Article

Connectomics of predicted *Sst* transcriptomic types in mouse visual cortex

https://doi.org/10.1038/s41586-025-08805-6

Received: 27 March 2023

Accepted: 18 February 2025

Published online: 9 April 2025

Open access

Check for updates

Clare R. Gamlin[1], Casey M. Schneider-Mizell[1], Matthew Mallory[1], Leila Elabbady[1,6], Nathan Gouwens[1], Grace Williams[1], Alice Mukora[1], Rachel Dalley[1], Agnes L. Bodor[1], Derrick Brittain[1], JoAnn Buchanan[1], Daniel J. Bumbarger[1], Emily Joyce[1], Daniel Kapner[1], Sam Kinn[1], Gayathri Mahalingam[1], Sharmishtaa Seshamani[1], Marc Takeno[1], Russel Torres[1], Wenjing Yin[1], Philip R. Nicovich[1], J. Alexander Bae[2,3], Manuel A. Castro[2], Sven Dorkenwald[2,4,7], Akhilesh Halageri[2], Zhen Jia[2,4], Chris Jordan[2], Nico Kemnitz[2], Kisuk Lee[2,5], Kai Li[4], Ran Lu[2], Thomas Macrina[2,4], Eric Mitchell[2], Shanka Subhra Mondal[2,3], Shang Mu[2], Barak Nehoran[2,4], Sergiy Popovych[2,4], William Silversmith[2], Nicholas L. Turner[2,4], William Wong[2], Jingpeng Wu[2], Szi-chieh Yu[2], Jim Berg[1], Tim Jarsky[1], Brian Lee[1], H. Sebastian Seung[2,4], Hongkui Zeng[1], R. Clay Reid[1], Forrest Collman[1]✉, Nuno Maçarico da Costa[1]✉ & Staci A. Sorensen[1]✉

Neural circuit function is shaped both by the cell types that comprise the circuit and the connections between them[1]. Neural cell types have previously been defined by morphology[2,3], electrophysiology[4], transcriptomic expression[5,6], connectivity[7–9] or a combination of such modalities[10–12]. The Patch-seq technique enables the characterization of morphology, electrophysiology and transcriptomic properties from individual cells[13–15]. These properties were integrated to define 28 inhibitory, morpho-electric-transcriptomic (MET) cell types in mouse visual cortex[16], which do not include synaptic connectivity. Conversely, large-scale electron microscopy (EM) enables morphological reconstruction and a near-complete description of a neuron's local synaptic connectivity, but does not include transcriptomic or electrophysiological information. Here, we leveraged morphological information from Patch-seq to predict the transcriptomically defined cell subclass and/or MET-type of inhibitory neurons within a large-scale EM dataset. We further analysed Martinotti cells—a somatostatin (*Sst*)-positive[17] morphological cell type[18,19]—which were classified successfully into *Sst* MET-types with distinct axon myelination and synaptic output connectivity patterns. We demonstrate that morphological features can be used to link cell types across experimental modalities, enabling further comparison of connectivity to gene expression and electrophysiology. We observe unique connectivity rules for predicted *Sst* cell types.

To understand neural circuit function, it is critical to determine both the cell types within the circuit and how they interconnect. Inhibitory cells regulate when and how information travels within a neural circuit[20] and, in the mouse primary visual cortex (VISp), the landscape of inhibitory neurons is complex. At least 28 inhibitory cell types (MET-types, that is, cell types defined by concordant morphology (M), electrophysiology (E) and transcriptomic expression (T)) have been identified by Patch-seq[21]. Patch-seq enables the morphology, transcriptomic profile and electrophysical signature to be determined for individual cells from in vitro slices[13–15,22]. However, this technique cannot resolve how these cells connect with each other. Such connectivity patterns can be identified with large-scale electron microscopy (EM)[23]. However, EM lacks the molecular information that could be used to relate connectivity data to transcriptomically defined cell types and their gene expression. We therefore performed an integrative morphological analysis to understand how these multi-modal cell types connect within the local circuit.

We leverage the morphological features of inhibitory MET-types to predict the MET-type or subclass identity of inhibitory cells in a large-scale EM dataset from mouse VISp. We then focused our study on Martinotti cells (MCs) in layers 4 and 5 (L4, L5) because they are well-accepted to be somatostatin (*Sst*) positive[17,24], are morphologically distinct from other subclasses (for example, vasoactive intestinal peptide (*Vip*) and parvalbumin (*Pvalb*)) and are morphologically diverse[16]. Previous functional studies have also shown that MCs can connect broadly in a 'blanket of inhibition'[25–27] or inhibit a subset of excitatory neuron targets within and across cortical layers[28–32]. We sought to determine whether differences in synaptic connectivity of individual cells were aligned to differences in MET-types.

[1]Allen Institute for Brain Science, Seattle, WA, USA. [2]Princeton Neuroscience Institute, Princeton University, Princeton, NJ, USA. [3]Electrical and Computer Engineering Department, Princeton University, Princeton, NJ, USA. [4]Computer Science Department, Princeton University, Princeton, NJ, USA. [5]Brain and Cognitive Sciences Department, Massachusetts Institute of Technology, Cambridge, MA, USA. [6]Present address: Department of Neurobiology and Biophysics, University of Washington, Seattle, WA, USA. [7]Present address: Allen Institute for Brain Science, Seattle, WA, USA. ✉e-mail: forrestc@alleninstitute.org; nunod@alleninstitute.org; stacis@alleninstitute.org

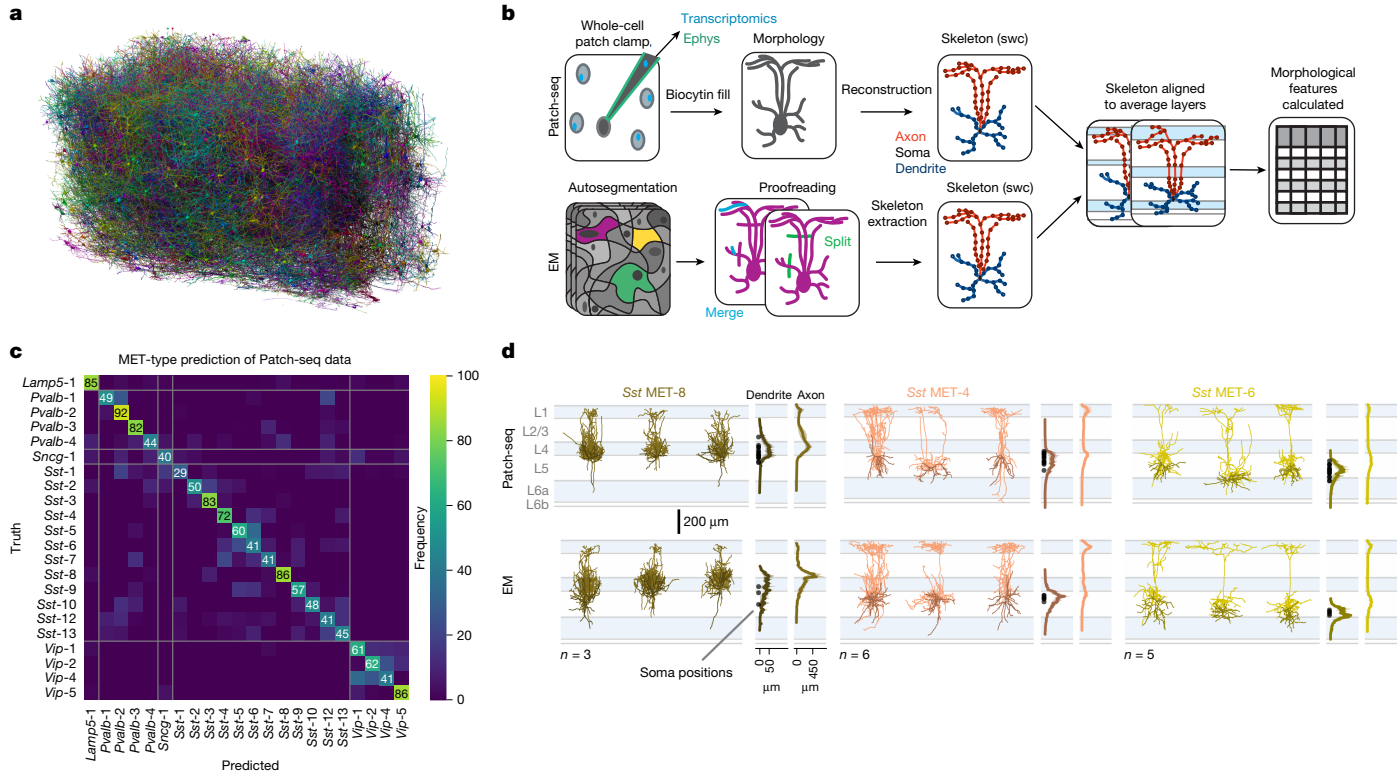


Fig. 1 | Comparison of EM and Patch-seq pipelines and morphologies. **a**, Subset of morphologies available from large-scale millimetre-squared EM dataset volume. Each cell is in a different colour. **b**, Schematic representation of the Patch-seq and EM pipelines for generating morphological reconstructions and comparison of features across pipelines. Ephys, electrophysiology; swc, file format of skeletonized neuronal reconstructions. **c**, Confusion matrix of RFC MET-type predictions showing the frequency with which the RFC correctly predicted the MET-type of Patch-seq cells (normalized by row). The RFC was trained on morphological features of inhibitory Patch-seq data ($n = 477$). We used this classifier to predict MET-type or subclass identity of EM cells ($n = 173$). **d**, Top, example Patch-seq morphology and average axon/dendrite histograms of MET-types. Bottom, example EM cells morphology and average axon/dendrite histograms grouped by predicted MET-type (MET-8, $n = 3$; MET-4, $n = 6$; MET-6, $n = 5$). Illustration in **a** adapted from refs. 33,60, Springer Nature Limited.

We reconstructed MCs from the EM, measured their synaptic connectivity and predicted their MET-type identity using a classifier trained on the morphology of inhibitory Patch-seq cells. We find that all of the EM MCs ($n = 16$) are predicted to belong to *Sst* MET-types. We also find that the predicted MET-types have distinct connectivity patterns. Individual cells within a predicted MET-type had largely consistent biases in connectivity onto excitatory targets. We also find that *Sst* MET-types can differ in the number and size of output synapses as well as axonal myelination patterns. These differences in synaptic features, connectivity, and myelination probably support distinct functional roles for inhibitory *Sst* MET-types.

By linking Patch-seq and EM data through neuron morphology, we have developed a method to explore the relationships between transcriptomically defined cell types, morphology, electrophysiology and synaptic connectivity. Using this approach, we find evidence that inhibitory MET-types participate in distinct cortical circuits and employ unique connectivity rules.

EM morphology is comparable with Patch-seq

A large-scale serial-section EM dataset, as described in the MICrONS consortium paper[23], was collected from an adult mouse VISp and some higher-order visual cortical areas (Fig. 1a). In brief, the largest portion of the structural dataset contains 82,247 neurons, which have been segmented initially using a machine-learning model to allow synapse-level characterization of the connectivity and ultrastructure. However, these neurons require further extensive proofreading to extract their complete morphological information. This manuscript and ref. 33 focus on proofread subsets of these neurons to enable this characterization. Schneider-Mizell et al.[33] proofread neurons in a 100 µm × 100 µm columnar sample, which contains 163 inhibitory cells. These neurons have morphological features that are consistent with most of the main inhibitory subclasses (*Lamp5*, *Sst*, *Pvalb*, *Vip*).

In this study, we focus on a dataset of curated MCs ($n = 16$) that are in the column and across the VISp region of the volume. MCs were identified by their sparsely spiny dendrites, an axon emerging from the pia side of the soma and a primary axon branch that reached L1 (ref. 19). The MCs were proofread using the same approach as the cells in the columnar sample, but every algorithmically identified branch and end point was checked manually and extended to create a comprehensive reconstruction. To date, these are some of the largest EM reconstructions of single neurons, with more than 5,000 µm of dendrites and more than 23,000 µm of axon reconstructed, and extending more than 600 µm vertically and horizontally across the cortex. The largest MC, in terms of output, makes 12,540 total output synapses across 3,937 unique postsynaptic targets.

To determine the feasibility of using a classifier trained on the morphological features from Patch-seq data to predict the identity of EM cells, we compared the mean and distribution of features of inhibitory neurons from both datasets (Extended Data Fig. 1). The Patch-seq data span all main inhibitory cell subclasses, but, unlike the EM dataset, cells were sampled using transgenic lines and targeted for electrophysiological recordings[16].

To compare the morphology of EM and Patch-seq cells, we aligned EM reconstructions to an average cortical layer space and calculated 43 morphological features using a pipeline developed for Patch-seq data[12,16] (Fig. 1b). Despite the different sampling methods, we find that the values from the inhibitory cells are similar between datasets.

Comparing *z* scores, the mean of the Patch-seq and EM distributions are less than 1 s.d. apart for most features (42 of 43, axon_depth_pc_04) (Extended Data Figs. 1 and 2a). All but one feature (axon_depth_pc_03) of the EM MCs (presumed *Sst*-positive) also fall within the range of *Sst* Patch-seq values (Extended Data Fig. 3). We therefore proceeded to use the morphological features to predict inhibitory subclass identity for all inhibitory neurons in the columnar sample and MET-type identity for EM MCs (Extended Data Fig. 4).

Classifier predicts *Sst* subclass of EM MCs

We trained a random forest classifier (RFC) on the morphological features of inhibitory Patch-seq cells to predict inhibitory MET-types. We achieved an estimated classifier accuracy of 59%, which is greater than expected if cells were assigned randomly to one of the 22 MET-types (around 4.5%) (Fig. 1c). At the subclass level, *Sst* Patch-seq cells were predicted correctly to be in the *Sst* subclass 87% of the time (Extended Data Fig. 5a).

To assess the relationship between the frequency and accuracy of a predicted label, we predicted a MET-type label for each Patch-seq cell over 500 iterations and compared the frequency versus the accuracy of that prediction (Extended Data Fig. 2b; Methods). We find that MET-type labels are more likely to be accurate if the label was predicted at least 54% of the time (Extended Data Fig. 2b). We then used this frequency as a threshold to select EM cells for connectivity analysis. All EM MCs were predicted to belong to an *Sst* MET-type (Fig. 1d and Extended Data Fig. 4) with a frequency of 79% or greater ('Reliability'; Extended Data Table 1).

EM MCs are assigned to several MET-types

MET-types[16] represent cell types with concordant morphology, electrophysiology and transcriptome. Some MET-types contain cells from a single transcriptomic type (t-type) whereas other MET-types contain cells across several transcriptomically similar t-types[6]. The L4 and L5 EM MCs are predicted to belong to one of five *Sst* MET-types described below.

Three EM cells are predicted to belong to the *Sst* MET-8 type (Extended Data Fig. 6). *Sst* MET-8 consists of cells from a single t-type, *Hpse Cbln4*, which have somas located in L4 and upper L5. *Hpse Cbln4* cells are described as non-MCs in somatosensory cortex[22,31], but have some L1 projection in primary visual cortex and hence are considered MCs[16,22].

Six EM cells are predicted to be *Sst* MET-4 type (Extended Data Fig. 6), which consists predominantly of cells from the *Sst Calb2 Pdlim5* t-type. The *Sst Calb2 Pdlim5* t-type splits across layers with some in L2 and L3 (*Sst* MET-3) and some in upper L5 (*Sst* MET-4)[16], but all *Sst Calb2 Pdlim5* cells are characterized by a L1-dominant axon lamination ('Martinotti') pattern[19,34], consistent with T-shaped MCs, although *Sst* MET-4 cells have the overall shape of a fanning-out MC[28,30].

Five EM cells are predicted to be *Sst* MET-6 (Extended Data Fig. 6), which are described as having 'T-shaped' morphology with an axon that reaches L1 but has dominant L5 innervation[30]. The *Sst* MET-6 type is composed of several t-types that are in proximity along the transcriptomic dendrogram, including *Sst Myh8 Etv1* and *Sst Chrna2 Glra3*.

