channel intensity in Kv2 raw versus RFP soma probability; R) variance of intensity in AIS probability; S) covariance of channel intensity in Ankryin G raw versus AIS probability; T) covariance of channel intensity in AIS probability.

Figure 5. *Analysis of synaptic subgroup spatial compactness reveals sequential spatial relationship between dendrite-targeting subgroups along proximal-to-distal axis of target dendrite.* (A, B) Modified Louvain analysis of x, y, z coordinates of synapses belong to each cluster (0-8) divided by random distribution and plotted as a enrichment matrix. (A) Enrichment matrix showing spatial relationship when clusters are ordered numerically from 0 to 8. (B) Re-ordered enrichment matrix showing spatial relationship when clusters re-order to maximize off-diagonal sum values. (C) Re-ordered enrichment matrix with cell compartments (AIS, soma, proximal, mid and distal

dendrite) denoted on plot. (D) Soma- and dendrite-targeting subgroups re-colored in red-orange-white heatmap. Boxed region rendered sequentially from proximal to distal dendrite (top to bottom). (E) Synaptic subgroups rendered in original colors and red-orange-white heatmap onto Sst^{Cre} dissociated culture. Spatially-segregated dendrite-targeting subgroups indicated on UMAP of synaptic clusters. (F) Synaptic subgroups rendered onto Sst^{Cre} cortical column. Spatially-segregated dendrite-targeting subgroups indicated on UMAP of synaptic clusters. Analysis of ~1.45 million synapses obtained from Nkx2-1^{Cre} n=4 and Sst^{Cre} n=3 cultures.

Figure 6. ***Soma-targeting synaptic subgroups label distinct subpopulations of postsynaptic neurons.*** (A) Soma targeting synaptic subgroups 4 (pink) and 5 (red) rendered onto original image without (left) and with (right) Kv2 labeling of soma and proximal dendrite (white). Boxed regions show postsynaptic cells that predominantly receive subgroup 4 synapses (type 4 cells), receive a mixture of subgroups 4 and 5 synapses (type 4/5 cells) and predominantly subgroup 5 synapses (type 5 cells). (B) Frequency histogram of proportions of synaptic subgroup 4 and 5 innervation showing the delineation or type 4, type 4/5 and type 5 postsynaptic cells. (C) Proportion of type 4, 4/5 and 5 cells found in Nkx2-1^{Cre} and Sst^{Cre} cultures. (D-G) Features of type 4, 4/5 and 5 postsynaptic cells. (D) Number of synapses received/soma for each type of postsynaptic cell. (E) Number of synapses received/soma separated by Cre driver. (F) Size of soma and proximal dendrite for each type of postsynaptic cell. (G) Intensity of Kv2 signal for each type of postsynaptic cell. Based on the analysis of 977 somata (Nkx2-1^{Cre} 600 somata from 4 cultures; Sst^{Cre} 377 somata from 3 cultures). One-way ANOVA, posthoc Tukey's test. * $p<0.05$, ** $p<0.01$, ***$p<0.001$, ****$p<0.0001$.

Figure 7. ***Analysis of Basket cell and somatostatin cell synapses along their axons reveal different strategies for enacting inhibitory coverage.*** Individually-labeled (B) Sst interneuron and (B) BC with synapse subgroups rendered in red-orange-white heatmap reflecting optimized synaptic subgroup sequence. (C) Calculating sequentiality score from successive synapses along axonal branch by their deviation from optimized sequence (See Supplemental Figure 17 for more details). (D) Frequency (left) distribution of the average sequentiality scores of individual axonal branches from Sst and BCs; 2-way ANOVA $p<0.0001$. (E) Sequentiality score for all (left; unpaired t-test p,0.0001) and individual cells (right) for BCs (red) and Sst (blue); 2-way ANOVA $p<0.0001$. (F) Sequentiality score for all (left; unpaired t-test $p=0.0016$) and individual cells (right) for BCs (red) and Sst (blue) where only contiguous dendrite-targeting subgroups were calculated; 2-way ANOVA $p=0.0013$. (D-F) Based on the analysis of 10 BCs and 12 Sst individually-labeled neurons with 5-9 axon branches analyzed per cell obtained from 3 BC cultures and 2 Sst cultures. (G) Schematic depicting interneuron (IN) targeting 3 postsynaptic neurons. Dashed lines indicate regions where synapses are formed successively onto the same postsynaptic neuron, which we term contiguous synaptic components. Image on right shows 3D reconstruction of IN (blue) forming contiguous synaptic components (pink) onto pyramidal neuron (gray). (H) Representative wire diagram depictions of axonal trees of dendrite-targeting cell, left

(putative Sst) and basket cell (putative Pv). Overlaid with single synapses (gray circles) and contiguous synaptic components ranging from 2-16 synapses (blue or red circles). (I) Frequency distribution of synaptic component sizes for dendrite-targeting (putative Sst, blue) and basket cells (putative Pv, red); Mann-Whitney U test ($p<10^{-300}$).

Supplemental Figure 1. ***MAP2 is not an ideal postsynaptic marker for dendrite-targeting synapses.*** (A) Schematic showing labeled compartments in panel B. Synaptophysin-tdT (red), drebrin labeling dendritic spine heads and necks (green), MAP2 labeling spine shafts (blue) and gephyrin (white). (B) Syn-tdT synapses colabeled with MAP2. Arrowheads indicate syn-tdT puncta that are not apposed to MAP2+ shaft. (B') Syn-tdT, MAP2 and gephyrin shows arrowhead-indicated puncta are apposed to inhibitory postsynapse. (B'') Syn-tdT, MAP2, drebrin shows arrowhead-indicated syn-tdT puncta are apposed to spin heads and necks labeled by drebrin and are therefore dendrite-targeting synapses. (B''') Addition of gephyrin to B'' shows that drebrin-apposed syn-tdT puncta are also apposed to gephyrin. (C) MAP2 signal is uneven in intensity (area of low intensity circumscribed by dotted line) in areas with high densities of syn-tdT puncta. (C'') Syn-tdT signal is dense in area of low MAP2. (C''') Just MAP2 signal showing uneven signal intensity.

Supplemental Figure 2. ***Cre driver lines allow for subclass-specific genetic labeling with Ai34 synaptophysin-tdTomato conditional reporter.*** (A-C) 3 Cre driver lines (A) Nkx2-1^{Cre}, (B) Sst^{Cre}, and (C) Pv^{Cre} cortical sections. Left panels show synaptophysin-tdTomato (Ai34) puncta alone and counterlabeled with Kv2 (red), Ankyrin G (green), gephyrin (blue). Right panels show same configuration with layer 1 outlined in dashed white lines. Cortical sections of Nkx2-1^{CreERT2}; Ai34 tamoxifen gavage at e17.5 showing synaptophysin-tdTomato puncta alone and counterlabeled with Ankyrin G (green) or Kv2 (red) in regions enriched for (D) chandelier cells and (E) basket cells.

Supplemental Figure 3. ***Late embryonic tamoxifen gavage of Nkx^{CreERT2}; Ai34 mice labels Pv lineage interneuron neurons with spatial segregation between chandelier cells and basket cells.*** (A) Coronal cortical sections from Nkx2-1^{CreERT2}; Ai9 e17.5 tamoxifen mice (n=8) immunolabeled for Pv and Sst. Coronal sections registered to Allen CCF and rendered as 3D model. Left panels shows all Pv (magenta) and Sst (green) cells identified. Right panel shows only tdTomato+ cells. (B) Quantification of tdTomato+ Pv and Sst cells for all sections analyzed. (C) (left) Sample field of tdtomato+ Ai9 signal in coronal section before and after (middle) ilastik processing to generate probability map for synaptic puncta and ImageJ macros to (right) distinguish between chandelier cell cartridge puncta (green) from remaining puncta (magenta). (D) Heatmap composite from 8 mice of regions enriched for chandelier cartridge puncta (green) and BC puncta (magenta). Cortex on the right shows approximate regions enriched for ChCs (green) and BCs (magenta), which informed our dissections for dissociated cultures and imaged regions for target assessment in subsequent experiments.

Supplemental Figure 4. ***Immunohistochemistry for pre- and post-synaptic features is an imperfect representation of features needed for image-based model analysis.*** (A) 4 channel confocal image of dissociated Sst^{Cre}; Ai34 cortical culture – synaptophysin tdTomato (white), Kv2 (red), Ankyrin G (green), gephyrin (blue). Yellow arrowhead in syn-tdT panel indicates tdTomato signal in soma Sst interneuron, with corresponding Kv2 signal (also yellow arrowhead). Kv2 signal is punctate and does not completely fill in soma and proximal dendrite. Ankyrin G signal is varied: high levels in AIS, lower levels in neurite processes. (B) Schematic showing ideal target compartments vs. immunolabeled compartments – more details in Supplemental Figure 5.

Supplemental Figure 5. ***Immunolabeled target compartments for soma/proximal dendrite and axon initial segment are defined by context and multiple channels.*** (A, B) Immunolabeled dissociated culture neurons - Kv2 (red), AnkyrinG (green), gephyrin (blue) and syn-tdT puncta (white). (A) General immunolabeling pattern for soma and proximal dendrite target compartment. Note punctate Kv2 signal (red), and presence of gephyrin puncta that define target compartment and syn-TdT puncta innervation of soma and proximal dendrite (white). (B) General staining pattern for axon initial segment, merged image of Kv2, Ankyrin G, gephyrin and syn-tdT puncta. (C) 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. (D) Kv2 signal at AIS. Note that Kv2 localizes to the AIS, where it is found in apposition to AIS-targeting syn-tdT puncta (white). (E) Lack of gephyrin immunolabel at AIS. Gephyrin clustering is a canonical marker of GABAergic post-synaptic densities (PSDs). The gephyrin antibody we use in this study (mAb7a from Synaptic Systems) 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 syn-tdT puncta.

Supplemental Figure 6. ***Target assignment for segmented syn-tdT puncta objects.*** (A) Target assignment (object classification) draws from inputs consisting of the raw and probability mapped target channels (collectively called the measurement basis image) and the object mask (segmented syn-tdT puncta). (B) Target assignment (Object probability score) for the five object classes: non-bouton, dendrite-, soma-, AIS-, and unknown-targeting. Object training of syn-tdT puncta objects. (C) user-trained ground truth data points. (D) (left) Heatmap depicting target probability for all 5 targets, (middle) assigned target based on probability score, (right) finalized assigned targets (non-bouton objects removed).

