Cell-type specific transcriptomic signatures of neocortical circuit organization and their relevance to autism

Anthony J. Moussa and Jason C. Wester

Department of Neuroscience, The Ohio State University College of Medicine, Columbus, OH 43210

Correspondence: jason.wester@osumc.edu

ABSTRACT

A prevailing challenge in neuroscience is understanding how diverse neuronal cell types select their synaptic partners to form circuits. In the neocortex, major subclasses of excitatory projection neurons and inhibitory interneurons are conserved across functionally distinct regions. There is evidence these subclasses form circuits that depend primarily on their identity; however, regional cues likely also influence their choice of synaptic partners. We mined the Allen Brain Institute's single-cell RNA-sequencing database of mouse cortical neurons to study the expression of cellular adhesion molecules (CAMs) necessary for synapse formation in two regions: the anterior lateral motor cortex (ALM) and the primary visual cortex (VISp). We used the Allen's metadata to parse cells by clusters representing major excitatory and inhibitory subclasses that are common to both ALM and VISp. We then performed two types of pairwise differential gene expression analysis: 1) between different neuronal subclasses within the same brain region (ALM or VISp), and 2) between the same neuronal subclass in ALM and VISp. We filtered our results for differentially expressed genes encoding CAMs and developed a novel bioinformatic approach to determine the sets uniquely enriched in each neuronal subclass in ALM, VISp, or both. This analysis provides an organized set of genes that may regulate circuit formation in a cell-type specific manner. Furthermore, it identifies candidate mechanisms for the formation of circuits that are conserved across functionally distinct cortical regions or that are region dependent. Finally, we used the SFARI Human Gene Module to identify CAMs from our analysis that are related to risk for autism spectrum disorder (ASD). From over 3,000 differentially expressed genes, we found 40 ASD-associated CAMs that are enriched in specific neuronal subclasses in both ALM and VISp. Our analysis provides clear molecular targets for future studies to understand neocortical circuit organization and abnormalities that underly autistic phenotypes.

INTRODUCTION

The diversity of cell-types and areas in the neocortex contributes to the breadth of computations that cortical networks perform. Among this diversity, neurons can be broadly divided into glutamatergic (excitatory) and GABAergic (inhibitory) classes. Glutamatergic projection neurons are born locally within each region and send long-range projections to cortical or subcortical targets (Greig et al., 2013;Tasic et al., 2018). GABAergic neurons are derived from extracortical areas, disperse throughout cortical regions, and primarily make short-range, local connections, reflecting their common name as interneurons (Batista-Brito and Fishell, 2009;Tasic et al., 2018). Within these classes, there exist subclasses - Groups of neurons that share features which delineate them from other neurons of the same class. Excitatory subclasses are generally defined by their projection target and the cortical layer of their cell body, while inhibitory subclasses are defined by the expression of unique molecular markers (Zeng and Sanes, 2017;Tasic et al., 2018). Although the definition of a cell-type is still debated, subclass distinctions separate neurons into groups that exhibit unique roles in neural circuits (Harris and Shepherd, 2015;Zeng and Sanes, 2017).

Interestingly, the same major subclasses of excitatory and inhibitory neurons are intermingled throughout the neocortex (Harris and Shepherd, 2015). While neurons from the same subclass can have distinct long-range inputs and outputs dependent on their cortical region (Greig et al., 2013;Harris and Shepherd, 2015;Pouchelon et al., 2021), how subclasses form local, intracortical circuits remains an open question. There is evidence of cell-types forming local synaptic connections according to their subclass identity (Harris and Shepherd, 2015;Ye et al., 2015;Wester et al., 2019). A prominent example is that across diverse cortical regions, intratelencephalic projection neurons preferentially form local synapses onto pyramidal tract projection neurons (Harris and Shepherd, 2015;Gutman-Wei and Brown, 2021). Among interneurons, stereotyped connections between pre- and post-synaptic cell-types have also been revealed (Harris and Shepherd, 2015;Gutman-Wei and Brown, 2021). In addition to cell-type biased connections, there is evidence of neuronal subclasses preferentially targeting specific subdomains on post-synaptic cells and exhibiting unique synaptic properties dependent on their partner's identity (Harris and Shepherd, 2015;Gutman-Wei and Brown, 2021). Together, these findings suggest molecular mechanisms for synaptic targeting and maintenance that are conserved within subclasses. However, it is likely that regional cues also shape synaptic connections for the unique functions of each region such as motor planning or visual processing. Recent work has even argued that major subclasses of interneurons form local circuits that are more predicted by their brain region than cell-type (Pouchelon et al., 2021). Thus, whether

intrinsic subclass properties or guidance from regional cues serves an organizing principle for the self-assembly of local circuits is unclear.

To investigate the extent that subclass and brain region influence cell-type differentiation and circuit organization, we analyzed the transcriptomic signatures of neocortical neurons from major subclasses and across functionally distinct regions. We found that excitatory and inhibitory neurons are primarily differentiated by their subclass rather than brain region. This pattern is reflected in transcripts for synapse formation and regulation, suggesting that across the neocortex, neurons integrate into local circuits based on their subclass identity. Building on this, we identified synaptic genes whose expression is conserved within specific subclasses across brain regions. However, our analysis also reveals synaptic genes that are enriched in unique or multiple subclasses in a region-specific manner. Together, these findings suggest that while subclass identity is the primary driver of cell-type differentiation, regional cues sculpt the synaptic machinery of each subclass. The presence of cell-types, each slightly tailored for area-specific functions, provides further evidence that neocortical circuits are variations on a theme (Harris and Shepherd, 2015). Labeling for genes implicated in autism reveals that disruptions in any major neuronal subclass likely contribute to the disorder. Thus, we identify clinically relevant candidate mechanisms for the organization of circuits that are conserved across the neocortex or that are region-dependent.

RESULTS

Transcriptomic profiles of both excitatory and inhibitory neurons are primarily differentiated by subclass rather than brain region

To compare neuronal subtypes between two functionally distinct cortical regions, we downloaded the Allen Institute's single-cell RNA-sequencing datasets of adult mouse cortical neurons collected from the anterior lateral motor cortex (ALM) and primary visual cortex (VISp) (Tasic et al., 2018) (**Fig. 1A**). These include metadata that define each cell's brain region (ALM or VISp), neurotransmitter (glutamatergic or GABAergic), and major subclass (e.g., L2/3 IT for excitatory neurons or Pvalb for inhibitory interneurons). We loaded these data into the R toolkit Seurat (Stuart et al., 2019) and ran the non-linear dimensionality reduction algorithm UMAP for unbiased clustering of excitatory projection neurons and inhibitory interneurons (**Figs. 1B, C**). We used the metadata to label each cell according to its brain region (**Figs. 1Bi, Ci**) or major subclass (**Figs. 1Bii, Cii**). As expected, this reproduced the findings of Tasic et al. (2018) that excitatory neurons cluster according to both brain region and subclass (**Fig. 1B**) but interneurons cluster only by subclass (**Fig. 1C**). This suggests that major subtypes of excitatory neurons (e.g., L2/3 IT) are not conserved across cortical regions but are distinct in ALM and VISp (Tasic et al., 2018). However, relative positions of clusters in UMAP space do not indicate the magnitude of differences between cell-types. Thus, we performed differential gene expression analyses to compare the number of genes that distinguish neuronal subtypes within and across brain regions (**Figs. 2A, B**).

For our analyses, we chose a subset of major excitatory and inhibitory neuronal subclasses that are shared between ALM and VISp and are relatively well studied (Harris and Shepherd, 2015a;Tremblay et al., 2016). For excitatory cells, these included layer 2/3 intratelencephalic (L2/3 IT), layer 5 intratelencephalic (L5 IT), layer 5 pyramidal tract (L5 PT), and layer 6 corticothalamic (L6 CT). For inhibitory cells, these included those expressing parvalbumin (PV), somatostatin (SST), or vasoactive intestinal peptide (VIP). We conducted two types of analyses: 1) pairwise differential expression tests between different excitatory or inhibitory subclasses within the same cortical region (ALM or VISp) (**Fig. 2A**) and 2) pairwise differential expression tests between the same subclasses across ALM and VISp (**Fig. 2B**). For each pairwise test, we counted the number of differentially expressed genes and normalized it to the total number compared (**Table X**). This allowed us to determine the proportions of differentially expressed genes among all the pairwise tests both within and across brain regions, which we visualized using a Circos plot (Krzywinski et al., 2009) (**Fig. 2C**). For both excitatory and inhibitory cells, we found that each major subclass is more similar between brain regions than between other subclasses within the same brain region. Thus, even though

excitatory neurons cluster by brain region (**Fig. 1B**), subclass identity best distinguishes them. Interestingly, we also found differences between major inhibitory subtypes across ALM and VISp, suggesting that regional cues also influence their differentiation (Pouchelon et al., 2021). Our results agree with previous work that major excitatory and inhibitory subclasses are primarily conserved across diverse cortical regions when considering transcriptomic profiles (Tasic et al., 2018;Yao et al., 2021).

Genes relevant to synaptic connectivity are primarily differentiated by subclass rather than brain region

Although major neuronal subclasses largely share transcriptomic profiles between ALM and VISp, these two cortical regions are functionally distinct. Thus, we next asked if genes related to synaptic connectivity and circuit organization are also shared for each neuronal subclass across the cortex. We filtered our differential expression results using the gene ontology (GO) software PANTHER 16.0 (Mi et al., 2021) to select genes labeled by the terms “Cell-cell adhesion”, “Regulation of cell-cell adhesion”, and “Regulation of trans-synaptic signaling”. These were chosen because they include genes related to the formation, maintenance, and modulation of synaptic connections among distinct neuronal subtypes (Fuccillo et al., 2015;Földy et al., 2016;Chowdhury et al., 2021). Throughout, we refer to these as “circuit-related” genes. For all three GO terms, we found lower proportions of differentially expressed circuit-related genes between the same subclass in ALM and VISp than between different subclasses, regardless of brain region **(Fig. 3)**. This matches the pattern observed for differential expression of all transcripts (**Fig. 2C**). Thus, these analyses suggest that despite the functional distinction of ALM and VISp, gene expression profiles that define cell-types and their markers for circuit organization are largely conserved across the neocortex.

Identifying the circuit-related genes that are differentially expressed between subclasses reveals that most are region-specific for excitatory projection neurons

The above analyses simply consider the proportions of differentially expressed genes to describe general trends in the data. Thus, by percentage most circuit-related genes are differentially expressed between neuronal subclasses within ALM and VISp rather than the same subclasses across these cortical regions (**Fig. 3**). However, this analysis does not consider if the genes that are differentially expressed between neuronal subclasses within ALM and VISp are different sets. This could obscure the true number of genes that provide region-specific instructions for circuit organization. Thus, for each differential expression test between neuronal subclasses within ALM or VISp, we categorized the result as either 1) ALM-specific, 2) VISp-specific, 3) conserved, or 4) unique

(**Fig. 4A**). For example, *rock1* is differentially expressed between VIP and PV cells in ALM but not VISp (ALM-specific, **Fig. 4Ai**). Similarly, *reln* is differentially expressed between VIP and PV cells in VISp but not ALM (VISp-specific, **Fig. 4Aii**). Importantly, for a gene that is differentially expressed between the same pair of neuronal subtypes in both ALM and VISp, it is necessary to consider which subtype had greater expression in each comparison. For example, *shisa6* exhibits greater expression in VIP cells relative to PV cells in both ALM and VISp (conserved, **Fig. 4Aiii)**. However, *calb1* exhibits greater expression in L5 IT cells relative to L5 PT cells in ALM but this relationship is reversed in VISp (unique, **Fig. 4Aiv)**.

In **Figure 4B**, we plotted the total number of circuit-related genes that demonstrated significant differential expression between neuronal subclasses in each category. Importantly, genes in the "unique" category were rare and only observed in a subset of tests between excitatory projection neurons (only 8 genes in 4 tests). Thus, the analysis from our differential expression tests yielded few results that would be difficult to interpret biologically. As expected, for inhibitory interneurons most differential expression test results were conserved between ALM and VISp. This is consistent with the observation that interneuron expression profiles do not cluster by brain region (**Fig. 1C**) and our Circos plots of proportions of differentially expressed genes within and across cortical regions (**Fig. 3**). However, we identified several genes for which differential expression results were ALM- or VISp-specific and consider them in detail in below (**Figs. 8 and 9**). Strikingly, for excitatory neurons we found that more than half of the circuit-related genes in each comparison were exclusive to ALM or VISp (except for the L2/3 IT vs L6 CT comparison). This appears incongruent with our above results that excitatory subclasses have a lower proportion of differentially expressed circuit-related genes across brain regions (**Fig. 3**). However, we emphasize that this analysis is unique from **Figures 2 and 3**. In **Figure 4**, we consider the *identity* of circuit-related genes that are differentially expressed between neuronal subclasses *within* each brain region (as in **Fig. 2A**) to determine if the gene sets and their differential expression results overlap. This reveals that for excitatory projection neurons most circuit-related genes that distinguish subclasses within ALM and VISp are exclusive to each region.