Finally, one cell is predicted to be in the *Sst* MET-5 type and one is predicted to be *Sst* MET-9 (Extended Data Fig. 6). *Sst* MET-5 is morphologically similar to *Sst* MET-6 (ref. 16), with both *Sst Myh8 Etv1* and *Sst Nr2f2 Necab1* cells (similar to fanning-out MCs)[16,28,30]. The *Sst* MET-9 type is made up primarily of the *Sst Tac2 Tacstd* t-type and has an axon that reaches L1 but has a peak in L5.

These results demonstrate that we can reliably assign a morphological, electrophysiological and transcriptomic identity to neurons characterized in an EM volume using local dendritic and axonal morphology. In the remainder of the paper, we will refer to these 'predicted MET-types' as MET-types and focus analysis on those with at least three cells (MET-8, MET-4 and MET-6).

To determine what proportion of the *Sst* subclass is represented by each of these MET-types, we referenced a published whole mouse brain MERFISH dataset (Methods), which facilitates quantification of the MET- and t-type spatial distribution across the mouse brain. We found that *Sst* MET-4, MET-6 and MET-8 represent 3.4%, 16.5% and 18.5% of the *Sst* subclass, respectively, in VISp (Extended Data Table 2). Combined with three other *Sst* MET-types, the L2 or L3 *Sst* MET-types MET-2 (16.9%) and MET-3 (11.8%) and the L6 *Sst* MET-12 (9.9%), these six MET-types constitute most cell types in the *Sst* subclass in VISp. Thus, with our MET-type classifier, we are able to describe the connectivity of the most abundant L4 and L5 MET-types in the *Sst* population.

Synapses and myelin vary by MET-type

From EM, we can measure a rich set of ultrastructural features, including synapse distributions onto postsynaptic cell types, what fraction of each cell type receives synapses, and myelination patterns. There is striking and strong diversity in each of these properties across the population and our classifier results facilitate examination of how these differences align with *Sst* MET-type predictions.

Output synapse features

In the EM volume, most output synapses (average 82.5% of all output synapses; Extended Data Table 1) along the reconstructed axons can be mapped onto a single postsynaptic target cell in the dataset. Each target cell was assigned a cell type based on somatic features using automated methods ('Identifying synapses'; Methods). This high rate of detection of individual target cells allowed us to confidently investigate the relationship between MET-type identity and connectivity statistics (Fig. 2a–c,f).

We observed that MET-8 cells form significantly more output synapses (9,621 ± 1,460; mean ± s.e.m.) than MET-4 (2,280 ± 250) or MET-6 (1,499 ± 121) cells (Fig. 2a–c; for quantification see Fig. 2d,f). Whereas MET-8 cells have more total axon, the MET-8 synapse density is nearly twice the density of MET-4 and MET-6, approximately 0.54 versus 0.25 synapses per micrometre (Extended Data Fig. 5c,d).

We also find significant differences in the size of output synapses between MET-types. The output synapses of MET-8 cells are significantly smaller than those from MET-4 cells. MET-8 cells form synapses that are 70% the size of MET-4 synapses and 78% of the size of MET-6 synapses (Fig. 2d). MET-8 cells also form significantly more synapses per connection than other MET-types (Fig. 2d). This finding is discussed in greater detail in 'Single versus multi-synaptic connections'.

Myelination patterns

EM data includes information about myelination, which can influence the biophysics of a cell[35]. Although inhibitory neurons[36,37], including MCs[38], have previously been shown to be myelinated, little is known so far regarding the relationship between myelination pattern, *Sst* cell types and biophysical properties. We found distinct myelination patterns across the EM MCs that varied with the three MET-types (Fig. 2a,e and Extended Data Fig. 6).

MET-4 cells (5/5 cells) are myelinated along their main ascending axon collateral. MET-8 cells (3/3) are myelinated near the soma along a short stretch of the primary axon branch, which rarely extends into L2/3. MET-6 cells are either not myelinated (1/6) or sparsely myelinated with a less clear pattern (5/6) (Fig. 2a,e and Extended Data Fig. 6). Most of the reconstructed cells had some portion of their axon myelinated, but the number of myelinated segments and total path length of the myelination varied by MET-type (Fig. 2e).

The ascending axon stalk of MET-4 cells is ensheathed by several segments (average 6.7 segments) separated by nodes of Ranvier. MET-4 cells have two-and-a-half times as many segments as MET-8 (2.7) and

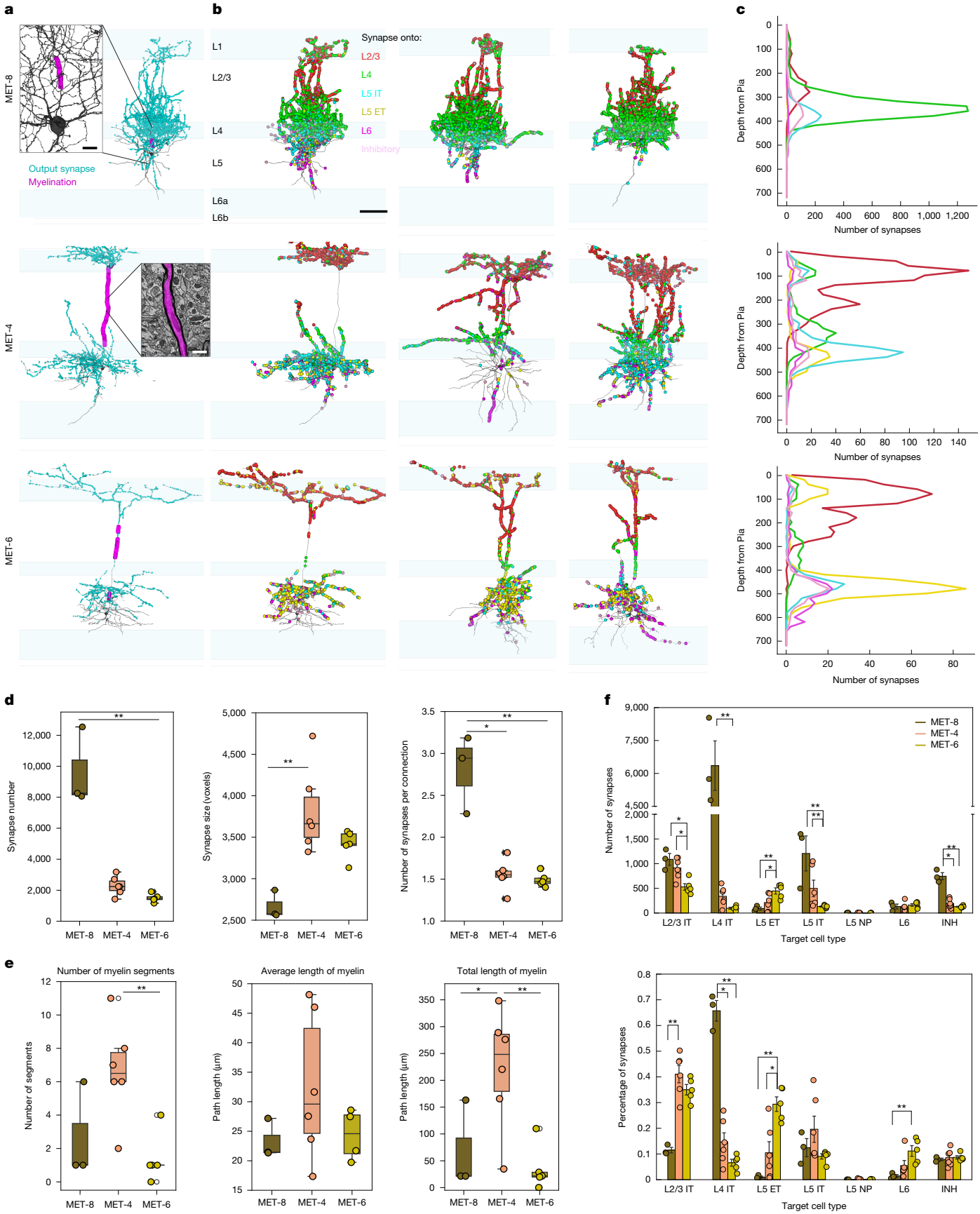


Fig. 2 | See next page for caption.

Fig. 2 | Output synapses and myelination of MET-types. a, Example cell from each predicted MET-type showing output synapses (cyan dots) and myelination (magenta). Insets, myelination of a predicted MET-8 cell and myelinated axon in EM. **b**, Same example cells (plus further examples) from each predicted MET-type with output synapses colour-coded by synapse target (as identified by a classifier trained on somatic features[60]). **c**, Average histogram of synapses onto targets by predicted MET-type. **d**, Quantification of number and size of output synapses by MET-type. **e**, Quantification of myelination features by MET-type. **f**, Quantification of the total number and percentage of output synapses onto each target cell type by MET-type. Bars indicate significant differences ($P < 0.05$ Kruskall–Wallis for group and Conover post hoc with Bonferroni correction, $^*P < 0.05$, $^{**}P < 0.01$; MET-8, $n = 3$; MET-4, $n = 6$; MET-6, $n = 5$). Scale bars, 10 µm (first inset), 1 µm (second inset) (**a**), 100 µm (**b**).

five times as many as MET-6 (1.4). We find no significant difference in the length of individual myelin segments between cell types. MET-4 cells have approximately 220 µm of total length of myelination, which is nearly three times the length of myelin of MET-8 (around 69 µm) and six times the length of myelin of MET-6 (36 µm) (Fig. 2e).

We hypothesize that such differences in output synapse number, size, density and myelination between these cells could underlie distinct signalling properties, both in terms of output strength and timing[20,39,40]. Further work to characterize the electrophysiological output of these neurons could shed light on the effects of these ultrastructural differences.

MET-types have distinct connection rules

Synaptic connectivity

After noting the distinct ultrastructural features between the inhibitory cells, we next examined whether MCs synapse onto similar cell types. Excitatory cortical cell types (pyramidal cells (PCs)) are named for both the layer in which they reside (for example, L2/3) and where they project their axons. Intra-telencephalic (IT) neurons project within the cortex-dominated anterior forebrain and extra-telencephalic (ET) neurons also project beyond the forebrain[41]. IT and ET (sometimes called PT) neurons are functionally distinct and differentially implicated in several diseases[42]. We determined the synaptic connectivity pattern of each *Sst* MET cell using automated methods to detect synapses and assign target type identity mentioned above (see 'Identifying synapses and postsynaptic targets' in Methods).

Diverse connectivity patterns amongst the inhibitory cells largely align with predicted MET-type (Fig. 2b,c,f and Extended Data Fig. 7). MET-8 cells preferentially target L4 IT PCs (65.7% of the total number of output synapses ± 4.0; mean ± s.e.m.). L5 IT and L2/3 PCs (12.5 ± 3.5% and 11.5 ± 1.1%, respectively) are the next main targets.

MET-4 cells target predominantly L2/3 PCs (41.0 ± 3.3%) followed by L5 IT PCs (19.6 ± 5.1%). L4 IT and L5 ET cells receive fewer synapses (14.7 ± 3.5% and 10.5 ± 4.3%, respectively). Four of the six predicted MET-4 followed this average pattern, although two MET-4 cells preferentially synapse with L5 ET rather than L5 IT cells, which may indicate variability either in targeting for this MET-type or misassignment of MET-type for these two cells (indicated by asterisks in Extended Data Fig. 7). Further characterization of the connectivity of MET-types will help to elucidate their intrinsic variability.

MET-6 cells form most synapses onto L2/3 PCs (35.0 ± 2.0%) followed by L5 ET neurons and L6 PCs (29.4 ± 2.8% and 11.2 ± 2.1%, respectively). Comparing the connectivity of MET-4 and MET-6 cells is of particular interest, because both groups have cells with classic Martinotti morphologies, yet differences in branching properties predict that they belong to distinct *Sst* MET-types and they have distinct connectivity profiles.

Despite having somas in L5, we find that the main targets for both *Sst* MET-4 and MET-6 are L2/3 PCs. This contrasts with the view that the primary role of L5 MCs is to inhibit the apical tufts of L5 excitatory neurons[43,44], although these synapses are present on the apical tufts of the preferred L5 target type of each type (Fig. 2c). MC axons have previously been described as overlapping with the basal dendrites of their L5 targets[19,28,45,46]; here we show that more synapses onto the preferred L5 targets are formed within L5 than L1 (peaks at 0–100 µm from pia; Fig. 2c), as opposed to purely targeting apical tufts.

Connectivity rates

The distinct output synapse distributions of the *Sst* MET-types suggests that *Sst* cells are not synapsing indiscriminately, but it does not measure directly whether all available cells are innervated, as predicted by a 'blanket' inhibition model[25]. Therefore, we calculated the percentage of neighbouring cells (within a given target cell type) that receive synapses from a presynaptic MC. We find that MCs form the most connections with neighbours within 100 µm. However, the percentages vary by MET-type and target cell type and rarely reach 100% (Fig. 3a).

MET-8, MET-4 and MET-6 cells synapse onto about 40% or fewer of the neighbouring cells from each target cell type within 50 µm, with a few notable exceptions as mentioned below. This result demonstrates that these inhibitory cells rarely synapse onto every available neuron. However, MET-8 cells do synapse onto nearly 100% of L4 IT cells within a 50-µm radius and 80% within a 100-µm radius, producing a 'blanket of inhibition' for this specific cell type. Similarly, MET-6 cells target nearly 60% of their preferred L5 postsynaptic target (L5 ET neurons within 100 µm), but not other cell types (they target fewer than 31% of the available L2/3 cells within 50-µm radial distance). MET-4 cells have a different targeting pattern. They target less than 40% of all available cells across types within 50 µm), including their preferred target cell types. Thus, each *Sst* MET-type uses a distinct inhibition pattern.

Single versus multi-synapse connections

Given that individual MCs do not connect to all neighbours, preferences in output synapse distribution could be due to differences in the number of synapses onto each target cell (number of synapses per connection). Therefore, we calculated the average number of synapses from each MC onto each target cell across MET-types (Fig. 3b). We found that most connections contain a single synapse; however, the fraction of single-synapse and multi-synapse connections differs by postsynaptic target cell type (Fig. 3b).

MET-8 cells make multi-synaptic contacts onto their preferred targets: L4, L5 IT and inhibitory cells. Specifically, more than 60% of all connections onto L4 cells are multi-synaptic and approximately 50% of connections onto L5 IT and inhibitory targets are multi-synaptic. Some targets receive more than 20 synapses from a single MET-8 cell (Fig. 3c). MET-4 cells form the fewest multi-synaptic connections and show the least modulation with respect to target cell subclass. MET-6 cells form predominantly single-synapse connections, including onto their preferred targets (L2/3 PCs), but do form highly multi-synaptic connections onto around 50% of targeted L5 ET PCs. In summary, MET-8 forms many multi-synaptic connections onto its preferred postsynaptic targets, MET-4 forms predominantly single-synapse connections and MET-6 forms multi-synaptic connections onto one specific target cell type.

Given the total number of output synapses from an individual neuron onto a limited number of targets of that type, a model that distributes those synapses randomly amongst the targets would produce a Poisson distribution of multi-synaptic connections. We find that we can reject the hypothesis that synapses from individual presynaptic cells onto a given target cell type are Poisson distributed, with only a few exceptions. For example, most (4/5 cells) MET-6 cells synapse onto L4 IT cells in a Poisson distribution (where $P > 0.05$) (Extended Data Fig. 5e).

