Comparative transcriptomics reveals shifts in cortical architecture at the metatherian/eutherian transition

Authors: Ryan Gorzek and Joshua T. Trachtenberg

Affiliations: Department of Neurobiology, David Geffen School of Medicine, University of California-Los Angeles; Los Angeles, CA 90095

Corresponding author: Joshua Trachtenberg, joshua.trachtenberg@gmail.com

Author Contributions: Single-cell experiments: RG, JTT; Spatial transcriptomics experiments: RG, JTT; bioinformatics: RG; anatomy and histology: RG, JTT; data analysis: RG; project management and supervision: JTT; and manuscript writing: RG, JTT.

Competing Interest Statement: Joshua Trachtenberg is the co-owner of Neurolabware, LLC

Classification: Biological Science; Neuroscience

Keywords: Neocortex, transcriptomics, comparative, evolution, column

This PDF file includes:

Main Text
Figure 1 to 4
Supplementary Figures 1 to 4
Supplementary Table

Abstract:

The neocortex, a layered structure unique to mammals, supports higher-order functions such as perception, learning, and decision-making. While its overall organization is broadly conserved, it remains unclear whether changes to the cell type-specific organization of the cortical column—the basic unit of cortical processing— occurred at the metatherian/eutherian split. To address this, we used single-nucleus RNA sequencing and spatial transcriptomics to compare gene expression, cell types, and laminar architecture in the primary visual cortex of a metatherian (*Monodelphis domestica*) and a eutherian mammal (*Mus musculus*). We find distinctions between supragranular (layer 2/3) and infragranular (layers 5 and 6) intratelencephalic (IT) neurons to be more pronounced in mice, with layer 2/3 neurons enriched for transcripts involved in synaptic transmission, adhesion, and dendritic development. Mouse cortex also exhibits expanded populations of disinhibitory interneurons, redistribution of perineuronal nets, and reduced oligodendrocyte density—features known to enhance cortical plasticity. These findings suggest the metatherian–eutherian split was accompanied by remodeling of the cortical column, highlighting potential substrates of neocortical evolution in early mammals.

Significance Statement:

The neocortex is central to perception and flexible behaviors in mammals, yet its evolutionary origins remain unclear. By comparing cell types and spatial organization in visual cortex of marsupials and placental mammals, we found that key neuron classes—especially those involved in intracortical communication—became more specialized and spatially distinct in placentals. These changes are paralleled by shifts in inhibitory circuitry and myelinating cells, suggesting an evolutionary trajectory favoring increased cortical plasticity. Our findings challenge the long-held view that cortical columns are tightly conserved and shed light on the cellular substrates of early neocortical evolution.

Main Text

Introduction

The neocortex is a hallmark of mammalian evolution, enhancing perception and action through its characteristic laminar and columnar structure (*1, 2*). Central to its evolutionary elaboration is cortical arealization, the subdivision of cortex into specialized functional regions that scale computational capacity as brains enlarge (*3-6*). However, in its cellular organization across the laminar depth of cortex – often termed a cortical column – comparative studies have underscored the stability of laminar organization, projection patterns, and major neuronal subclasses (*7, 8*). This conservation has shaped the prevailing view that local cortical architecture is conserved across mammalian lineages (*9*), with primate V1 being a notable exception (*10-12*). Whether fundamental changes in the cortical column occurred during early mammalian evolution remains poorly understood.

One of the earliest mammalian divergences – the split between metatherians (marsupials) and eutherians (placentals) approximately 160 to 180 million years ago (*13-15*) – offers insight into this question. The gray short-tailed opossum (*Monodelphis domestica*), a marsupial, retains many features thought to resemble the ancestral therian brain, including a lissencephalic neocortex, relatively small and unlaminated thalamic nuclei, and prolonged postnatal development, which recapitulates aspects of early mammalian developmental patterns.

To determine whether this deep evolutionary split was accompanied by changes in the laminar and cell type–specific organization of neocortex, we used single-nucleus RNA sequencing (snRNA-seq) and spatial transcriptomics (Stereo-seq) to compare the primary visual cortex of opossums and mice. Mice were chosen to minimize ecological variability and focus on evolutionary divergence of cortical architecture. Both species are small, ground-dwelling nocturnal foragers and opportunistic omnivores (*16-18*); they are largely solitary outside of mating, rear their young in nests or burrows, and occupy similar trophic positions as both predator and prey, with short lifespans and brains constituting roughly 1–1.5% of adult body weight. Retinal cell types have been sequenced in both species, revealing broadly conserved classes of retinal ganglion cells and thus comparable information streams relayed through the thalamus to primary visual cortex (*19*). Although direct comparative behavioral evidence is limited, opossums and rodents diverge in tasks requiring learning, memory, and behavioral flexibility (*20-22*), highlighting potential differences that may be rooted in evolutionary changes to cortical structure and function.

Our data reveal that although cell subclasses and laminar structures are generally conserved, intratelencephalic (IT) glutamatergic neurons differ dramatically in transcriptomic identity and spatial organization across species. Consistent with the concept of evolutionary specialization and division of labor (*23-25*), opossum cortex exhibits generalized, overlapping functional identities among IT neurons, while mouse cortex displays clearly defined transcriptomic and spatial specializations. Furthermore, differences in inhibitory interneuron populations, perineuronal net distribution, and oligodendrocyte abundance collectively suggest a broader evolutionary shift towards

increased cortical plasticity (*26-30*). These findings demonstrate that fundamental changes in cortical organization, particularly within IT neuron populations, occurred early in mammalian evolution, possibly in the metatherian/eutherian split, highlighting potential mechanisms by which the neocortex acquired new functional capacities.

Results

A single-nucleus transcriptomic atlas of the primary visual cortex in the gray short-tailed opossum

We generated a single-nucleus RNA sequencing (snRNA-seq) atlas of the adult gray short-tailed opossum (*Monodelphis domestica*) V1 using 10x Chromium (Fig. 1A, B, see Materials and Materials and Methods). Tissue from both hemispheres was collected from two animals, pooled by animal, and used to generate four nuclear transcriptomic libraries. After preprocessing (see Materials and Methods, fig. S1E–P), we obtained 32,764 high-quality nuclear transcriptomes (Fig. 1C). We compare these data to our previously reported snRNA-seq dataset from adult (P38) mice (*31*), generated using similar protocols (Fig. 1D).

Following dimensionality reduction and clustering, we used canonical marker genes (Fig. 1E, fig. S1I–J) and whole-transcriptome classification (fig. S1K–P) to define major cell classes—glutamatergic, GABAergic, and non-neuronal—and their subclasses in opossum V1. Cross-species integration and co-clustering revealed that most subclasses were conserved, and transcriptomic hierarchies were broadly similar across species, suggesting conserved cellular taxonomies (Fig. 1F). However, subclass-level frequencies diverged. While glutamatergic neurons were more abundant in the mouse (Fig. 1G), the relative proportions of glutamatergic subclasses were nearly identical (Fig. 1H). By contrast, GABAergic composition differed markedly: opossum inhibition was dominated by medial ganglionic eminence (MGE)-derived interneurons expressing *Pvalb* or *Sst*, with parvalbumin (PV)-positive cells comprising nearly 50% of the GABAergic population—far more than in mouse (~35%, Fig. 1J; confirmed by histology, Fig. 4A–C). In mice, inhibition was biased towards caudal ganglionic eminence (CGE)-derived subtypes expressing *Vip*, *Lamp5*, or *Frem1* (Fig. 1I). Notably, MGE-derived interneurons are typically involved in feed-forward inhibition, driven by early thalamic and intracortical input (*32, 33*), whereas CGE-derived interneurons are recruited later via feedback from higher cortical areas and participate in modulatory or disinhibitory circuits (*34-36*). Among non-neuronal types, oligodendrocytes were also more abundant in opossums than mice (Fig. 1K), even when excluding endothelial cells (fig. S1Q), which are particularly sensitive to dissociation (*37*). These differences are examined in greater detail below.

Intratelencephalic glutamatergic neurons have poor cross-species correspondence

While we successfully identified most glutamatergic, GABAergic, and non-neuronal subclasses in opossums, IT neurons posed a unique challenge. These neurons project within the cortex and to the striatum and are subdivided in mice and other eutherians into layer-specific populations based on gene expression (*31*). In transcriptomic space, mouse, IT neurons form a well-structured continuum (Fig. 2A, center panel) where unsupervised clustering and canonical markers (Fig. S2A–D) resolve distinct subclasses

in layers 2/3, 4, 5, and 6. Opossum IT neurons also formed a continuum but lacked coherent expression of canonical layer markers (Fig. 2B, fig. S2H, L), precluding assignment to homologous subclasses. We therefore defined four opossum IT subclasses (IT_A–D) using unsupervised clustering (fig. S2E–M) and assessed their transcriptomic correspondence to the layer-specific IT types in the mouse.

To do so, we restricted our analysis to genes with one-to-one orthologues (fig. S2N). This filtering had minimal impact on subclass separability within each species, which remained highly distinct (fig. S2O). We then applied Seurat's MapQuery pipeline to train models on mouse subclass labels and predict labels for opossum cells based on gene expression (Fig. 2A). Some opossum subclasses mapped strongly onto mouse counterparts—for example, IT_B corresponded to thalamorecipient L4 and IT_D to L6IT. In contrast, IT_A, IT_C, and L5NP showed weak and inconsistent correspondence to any mouse IT subclass (Fig. 2C, fig. S2P). By comparison, other glutamatergic populations (e.g., L5PT, L6CT, and L6b), as well as most GABAergic and non-neuronal subclasses, showed strong cross-species alignment (Fig. 2D–E, fig. S2Q–R). Notably, the distinction between *Vip*- and *Frem1*-expressing interneurons—CGE-derived GABAergic subclasses that, like IT neurons, form a continuum—was substantially blurred across species (Fig. 2D). While this reflects the difficulty of defining 'types' within continuous cell populations, closer examination (described below) revealed fundamental differences between the structure of IT and CGE continua across species.

<u>Gene expression continua reveal conserved and divergent features of IT neurons</u>

Gene expression continua, as seen for IT cells, arise naturally during development but also exist in differentiated, multifunctional tissues that divide labor among specialized cells. According to tissue multi-tasking theory (*23-25*), such division of labor drives cell types to evolve along Pareto fronts—geometric boundaries that capture trade-offs between competing functions—resulting in predictable patterns of differentiation and spatial organization. As such, cells in different neocortical layers are specialized, with those projecting to subcortical targets (e.g., L5PT, L6CT) having distinct transcriptomic profiles from IT neurons (*31, 38*). Among IT neurons, all of which project within the cortex, gene expression patterns are graded. These gradients suggest partial functional overlap and division of labor across layers. Given the imperfect correspondence of IT subclasses between opossum and mouse, we examined the structure of IT continua in species-specific principal component (PC) spaces. Multi-tasking theory predicts that tissues with D-dimensional spatial zonation (here, D=4 cortical layers: 2/3, 4, 5, 6) will lie on a D-1-dimensional manifold in gene expression space (*23*). Indeed, both mouse and opossum IT neurons formed tetrahedral structures in the first three PCs (Fig. 2F–G, fig. S3A–B). Projecting opossum IT neurons into mouse PC space preserved this structure (Fig. 2H–I, fig. S3C), but only L4 and IT_B showed substantial cross-species overlap (Fig. 2J), consistent with their unique alignment discussed above.

Moreover, the mouse IT-specific weighted gene co-expression network analysis (WGNCA) module with the greatest cross-species preservation was large (>500 genes) and L4-specific (Fig. 2K–L). Other conserved modules (M7 and M10) were shared across L4 / IT_B and L6IT / IT_D but were weakly represented in opossum IT_A despite strong L2/3 representations in mice (Fig. 2L), highlighting the divergence of upper-layer

IT neurons across species. Gene ontology analysis of the conserved M4 and M7 revealed an enrichment for genes regulating synaptic plasticity, cell-cell adhesion, synapse assembly, and dendritic arborization, suggesting evolutionarily conserved functional and anatomical properties in the thalamorecipient layer (Fig. 2M).

Principal component analysis further revealed that while mouse IT cells varied continuously between L4 and other IT subclasses, L2/3, L5IT, and L6IT were themselves more discrete and occupied distinct vertices of the tetrahedron (Fig. 2F, N, fig. S3A). By contrast, in opossum, the tetrahedral IT structure comprised a triangular plane containing a continuous gradient of IT_A, IT_B, and IT_C cells (Fig. 2N–O, fig. S3B), with IT_D cells partially overlapping with the L4-like IT_B (fig. S3B, bottom left). Importantly, this gradient phenomenon is unique to the IT continua (fig. S3D–E) and does not arise from low-quality or noisy gene expression data in opossums (fig. S3F–H, see Materials and Methods). In line with the unique variation between IT_A and IT_C in opossums (Fig. 2N), the correlation of their top highly variable genes was far higher than mouse L2/3 and L5IT (Fig. 2P). The top within-species DEGs between these subclasses were enriched for genes involved in neuronal signaling, including ionotropic receptor subunits, and synaptic release (Fig. 2Q). However, mouse-specific DEGs (between L2/3 and L5IT) were also enriched for genes related to adhesion and dendritic development, suggesting that additional specialization in connectivity and morphology evolved in eutherian mammals.

<u>Opossum IT_A neurons display reduced laminar spatial organization</u>

Pareto optimality, when applied to tissues, predicts that cells that are transcriptomically specialized—those at the vertices of a tetrahedron in PC space—should also be spatially segregated (*23*). Given the poor cross-species correspondence among IT subclasses and the less distinct transcriptomic clustering in opossums, we asked whether these differences extended to laminar organization. We performed whole-transcriptome spatial profiling of mouse and opossum V1 using Stereo-seq (*39*).

Following multi-embedding of fresh-frozen coronal sections and the Stereo-seq pipeline (see Materials and Methods), we identified layers 1-6 within V1 using anatomical landmarks and canonical marker genes (Fig. 3A–B, fig. S4A–B), extracted cortical columns from multiple sections, and assigned cell subclass identities via co-clustering with matched snRNA-seq data (Fig. 3C, fig. S4E–H). Subpial density profiles reveal varying degrees of spatial organization in glutamatergic subclasses from both species (Fig. 3D). Mouse IT subclasses L2/3, L4, and L6IT had clear laminar peaks, while L5IT and L5NP had broader distributions that were appropriately centered in layer 5. In opossum, IT_A, IT_B, and L5NP had prominent peaks, but their distributions differed from expectations. For instance, while IT_A displayed a prominent peak above IT_B, a secondary, smaller peak was also present below it. In addition, IT_B was more broadly distributed than mouse L4. L5NP density peaked below L5PT, reversing the typical layer 5a/5b relationship. IT_C and IT_D were broadly distributed; IT_C peaked in putative layer 5, while IT_D peaked much higher than expected given its cross-species correspondence to L6IT.