Supplemental Figure 7. ***Representative performance of image-based model for synaptic target assignment in dissociated cortical cultures and in cortical sections.*** Confocal tiles of (A) dissociated cortical culture and (B) cortical sections processed through various stages of image-based model for assigning synaptic targets. Left, raw confocal channels converted into 7 postsynaptic (top) and 4 presynaptic

(bottom) probability channels. Presynaptic channels used for object segmentation of syn-tdT+ puncta. Right, syn-tdT objects assigned targets (red=soma/proximal dendrite; green=AIS; blue=dendrite).

Supplemental Figure 8. ***Comparison of model performance versus user assignment of synaptic targets shows strong concordance.*** (A) Example field of randomly-selected syn-tdT puncta used for comparison between human user and model following target assignment (red=soma/proximal dendrite; green=AIS; blue=dendrite). Orange puncta circled in yellow indicate mis-matches between model and user target assignment. (B) Overall concordance (accuracy) of model vs. human target assignment for each of the target categories (dendrite, soma, AIS, unknown and non-bouton). Boutons is accuracy excluding non-bouton assignment. (C) Confusion matrix shows how user and model target calls differed when there was disagreement. False discovery rate read in rows and precision along the diagonal.

Supplemental Figure 9. ***Reeler mutant cortex exhibits inverted pyramidal cell layering, defects in polarity, and randomization of synaptic compartmental target across cortical layers.*** Coronal cortical sections of P28 Sst^{Cre} mice that are (A) wildtype and (A') *Reeler* homozygous mutants immunolabled for layer 5-enriched pyramidal cell marker, Ctip2 (green) shows disordered and largely inverted cortex. Kv2 immunolabeling of (B) wildtype and (B') *Reeler* homozygous mutants shows disruptions in pyramidal cell polarity in mutants. Target assignment of syn-tdT puncta in superficial layer of (C) wildtype and (C') *Reeler* homozygous mutant cortex. White dashed lines denote layer 1 in wildtype and cortical surface in *Reeler* homozygous mutant.

Supplemental Figure 10. ***Empirical determination of K and resolution results in the identification of 9 synaptic subgroups.*** (A) Computed Calinski-Harabasz (CH) score as a function of varying k neighbor value and resolution. (B) Computed Davies-Bouldin (DB) score as a function of varying k neighbor value and resolution. (A, B) To determine optimal k and resolution, we determined where the local maximum for DB coincided with the local minimum for CH – K=10, resolution = 0.4. (C) Computed of number clusters resulting from vary k neighbor values as a function of resolution. (D) Computed number of clusters where the synaptic puncta within the cluster constituted a minimum of 0.5% of the total number of puncta (termed 'retained clusters, a cut-off for noise). (E) Frequency stacked plot showing the relative abundance of each cluster and their target assignment composition. Note that past cluster 8 cluster abundance is negligible. (F) Target assignment score stacked plot set to 100%.

Supplemental Figure 11. ***Top-20 highest variance ilastik-derived synaptic features across all 9 identified subgroups.***

Supplemental Figure 12. ***Spatial correlation (spatial compactness) enrichment matrix re-ordering process based on maximal sum of off-diagonal values.*** (A) Modified Louvain analysis (10 nearest neighbors divided by nearest neighbors assuming random distribution) plotted as an enrichment matrix. Values greater than 0 represent spatial correlation above random distribution and values less than 0 represent

spatial anti-correlation vs. random. 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 column i, row j. That is, the diagonal values (dark green box) represent the spatial compactness within a cluster whereas the off-diagonal values (light green box) represent the spatial compactness across adjacent clusters. Additionally, cluster compactness across non-adjacent clusters can be found cross-referencing column and row, which shows that some clusters are spatially-correlated (outlined in red box) or spatially anti-correlated (outlined in blue box). We hypothesized that there is an optimal spatial sequence across all the clusters and therefore empirically determined the (B) optimal cluster sequence by calculating the sum of the off-diagonal values for all cluster permutations. In (A), the off-diagonal sum for clusters arranged in numerical order is -12.26. (B) In the optimized cluster sequence, the off-diagonal sum is 1.23.

Supplemental Figure 13. ***Assessing the robustness and reproducibility of unsupervised analysis of synaptic feature dataset.*** (A) Stacked bar plot composition of autoencoder+Louvain derived cluster composition for each PCA+K-means derived cluster. Note the general enrichment of soma, proximal, mid and distal dendrite autoencoder+Louvain clusters distributed onto PCA+K-means clusters with similar spatial segregation (B) Enrichment matrix of spatial compactness scores arranged according to empirically-determined optimized sequence for PCA+K-means clusters (top) and autoencoder+Louvain clusters (bottom). (C) Representative Nkx2-1^{Cre}; Ai34 culture overlaid with PCA+K-means clusters (top) and autoencoder+Louvain clusters (bottom) colored as a red-orange-white heatmap according to optimized sequence in (B).

Supplemental Figure 14. ***Spatial arrangement of soma-targeting synaptic subgroups.*** (A) Schematic of 2 competing models for how soma-targeting subgroups might be spatially arranged on target soma. The model on the right depicts subgroup 4 and 5 synapses arranged on somata in a graded fashion with enrichment across different subcompartments of the soma. The model on the left depicts differential enrichment of subgroup 4 and 5 across different populations of postsynaptic neurons. (B) Subgroups 4 (pink) and 5 (dark pink) mapped back onto original image without Kv2 signal. Arrangement of subgroups 4 and 5 is clumped, supporting the model on the left in (A). (C) ilastik-derived metric showing spatial correlation between each synaptic subgroup and the probability map for AIS. No significant difference (1-way ANOVA) between subgroups 4 and 5 suggest that mapping of 4 and 5 onto individual soma is not spatially enriched across soma compartments.

Supplemental Figure 15. ***Individually-labeled basket cells form synapses onto types 4, 4/5 and 5 somata.*** (A) Individually-labeled BC cultured from BC-rich region of

Nkx2-1^{CreERT2}; Ai34 mice – Syn-tdT puncta (white), Kv2 (red). (B) All synapse subgroups mapped onto the same individually-labeled BC. Just subgroups 4 and 5 mapped on (C) without and (D) with counterlabel Kv2 (white). (D') Enlarged region of box in (D) shows that the same BC forms synapses onto type 4, type 4/5 and type 5 somata.

Supplemental Figure 16. ***Features of types 4, 4/5 and 5 postsynaptic cells are distinct.*** A sizeable fraction of type 4 and type 5 postsynaptic cells contain exclusively subgroup 4 and 5 synapses, respectively. These cells are named type 4 only and type 5 only. (A-C) Analysis of features of type 4 only, mainly type 4, type 4/5, mainly type 5 and type 5 only cells for the following measurements: (A) # synapses/soma; (B) size of soma and proximal dendrite; (C) intensity of Kv2 signal. Based on the analysis of 977 somata (Nkx2-1^{Cre} 600 somata from 4 cultures; Sst^{Cre} 377 somata from 3 cultures). 1-way ANOVA with post hoc Tukey's test. * $p<0.05$, ** $p<0.01$, ***$p<0.001$, ****$p<0.0001$.

Supplemental Figure 17. ***Relative proportion of synaptic subgroups employed by Sst, BCs and ChCs.***

Supplemental Figure 18. ***Calculation of sequentiality score of axonal branches.*** (A) Optimal subgroup sequence (as determined previously in Figure 4, Supplemental Figure 13) is used as the basis for calculating sequentiality score. The subgroup identity of synapses is compared to each successive synapse to count the 'jump' from optimal sequence. (B) Formula for calculating average sequentiality score for an axonal branch ($\bar{S}_{\text{axonal branch}}$). (C) Example of axonal branch with synaptic subgroup identity mapped onto original image. Box on the right is enlarge image of boxed region on left. White arcs depict comparison of subgroup identity across successive synapses and the 'jump' in sequentiality score. Bottom right shows calculation of sequentiality score for stretch of axonal branch evaluated.

Supplemental Figure 19. ***Summary of findings related to the spatial distribution of synaptic subgroups.*** (left) Schematic of interneuron subclasses forming canonical synapses onto a postsynaptic neuron. The panels on the right show how the model has changed with each successive finding: (1) We found dendrite targeting sub-groups that were spacially localized to proximal, mid, and distal subdomains of the dendrite, revealing a more complex postsynaptic topology. (2) Soma-targeting subgroups were distributed in different proportions across 3 groups of postsynaptic neurons that exhibited distinct features. (3) BCs and Sst interneurons enact inhibitory coverage via different strategies. BCs form numerous synapses along the postsynaptic neuron while Sst interneurons form fewer synapses before 'skipping' to another process.