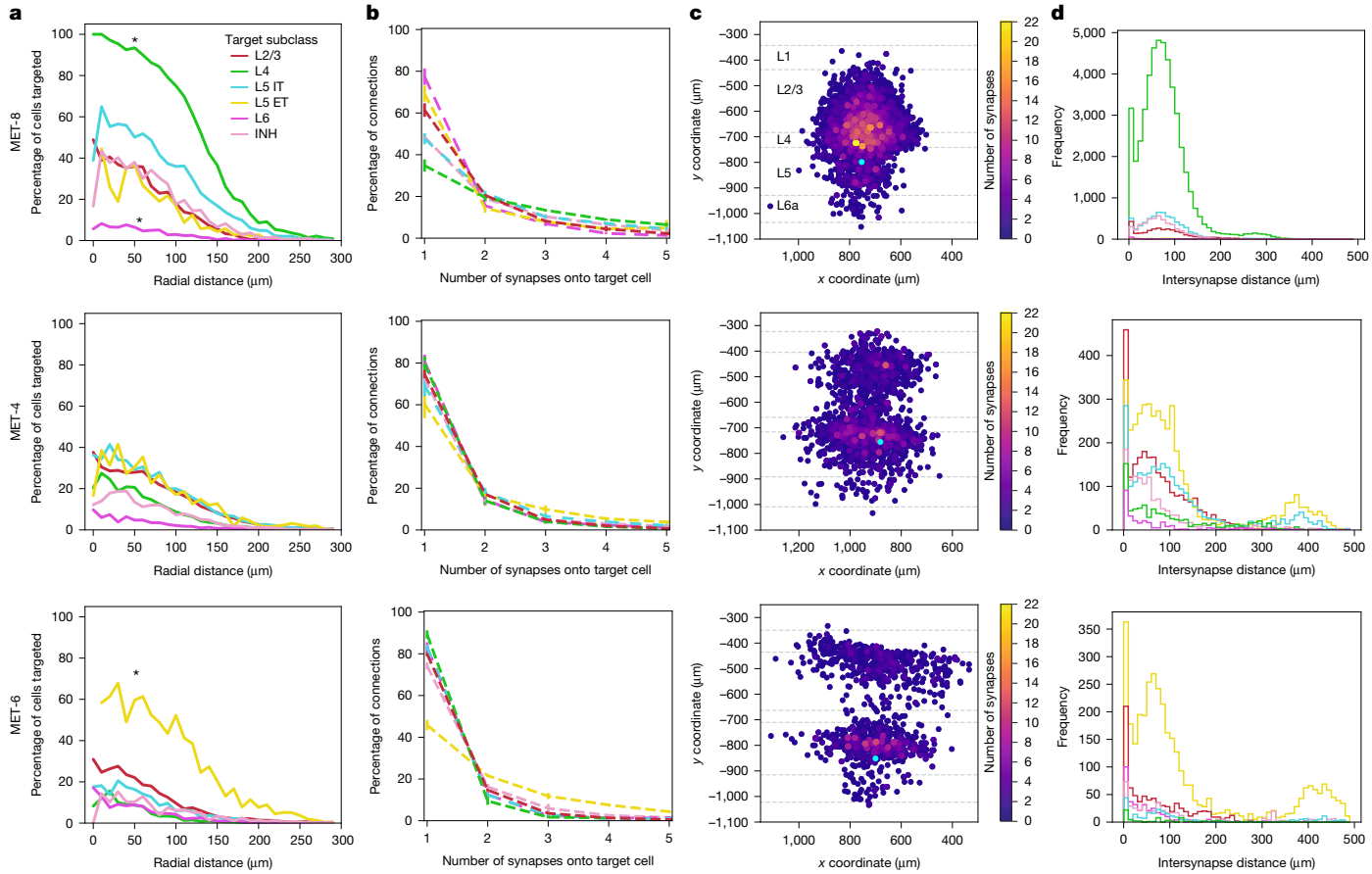


Fig. 3 | Exploring pairwise connectivity of MET-types and postsynaptic targets. a, Average histogram of percentage of each cell type that receives synapses as a function of radial (*xz*) distance between presynaptic EM cell and postsynaptic target. All distributions were compared pairwise by a Kolmogorov–Smirnov test with a false discovery rate correction. *$P < 0.05$ for all comparisons of that distribution to others. **b**, Percentage of connections that contain a single versus multiple synapses (shown up to five) for each target cell type across predicted MET-types (for example, predicted MET-6 cells form more multi-synaptic contacts onto L5 ET targets than onto other cell types). **c**, Example cell from each predicted MET-type showing soma locations of postsynaptic targets. Somas are colour-coded to indicate the number of synapses that cell receives from the presynaptic cell (soma in cyan). All examples use the same scale. **d**, Histograms of inter-synaptic distances onto target types (distances calculated per postsynaptic target). MET-8, -4 and -6 ($n = 3$, 6 and 5, respectively).

In Fig. 3c, we provide examples of the soma locations of postsynaptic targets (colour-coded by the number of synapses received) from each MET-type. These plots illustrate the differences in the target location and range in synapse number per connection for each MET-type. MET-8 cells form multi-synaptic connections with L4 and L5 IT PCs, visualized by the cluster of orange–yellow somas directly above the presynaptic (cyan) cell soma. MET-4 cells target diffusely across most targets but form a few multi-synapse connections in L5. Finally, MET-6 cells target diffusely across most targets, but form multi-synaptic connections with L5 ET cells located just above the presynaptic cell soma (Fig. 3c).

Dendritic targeting of MCs onto L5 PCs

To determine wh tonization, and hierarchical d ether *Sst* MET-types form spatially clustered synapses onto a postsynaptic target, we quantified the inter-synapse distances of a given pre–postsynaptic pair. We find that most synapses are formed within 150 μm of each other (Euclidean distance) (Fig. 3d). However, both MET-4 and MET-6 cells have many synapses that are more than 300 μm apart. These distances may be due to synapses formed onto both the apical and basal dendrites of a target cell or across a wide lateral extent of basal dendrites.

One hypothesis regarding L5 MC connectivity is that they form synapses onto both the apical and basal dendrites of L5 excitatory neurons (see histograms in Fig. 2c and more than 300-μm distances in Fig. 3d) to coordinate inhibition across compartments of individual cells. We determined the location of synapses from MET-4 and MET-6 cells onto L5 IT and L5 ET targets (Extended Data Table 3). Both MET-4 and MET-6 cells form approximately half of their synapses onto apical dendrites (around 51% and around 47%, respectively). Approximately 42% of MET-4 connections are apical only, around 46% are non-apical and around 11% are onto both apical and non-apical (probably basal) dendrites. We find about 34% of MET-6 connections are apical only, around 50% are non-apical and around 16% are onto both apical and basal dendrites. Thus, the coordinated inhibition across compartments of L5 PCs by individual MCs occurs in only a small fraction of connections to postsynaptic targets; however, most connections are made onto a single compartment, and both apical and non-apical dendrites receive similar numbers of synapses.

Generalization of the morpho-classifier

After characterizing the ultrastructural differences that align with the predicted MC MET-types, we tested the generalizability of the classifier. We compared the predicted MET-subclass for inhibitory cells within the columnar sample to the 'targeting type' of the cells defined by Schneider-Mizell et al.[33]. The predicted subclasses largely align with the targeting type (Extended Data Fig. 2c). Most predicted *Sst* cells (64%) are distal dendrite targeting cells, most predicted *Pvalb* cells (78%) are perisomatic targeting cells, and *Vip* cells (91%) are inhibitory targeting cells. These patterns match with published connectivity patterns for these subclasses[9]. We observe some off-diagonal mapping,

although further work is needed to determine whether this is driven by methodology or biology. Patch-seq and EM cells grouped by predicted subclass and targeting type are shown in Extended Data Fig. 4.

Limitations

There are some technical artifacts that will have minor effects on the quantifications of connectivity we report. First, we focused our analysis on a small number of reconstructed EM cells. These cells are highlighted because they are comprehensively proofread and are within the same cortical region (VISp) as the Patch-seq data. It is possible that we have only sub-sampled MET-types and our findings are an incomplete description of the diversity within the cell type. Another possibility is that some MET-types are more intrinsically variable. The consistency of MET-type connectivity will need to be evaluated with further studies. The top 10 µm of the cortical surface is also not included in these reconstructions owing to segmentation errors. However, we find that synapses onto L5 targets peak at a lower depth compared with those onto L2/3 targets. We therefore do not expect either of these effects to be large enough to change the overall conclusions reported here.

There are potential differences in the morphological feature values that are driven by differences in the Patch-seq and EM experimental procedures. Theoretically, the classifier trained on Patch-seq data could be less accurate for EM data than it appears to be in cross-validation experiments applied to Patch-Seq. However, the differences in morphology we observe between MET-types, which we focus on here, are clear to a trained anatomist, and differences in connectivity are largely consistent from cell to cell. We take the consistency between the MET predictions and stark shifts in connectivity as support for true biological differences between these groups. It is also possible that closely related MET-types (for example, *Sst* MET-5 and *Sst* MET-6) can have similar targeting and that this modality can be used to further refine or collapse cell type(s). Thus, we believe that exploring the synaptic connectivity across all cortical MET-types is an exciting future application of this type of work.

Discussion

By establishing a morphological feature set aligned across EM and a multi-modal Patch-seq dataset, we can map cells between the two datasets to predict the connectivity of different types of inhibitory neuron. We can also predict the molecular identity and electrophysiological profile of neurons sampled in the large volume EM data. These predictions reveal that the MET-types analysed consistently differ in their myelination, synaptic features and target cell subclass connectivity profiles, which suggests that they play unique roles in cortical circuitry (Fig. 4). We believe this approach could be a powerful tool to leverage existing and future Patch-seq and EM datasets to predict cellular properties across methodologies.

Integration of MET-type and connectivity

Previously defined MET-types[16] have been shown to have distinct morphological, electrical and transcriptomic properties (Fig. 4a). We can now see how *Sst* MET-8, MET-4 and MET-6 target distinct excitatory neuron subclasses using different connectivity motifs (Fig. 4a). We can also generate hypotheses about what features measured in Patch-seq (for example, electrophysical properties, transcriptomic expression patterns, morphology) may correlate with the connectivity patterns observed for cells with similar morphology in EM (Fig. 4b). We show expression patterns of the highest differentially expressed genes between the principal *Sst* t-types that comprise the MET-types (Fig. 4c). These genes could play a role in setting up and/or maintaining the distinct morphology and connectivity of these MET-types. Several of the genes identified above are involved in cell adhesion (*Cdh12* (ref. 47), *Pcdh8* (ref. 48)) or both adhesion and inhibitory synapse formation (*Ptprd*)[49,50]. Calsyntenin-2 (encoded by *Clstn2*) is also found at GABAergic synapses, and knocking-out the gene leads to fewer parvalbumin neurons and inhibitory synapses in the hippocampus[51] and loss of inhibitory synapses in cortex[52]. Identifying subcellular protein expression patterns and loss-of-function studies of these proteins during development would shed light on their role in synaptic patterning in these MET-types.

Potential role of diverse myelination

Although myelination is frequently described for excitatory neurons, previous studies have shown that inhibitory neurons[36,53], including MCs[38], can be myelinated. Here we find significant differences in the number and total length of myelinated segments across *Sst* MET-types (Fig. 2e). Reduced myelination onto *Pvalb* neurons can reduce firing rates and conduction velocity[36,54]. Differences in the myelination of *Sst* cells may similarly influence their firing properties. For example, the myelination of the ascending axon stalk of MET-4 cells could help to synchronize inhibition of L2/3 and L5 IT PCs and/or increase the speed of inhibition onto L2/3 targets of MET-4 versus MET-6 cells (Fig. 2a,e). Oligodendrocytes—the glial cells that myelinate axons—have also been shown to metabolically support the axons they ensheath[55,56]. Thus, myelination onto MET-4 cells may indicate glial support of their larger output synapses (Fig. 2d). Our observations suggest that myelination may vary in a cell-type-specific manner and could provide another marker of cell-type identity in EM datasets.

Potential circuit roles of synapse size

We find significant differences in output synapse sizes across the observed MET-types (Fig. 2d). Correlated slice electrophysiology and EM performed on excitatory cells demonstrated a linear relationship between chemical synapse size and strength[39]. Thus, an individual MET-8 synapse could be weaker than an individual MET-4 or MET-6 synapse. However, we also find that MET-8 cells form significantly more output synapses, so a single synapse analysis might be under-counting their inhibitory influence on the circuit.

MCs can form several synapses onto a single target, so we calculated the average connection size of a cell (Extended Data Fig. 5c). We find that MET-8 cell connections are larger than those of MET-6 (not MET-4). It is possible that the greater number of synapses onto individual targets may be related, in part, to the smaller synapse size. The number and size of inhibitory synapses onto a target may balance the excitation impinging onto that cell and shape its response. In the chicken auditory nucleus, inhibition varies along the tonotopic axis, which shapes the timing and dynamic range of postsynaptic responses[57]. Thus, inhibitory synapse size may be a function of homeostatic plasticity and/or be suited to the features encoded by the postsynaptic targets of each MET-type.

Inhibitory connectivity rules

Each *Sst* MET-type uses a distinct inhibition pattern: inhibit cell types in proportion to preference, culminating in targeting nearly 100% of the neighbouring preferred cell type (MET-8); inhibit a small fraction of all target cell types (MET-4); or inhibit a small fraction of most target cell types except the preferred 'local' target (MET-6). These findings align well with paired recordings showing high connection probability from L4 *Sst* cells to L4 PCs and L5 *Sst* cells to L5 ET cells[58]. Our radial distance findings also align with studies from brain slices showing that most connections from *Sst* cells are onto excitatory targets within a 200-µm lateral distance[58,59].

Previous studies found dense connectivity from L2/3 *Sst* cells onto L2/3 PC targets, creating 'blanket inhibition' in the cortex[25], and predicted 'non-specific' connectivity onto most cell types[10]. We observe 'blanket inhibition' from the MET-8 cells onto L4 PCs, but not from MET-8 onto other target subclasses or from other *Sst* MET-types. Thus, 'blanket inhibition' is present in specific contexts, but it is not a universal

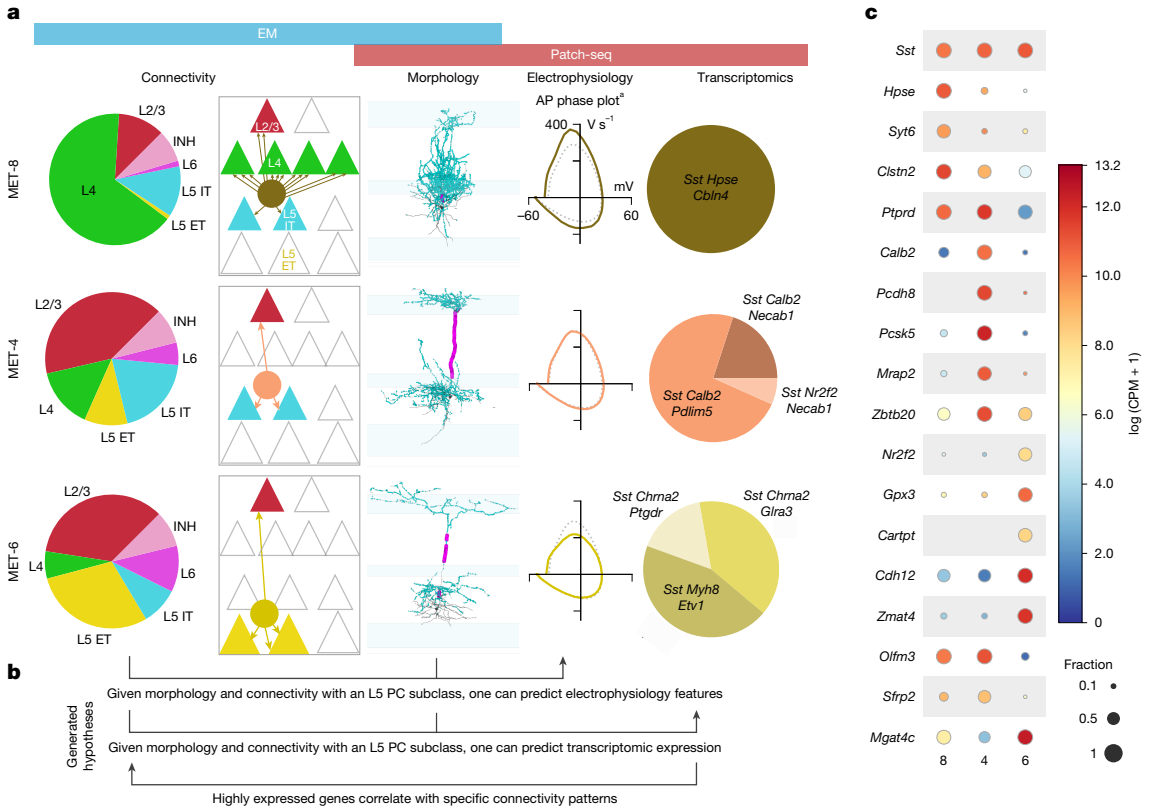


Fig. 4 | Integrated view of MET-types including modalities from Patch-seq and EM. a, First column, average percentage of synapses onto each postsynaptic target group. Second column, schematic summarizing the connectivity motifs observed from EM data. Third column, example cells from the EM dataset. Fourth column, average action potential (AP) traces per MET-type. Fifth column, transcriptomic cell types that comprise previously defined MET-types. **b**, This integrated view of MET-types now allows us to generate hypotheses such as the role of highly expressed genes in a given transcriptomic type on connectivity patterns. **c**, Dot plot showing the fraction of cells within the MET-type (circle size) and expression level (red to blue scale bar) of the genes that are differentially expressed across the transcriptomic types in MET-8, 4 and 6 (n = 18, 15 and 18, respectively) and present in at least 50% of one MET-type. Genes listed are the top five upregulated or downregulated genes by pairwise comparison (minus duplicates). [a]Data previously published[16]. Credits: plots in **a** reproduced with permission from ref. 16, Cell Press.

connectivity pattern for *Sst* cells, highlighting the diversity of connectivity motifs within this subclass.