The results from individual differential expression tests can be combined to identify circuit-related genes relevant to each major neuronal subclass

In the above analysis, we found 1452 significant differential expression test results for 394 circuit-related genes that were exclusive to ALM, exclusive to VISp, or conserved between cortical regions (**Fig. 4B, Table X**). This makes it challenging to derive practical biological insights from these

data. Thus, we developed a novel bioinformatic approach to narrow our results to a tractable set of genes that have a high likelihood to play key roles in cortical circuit organization. Our strategy identified circuit-related genes for which differential expression results were consistent across all tests for each neuronal subclass. As an example, we consider a subset of seven circuit-related genes with significant differential expression among inhibitory neuronal subclasses in **Figure 5**. *Cyfip1* has greater expression in VIP cells than PV cells in ALM, but not in VISp (**Fig. 5A**); thus, we classified this gene as ALM-specific (**Fig. 5Bi**). *Dlg2* has greater expression in VIP cells than PV cells in *both* ALM and VISp (**Fig. 5A**); thus, we classified this gene as Conserved (**Fig. 5Bii**). Finally, *kctd13* has greater expression in VIP cells than PV cells in VISp, but not in ALM (**Fig. 5A**); thus, we classified this gene as VISp-specific (**Fig. 5Biii**). For these three genes, differential expression between SST and PV cells is not different, thus, we would predict greater expression in VIP cells relative to SST cells for each. However, this was not the case (**Fig. 5A**), which renders the biological relevance of these results unclear. In contrast, *nrx1n*, *pten*, *snap25*, and *socs5*, all demonstrate consistent differential expression results among all three tests (**Fig. 5A**). Thus, we sorted these genes as ALM-specific (*socs5*), VISp-specific (*pten*), or Conserved (*nrx1n* and *snap25*), as above (**Fig. 5B**). We then kept these four genes for further analysis, and discarded *cyfip1*, *dlg2*, and *kctd13* (**Fig. 5C**). Next, we defined *nrx1n*, *pten*, *snap25*, and *socs5* as "subclass-relevant" to VIP cells and denoted whether expression was higher or lower relative to the other subclasses with up and down arrows, respectively (see **Fig. 5Cii**). Finally, dysregulation of circuit-related genes directly contributes to neurological disorders (Chowdhury et al., 2021). Thus, we highlighted candidate risk genes for autism using the SFARI Human Gene Module (https://gene.sfari.org/database/human-gene/) in green (**Fig. 5C**).

For some subclass-relevant genes, a subset of their differential expression results was shared between ALM and VISp, and a subset was exclusive to one cortical region. As an example, we include the analysis of *egr1* in **Figure 6**. *Erg1* demonstrates significant differential expression in all three tests between interneuron subclasses in ALM, but in only two tests in VISp (**Fig. 6A**). Thus, in ALM *erg1* has greater expression in VIP cells relative to SST cells (**Fig. 6Bi**), while in both ALM and VISp *erg1* has greater expression in SST cells relative to PVs cells and in VIP cells relative to PV cells (**Fig. 6Bii**). These differences are important for interpreting the relevance of *erg1* expression among neuronal subclasses. In ALM, we define *erg1* as having greatest expression in VIP cells relative to PV and SST cells (**Fig. 6Ci**). However, we indicate with an asterisk (*) that a subset of these results is conserved in VISp, where we define *erg1* as having lower expression in PV cells relative to VIP and SST cells (**Fig. 6Cii**).

In summary, our strategy combined pairwise differential expression test results to relate each circuit-related gene to a major neuronal subclass. This allowed us to create an organized and tractable list of genes that are candidates to mediate synaptic connections between specific neuronal subtypes across the cortex or dependent on cortical region (**Figs. 7 – 9**). Of the 394 circuit-related genes in the ALM-specific, VISp-specific, and Conserved subsets, 71% (280) mapped to one or more subclass. Furthermore, 26% of these are risk genes for autism and may inform future studies investigating the circuit mechanisms of this disorder. We describe these data in greater detail below and provide a searchable spreadsheet that identifies each gene-subclass relationship (**Supplemental Table X**).

Most circuit-related genes identified as Conserved are subclass-relevant to inhibitory interneurons

Our analysis identified 57 circuit-related genes relevant to excitatory neuronal subclasses that were conserved in both ALM and VISp (**Fig. 7A, B**). A subset of 9 genes was relevant to both L2/3 IT and L5 IT subclasses; thus, we created a new subclass “IT” for these genes (**Fig. 7A**). Indeed, most circuit-related genes that were conserved among excitatory neurons in both ALM and VISp were relevant to the L2/3 IT, L5 IT, and combined IT subclasses. Only a small number of genes (5) were relevant to other combinations of multiple subclasses (e.g., L2/3 IT and L6 CT) in both ALM and VISp (**Fig. 7B**). Interestingly, twice as many genes (124) were conserved for interneuron subclasses compared to excitatory subclasses (**Fig. 7C, D**). Furthermore, most of these genes were unique to interneurons; only 23 genes were identified as subclass-relevant in both the excitatory and inhibitory populations. Thus, the mechanisms that regulate synaptic connections in a subclass-specific manner may be mostly distinct between excitatory and inhibitory cell types. Among interneurons, most were relevant to VIP cells and distinguished them from PV and SST cells (**Fig. 7C**). Thus, the mechanisms underlying the formation of disinhibitory circuits appear to be conserved across cortical regions. Finally, SFARI autism risk genes were distributed uniformly among all excitatory and inhibitory subclasses and represented 23% and 28% of the total genes for each group, respectively. Thus, we could not identify a specific set of neuronal subtypes that may be uniquely vulnerable in autism. In summary, we identified circuit-related genes that are conserved between ALM and VISp for each neuronal subclass and may act as molecular mechanisms underlying stereotyped “canonical circuits” found across the cortex (Kubota, 2014;Harris and Shepherd, 2015b;Gutman-Wei and Brown, 2021). Furthermore, many of them are candidates to disrupt the balance between excitation and inhibition throughout the cortex in developmental disorders (Sohal and Rubenstein, 2019).

Most circuit-related genes that are identified as specific to subclasses in ALM or VISp are in excitatory projection neurons

We next identified circuit-related genes that were relevant to neuronal subclasses exclusively in ALM (**Fig. 8**) or VISp (**Fig. 9**). In ALM, we found 92 genes relevant to excitatory subclasses (**Fig. 8A**), 11 of which shared a subset of differential expression results that were also relevant to subclasses in VISp (**Fig. 8A,** genes denoted by * are also found in **Fig. 7A**). Only a small number of genes (6) were relevant to combinations of multiple subclasses (e.g., L5 IT and L5 PT) exclusively in ALM (**Fig. 8B**). For interneurons, we found 47 subclass-relevant genes (**Fig. 8C)**, 11 of which shared a subset of differential expression results that were also relevant to subclasses in VISp (**Fig. 8C,** genes denoted by * are also found in **Fig. 7C**). Like our analyses of conserved genes (**Fig. 7**), few were shared between excitatory and inhibitory subclasses (12 total). However, in contrast to conserved genes, most were subclass relevant to excitatory cells. Furthermore, among excitatory cells most genes were relevant specifically to the IT subclasses. Our analysis of genes that were subclass relevant in VISp yielded similar results (**Fig. 9**). In VISp, we found 79 genes relevant to excitatory subclasses (**Fig. 9A**), 12 of which shared a subset of differential expression results that were also relevant to subclasses in ALM (**Fig. 9A,** genes denoted by * are also found in **Fig. 7A**). Only 2 genes were relevant to combinations of multiple subclasses (e.g., L5 PT and L6 CT) exclusively in VISp (**Fig. 9B**). For interneurons, we found 28 subclass-relevant genes (**Fig. 9C)**, 8 of which shared a subset of differential expression results that were also relevant to subclasses in ALM (**Fig. 9C,** genes denoted by * are also found in **Fig. 7C**). Like in ALM, most genes in VISp were relevant to excitatory IT-type subclasses, and few overlapped between excitatory and inhibitory cells (9 total).

In summary, we found that most of the subclass-relevant circuit-related genes that are dependent on cortical region are found in excitatory neurons. Furthermore, most of these are relevant to IT-type cells and many are candidate risk genes for autism. These data agree with previous work in humans that IT-type cells may be a key contributor to autistic phenotypes (Parikshak et al., 2013). However, our results highlight that the development of therapies to target excitatory cells may not work uniformly across the cortex.

A small set of genes demonstrate greater differential expression exclusively in ALM or VISp

Finally, we analyzed the differential expression results for each subclass between ALM and VISp (**Fig. 10**; see **Fig. 2B** for schematic). We hypothesized there are genes that are uniquely expressed between these functionally distinct and spatially distant cortical regions. Thus, we filtered

genes that demonstrated consistent differential expression results for all neuronal subclass between ALM and VISp. Strikingly, this analysis identified few genes. Among excitatory subclasses, only 4 genes were biased to ALM (*lmo4*, *lphn2*, *neurod6*, and *tmeff1*) and only 4 were biased to VISp (*brinp3*, *id2*, *spock3*, *tenm2*) (**Fig. 10A**). Only 1 of these genes, *tenm2*, was circuit-related (**Fig. 10C**). Among inhibitory subclasses, we found only 1 gene (*cenpa*) that was biased to VISp, but it was not circuit-related (**Fig. 10B, D**). Finally, PV and SST interneurons share their developmental origin from the medial ganglionic eminence (Xu et al., 2004;Butt et al., 2005), but VIP interneurons are derived from the caudal ganglionic eminence (Lee et al., 2010;Miyoshi et al., 2010). Thus, we asked if there are circuit-related genes that distinguish interneurons in ALM and VISp dependent on embryonic origin. We found none for PV/SST subclasses and only 4 such genes (*calb2*, *igf1*, *npy2r*, and *pvrl3*) for the VIP subclass (**Fig. 10B, E**). Of note, a previous study showed that *igf1* expression in VIP interneurons regulates experience-dependent plasticity in primary visual cortex (Mardinly et al., 2016). Our analysis suggests this mechanism may be unique to VISp (**Fig. 10E**). In summary, surprisingly few genes distinguish excitatory and inhibitory neuronal subclasses between the rostral and caudal poles of the cortex.

DISCUSSION

The sheer number of neurons and synapses in the brain remains a prevailing challenge to understanding its organization and the etiology of neurological disorders. Although there has been progress in identifying neuronal subtypes (Tasic et al., 2018), this has underscored the potential complexity of how neural circuits are structurally and functionally organized. The diversity of brain areas provides an additional consideration of how circuits are regionally specialized and whether arealization accounts for distinct cell types across regions. In the neocortex, the presence of many regions, each populated with the same major subclasses, is a perquisite for local circuits potentially being globally stereotyped by interactions between subclasses. However, it does not necessitate such an organization - even subclasses of neurons that are thought to integrate into circuits in a conserved manner have shown unique patterns of connectivity based on the region they inhabit (Pouchelon et al., 2021). It is unlikely that higher order features such as subclass identity or brain region serve as the only organizing principle. However, identifying the primary mediator of circuit organization and the extent it is modified will guide theories for understanding and modeling neural circuits.

By leveraging single-cell RNA-sequencing datasets, we show that both glutamatergic and GABAergic neurons are primarily differentiated by their cell-type identity across neocortical regions. This is consistent with recent analyses of subclasses beyond ALM and VISp (Yao et al., 2021a) and at the class level in the developing human brain (Bhaduri et al., 2021). We considered the possibility that specific gene modules are directed by regional cues and underlie features that are primarily differentiated by brain region. We focused on CAM-related genes due to their role in mediating and specifying synaptic connections. Filtering our differential expression data, we found that the expression of these genes is also differentiated in a subclass conserved manner. This is surprising as it suggests that the self-assembly of neural circuits is primarily mediated by intrinsic subclass properties rather than regional or other extrinsic cues. Notable differences between PT cells across brain regions that we and others (Yao et al., 2021a) observe may be due to unique CAM sets for distinct subcortical projection targets (optic tectum for VISp PT cells versus brainstem and spinal cord for ALM PT cells).

By identifying genes that describe subclasses, we found that GABAergic neurons conserve the most subclass features across regions. These results align with recent studies showing that cell-types vary along a continuum across cortical regions and that glutamatergic neurons exhibit greater differences than interneurons (Bhaduri et al., 2021;Nano et al., 2021;Yao et al., 2021a). The distinct clustering of pyramidal cells between ALM and VISp is thought to be accounted for by a frontal to

occipital pole gradient in cell-types (Yao et al., 2021a). This has interesting implications for our analysis: subclass enrichment that is conserved in ALM and VISp may also be conserved in the same subclasses in areas between, and subclass enrichment that is ALM or VISp-specific may be conserved in adjacent areas. This suggests a framework for potentially extrapolating extensive studies on local circuit motifs (Campagnola et al., 2022) to regions where they were not directly studied. Together, these findings also set up a theory for cross-modal plasticity in which slight variations in local circuitry are adopted among functionally distinct regions. This may underlie adaption to sensory deprivation (e.g. improvements in hearing after blindness) (Ewall et al., 2021).

Multiple mechanisms likely contribute to the degree of regional-specification between classes. The birth of glutamatergic neurons locally in developing cortical regions versus the origin of GABAergic neurons from the ganglionic eminences suggests arealization at the level of progenitors. However, there is evidence of similarities between progenitors for different GABAergic subtypes, the potential to produce multiple subtypes from the same progenitor, and the presence of subclass-specific features after cell-cycle exit (Mayer et al., 2018;Bandler et al., 2022). This suggests that conservation of GABAergic subclass features primarily arises from developmental programs that take precedence post-mitotically. Future studies can investigate the mechanisms guiding interneurons of the same subclass to invade the same versus different brain regions and how they become specialized in each region. For glutamatergic progenitors, local cues modify several of their features and are thought to adapt them for region-specific architectures (Llorca and Marin, 2021). While glutamatergic progenitors vary across cortical regions (Bhaduri et al., 2021), how this variation is reflected in progenitors fated to produce the same subclass remains to be investigated.

The diversity among excitatory and inhibitory neurons is likely tied to how these cell-types evolved from primordial neural circuits. Indeed, a study showed that neocortical glutamatergic neurons appear to have diverged more throughout evolution than GABAergic neurons (Tosches et al., 2018), suggesting, like others have speculated (Pouchelon et al., 2021), interneurons serve a fundamental role in safeguarding local circuit dynamics. This aligns with our analysis and previous studies: while interneuron subclasses exhibit limited genetic variation across cortical areas, pyramidal subclasses are regionally diverse (Tasic et al., 2018;Yao et al., 2021a). Despite the observation that L5 PT cells have the most differences across brain regions, we find that IT cells have the most region-specific features based on our enrichment analysis. Interestingly, this also has evolutionary relevance: only eutherians developed the corpus callosum (Suarez et al., 2014), suggesting IT cells are a prominent circuit element which expanded in evolution. Thus, IT cells may underlie advances in the computational ability of neural networks by helping specialize local circuits and facilitating their coordination across regions.