References:

1. Gouwens, N.W., Sorensen, S.A., Baftizadeh, F., Budzillo, A., Lee, B.R., Jarsky, T., Alfiler, L., Baker, K., Barkan, E., Berry, K., et al. (2020). Integrated Morphoelectric and Transcriptomic Classification of Cortical GABAergic Cells. Cell *183*, 935-953.e919. 10.1016/j.cell.2020.09.057.
2. Tasic, B., Yao, Z., Graybuck, L.T., Smith, K.A., Nguyen, T.N., Bertagnolli, D., Goldy, J., Garren, E., Economo, M.N., Viswanathan, S., et al. (2018). Shared and distinct transcriptomic cell types across neocortical areas. Nature *563*, 72-78. 10.1038/s41586-018-0654-5.
3. Yao, Z., van Velthoven, C.T.J., Nguyen, T.N., Goldy, J., Sedeno-Cortes, A.E., Baftizadeh, F., Bertagnolli, D., Casper, T., Chiang, M., Crichton, K., et al. (2021). A taxonomy of transcriptomic cell types across the isocortex and hippocampal formation. Cell *184*, 3222-3241.e3226. 10.1016/j.cell.2021.04.021.
4. Contractor, A., Ethell, I.M., and Portera-Cailliau, C. (2021). Cortical interneurons in autism. Nat Neurosci *24*, 1648-1659. 10.1038/s41593-021-00967-6.
5. Dienel, S.J., and Lewis, D.A. (2019). Alterations in cortical interneurons and cognitive function in schizophrenia. Neurobiol Dis *131*, 104208. 10.1016/j.nbd.2018.06.020.
6. Hashimoto, T., Arion, D., Unger, T., Maldonado-Avilés, J.G., Morris, H.M., Volk, D.W., Mirnics, K., and Lewis, D.A. (2008). Alterations in GABA-related transcriptome in the dorsolateral prefrontal cortex of subjects with schizophrenia. Mol Psychiatry *13*, 147-161. 10.1038/sj.mp.4002011.
7. Pantazopoulos, H., Woo, T.U., Lim, M.P., Lange, N., and Berretta, S. (2010). Extracellular matrix-glial abnormalities in the amygdala and entorhinal cortex of subjects diagnosed with schizophrenia. Arch Gen Psychiatry *67*, 155-166. 10.1001/archgenpsychiatry.2009.196.
8. Wang, A.Y., Lohmann, K.M., Yang, C.K., Zimmerman, E.I., Pantazopoulos, H., Herring, N., Berretta, S., Heckers, S., and Konradi, C. (2011). Bipolar disorder type 1 and schizophrenia are accompanied by decreased density of parvalbumin- and somatostatin-positive interneurons in the parahippocampal region. Acta Neuropathol *122*, 615-626. 10.1007/s00401-011-0881-4.
9. Gonchar, Y., Turney, S., Price, J.L., and Burkhalter, A. (2002). Axo-axonic synapses formed by somatostatin-expressing GABAergic neurons in rat and monkey visual cortex. J Comp Neurol *443*, 1-14. 10.1002/cne.1425.
10. Melchitzky, D.S., and Lewis, D.A. (2008). Dendritic-targeting GABA neurons in monkey prefrontal cortex: comparison of somatostatin- and calretinin-immunoreactive axon terminals. Synapse *62*, 456-465. 10.1002/syn.20514.
11. DeFelipe, J. (1997). Types of neurons, synaptic connections and chemical characteristics of cells immunoreactive for calbindin-D28K, parvalbumin and calretinin in the neocortex. J Chem Neuroanat *14*, 1-19. 10.1016/s0891-0618(97)10013-8.
12. Hendry, S.H., Jones, E.G., Emson, P.C., Lawson, D.E., Heizmann, C.W., and Streit, P. (1989). Two classes of cortical GABA neurons defined by differential

calcium binding protein immunoreactivities. Exp Brain Res *76*, 467-472. 10.1007/BF00247904.

13. DeFelipe, J., Hendry, S.H., and Jones, E.G. (1989). Visualization of chandelier cell axons by parvalbumin immunoreactivity in monkey cerebral cortex. Proc Natl Acad Sci U S A *86*, 2093-2097. 10.1073/pnas.86.6.2093.
14. Somogyi, P., Freund, T.F., Hodgson, A.J., Somogyi, J., Beroukas, D., and Chubb, I.W. (1985). Identified axo-axonic cells are immunoreactive for GABA in the hippocampus and visual cortex of the cat. Brain Res *332*, 143-149. 10.1016/0006-8993(85)90397-x.
15. Tremblay, R., Lee, S., and Rudy, B. (2016). GABAergic Interneurons in the Neocortex: From Cellular Properties to Circuits. Neuron *91*, 260-292. 10.1016/j.neuron.2016.06.033.
16. Bishop, H.I., Guan, D., Bocksteins, E., Parajuli, L.K., Murray, K.D., Cobb, M.M., Misonou, H., Zito, K., Foehring, R.C., and Trimmer, J.S. (2015). Distinct Cell- and Layer-Specific Expression Patterns and Independent Regulation of Kv2 Channel Subtypes in Cortical Pyramidal Neurons. J Neurosci *35*, 14922-14942. 10.1523/jneurosci.1897-15.2015.
17. Kordeli, E., Lambert, S., and Bennett, V. (1995). AnkyrinG. A new ankyrin gene with neural-specific isoforms localized at the axonal initial segment and node of Ranvier. J Biol Chem *270*, 2352-2359. 10.1074/jbc.270.5.2352.
18. Iascone, D.M., Li, Y., Sümbül, U., Doron, M., Chen, H., Andreu, V., Goudy, F., Blockus, H., Abbott, L.F., Segev, I., et al. (2020). Whole-Neuron Synaptic Mapping Reveals Spatially Precise Excitatory/Inhibitory Balance Limiting Dendritic and Somatic Spiking. Neuron *106*, 566-578.e568. 10.1016/j.neuron.2020.02.015.
19. Karimi, A., Odenthal, J., Drawitsch, F., Boergens, K.M., and Helmstaedter, M. (2020). Cell-type specific innervation of cortical pyramidal cells at their apical dendrites. eLife *9*, e46876. 10.7554/eLife.46876.
20. Madisen, L., Zwingman, T.A., Sunkin, S.M., Oh, S.W., Zariwala, H.A., Gu, H., Ng, L.L., Palmiter, R.D., Hawrylycz, M.J., Jones, A.R., et al. (2010). A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. Nat Neurosci *13*, 133-140. 10.1038/nn.2467.
21. Taniguchi, H., Lu, J., and Huang, Z.J. (2013). The spatial and temporal origin of chandelier cells in mouse neocortex. Science *339*, 70-74. 10.1126/science.1227622.
22. Miyoshi, G., Butt, S.J., Takebayashi, H., and Fishell, G. (2007). Physiologically distinct temporal cohorts of cortical interneurons arise from telencephalic Olig2-expressing precursors. J Neurosci *27*, 7786-7798. 10.1523/jneurosci.1807-07.2007.
23. Ango, F., di Cristo, G., Higashiyama, H., Bennett, V., Wu, P., and Huang, Z.J. (2004). Ankyrin-based subcellular gradient of neurofascin, an immunoglobulin family protein, directs GABAergic innervation at purkinje axon initial segment. Cell *119*, 257-272. 10.1016/j.cell.2004.10.004.

24. Favuzzi, E., Deogracias, R., Marques-Smith, A., Maeso, P., Jezequel, J., Exposito-Alonso, D., Balia, M., Kroon, T., Hinojosa, A.J., E, F.M., and Rico, B. (2019). Distinct molecular programs regulate synapse specificity in cortical inhibitory circuits. Science *363*, 413-417. 10.1126/science.aau8977.
25. Hayano, Y., Ishino, Y., Hyun, J.H., Orozco, C.G., Steinecke, A., Potts, E., Oisi, Y., Thomas, C.I., Guerrero-Given, D., Kim, E., et al. (2021). IgSF11 homophilic adhesion proteins promote layer-specific synaptic assembly of the cortical interneuron subtype. Sci Adv *7*. 10.1126/sciadv.abf1600.
26. Tai, Y., Gallo, N.B., Wang, M., Yu, J.R., and Van Aelst, L. (2019). Axo-axonic Innervation of Neocortical Pyramidal Neurons by GABAergic Chandelier Cells Requires AnkyrinG-Associated L1CAM. Neuron *102*, 358-372.e359. 10.1016/j.neuron.2019.02.009.
27. Tyagarajan, S.K., and Fritschy, J.-M. (2014). Gephyrin: a master regulator of neuronal function? Nature Reviews Neuroscience *15*, 141-156. 10.1038/nrn3670.
28. Berg, S., Kutra, D., Kroeger, T., Straehle, C.N., Kausler, B.X., Haubold, C., Schiegg, M., Ales, J., Beier, T., Rudy, M., et al. (2019). ilastik: interactive machine learning for (bio)image analysis. Nat Methods *16*, 1226-1232. 10.1038/s41592-019-0582-9.
29. Sommer, C., Straehle, C., Köthe, U., and Hamprecht, F.A. (2011). Ilastik: Interactive learning and segmentation toolkit. 30 March-2 April 2011. pp. 230-233.
30. Boyle, M.P., Bernard, A., Thompson, C.L., Ng, L., Boe, A., Mortrud, M., Hawrylycz, M.J., Jones, A.R., Hevner, R.F., and Lein, E.S. (2011). Cell-type-specific consequences of Reelin deficiency in the mouse neocortex, hippocampus, and amygdala. J Comp Neurol *519*, 2061-2089. 10.1002/cne.22655.
31. The, M.C., Bae, J.A., Mahaly, B., Caitlyn, A.B., Agnes, L.B., Derrick, B., JoAnn, B., Daniel, J.B., Manuel, A.C., Brendan, C., et al. (2023). Functional connectomics spanning multiple areas of mouse visual cortex. bioRxiv, 2021.2007.2028.454025. 10.1101/2021.07.28.454025.
32. Schneider-Mizell, C.M., Bodor, A.L., Brittain, D., Buchanan, J., Bumbarger, D.J., Elabbady, L., Gamlin, C., Kapner, D., Kinn, S., Mahalingam, G., et al. (2024). Cell-type-specific inhibitory circuitry from a connectomic census of mouse visual cortex. bioRxiv. 10.1101/2023.01.23.525290.
33. Gour, A., Boergens, K.M., Heike, N., Hua, Y., Laserstein, P., Song, K., and Helmstaedter, M. (2021). Postnatal connectomic development of inhibition in mouse barrel cortex. Science *371*, eabb4534. doi:10.1126/science.abb4534.
34. Patz, S., Grabert, J., Gorba, T., Wirth, M.J., and Wahle, P. (2004). Parvalbumin expression in visual cortical interneurons depends on neuronal activity and TrkB ligands during an Early period of postnatal development. Cereb Cortex *14*, 342-351. 10.1093/cercor/bhg132.
35. Schneider-Mizell, C.M., Bodor, A.L., Collman, F., Brittain, D., Bleckert, A., Dorkenwald, S., Turner, N.L., Macrina, T., Lee, K., Lu, R., et al. (2021). Structure and function of axo-axonic inhibition. Elife *10*. 10.7554/eLife.73783.