Alignment with other studies

Previous studies have shown that *Hpse Cbln4* cells in mouse primary somatosensory and visual cortex synapse predominantly onto L4 and not L5 PCs[22,31]. Our EM results not only recapitulate the connectivity bias observed in mouse sensory cortices (L4 > L5), but also reveal the degree to which MET-8 cells synapse across cell types.

Previous studies also found that *Chrna2*$^+$ MCs in L5 of auditory cortex preferentially inhibit thick-tufted (probably L5 ET cells) but not thin-tufted (probably L5/L6 IT) cells[29], and optogenetic mapping of *Sst Myh8* cells (genetically targeted using *Chrna2*-cre) shows that they more strongly inhibit L5 ET than IT cells[32]. MET-6 cells, which include *Chrna2*$^+$ cells[16], form more synapses onto L5 ET than IT cells. Thus, the MET-6-type structural connectivity we characterized aligns with the probed physiology of *Chrna2*$^+$ MCs[29].

Recent work in mouse VISp finds that bulk optogenetic activation of virally labelled *Sst Calb2* cells (*Sst* MET-3 and *Sst* MET-4) produces larger amplitude iPSCs onto L5 ET than L5 IT cells[32], whereas we find that the MET-4 group forms more output synapses onto L5 IT cells and has similar connection probabilities to both populations. Thus differences in connection strength could be driven by differences in the convergence of connections or the output connectivity of MET-3 and MET-4 types, or could reflect that MET-4/*Sst Calb2* cells have functionally stronger individual synapses onto ET than IT cells. Predicting a molecular cell type for EM data enables the comparison of circuit and synaptic features to physiological features. Future targeted studies to examine the functional dynamics of individual inhibitory synapses onto distinct excitatory target populations would help to resolve observed differences such as this.

Finally, we have demonstrated that local morphological features can be used to link cell types across datasets. Linking these cell type identities enables the investigation of synaptic connectivity with respect to morphology, electrophysiology and transcriptomic expression in advance of the availability of technologies that allow the direct measurement of these properties from the same cell. As larger electron microscopy datasets are generated, this approach could be extended to cell types across the whole mouse brain and other species. Measuring the synaptic connectivity of transcriptomic cell types will facilitate future work to characterize the behaviour of these types, in local and brain-wide circuits, using genetic tools.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-025-08805-6.

1. Zeng, H. & Sanes, J. R. Neuronal cell-type classification: challenges, opportunities and the path forward. *Nat. Rev. Neurosci.* **18**, 530–546 (2017).
2. Cajal, S. R. Y., Azoulay, D. L., Swanson, N. & Swanson, L. W. *Histology Of The Nervous System: Of Man And Vertebrates* (Oxford Univ. Press, 1995).

3. De Carlos, J. A. & Borrell, J. A historical reflection of the contributions of Cajal and Golgi to the foundations of neuroscience. *Brain Res. Rev.* **55**, 8–16 (2007).
4. Hubel, D. H. & Wiesel, T. N. Receptive fields of single neurones in the cat's striate cortex. *J. Physiol.* **148**, 574–591 (1959).
5. Zeisel, A. et al. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science* **347**, 1138–1142 (2015).
6. Tasic, B. et al. Shared and distinct transcriptomic cell types across neocortical areas. *Nature* **563**, 72–78 (2018).
7. Somogyi, P., Tamás, G., Lujan, R. & Buhl, E. H. Salient features of synaptic organisation in the cerebral cortex. *Brain Res. Brain Res. Rev.* **26**, 113–135 (1998).
8. Kubota, Y. Untangling GABAergic wiring in the cortical microcircuit. *Curr. Opin. Neurobiol.* **26**, 7–14 (2014).
9. Tremblay, R., Lee, S. & Rudy, B. GABAergic interneurons in the neocortex: from cellular properties to circuits. *Neuron* **91**, 260–292 (2016).
10. Jiang, X. et al. Principles of connectivity among morphologically defined cell types in adult neocortex. *Science* **350**, aac9462 (2015).
11. Markram, H. et al. Reconstruction and simulation of neocortical microcircuitry. *Cell* **163**, 456–492 (2015).
12. Gouwens, N. W. et al. Classification of electrophysiological and morphological neuron types in the mouse visual cortex. *Nat. Neurosci.* **22**, 1182–1195 (2019).
13. Cadwell, C. R. et al. Multimodal profiling of single-cell morphology, electrophysiology, and gene expression using Patch-seq. *Nat. Protoc.* **12**, 2531–2553 (2017).
14. Földy, C. et al. Single-cell RNAseq reveals cell adhesion molecule profiles in electrophysiologically defined neurons. *Proc. Natl Acad. Sci. USA* **113**, E5222–E5231 (2016).
15. Fuzik, J. et al. Integration of electrophysiological recordings with single-cell RNA-seq data identifies neuronal subtypes. *Nat. Biotechnol.* **34**, 175–183 (2016).
16. Gouwens, N. W. et al. Integrated morphoelectric and transcriptomic classification of cortical GABAergic cells. *Cell* **183**, 935–953.e19 (2020).
17. Wahle, P. Differential regulation of substance P and somatostatin in Martinotti cells of the developing cat visual cortex. *J. Comp. Neurol.* **329**, 519–538 (1993).
18. Martinotti, C. Contributo allo studio della corteccia cerebrale, ed all'origine centrale dei nervi. *Ann. Freniatr. Sci. Affini* **1**, 14–381 (1889).
19. Wang, Y. et al. Anatomical, physiological and molecular properties of Martinotti cells in the somatosensory cortex of the juvenile rat. *J. Physiol.* **561**, 65–90 (2004).
20. Isaacson, J. S. & Scanziani, M. How inhibition shapes cortical activity. *Neuron* **72**, 231–243 (2011).
21. Lee, B. R. et al. Scaled, high fidelity electrophysiological, morphological, and transcriptomic cell characterization. *eLife* **10**, e65482 (2021).
22. Scala, F. et al. Layer 4 of mouse neocortex differs in cell types and circuit organization between sensory areas. *Nat. Commun.* **10**, 4174 (2019).
23. The MICrONS Consortium. Functional connectomics spanning multiple areas of mouse visual cortex. *Nature* https://doi.org/10.1038/s41586-025-08790-w (2025).
24. Meinecke, D. L. & Peters, A. Somatostatin immunoreactive neurons in rat visual cortex: a light and electron microscopic study. *J. Neurocytol.* **15**, 121–136 (1986).
25. Fino, E. & Yuste, R. Dense inhibitory connectivity in neocortex. *Neuron* **69**, 1188–1203 (2011).
26. Silberberg, G. & Markram, H. Disynaptic inhibition between neocortical pyramidal cells mediated by Martinotti cells. *Neuron* **53**, 735–746 (2007).
27. Berger, T. K., Perin, R., Silberberg, G. & Markram, H. Frequency-dependent disynaptic inhibition in the pyramidal network: a ubiquitous pathway in the developing rat neocortex. *J. Physiol.* **587**, 5411–5425 (2009).
28. Nigro, M. J., Hashikawa-Yamasaki, Y. & Rudy, B. Diversity and connectivity of layer 5 somatostatin-expressing interneurons in the mouse barrel cortex. *J. Neurosci.* **38**, 1622–1633 (2018).
29. Hilscher, M. M., Leão, R. N., Edwards, S. J., Leão, K. E. & Kullander, K. Chrna2–Martinotti cells synchronize layer 5 type A pyramidal cells via rebound excitation. *PLoS Biol.* **15**, e2001392 (2017).
30. Muñoz, W., Tremblay, R., Levenstein, D. & Rudy, B. Layer-specific modulation of neocortical dendritic inhibition during active wakefulness. *Science* **355**, 954–959 (2017).
31. Naka, A. et al. Complementary networks of cortical somatostatin interneurons enforce layer specific control. *eLife* **8**, e43696 (2019).
32. Wu, S. J. et al. Cortical somatostatin interneuron subtypes form cell-type-specific circuits. *Neuron* **111**, 2675–2692.e9 (2023).
33. Schneider-Mizell, C. M. et al. Inhibitory specificity from a connectomic census of mouse visual cortex. *Nature* https://doi.org/10.1038/s41586-024-07780-8 (2025).
34. DeFelipe, J. et al. New insights into the classification and nomenclature of cortical GABAergic interneurons. *Nat. Rev. Neurosci.* **14**, 202–216 (2013).
35. Waxman, S. G. Determinants of conduction velocity in myelinated nerve fibers. *Muscle Nerve* **3**, 141–150 (1980).
36. Micheva, K. D. et al. A large fraction of neocortical myelin ensheathes axons of local inhibitory neurons. *eLife* **5**, e15784 (2016).
37. Mazuir, E., Fricker, D. & Sol-Foulon, N. Neuron–oligodendrocyte communication in myelination of cortical GABAergic cells. *Life (Basel)* **11**, 216 (2021).
38. Zonouzi, M. et al. Individual oligodendrocytes show bias for inhibitory axons in the neocortex. *Cell Rep* **27**, 2799–2808.e3 (2019).
39. Holler, S., Köstinger, G., Martin, K. A. C., Schuhknecht, G. F. P. & Stratford, K. J. Structure and function of a neocortical synapse. *Nature* **591**, 111–116 (2021).
40. Wehr, M. & Zador, A. M. Balanced inhibition underlies tuning and sharpens spike timing in auditory cortex. *Nature* **426**, 442–446 (2003).
41. Saiki, A. et al. In vivo spiking dynamics of intra- and extratelencephalic projection neurons in rat motor cortex. *Cereb. Cortex* **28**, 1024–1038 (2018).
42. Shepherd, G. M. G. Corticostriatal connectivity and its role in disease. *Nat. Rev. Neurosci.* **14**, 278–291 (2013).
43. Murayama, M. et al. Dendritic encoding of sensory stimuli controlled by deep cortical interneurons. *Nature* **457**, 1137–1141 (2009).
44. Zhou, X., Mansori, I., Fischer, T., Witte, M. & Staiger, J. F. Characterizing the morphology of somatostatin-expressing interneurons and their synaptic innervation pattern in the barrel cortex of the GFP-expressing inhibitory neurons mouse. *J. Comp. Neurol.* **528**, 244–260 (2020).
45. Packer, A. M., McConnell, D. J., Fino, E. & Yuste, R. Axo-dendritic overlap and laminar projection can explain interneuron connectivity to pyramidal cells. *Cereb. Cortex* **23**, 2790–2802 (2013).
46. Urban-Ciecko, J. & Barth, A. L. Somatostatin-expressing neurons in cortical networks. *Nat. Rev. Neurosci.* **17**, 401–409 (2016).
47. Guo, B. et al. Cadherin-12 regulates neurite outgrowth through the PKA/Rac1/Cdc42 pathway in cortical neurons. *Front. Cell Dev. Biol.* **9**, 768970 (2021).
48. Pancho, A., Aerts, T., Mitsogiannis, M. D. & Seuntjens, E. Protocadherins at the crossroad of signaling pathways. *Front. Mol. Neurosci.* **13**, 117 (2020).
49. Uhl, G. R. & Martinez, M. J. PTPRD: neurobiology, genetics, and initial pharmacology of a pleiotropic contributor to brain phenotypes. *Ann. N. Y. Acad. Sci.* **1451**, 112–129 (2019).
50. Takahashi, H. et al. Selective control of inhibitory synapse development by Slitrk3-PTPδ trans-synaptic interaction. *Nat. Neurosci.* **15**, 389–398 (2012).
51. Lipina, T. V. et al. Cognitive deficits in calsyntenin-2-deficient mice associated with reduced GABAergic transmission. *Neuropsychopharmacology* **41**, 802–810 (2016).
52. Ranneva, S. V., Maksimov, V. F., Korostyshevskaja, I. M. & Lipina, T. V. Lack of synaptic protein, calsyntenin-2, impairs morphology of synaptic complexes in mice. *Synapse* **74**, e22132 (2020).
53. Call, C. L. & Bergles, D. E. Cortical neurons exhibit diverse myelination patterns that scale between mouse brain regions and regenerate after demyelination. *Nat. Commun.* **12**, 4767 (2021).
54. Benamer, N., Vidal, M., Balia, M. & Angulo, M. C. Myelination of parvalbumin interneurons shapes the function of cortical sensory inhibitory circuits. *Nat. Commun.* **11**, 5151 (2020).
55. Lee, Y. et al. Oligodendroglia metabolically support axons and contribute to neurodegeneration. *Nature* **487**, 443–448 (2012).
56. Fünfschilling, U. et al. Glycolytic oligodendrocytes maintain myelin and long-term axonal integrity. *Nature* **485**, 517–521 (2012).
57. Al-Yaari, M., Yamada, R. & Kuba, H. Excitatory–inhibitory synaptic coupling in avian nucleus magnocellularis. *J. Neurosci.* **40**, 619–631 (2020).
58. Campagnola, L. et al. Local connectivity and synaptic dynamics in mouse and human neocortex. *Science* **375**, eabj5861 (2022).
59. Hage, T. A. et al. Synaptic connectivity to L2/3 of primary visual cortex measured by two-photon optogenetic stimulation. *eLife* **11**, e71103 (2022).
60. Elabbady, L. et al. Perisomatic ultrastructure efficiently classifies cells in mouse cortex. *Nature* https://doi.org/10.1038/s41586-024-07765-7 (2025).

Methods

EM dataset generation and image alignment

The EM dataset described above is from tissue collected from the visual cortex (including primary and higher-order visual areas) dissected from a P87 (postnatal day) male SLC17a7-CrexAi162 heterozygous mouse[23]. In brief, a large-scale serial-section EM dataset was collected and imaged using automated transmission electron microscopes[61]. The data above are from a sub-volume representing 65% of the original EM volume with images of approximately 4 × 4 × 40 nm per pixel resolution. These images were segmented into meshes using convolutional neural networks and subsequent agglomeration[62]. The EM images and meshes are visualized in Neuroglancer. These meshes can be proofread (merged or split) within the ChunkedGraph system[33,63] Neuroglancer framework to facilitate proofreading of cells.

Correcting and generating representations of cells

Meshes underwent skeletonization (skeleton originated from a defined soma point) to generate a list of branch and end points for each mesh, visible in Neuroglancer[33]. Each branch point was inspected manually. True branch points were left alone and false branch points (often due to overlapping processes from distinct cells) were split using Neuroglancer tools. Subsequently, each end point was inspected manually. True endpoints were left alone and false endpoints (premature end of a process) were extended by an expert annotator, who would follow the process along the EM imagery to a natural ending (bouton, tapered end) or until the process could no longer be extended reliably (for example, edge of block). The total number of EM reconstructions of inhibitory cells used in this study is 173 (including clean and comprehensive reconstructions).

Morphological analysis and MET-type prediction

Determining cortical layers. Soma position, pia, white matter and laminar borders were drawn manually. For Patch-seq cells (data previously collected) a ×20 brightfield and fluorescent image of 4′,6-diamidino-2-phenylindole-stained tissue was used[16] and for EM cells, an expert annotator drew layer boundaries based on soma density near a *z* plane that contained the soma of the cell of interest. Drawings were made in Neuroglancer on a single EM *z* plane for the 16 MCs. Cortical layers were drawn at a central *z* plane of the cortical column and were applied to all cells from the column that were not the curated MCs. These polygons were then exported to be used for feature calculation.