By contrast, the extratelencephalic projection neuron subclasses L5PT and L6CT displayed prominent and conserved laminar localization in both species. GABAergic and

non-neuronal subclasses largely followed expected depth distributions (fig. S4I–J). To quantify spatial restriction, we computed Gini indices from the subpial depths distributions of each glutamatergic subclass and plotted cross-species comparisons (Fig. 3E). This measures the degree to which cells are restricted to a single subpial depth, as is the expectation for the laminated neocortical glutamatergic subclasses. Subclasses with input and output roles—L4 / IT_B, L5PT and L6CT—showed the highest spatial organization in both species, consistent with the notion that these subclasses are the most tightly conserved throughout cortical evolution.

To further explore the relationship between spatial and transcriptomic variation in IT neurons, we integrated Stereo-seq and snRNA-seq data from L2/3 / IT_A, L4 / IT_B, and L5IT / IT_C cells and projected them onto shared principal components (Fig. 3F). In both species, Stereo-seq cells formed triangular gradients in PC space, reminiscent of our snRNA-seq findings (Fig. 2N). We then pseudocolored Stereo-seq IT cells in the gradient by their distance from each of three triangle vertices—A (L2/3-like), B (L4-like), and C (L5IT-like)—defined in PC2-PC3 space (Fig. 3G). Spatial mapping of these pseudocolored cells showed that IT_A cells were well-represented both in putative L2/3 and L5. (Fig. 3H, I). The spatial link suggests that IT_A and L5IT-like IT_C may have diverged into functionally distinct populations during the metatherian-eutherian transition. Supporting this, IT_A cells closest to transcriptomic vertex A (bright cyan) were concentrated in putative layer 2/3 (Fig. 2H, see also fig. S4N–P). In mice, L2/3 cells show a continuous gradient toward L4 in both transcriptomic and physical space, such that deeper L2/3 cells have gene expression profiles more similar to L4 cells (Fig. 2J). This relationship also holds true for opossum IT_A, which varies linearly with depth relative to vertex B. Critically, opossum IT_A varies similarly with respect to vertex C (L5IT-like), a pattern absent in mouse L2/3. This suggests that IT_A and IT_C form a biologically meaningful gradient in opossums that has no clear analog in eutherians—implying a fundamental shift in spatio-transcriptomic organization across species.

<u>PV-positive GABAergic neuron and oligodendrocyte populations are denser in opossum V1 without significant changes in total neuronal packing density</u>

To validate the cross-species differences in GABAergic and non-neuronal subclass proportions observed in our transcriptomic data (Fig. 1J, K), we performed immunofluorescent staining in mouse and opossum V1. Consistent with our RNA-seq findings, the density of PV-positive neurons was ~20% higher in opossums than in mice (Fig. 4A, B), with the most pronounced difference in layer 5—the layer with the highest PV density in both species. Nonetheless, PV neuron densities were broadly scaled across layers in opossums (Fig. 4C). Given the known role of PV neurons in regulating critical period plasticity (*28, 40*), we also stained for perineuronal nets (PNNs) using Wisteria floribunda agglutinin (WFA; Fig. 4A). PNNs are associated with structural stabilization and are enriched in layer 4 of mouse primary sensory cortex, where their density correlates with thalamic input strength (*41*). As reported previously (*42*), WFA staining was densest in mouse layer 4. In contrast, staining in opossum was restricted to layers 5b and 6, appearing to surround the apical dendrites of layer 5 pyramidal tract neurons (Fig. 4A). Outside of V1, both PV neuron density and WFA intensity dropped sharply in both species (Fig. 4A, D, white arrows). We also examined oligodendrocyte density using CC1 immunostaining (Fig. 4E). In agreement with our snRNA-seq data

(Fig. 1K), CC1-positive oligodendrocytes were found to be dramatically more abundant in opossums than mice (Fig. 4F). These differences in PV and oligodendrocyte density were not explained by global differences in neuronal packing density (Fig. 4G, H), though NeuN staining curiously revealed that layer 6 is significantly denser in opossums than mice.

Discussion

Metatherian and eutherian mammals diverged roughly 160 to 180 million years ago during the Middle to Late Jurassic, marking a critical transition in mammalian brain evolution (*13-15*). Our data reveal fundamental differences in the radial unit of cortical organization—often referred to as a cortical column—between the short-tailed South American opossum, which retains many features thought to resemble the ancestral therian brain, and mice, a representative eutherian rodent. Using comparative transcriptomics, we identify specific cell populations most affected by evolutionary pressures. Most notable is the transcriptomic divergence of L2/3 and L5IT subclasses and their increased segregation in both gene expression space (the vertices of a tetrahedron in the top PCs) and physical (laminar) space in mouse cortex. This emergent organization supports a view in which cell types evolve along Pareto fronts, with greater specialization in eutherian cortex potentially emerging from initially less differentiated, generalist cell types in metatherian cortex (*23-25*). Pareto front theory predicts the emergence of specialized transcriptomic states through trade-offs between functional demands, but the specific "tasks" driving specialization in the neocortex—likely related to sensory and motor integration, inter-areal communication, and experience-dependent adaptability—remain to be defined. The conserved identity and organization of thalamorecipient neurons across species suggests that cortical specialization evolved via remodeling of intracortical circuits, rather than shifts in thalamic input.

The case of L2/3 IT neurons exemplify this shift most clearly. In mice, these neurons are crucial for intracortical communication, sensory integration, and top-down modulation. Compared with opossums, they exhibit greater transcriptomic specialization, including enrichment for genes associated with synaptic transmission, connectivity, and plasticity. This pattern is consistent with an evolutionary sharpening of L2/3's role in eutherian cortical processing. In contrast, the more diffuse transcriptomic profile of opossum L2/3 neurons suggests a broader functional repertoire within a less modular metatherian cortical framework.

The spatial organization of these transcriptomic continua supports this view. In both species, IT neurons form low-dimensional manifolds in gene expression space, but these manifolds align more closely with cortical layers in mice. The most conserved subclasses – thalamorecipient L4, extratelencephalic L5b, and corticothalamic L6 – remained transcriptomically and spatially stable across nearly 180 million years of evolutionary divergence. By contrast, IT neurons appear to bear the brunt of evolutionary pressures, even relative to other transcriptomically continuous populations, a pattern that may reflect their comparatively flexible and experience-dependent connectivity (*43*). This contast – plasticity in IT neurons and stability in input/output neurons – differs sharply from the retina, where the 'output' ganglion cells are most affected by evolutionary change (*19*).

We also uncovered intriguing evolutionary shifts within populations of inhibitory neurons. Mice possess relatively fewer PV interneurons in relation to CGE-derived interneurons, including VIP-expressing cells. This may suggest an evolutionary preference for enhanced flexibility, disinhibition, and neuromodulation, reducing rigid temporal precision and promoting adaptable cortical computations. PV interneurons receive strong feedforward excitation and provide strong perisomatic intracortical inhibition that sharpens spike timing and enhances gamma oscillations (*32, 44, 45*). Their relative scarcity in mice suggests a somewhat weakened inhibitory control over pyramidal neuron output. This could result in a cortex that is less reliant on tight temporal precision and synchronous firing, potentially allowing for more flexible, exploratory, or integrative processing strategies. The higher relative abundance of CGE-derived interneurons suggests an increased role for disinhibition and neuromodulation. Many CGE-derived interneurons, such as VIP cells, are strongly modulated by behavioral state (*46-48*). By targeting SST and PV cells, CGE-derived inhibition gates top-down intracortical feedback via disinhibition (*36, 49, 50*). This balance in eutherians likely enhances flexible, context-dependent cortical processing, whereas opossums—with fewer CGE-derived interneurons and stronger perisomatic inhibition—may have a more feedforward-driven, bottom-up sensory processing style with less cortical modulation of sensory inputs. Notably, variation in interneuron proportions is observed across multiple mammalian lineages (*38*), but we stress the importance of making these comparisons in ethologically relevant contexts. Even among primates such as marmosets and humans, relative differences in inhibitory neuron classes are limited; instead, distinctions in excitatory populations like L2/3 IT neurons appear more pronounced.

Additionally, the relocation of perineuronal nets (PNNs) – from L5b pyramidal neurons in opossums to L4-localized PV interneurons in mice – likely reflects changing strategies in cortical plasticity and circuit stabilization. In most eutherian species, PNNs are closely associated with PV interneurons, where they stabilize inhibitory circuits and contribute to the closure of critical periods and memory consolidation (*29, 30*). During early development, inputs to PV interneurons are highly plastic, allowing experience-dependent refinements in feedforward drive to cortical circuits (*28*). The subsequent consolidation of feedforward inputs at the closure of the critical period stabilizes fundamental response properties, while leaving intact the cortical plasticity needed for lifelong learning (*51*). In contrast, layer 5 pyramidal neurons integrate intracortical inputs and send long-range outputs to subcortical structures. The association of PNNs with these neurons in metatherians may constrain the flexibility of cortical outputs, whereas in eutherians, restricting PNNs to input circuits could facilitate greater adaptability in modifying output pathways based on learning and experience.

Our finding of reduced oligodendrocyte numbers in mice compared to opossums further supports the view that critical period regulation was shaped early in cortical evolution. Oligodendrocytes produce myelin-associated proteins such as Nogo receptor and myelin basic protein, known modulators of synaptic plasticity and critical period closure (*52*). This aligns with evidence that preventing oligodendrocyte maturation and myelination can extend plasticity windows in the visual cortex (*26*). Together, these observations are consistent with a model in which oligodendrocyte-derived myelination acts as a developmental brake, stabilizing cortical circuits and constraining experience-dependent

plasticity as mammals mature. The relative paucity of oligodendrocytes in eutherian cortex may, therefore, help sustain prolonged plasticity and enhance adaptive responses to environmental challenges.

Together, these findings suggest the metatherian-eutherian split involved a fundamental reorganization of cortical architecture toward increased transcriptomic specialization, spatial differentiation, and greater capacity for plasticity. Such innovations in the neocortex likely supported enhanced modularity, computational complexity, and behavioral flexibility during mammalian evolution.

Materials and Methods

Animals

Mouse breeding and husbandry were performed in accordance with the institutional animal care and use committee protocol number ARC-2009-031 at the University of California, Los Angeles. Mice were given food and water ad libitum and lived in a reverse 12-hour light/dark cycle with fewer than five adult animals per cage. One adult (~P40) male wild-type C57BL/6J mouse was used for spatial transcriptomics experiments in this study. Two additional adult (~P40) males were used for immunofluorescence. Single-nucleus RNA sequencing data from adult mice (P38) was obtained from our previous study (30), which utilized identical breeding, husbandry, and genetic conditions.

Opossum (*Monodelphis domestica*) breeding and husbandry were performed in accordance with the institutional animal care and use committee protocol number ARC-2009-031 at the University of California, Los Angeles. Opossums were given food and water ad libitum and lived in a 12-hour light/dark cycle without cagemates. Two adult (1.5–2 years) male opossums were used for single-nucleus RNA sequencing. One juvenile (~6 months) male opossum was used for spatial transcriptomics. Two additional adult (1.5–2 years) males were used for immunofluorescence.

V1 dissection

Opossums were deeply anesthetized with isoflurane followed by decapitation, and the brain was immediately removed and submerged in ice-cold Hibernate A (BrainBits Cat. # HACA). Anatomical landmarks (52) were used to identify area 17, the primary visual cortex (V1). Extracted brains were placed in an ice-cold metal mold (Kent Scientific Cat. # RBMS-300C) and a 2 mm coronal slice was obtained by bisecting the superior colliculus with a razor blade (Personna Super) and inserting a second blade 2 mm anterior to the first. This slice was immediately placed in Hibernate A and cuts were made lateral to V1 (~3 mm from the midline) in both hemispheres. V1 (from both hemispheres) was then removed from the underlying white matter and immediately transferred to a cryogenic vial (Millipore Sigma Cat. # CLS431417) and flash frozen in liquid nitrogen. Tissue was stored in liquid nitrogen for ~1 year before single-nucleus RNA-sequencing was performed.

Nuclei isolation and droplet-based single-nucleus RNA sequencing

Nuclei were isolated from each biological replicate using the 10x Genomics Chromium Nuclei Isolation Kit. Both biological replicates were processed in parallel by two different experimenters. To remove excess debris (previously observed in adult opossum nuclei preparations), we performed fluorescence-aided nuclear sorting (FANS) for Draq5-positive nuclei at the UCLA Flow Cytometry Core. After FANS, nuclei suspensions were counted on a hemocytometer using an ethidium bromide counterstain. Nuclei suspensions were examined for blebbing under a brightfield microscope before dilution to 700-1200 cells per µL. Nuclei suspensions from each biological replicate were loaded into two wells of a 10x Chromium 3' v3 chip, each targeting 10,000 cells. We refer to these as library replicates. Libraries were pooled into a single lane of an Illumina NovaSeq™ X Plus 10B and sequenced to a depth of ~30,000 reads (2x50 bp) per cell. All library preparation and sequencing was performed by the UCLA Technology Center for Genomics and Bioinformatics (TCGB).

Spatial transcriptomics (Stereo-seq v1.3)

Animals were deeply anesthetized with isoflurane followed by decapitation, and brains were immediately removed and submerged in ice-cold 1x PBS containing RNase inhibitor (1 unit per µL; Thermo Fisher Cat. # AM2696). Extracted brains were placed in an ice-cold metal mold (opossum: Kent Scientific Cat. # RBMS-300C; mouse: RBMS-205C) and in the opossum, 2 mm coronal slices were obtained by bisecting the superior colliculus with one blade and inserting additional blades 2 mm anterior and posterior to the first (52). In the mouse, one blade was inserted at the cortical margins of the midline (below the lambdoid suture) and additional blades 1 mm anterior and posterior to the first (53). These slices were immediately removed from the mold and placed back into ice-cold PBS. To embed two coronal slices (containing four total V1 hemispheres) in the opossum, cuts were made lateral to V1 (~3 mm from the midline) in both hemispheres, as well as horizontal (~4 mm from the highest point of cortical surface). Because of the difference in brain size, these cuts were not necessary in the mouse. Coronal slices were arranged in an ice-cold metal base mold (Azer Scientific Cat. # ES4440) containing a drop of OCT (Tissue-Tek), then covered in OCT and flash-frozen in liquid nitrogen. Blocks were stored at -80°C and sectioned at -20°C. The Stereo-seq pipeline was performed according to user manuals provided by the vendor (Complete Genomics). Sectioning and imaging was performed by the UCLA Translational Pathology Core Laboratory (TPCL). RNA integrity (RIN) analysis, permeabilization, library preparation, and sequencing was performed by the UCLA TCGB. First, RNA integrity numbers (RINs) were obtained from shavings of the embedded tissue blocks. Due to skull thickness and the difficulty of removing the brain quickly in opossums, we obtained a RIN of 7.1 for opossum tissue and a RIN of >9 for mouse tissue. Despite this disparity, both were well above the RIN threshold of 4.0 required for Stereo-seq v1.3. Next, the optimal tissue permeabilization length for each sample was determined according to the Stereo-seq Permeabilization Set for Chip-on-a-slide User Manual. Briefly, four 10-µm sections were cut and mounted onto separate 10 mm x 10 mm chips on the Stereo-seq Chip P slide using the 'warm' method. Chip P slides were then dried, fixed in pre-cooled methanol for 30 minutes, and re-dried. 1X Permeabilization Reagent Solution was pre-warmed to 37°C and applied to different chips for 6, 12, 18, or 24 minutes. Following a wash, Chip P slides underwent in situ reverse transcription. RT QC

Mix was added to each chip and incubated at 45°C for 60 minutes. Then, Tissue Removal Mix was incubated for 60 minutes and the slide was washed until no tissue traces remained on the chips. Slides were imaged at 10x magnification using an EVOS™ M7000. The optimal permeabilization time was independently chosen to be 12 minutes for both mouse and opossum samples based on maximal fluorescence signal with minimal signal diffusion. Next, the transcriptomics procedure was performed according to the Stereo-seq Transcriptomics Set for Chip-on-a-slide User Manual. First, a single 10-µm section was cut and mounted onto the 10 mm x 10 mm chip on the Stereo-seq Chip T slide. Chip T slides were then dried, fixed in pre-cooled methanol for 30 minutes, and re-dried before staining with nucleic acid dye (Qubit ssDNA Assay Kit, Invitrogen Cat. # Q10212) and imaging at 10x magnification (EVOS M7000; fluorescein isothyocyanate channel). Chip T slides then underwent tissue permeabilization (12 minutes) and in situ reverse transcription before cDNAs were released, denatured, amplified, and purified. The resulting libraries were sequenced on a full MGI DNBSEQ-T7 flow cell with paired-end, 75 bp reads. Stereo-seq data processing is discussed in detail below.

Immunofluorescence

Animals were deeply anesthetized with isoflurane followed by decapitation. Brains were immediately removed and submerged in ice-cold 1x PBS for 5 minutes, then transferred to ice-cold 4% paraformaldehyde (PFA) for another 5 minutes. Extracted brains were transferred to a fresh tube of 4% PFA and post-fixed overnight at 4°C, then transferred to 20% sucrose until they sank to the bottom of the tube. Brains were placed in a plastic mold (Sigma-Aldrich Cat. # E6032), covered in OCT, and frozen on dry ice. Blocks were stored at -80°C until cryosectioning at -20°C to obtain 40-µm free-floating coronal sections encompassing the rostrocaudal extent of V1. Sections were stored in 1x PBS at 4°C before and after the immunofluorescence protocol. Sections were permeabilized and blocked with 0.2% Triton X-100 and 10% normal goat serum (NGS) in 1x PBS for 1 hour at room temperature (RT). Sections were then incubated in primary antibodies prepared with 0.2% Triton X-100 and 10% NGS in 1x PBS overnight at 4°C. Following three 10-minute washes in 1x PBS, sections were incubated in secondary antibodies prepared with 10% NGS in 1x PBS and 1x PBS for 2 hours at RT. Before imaging, sections underwent three 5-minute washes in 1x PBS at RT. Primary antibodies and concentrations used were as follows: mouse anti-PV (1:1000; Millipore Sigma Cat. # P3088), wisteria floribunda agglutinin (WFA; 1:500; Vector Laboratories Cat. # FL-1351-2), mouse anti-CC1 (1:200; Millipore Sigma Cat. # OP80), rat anti-myelin basic protein (MBP; 1:200; Bio-Rad Cat. # MCA409S), and rabbit anti-NeuN (1:1000; Abcam Cat. # ab177487). Secondary antibodies and concentrations used were as follows: goat anti-mouse Alexa Fluor™ 594 (1:1000; Thermo Fisher Cat. # A-11005), goat anti-rabbit Alexa Fluor™ 488 (1:1000; Thermo Fisher Cat. # A-11008), and goat anti-rat Alexa Fluor™ 488 (1:1000; Thermo Fisher Cat. # A-11006).

Microscopy

Following the immunofluorescence protocol, free-floating sections were mounted onto glass slides (Fisher Cat. # 22-034-486) and coverslipped (VWR Cat. # 48404-452) with antifade mounting media (Vector Laboratories Cat. # H-1000-10) and nail polish.

Anatomical landmarks in mice and opossums were used to identify sections containing V1. A combination of images at 2x, 4x, and 10x magnification were taken with an Olympus BX51WI Upright Fluorescence Microscope equipped with a Prior Lumen 200 Fluorescence Illuminator and a Basler Ace Classic camera (acA2040-90um). Multiple sections from each animal were imaged, and two high-magnification images of the cortex (medial and lateral) were taken in each hemisphere from each section where possible. Images were binned every 2 pixels in both directions during acquisition, resulting in a 2048 x 2048 output resolution.

Analysis of single-nucleus transcriptomics data

Mouse snRNA-seq data (30) was obtained from publicly available sources. Raw counts (CSV) were downloaded from the Gene Expression Omnibus (GSE190940; P38 normally reared samples GSM5754081–3) and matched with metadata (including 'Class' and 'Subclass' fields) from processed H5AD objects (https://github.com/-shekharlab/mouseVC). These 3 samples consisted of 2 biological replicates (mice), one of which was split into 2 library replicates. Seurat v4.3.0 (54) objects were constructed from this data and processed in an identical fashion to opossum snRNA-seq data, without the need for gene ortholog mapping or class and subclass assignment (discussed in detail below).

Raw reads (FASTQ) from opossum snRNA-seq libraries were aligned to the *Monodelphis domestica* reference genome ASM229v1 (55) (Ensembl Release 110) with Cell Ranger v6.1.1 (10x Genomics) on the UCLA Hoffman2 computing cluster. The --include-introns option was set to true. The reference package was generated from FASTA and GTF files (Ensembl) using Cell Ranger's mkref function. We immediately noticed a drastic cross-species difference in the 'percentage of reads mapping confidently to transcriptome' (fig. S1A, Original vs. Mouse), with opossum alignments containing considerably more 'intergenic' (not mapped to transcriptome) reads. Examination of the opossum reference genome revealed that many genes contained only coding sequences and lacked introns and 5' or 3' untranslated regions (UTRs). In fact, for several genes we inspected, there were large 'intergenic' read pileups directly downstream of the 3' end of any annotation. We used Samtools to obtain all intergenic reads from the binary alignment map (BAM) output of Cell Ranger and examined their proximity to the nearest gene (same strand) across the entire genome using BEDtools. This revealed a clear genome-wide pileup within 5 kbp downstream of the nearest gene, as well as a peak 1 kbp upstream (fig. S1B). Notably, most intergenic reads were far (>20 kbp) from the nearest gene (fig. S1B). We then created artificial 3' UTRs extending 5 kbp or 10 kbp downstream from the 3' end of annotation for every gene lacking an annotated 3' UTR and farther than 10 kbp upstream of the nearest gene (same strand) and re-aligned reads to this custom genome (56). This provided a small increase in the percentage of reads mapping confidently to transcriptome (fig. S1A), as well as the number of genes and unique molecular identifiers (UMIs, i.e., transcripts) per cell (fig. S1C–D). Although the overall benefit to quality control measures was small, we performed all downstream analyses (including spatial transcriptomics) on data aligned to the genome containing custom 5-kbp 3' UTRs.

All analyses were performed in R (v4.3.0) and primarily based on the Seurat (v4.3.0) package using default parameters, unless otherwise specified. First, gene expression matrices (GEMs) obtained from Cell Ranger were built into Seurat objects. We obtained 4 opossum samples, consisting of 2 biological replicates (opossums) each split into 2 library replicates. As in (30), we first removed cells that did not meet the following criteria: 700 < number of genes < 6500 and number of UMIs < 40,000. We also removed genes that were expressed in 8 or fewer cells. Mitochondrial genes are not annotated in the opossum genomes, so we did not remove cells based on the typical criterion of mitochondrial gene percentage < 1%. This resulted in a total of 32,764 cells and 26,440 genes. Crucially, we then mapped opossum gene names (e.g., RORB) to mouse gene names (e.g., Rorb) if they had a one-to-one ortholog (orthology tables were obtained from Ensembl BioMart). We retained all opossum- and mouse-specific genes in their respective Seurat objects and conducted all downstream analyses on full gene sets, unless otherwise specified.

Cells were normalized using Seurat's SCTransform (v2) workflow (57), and the first 30 principal components (PCs) were used to construct a nearest-neighbor graph. The graph was clustered using the Leiden algorithm (58) (resolution = 1) and visualized by embedding cells in the first two uniform manifold approximation (UMAP) dimensions (59). In opossums, this yielded 27 initial clusters (fig. S1E). All samples were represented in each cluster (fig. S1F–G) and no cluster was composed of more than 15% putative doublets (fig. S1H) identified with an R implementation of Scrublet (60). Therefore, all 27 initial clusters were divided into glutamatergic, GABAergic, and non-neuronal classes on the basis of canonical marker genes (fig. S1I–J). We normalized and clustered again with each class (fig. S1K–M) and examined cross-species correspondence by label transfer (see below) and predicted doublet composition of the resulting clusters (fig. S1N–P). Seven GABAergic clusters (10, 11, 12, 14, 15, 17, and 18) and five non-neuronal clusters (7, 10, 11, 14, 16) were removed from downstream analyses due to size, ambiguous cross-species correspondence, and/or putative doublet fractions greater than 15% (fig. S1N–P).

We then identified cell subclasses within each class. We re-normalized and clustered (as described above) all cells within each class separately using 3 different resolutions in the Leiden algorithm: 0.5, 1, and 2. We examined canonical subclass marker genes (Fig. 1E) and partitioned clusters at each resolution into subclasses. Opossums lacked homologs to the mouse GABAergic subclass Stac and non-neuronal vascular and leptomeningeal cells (VLMC). As discussed further below, glutamatergic intratelencephalic (IT) neurons had particularly ambiguous marker gene expression, so we labeled 4 opossum IT subclasses (IT_A–D; matching the number of mouse IT subclasses) based on unsupervised clustering (fig. S2E–M). We benchmarked all subclass assignments by 1) label transfer (Seurat's MapQuery, see below) from one half of the data to the other and 2) differential gene expression analysis (see below). These analyses yielded measures of subclass consistency and separability in opossums that reflected our observations in mice. We labeled cell types in a similar fashion, but do not discuss them further due to lack of interpretability in opossums.

We performed cross-species integration as part of various analyses (Fig. 1F, 3F). First, we subsetted opossum and mouse Seurat objects such that they contained only one-to-

one orthologs (i.e., they exist in the same gene expression space). We normalized and clustered (as described above) each object separately in this cross-species space, then applied Seurat's SCTransform integration pipeline (61) (SelectIntegrationFeatures, PrepSCTIntegration, FindIntegrationAnchors, and IntegrateData). The 'normalization.method' argument in FindIntegrationAnchors and IntegrateData was set to 'SCT', and 'features.to.integrate' in IntegrateData was passed a vector of orthologous gene names. We performed PCA, constructed and clustered a nearest-neighbor graph, and obtained UMAP dimensions from the resulting integrated Seurat object. To calculate the fractional overlap of subclasses in this integrated space (Fig. 1F), we summed the minimum fraction of each pair of subclasses in each integrated cluster. For example, if two subclasses overlap in a single integrated cluster, the minimum fraction present in said cluster is 1. Analyses in Fig. 3F conducted in the integrated space are discussed below.

We performed within- and cross-species label transfer as part of various analyses (fig. S1N–P, Fig. 2A, C–E, fig. S2C, F–G, J–K, O). In the within-species configuration (fig. S2C, F, G, O), we randomly divided opossum or mouse cells into two Seurat objects and maintained full gene sets (except for fig. S2O). Each object was normalized and clustered (as described above), then used as a reference for label transfer to the other (query) object. The labels (predicted and true) and prediction probabilities were pooled for statistics and visualization. In the cross-species configuration (Fig. 2A, fig. S2G, K), we subsetted genes before normalization and clustering, as in the integration workflow described above. Mouse cells were used as a reference for label transfer to opossum cells. In the cross-species configuration, we performed 100 iterations of label transfer where each subclass was represented by a maximum of 100 random cells. After each iteration, cells were removed from the pool of cells available in the next iteration until no available cells remained for their subclass, when the pool was reset. The results from all iterations were pooled for statistics and visualization. Notably, this balancing strategy improved label transfer for IT cells. In each iteration, we applied Seurat's MapQuery (61) pipeline (FindTransferAnchors and MapQuery). The 'reference.reduction' argument in these functions was set to 'pca'. The 'dims' argument in FindTransferAnchors was set to 1:30, and the 'reduction.model' argument in MapQuery was set to 'umap'. Confusion matrices obtained from comparing predicted (rows) and true (columns) labels were row-normalized.

We examined opossum and mouse IT cells in a 3-dimensional space comprising their first 3 within-species PCs (Fig. 2F–G). These PCs were computed from the full gene sets for each species, but their organization is maintained in ortholog space (Fig. 2H, fig. S3A–C). In order to construct an unbiased PC space, we randomly subsampled each IT subclass to match the count of the least common subclass in each species (L6IT and IT_D) before computing PCs, then projected all IT cells onto these PCs. We used a MATLAB (r2023a) implementation of the Pareto Task Inference method (24) (ParTI; https://github.com/AlonLabWIS/ParTI) to fit a polytope to IT cells in each species-specific PC space (dimensions 1-30). This method first estimates the optimal number of vertices by calculating variance explained by polytopes with a range of vertices (2 to 10) fit by the principal convex hull algorithm (PCHA). Based on the elbow method, both opossum and mouse IT neurons were best fit by a tetrahedron (n=4

vertices). We projected the vertices of these tetrahedra into PC1–3 space (Fig. 2F–H). Separately, we computed pairwise correlations between gene loadings of the top 4 within-species PCs in mice and opossums (Fig. 2I) by subsetting one-to-one orthologs from the loadings vectors. Using a previously established shuffling procedure (62), we validated that opossum and mouse IT cells form a genuine continuum and are not discrete subclasses that appear continuous due to noisy gene expression (fig. S3F–G). We also examined whether the cross-species differences in IT continua were attributable to differences in sequencing depth and transcriptome coverage (fig. S1D, S3H). To do this, we downsampled the mouse dataset to match the UMI count distribution of opossum cells. For each mouse cell, we randomly sampled a target UMI count from the empirical distribution of total UMI counts observed in opossum. Mouse gene expression profiles were then downsampled using multinomial resampling: for each cell, gene-level counts were proportionally reduced to the sampled target UMI total while preserving the original relative expression profile within the cell. This procedure was implemented using the raw count matrix and performed independently for each cell. The resulting downsampled mouse dataset closely matched the overall UMI distribution of the opossum dataset, allowing downstream analyses to control for differences in sequencing depth. (fig. S3H).

We also applied an R-based implementation of weighted gene co-expression network analysis (WGCNA) called hdWGCNA (63) (https://smorabit.github.io/hdWGCNA) to opossum and mouse IT cells (Fig. 2K–N). The standard pipeline yielded 10 modules in mouse IT cells. After subsetting one-to-one orthologs, we projected these modules onto opossum IT cells using ProjectModules and examined their preservation statistics using ModulePreservation. Mouse modules made up of fewer than 50% one-to-one orthologs (M2) or with extremely poor cross-species preservation (M8) were dropped from the analysis. We visualized subclass-specific preservation of modules (Fig. 2L) using ModuleRadarPlot on the post-projection opossum Seurat object. We visualized hub genes from highly preserved modules (Fig. 2M) using HubGeneNetworkPlot and performed gene ontology (GO) analysis and visualization (Fig. 2N) with RunEnrichr and EnrichrBarPlot, respectively.

In Fig. 2P, we computed pairwise, within-species correlations between the top 3,000 highly variable genes in opossum and mouse. Subclass-level average expression profiles were computed from log-normalized UMI counts using Seurat. Gene sets were selected using Seurat's FindVariableFeatures function. For each subclass, average expression was calculated per gene, yielding subclass-by-gene matrices. Correlation matrices were then generated within each species by computing Spearman correlations between subclass expression profiles. To compare differentiation across species, we plotted pairwise combinations of each orthologous subclass mapping (e.g., IT_A / L2/3, IT_B / L4, IT_C / L5IT, IT_D / L6IT) against one another for opossum and mouse.

We performed differential expression testing as part of Fig. 2Q using Seurat's FindMarkers function and applied a log2FC threshold of 0.75 (and previous thresholds of 25% expressing and p-value < 0.05) to DEGs in separate IT_A vs. IT_C and L2/3 vs. L5IT comparisons. We then conducted GO analysis using the R package GOfuncR (https://github.com/sgrote/GOfuncR). Importantly, GO analysis was limited to orthologous DEGs, and we used the *Mus musculus* database available on Bioconductor

(see go_enrich function). We selected 'biological process' terms with family-wise error rates less than 0.05. We summarize top GO biological processes and display several associated genes for opossum- and mouse-specific comparisons. Transcripts encoding cell-cell adhesion molecules and proteins that regulate dendrite development were enriched only in the mouse-specific comparison, while those encoding signaling molecules were enriched in both opossum- and mouse-specific comparisons.

Analysis of spatial transcriptomics data

We obtained spatial GEMs from Stereo-seq output FASTQ files with the Stereo-seq Analysis Workflow (SAW v7.1.2) using manual tissue alignment, aided by the UCLA TCGB. Opossum Stereo-seq reads were aligned to the same custom genome as snRNA-seq reads (discussed above). Mouse Stereo-seq reads were aligned to GRCm39 (GCA_000001635.9). 1 micron is equal to 2 spatial units in Stereo-seq output data. We removed cells that did not meet the following criteria: number of genes > 50 and number of UMIs > 100. We also removed genes that were expressed in 8 or fewer cells.

We plotted the remaining Stereo-seq cells in space and manually selected V1 regions of interest (ROIs) using anatomical landmarks in opossums and mice (fig. S4A–B). These ROIs extended from the top of layer 2/3 to the bottom of layer 6 (beginning of white matter). Because of the curvature of the cortex, we selected multiple adjacent ROIs (i.e., cortical columns) from some sections. Within each species, we combined these columns into a 'pseudocolumn' by scaling every ROI to the width (along any given cortical layer) and height (along the pial-ventricular axis) of the largest ROI. Importantly, the physical height of each column was similar (i.e., the height underwent little scaling; fig. S4A–B). Gene expression patterns and subclass distributions were highly consistent along the width of each ROI (fig. S3B, S4A–B).

We clustered Stereo-seq pseudocolumns obtained from opossums and mice in transcriptomic space and found little structure or coherent marker gene expression patterns (fig. S4C–D). Thus, we integrated Stereo-seq pseudocolumn data with snRNA-seq data across all classes within each species (fig. S4E–F). Despite many Stereo-seq cells appearing well-mixed with snRNA-seq cells in the integrated UMAP space, global differences in gene expression quality drove them to cluster separately when using the standard Leiden algorithm. To transfer labels to Stereo-seq cells, we performed k-nearest neighbor clustering on cells in the integrated UMAP embeddings and assigned them labels if a sufficient percentage of their nearest neighbors were snRNA-seq cells from a single class or subclass. First, we performed class assignment (glutamatergic, GABAergic, or non-neuronal) requiring at least 5% of a Stereo-seq cell's 500 nearest neighbors to belong to a given class (fig. S4E–F). Many cells' nearest neighbors were almost entirely other Stereo-seq cells, or they fell on the boundary between classes, and were removed from subsequent analyses. After removing ambiguous cells, we re-integrated Stereo-seq and snRNA-seq datasets and repeated the label transfer process for all subclasses. Here, we required that at least 25% of a Stereo-seq cell's nearest 100 neighbors belong to a given subclass (fig. S4G–H). Some ambiguous cells were also removed after this step. Critically, Stereo-seq subclasses labeled with this method closely matched the expected spatial distributions, particularly in mice (fig. S3A, C, S4I–

J). Normalized density profiles along the pial-ventricular axis were computed using ggplot2's geom_density function.

In the opossum snRNA-seq data, IT_A is made up of two (resolution = 0.2) or three (resolution = 0.3) Leiden clusters (fig. S2E, I) that were combined during initial subclass assignment. We examined whether these clusters separately explain the upper (above IT_B) and lower (below IT_B) peaks in the pial-ventricular distribution of IT_A observed with Stereo-seq. Critically, at both clustering resolutions, a spatially bimodal cluster of IT_A is present (fig. S3N; cluster 1 at 0.2, cluster 2 at 0.3), in addition to clusters that have a single upper (cluster 2 at 0.2, cluster 1 at 0.3) or lower (cluster 5 at 0.3) peak. We had also previously observed that IT_A subcluster 2 (at resolution 0.2) was enriched for long non-coding RNAs (lncRNAs; fig. S4O), which are typically species-specific and could explain the novel spatial arrangement of IT_A. Instead, lncRNA-enriched opossum IT_A subclusters display a highly mouse-like density profile with a singular peak above IT_B (fig. S4N). Additionally, all IT_A subclusters have the highest gene expression correlations with mouse L2/3 (fig. S4P). Note the graded spatial (peaks in fig. S4N) and transcriptomic (correlation in fig. S4P) relationship with L5IT among these IT_A subclusters, which is explored directly in Fig. 3G–J. Thus, the subclustering of IT_A might be best conceptualized as cell types with differing spatial arrangements, as is the case for mouse L2/3 (*31, 53, 54*).

We examined transcriptomic IT gradients in Stereo-seq cells (Fig. 3F) by subsetting IT subclasses (L2/3 / IT_A, L4 / IT_B, and L5IT / IT_C) from integrated objects before computing PCA. Subclass labels shown for Stereo-seq cells in Fig. 3F were transferred by nearest-neighbor clustering in UMAP space, as described above. We fit triangles to the Stereo-seq cells in PC space using a Python (v3.9.19) implementation of the PCHA (https://github.com/ulfaslak/py_pcha; noc = 3, delta = 0.1). In Fig. 3G–H, cells were colored using a barycentric interpolation scheme based on their position within this triangle. Each triangle vertex was assigned a color, and the color of each cell was computed as a weighted mixture of the vertex colors according to its barycentric coordinates.

Analysis of immunofluorescence data

We used Cellpose (64) (https://github.com/mouseland/cellpose) to produce automated segmentations of PV-, CC1-, and NeuN-positive cells in immunofluorescence images. We seeded masks with a cellprob_threshold of -2.5 and used the cyto2 model with a cell diameter of 8 for PV and NeuN (quantified at 4x magnification) and 20 for CC1 (quantified at 10x). Layer-specific PV cells densities (Fig. 4C) were obtained by manual selection of the neocortical layers using the co-labeled NeuN channel.

Data and materials availability: All code is currently available on GitHub (https://github.com/ryan-gorzek/trachtenberg-lab/tree/main/transcriptomics/V1, https://github.com/ryan-gorzek/trachtenberg-lab/tree/main/transcriptomics/Stereo-seq) and will be curated in a repository specific to this manuscript. snRNA-seq and Stereo-seq data are available on NCBI Gene Expression Omnibus (opossum snRNA-seq processed samples: GSM9038449 and GSM9038450; opossum processed Stereo-seq: GSM9038419; mouse processed Stereo-seq: GSM9038420) by reviewer access code. Immunofluorescence images will be deposited onto Figshare.

Acknowledgments: We thank S. Jain, F. Xie, S.L. Zipursky, and K. Shekhar for analytical support and critical feedback. We thank the UCLA Technology Center for Genomics (TCGB) and Bioinformatics and Translational Pathology Core Laboratory (TPCL), as well as E. Tring, S. Jain, and J. Yoo for experimental support. This work used computational and storage services associated with the Hoffman2 Cluster which is operated by the UCLA Office of Advanced Research Computing's Research Technology Group. Select images were obtained from BioRender. This work was supported by National Institutes of Health grant R01 EY023871 (JTT).

References:

1. V. B. Mountcastle, The columnar organization of the neocortex. *Brain* **120 (Pt 4)**, 701-722 (1997).
2. P. Rakic, Evolution of the neocortex: a perspective from developmental biology. *Nat Rev Neurosci* **10**, 724-735 (2009).
3. L. Krubitzer, J. Kaas, The evolution of the neocortex in mammals: how is phenotypic diversity generated? *Curr Opin Neurobiol* **15**, 444-453 (2005).
4. D. D. O'Leary, S. J. Chou, S. Sahara, Area patterning of the mammalian cortex. *Neuron* **56**, 252-269 (2007).
5. J. H. Kaas, Neocortex in early mammals and its subsequent variations. *Ann N Y Acad Sci* **1225**, 28-36 (2011).
6. J. H. Kaas, The origin and evolution of neocortex: From early mammals to modern humans. *Prog Brain Res* **250**, 61-81 (2019).
7. J. R. Wei *et al.*, Identification of visual cortex cell types and species differences using single-cell RNA sequencing. *Nat Commun* **13**, 6902 (2022).
8. S. L. Ding, Lamination, Borders, and Thalamic Projections of the Primary Visual Cortex in Human, Non-Human Primate, and Rodent Brains. *Brain Sci* **14**, (2024).
9. I. K. Suzuki, T. Hirata, Neocortical neurogenesis is not really "neo": a new evolutionary model derived from a comparative study of chick pallial development. *Dev Growth Differ* **55**, 173-187 (2013).
10. M. Cartmill, Rethinking primate origins. *Science* **184**, 436-443 (1974).
11. S. Herculano-Houzel, C. E. Collins, P. Wong, J. H. Kaas, R. Lent, The basic nonuniformity of the cerebral cortex. *Proc Natl Acad Sci U S A* **105**, 12593-12598 (2008).
12. R. W. Sussman, D. Tab Rasmussen, P. H. Raven, Rethinking primate origins again. *Am J Primatol* **75**, 95-106 (2013).
13. Z. X. Luo, C. X. Yuan, Q. J. Meng, Q. Ji, A Jurassic eutherian mammal and divergence of marsupials and placentals. *Nature* **476**, 442-445 (2011).
14. T. E. Williamson, S. L. Brusatte, G. P. Wilson, The origin and early evolution of metatherian mammals: the Cretaceous record. *Zookeys*, 1-76 (2014).
15. S. Bi *et al.*, An Early Cretaceous eutherian and the placental-marsupial dichotomy. *Nature* **558**, 390-395 (2018).
16. D. M. Kahn, K. J. Huffman, L. Krubitzer, Organization and connections of V1 in Monodelphis domestica. *J Comp Neurol* **428**, 337-354 (2000).
17. T. B. Rowe, T. E. Macrini, Z. X. Luo, Fossil evidence on origin of the mammalian brain. *Science* **332**, 955-957 (2011).

18. J. C. Dooley, J. G. Franca, A. M. Seelke, D. F. Cooke, L. A. Krubitzer, A connection to the past: Monodelphis domestica provides insight into the organization and connectivity of the brains of early mammals. *J Comp Neurol* **521**, 3877-3897 (2013).
19. J. Hahn *et al.*, Evolution of neuronal cell classes and types in the vertebrate retina. *Nature* **624**, 415-424 (2023).
20. D. Kimble, I. Q. Whishaw, Spatial behavior in the Brazilian short-tailed opossum (Monodelphis domestica): comparison with the Norway rat (Rattus norvegicus) in the Morris water maze and radial arm maze. *J Comp Psychol* **108**, 148-155 (1994).
21. M. Wesierska, K. Turlejski, Spontaneous behavior of the gray short-tailed opossum (Monodelphis domestica) in the elevated plus-maze: comparison with Long-Evans rats. *Acta Neurobiol Exp (Wars)* **60**, 479-487 (2000).
22. M. Wesierska, G. Walasek, J. Kilijanek, R. L. Djavadian, K. Turlejski, Behavior of the gray short-tailed opossum (Monodelphis domestica) in the open field and in response to a new object, in comparison with the rat. *Behav Brain Res* **143**, 31-40 (2003).
23. M. Adler, Y. Korem Kohanim, A. Tendler, A. Mayo, U. Alon, Continuum of Gene-Expression Profiles Provides Spatial Division of Labor within a Differentiated Cell Type. *Cell Syst* **8**, 43-52.e45 (2019).
24. M. Adler *et al.*, Emergence of division of labor in tissues through cell interactions and spatial cues. *Cell Rep* **42**, 112412 (2023).
25. Y. Hart *et al.*, Inferring biological tasks using Pareto analysis of high-dimensional data. *Nat Methods* **12**, 233-235, 233 p following 235 (2015).
26. W. Xin *et al.*, Oligodendrocytes and myelin limit neuronal plasticity in visual cortex. *Nature* **633**, 856-863 (2024).
27. C. Stephany, M. G. Frantz, A. W. McGee, Multiple Roles for Nogo Receptor 1 in Visual System Plasticity. *Neuroscientist* **22**, 653-666 (2016).
28. S. J. Kuhlman *et al.*, A disinhibitory microcircuit initiates critical-period plasticity in the visual cortex. *Nature* **501**, 543-546 (2013).
29. T. Pizzorusso *et al.*, Reactivation of ocular dominance plasticity in the adult visual cortex. *Science* **298**, 1248-1251 (2002).
30. N. Gogolla, P. Caroni, A. Lüthi, C. Herry, Perineuronal nets protect fear memories from erasure. *Science* **325**, 1258-1261 (2009).
31. S. Cheng *et al.*, Vision-dependent specification of cell types and function in the developing cortex. *Cell* **185**, 311-327.e324 (2022).
32. X. Y. Ji *et al.*, Thalamocortical Innervation Pattern in Mouse Auditory and Visual Cortex: Laminar and Cell-Type Specificity. *Cereb Cortex* **26**, 2612-2625 (2016).
33. C. K. Pfeffer, M. Xue, M. He, Z. J. Huang, M. Scanziani, Inhibition of inhibition in visual cortex: the logic of connections between molecularly distinct interneurons. *Nat Neurosci* **16**, 1068-1076 (2013).
34. J. L. Chen, S. Carta, J. Soldado-Magraner, B. L. Schneider, F. Helmchen, Behaviour-dependent recruitment of long-range projection neurons in somatosensory cortex. *Nature* **499**, 336-340 (2013).
35. M. M. Karnani *et al.*, Cooperative Subnetworks of Molecularly Similar Interneurons in Mouse Neocortex. *Neuron* **90**, 86-100 (2016).

36. A. J. Keller *et al.*, A Disinhibitory Circuit for Contextual Modulation in Primary Visual Cortex. *Neuron* **108**, 1181-1193.e1188 (2020).
37. A. C. Yang *et al.*, A human brain vascular atlas reveals diverse mediators of Alzheimer's risk. *Nature* **603**, 885-892 (2022).
38. T. E. Bakken *et al.*, Comparative cellular analysis of motor cortex in human, marmoset and mouse. *Nature* **598**, 111-119 (2021).
39. A. Chen *et al.*, Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. *Cell* **185**, 1777-1792.e1721 (2022).
40. D. G. Southwell, R. C. Froemke, A. Alvarez-Buylla, M. P. Stryker, S. P. Gandhi, Cortical plasticity induced by inhibitory neuron transplantation. *Science* **327**, 1145-1148 (2010).
41. L. Lupori *et al.*, A comprehensive atlas of perineuronal net distribution and colocalization with parvalbumin in the adult mouse brain. *Cell Rep* **42**, 112788 (2023).
42. G. Brückner *et al.*, Cortical perineuronal nets in the gray short-tailed opossum (Monodelphis domestica): a distribution pattern contrasting with that shown in placental mammals. *Anat Embryol (Berl)* **197**, 249-262 (1998).
43. J. T. Trachtenberg, M. P. Stryker, Rapid anatomical plasticity of horizontal connections in the developing visual cortex. *J Neurosci* **21**, 3476-3482 (2001).
44. H. A. Swadlow, A. G. Gusev, The impact of 'bursting' thalamic impulses at a neocortical synapse. *Nat Neurosci* **4**, 402-408 (2001).
45. I. Onorato *et al.*, Distinct roles of PV and Sst interneurons in visually induced gamma oscillations. *Cell Rep* **44**, 115385 (2025).
46. J. Jackson, I. Ayzenshtat, M. M. Karnani, R. Yuste, VIP+ interneurons control neocortical activity across brain states. *J Neurophysiol* **115**, 3008-3017 (2016).
47. K. A. Ferguson *et al.*, VIP interneurons regulate cortical size tuning and visual perception. *Cell Rep* **42**, 113088 (2023).
48. Y. Fu *et al.*, A cortical circuit for gain control by behavioral state. *Cell* **156**, 1139-1152 (2014).
49. H. J. Pi *et al.*, Cortical interneurons that specialize in disinhibitory control. *Nature* **503**, 521-524 (2013).
50. D. J. Millman *et al.*, VIP interneurons in mouse primary visual cortex selectively enhance responses to weak but specific stimuli. *Elife* **9**, (2020).
51. C. E. Yaeger *et al.*, A dendritic mechanism for balancing synaptic flexibility and stability. *Cell Rep* **43**, 114638 (2024).
52. A. W. McGee, Y. Yang, Q. S. Fischer, N. W. Daw, S. M. Strittmatter, Experience-driven plasticity of visual cortex limited by myelin and Nogo receptor. *Science* **309**, 2222-2226 (2005).
53. F. Xie *et al.*, Spatial profiling of the interplay between cell type- and vision-dependent transcriptomic programs in the visual cortex. *Proc Natl Acad Sci U S A* **122**, e2421022122 (2025).
54. J. Yoo *et al.*, Establishing a continuum of cell types in the visual cortex. *bioRxiv*, 2025.2009.2022.677893 (2025).

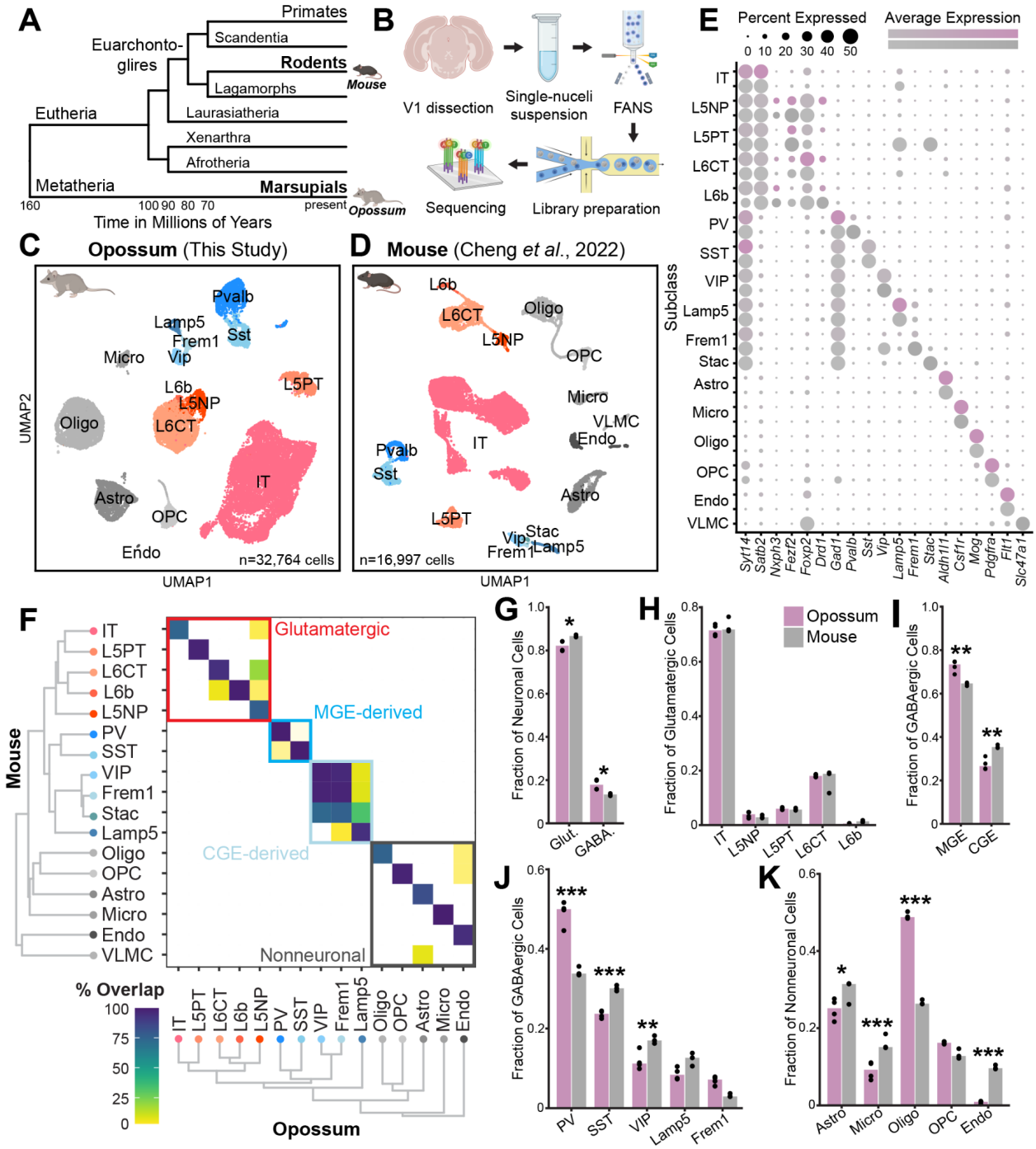

Figure 1. Cross-species transcriptomic comparison in primary visual cortex (V1) of the metatherian gray short-tailed opossum (*Monodelphis domestica*) and eutherian mouse.

(**A**) Cladogram of therian mammalian evolution. (**B**) Experimental workflow of single-nucleus RNA sequencing (snRNA-seq) in opossum V1. (**C**) Opossum cells embedded in uniform manifold approximation (UMAP) dimensions, labeled by subclass. (**D**) Same as (C) but mouse cells from our previous study (*30*). E, Canonical marker gene expression used in part to identify opossum subclasses. (**F**) Mouse and opossum subclass-level dendrograms and co-clustering in

an integrated cross-species transcriptomic space. (**G–K**) Relative proportions of some transcriptomic groupings (G, class; H, glutamatergic subclass; I, GABAergic origin structure; J, GABAergic subclass; K, non-neuronal subclass) are significantly different between species (ANOVA followed by two-sided Tukey's HSD test and Bonferroni correction; *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$; table S1). Opossum data points include two biological replicates with two technical replicates each, mouse data points include three biological replicates.

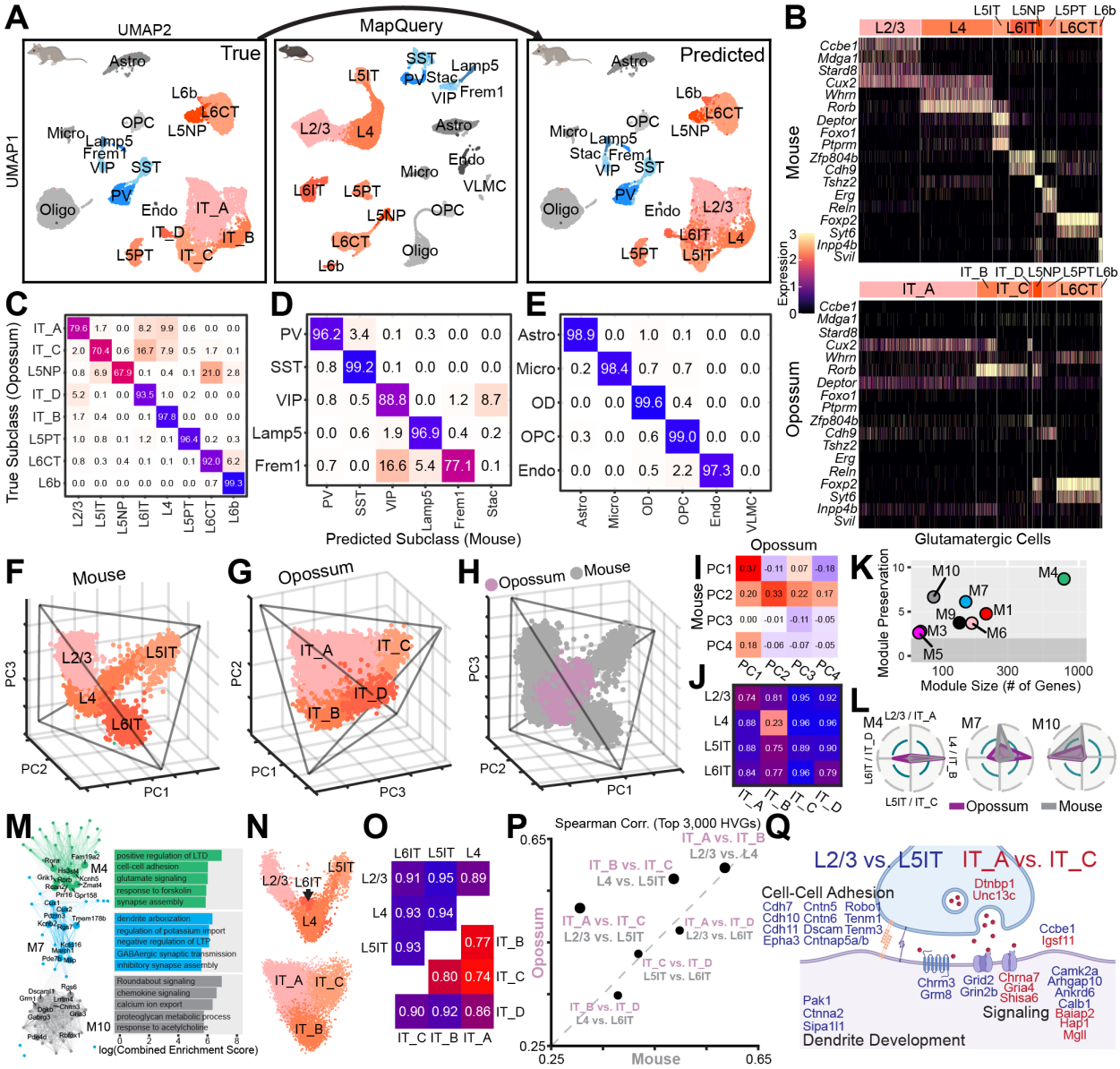

Figure 2. Intratelencephalic (IT) glutamatergic neurons have poor cross-species correspondence.

(**A**) Opossum cells, shown by subclass in UMAP space (left) were assigned mouse subclass labels (right) using mouse cells (center) as a reference in Seurat's MapQuery pipeline. (**B**) Canonical glutamatergic subclass marker gene expression in mouse and opossum glutamatergic cells (columns) organized by subclass. (**C–E**) Confusion matrices (row-normalized) showing the relationship between true and predicted opossum subclass labels described in (A) for glutamatergic (C), GABAergic (D), and non-neuronal (E) subclasses, which were classified independently. (**F–G**) IT neurons in mice (F) and opossums (G) form tetrahedral gradients when projected onto their top three principal components (PCs) computed from full gene sets within each species. This organization is maintained in ortholog space (fig. S3A–B). Tetrahedrons were fit using a published Pareto task inference method. (**H**) Opossum IT neurons

projected into mouse IT PC space. (**I**) Correlation matrix between the loadings (gene weights) of the top four within-species PCs from mouse and opossum IT neurons. (**J**) Average silhouette distances between cell embeddings of mouse and opossum IT subclasses in (H). (**K**) Weighted gene co-expression network analysis (WGCNA) modules identified in mouse IT neurons and projected onto opossum IT neurons, shown by size and cross-species correspondence (preservation, see Materials and Methods). (**I**) Subclass specificity of the three most highly preserved WGCNA modules (M4, M7, and M10) in mouse and opossums. (**M**) Top 20 hub genes associated with the modules in (L) shown as network diagrams and module-enriched gene ontology (GO) terms. (**N**) View of a single plane in species-specific mouse (top) and opossum (bottom) IT PC spaces containing gradients of corresponding subclasses. Opossum IT_D is fully obscured from this view. (**O**) Average silhouette distances between cell embeddings from mouse and opossum IT subclasses in species-specific (three-dimensional) PC spaces shown in (F–G). (**P**) Spearman correlations between the top 3,000 highly variable genes in pairwise, within-species combinations of IT subclasses. Show as mouse vs. opossum against identity line to emphasize cross-species differences. (**Q**) DEGs and enriched GO terms in mouse L2/3 vs. L5IT (blue) and opossum IT_A vs. IT_C (red) shown in their expected cellular compartment.

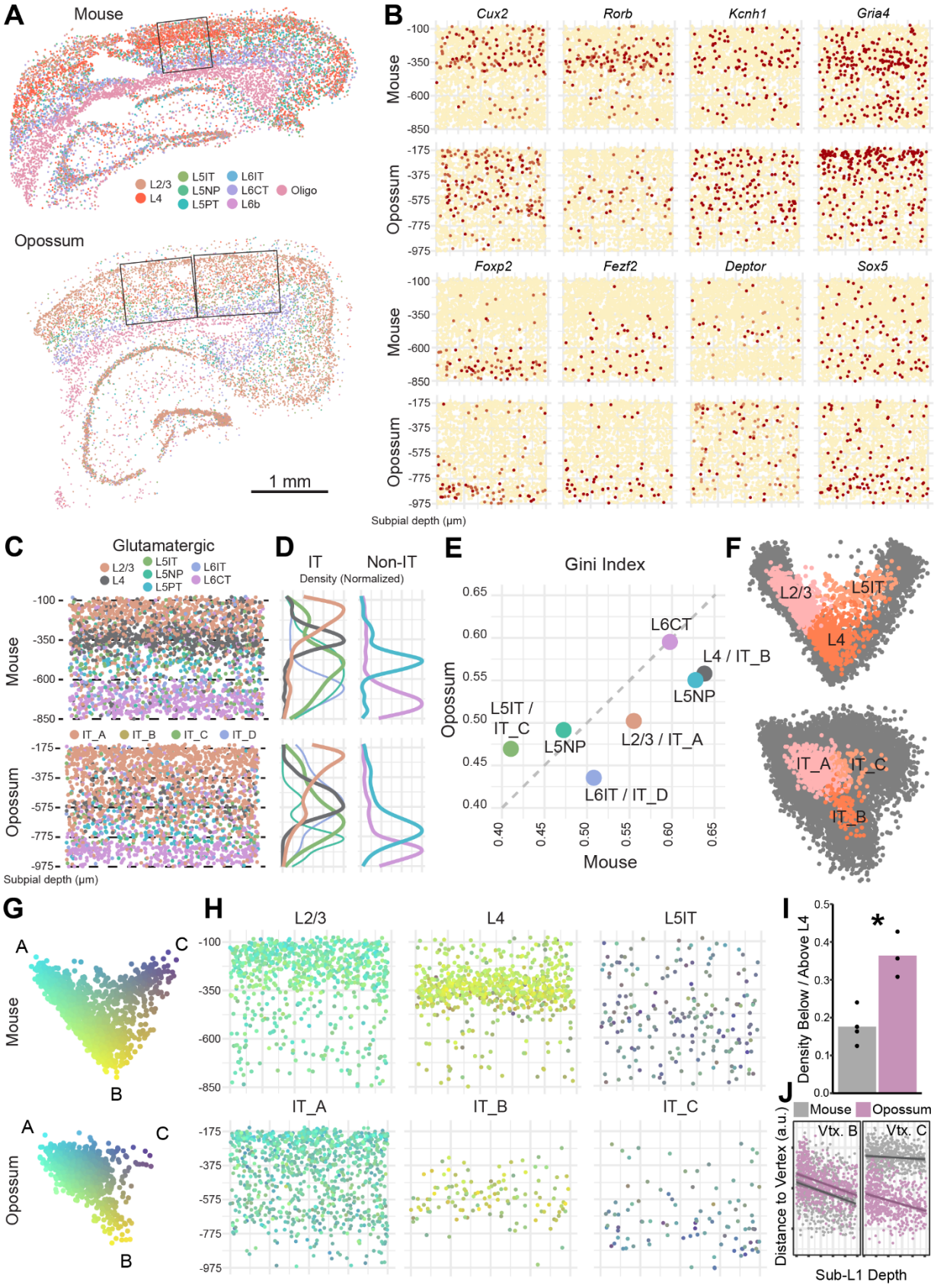
A
Mouse
L2/3
L4
L5IT
L5NP
L5PT
L6IT
L6CT
L6b
Oligo
Opossum
1 mm
B
Cux2
Rorb
Kcnh1
Gria4
Foxp2
Fezf2
Deptor
Sox5
Mouse
Opossum
Subpial depth (µm)
C
Glutamatergic
IT_A
IT_B
IT_C
IT_D
Subpial depth (µm)
D
IT
Non-IT
Density (Normalized)
E
Gini Index
L6CT
L4 / IT_B
L5NP
L5IT / IT_C
L5NP
L2/3 / IT_A
L6IT / IT_D
Opossum
Mouse
F
L2/3
L5IT
L4
IT_A
IT_C
IT_B
G
A
B
C
Mouse
Opossum
H
L2/3
L4
L5IT
IT_A
IT_B
IT_C
I
Density Below / Above L4
J
Mouse
Opossum
Vtx. B
Vtx. C
Distance to Vertex (a.u.)
Sub-L1 Depth

Figure 3. Stereo-seq reveals poor laminar specificity of IT_A in opossums.

(**A**) Representative Stereo-seq sections from mouse (top) and opossum (bottom), displaying all glutamatergic subclasses and oligodendrocytes. Subclass labels were predicted from species-specific snRNA-seq atlases using cross-species integration and co-clustering (see Materials and Methods). V1 was identified using cytoarchitectural features and the prominent presence of layer 4. Boxes indicate regions of interest (ROIs) from these sections used for downstream analyses. (**B**) Spatially variable genes include canonical upper- and lower-layer glutamatergic subclass markers. (**C**) Distribution of glutamatergic subclasses from pooled V1 regions. (**D**) Subpial densities of IT (left) and non-IT (right) glutamatergic subclasses, collapsed along the x-axis in (C). (**E**) Gini index of subpial position for glutamatergic subclasses in mice and opossums. (**F**) IT cells identified in Stereo-seq (colored) shown against snRNA-seq IT cells (gray) in pooled PC space (similar to Fig. 2N). (**G**) Stereo-seq IT cells in PC space as in (F), pseudocolored by distance from the vertices of a fitted triangle. (**H**) Pseudocolored cells from (G) shown in physical space as in (C), grouped by subclass labels shown in (F). (**I**) Ratio of densities of L2/3 (mouse) and IT_A (opossum) cells below and above the species-specific mean subpial depth of L4 / IT_B. This ratio (below / above) is significantly higher in opossums (Student's t-test; *$p < 0.05$; table S1). Opossum data points include three ROIs from two sections, and mouse data points include four ROIs from three sections. (**J**) Relationships between the depth of L2/3 (mouse) and IT_A (opossum) cells and their distance from B (left) and C (right) vertices in PC space (g) for the first 250 μm below layer 1. Lines are simple linear regression fits.

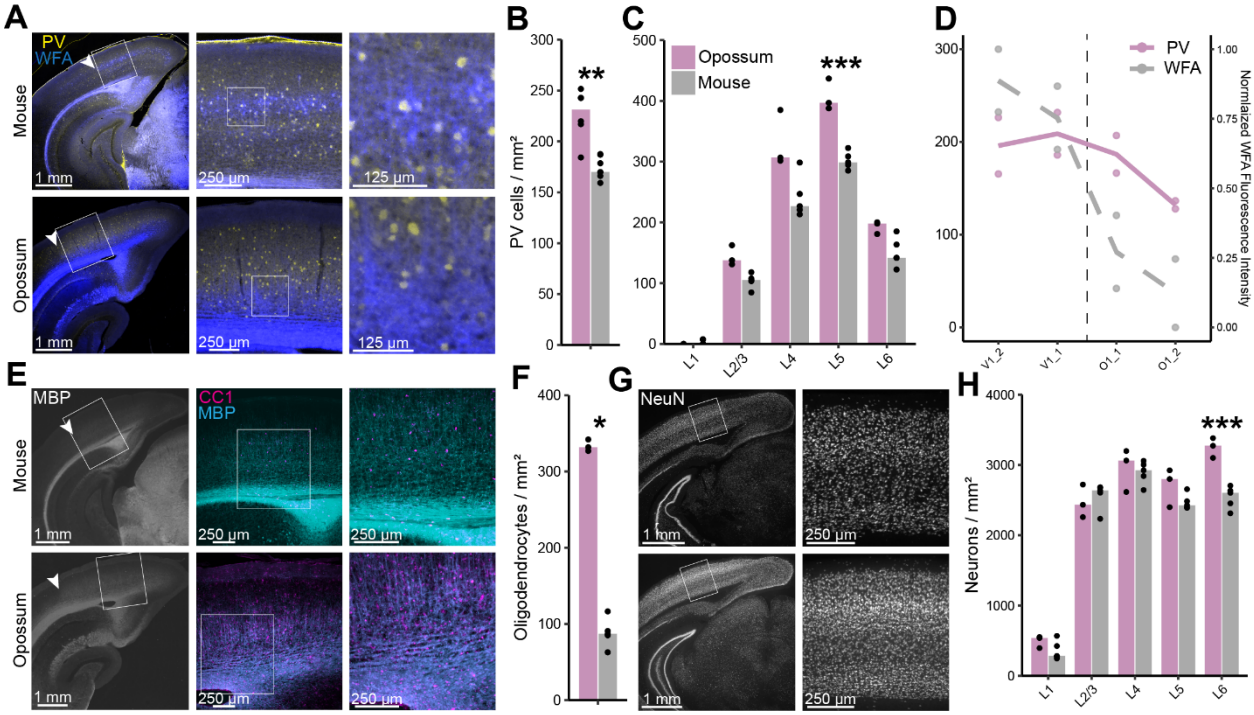

Figure 4. Immunolabeling confirms changes in PV and oligodendrocyte densities and reveals altered peri-neuronal net localization in opossums

(**A**) Parvalbumin (PV; blue) and Wisteria floribunda agglutinin (WFA; yellow) labeling in representative mouse and opossum coronal sections. Arrows mark the lateral edge of V1 and boxes indicate the magnified area shown on the right. (**B**) PV-positive cell densities in V1 are significantly different (ANOVA followed by two-sided Tukey's HSD test and Bonferroni correction; *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$; table S1; data points are single fields of view from separate sections). Opossum and mouse data points include five sections from each species. (**C**) layer-specific PV cell densities (identified cytoarchitecturally with NeuN counterstaining) are significantly different in layer 5. (**D**) PV density and WFA fluorescence intensity in bins located inside and outside of V1. Bins are adjacent and surround the lateral edge of V1. (**E**) CC1 (magenta) and myelin basic protein (MBP; cyan) labeling as in (A). (**F**) Oligodendrocyte (CC1-positive cell) densities in V1 are significantly different. Opossum and mouse data points include three and four sections, respectively. (**G**) NeuN labeling. (**H**) layer-specific neuron densities in V1 reveal significantly higher packing density in opossum layer 6. Opossum and mouse data points include three and five sections, respectively.

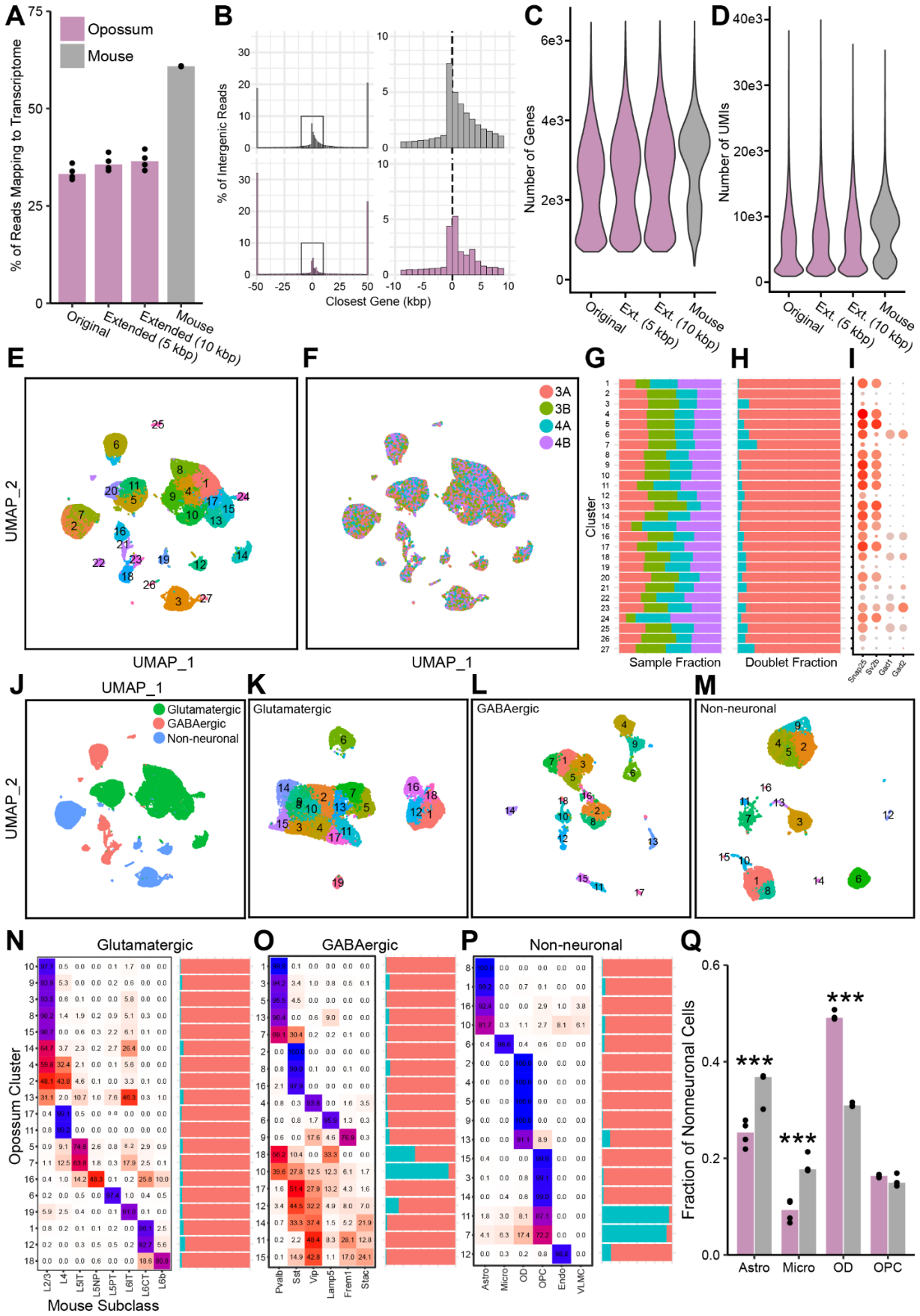
A
Opossum
Mouse
% of Reads Mapping to Transcriptome
Original
Extended (5 kbp)
Extended (10 kbp)
Mouse
B
% of Intergenic Reads
Closest Gene (kbp)
C
Number of Genes
Original
Ext. (5 kbp)
Ext. (10 kbp)
Mouse
D
Number of UMIs
Original
Ext. (5 kbp)
Ext. (10 kbp)
Mouse
E
UMAP_2
UMAP_1
F
3A
3B
4A
4B
UMAP_1
G
Cluster
Sample Fraction
H
Doublet Fraction
I
Snap25
Syp6
Gad1
Gad2
J
UMAP_1
UMAP_2
Glutamatergic
GABAergic
Non-neuronal
K
Glutamatergic
L
GABAergic
M
Non-neuronal
N
Glutamatergic
Opossum Cluster
L2/3
L4
L5IT
L5NP
L5PT
L6IT
L6CT
L6b
Mouse Subclass
O
GABAergic
Pvalb
Sst
Vip
Lamp5
Frem1
Stac
P
Non-neuronal
Astro
Micro
OD
OPC
Endo
VLMC
Q
Fraction of Nonneuronal Cells
***
***
***
Astro
Micro
OD
OPC

Figure S1. Opossum snRNA-seq data preprocessing.

(**A**) Percentage of snRNA-seq reads mapping confidently to the transcriptome (exons, introns, and untranslated regions) in opossums and mice. Opossum genes lacking 3' untranslated regions (UTRs) were extended 5 or 10 kilobase pairs (kbp) in the 3' direction. (**B**) Intergenic (mapping confidently to regions outside of exons, introns, or UTRs) read locations relative to the closest gene (negative is upstream or 5', positive is downstream or 3') in mice and opossums. (**C–D**) Number of genes (C) and unique molecular identifiers (UMIs, i.e., transcripts; D) per cell in mice and opossums (original and extended genome configurations). (**E**) Unsupervised Leiden clustering of all opossum cells passing initial quality control (see Materials and Methods) in UMAP space. (**F–G**) UMAP (F) and cluster-wise (G) representations of samples in opossum snRNA-seq data. Numbers (3 and 4) represent biological replicates and letters (A and B) represent technical replicates. (**H**) Cluster-wise representations of putative doublets (see Materials and Methods). (**I**) Canonical neuronal (*Snap25*), glutamatergic (*Sv2b*), and GABAergic (*Gad1*, *Gad2*) marker genes used to divide unsupervised clusters into cell classes. (**J**) All opossum cells divided into cell classes. (**K–M**) Unsupervised clustering of glutamatergic (K), GABAergic (L), and non-neuronal (M) cells prior to removal of ambiguous or doublet-enriched clusters. (**N–P**) Cluster-wise cross-species correspondence (left) and doublet representation (right) for glutamatergic (N), GABAergic (O), and non-neuronal (P) classes used to guide removal of ambiguous/artifact clusters. (**Q**) Relative proportions of non-neuronal subclasses without considering endothelial cells, which are uncommon in opossums (ANOVA followed by two-sided Tukey's HSD test and Bonferroni correction; *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$; table S1). Related to Fig. 1K.

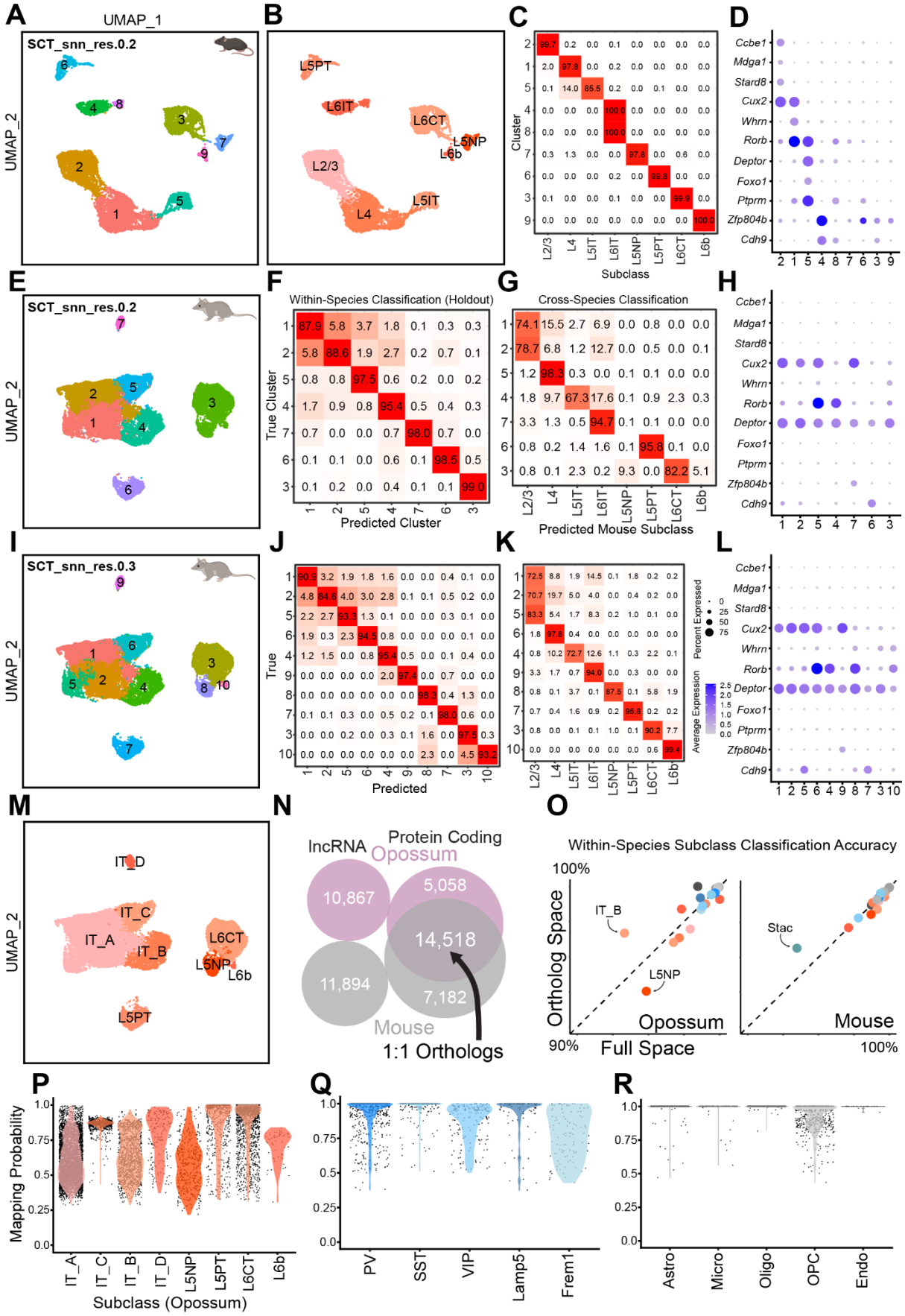
A
UMAP_1
SCT_snn_res.0.2
UMAP_2
B
L5PT
L6IT
L6CT
L5NP
L6b
L2/3
L4
L5IT
C
Cluster
Subclass
D
Ccbe1
Mdga1
Stard8
Cux2
Whrn
Rorb
Deptor
Foxo1
Ptprm
Zfp804b
Cdh9
E
SCT_snn_res.0.2
UMAP_2
F
Within-Species Classification (Holdout)
True Cluster
Predicted Cluster
G
Cross-Species Classification
Predicted Mouse Subclass
H
I
SCT_snn_res.0.3
UMAP_2
J
True
Predicted
K
L
Percent Expressed
Average Expression
M
IT_D
IT_C
IT_A
IT_B
L6CT
L5NP
L6b
L5PT
UMAP_2
N
lncRNA
Protein Coding
Opossum
10,867
5,058
14,518
11,894
7,182
Mouse
1:1 Orthologs
O
Within-Species Subclass Classification Accuracy
100%
Ortholog Space
IT_B
L5NP
Opossum
90%
Full Space
Stac
Mouse
100%
P
Mapping Probability
Subclass (Opossum)
Q
PV
SST
VIP
Lamp5
Frem1
R
Astro
Micro
Oligo
OPC
Endo

Figure S2. Semi-supervised labeling of opossum intratelencephalic (IT) subclasses and cross-species correspondence.

(**A**) Mouse glutamatergic cells in UMAP space, clustered at low resolution (0.2) with the Leiden algorithm. (**B**) Mouse glutamatergic cells in UMAP space, labeled by subclass (*30*). (**C**) Confusion matrix between low-resolution unsupervised clustering and subclass labels in mice. (**D**) Canonical glutamatergic subclass marker gene expression in unsupervised clusters shown in (A). (**E**) Opossum glutamatergic cells in UMAP space, clustered at low resolution (0.2). (**F**) Confusion matrix derived from 50% train-test splits of unsupervised clusters shown in (E). (**G**) Confusion matrix derived from cross-species label transfer (see Materials and Methods) between mouse glutamatergic subclasses and unsupervised clusters shown in (E). (**H**) Canonical glutamatergic subclass marker gene expression in unsupervised clusters shown in (E). (**I–L**) Same as (E–H), but with Leiden resolution set to 0.3. (**M**) Final glutamatergic subclass assignment in opossums. Clusters 1 and 2 in (E) were merged to form IT_A. L6CT, L5NP, and L6b correspond to clusters 3, 8, and 10 in (I). (**N**) Composition of the opossum genome. Long noncoding RNAs (lncRNAs) are species-specific and one-to-one orthologs are exclusively protein-coding genes. (**O**) Classification accuracy of 50% train-test splits for each subclass in species-specific (full) and orthologous gene expression space in opossums and mice. (**P–R**) Cross-species mapping quality (obtained from MapQuery, see Materials and Methods) scores for opossum cells shown by glutamatergic (P), GABAergic (Q), and non-neuronal (R) subclasses. Related to Fig. 2A, C–E.

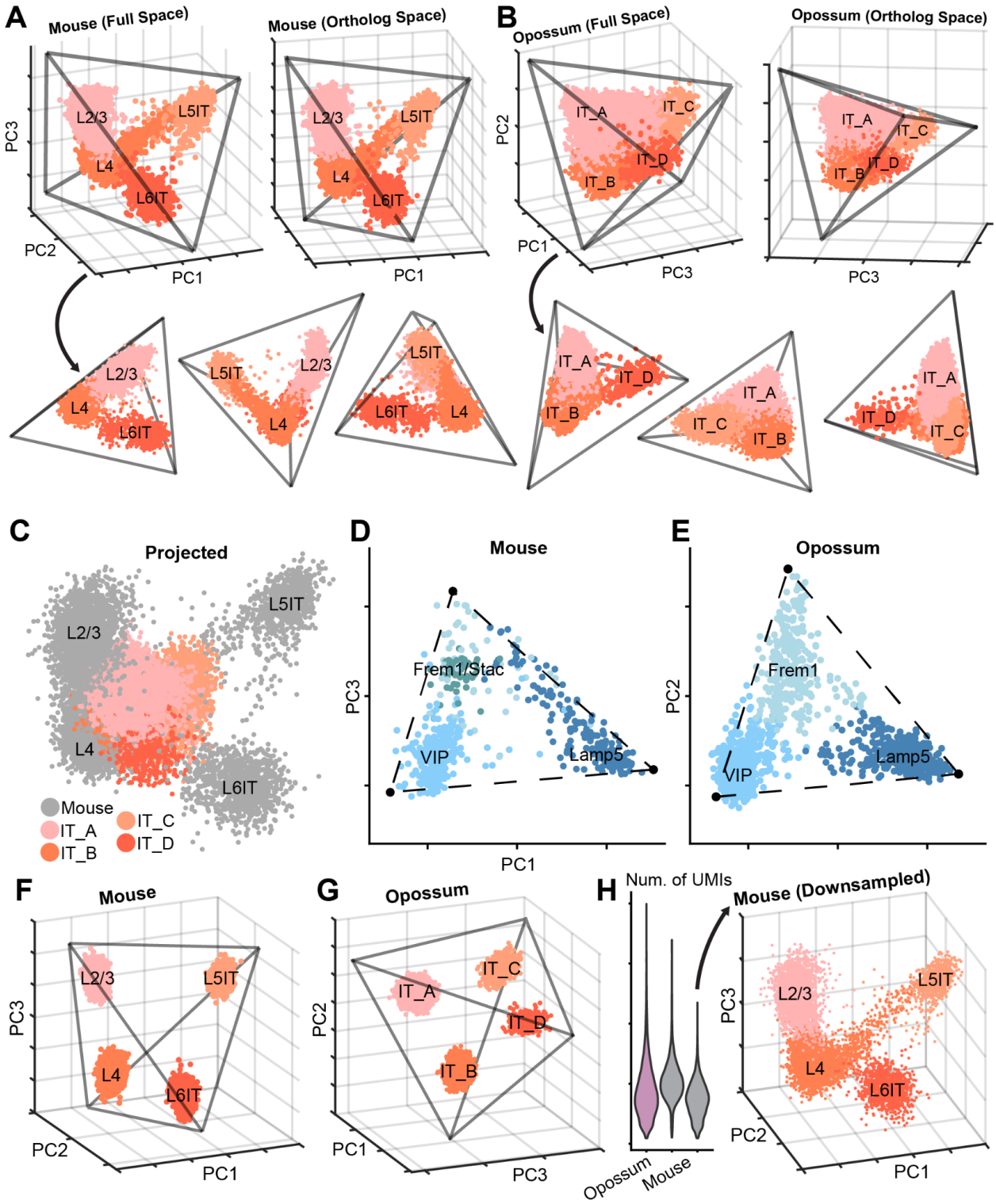

Figure S3. IT principal component (PC) spaces, other gradients, and quality controls.

(**A–B**) Tetrahedral IT gradients in mouse (A) and opossum (B) principal component spaces, shown from various angles. Tetrahedrons were fit using established algorithms (see Materials

and Methods). Related to Fig. 2F–G. (**C**) Opossum IT neurons (colored by subclass) projected into mouse IT (gray) PC space. IT_B is obscured by L4 due to considerable overlap. (**D–E**) GABAergic neurons derived from the caudal ganglionic eminence (CGE), which form a continuum in gene expression space, display similar triangular gradients in mice (D) and opossums (E). (**F–G**) Randomly shuffling gene expression within subclasses (see Materials and Methods) confirms that mouse (F) and opossum (G) IT cells form continua that are not driven by noise (see Materials and Methods). (**H**) The mouse IT continuum retains its structure after downsampling UMI (transcript) counts to match opossum levels (see Materials and Methods).

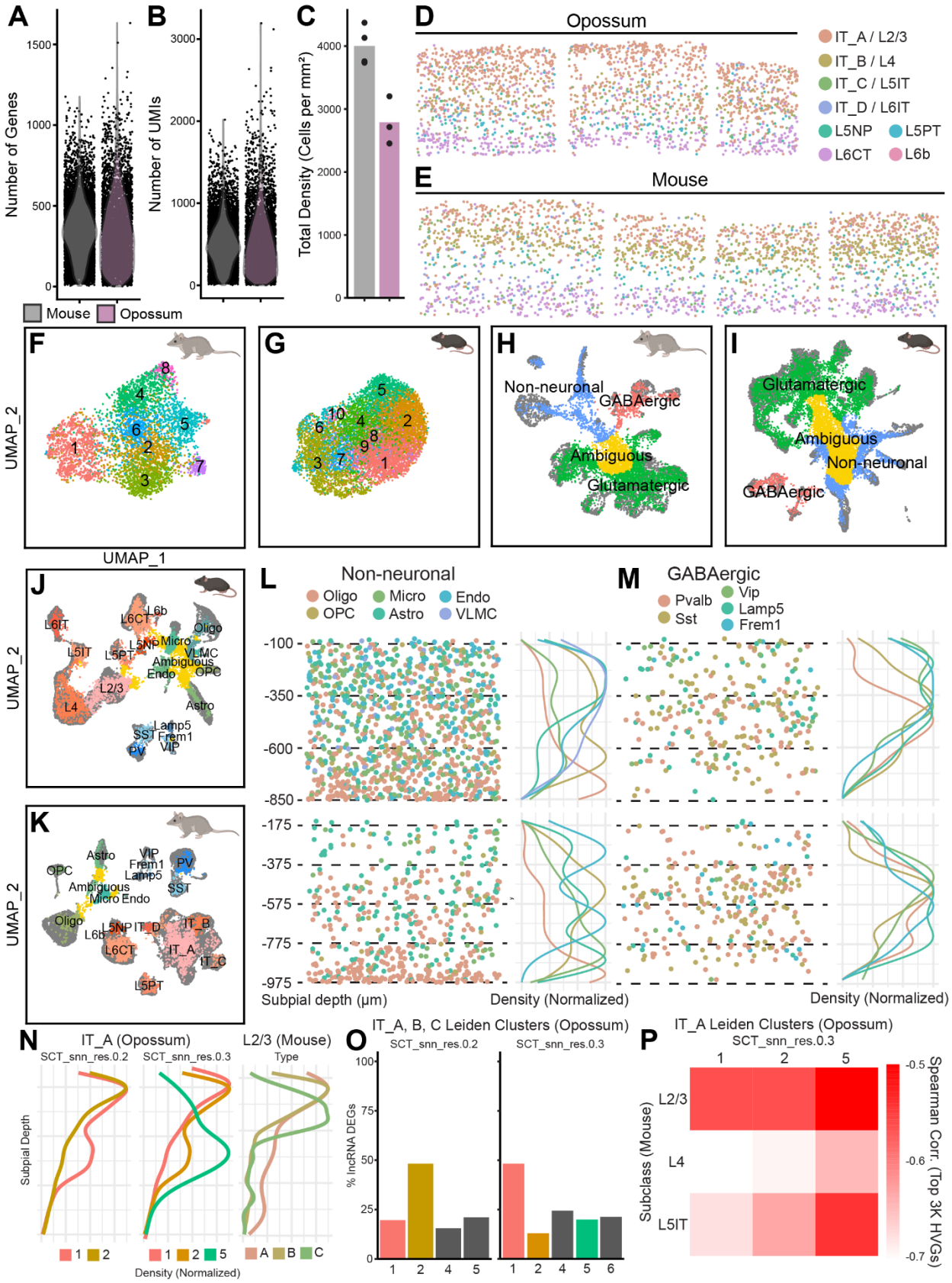
A
Number of Genes
Mouse
Opossum
B
Number of UMIs
C
Total Density (Cells per mm²)
D
Opossum
IT_A / L2/3
IT_B / L4
IT_C / L5IT
IT_D / L6IT
L5NP
L5PT
L6CT
L6b
E
Mouse
F
UMAP_2
UMAP_1
G
H
Non-neuronal
GABAergic
Ambiguous
Glutamatergic
I
Glutamatergic
Ambiguous
Non-neuronal
GABAergic
J
L6IT
L6CT
L6b
Oligo
L5IT
L5NP
L5PT
Micro
VLMC
Ambiguous
Endo
OPC
L2/3
L4
Astro
Lamp5
SST
Frem1
Pv
VIP
K
Astro
VIP
Frem1
PV
OPC
Lamp5
Ambiguous
SST
Micro
Endo
Oligo
L6b
L5NP
IT_D
IT_B
L6CT
IT_A
IT_C
L5PT
L
Non-neuronal
Oligo
Micro
Endo
OPC
Astro
VLMC
Subpial depth (µm)
Density (Normalized)
M
GABAergic
Pvalb
Vip
Sst
Lamp5
Frem1
Density (Normalized)
N
IT_A (Opossum)
SCT_snn_res.0.2
SCT_snn_res.0.3
L2/3 (Mouse)
Type
Subpial Depth
Density (Normalized)
O
IT_A, B, C Leiden Clusters (Opossum)
SCT_snn_res.0.2
SCT_snn_res.0.3
% lncRNA DEGs
P
IT_A Leiden Clusters (Opossum)
SCT_snn_res.0.3
Subclass (Mouse)
L2/3
L4
L5IT
Spearman Corr. (Top 3K HVGs)

Figure S4. Stereo-seq preprocessing and non-glutamatergic cell subclass distributions.

(**A–B**) Number of genes (A) and UMIs (B) per cell in all Stereo-seq cells from mice and opossums. (**C**) Total cell density (cells per mm^2) for mouse and opossum Stereo-seq data prior to processing. (**D–E**) Individual cortical columns extracted from opossum (D) and mouse (E) Stereo-seq sections. Only glutamatergic cells are shown, colored by subclass. (**F–G**) Unsupervised clustering of opossum (F) and mouse (G) Stereo-seq cells. (**H–I**) UMAP representation of integrated snRNA-seq (gray) and Stereo-seq (colored by class) cells in opossums (H) and mice (I). Stereo-seq class labels were assigned by nearest-neighbor clustering in this UMAP space (see Materials and Methods). (**J–K**) Same as (H–I), but with Stereo-seq cells labeled by subclass. Cells labeled 'Ambiguous' in (H–I) were removed before the integration step whose results are shown here. (**L–M**) Distribution of non-neuronal (L) and GABAergic (M) subclasses from pooled V1 regions. Related to Fig. 3C–D. (**N**) Spatial distributions of IT_A subclusters at Leiden resolution 0.2 (left; fig. S2E–H) and 0.3 (center; fig. S2I–L) and mouse L2/3 types (right). (**O**) Fraction of IT subcluster DEGs that are long non-coding RNAs (lncRNAs) at Leiden resolution 0.2 (left; fig. S2E–H) and 0.3 (right; fig. S2I–L). See Methods for further discussion. (**P**) Spearman correlation between the top 3,000 highly variable genes (HVGs) of opossum IT_A subclusters (Leiden resolution 0.3) and mouse IT subclasses. See Methods for further discussion.

Figure	Panel	Test	Statistic	P-Value
1	G	ANOVA (fraction ~ species * cell.class)	F = 19.74 (species:cell.class)	0.00125
		Tukey MC T-test (Mouse:Glut. - Opossum:Glut.)		0.0438522
		Tukey MC T-test (Mouse:GABA. - Opossum:GABA.)		0.0438522
	H	ANOVA (fraction ~ species * cell.subclass)	F = 0.917 (species:cell.subclass)	0.47
	I	ANOVA (fraction ~ species * cell.origin)	F = 48.46 (species:cell.origin)	0.0000389
		Tukey MC T-test (Mouse:MGE - Opossum:MGE)		0.002813
		Tukey MC T-test (Mouse:CGE - Opossum:CGE)		0.002813
	J	ANOVA (fraction ~ species * cell.subclass)	F = 47.93 (species:cell.subclass)	2.07E-13
		Tukey MC T-test (Mouse:PV - Opossum:PV)		0
		Tukey MC T-test (Mouse:SST - Opossum:SST)		0.0002148
		Tukey MC T-test (Mouse:VIP - Opossum:VIP)		0.0041064
		Tukey MC T-test (Mouse:Lamp5 - Opossum:Lamp5)		0.1066756
		Tukey MC T-test (Mouse:Frem1 - Opossum:Frem1)		0.1007594
	K	ANOVA (fraction ~ species * cell.subclass)	F = 92.14 (species:cell.subclass)	2E-16
		Tukey MC T-test (Mouse:Astro - Opossum:Astro)		0.0181231
		Tukey MC T-test (Mouse:Micro - Opossum:Micro)		0.0001084
		Tukey MC T-test (Mouse:Oligo - Opossum:Oligo)		0
		Tukey MC T-test (Mouse:OPC - Opossum:OPC)		0.4014916
		Tukey MC T-test (Mouse:Endo - Opossum:Endo)		0.0000018
S1	Q	ANOVA (fraction ~ species * cell.subclass)	F = 70.95 (species:cell.subclass)	7.76E-11
		Tukey MC T-test (Mouse:Astro - Opossum:Astro)		0.0001341
		Tukey MC T-test (Mouse:Micro - Opossum:Micro)		0.0001052
		Tukey MC T-test (Mouse:Oligo - Opossum:Oligo)		0
		Tukey MC T-test (Mouse:OPC - Opossum:OPC)		0.9983726
3	I	Student's T-test (Two-sided; Mouse - Opossum)	t = -4.4619	0.01259
4	A	ANOVA (density ~ species * layer)	F = 5.387 (species:layer)	0.00216
		Tukey MC T-test (Mouse:L1 - Opossum:L1)		0.9878207
		Tukey MC T-test (Mouse:L2/3 - Opossum:L2/3)		0.9990372
		Tukey MC T-test (Mouse:L4 - Opossum:L4)		0.9999628
		Tukey MC T-test (Mouse:L5 - Opossum:L5)		0.7238365
		Tukey MC T-test (Mouse:L6 - Opossum:L6)		0.0002752
	B	Wilcoxon Rank-Sum	W = 40	0.004662
	C	Wilcoxon Rank-Sum	W = 15	3.57E-02

Table S1. Statistical tests, test statistics, and p-values for all main and supplementary figures.