36. Rall, W. (1962). Electrophysiology of a dendritic neuron model. Biophys J *2*, 145-167. 10.1016/s0006-3495(62)86953-7.
37. Rall, W. (1962). Theory of physiological properties of dendrites. Ann N Y Acad Sci *96*, 1071-1092. 10.1111/j.1749-6632.1962.tb54120.x.
38. Stuart, G.J., and Spruston, N. (2015). Dendritic integration: 60 years of progress. Nature Neuroscience *18*, 1713-1721. 10.1038/nn.4157.
39. Beniaguev, D., Segev, I., and London, M. (2021). Single cortical neurons as deep artificial neural networks. Neuron *109*, 2727-2739.e2723. 10.1016/j.neuron.2021.07.002.
40. Losonczy, A., Makara, J.K., and Magee, J.C. (2008). Compartmentalized dendritic plasticity and input feature storage in neurons. Nature *452*, 436-441. 10.1038/nature06725.
41. Hao, J., Wang, X.D., Dan, Y., Poo, M.M., and Zhang, X.H. (2009). An arithmetic rule for spatial summation of excitatory and inhibitory inputs in pyramidal neurons. Proc Natl Acad Sci U S A *106*, 21906-21911. 10.1073/pnas.0912022106.
42. Gidon, A., and Segev, I. (2012). Principles governing the operation of synaptic inhibition in dendrites. Neuron *75*, 330-341. 10.1016/j.neuron.2012.05.015.
43. Julie, J., Giuseppe, C., Tim, K., Fursham, H., Stella, S., Teresa, G., Patricia, M., Maddalena, B., and Beatriz, R. (2023). Cadherins orchestrate specific patterns of perisomatic inhibition onto distinct pyramidal cell populations. bioRxiv, 2023.2009.2028.559922. 10.1101/2023.09.28.559922.
44. Wu, S.J., Sevier, E., Dwivedi, D., Saldi, G.A., Hairston, A., Yu, S., Abbott, L., Choi, D.H., Sherer, M., Qiu, Y., et al. (2023). Cortical somatostatin interneuron subtypes form cell-type-specific circuits. Neuron *111*, 2675-2692.e2679. 10.1016/j.neuron.2023.05.032.
45. Uezu, A., Kanak, D.J., Bradshaw, T.W., Soderblom, E.J., Catavero, C.M., Burette, A.C., Weinberg, R.J., and Soderling, S.H. (2016). Identification of an elaborate complex mediating postsynaptic inhibition. Science *353*, 1123-1129. 10.1126/science.aag0821.
46. Packer, A.M., and Yuste, R. (2011). Dense, unspecific connectivity of neocortical parvalbumin-positive interneurons: a canonical microcircuit for inhibition? J Neurosci *31*, 13260-13271. 10.1523/jneurosci.3131-11.2011.
47. Fino, E., and Yuste, R. (2011). Dense inhibitory connectivity in neocortex. Neuron *69*, 1188-1203. 10.1016/j.neuron.2011.02.025.
48. Packer, A.M., McConnell, D.J., Fino, E., and Yuste, R. (2013). Axo-dendritic overlap and laminar projection can explain interneuron connectivity to pyramidal cells. Cereb Cortex *23*, 2790-2802. 10.1093/cercor/bhs210.
49. Miles, R., Tóth, K., Gulyás, A.I., Hájos, N., and Freund, T.F. (1996). Differences between somatic and dendritic inhibition in the hippocampus. Neuron *16*, 815-823. 10.1016/s0896-6273(00)80101-4.
50. Kapfer, C., Glickfeld, L.L., Atallah, B.V., and Scanziani, M. (2007). Supralinear increase of recurrent inhibition during sparse activity in the somatosensory cortex. Nat Neurosci *10*, 743-753. 10.1038/nn1909.

51. Silberberg, G., and Markram, H. (2007). Disynaptic inhibition between neocortical pyramidal cells mediated by Martinotti cells. Neuron *53*, 735-746. 10.1016/j.neuron.2007.02.012.
52. Lee, S., Kruglikov, I., Huang, Z.J., Fishell, G., and Rudy, B. (2013). A disinhibitory circuit mediates motor integration in the somatosensory cortex. Nat Neurosci *16*, 1662-1670. 10.1038/nn.3544.
53. Pfeffer, C.K., Xue, M., He, M., Huang, Z.J., and Scanziani, M. (2013). Inhibition of inhibition in visual cortex: the logic of connections between molecularly distinct interneurons. Nat Neurosci *16*, 1068-1076. 10.1038/nn.3446.
54. Gross, G.G., Junge, J.A., Mora, R.J., Kwon, H.B., Olson, C.A., Takahashi, T.T., Liman, E.R., Ellis-Davies, G.C., McGee, A.W., Sabatini, B.L., et al. (2013). Recombinant probes for visualizing endogenous synaptic proteins in living neurons. Neuron *78*, 971-985. 10.1016/j.neuron.2013.04.017.
55. van Oostrum, M., Blok, T.M., Giandomenico, S.L., tom Dieck, S., Tushev, G., Fürst, N., Langer, J.D., and Schuman, E.M. (2023). The proteomic landscape of synaptic diversity across brain regions and cell types. Cell *186*, 5411-5427.e5423. https://doi.org/10.1016/j.cell.2023.09.028.
56. Berning, M., Boergens, K.M., and Helmstaedter, M. (2015). SegEM: Efficient Image Analysis for High-Resolution Connectomics. Neuron *87*, 1193-1206. 10.1016/j.neuron.2015.09.003.
57. Hua, Y., Loomba, S., Pawlak, V., Voit, K.M., Laserstein, P., Boergens, K.M., Wallace, D.J., Kerr, J.N.D., and Helmstaedter, M. (2022). Connectomic analysis of thalamus-driven disinhibition in cortical layer 4. Cell Rep *41*, 111476. 10.1016/j.celrep.2022.111476.
58. Turner, N.L., Macrina, T., Bae, J.A., Yang, R., Wilson, A.M., Schneider-Mizell, C., Lee, K., Lu, R., Wu, J., Bodor, A.L., et al. (2022). Reconstruction of neocortex: Organelles, compartments, cells, circuits, and activity. Cell *185*, 1082-1100.e1024. 10.1016/j.cell.2022.01.023.
59. Zhu, F., Cizeron, M., Qiu, Z., Benavides-Piccione, R., Kopanitsa, M.V., Skene, N.G., Koniaris, B., DeFelipe, J., Fransén, E., Komiyama, N.H., and Grant, S.G.N. (2018). Architecture of the Mouse Brain Synaptome. Neuron *99*, 781-799.e710. https://doi.org/10.1016/j.neuron.2018.07.007.

Materials and methods:

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^{CreERT2} homozygotes are embryonic lethal, and thus were maintained as heterozygotes. For all experiments involving Reln$^{RL/RL}$ mutants, 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).

Genotyping

DNA was sourced from tails and digested in proteinase K (prot K) 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 prot K was heat inactivated at 95°C for 10 minutes prior to use for genotyping reactions. 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.

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. Experiments involving Nkx2-1-CreER mice were performed according to [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.

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 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 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 [2]. 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]. 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] 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.

Cortical dissection and dissociation

Dissection and dissociation were performed as outlined in [3] using buffers 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 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.

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 [4]. All tissue culture was performed using 24-well optical plastic tissue culture plates [Ibidi]. Plates were pretreated with 20 μg/mL poly-D lysine [Sigma Aldrich] in borate buffer (10 mM Sodium Borate, pH 8.5 [sodium tetraborate, S9640, and sodium tetraborate, S9640, both Sigma]) 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).

Immunostaining

Cryopreserved tissue slices were processed as free-floating sections according to previous work [2] 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 2 and Table 3, 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 in TrueBlack mounting media [Biotium], 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] 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 [5].

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 afore-mentioned resolution. Initial focal points were established by the addition of support points and verified by autofocus prior to imaging.

Image processing

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 [6] plugin in FIJI [7], 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 [8]. 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) [9]. Unsupervised classification and clustering - Dylan. Unsupervised classification results were mapped onto boutons using custom code written in FIJI/ImageJ using the CLIJ2 plugin [10].

Custom code:

At most stages in the pipeline, custom code was written to facilitate image batch processing and analysis.

Metadata extraction

Metadata 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 [6]. 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).

Stitching

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 [11] to automatically stitch tiles at any stage in the pipeline into their respective multi-tile scenes using this generalized tile configurations map.

File Grouping

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.

Image standardization

We built in multiple checkpoints for image standardization to ensure that the inputs meet expectations for the given pipeline. 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).

Image Correction

We implemented background and shading correction in the X/Y dimensions by adapting the previously-published BaSiC plugin [12] to run efficiently for large, multi-channel tiled Z-stacks. 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 [10]. 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 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 [13] in 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.

Pixel Classification

Corrected and uncorrected images were converted into HDF5 format using the Ilastik4imageJ plugin [8] 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.

Image Concatenation

After pixel classification, the input used for 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.

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) [9]. 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.

Object classification

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.

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 [14], 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 [15] implemented in the Keras library in Python [16]. 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 [17].

Unsupervised clustering

Unsupervised object clustering was performed in Python using the Scanpy library [18], using GPU acceleration provided by the RAPIDS library . 16-dimensional autoencoder embeddings were loaded as features of an Anndata object [19], 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 [20] of objects with the *scanpy.tl.umap* function. Louvain clustering [21] 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 [22] and minimizing Davies-Bouldin [23] 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.

To verify the stability and information loss of our autoencoder embedding, we evaluate reconstruction loss on a held-out validation set with multiple models of identical architecture but different random seeds. We perform a forward pass through the autoencoder and obtain the decoder-reconstructed data in the original feature space. We measure reconstruction loss with Mean Squared Error (MSE). Computationally, for a dataset comprising N samples each with M features, the MSE is calculated by first determining the squared difference between the reconstructed and actual values for each feature within a sample. These squared differences are then averaged across all features for each sample, producing a single error value per sample. The mean of these error values across all NN samples provides the MSE.

Spatial Compactness 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.

Optimal Consensus Sequence:

The inter-cluster relationships are encoded in the off-diagonal entries. There are many possible perspectives in which these off-diagonal entries could be analyzed. We introduce an approach, referred to as Radial Permutation Analysis, that concentrates on the superdiagonal and subdiagonal elements of our spatial correlation matrix. We consider permutations of the Louvain cluster labels with a fixed pivot, where a pivot is a cluster label that serves as the first (fixed) element in the permutation. We investigate

how the clusters are related by expanding from the pivot using pairwise correlations computed in our spatial correlation matrix. From this, we sought to identify permutations/orderings/organizations of the clusters that present the strongest sequential spatial correlation. More specifically, we focus on the relationship of adjacent clusters in the permutation given a fixed anchor, and we evaluate sequential spatial correlation as a function of all the possible permutations of the cluster's labels. Iteratively, we permute the non-fixed elements in the sequence, record the sum of the superdiagonal and the subdiagonal, and select the top n permutations based on the sum. On the selected permutations, we compute the frequency of occurrences of each cluster label at each index across the top n permutations, before normalizing these values to percentages. For the first non-fixed element at the current iteration, we select the cluster with the highest normalized percentage, fix it as the next element in our final sequence. We repeat this process until all clusters have been assigned a position in the sequence. Lastly, we reorder the previously computed spatial correlation matrix using this ordering.

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 [10], and could be adapted to map any numerical values in a csv onto a corresponding binarized mask.

EM Analysis of Contiguous Components:

For the morphological analysis from electron microscopy, we used neuronal skeletons from MICrONs dataset of mouse visual cortex. Data was taken from materialization version 795 (*doi*: *10.5281/zenodo.10463756*). We used the skeletonizations of inhibitory neurons classified as either “Perisomatic Targeting Cells” (PeriTC) or “Distal Dendrite Targeting Cells” (DistTC). Because these are anatomical and not molecular classes, PeriTCs can include both parvalbumin basket cells as well as other molecular classes (CCK+/VIP+, CCK+/VIP-, etc) and DistTCs could include SST neurons as well as other cell types. However, expert labeling based on reconstructions suggested that these categories are dominated by PV and SST molecular classes [24].

To identify contiguous targeting domains, we mapped all synaptic outputs from interneuronal axons to segmentation IDs. We identified the subset of those synapses that were onto cells classified as excitatory neurons based on perisomatic features (table name `*aibs_metamodel_celltypes_v661*` [25]). Next, for each connection from an interneuron onto an excitatory cell, we computed the graph connected components of the axonal skeleton after removing skeleton vertices with synapses onto other targets. For each connected component, we then counted the number of synapses within that connection. If two synapses onto different targets were close enough to be part of the same skeleton vertex (mean distance between skeleton vertices is approximately 1.5 μm), they were treated as distinct single-synapse components.

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	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
	Nex-Cre	Math2/Cre, Neurod6-cre, Neurod6/Cre, NexCre, NEX-Cre	Neurod6$^{tm1(cre)Kan}$	MGI:2668659	Schwab, et al., 2000
Reporter Lines	Ai34	Ai34D or Ai34(RCL-Syp/tdT)-D	B6;129S-Gt(ROSA)26Sor$^{tm34.1(CAG-Syp/tdTomato)Hze}$/J	012570	Jax
	Ai14	Ai14 , Ai14D or 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: Primary Antibodies

Target	Host	mono/ poly	clone	Subtype	Source	Catalog	Dilution
Ankyrin G	Rb	pAb	n/a	IgG	Synaptic Systems	386 003	1:1000
Gephyrin (pS270)	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	n/a	IgY	Synaptic Systems	188 006	1:1000
PV	Chk	pAb	n/a	IgY	Novus Biologicals	NBP2-50036	1:2000
Sst	Rt	mAb	YC7	IgG2b	Chemicon	MAB354	1:300
Synaptophysin	Ms	mAb	SVP-38	IgG1	Sigma	S5768	1:200
VGAT	Rb	pAb	n/a	IgG	Synaptic Systems	131 002	1:500

Table 3: 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

Appendix A

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:

Artificial 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

Materials and Methods References:

1. Taniguchi, H., Lu, J., and Huang, Z.J. (2013). The spatial and temporal origin of chandelier cells in mouse neocortex. Science *339*, 70-74. 10.1126/science.1227622.
2. Notter, T., Panzanelli, P., Pfister, S., Mircsof, D., and Fritschy, J.M. (2014). A protocol for concurrent high-quality immunohistochemical and biochemical analyses in adult mouse central nervous system. Eur J Neurosci *39*, 165-175. 10.1111/ejn.12447.
3. Nunnelly, L.F., Campbell, M., Lee, D.I., Dummer, P., Gu, G., Menon, V., and Au, E. (2022). St18 specifies globus pallidus projection neuron identity in MGE lineage. Nat Commun *13*, 7735. 10.1038/s41467-022-35518-5.
4. Barron, J.J., Mroz, N.M., Taloma, S.E., Dahlgren, M.W., Ortiz-Carpena, J., Dorman, L.C., Vainchtein, I.D., Escoubas, C.C., Molofsky, A.B., and Molofsky, A.V. (2023). Group 2 innate lymphoid cells promote inhibitory synapse development and social behavior. bioRxiv. 10.1101/2023.03.16.532850.
5. Gallerani, N., and Au, E. (2020). Loss of Clustered Protocadherin Diversity Alters the Spatial Distribution of Cortical Interneurons in Mice. Cereb Cortex Commun *1*, tgaa089. 10.1093/texcom/tgaa089.
6. Linkert, M., Rueden, C.T., Allan, C., Burel, J.M., Moore, W., Patterson, A., Loranger, B., Moore, J., Neves, C., Macdonald, D., et al. (2010). Metadata matters: access to image data in the real world. J Cell Biol *189*, 777-782. 10.1083/jcb.201004104.
7. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nature Methods *9*, 676-682. 10.1038/nmeth.2019.
8. Berg, S., Kutra, D., Kroeger, T., Straehle, C.N., Kausler, B.X., Haubold, C., Schiegg, M., Ales, J., Beier, T., Rudy, M., et al. (2019). ilastik: interactive machine learning for (bio)image analysis. Nature Methods *16*, 1226-1232. 10.1038/s41592-019-0582-9.
9. Ollion, J., Cochennec, J., Loll, F., Escudé, C., and Boudier, T. (2013). TANGO: a generic tool for high-throughput 3D image analysis for studying nuclear organization. Bioinformatics *29*, 1840-1841. 10.1093/bioinformatics/btt276.
10. Haase, R., Royer, L.A., Steinbach, P., Schmidt, D., Dibrov, A., Schmidt, U., Weigert, M., Maghelli, N., Tomancak, P., Jug, F., and Myers, E.W. (2020). CLIJ: GPU-accelerated image processing for everyone. Nature Methods *17*, 5-6. 10.1038/s41592-019-0650-1.
11. Preibisch, S., Saalfeld, S., and Tomancak, P. (2009). Globally optimal stitching of tiled 3D microscopic image acquisitions. Bioinformatics *25*, 1463-1465. 10.1093/bioinformatics/btp184.
12. Peng, T., Thorn, K., Schroeder, T., Wang, L., Theis, F.J., Marr, C., and Navab, N. (2017). A BaSiC tool for background and shading correction of optical microscopy images. Nature Communications *8*, 14836. 10.1038/ncomms14836.

13. Graf, J., Cho, S., McDonough, E., Corwin, A., Sood, A., Lindner, A., Salvucci, M., Stachtea, X., Van Schaeybroeck, S., Dunne, P.D., et al. (2021). FLINO: a new method for immunofluorescence bioimage normalization. Bioinformatics *38*, 520-526. 10.1093/bioinformatics/btab686.
14. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Louppe, G., Prettenhofer, P., Weiss, R., et al. (2011). Scikit-learn: Machine Learning in Python. ArXiv *abs/1201.0490*.
15. Kramer, M.A. (1991). Nonlinear principal component analysis using autoassociative neural networks. AIChE Journal *37*, 233-243. https://doi.org/10.1002/aic.690370209.
16. Chollet, F.a.o. (2015). https://keras.io.
17. Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y., Brenner, M., Loh, P.-r., and Raychaudhuri, S. (2019). Fast, sensitive and accurate integration of single-cell data with Harmony. Nature Methods *16*, 1289-1296. 10.1038/s41592-019-0619-0.
18. Wolf, F.A., Angerer, P., and Theis, F.J. (2018). SCANPY: large-scale single-cell gene expression data analysis. Genome Biology *19*, 15. 10.1186/s13059-017-1382-0.
19. Virshup, I., Rybakov, S., Theis, F.J., Angerer, P., and Wolf, F.A. (2021). anndata: Annotated data. bioRxiv, 2021.2012.2016.473007. 10.1101/2021.12.16.473007.
20. McInnes, L., and Healy, J. (2018). UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. ArXiv *abs/1802.03426*.
21. Blondel, V.D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. Journal of Statistical Mechanics: Theory and Experiment *2008*, 10008. 10.1088/1742-5468/2008/10/P10008.
22. Caliński, T., and Harabasz, J. (1974). A dendrite method for cluster analysis. Communications in Statistics *3*, 1-27. 10.1080/03610927408827101.
23. Davies, D.L., and Bouldin, D.W. (1979). A Cluster Separation Measure. IEEE Transactions on Pattern Analysis and Machine Intelligence *PAMI-1*, 224-227. 10.1109/TPAMI.1979.4766909.
24. Schneider-Mizell, C.M., Bodor, A.L., Brittain, D., Buchanan, J., Bumbarger, D.J., Elabbady, L., Gamlin, C., Kapner, D., Kinn, S., Mahalingam, G., et al. (2024). Cell-type-specific inhibitory circuitry from a connectomic census of mouse visual cortex. bioRxiv. 10.1101/2023.01.23.525290.
25. Leila, E., Sharmishtaa, S., Shang, M., Gayathri, M., Casey, S.-M., Agnes, L.B., Bae, J.A., Derrick, B., JoAnn, B., Daniel, J.B., et al. (2024). Perisomatic Features Enable Efficient and Dataset Wide Cell-Type Classifications Across Large-Scale Electron Microscopy Volumes. bioRxiv, 2022.2007.2020.499976. 10.1101/2022.07.20.499976.