Morphological feature calculation. Morphological features were calculated as previously described[12], using features that were derived from previous studies[11,64]. Features were calculated using the skeleton keys Python package (https://github.com/AllenInstitute/skeleton_keys). First, neuron skeletons were extracted, resampled (at 1,550), and exported from the EM dataset using the pcg_skel python package (https://github.com/CAVEconnectome/pcg_skel). Features were extracted from neurons aligned in the direction perpendicular to pia and white matter. Laminar axon histograms (bin size of 5 µm) and earth movers' distance features require a layer-aligned version of the morphology where node depths are registered to an average laminar depth template. These features include measurements such as total axon length, maximum path distance and total number of branches for both axon and dendrites from skeletonized representations.

RFC details. An RFC, support vector machine and logistic regression model were assessed for performance in predicting MET-type labels for Patch-seq cells using the morphological features of inhibitory cell types from a previously published Patch-seq dataset (Patch-seq $n = 477$, *Sst* $n = 236$, *Pvalb* $n = 89$, *Vip* $n = 79$, *Lamp5* $n = 44$, *Sncg* $n = 29$)[16]. Specifically, we used a stratified fivefold cross-validation approach to assess classification accuracy. The data were split randomly into five partitions while maintaining the distribution of MET-type labels in each partition. This method iteratively rotates which partition is withheld from training and used to validate the model. MET-types with fewer than five morphological reconstructions (*Lamp5* MET-2, *Pvalb* MET-5, *Sncg* MET-2, *Sncg* MET-3, *Sst* MET-11 and *Vip* MET-3) were omitted. An RFC with 250 estimators, a maximum depth of ten, balanced class weights, a minimum of ten samples per split and at least five samples per leaf node outperformed logistic regression and support vector machine models. Fivefold stratified cross-validation with shuffling was repeated 20 times and achieved a mean accuracy of 58.9 ± 4.1% (s.e.m.), far exceeding the expected chance accuracy for 22 categories (4.5%). Classifier accuracy was determined by how frequently the model correctly predicted the MET-type label of held out data (not used in training the model) (Fig. 1c and Extended Data Fig. 5a). We calculated an overall F_1 score of 0.58, based on averaging F_1 scores for each MET-type, based on classifying that type versus any other type. The cumulative confusion matrix for hold-out validation data was recorded in Fig. 1c.

EM MET-type inference. For 500 iterations, a random subsample (95%) of the Patch-seq data was selected with probabilities according to MET-type class size (a Patch-seq cell from a well-represented MET-type was more likely to be omitted). MET-types with five or fewer specimens were exempt from subsampling. In each iteration, a new RFC with the aforementioned parameters was fitted with sub-sampled Patch-seq data and MET labels were predicted for EM cells. The final MET assignment was given as the most frequently predicted MET label for each cell (Extended Data Table 1). We used these predicted MET-type labels to group cells for subsequent analysis.

Calculating a reliability threshold. A reliability metric was quantified as the fraction of iterations each sample was predicted as its final MET assignment out of all predictions (for example, a cell was predicted into this MET-type 80% of the time). To determine an appropriate threshold, we explored reliability scores in the Patch-seq data. We applied random subsampling iterations in a leave-one-out manner to the Patch-seq data. Here we set one Patch-seq sample aside and use the rest as training data. In each iteration, the training data are sub-sampled randomly as described above and used to fit a new RFC. The classifier predicts the MET-type label of the single left-out Patch-seq sample. This process was repeated 500 times for every Patch-seq cell until each had a predicted MET-type label and corresponding reliability metric. We then plotted a cumulative histogram of the correctly and incorrectly predicted labels versus the reliability score. We found that a reliability score of more than 0.54 is the most inclusive value at which Patch-seq samples are more frequently predicted correctly than incorrectly (Extended Data Fig. 2b).

Morphological feature *z* score analysis

The *z* scores for Patch-seq and EM morphological features were calculated; *z* scores for EM morphological features were derived using the mean and s.d. of the Patch-seq data.

Poisson distribution analysis

We fit a simple Poisson model to match the total number of connections and mean number of synapses per connection for a given cell type distribution. We used a chi-squared test to probe whether the measured distributions were significantly different from the Poisson model. Specifically, the total number and distribution of output synapses onto target cell types were calculated for each cell. We calculated the number of single-, double- and triple-synapse (up to 17-synapse) containing connections for each presynaptic MC onto each target cell type (L2/3, L4, L5 ET, and so on), then built a Poisson distribution for each presynaptic cell and postsynaptic target cell type using the total number of connections and the average number of synapses per connection. We then compared the real and Poisson distributions using

a Chi-squared test (degrees of freedom = 0) to determine whether we could reject the null hypothesis (that the distributions were the same, $P < 0.05$).

Subclass prediction and cell targeting for EM column data

A MET-type prediction was made for each inhibitory neuron from a defined 100 µm × 100 µm cortical column (these cells are analysed in greater detail in ref. 33). Given the diverse population of interneurons captured in these data and the *Sst*-centric focus of this paper, the data are presented at the subclass level (apart from Extended Data Fig. 4, which shows Patch-seq and EM morphologies mapped to *Sst* MET-types or non-*Sst* subclasses). Subclass labels were assigned based on MET-type classifications (for example, all *Pvalb* MET-types are categorized under the *Pvalb* subclass).

Connectivity based inhibitory subclasses were defined by Schneider-Mizell et al.[33]. In brief, the dendritic arbour of excitatory neurons was assigned to different compartments: soma, proximal (less than 50 µm from the soma), apical and distal basal dendrite. Inhibitory cells were considered a fifth compartment. Synapses were then assigned a compartment based on where they were located on the postsynaptic target. The inhibitory subclass of a cell was then defined by where most of its synapses were located: perisomatic targeting cells that form synapses with the soma or proximal dendrites of postsynaptic cells, distal dendrite targeting cells synapse onto distal and/or apical dendrites, sparsely targeting cells do not form many multi-synaptic connections and inhibitory targeting cells predominantly synapse onto inhibitory cells.

Identifying synapses and postsynaptic targets

Synapses and their pre- and postsynaptic meshes in the EM dataset were previously algorithmically detected[65]. These data also included the automatically detected synapse size (number of voxels per synapse)[66]. In brief, a convolutional network was trained to identify synaptic clefts and assign voxels to each side of the cleft. Inference on the MICrONS volume was processed using the methods described in Wu et al.[65] using $8 \times 8 \times 40$ nm^3 images. Subsequently, a second convolutional network was trained to perform pre- and postsynaptic partner assignment for the voxels for each detected synaptic cleft[66]. At the dataset level, manual inspection of the algorithmically identified synapses across subvolumes provides an estimated precision of 97% (correctly identifying a synapse) and recall of 89% (identifying all synapses)[23].

To confirm that these synapse detection rates apply for the cells in this study specifically, we manually inspected 5% of output synapses per cell (3,180 synapses across 11 of 16 MCs) to confirm the presence of true synapses and to determine the postsynaptic target cell identity. We found that overall 4% of synapses were false positives, which aligns with the estimate for the entire volume. We did observe some differences across individual cells (5.6 ± 4.3% s.d. false positives per cell). To determine a cell's average connection size, we calculated the sum of synapse sizes onto each individual postsynaptic target and then divided by the number of postsynaptic targets.

Cell subclass identities were assigned to all meshes with single somas (individual cells) in the EM dataset using a support vector machine classifier trained on somatic and nuclear features[60]. This classifier was then applied across the EM dataset to generate predicted cell type identities for most cells. We use these assigned types in all plots shown above. There was broad agreement between the automated and manual cell typing except for a specific disagreement of L2/3 versus L4 identity for targets of the predicted MET-8 cell type due to differences in layer boundaries used by manual versus automated methods (Extended Data Fig. 5b). Labels shown in the figures are from automated assignment.

Most synapses from individual cells were onto postsynaptic targets that contained single somas in the reconstruction (Extended Data Table 1). Of the synapses that were not onto single-soma targets, most were 'orphans', referring to objects that cannot be connected to any somas, either because the soma of the cell is outside the volume of the dataset or because the dendrite or dendritic compartment was not connected to a soma.

Quantification of myelination

Using automatically detected synapses, annotators visualized all output synapses on a given presynaptic cell in Neuroglancer. Regions lacking synapses were manually inspected in the EM imagery. If myelination was seen, an annotator marked the start and end point of each myelinated segment in Neuroglancer to generate a line. The number of these annotations was summed to determine the number of myelinated segments per cell. The length of each annotation was summed to determine the distance of myelinated axon per cell.

Apical synapse location analysis

The details of defining an apical dendrite for excitatory neurons in the EM dataset are described in ref. 33. In brief, the apical dendrite of skeletonized excitatory neurons was identified algorithmically by an RFC trained on expert annotation data. The RFC predicts whether a skeletal vertex is apical from features of the vertex such as dendritic length (for example, distance of vertex from soma along the skeleton) and complexity (for example, number of branch points). The apical label is then associated with synapses near that vertex. For our analysis, we queried the output synapses of each MET-4 and MET-6 MC. We restricted our analysis to the subset of synapses that were made onto L5 excitatory targets (algorithmically defined as L5 ET or L5 IT by Elabbady et al.[60]) that had identified apical dendrites. We calculated the percentage of output synapses that each MET-4 and MET-6 cell formed onto the apical dendrites versus non-apical locations.

MET-type proportions derived from MERFISH data

Proportions of MET-types were estimated from a recently published MERFISH dataset[67]. Cell counts were calculated for each *Sst* t-type in VISp in the MERFISH dataset. As the cells in the MERFISH dataset were mapped to a whole-brain taxonomy[67], whereas the MET-types were based on a VISp-specific taxonomy[6], we identified correspondences between t-types across several taxonomies. Correspondences between the original taxonomy of Tasic et al.[6] and the cortex or hippocampal formation (CTX/HPF) taxonomy were identified by finding the Tasic et al.[6] t-types that had the highest number of shared cells for each CTX/HPF t-type[68]. Correspondences between the CTX/HPF and whole-brain taxonomy were taken directly from ref. 67. We assigned Tasic et al.[6] t-types to the MET-types to which the largest number of cells belonged[16], except for the t-type *Sst Calb2 Pdlim5*, for which corresponding cells were assigned to either *Sst* MET-3 (if located in L4 or above) or *Sst* MET-4 (if located in L2/3). Note that cells from the *Sst Chodl* subclass (*Sst* MET-1) were not analysed here. In addition, no clear t-type correspondences were identified for the MET-type *Sst* MET-11, which also contained a small number of cells in the original study[16].

Gene expression analysis

We analysed differentially expressed genes using scrattch.hicat (https://github.com/AllenInstitute/scrattch.hicat/tree/master) on the reference fluorescence-activated cell sorting-collected single-cell RNA sequencing dataset[6] to compare the main transcriptomic types that make up each MET-type (MET-8: *Sst Hpse Cbln4*; MET-4: *Sst Calb2 Necab1* and *Sst Calb2 Pdlim5*; MET-6: *Sst Chrna2 Glra3*, *Sst Chrna2 Ptdg4* and *Sst Myh8 Etv1*). Pairwise differentially expressed genes were identified as previously described[6] using the limma package[69] and selecting genes with at least a twofold change in expression and an adjusted *P* value of less than 0.01. The top five upregulated and downregulated genes for each pairwise comparison were selected for visualization (ranked by adjusted *P* value). The average expression of these genes and the fraction of cells with non-zero expression was calculated for the neurons of the three *Sst* MET-types in the Patch-seq data[16] and presented as a

dot plot. Only genes expressed in at least 50% of cells in at least one MET-type were selected for visualization.

Statistics

Comparisons across several MET-types were performed using non-parametric Kruskal–Wallis tests followed by Conover post hoc tests with Bonferroni corrections for pairwise comparisons. *P* values are indicated on plots if the Kruskal–Wallis *P* values < 0.05 and post hoc tests for pairwise comparisons are $P < 0.05$ or $P < 0.01$. Errors reported are s.e.m. unless otherwise indicated. Comparisons of the fraction of cell types targeted by a MET-type (from 0 to 300 µm) (Fig. 3a) were performed using a non-parametric Kolmogorov–Smirnov test with *P* values adjusted by a false discovery rate (Benjamini–Hochberg) correction. Boxplot whiskers indicate the range (maximum/minimum) of data; $n = 173$ inhibitory neurons from the EM dataset; $n = 16$ extended MCs and, of those, $n = 3$ were predicted to be MET-8, $n = 6$ were predicted to be MET-4 and $n = 5$ were predicted to be MET-6. Measurements are taken repeatedly from these cells.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The mm^3 EM dataset is publicly available at: https://www.microns-explorer.org/cortical-mm3. The Patch-seq gene expression data 'Mouse Patch-seq VIS–RNAseq' are publicly available at: https://knowledge.brain-map.org/data/1HEYEW7GMUKWIQW37BO/summary and the associated cell metadata are available at https://portal.brain-map.org/cell-types/classes/multimodal-characterization.

Code availability

We analysed data for this paper using v.661 of the dataset, a snapshot taken on 9 September 2023 at 4:00 a.m. UTC. Analysis code is publicly available at Allen Institute GitHub repository (https://github.com/AllenInstitute/EM_MC_paper_figures). Analysis used Python v.3.10.09 and made extensive use of the following packages and libraries: CAVEclient (https://github.com/seung-lab/CAVEclient), CloudVolume[70], MeshParty[71] and pcg_skel[72] (to extract neuronal skeletons from EM), skeleton_keys (https://github.com/AllenInstitute/skeleton_keys) (to extract and analyse morphological features from skeletons). We used the following libraries for visualization and analysis: Matplotlib[73], Seaborn[74], Numpy[75], Pandas[76], VTK[77], Scipy[78], Scikit-posthocs[79], Scikit-learn[80], scrattch-hicat (https://github.com/AllenInstitute/scrattch.hicat/) and statsmodels[81].

61. Yin, W. et al. A petascale automated imaging pipeline for mapping neuronal circuits with high-throughput transmission electron microscopy. *Nat. Commun.* **11**, 4949 (2020).
62. Macrina, T. et al. Petascale neural circuit reconstruction: automated methods. Preprint at *bioRxiv* https://doi.org/10.1101/2021.08.04.455162 (2021).
63. Dorkenwald, S. et al. FlyWire: online community for whole-brain connectomics. *Nat. Methods* **19**, 119–128 (2022).
64. Scorcioni, R., Polavaram, S. & Ascoli, G. A. L-Measure: a web-accessible tool for the analysis, comparison and search of digital reconstructions of neuronal morphologies. *Nat. Protoc.* **3**, 866–876 (2008).
65. Wu, J., Silversmith, W. M., Lee, K. & Seung, H. S. Chunkflow: hybrid cloud processing of large 3D images by convolutional nets. *Nat. Methods* **18**, 328–330 (2021).
66. Turner, N. L. et al. Synaptic partner assignment using attentional voxel association networks. In *Proc. 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)* 1–5 (IEEE, 2020).
67. Yao, Z. et al. A high-resolution transcriptomic and spatial atlas of cell types in the whole mouse brain. *Nature* **624**, 317–332 (2023).
68. Yao, Z. et al. A taxonomy of transcriptomic cell types across the isocortex and hippocampal formation. *Cell* **184**, 3222–3241.e26 (2021).
69. Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
70. Silversmith, W. et al. Igneous: Distributed dense 3D segmentation meshing, neuron skeletonization, and hierarchical downsampling. *Front. Neural Circuits* **16**, 977700 (2022).
71. Dorkenwald, S., Schneider-Mizell & Collman, F. F. sdorkenw/MeshParty. *GitHub* https://github.com/CAVEconnectome/MeshParty (2020).
72. pcg-skel: skeletonization using the pychunkedgraph. *GitHub* https://github.com/CAVEconnectome/pcg_skel (2025).
73. Hunter, J. D. Matplotlib: a 2D graphics environment. *Comput. Sci. Eng.* **9**, 90–95 (2007).
74. Waskom, M. L. seaborn: statistical data visualization. *J. Open Source Softw.* **6**, 3021 (2021).
75. Harris, C. R. et al. Array programming with NumPy. *Nature* **585**, 357–362 (2020).
76. The pandas development team. pandas-dev/pandas: Pandas. *GitHub* https://github.com/pandas-dev/pandas (2021).
77. Schroeder, W., Martin, K. & Lorensen, B. *Visualization Toolkit: an Object-Oriented Approach to 3D Graphics*, 4th edn. (Kitware Inc., 2006).
78. Virtanen, P. et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).
79. Terpilowski, M. A. scikit-posthocs: pairwise multiple comparison tests in Python. *J. Open Source Softw.* **4**, 1169 (2019).
80. Pedregosa, F. et al. Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
81. Seabold, S. & Perktold, J. Statsmodels: econometric and statistical modeling with python. In *Proc. of the 9th Python in Science Conference* (2010).