Epigenetic regulation is expected to underlie the presence of cell-type markers. Indeed, we predict PV interneurons are enriched with Mef2c across cortical regions and a recent study revealed that the transcription factors Maf and Mafb drive Mef2c expression which promotes differentiation of PV interneurons (Pai et al., 2020). Additional evidence has demonstrated that Mef2c is essential for the differentiation of PV interneurons and persists in expression into adulthood (Mayer et al., 2018). These data suggest that early enrichment of CAM-related genes can be necessary for cell fate specification. Thus, applying our findings to guide studies of transcription factors and chromatin remodelers may reveal regulators that are upstream of cell-type specific CAM enrichment and investigation of whether such regulation is shared within cell-types across regions. This will enable the study of how cell-autonomous and non-cell autonomous mechanisms may compete for or converge on trajectories of neuronal differentiation.

Although there are inherit limitations to RNA-sequencing analysis, we identified cases where predicted enrichment has been verified by at the protein level. This includes subclass markers such as Cux2 in L2/3 IT cells and Mef2c in PV cells (Cubelos et al., 2010;Mayer et al., 2018). We also identify class markers that have been shown to delineate regions such as Lmo4 in motor cortex or Tenm2 and Id2 in visual cortex (Rubenstein et al., 1999;Cederquist et al., 2013;Young et al., 2013). Strikingly, Tenm2 knock out mice display visual but not motor deficits, suggesting, as predicted, Tenm2 is critical for visual circuitry (Young et al., 2013).

There are many genes we identify as being markers of subclasses or brain regions that have not been previously reported. This includes the histone H3 variant CENP-A which we found as a putative marker of interneurons in ALM. CENP-A is primarily thought to localize in centromeric nucleosomes, supporting chromatin compaction in preparation for kinetochore binding and proper chromosomal segregation during cell division (Zhou et al., 2022). Thus, it is surprising to find this gene expressed in presumably post-mitotic neurons. Also, it is interesting that this is the only gene differentially expressed between interneurons across brain regions despite being a modifier of chromatin structure and potentially influencing the transcription of other genes. Assuming CENP-A is not localized at the centromere, it possible that it has non-centromeric functions. Indeed, there are mechanisms for CENP-A deposition at non-centromeric regions and findings that its expression is essential for post-mitotic cells of the Drosophila midgut (Garcia Del Arco et al., 2018;Demirdizen et al., 2019). Recent work also suggests that it is likely repurposed in neurons along with kinetochore and cytoskeleton proteins important for cell-division (Del Castillo et al., 2019;Zhao et al., 2019). Strikingly, CENP-A mutants show disruptions in synaptic development and neurite growth with compelling evidence that this is not explained by abnormal chromosome segregation (Zhao et al., 2019). The role of CENP-A in cortical circuits is unknown. In the context of our analysis, we speculate

that greater expression in adult interneurons of ALM is likely a remnant of development, serves a fundamental function in neurite or synaptic organization, or implicates an unstudied role in specializing GABAergic signaling in motor circuitry.

A challenge to understanding ASD is its broad range of conditions involving impairments in social-communication, cognitive, and sensorimotor functions (Marco et al., 2011;Whyatt and Craig, 2013;Donovan and Basson, 2017). This is likely attributed to disruptions in neural circuits across multiple brain regions including the neocortex (Marco et al., 2011;Donovan and Basson, 2017;Velmeshev et al., 2019;Chowdhury et al., 2021). Identifying genetic disruptions that impair circuitry across regions and pinpointing specific types of neurons that are affected is crucial. An important part of this goal is determining if ASD CAMs are uniquely expressed in select subclasses and whether this pattern is conserved across regions. By performing enrichment analysis, we identify many ASD risk genes that are specific to subclasses in one or more brain regions. For genes that describe a subclass in multiple brain regions, it is possible that in patients with ASD, these genes are only impaired in a specific region. Thus, to understand the translational prospects of putative subclass features, we compared our findings to a published study of single-nucleus RNA sequencing of cortical neurons from healthy individuals and ASD patients (Velmeshev et al., 2019). We identified multiple examples where subclass-specific features are identified as markers of those same subclasses in humans. Additionally, we found two cases where subclass-specific features that are predicted by our analysis are disrupted in the predicted subclasses across brain regions in patients: CAMs Grik1 and Fut9 are predicted to exhibit lesser expression in VIP cells relative to PV and SST cells **(Fig. 7B)**. In ASD patients, these CAMs are upregulated in VIP cells compared to controls. Velmeshev and colleagues also identify genes in projection neuron and interneuron cell types that are dysregulated in a region-specific manner. This highlights the relevance of CAM genes that are enriched in subclasses only in ALM or VISp. Our results provide insights to evaluate how changes in CAM expression may map to disruptions of specific cell types and impair circuit organization in one or more regions. Although CAM dysregulation may not directly translate between murine and human cell-types, there are likely homologous wiring principles (Campagnola et al., 2022) that may be disrupted. This has the potential to support the design of targeted, brain-wide treatments for ASD.

Similarities in local circuitry across neocortical regions provides promise to distill circuits into common elements with stereotyped interactions. Rather than belying the complexity of the neocortex, such a connectivity code would support investigation into its generality and the mechanisms by which it is established, tailored for multimodal computations, and disrupted in development. Combining our findings with circuit mapping techniques and growing lists of CAM interactions will guide studies into the mediators of synaptic biases. Our analysis provides a framework for applying transgenic mouse

models and testing hypotheses about subclass-specific connectivity patterns, their molecular underpinnings across brain regions, and relevance in autism and other neurodevelopmental disorders.

METHODS

***Datasets*.** We downloaded the Allen's ALM and V1 – SMART-seq (2018) datasets (https://portal.brain-map.org/atlases-and-data/rnaseq). Using the scRNA-seq toolkit Seurat (Stuart et al., 2019) and provided metadata, we filtered for glutamatergic and GABAergic neurons from each brain region to create four separate Seurat objects. We merged glutamatergic and GABAergic objects from separate regions to create two final objects for analysis. We followed the Seurat standard pre-processing workflow for QC, normalization, identification of highly variable features, scaling, dimensionality reduction, and UMAP clustering.

***Conducting differential expression tests*.** Using brain region and subclass metadata, expression profiles of cells from each brain region (ALM or VISp) and selected major subclasses (L2/3 IT, L5 IT, L5 PT, and L6 CT for glutamatergic object; Pvalb, Sst, and Vip for GABAergic object) were compared in a pairwise manner (**Figs. 2A, B)**. Comparisons were performed using the Seurat function FindMarkers which implements a non-parametric Wilcoxon rank-sum test for differential expression. For comparisons between different subclasses within a brain region (**Fig. 2A**), we subsetted each object for cells in either ALM or VISp before performing FindMarkers. For example, to compare SST and PV cells in ALM, we first subsetted the GABAergic object for cells in ALM and used the subclass metadata to perform the test. For comparisons between the same subclasses across brain regions (**Fig. 2B**), we subsetted the glutamatergic or GABAergic object for the subclass of interest before performing FindMarkers. For example, to compare SST cells across ALM and VISp, we subsetted the GABAergic object for SST cells and used the region metadata to perform the test. Equivalent comparisons between different subclasses within ALM or VISp were conducted in the same order (e.g. GABA ALM SST vs PV is equivalent to GABA VISp SST vs PV). This allows the log fold change values of differentially expressed genes to be directly compared for equivalent tests in ALM and VISp. Differentially expressed genes were selected with an adjusted P-value cut-off of <0.05 based on Bonferroni correction using all genes in the dataset.

***Approach for normalizing differential expression counts*.** To determine the maximum number of genes that could be differentially expressed in each pairwise comparison, we used the AverageExpression function in Seurat. This function returns the averaged expression of each gene across all cells in an object (log normalized expression values were exponentiated before averaging).

For comparisons between different subclasses within a brain region (**Fig. 2A**), we subsetted the glutamatergic or GABAergic object for a pair of cell-types before performing AverageExpression. For example, to determine the number of unique genes expressed by SST and PV cells, we first subsetted the GABAergic object for both subclasses and used the region metadata to calculate average expression in ALM and VISp. For comparisons between the same subclass across brain regions (**Fig. 2B**), we used the subclass metadata to identify cells in the glutamatergic or GABAergic object before performing AverageExpression. For example, to determine the number of unique genes expressed by SST across ALM and VISp, we labeled cells in the GABAergic object by subclass and used the averaged expression results pertaining to SST cells. For each comparison, we removed genes with an average expression value of 0. This enabled us to determine the total number of unique genes that are expressed by each pair of cell-types in **Figures 2A, B**. We term these genes "union" since they are the union of the set of genes that are expressed by a pair of cell-types. Using the differential expression and union data for each pairwise test, we calculated the proportion of genes that are differentially expressed (e.g. out of 33581 unique genes expressed by L2/3 IT and L5 IT cells in ALM, 1.4% (465 genes) are differentially expressed). We organize these results in a matrix presented in the Circos plot in **Figure 2C** (Krzywinski et al., 2009).

Filtering for circuit-related genes. To identify genes that are circuit-related, we used the classification software PANTHER 16.0 (Mi et al., 2021). We uploaded differentially expressed and union genes, selected Mus musculus as the reference organism, and ran a statistical overrepresentation test (GO biological process complete) with no corrections. Genes labeled under the gene ontology categories "Cell-cell adhesion", "Regulation of cell-cell adhesion", and "Regulation of trans-synaptic signaling" were selected. We annotated and filtered for selected genes in the differential expression and union gene sets. For every GO category and pairwise test, we normalized the number of differentially expressed genes to the number of GO union genes in the manner described above. We present these results in the Circos plots in **Figure 3** (Krzywinski et al., 2009).

Labeling for clinically relevant genes*.* We downloaded the 2021 list of autism risk genes from the SFARI Human gene module and labeled differentially expressed genes that are risk genes. Risk genes from all categories ("Suggestive Evidence", "Strong Candidate", "High Confidence", and "Syndromic") were considered.

Strategy for determining the regional-specificity of each gene. For our analysis, we consider a differential expression datapoint as a coordinate of the following features: Comparison (Glut/GABA ALM/VISp Subclass 1 vs Subclass 2 or Glut/GABA ALM vs VISp Subclass 1), gene, log fold change (logFC) value, relevant GO terms, and relevant conditions. We compiled results for glutamatergic and GABAergic differential expression data that were filtered for circuit-related genes. We created the tool

DE_Collapser in Python to identify differential expression data that is “ALM-specific”, “VISp-specific”, “conserved” or “unique” (**examples in Fig 4A**). This was accomplished using the following algorithm: Analyzing the compiled differential expression data, DE_Collapser iterates through pairs of comparisons to identify comparisons that are equivalent across brain regions (e.g. GABA ALM SST vs PV is equivalent to GABA VISp SST vs PV). It then subsets for their corresponding differential expression data and identifies the intersection of the sets of genes that are differentially expressed in each comparison. For each of these genes, DE_Collapser determines if the relative expression between cell-types is either consistent across comparisons or different by using the sign of the logFC. The logFC values can be used because subclasses were compared in the same order in equivalent differential expression tests. Genes exhibiting a consistent relationship across regions are labeled ‘conserved’. For example, *shisa6* is enriched in VIP cells relative to PV cells in ALM and VISp (**Fig. 4Aiii**). This is indicated by a positive logFC value in the ALM and VISp comparisons. Genes exhibiting a relationship that is reversed across regions are labeled ‘unique’. For example, L5 IT has greater expression of *calb1* relative to L5 PT in ALM but this relationship is reversed in VISp (**Fig. 4Aiv**). This is indicated by opposite signs of the logFC values for the comparison in ALM versus that in VISp. For conserved and unique genes, DE_Collapser computes the average of the logFC values from equivalent comparisons and converts a pair of differential expression datapoints for that gene into one datapoint with the average logFC value. Datapoints for unique genes are also separated into a distinct subset that preserves raw logFC values for comparisons in each region. For genes that are differentially expressed in only one of two equivalent comparisons, DE_Collapser identifies if they are differentially expressed in the ALM comparison (label ‘ALM-specific’; **Fig. 4Ai**) or VISp comparison (label ‘VISp-specific’; **Fig. 4Aii**). Examples of applying DE_Collapser on raw datapoints is represented by the set of arrows between **Figure 5A** and **Figure 5B**. The output of DE_Collapser for each pair of comparisons is compiled into a new dataset combining conserved and unique data with averaged logFC values and ALM- or VISp-specific data with raw logFC values. Raw logFC values for ‘Unique’ genes remain accessible in a separate output. We calculated the total number of conserved, unique, and region-specific genes in each comparison (**Figure 4B**).

Strategy for subclass-relevance analysis. All of the following subclass analyses presented in **Figures 5-10** were performed separately for glutamatergic and GABAergic data. To determine if differentially expressed genes map to subclass features, we created Subclass Identifier (SCID). This algorithm takes the output of DE_Collapser including conserved or region-specific data. It separates comparisons into the order of their constituent subclasses (e.g., VIP vs PV – First subclass = VIP; Second subclass = PV). Based on the logFC values, SCID identifies which subclass has greater or lesser relative expression of each gene: If the logFC > 0 for gene A in a VIP vs PV comparison, then

the first subclass (VIP) has greater expression and second subclass (PV) has lesser expression. If the logFC < 0 for gene B in a VIP vs PV comparison, then the second subclass (PV) has greater expression and first subclass (VIP) has lesser expression.

To assist with classification, we define sets for the minimum comparisons that a gene needs to be differentially expressed in to be in consideration for mapping to a subclass:

L2/3 IT = {'L2/3 IT vs L5 IT', 'L2/3 IT vs L6 CT', 'L2/3 IT vs L5 PT'}

L5 IT = {'L2/3 IT vs L5 IT','L5 IT vs L5 PT','L5 IT vs L6 CT'}

IT = {'L2/3 IT vs L5 PT', 'L2/3 IT vs L6 CT', 'L5 IT vs L5 PT','L5 IT vs L6 CT'}

L5 PT = {'L2/3 IT vs L5 PT','L5 IT vs L5 PT','L5 PT vs L6 CT'}

L6 CT = {'L2/3 IT vs L6 CT','L5 IT vs L6 CT','L5 PT vs L6 CT'}

VIP = {'VIP vs PV', 'VIP vs SST'}

SST = {'SST vs PV', 'VIP vs SST'}

PV = {'SST vs PV', 'VIP vs PV'}

The following steps comprise the key computations of the algorithm which we demonstrate with an example considering a hypothetical gene X. We refer to a comparison set as all of the comparisons that a gene appears differentially expressed in.