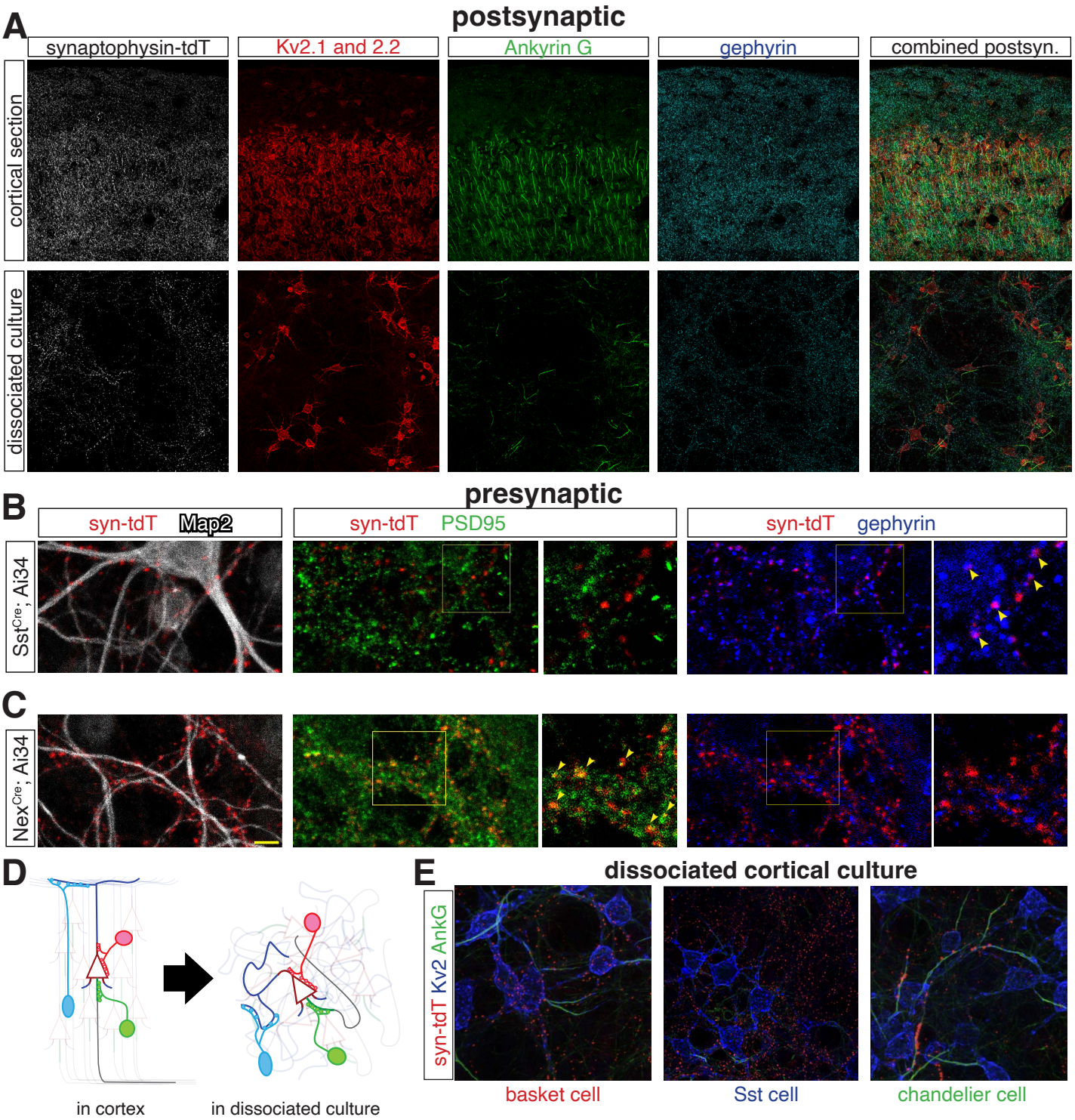

A
postsynaptic
synaptophysin-tdT
Kv2.1 and 2.2
Ankyrin G
gephyrin
combined postsyn.
cortical section
dissociated culture
B
presynaptic
syn-tdT Map2
syn-tdT PSD95
syn-tdT gephyrin
Sst^{Cre}; Ai34
C
Nex^{Cre}; Ai34
D
in cortex
in dissociated culture
E
dissociated cortical culture
syn-tdT Kv2 AnkG
basket cell
Sst cell
chandelier cell

Figure 1

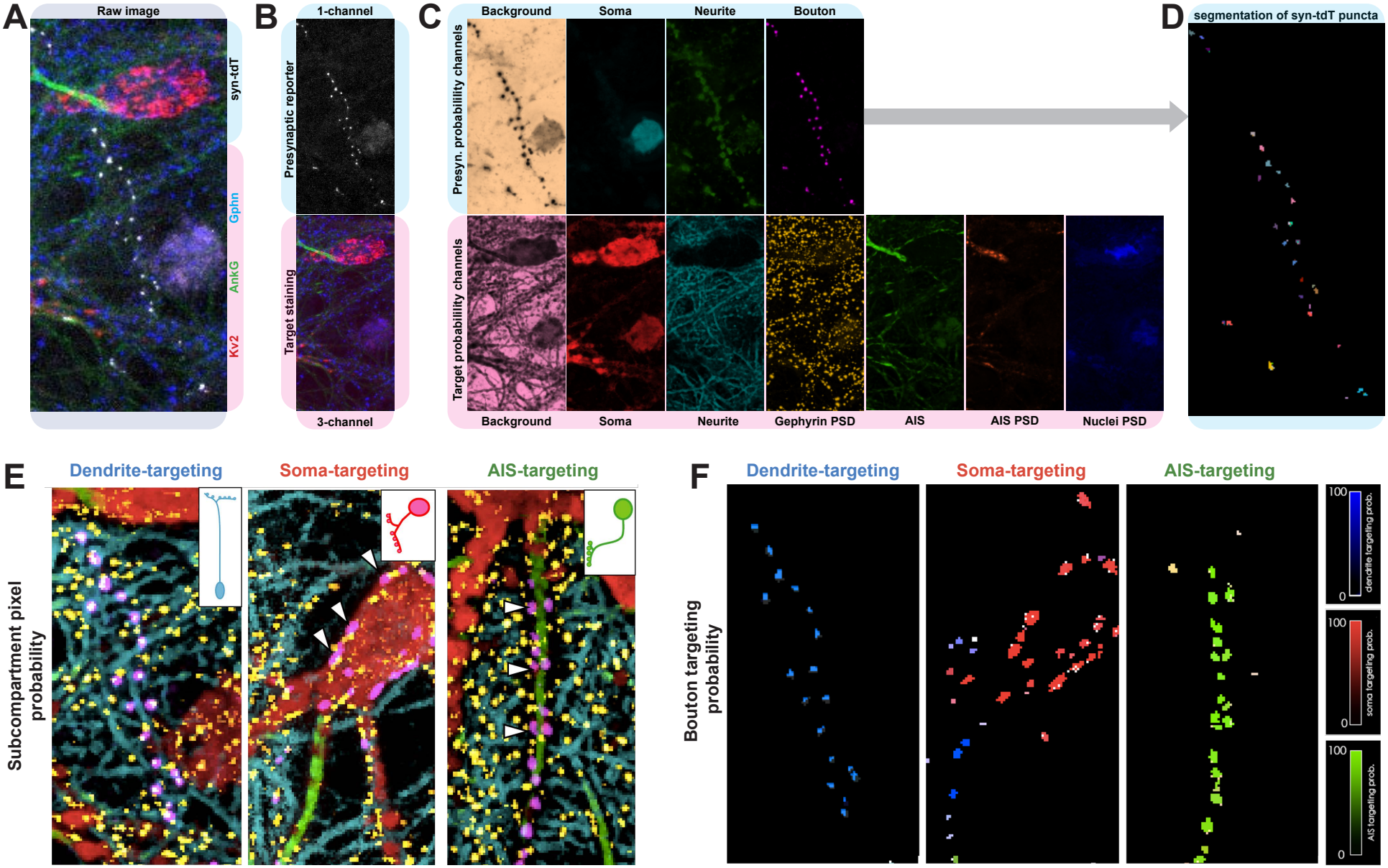

A
Raw image
syn-tdT
Gphn
AnkG
Kv2
B
1-channel
Presynaptic reporter
Target staining
3-channel
C
Background
Soma
Neurite
Bouton
Presyn. probabilility channels
Target probabilility channels
Background
Soma
Neurite
Gephyrin PSD
AIS
AIS PSD
Nuclei PSD
D
segmentation of syn-tdT puncta
E
Dendrite-targeting
Soma-targeting
AIS-targeting
Subcompartment pixel probability
F
Dendrite-targeting
Soma-targeting
AIS-targeting
Bouton targeting probability
100
0
dendrite targeting prob.
soma targeting prob.
AIS targeting prob.

Figure 2

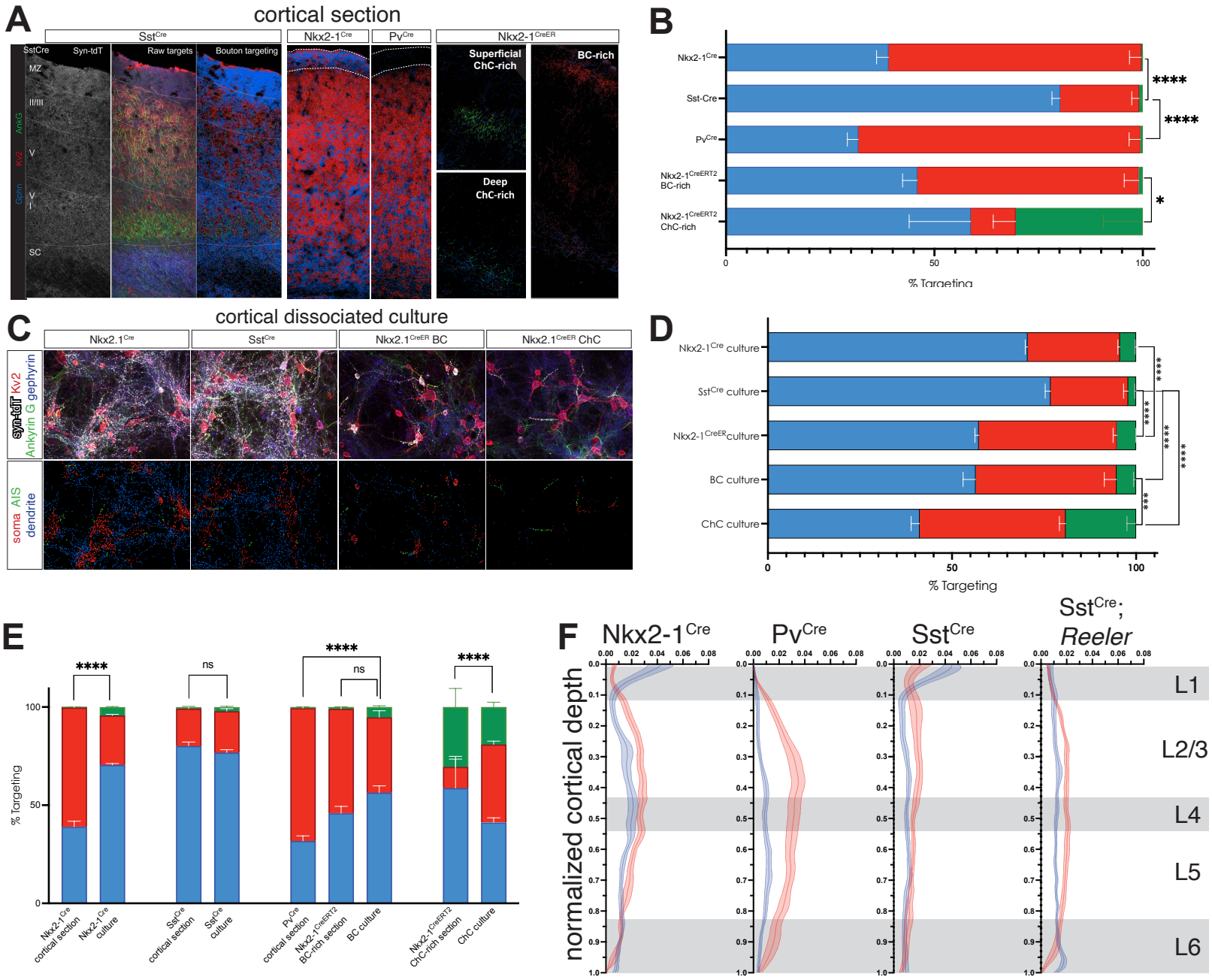

A
cortical section
Sst^Cre
Nkx2-1^Cre
Pv^Cre
Nkx2-1^CreER
SstCre
Syn-tdT
Raw targets
Bouton targeting
MZ
II/III
V
VI
SC
AnkG
Kv2
Gphn
Superficial ChC-rich
Deep ChC-rich
BC-rich
B
Nkx2-1^Cre
Sst-Cre
Pv^Cre
Nkx2-1^CreERT2 BC-rich
Nkx2-1^CreERT2 ChC-rich
****
****
*
0
50
100
% Targeting
C
cortical dissociated culture
Nkx2.1^Cre
Sst^Cre
Nkx2.1^CreER BC
Nkx2.1^CreER ChC
syn-tdT Kv2
Ankyrin G gephyrin
soma AIS
dendrite
D
Nkx2-1^Cre culture
Sst^Cre culture
Nkx2-1^CreER culture
BC culture
ChC culture
****
***
0
50
100
% Targeting
E
****
ns
****
ns
****
100
50
0
% Targeting
Nkx2.1^Cre cortical section
Nkx2-1^Cre culture
Sst^Cre cortical section
Sst^Cre culture
Pv^Cre cortical section
Nkx2.1^CreERT2 BC-rich section
BC culture
Nkx2.1^CreERT2 ChC-rich section
ChC culture
F
Nkx2-1^Cre
Pv^Cre
Sst^Cre
Sst^Cre; Reeler
normalized cortical depth
L1
L2/3
L4
L5
L6

Figure 3

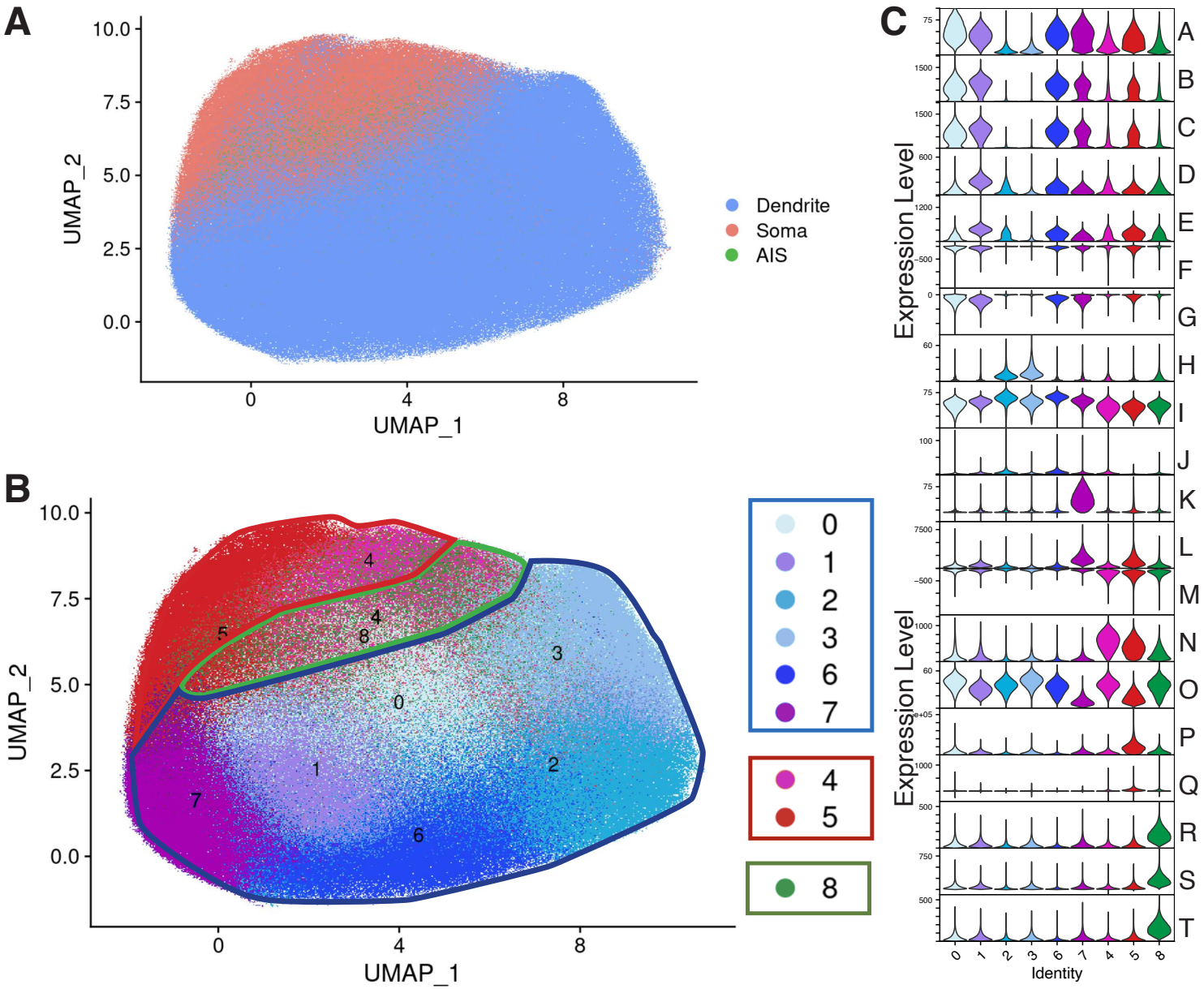
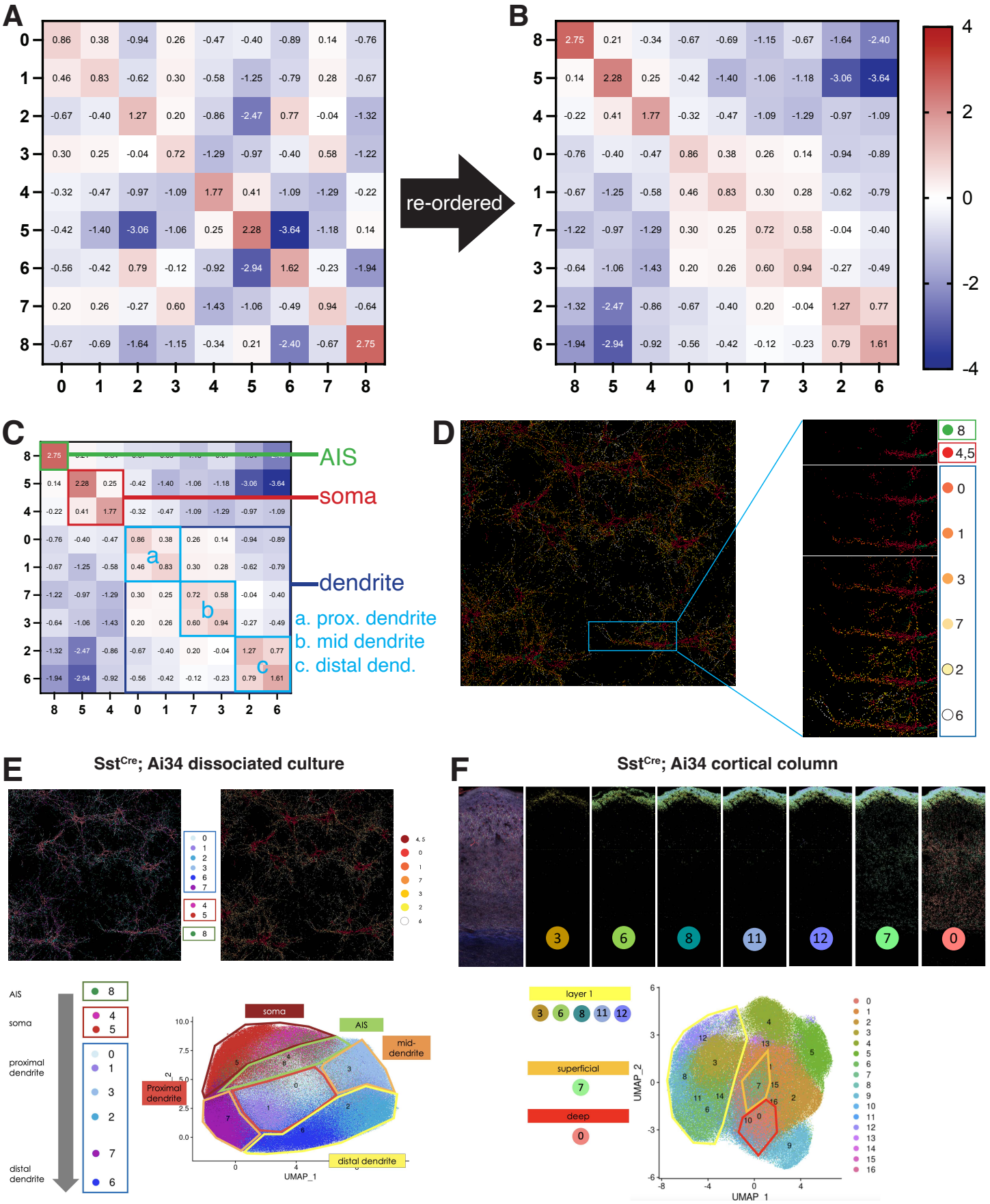
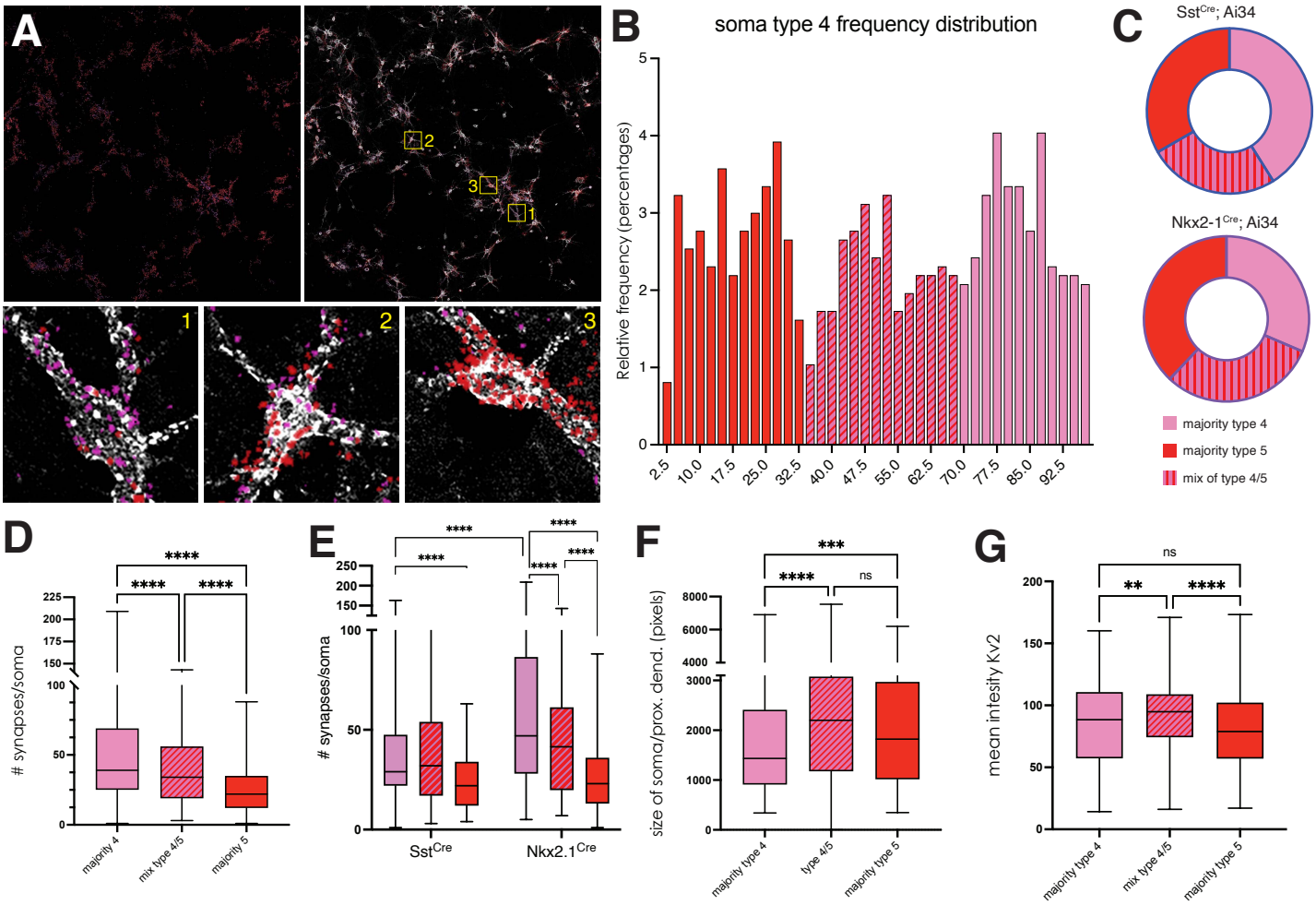

A
UMAP_2
UMAP_1
Dendrite
Soma
AIS
B
0
1
2
3
6
7
4
5
8
C
Expression Level
Identity
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T

Figure 4


A
B
re-ordered
C
AIS
soma
dendrite
a. prox. dendrite
b. mid dendrite
c. distal dend.
D
E
Sst^Cre; Ai34 dissociated culture
AIS
soma
proximal dendrite
distal dendrite
Proximal dendrite
mid-dendrite
distal dendrite
UMAP_1
UMAP_2
F
Sst^Cre; Ai34 cortical column
layer 1
superficial
deep
UMAP 1
UMAP_2


A
1
2
3
B
soma type 4 frequency distribution
Relative frequency (percentages)
0
1
2
3
4
5
2.5
10.0
17.5
25.0
32.5
40.0
47.5
55.0
62.5
70.0
77.5
85.0
92.5
C
Sst^{Cre}; Ai34
Nkx2-1^{Cre}; Ai34
majority type 4
majority type 5
mix of type 4/5
D
synapses/soma
majority 4
mix type 4/5
majority 5
E
synapses/soma
Sst^{Cre}
Nkx2.1^{Cre}
F
size of soma/prox. dend. (pixels)
majority type 4
type 4/5
majority type 5
ns
G
mean intesity Kv2
majority type 4
mix type 4/5
majority type 5
ns

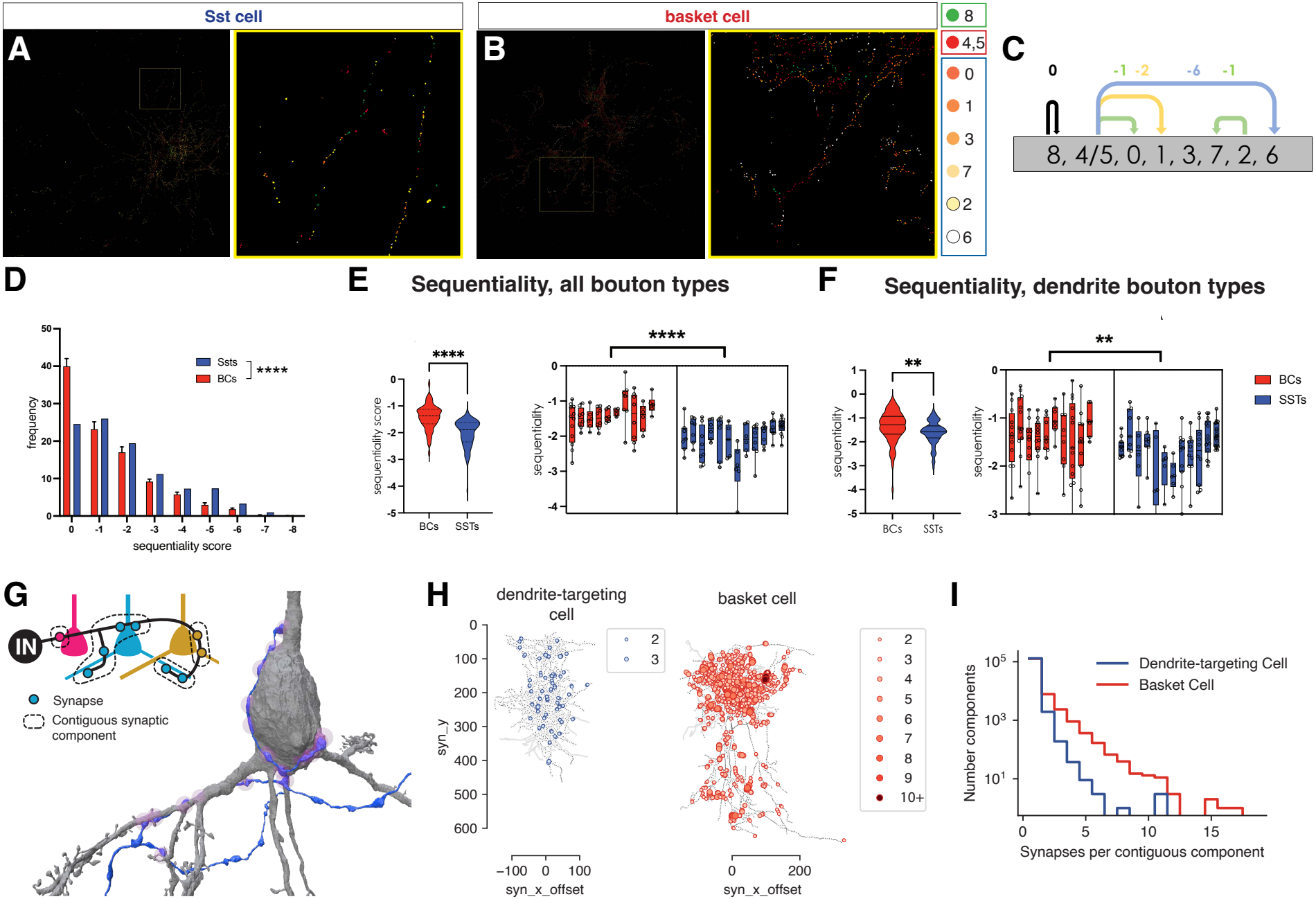

Sst cell
basket cell
A
B
C
8
4,5
0
1
3
7
2
6
0
-1 -2
-6
-1
8, 4/5, 0, 1, 3, 7, 2, 6
D
E
Sequentiality, all bouton types
F
Sequentiality, dendrite bouton types
Ssts
BCs
****
frequency
sequentiality score
sequentiality score
sequentiality
BCs
SSTs
**
G
IN
Synapse
Contiguous synaptic component
H
dendrite-targeting cell
basket cell
syn_y
syn_x_offset
syn_x_offset
10+
I
Dendrite-targeting Cell
Basket Cell
Number components
Synapses per contiguous component

Figure 7