Acknowledgements We thank A. Bhandiwad, Y. Browning, T. Chartrand, T. Hage, J. Miller, J. Schardt and S. Seeman for their constructive feedback and conversations. We thank A. Jones and R. Costa for leadership and guidance. We wish to thank Allen Institute founders, P. G. Allen and J. Allen, for their vision, encouragement and support. The work was supported by the Intelligence Advanced Research Projects Activity (IARPA) via Department of Interior/Interior Business Center (DoI/IBC) contract numbers D16PC00003, D16PC00004 and D16PC00005. The US Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon. NMdC, FC and RCR also acknowledge support from National Institutes of Health (NIH) RF1MH125932 and from NSF NeuroNex 2 award 2014862. The research was also supported by several grant awards from institutes under the NIH, including award numbers R01EY023173 from The National Eye Institute and U01MH105982 from the National Institute of Mental Health and Eunice Kennedy Shriver National Institute of Child Health and Human Development to H.Z. Disclaimer: The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA, DoI/IBC, or the U.S. Government. This work was also funded by the Allen Institute for Brain Science.

Author contributions C.R.G., F.C., N.M.d.C. and S.A.S. conceived the project. C.M.S.-M., L.E., A.L.B., D.B., J.B., D.J.B., D.K., S.K., G.M., M.T., R.T., W.Y., J.A.B., M.A.C., S.D., A.H., Z.J., C.J., N.K., K. Lee, K. Li, R.L., T.M., E.M., S.S.M., S.M., B.N., S.P., W.S., N.L.T., W.W., J.W., S.Y., H.S.S., R.C.R., F.C. and N.M.d.C. contributed to the design, collection, alignment, segmentation and development of software to analyse the electron microscopy dataset. J.B., N.G., T.J., B.L., P.R.N., S.A.S. and H.Z. contributed to the design and J.B., N.G., T.J., B.L., S.A.S. and H.Z. contributed to the leadership on the generation of the Patch-seq dataset. C.R.G., C.M.S.-M., R.D., G.W. and A.L.B. contributed to data curation. A.M., E.J., M.M. and S.S. contributed code for analysis. C.R.G., M.M., N.G. and F.C. contributed to data analysis. C.R.G. wrote the manuscript. C.R.G., C.M.S.-M., R.D., T.J., H.Z., F.C., N.M.d.C. and S.A.S. revised the manuscript.

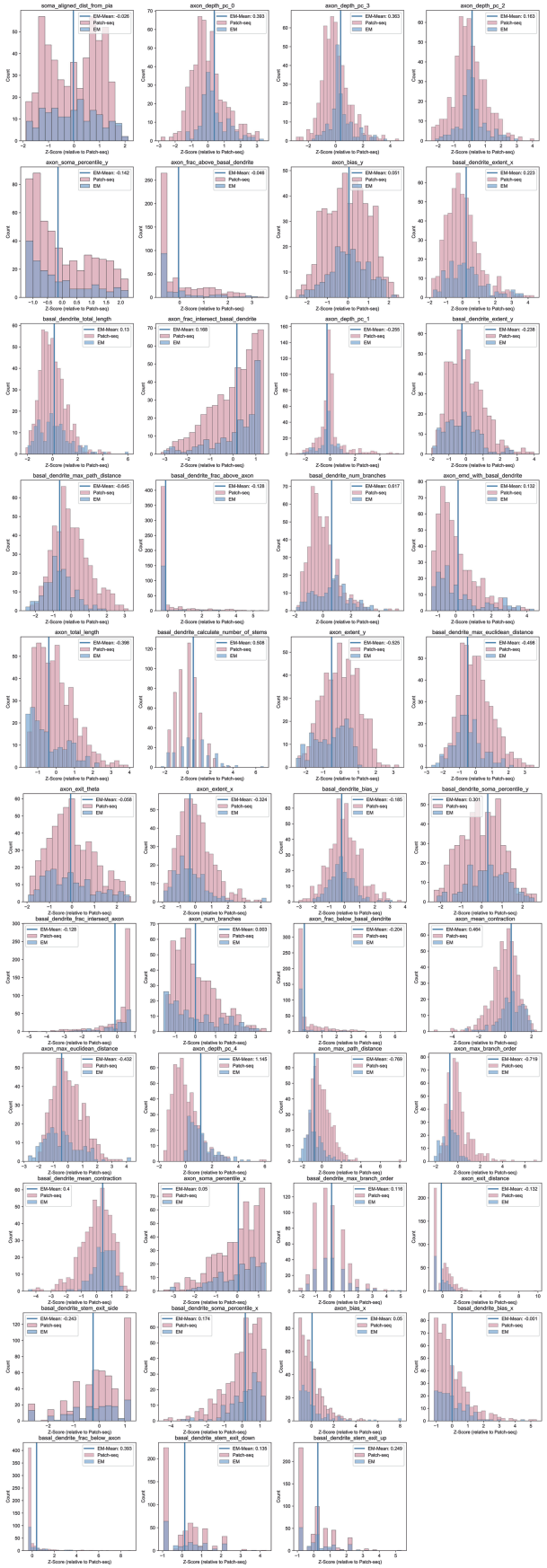
Competing interests N.K., K. Lee, T.M. and H.S.S. disclose financial interests in Zetta AI LLC. The remaining authors declare no competing interests.

Additional information

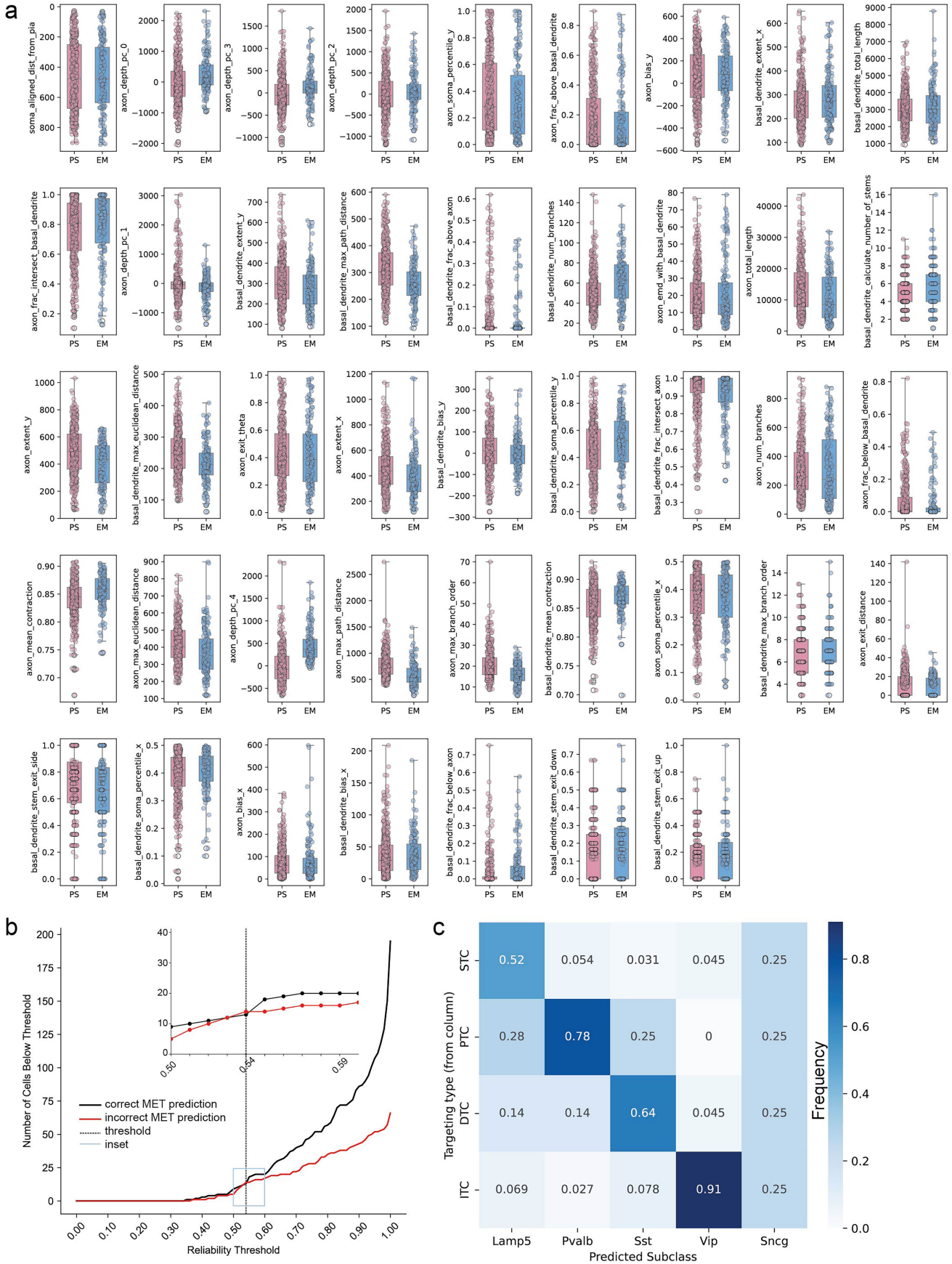
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-025-08805-6.

Correspondence and requests for materials should be addressed to Forrest Collman, Nuno Maçarico da Costa or Staci A. Sorensen.

Peer review information *Nature* thanks Matthew Grubb, Kristina Micheva and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

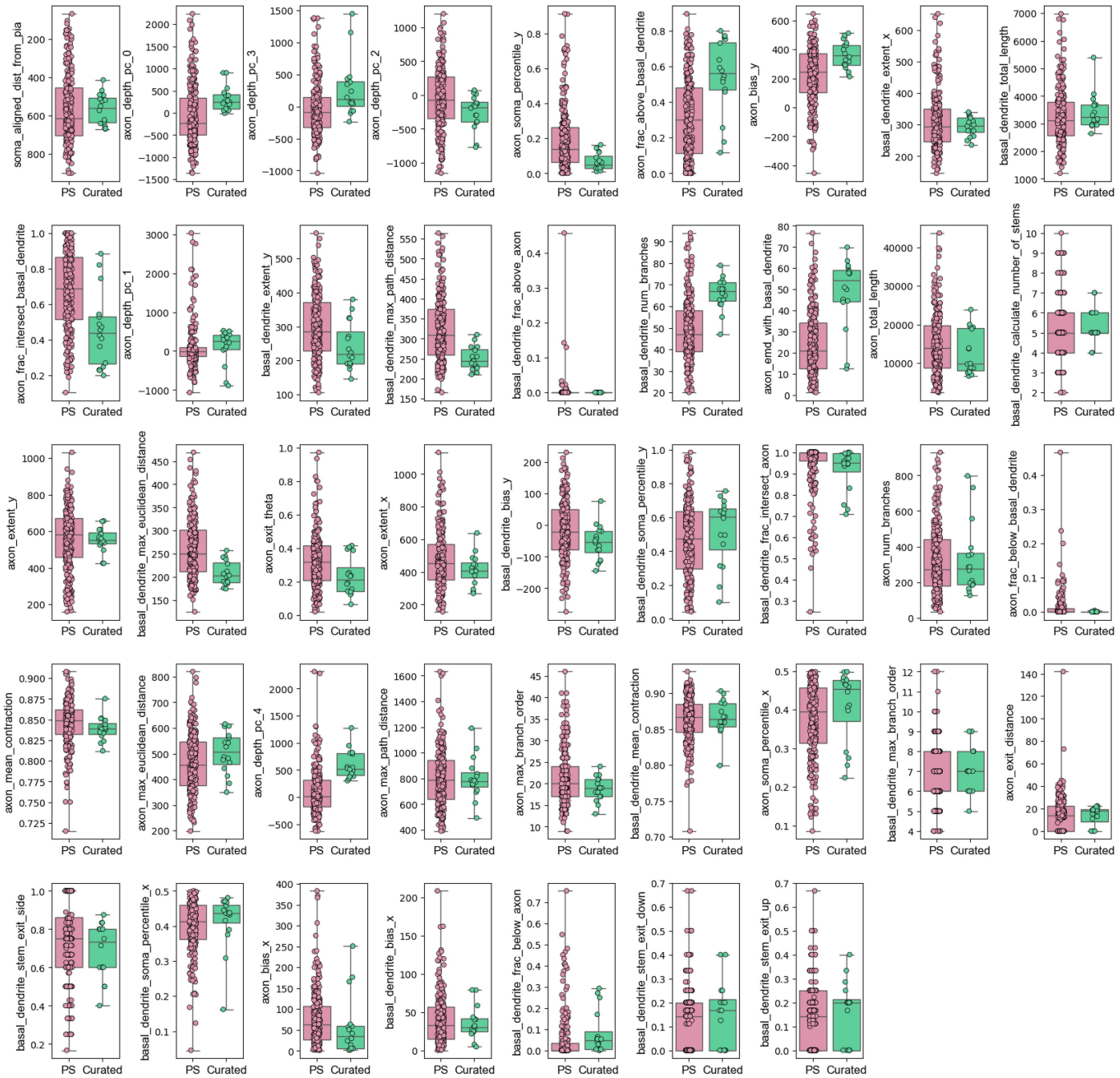


Extended Data Fig. 1 | Patch-seq versus EM feature Z-score comparison. Z-score distributions for Patch-seq (pink) versus EM (blue) data and the EM distribution mean (blue line) for all morphological features calculated (organized in decreasing Gini index score). We find that only one feature for which the EM mean is greater than 1 standard deviation away from the Patch-seq mean (axon_depth_pc_04). Patch-seq cell (n = 477); EM (n = 173).

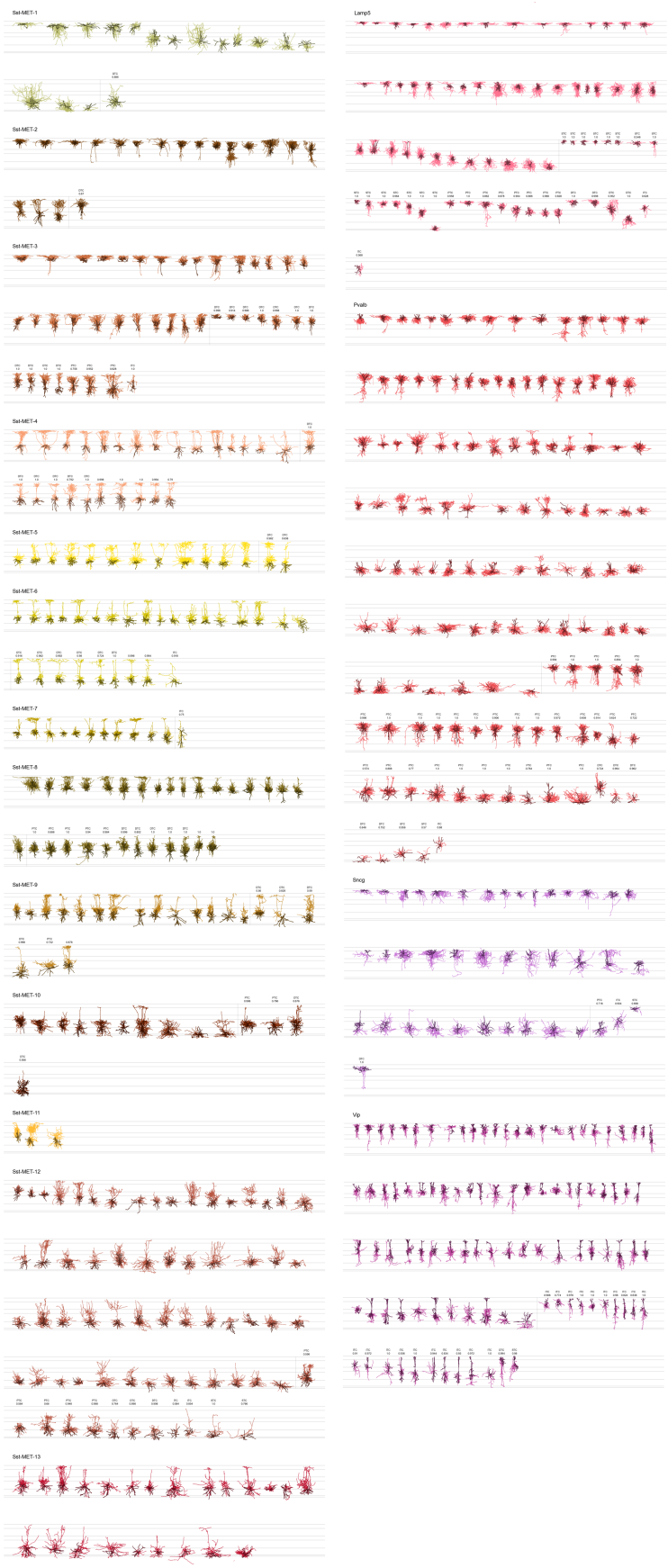


Extended Data Fig. 2 | See next page for caption.