1) SCID iterates through the differential expression data for each gene and determines the list of subclasses that differentially express the gene. If L2/3 IT and L5 IT are identified, IT is added to the subclass list.

   Example: If gene X is differentially expressed in the comparison set {‘VIP vs PV', 'VIP vs SST', ‘SST vs PV’}, then SCID identifies SST, VIP, and PV as the subclass list.

2) For each subclass in the list, SCID iterates through pertinent subclass sets that we define above. It identifies subclass sets that are a subset of the comparison set associated with the gene.

   Example: For gene X, the VIP, {'VIP vs PV', 'VIP vs SST'}; SST, {'SST vs PV', 'VIP vs SST'}; and PV {'SST vs PV', 'VIP vs PV'} subclass sets would each be identified as a subset of the comparison set {'VIP vs PV', 'VIP vs SST', ‘SST vs PV’}.

3) For each subclass set, SCID filters the differential expression data for comparisons that are included in the set.

   Example: For the VIP set, SCID filters for Gene X data involving 'VIP vs PV', 'VIP vs SST' comparisons; for the SST set, it filters for Gene X data involving 'SST vs PV', 'VIP vs SST' comparisons; and for the PV set, it filters for Gene X data involving 'SST vs PV', 'VIP vs PV' comparisons.

4) In each filtered dataset, SCID evaluates the information on which subclass has greater or lesser expression of a gene in each comparison. This enables it to determine if the subclass (e.g. VIP) corresponding to the subclass set (e.g. {VIP vs PV', 'VIP vs SST'}) has consistent expression relative to all other subclasses (e.g. PV and SST). We refer to these as 'subclass-relevant' genes. We label subclass-relevant genes for their subclass features and the corresponding expression relative to all other subclasses for each feature **(Figs. 5C, 7-10)**. Genes that are not subclass-relevant remain unlabeled.
   Example: If gene X is expressed greater in VIP cells relative to PV and SST cells, and expressed less in SST cells relative to VIP and PV cells, gene X is labeled accordingly: up in VIP cells and down in SST cells.

Note: Given the low number of genes in the unique subset, we searched for subclass-relevant genes directly. No genes were identified. Also, per the defined subclass lists and differential expression analysis, SCID identified features are relative to either glutamatergic or GABAergic subclasses.

Strategy to identify subclass-relevant genes that involve conserved and region-specific data
To account for subclass features that may be overlooked if only considering region-specific data (**Fig. 6**), we created CortexSCID. This algorithm considers the results of performing SCID on both the conserved data and the raw ALM or VISp data. The latter contains all of the ALM or VISp data not just that which is region-specific. CortexSCID identifies genes that are present in the conserved set and ALM or VISp sets and removes genes that have the same comparison sets and subclass features. The remaining genes have one or more subclass features that are distinct to ALM or VISp. We refer to these as * genes. CortexSCID identifies * genes, their subclass features, and corresponding relative expression. It updates the ALM or VISp differential expression data by preserving only genes and their subclass relevance that are region-specific. * genes are only identified if the same gene is relevant to different subclasses in a conserved and region-specific manner (e.g. *egr1* in **Fig. 6**). There are genes that describe subclasses in a region-specific manner that we do not denote as * genes even though they consider conserved differential expression datapoints. This is because the conserved differential expression data was insufficient to describe subclass features that are shared in ALM and VISp.
Note: We do not label genes with an * that are relevant to subclasses of different classes in a region-specific and conserved manner.
Identifying genes that are biased to ALM or VISp. Using the differential expression data for comparisons between the same subclass across ALM and VISp **(Fig. 2B)**, we queried for genes that are enriched across multiple glutamatergic or GABAergic subclasses in ALM relative to VISp or vice

versa **(Fig. 10)**. For example, we identified *cenpa* as being enriched in VIP, PV, and SST cells in VISp relative to their equivalents in ALM (**Fig. 10B left, Fig. 10D**). We also searched for genes that are differentially expressed by VIP cells but not PV or SST cells across brain regions (**Fig. 10B right, Fig. 10 E**) or PV and SST cells but not VIP cells across brain regions (no such genes identified). The Violin plots for *tenm2*, *cenpa*, and *igf1* were created in Seurat (**Fig. 10C-E**).

Heatmap creation. We created the function Morpheus Prepper (MorphPrep) to streamline subclass and clinical analyses and organize results as matrices for heatmap visualization in Morpheus (https://software.broadinstitute.org/morpheus

FIGURE LEGENDS

Figure 1. Single-cell RNA-seq data for major neuronal subclass in ALM and VISp. (A) Anatomical locations of ALM (blue) and VISp (green) with corresponding coronal sections. **(B)** UMAP dimensionality reduction for glutamatergic cells color-coded using metadata for brain region **(Bi)** or subclass **(Bii)**. **(C)** UMAP dimensionality reduction for GABAergic cells color-coded using metadata for brain region **(Ci)** or subclass **(Cii)**. L2/3 IT, L5 IT, L5 PT, L6 CT, Pvalb, Sst, and Vip clusters were selected for downstream analysis. CR, Cajal–Retzius; IT, intratelencephalic; PT, pyramidal tract; CT, corticothalamic; NP, near-projecting; Pvalb, parvalbumin; Sst, somatostatin; Vip, vasoactive intestinal peptide. Coronal sections adapted from the Allen Mouse Brain Atlas (https://mouse.brain-map.org/static/atlas).

Figure 2. Differential expression tests for major neuronal subclasses either within or between cortical regions. (A) Pairwise differential expression tests conducted between excitatory or inhibitory neuronal subclasses in ALM or VISp. **(B)** Pairwise differential expression tests conducted between the same neuronal subclasses across ALM or VISp. **(C)** Circos plot summary of all differential expression test results. These include results for all transcripts in the dataset, in contrast to Figure 3 in which genes are filtered by gene ontology labels. Subclasses are listed along the circumference of the plot in their respective brain region (blue - ALM; green - VISp). Bands connect subclasses that were compared in a pairwise differential expression test. Band thickness represents the proportion of differentially expressed genes in each test: thinner bands indicate a small percentage of distinguishing genes and thicker bands indicate a larger percentage of distinguishing genes.

Figure 3. Circuit-related genes are primarily differentiated by subclass rather than cortical region. Circos plots of differential expression results filtered for circuit-related genes using PANTHER gene ontology labels “cell-cell adhesion”, “regulation of cell-cell adhesion”, and “regulation of trans-synaptic signaling”. Subclasses are listed along the circumference of the plot in their respective brain region (blue - ALM; green - VISp). Bands connect subclasses that were compared in a pairwise differential expression test. Band thickness represents the proportion of differentially expressed genes in each test: thinner bands indicate a small percentage of distinguishing genes and thicker bands indicate a larger percentage of distinguishing genes. No genes involved in Regulation of cell-cell adhesion were differentially expressed between SST or PV cells across regions.

Figure 4. Categorization of circuit-related genes that are differentially expressed between neuronal subclasses within ALM or VISp. (A) Examples of differential expression test results used to categorize circuit-related genes as: ALM-specific **(Ai)**, VISp-specific **(Aii)**, Conserved **(Aiii)**, or Unique **(Aiv)**. The genes and results are real examples from the analysis. **(B)** Stacked bar plot of the total number of circuit-related genes in each category for every pairwise differential expression test between subclasses. "Unique" genes were only observed in L2/3 IT vs L5 PT; L5 IT vs L5 PT; L5 IT vs L6 CT; and L5 PT vs L6 CT comparisons.

Figure 5. Strategy to determine circuit-related genes relevant to specific neuronal subclasses. (A) Examples of differential expression rests results for seven circuit-related genes in interneuron subclasses in ALM and VISp. **(B)** Separation of genes from *(A)* into ALM-specific **(Bi),** Conserved **(Bii)**, and VISp-specific **(Biii)** subsets. **(C)** Subclass-relevant genes from *(B)* identified by their consistent differential expression results among all subclass tests. For each gene, we provide a subclass-relevant label and an arrow indication whether it is expressed higher or lower relative to other subclasses. Subclass-relevant genes listed on the SFARI Human Gene Module as implicated in autism were highlighted green.

Figure 6. Strategy to identify subclass-relevant genes that describe both conserved and region-specific relationships. (A) Example differential expression test results for *egr1* in both ALM and VISp. **(B)** Separation of differential expression results into subsets that are ALM-specific **(Bi)** and conserved in both ALM and VISp **(Bii). (C)** Annotation of *erg1* using an asterisk (*) to denote that a subset of ALM-specific differential expression results **(Ci)** is also found within the set labeled as "Conserved" **(Cii)**. Note that for *erg1*, the relevant subclass and relative expression-level are different in *Ci* and *Cii* because the combinations of differential expression results among neuronal subclasses are unique.

Figure 7. Circuit-related genes that are relevant to the same neuronal subclasses in ALM and VISp. (A) Circuit-related genes relevant to single subclasses of excitatory projection neurons. We created the new subclass "IT" for genes was relevant to both L2/3 IT and L5 IT subclasses. **(B)** Circuit-related genes relevant to multiple subclasses of excitatory projection neurons. **(C)** Circuit-related genes relevant to single subclasses of inhibitory interneurons. **(D)** Circuit-related genes relevant to multiple subclasses of inhibitory interneurons. CCA = cell-cell adhesion; RCCA = regulation of cell-cell adhesion; RTSS = regulation of trans-synaptic signaling.

Figure 8. Circuit-related genes only relevant to neuronal subclasses in ALM. A) Circuit-related genes relevant to single subclasses of excitatory projection neurons. We created the new subclass "IT" for genes was relevant to both L2/3 IT and L5 IT subclasses. **(B)** Circuit-related genes relevant to multiple subclasses of excitatory projection neurons. **(C)** Circuit-related genes relevant to single subclasses of inhibitory interneurons. CCA = cell-cell adhesion; RCCA = regulation of cell-cell adhesion; RTSS = regulation of trans-synaptic signaling.

Figure 9. Circuit-related genes only relevant to neuronal subclasses in VISp. A) Circuit-related genes relevant to single subclasses of excitatory projection neurons. We created the new subclass "IT" for genes was relevant to both L2/3 IT and L5 IT subclasses. **(B)** Circuit-related genes relevant to multiple subclasses of excitatory projection neurons. **(C)** Circuit-related genes relevant to single subclasses of inhibitory interneurons. CCA = cell-cell adhesion; RCCA = regulation of cell-cell adhesion; RTSS = regulation of trans-synaptic signaling.

Figure 10. Genes that have biased expression for all excitatory or inhibitory neuronal subclasses in ALM or VISp. (A) Genes with biased expression between ALM or VISp for excitatory projection neuron subclasses. Note that only 1 gene, *tenm2*, is circuit-related. **(B)** Genes with biased expression between ALM or VISp for inhibitory interneuron subclasses. (*Left*) Only 1 gene, *cenpa*, had biased expression among all subclasses. (*Right*) Only 4 genes demonstrated biases expression when comparing VIP to PV/SST subclasses. **(C)** Violin plots of *tenm2* single-cell expression for excitatory neuronal subclasses in ALM and VISp. **(D)** Violin plots of *cenpa* single-cell expression for inhibitory interneuron subclasses in ALM and VISp. **(E)** Violin plots of *igf1* single-cell expression for inhibitory interneuron subclasses in ALM and VISp.

FIGURE 1

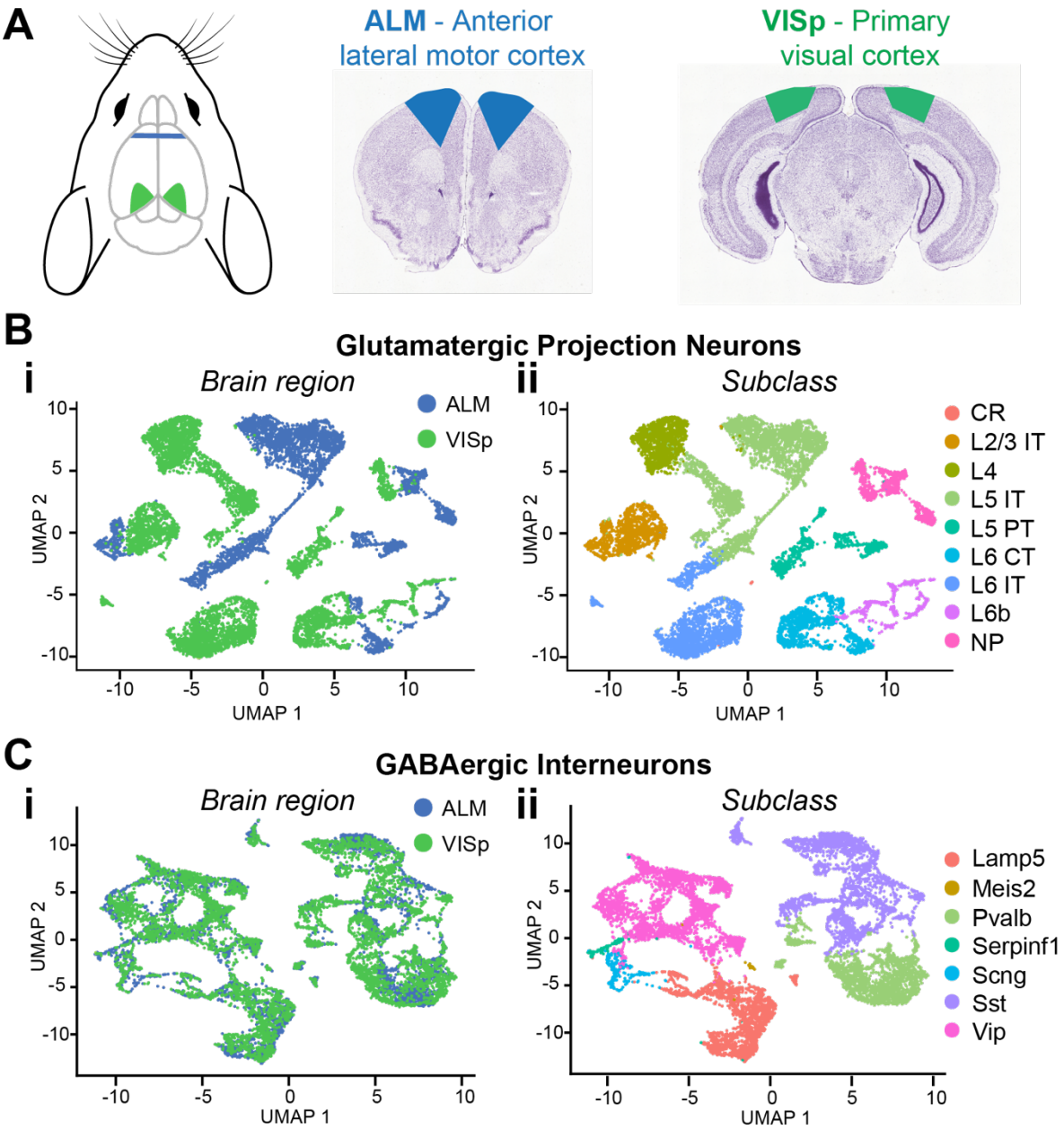

A
ALM - Anterior lateral motor cortex
VISp - Primary visual cortex
B
Glutamatergic Projection Neurons
i
Brain region
ALM
VISp
ii
Subclass
CR
L2/3 IT
L4
L5 IT
L5 PT
L6 CT
L6 IT
L6b
NP
UMAP 1
UMAP 2
C
GABAergic Interneurons
i
Brain region
ALM
VISp
ii
Subclass
Lamp5
Meis2
Pvalb
Serpinf1
Scng
Sst
Vip
UMAP 1
UMAP 2

FIGURE 2

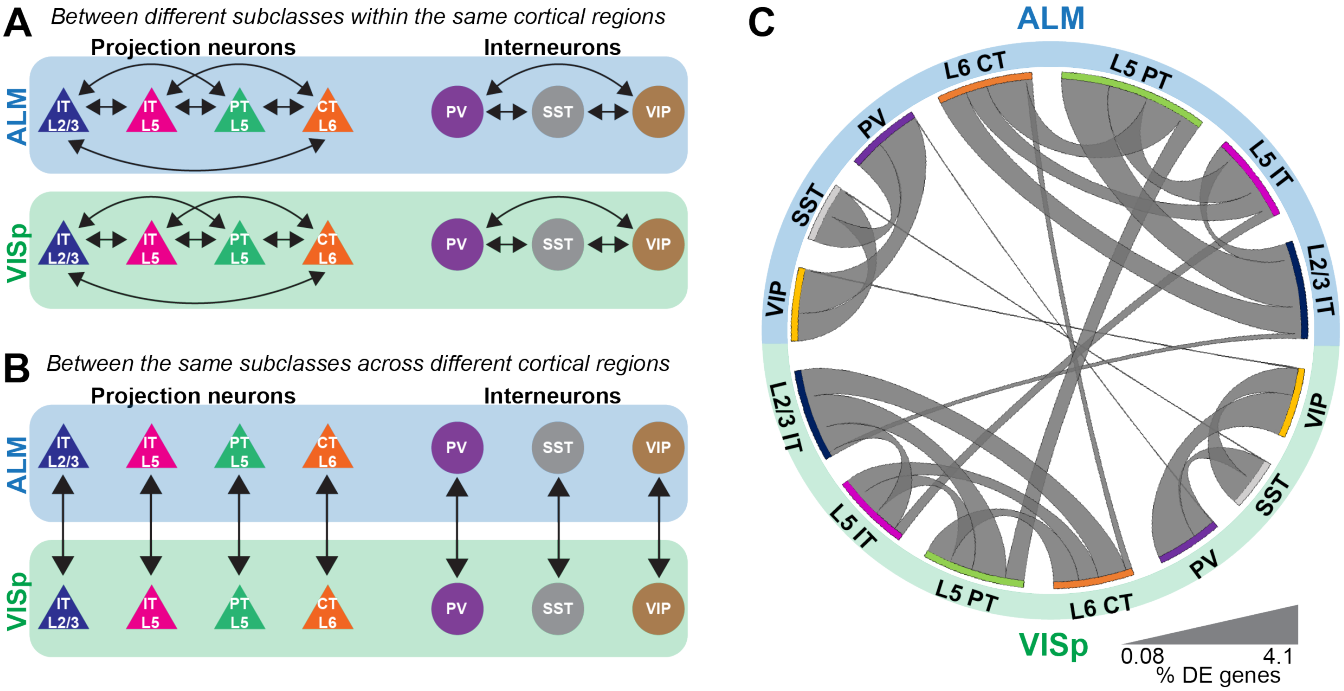

A
Between different subclasses within the same cortical regions
Projection neurons
Interneurons
ALM
VISp
IT L2/3
IT L5
PT L5
CT L6
PV
SST
VIP
B
Between the same subclasses across different cortical regions
Projection neurons
Interneurons
ALM
VISp
IT L2/3
IT L5
PT L5
CT L6
PV
SST
VIP
C
ALM
VISp
L6 CT
L5 PT
L5 IT
L2/3 IT
VIP
SST
PV
0.08
4.1
% DE genes

FIGURE 3

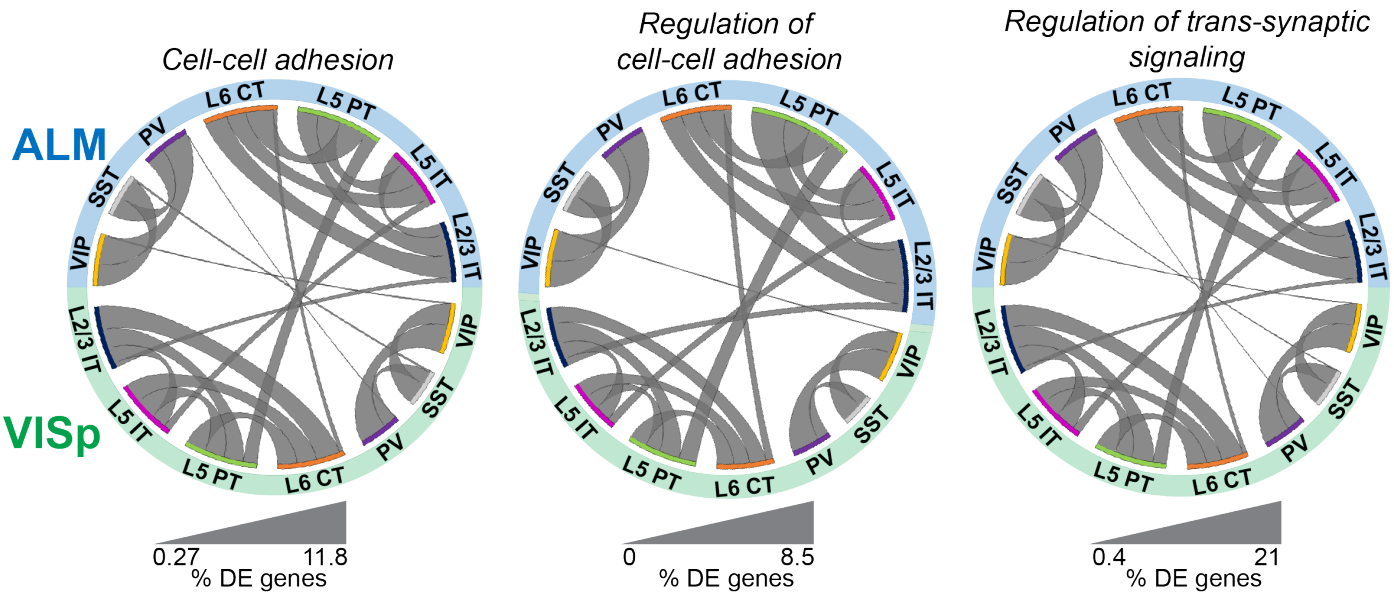

Cell-cell adhesion
Regulation of cell-cell adhesion
Regulation of trans-synaptic signaling
ALM
VISp
L6 CT
L5 PT
L5 IT
L2/3 IT
VIP
SST
PV
0.27
11.8
0
8.5
0.4
21
% DE genes

FIGURE 4

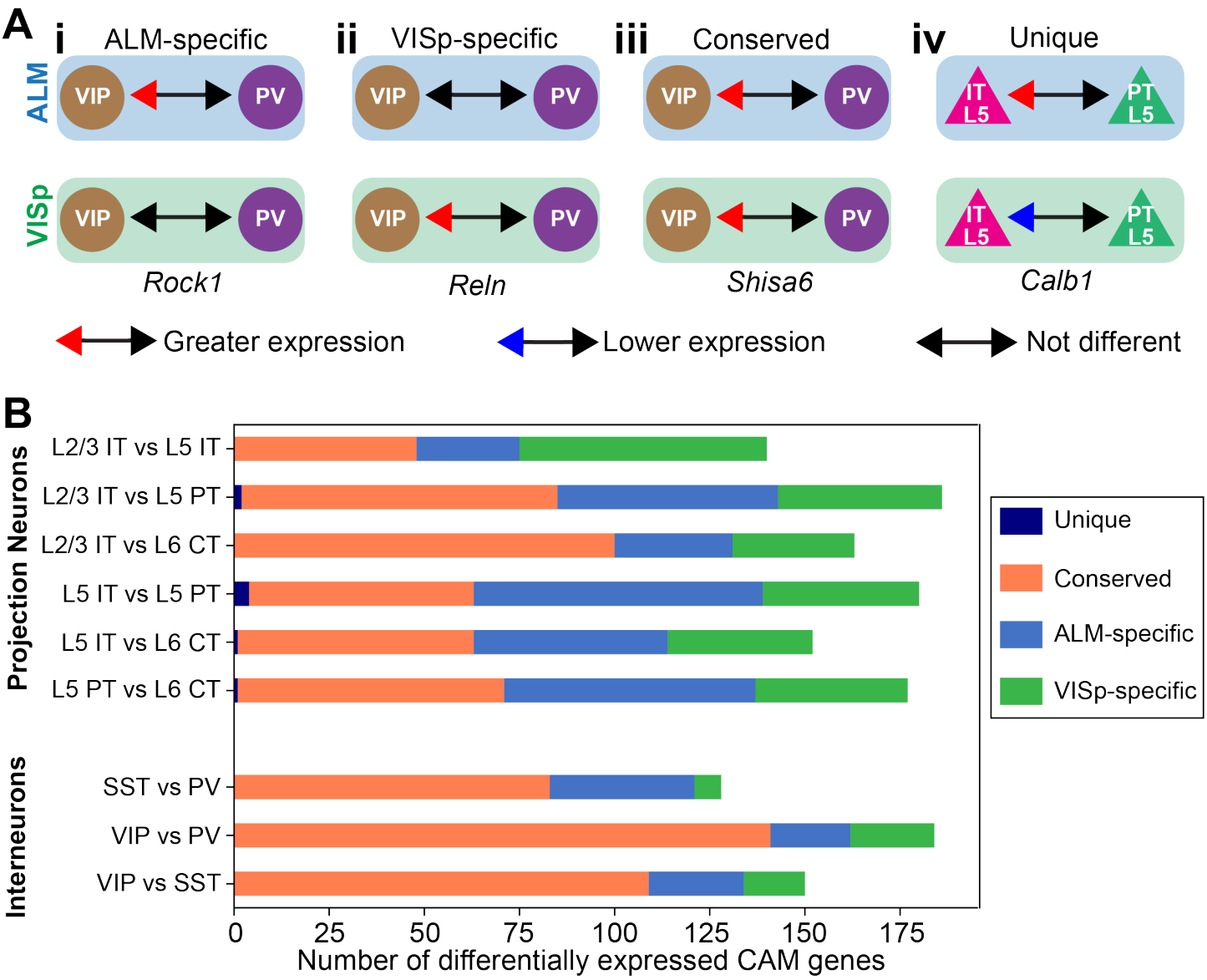

A
i ALM-specific
ii VISp-specific
iii Conserved
iv Unique
ALM
VISp
VIP
PV
IT L5
PT L5
Rock1
Reln
Shisa6
Calb1
Greater expression
Lower expression
Not different
B
Projection Neurons
Interneurons
L2/3 IT vs L5 IT
L2/3 IT vs L5 PT
L2/3 IT vs L6 CT
L5 IT vs L5 PT
L5 IT vs L6 CT
L5 PT vs L6 CT
SST vs PV
VIP vs PV
VIP vs SST
0
25
50
75
100
125
150
175
Number of differentially expressed CAM genes
Unique
Conserved
ALM-specific
VISp-specific

FIGURE 5

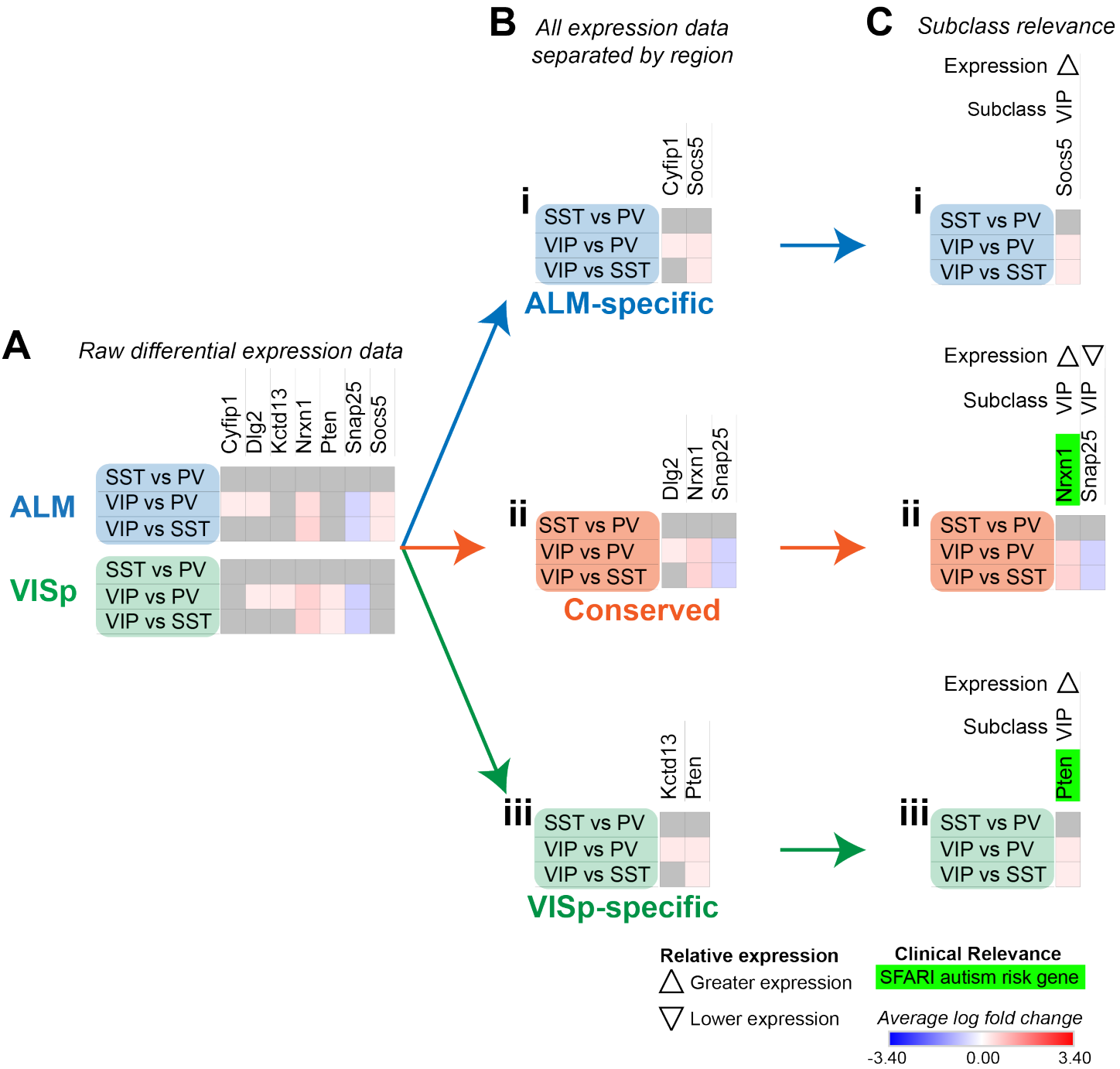

A
Raw differential expression data
Cyfip1
Dlg2
Kctd13
Nrxn1
Pten
Snap25
Socs5
ALM
SST vs PV
VIP vs PV
VIP vs SST
VISp
SST vs PV
VIP vs PV
VIP vs SST
B
All expression data separated by region
i
Cyfip1
Socs5
SST vs PV
VIP vs PV
VIP vs SST
ALM-specific
ii
Dlg2
Nrxn1
Snap25
SST vs PV
VIP vs PV
VIP vs SST
Conserved
iii
Kctd13
Pten
SST vs PV
VIP vs PV
VIP vs SST
VISp-specific
C
Subclass relevance
i
Expression
Subclass
VIP
Socs5
SST vs PV
VIP vs PV
VIP vs SST
ii
Expression
Subclass
VIP
VIP
Nrxn1
Snap25
SST vs PV
VIP vs PV
VIP vs SST
iii
Expression
Subclass
VIP
Pten
SST vs PV
VIP vs PV
VIP vs SST
Relative expression
Greater expression
Lower expression
Clinical Relevance
SFARI autism risk gene
Average log fold change
-3.40
0.00
3.40

FIGURE 6

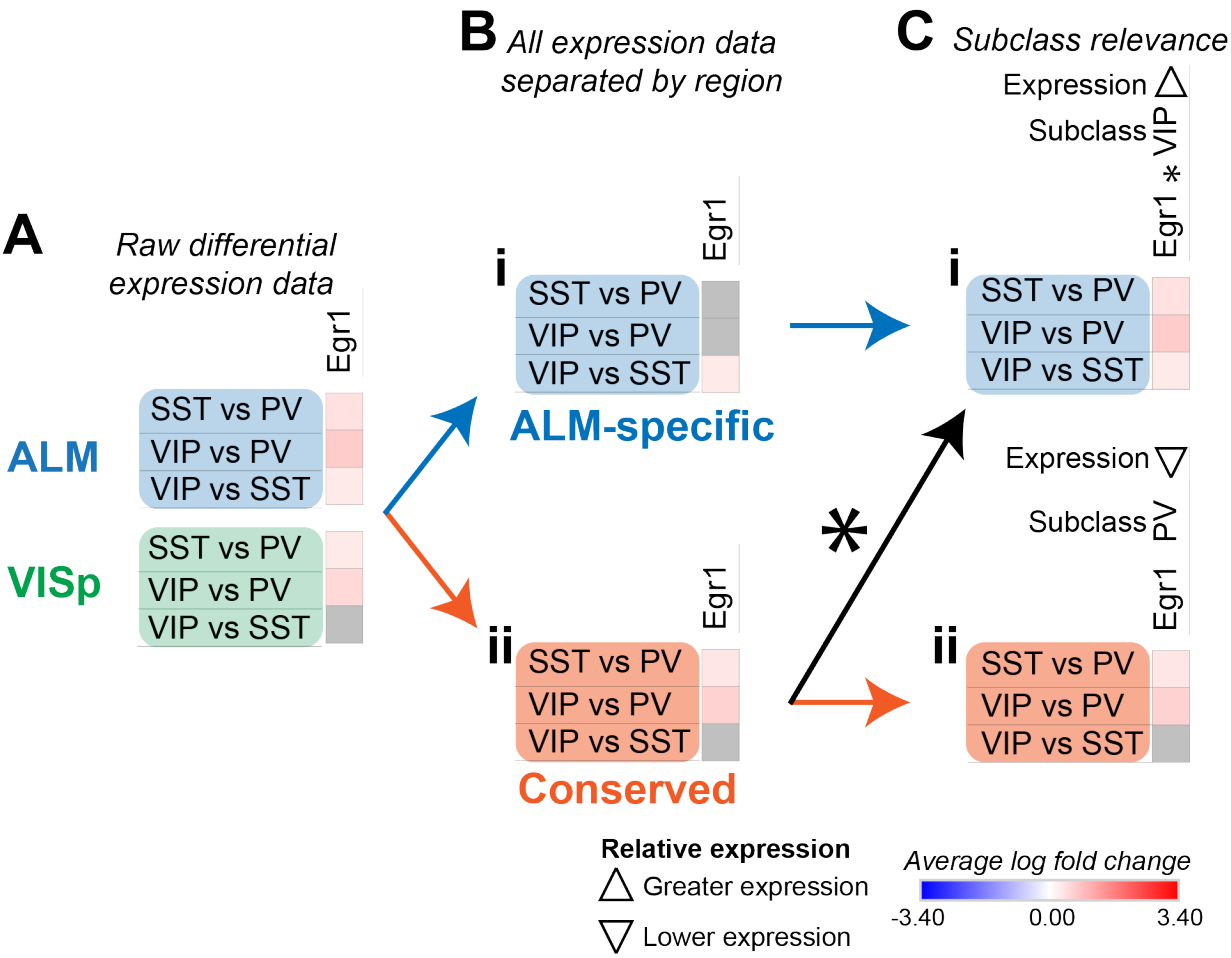
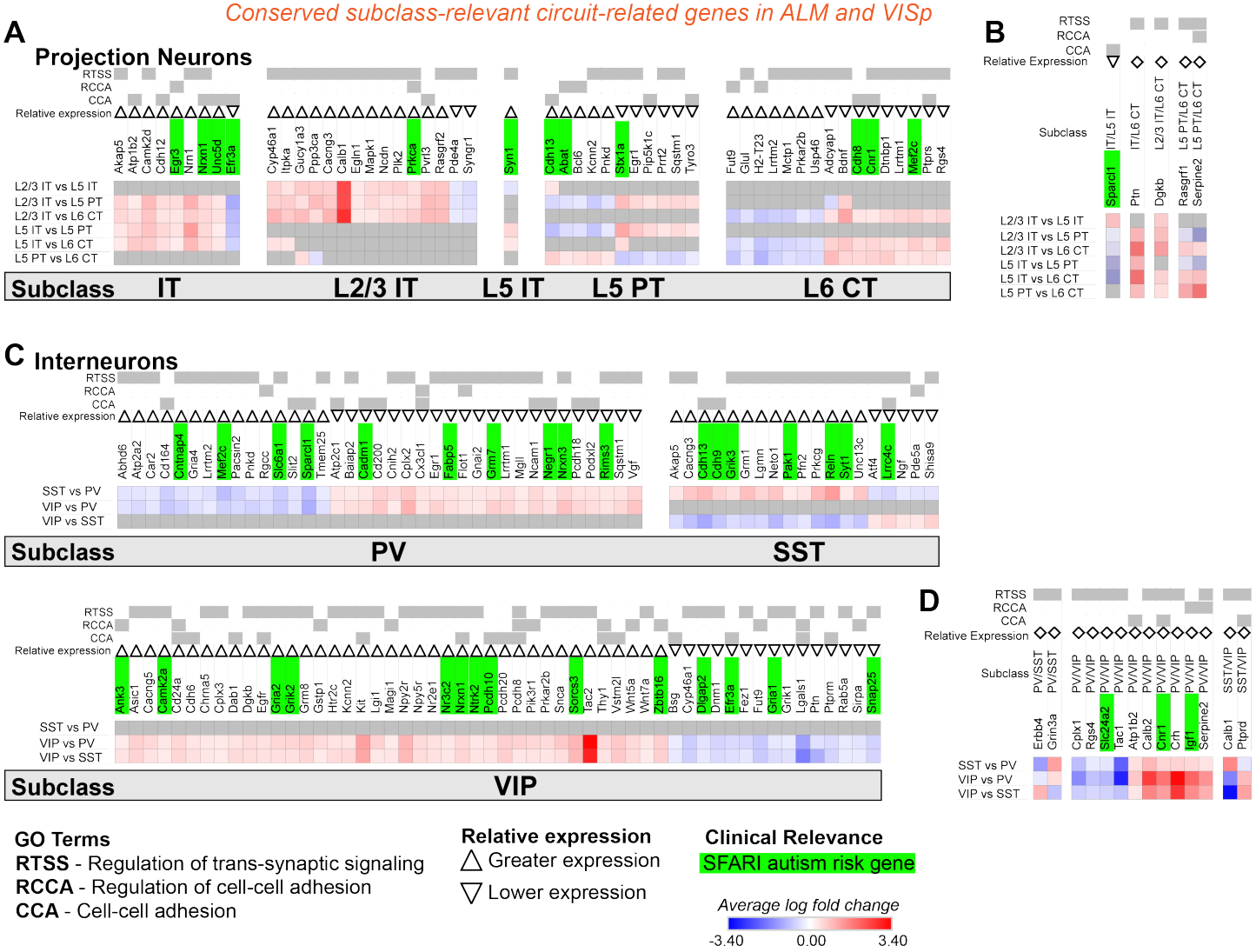

A
Raw differential expression data
Egr1
ALM
SST vs PV
VIP vs PV
VIP vs SST
VISp
SST vs PV
VIP vs PV
VIP vs SST
B
All expression data separated by region
i
Egr1
SST vs PV
VIP vs PV
VIP vs SST
ALM-specific
ii
Egr1
SST vs PV
VIP vs PV
VIP vs SST
Conserved
C
Subclass relevance
Expression
Subclass
VIP
Egr1 *
i
SST vs PV
VIP vs PV
VIP vs SST
Expression
Subclass
PV
Egr1
ii
SST vs PV
VIP vs PV
VIP vs SST
*
Relative expression
Greater expression
Lower expression
Average log fold change
-3.40
0.00
3.40

FIGURE 7

Conserved subclass-relevant circuit-related genes in ALM and VISp



GO Terms
RTSS - Regulation of trans-synaptic signaling
RCCA - Regulation of cell-cell adhesion
CCA - Cell-cell adhesion

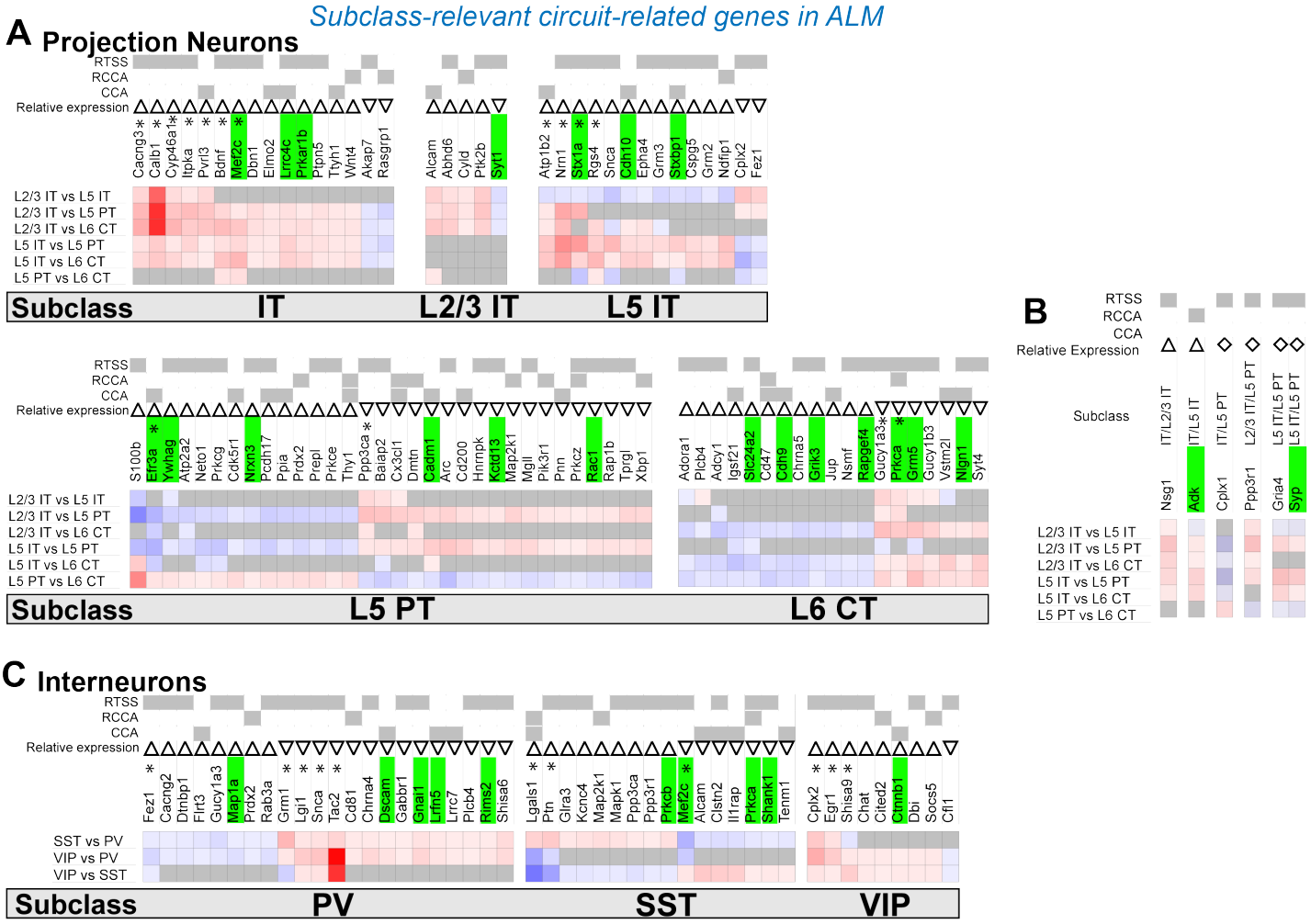
Relative expression
△ Greater expression
▽ Lower expression

Clinical Relevance
SFARI autism risk gene

Average log fold change
-3.40 0.00 3.40

FIGURE 8

Subclass-relevant circuit-related genes in ALM



GO Terms
RTSS - Regulation of trans-synaptic signaling
RCCA - Regulation of cell-cell adhesion
CCA - Cell-cell adhesion

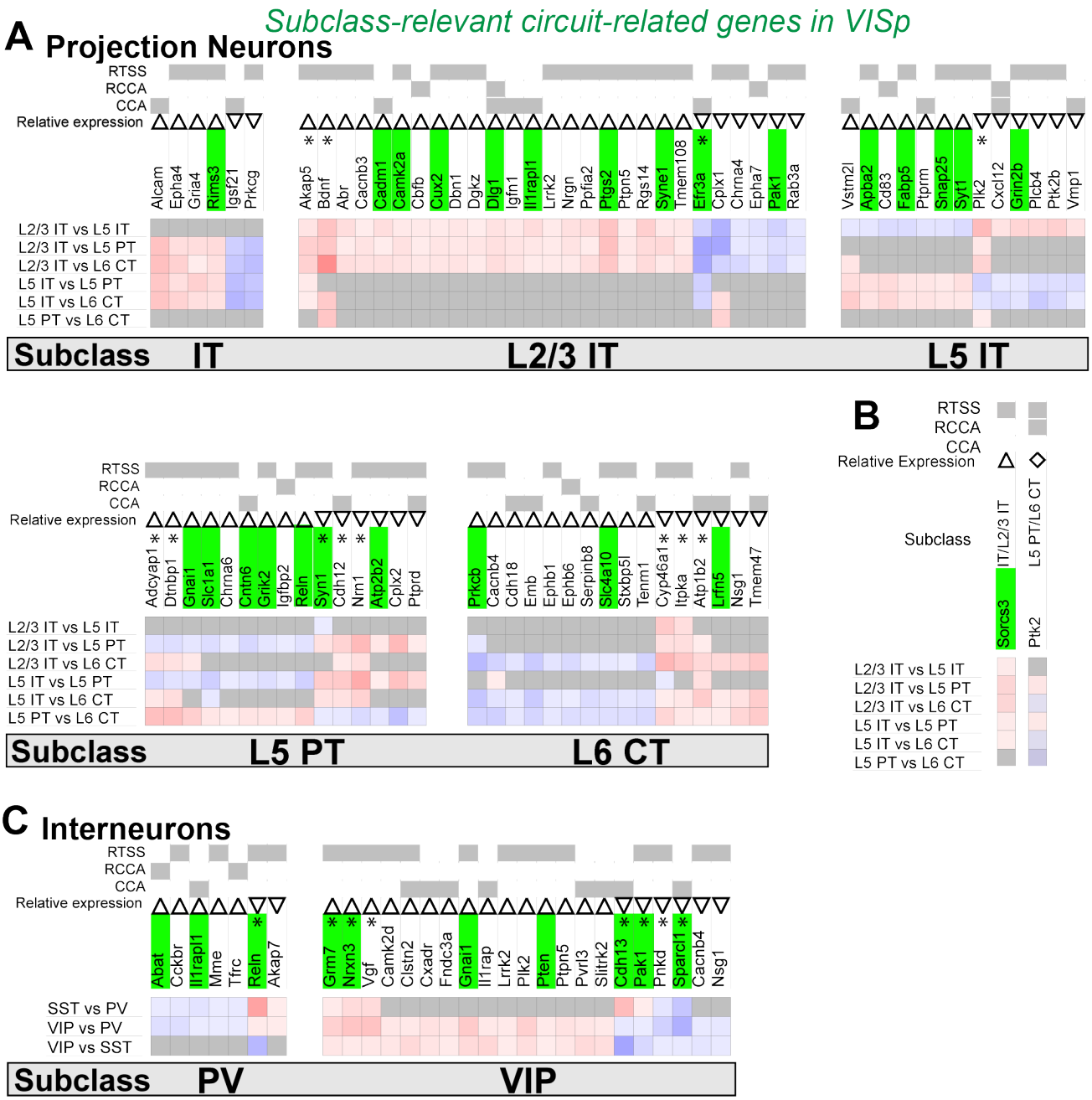
Relative expression
△ Greater expression
▽ Lower expression

Clinical Relevance
SFARI autism risk gene

Average log fold change
-3.40 0.00 3.40

Other
* Some expression tests conserved in ALM and VISp

FIGURE 9



GO Terms

RTSS - Regulation of trans-synaptic signaling

RCCA - Regulation of cell-cell adhesion

CCA - Cell-cell adhesion

Relative expression

△ Greater expression

▽ Lower expression

Clinical Relevance

SFARI autism risk gene

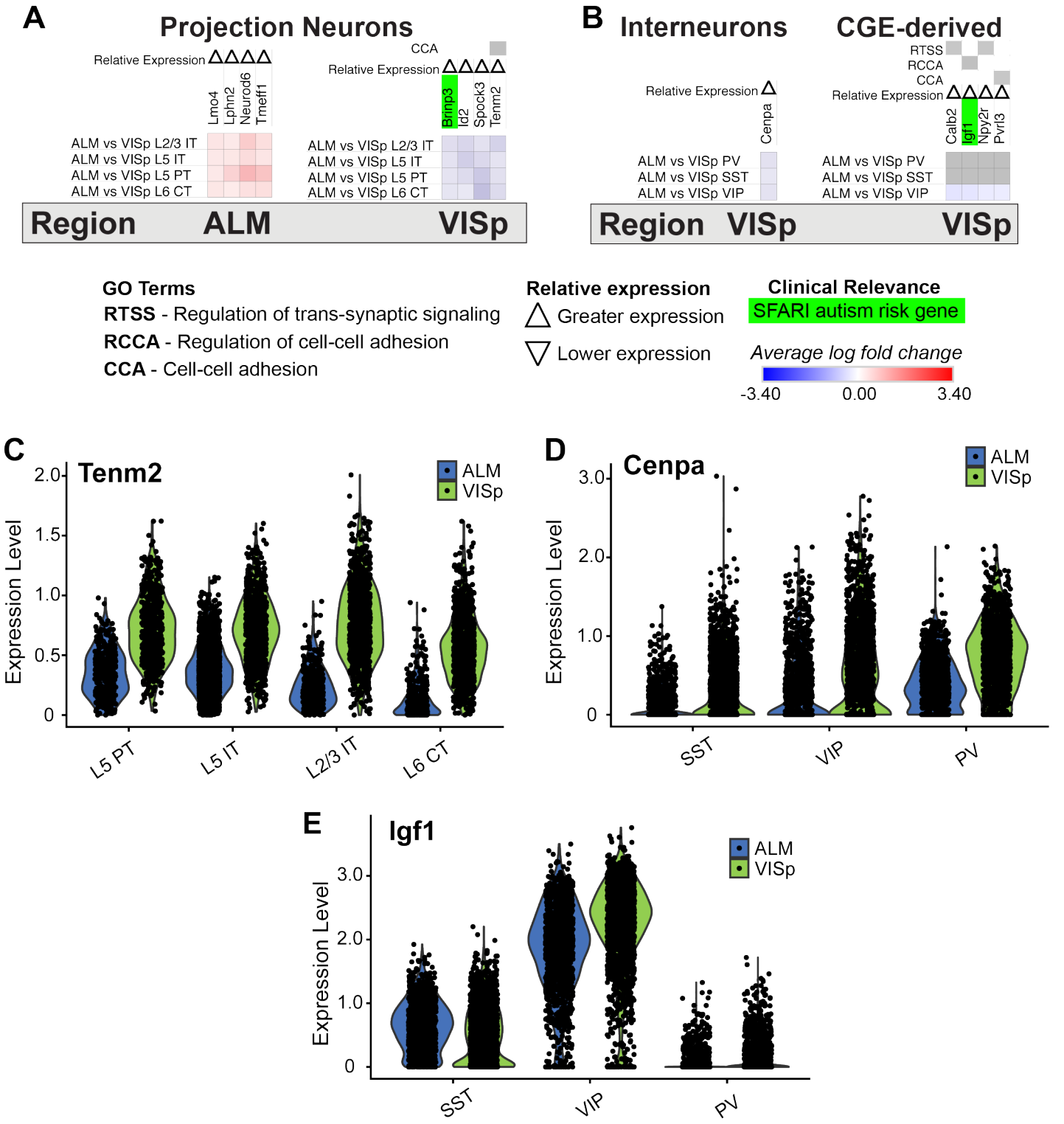
Average log fold change

-3.40 0.00 3.40

Other

* Some expression tests conserved in ALM and VISp

FIGURE 10


A
Projection Neurons
Relative Expression
Lmo4
Lphn2
Neurod6
Tmeff1
ALM vs VISp L2/3 IT
ALM vs VISp L5 IT
ALM vs VISp L5 PT
ALM vs VISp L6 CT
CCA
Relative Expression
Brinp3
Id2
Spock3
Tenm2
ALM vs VISp L2/3 IT
ALM vs VISp L5 IT
ALM vs VISp L5 PT
ALM vs VISp L6 CT
Region ALM VISp
B
Interneurons
Relative Expression
Cenpa
ALM vs VISp PV
ALM vs VISp SST
ALM vs VISp VIP
CGE-derived
RTSS
RCCA
CCA
Relative Expression
Calb2
Igf1
Npy2r
Pvrl3
ALM vs VISp PV
ALM vs VISp SST
ALM vs VISp VIP
Region VISp VISp
GO Terms
RTSS - Regulation of trans-synaptic signaling
RCCA - Regulation of cell-cell adhesion
CCA - Cell-cell adhesion
Relative expression
Greater expression
Lower expression
Clinical Relevance
SFARI autism risk gene
Average log fold change
-3.40
0.00
3.40
C
Tenm2
ALM
VISp
Expression Level
2.0
1.5
1.0
0.5
0
L5 PT
L5 IT
L2/3 IT
L6 CT
D
Cenpa
ALM
VISp
Expression Level
3.0
2.0
1.0
0
SST
VIP
PV
E
Igf1
ALM
VISp
Expression Level
3.0
2.0
1.0
0
SST
VIP
PV

REFERENCES

Bandler, R.C., Vitali, I., Delgado, R.N., Ho, M.C., Dvoretskova, E., Ibarra Molinas, J.S., Frazel, P.W., Mohammadkhani, M., Machold, R., Maedler, S., Liddelow, S.A., Nowakowski, T.J., Fishell, G., and Mayer, C. (2022). Single-cell delineation of lineage and genetic identity in the mouse brain. *Nature* 601**,** 404-409.

Batista-Brito, R., and Fishell, G. (2009). The developmental integration of cortical interneurons into a functional network. *Curr Top Dev Biol* 87**,** 81-118.

Bhaduri, A., Sandoval-Espinosa, C., Otero-Garcia, M., Oh, I., Yin, R., Eze, U.C., Nowakowski, T.J., and Kriegstein, A.R. (2021). An atlas of cortical arealization identifies dynamic molecular signatures. *Nature* 598**,** 200-204.

Butt, S.J., Fuccillo, M., Nery, S., Noctor, S., Kriegstein, A., Corbin, J.G., and Fishell, G. (2005). The temporal and spatial origins of cortical interneurons predict their physiological subtype. *Neuron* 48**,** 591-604.

Campagnola, L., Seeman, S.C., Chartrand, T., Kim, L., Hoggarth, A., Gamlin, C., Ito, S., Trinh, J., Davoudian, P., Radaelli, C., Kim, M.H., Hage, T., Braun, T., Alfiler, L., Andrade, J., Bohn, P., Dalley, R., Henry, A., Kebede, S., Alice, M., Sandman, D., Williams, G., Larsen, R., Teeter, C., Daigle, T.L., Berry, K., Dotson, N., Enstrom, R., Gorham, M., Hupp, M., Dingman Lee, S., Ngo, K., Nicovich, P.R., Potekhina, L., Ransford, S., Gary, A., Goldy, J., Mcmillen, D., Pham, T., Tieu, M., Siverts, L., Walker, M., Farrell, C., Schroedter, M., Slaughterbeck, C., Cobb, C., Ellenbogen, R., Gwinn, R.P., Keene, C.D., Ko, A.L., Ojemann, J.G., Silbergeld, D.L., Carey, D., Casper, T., Crichton, K., Clark, M., Dee, N., Ellingwood, L., Gloe, J., Kroll, M., Sulc, J., Tung, H., Wadhwani, K., Brouner, K., Egdorf, T., Maxwell, M., Mcgraw, M., Pom, C.A., Ruiz, A., Bomben, J., Feng, D., Hejazinia, N., Shi, S., Szafer, A., Wakeman, W., Phillips, J., Bernard, A., Esposito, L., D'orazi, F.D., Sunkin, S., Smith, K., Tasic, B., Arkhipov, A., Sorensen, S., Lein, E., Koch, C., Murphy, G., Zeng, H., and Jarsky, T. (2022). Local connectivity and synaptic dynamics in mouse and human neocortex. *Science* 375**,** eabj5861.

Cederquist, G.Y., Azim, E., Shnider, S.J., Padmanabhan, H., and Macklis, J.D. (2013). Lmo4 establishes rostral motor cortex projection neuron subtype diversity. *J Neurosci* 33**,** 6321-6332.

Chowdhury, D., Watters, K., and Biederer, T. (2021). Synaptic recognition molecules in development and disease. *Curr Top Dev Biol* 142**,** 319-370.

Cubelos, B., Sebastian-Serrano, A., Beccari, L., Calcagnotto, M.E., Cisneros, E., Kim, S., Dopazo, A., Alvarez-Dolado, M., Redondo, J.M., Bovolenta, P., Walsh, C.A., and Nieto, M. (2010). Cux1 and Cux2 regulate dendritic branching, spine morphology, and synapses of the upper layer neurons of the cortex. *Neuron* 66**,** 523-535.

Del Castillo, U., Norkett, R., and Gelfand, V.I. (2019). Unconventional Roles of Cytoskeletal Mitotic Machinery in Neurodevelopment. *Trends Cell Biol* 29**,** 901-911.

Demirdizen, E., Spiller-Becker, M., Fortsch, A., Wilhelm, A., Corless, S., Bade, D., Bergner, A., Hessling, B., and Erhardt, S. (2019). Localization of Drosophila CENP-A to non-centromeric sites depends on the NuRD complex. *Nucleic Acids Res* 47**,** 11589-11608.

Donovan, A.P., and Basson, M.A. (2017). The neuroanatomy of autism - a developmental perspective. *J Anat* 230**,** 4-15.

Ewall, G., Parkins, S., Lin, A., Jaoui, Y., and Lee, H.K. (2021). Cortical and Subcortical Circuits for Cross-Modal Plasticity Induced by Loss of Vision. *Front Neural Circuits* 15**,** 665009.

Földy, C., Darmanis, S., Aoto, J., Malenka, R.C., Quake, S.R., and Südhof, T.C. (2016). Single-cell RNAseq reveals cell adhesion molecule profiles in electrophysiologically defined neurons. *Proc Natl Acad Sci U S A* 113**,** E5222-5231.

Fuccillo, M.V., Földy, C., Gökce, Ö., Rothwell, P.E., Sun, G.L., Malenka, R.C., and Südhof, T.C. (2015). Single-Cell mRNA Profiling Reveals Cell-Type-Specific Expression of Neurexin Isoforms. *Neuron* 87**,** 326-340.

Garcia Del Arco, A., Edgar, B.A., and Erhardt, S. (2018). In Vivo Analysis of Centromeric Proteins Reveals a Stem Cell-Specific Asymmetry and an Essential Role in Differentiated, Non-proliferating Cells. *Cell Rep* 22**,** 1982-1993.

Greig, L.C., Woodworth, M.B., Galazo, M.J., Padmanabhan, H., and Macklis, J.D. (2013). Molecular logic of neocortical projection neuron specification, development and diversity. *Nat Rev Neurosci* 14**,** 755-769.

Gutman-Wei, A.Y., and Brown, S.P. (2021). Mechanisms Underlying Target Selectivity for Cell Types and Subcellular Domains in Developing Neocortical Circuits. *Front Neural Circuits* 15**,** 728832.

Harris, K.D., and Shepherd, G.M. (2015a). The neocortical circuit: themes and variations. *Nat Neurosci* 18**,** 170-181.

Harris, K.D., and Shepherd, G.M. (2015b). The neocortical circuit: themes and variations. *Nat Neurosci* 18**,** 170-181.

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S.J., and Marra, M.A. (2009). Circos: an information aesthetic for comparative genomics. *Genome Res* 19**,** 1639-1645.

Kubota, Y. (2014). Untangling GABAergic wiring in the cortical microcircuit. *Curr Opin Neurobiol* 26**,** 7-14.

Lee, S., Hjerling-Leffler, J., Zagha, E., Fishell, G., and Rudy, B. (2010). The largest group of superficial neocortical GABAergic interneurons expresses ionotropic serotonin receptors. *J Neurosci* 30**,** 16796-16808.

Llorca, A., and Marin, O. (2021). Orchestrated freedom: new insights into cortical neurogenesis. *Curr Opin Neurobiol* 66**,** 48-56.

Marco, E.J., Hinkley, L.B., Hill, S.S., and Nagarajan, S.S. (2011). Sensory processing in autism: a review of neurophysiologic findings. *Pediatr Res* 69**,** 48R-54R.

Mardinly, A.R., Spiegel, I., Patrizi, A., Centofante, E., Bazinet, J.E., Tzeng, C.P., Mandel-Brehm, C., Harmin, D.A., Adesnik, H., Fagiolini, M., and Greenberg, M.E. (2016). Sensory experience regulates cortical inhibition by inducing IGF1 in VIP neurons. *Nature* 531**,** 371-375.

Mayer, C., Hafemeister, C., Bandler, R.C., Machold, R., Batista Brito, R., Jaglin, X., Allaway, K., Butler, A., Fishell, G., and Satija, R. (2018). Developmental diversification of cortical inhibitory interneurons. *Nature* 555**,** 457-462.

Mi, H., Ebert, D., Muruganujan, A., Mills, C., Albou, L.P., Mushayamaha, T., and Thomas, P.D. (2021). PANTHER version 16: a revised family classification, tree-based classification tool, enhancer regions and extensive API. *Nucleic Acids Res* 49**,** D394-D403.

Miyoshi, G., Hjerling-Leffler, J., Karayannis, T., Sousa, V.H., Butt, S.J., Battiste, J., Johnson, J.E., Machold, R.P., and Fishell, G. (2010). Genetic fate mapping reveals that the caudal ganglionic eminence produces a large and diverse population of superficial cortical interneurons. *J Neurosci* 30**,** 1582-1594.

Nano, P.R., Nguyen, C.V., Mil, J., and Bhaduri, A. (2021). Cortical Cartography: Mapping Arealization Using Single-Cell Omics Technology. *Front Neural Circuits* 15**,** 788560.

Pai, E.L., Chen, J., Fazel Darbandi, S., Cho, F.S., Chen, J., Lindtner, S., Chu, J.S., Paz, J.T., Vogt, D., Paredes, M.F., and Rubenstein, J.L. (2020). Maf and Mafb control mouse pallial interneuron fate and maturation through neuropsychiatric disease gene regulation. *Elife* 9.

Parikshak, N.N., Luo, R., Zhang, A., Won, H., Lowe, J.K., Chandran, V., Horvath, S., and Geschwind, D.H. (2013). Integrative functional genomic analyses implicate specific molecular pathways and circuits in autism. *Cell* 155**,** 1008-1021.

Pouchelon, G., Dwivedi, D., Bollmann, Y., Agba, C.K., Xu, Q., Mirow, A.M.C., Kim, S., Qiu, Y., Sevier, E., Ritola, K.D., Cossart, R., and Fishell, G. (2021). The organization and development of cortical interneuron presynaptic circuits are area specific. *Cell Rep* 37**,** 109993.

Rubenstein, J.L., Anderson, S., Shi, L., Miyashita-Lin, E., Bulfone, A., and Hevner, R. (1999). Genetic control of cortical regionalization and connectivity. *Cereb Cortex* 9**,** 524-532.

Sohal, V.S., and Rubenstein, J.L.R. (2019). Excitation-inhibition balance as a framework for investigating mechanisms in neuropsychiatric disorders. *Mol Psychiatry* 24**,** 1248-1257.

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., 3rd, Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive Integration of Single-Cell Data. *Cell* 177**,** 1888-1902 e1821.

Suarez, R., Gobius, I., and Richards, L.J. (2014). Evolution and development of interhemispheric connections in the vertebrate forebrain. *Front Hum Neurosci* 8**,** 497.

Tasic, B., Yao, Z., Graybuck, L.T., Smith, K.A., Nguyen, T.N., Bertagnolli, D., Goldy, J., Garren, E., Economo, M.N., Viswanathan, S., Penn, O., Bakken, T., Menon, V., Miller, J., Fong, O., Hirokawa, K.E., Lathia, K., Rimorin, C., Tieu, M., Larsen, R., Casper, T., Barkan, E., Kroll, M., Parry, S., Shapovalova, N.V., Hirschstein, D., Pendergraft, J., Sullivan, H.A., Kim, T.K., Szafer, A., Dee, N., Groblewski, P., Wickersham, I., Cetin, A., Harris, J.A., Levi, B.P., Sunkin, S.M., Madisen, L., Daigle, T.L., Looger, L., Bernard, A., Phillips, J., Lein, E., Hawrylycz, M., Svoboda, K., Jones, A.R., Koch, C., and Zeng, H. (2018). Shared and distinct transcriptomic cell types across neocortical areas. *Nature* 563**,** 72-78.

Tosches, M.A., Yamawaki, T.M., Naumann, R.K., Jacobi, A.A., Tushev, G., and Laurent, G. (2018). Evolution of pallium, hippocampus, and cortical cell types revealed by single-cell transcriptomics in reptiles. *Science* 360**,** 881-888.

Tremblay, R., Lee, S., and Rudy, B. (2016). GABAergic Interneurons in the Neocortex: From Cellular Properties to Circuits. *Neuron* 91**,** 260-292.

Velmeshev, D., Schirmer, L., Jung, D., Haeussler, M., Perez, Y., Mayer, S., Bhaduri, A., Goyal, N., Rowitch, D.H., and Kriegstein, A.R. (2019). Single-cell genomics identifies cell type-specific molecular changes in autism. *Science* 364**,** 685-689.

Wester, J.C., Mahadevan, V., Rhodes, C.T., Calvigioni, D., Venkatesh, S., Maric, D., Hunt, S., Yuan, X., Zhang, Y., Petros, T.J., and Mcbain, C.J. (2019). Neocortical Projection Neurons Instruct Inhibitory Interneuron Circuit Development in a Lineage-Dependent Manner. *Neuron* 102**,** 960-975 e966.

Whyatt, C., and Craig, C. (2013). Sensory-motor problems in Autism. *Front Integr Neurosci* 7**,** 51.

Xu, Q., Cobos, I., De La Cruz, E., Rubenstein, J.L., and Anderson, S.A. (2004). Origins of cortical interneuron subtypes. *J Neurosci* 24**,** 2612-2622.

Yao, Z., Van Velthoven, C.T.J., Nguyen, T.N., Goldy, J., Sedeno-Cortes, A.E., Baftizadeh, F., Bertagnolli, D., Casper, T., Chiang, M., Crichton, K., Ding, S.L., Fong, O., Garren, E., Glandon, A., Gouwens, N.W., Gray, J., Graybuck, L.T., Hawrylycz, M.J., Hirschstein, D., Kroll, M., Lathia, K., Lee, C., Levi, B., Mcmillen, D., Mok, S., Pham, T., Ren, Q., Rimorin, C.,

Shapovalova, N., Sulc, J., Sunkin, S.M., Tieu, M., Torkelson, A., Tung, H., Ward, K., Dee, N., Smith, K.A., Tasic, B., and Zeng, H. (2021a). A taxonomy of transcriptomic cell types across the isocortex and hippocampal formation. *Cell* 184**,** 3222-3241.e3226.

Yao, Z., Van Velthoven, C.T.J., Nguyen, T.N., Goldy, J., Sedeno-Cortes, A.E., Baftizadeh, F., Bertagnolli, D., Casper, T., Chiang, M., Crichton, K., Ding, S.L., Fong, O., Garren, E., Glandon, A., Gouwens, N.W., Gray, J., Graybuck, L.T., Hawrylycz, M.J., Hirschstein, D., Kroll, M., Lathia, K., Lee, C., Levi, B., Mcmillen, D., Mok, S., Pham, T., Ren, Q., Rimorin, C., Shapovalova, N., Sulc, J., Sunkin, S.M., Tieu, M., Torkelson, A., Tung, H., Ward, K., Dee, N., Smith, K.A., Tasic, B., and Zeng, H. (2021b). A taxonomy of transcriptomic cell types across the isocortex and hippocampal formation. *Cell* 184**,** 3222-3241 e3226.

Ye, Z., Mostajo-Radji, M.A., Brown, J.R., Rouaux, C., Tomassy, G.S., Hensch, T.K., and Arlotta, P. (2015). Instructing Perisomatic Inhibition by Direct Lineage Reprogramming of Neocortical Projection Neurons. *Neuron* 88**,** 475-483.

Young, T.R., Bourke, M., Zhou, X., Oohashi, T., Sawatari, A., Fassler, R., and Leamey, C.A. (2013). Ten-m2 is required for the generation of binocular visual circuits. *J Neurosci* 33**,** 12490-12509.

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

Zhao, G., Oztan, A., Ye, Y., and Schwarz, T.L. (2019). Kinetochore Proteins Have a Post-Mitotic Function in Neurodevelopment. *Dev Cell* 48**,** 873-882 e874.

Zhou, K., Gebala, M., Woods, D., Sundararajan, K., Edwards, G., Krzizike, D., Wereszczynski, J., Straight, A.F., and Luger, K. (2022). CENP-N promotes the compaction of centromeric chromatin. *Nat Struct Mol Biol* 29**,** 403-413.