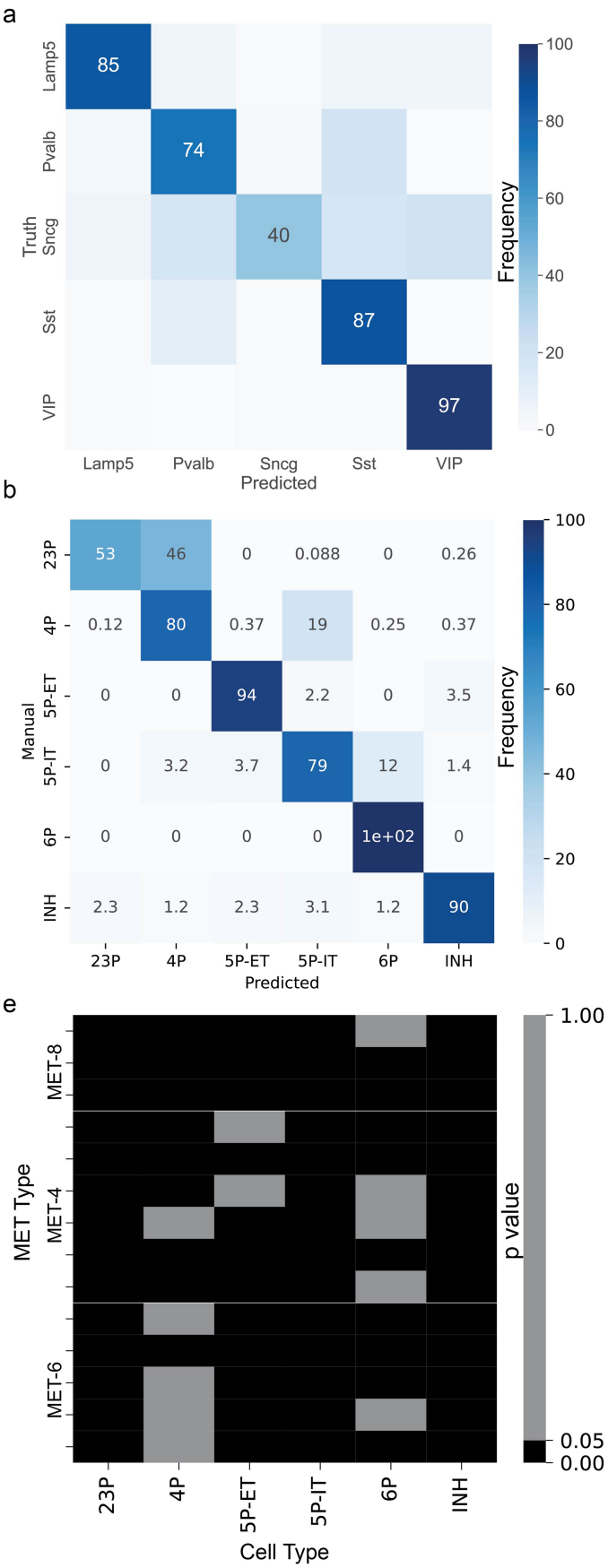
Extended Data Fig. 2 | Comparison of EM to Patch-seq features, classifier performance, and predicted MET-subclass relative to connectivity-defined subclass. a) Features of all inhibitory Patch-seq (PS) (n = 477) versus all inhibitory EM neurons (columnar sample + curated MCs n = 173). Boxplot whiskers indicate the range (max/min) of the data. Features are presented by decreasing Gini index as used by the classifier to predict MET-type identity. b) Cumulative histogram for reliability metric. Patch-seq cells which are correctly predicted with their known MET-type are labeled as "Correct MET prediction" and Patch-seq cells which the classifier incorrectly identifies are labeled as "Incorrect MET prediction". An inset is provided to clearly show where the reliability threshold is defined. c) Confusion matrix for connectivity-defined subclass (labels from Schneider-Mizell.[33]) versus predicted MET subclass for EM cells from the columnar sample (n = 163) (frequencies are normalized within subclass). Connectivity-defined targeting labels: Perisomatic targeting cells (PTC) target soma or proximal dendrites, distal dendrite targeting cells (DTC) primarily target apical and/or distal dendrites, sparsely targeting cells (STC) have relatively sparse multisynaptic connections, and inhibitory targeting cells (ITC) target other inhibitory neurons. These subclass definitions correspond to "coarse classical or molecular subclasses…but there is not a one-to-one match"[33]. Here they provide a useful benchmark for molecular subclass predictions derived from the Patch-seq classifier. For perspective on the *Sst* MET-type mapping of column neurons, morphologies are presented by predicted MET-type in Extended Data Fig. 4.



Extended Data Fig. 3 | Comparison of features of *Sst* Patch-seq versus EM Martinotti. Features plotted in order of decreasing gini index of features of *Sst* Patch-seq (n = 236) versus EM Martinotti cells (curated EM n = 16) features. Boxplot whiskers indicate the range (max/min) of the data. We see for all but one EM cell for one feature (axon_depth_pc_3) the values from EM fall within the range observed for Patch-seq data. We expect these values to not be the same as the Patch-seq data contains all *Sst* MET-types, whereas the curated EM cells are only a subset of Martinotti cells.

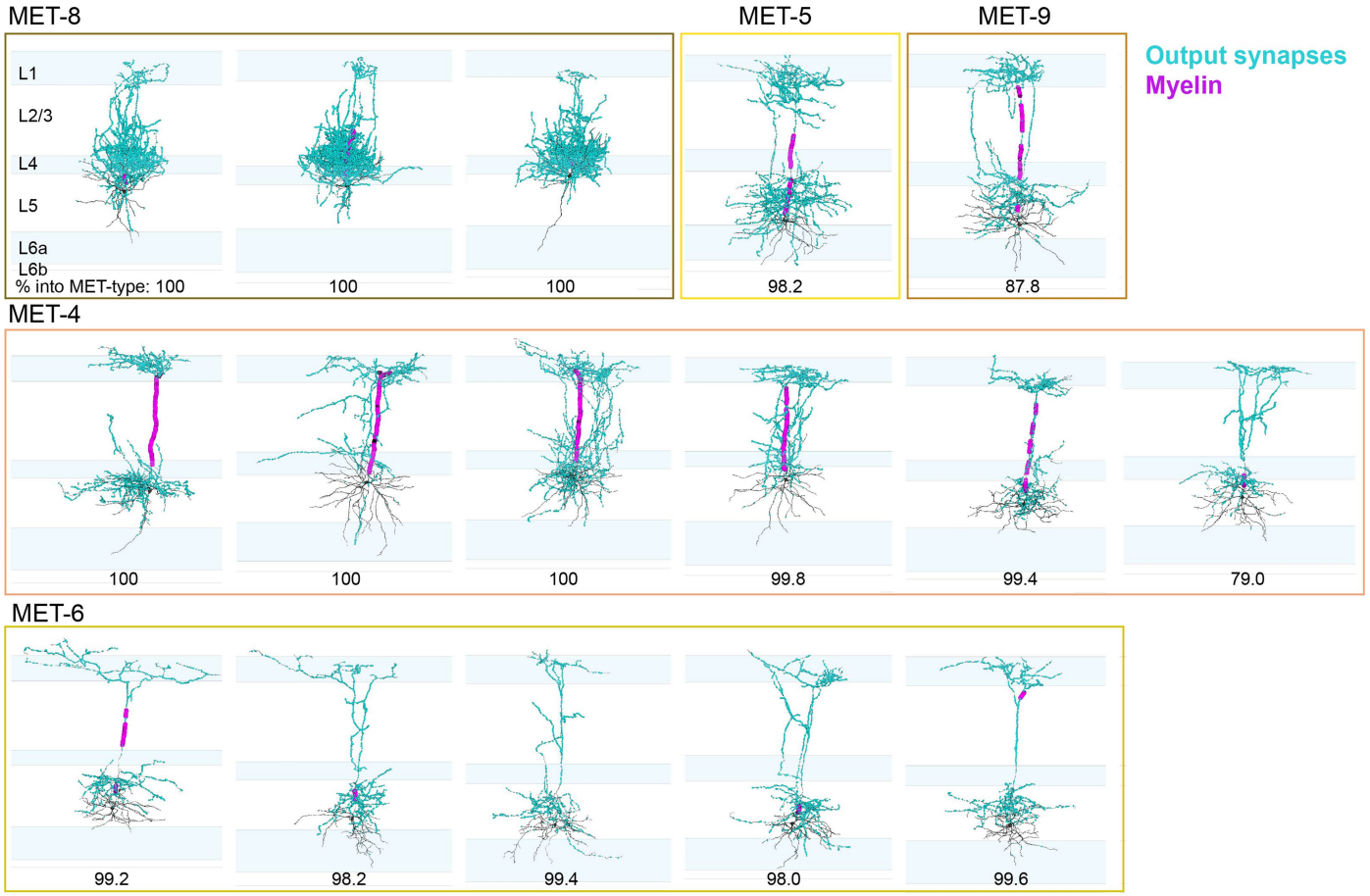


Extended Data Fig. 4 | Morphology of Patch-seq and EM cells grouped by subclass or *Sst* MET-type. Morphology of Patch-seq (before vertical line) and EM cells (after vertical line, including from the columnar sample and curated MCs) grouped by subclass or *Sst* MET-type. Abbreviations above cells indicate target cell type (described in Extended Data Fig. 2). Perisomatic targeting cells (PTC), distal dendrite targeting cells (DTC), sparsely targeting cells (STC), and inhibitory targeting cells (ITC). The number indicates the frequency with which that cell was predicted to that MET-type. For *Sst* MET-types, at least one EM cell is predicted to belong to 11 of 13 *Sst* MET-types. *Sst* MET-types: *Sst* MET-1, -2, -5, and -7 contain only a few EM neurons, but visual inspection demonstrates that morphologies are quite similar between Patch-seq cells and EM; these neurons are largely distal targeting cells (DTCs), consistent with the *Sst* subclass prediction. *Sst* MET-3 EM column neurons (n = 15) are found in L2/3 and L4 and largely have DTC connectivity profiles (11/15 cells). These cells have one longer, descending dendrite, and a "fanning" axonal profile that widens in L1 and L2/3. *Sst* MET-4 EM column and curated neurons (n = 11) have somas in L5, and dominant L1 axon. All column cells (6/6 cells) are also DTCs. *Sst* MET-6 EM column and curated neurons (n = 9) have somas in deep L5 and a "T-shaped" axonal profile. Six of seven column neurons are DTCs. *Sst* MET-8 column and curated neurons (n = 12) have somas in L4 and upper L5 and dominant L4 axon. Five out of ten column neurons have an *Sst*-consistent, DTC label. All of these neurons, like the Patch-seq neurons in this type, have axon in L1. Neurons that map to other connection types lack axon in L1. *Sst* MET-9 column and curated neurons (n = 6) have somas in upper L5 or L6 and large dendritic trees. Four of the five column neurons are DTCs. *Sst* MET-10 column neurons (n = 4) have somas in upper L5 or L6 and large dendritic trees with dominant axon in L5 and/or L6. Half of the cells are DTCs with axons that extend to superficial layers, but not necessarily to L1. *Sst* MET-12 column neurons (n = 12) have somas in L6. Most cells have large dendritic trees and diverse axonal phenotypes. EM DTCs that map to this type lack L1-projections and are thus non-Martinotti cells (3/12). EM PTCs also frequently map to *Sst* MET-types that are dominated by non-Martinotti neurons (5/12 cells are PTCs). *Lamp5* subclass: Patch-seq *Lamp5* neurons most characteristically have small dendrites and a wide, dense, horizontally extending axon in L1. All L1 EM cells (9/9) have small dendrites and the sparse targeting cell (STC) profile that is suggested to correspond to *Lamp5*/neurogliaform cells (NGCs). *Lamp5* neurons found in deeper layers also have dense axon overlapping with their dendrites, but it's distributed across layers. Only 7/21 deep *Lamp5* cells are STCs. *Pvalb* subclass: Patch-seq *Pvalb* neurons have typical *Pvalb*/basket cell morphologies with stellate dendrites and abundant local axon in each cell's soma layer. Most cells extend axon across multiple layers, but avoid L1. Thirty of 37 EM column cells are proximal targeting cells (PTCs). *Scng* subclass: Patch-seq *Scng* neurons have widely branching bitufted or multipolar dendrites and axons. EM column cells that map to this type (n = 4) each have a different connectivity type. *Vip* subclass: Patch-seq *Vip* neurons frequently have bipolar primary dendrites and axon with a narrow horizontal and long vertical extent. Twenty out of 22 EM column cells that map to *Vip* have an inhibitory cell targeting type (ITCs), which is consistent with *Vip* neurons. Figure adapted with permission from ref. 16, Cell Press.

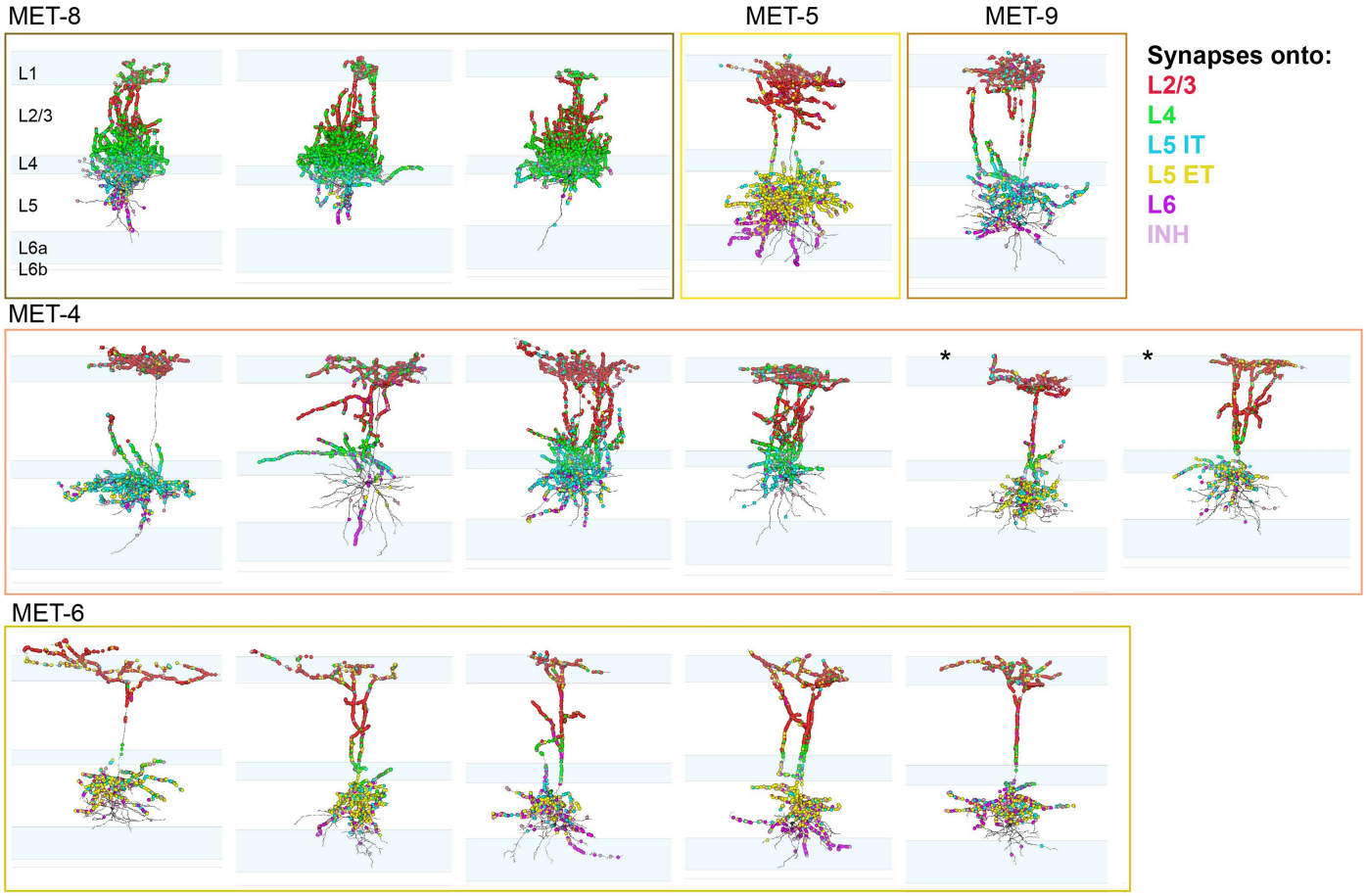


Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | Comparison of classifiers performance versus Patch-seq or manual analysis and additional details of synaptic connections by MET-type. a) Confusion matrix showing correct versus predicted labels for Patch-seq cells from the random forest classifier (n = 477). This is data shown in Fig. 1c, but is shown here collapsed along subclass. The classifier largely correctly predicts a cell's subclass for *Lamp5*, *Pvalb*, *Sst* and *Vip* cells. The frequency of a correct prediction appears in each square along the diagonal. b) Confusion matrix showing differences between manual (expert annotator) versus predicted subclass calls from Elabbady et al.[60] (n = 2454 cells). The largest disagreement results from the human annotator referencing layer drawings from different z-planes throughout the dataset which features a larger L2/3 and more restricted L4, whereas the classifier used labels derived from a columnar sample which resulted in more cells being called L2/3 that the classifier predicted at L4. Overall we observe a large degree of alignment between manual and predicted labels. c) Average connection size: the sum of the size (voxel) of all synapses onto an individual postsynaptic target. The connection size of MET-8 cells is significantly larger than MET-6 cells (K-W test $p < 0.05$; post hoc $^{**}p < 0.01$). MET- 8, 4, 6 (n = 3, 6, 5). d) Overall synaptic density: Total number of output synapses divided by total axon length. MET-8 cells have significantly higher output synapse density compared to MET-4 and MET-6 cells (K-W test $p < 0.05$; post hoc $^{**}p < 0.01$). e) Plot of the p-values comparing the actual and Poisson distribution of synapses from individual MCs onto target cell classes. We highlight the connections where the null hypothesis (the distributions were Poisson distributed) could not be rejected (grey).



Extended Data Fig. 6 | Visualization of myelin and output synapses for predicted MET-types. Output synapses and myelination patterns of all EM MCs grouped by predicted MET-type. The reliability metric (% out of 500 runs) that the cell was predicted into the MET-type displayed below.



Extended Data Fig. 7 | Visualization of output synapses by target cell type for predicted MET-types. Output synapses colour-coded by postsynaptic cell subclass of all EM MCs grouped by predicted MET-type. * denotes cells with dominant L5 ET targeting in contrast to the rest of the MET-4 cells.

Extended Data Table 1 | Classifier output, reliability, and fraction of synapses analyzed for EM Martinotti cells

Cell ID	Predicted MET-type	Reliability	Fraction of synapses analyzed
864691135544588584	Sst-MET-8	1.0	0.803
864691135572082925	Sst-MET-8	1.0	0.871
864691135699487522	Sst-MET-8	1.0	0.895
Average	-	-	0.856
Std Dev	-	-	0.039

Cell ID	Predicted MET-type	Reliability	Fraction of synapses analyzed
864691135013417622	Sst-MET-4	1.0	0.804
864691135969447781	Sst-MET-4	1.0	0.84
864691136118647832	Sst-MET-4	1.0	0.837
864691135577202181	Sst-MET-4	0.998	0.802
864691135785592004	Sst-MET-4	0.994	0.666
864691135118298333	Sst-MET-4	0.79	0.798
Average	-	-	0.791
Std Dev	-	-	0.058

Cell ID	Predicted MET-type	Reliability	Fraction of synapses analyzed
864691136618564493	Sst-MET-6	0.996	0.797
864691136116457636	Sst-MET-6	0.994	0.845
864691136238652476	Sst-MET-6	0.992	0.801
864691135404765166	Sst-MET-6	0.982	0.882
864691135697251738	Sst-MET-6	0.98	0.868
Average	-	-	0.839
Std Dev	-	-	0.035

Cell ID	Predicted MET-type	Reliability	Fraction of synapses analyzed
864691136195284556	Sst-MET-5	0.982	0.893

Cell ID	Predicted MET-type	Reliability	Fraction of synapses analyzed
864691135341516741	Sst-MET-9	0.878	0.805

Cell IDs, their predicted MET-type, the reliability at which they were predicted to that MET-type, and the fraction of output synapses analyzed. We report the fraction of output synapses that were made on to single-soma objects with cell type predictions and the average and standard deviation of those fractions by MET-type when possible. Loss of targets may be due to cell location relative to tissue boundaries. We find that ~83% of all output synapses have a defined single-soma target within tissue boundaries. We report the average and standard deviations of the fraction of synapses that have postsynaptic targets for MET-types with n>1.

Extended Data Table 2 | Estimated MET-type abundance from MERFISH data

Sst MET-type	Contributing t-types (Tasic et al. 2018)	Corresponding whole-brain taxonomy clusters (Yao et al. 2023)	Number of cells (Yao et al. 2023 MERFISH)	Fraction of Sst population (Yao et al. 2023 MERFISH)
Sst MET-2			*247*	***16.9%***
	Sst Tac1 Htr1d	795, 796, 797, 821	247	16.9%
Sst MET-3			*172*	***11.8%***
	Sst Calb2 Pdlim5 (*L4 and above*)	794, 802, 803, 804	172	11.8%
Sst MET-4			*50*	***3.4%***
	Sst Calb2 Pdlim5 (*below L4*)	794, 803, 804, 826	50	3.4%
Sst MET-5			*25*	***1.7%***
	Sst Nr2f2 Necab1	766, 806	25	1.7%
Sst MET-6			*241*	***16.5%***
	Sst Myh8 Etv1	805	14	1.0%
	Sst Chrna2 Glra3	778, 779, 780	184	12.6%
	Sst Chrna2 Ptgdr	776, 801	43	2.9%
Sst MET-7			*89*	***6.1%***
	Sst Hpse Sema3c	798, 799, 800	89	6.1%
Sst MET-8			*270*	***18.5%***
	Sst Hpse Cbln4	792, 793, 816	270	18.5%
Sst MET-9			*96*	***6.6%***
	Sst Tac2 Tacstd2	758, 760	82	5.6%
	Sst Tac2 Myh4	756	14	1.0%
Sst MET-10			*31*	***2.1%***
	Sst Rxfp1 Prdm8	757, 761	31	2.1%
Sst MET-12			*145*	***9.9%***
	Sst Crhr2 Efemp1	782, 818, 819, 820	72	4.9%
	Sst Rxfp1 Eya1	759	47	3.2%
	Sst Crh 4930553C11Rik	781, 783, 784	26	1.8%
Sst MET-13			*47*	***3.2%***
	Sst Nts	813, 814	24	1.6%
	Sst Esm1	810, 811	23	1.6%

Estimate of *Sst* MET-type abundance from MERFISH quantification of T-types[6] (Yao et al.[67]). The table shows *Sst* T-types, the number of cells mapping to that T-type identified in mouse primary visual cortex (VISp), the percentage of the *Sst* subclass that T-type represents, and the corresponding MET-type that the T-type maps into.

Extended Data Table 3 | Quantification of apical dendrite synapses for Layer 5 Martinotti and pyramidal cell targets

	Cell ID	Total syn	L5 PC syn	% L5 PCs	Apical assessed syn	% apical assessed syn	Apical syn	% Apical syn	MET-type
0	864691135118298333	2797.0	530.0	0.189488738	517.0	0.975471698	279.0	0.539651838	4
1	864691135577202181	2793.0	384.0	0.137486574	360.0	0.9375	270.0	0.75	4
2	864691135013417622	2359.0	187.0	0.079270877	177.0	0.946524064	113.0	0.638418079	4
3	864691135785592004	2141.0	489.0	0.228397945	487.0	0.99591002	221.0	0.453798768	4
4	864691135969447781	3783.0	869.0	0.229711869	766.0	0.881472957	313.0	0.408616188	4
5	864691136118647832	3235.0	1113.0	0.344049459	1084.0	0.973944295	300.0	0.276752768	4
6	Mean	2851.333333	595.3333333	0.20140091	565.1666667	0.951803839	249.3333333	0.511206273	-
7	Std dev	594.7307514	337.5930489	0.090577752	319.5993846	0.040434221	73.9152668	0.168998953	-
8	864691135404765166	1786.0	568.0	0.318029115	470.0	0.827464789	265.0	0.563829787	6
9	864691135697251738	2196.0	446.0	0.203096539	385.0	0.8632287	160.0	0.415584416	6
10	864691136238652476	1804.0	486.0	0.26940133	435.0	0.895061728	231.0	0.531034483	6
11	864691136618564493	1761.0	409.0	0.232254401	387.0	0.946210269	182.0	0.470284238	6
12	864691136116457636	1380.0	343.0	0.248550725	334.0	0.973760933	123.0	0.368263473	6
13	Mean	1785.4	450.4	0.254266422	402.2	0.901145284	192.2	0.469799279	-
14	Std dev	288.9131357	84.20985691	0.043099599	52.08358667	0.059597353	56.42428555	0.080374892	-

Table showing the total number of output synapses (Total syn), number of synapses that are formed onto L5 pyramidal cell targets (L5 PC syn) and the percentage of synapses that are formed onto the L5 PCs (% L5 PCs), calculated the previous column divided by total synapse number. Next we show the number of synapses that had a postsynaptic L5 cell whose apical dendrite had been assessed and determined algorithmically (Apical assessed syn) and the fraction of synapses onto L5 PCs that had this assessment (% apical assessed syn), calculated by "Apical assessed syn" divided by "L5 PC syn". Lastly we show the number of synapses that were formed onto apical dendrites (Apical syn) and the fraction of synapses onto L5 PCs that are onto apical dendrites (% Apical syn), calculated by "Apical syn" divided by "L5 PC syn". These cells are shown grouped by their MET-type and the mean and standard deviation of these values are shown below.

nature portfolio

Corresponding author(s): Forrest Collman, Nuno da Costa, Staci Sorensen

Last updated by author(s): 02/05/2024

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Analysis was performed using Python 3.10.9 and made extensive use of the following packages and libraries: CAVEclient (https://github.com/seung-lab/CAVEclient), CloudVolume, MeshParty, pcg_skel, skeleton_keys (https://github.com/AllenInstitute/skeleton_keys) to extract and analyze morphological features.
Data analysis	We used the following libraries for visualization and analysis: Matplotlib , Seaborn, Numpy, Pandas, VTK, Scipy, Scikit-posthocs, Scikit-learn, scrattch-hicat, statsmodels.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The mm3 EM dataset is publicly available at https://www.microns-explorer.org/cortical-mm3. Mouse Patch-seq data is publicly available at: https://

knowledge.brain-map.org/data/1HEYEW7GMUKWIQW37BO/summary. The associated metadata for Patch-seq cells is available at: https://portal.brain-map.org/cell-types/classes/multimodal-characterization.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	*Use the terms sex (biological attribute) and gender (shaped by social and cultural circumstances) carefully in order to avoid confusing both terms. Indicate if findings apply to only one sex or gender; describe whether sex and gender were considered in study design whether sex and/or gender was determined based on self-reporting or assigned and methods used. Provide in the source data disaggregated sex and gender data where this information has been collected, and consent has been obtained for sharing of individual-level data; provide overall numbers in this Reporting Summary. Please state if this information has not been collected. Report sex- and gender-based analyses where performed, justify reasons for lack of sex- and gender-based analysis.*
Population characteristics	*Describe the covariate-relevant population characteristics of the human research participants (e.g. age, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above."*
Recruitment	*Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results.*
Ethics oversight	*Identify the organization(s) that approved the study protocol.*

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The EM dataset is of unprecedented size and only one was produced. We identified 16 neurons with Martinotti-like morphology based on previous literature (at least three cells per gross morphological shape) across the depth of L5, restricted to the primary visual cortex region of the dataset. We included additional inhibitory neurons from a columnar sample (100 um x 100 um) (n = 163) for testing our classifier. Some Martinotti cells were present in the columnar sample so the total number of analyzed EM cells is n = 173. We trained the Random Forest Classifier on inhibitory Patch-seq data (previously published) (Total n = 477: Sst n = 236, Pvalb n = 89, Vip n = 79, Lamp5 n = 44, Sncg n = 29).
Data exclusions	We limited our analysis in EM to cells with Martinotti-like morphology (soma in L5, sparsely spiny dendrites, axon emerging from pia side of cell soma or dendrite, axon branch reaches layer 5). We did not include cells from MET-types with n < 5 when training the Random Forest Classifier.
Replication	Replication across EM datasets is not possible at this time due to the unprecedented large size of these data.
Randomization	Randomization is not relevant to this study as there was a single condition for all acquired data.
Blinding	Blinding is not relevant to this study as there was a single condition for all acquired data.

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	*Briefly describe the study type including whether data are quantitative, qualitative, or mixed-methods (e.g. qualitative cross-sectional, quantitative experimental, mixed-methods case study).*
Research sample	*State the research sample (e.g. Harvard university undergraduates, villagers in rural India) and provide relevant demographic information (e.g. age, sex) and indicate whether the sample is representative. Provide a rationale for the study sample chosen. For studies involving existing datasets, please describe the dataset and source.*

Sampling strategy	*Describe the sampling procedure (e.g. random, snowball, stratified, convenience). Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient. For qualitative data, please indicate whether data saturation was considered, and what criteria were used to decide that no further sampling was needed.*
Data collection	*Provide details about the data collection procedure, including the instruments or devices used to record the data (e.g. pen and paper, computer, eye tracker, video or audio equipment) whether anyone was present besides the participant(s) and the researcher, and whether the researcher was blind to experimental condition and/or the study hypothesis during data collection.*
Timing	*Indicate the start and stop dates of data collection. If there is a gap between collection periods, state the dates for each sample cohort.*
Data exclusions	*If no data were excluded from the analyses, state so OR if data were excluded, provide the exact number of exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.*
Non-participation	*State how many participants dropped out/declined participation and the reason(s) given OR provide response rate OR state that no participants dropped out/declined participation.*
Randomization	*If participants were not allocated into experimental groups, state so OR describe how participants were allocated to groups, and if allocation was not random, describe how covariates were controlled.*

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	*Briefly describe the study. For quantitative data include treatment factors and interactions, design structure (e.g. factorial, nested, hierarchical), nature and number of experimental units and replicates.*
Research sample	*Describe the research sample (e.g. a group of tagged Passer domesticus, all Stenocereus thurberi within Organ Pipe Cactus National Monument), and provide a rationale for the sample choice. When relevant, describe the organism taxa, source, sex, age range and any manipulations. State what population the sample is meant to represent when applicable. For studies involving existing datasets, describe the data and its source.*
Sampling strategy	*Note the sampling procedure. Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient.*
Data collection	*Describe the data collection procedure, including who recorded the data and how.*
Timing and spatial scale	*Indicate the start and stop dates of data collection, noting the frequency and periodicity of sampling and providing a rationale for these choices. If there is a gap between collection periods, state the dates for each sample cohort. Specify the spatial scale from which the data are taken*
Data exclusions	*If no data were excluded from the analyses, state so OR if data were excluded, describe the exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.*
Reproducibility	*Describe the measures taken to verify the reproducibility of experimental findings. For each experiment, note whether any attempts to repeat the experiment failed OR state that all attempts to repeat the experiment were successful.*
Randomization	*Describe how samples/organisms/participants were allocated into groups. If allocation was not random, describe how covariates were controlled. If this is not relevant to your study, explain why.*
Blinding	*Describe the extent of blinding used during data acquisition and analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.*

Did the study involve field work? ☐ Yes ☒ No

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging