

Translational Insights From Cell Type Variation Across Amygdala Subnuclei in Rhesus Monkeys and Humans

Shawn Kamboj, B.S., Erin L. Carlson, M.Sc., Bradley P. Ander, Ph.D., Kari L. Hanson, Ph.D., Karl D. Murray, Ph.D., Julie L. Fudge, M.D., Melissa D. Bauman, Ph.D., Cynthia M. Schumann, Ph.D., Andrew S. Fox, Ph.D.

Objective: Theories of amygdala function are central to our understanding of psychiatric and neurodevelopmental disorders. However, limited knowledge of the molecular and cellular composition of the amygdala impedes translational research aimed at developing new treatments and interventions. The aim of this study was to characterize and compare the composition of amygdala cells to help bridge the gap between preclinical models and human psychiatric and neurodevelopmental disorders.

Methods: Tissue was dissected from multiple amygdala subnuclei in both humans ($N=3$, male) and rhesus macaques ($N=3$, male). Single-nucleus RNA sequencing was performed to characterize the transcriptomes of individual nuclei.

Results: The results reveal substantial heterogeneity between regions, even when restricted to inhibitory or excitatory neurons. Consistent with previous work, the data highlight the complexities of individual marker genes for uniquely targeting specific cell types. Cross-species analyses

suggest that the rhesus monkey model is well-suited to understanding the human amygdala, but also identify limitations. For example, a cell cluster in the ventral lateral nucleus of the amygdala (vLa) is enriched in humans relative to rhesus macaques. Additionally, the data describe specific cell clusters with relative enrichment of disorder-related genes. These analyses point to the human-enriched vLa cell cluster as relevant to autism spectrum disorder, potentially highlighting a vulnerability to neurodevelopmental disorders that has emerged in recent primate evolution. Further, a cluster of cells expressing markers for intercalated cells is enriched for genes reported in human genome-wide association studies of neuroticism, anxiety disorders, and depressive disorders.

Conclusions: Together, these findings shed light on the composition of the amygdala and identify specific cell types that can be prioritized in basic science research to better understand human psychopathology and guide the development of potential treatments.

AJP in Advance (doi: 10.1176/appi.ajp.20230602)

Translational psychiatry is aimed at leveraging insights from basic science to understand and treat psychiatric disorders. Understanding the function of specific cell types in disorder-relevant brain regions has the potential to guide the development of new treatments. However, there is an increasingly large gap in our collective understanding of molecules and cells in humans relative to rodents, providing a barrier for cross-species translational research. In particular, there remains a disconnect in how within-brain heterogeneity is considered across studies of humans and nonhuman animals. Human studies often focus on the contributions of specific brain regions, while animal studies are increasingly focused on cellular and molecular heterogeneity within anatomically specific subregions containing individual cell types that have distinct and opposing functions. Translating insights from animal studies to human treatments will require changing how we consider cellular heterogeneity. Here, we demonstrate how cross-species single-nucleus RNA sequencing

(snRNA-seq) data can be used to guide future translational and reverse translational research in animal models and increase our understanding of disease-relevant cellular heterogeneity in humans. These insights are critical for increasing the effectiveness of translational research, as optimal treatments for psychopathology are likely to require both regional and cell type specificity. We present new data and analyses focusing on the human and non-human primate amygdala to glean insight into amygdala-relevant disorders and identify barriers that will need to be considered to leverage insights from basic science in the clinic.

The amygdala is central to modulating socioemotional brain circuitry and has been implicated in disorders from anxiety and depression to autism spectrum disorder (ASD) and psychosis spectrum disorders (1–7). As a result, many theories about the pathophysiology and potential treatments for psychiatric and neurodevelopmental disorders

have focused on the amygdala (5, 8, 9). These theories have motivated efforts to use neuroimaging to identify amygdala biomarkers in humans (e.g., 10, 11) and develop a mechanistic understanding of amygdala cell types in rodents (12, 13). Thus far these efforts have largely failed to produce actionable treatment targets. In part, this reflects our limited knowledge of the regional and molecular heterogeneity within the amygdala and limited cross-pollination between animal and human studies. Here, we take a translational approach to understand amygdala heterogeneity and its relevance to psychiatric and neurodevelopmental disease.

The amygdala is an anatomical concept used to describe an almond-shaped mass of gray matter in the medial temporal lobe. As early as 1923, it was recognized that the amygdala is comprised of a heterogeneous set of approximately 13 subnuclei, including the central amygdala (Ce) and lateral amygdala (La), as well as “intercalated” cell masses (ITCs) that are distributed throughout the amygdala and not confined to a particular subnucleus (14). By the late 20th century, amygdala subnuclei were classified based on their cellular composition (cell type, markers, developmental lineage) (15). Different amygdala subnuclei can be parcelled into at least three distinct functional units that are molecularly distinct and send and receive projections from different brain structures (15). More recently, it has been shown that amygdala subnuclei contain substantive cellular and molecular heterogeneity (12, 13, 16). The functional and molecular heterogeneity is a major limitation for translational research, which often focuses on the amygdala as a single functional unit, overlooking the potential for individual molecules to play distinct functional roles across amygdala subnuclei. Studies of the developmental heterogeneity of amygdala subnuclei (17–20), along with recent efforts to dissect amygdala microcircuits (12, 13), reinforce the importance of distinguishing between subnuclei.

Amygdala subnuclei contain cells that develop from distinct progenitors across multiple embryonic divisions, resulting in robust differences in cytoarchitecture and neuronal composition. For example, the La, basal nucleus (Ba), and periamygdaloid cortices (PAC) largely receive neuronal progenitors from the pallium, while the Ce, ITCs, and medial nucleus of the amygdala (Me) primarily derive from ganglionic eminences within the subpallium (20). As a result, La, Ba, and PAC are primarily “cortical-like” regions that contain glutamatergic neurons and a significantly smaller proportion of GABAergic cell types. In contrast, the reverse is true for Ce, ITC, and Me, which primarily contain “striatal-like” GABAergic neurons. In addition, even within a specific subnucleus, different cell types can perform different functions. For example, when stimulated in the same task, Ce-expressing somatostatin and corticotropin-releasing hormone neurons can elicit mutually exclusive behaviors (i.e., freezing and escape, respectively) (13). Similarly, stimulation studies, in some contexts, suggest that distinct sets of intermingled cells in rodent La and Ba are differentially involved in threat and reward learning (12, 21). Although it is

unclear whether these findings are temporally and contextually specific, it is clear that translational research aiming to identify treatments for psychiatric disorders may have to take both subnucleus and cell type into account.

Further complicating translational research is that very little is known about the extent to which the molecular and cellular composition of the amygdala is conserved across species (22–24). The rodent and primate lineages diverged ~75 million years ago, allowing for substantial divergence in the composition of specific brain regions, making it unclear whether findings in rodents, which are often based on molecular marker genes, will effectively translate to humans. The rhesus monkey is one of the most common research primates, and it shares a common evolutionary ancestor with humans, only 25 million years ago (25). This recent evolutionary divergence has resulted in conserved social-behavioral repertoire, physiology, and neurobiology between humans and rhesus monkeys (26, 27), underscoring the potential of the rhesus monkey model for translational research.

Here, to gain insight into the cellular and molecular composition of different regions of the amygdala, we performed snRNA-seq from the “striatal-like” Ce and the “cortical-like” ventral La (vLa), in human and rhesus macaque. Although these regions are often combined in human amygdala studies, rodent studies demonstrate that they have distinct developmental origins, and their distinct spatial locations allow them to be examined separately using neuroimaging. These data allow for comparisons of the cellular composition between regions and species, with conceptual relevance for the development of amygdala-focused treatments. Finally, we use these data to identify potential cellular and molecular targets that have been implicated in psychiatric and developmental disorders. We show that these data, although not capturing a comprehensive taxonomy of amygdala cell types, can provide guidance for translational research by motivating and prioritizing mechanistic studies aimed at developing novel approaches to treatment and prevention of amygdala-implicated disorders.

METHODS

Summary

Single-nucleus RNA sequencing analyses were performed in the La and Ce across three humans and three rhesus macaques (see Table S1 in the online supplement) (detailed methods are described below). Individual samples were collected by performing small punches from frozen slabs and processed using the Parse Biosciences Nuclei Fixation and Whole Transcriptome Kits (SplitBio, SB1003; EC-W01030). Reads were aligned to the RheMac10 genome assembly and annotated using custom annotation built using the Comparative Annotation Toolkit (CAT). Data were processed in Python using Scanpy. Cell clusters were annotated according to the following: [Region or Species][Leiden Cluster Number][Cell Class][Developmental Origin (if Neuron)][Unique Marker or Overexpressed Gene]. Cells named based on

overexpressed genes, rather than a unique marker, are indicated by an asterisk. Cell cluster-specific differences in gene expression were computed using differential expression analyses between cells or cell clusters. Cell proportions were compared using Bayesian estimates of cell proportion (scCODA). Cross-species correspondence was assessed using rank-rank hypergeometric testing and MetaNeighbor. Permutation tests were performed to identify disease-relevant cell clusters using genes from previously published genome-wide gene-association studies. RNAscope *in situ* hybridization was performed to confirm coexpression of ITC markers.

Subjects

Subjects included three from each species in the final analyses. Samples were collected from four male rhesus macaques (*Macaca mulatta*; age range, 3.5–4 years) and four male human donors (age range, 15–19 years; mean post-mortem interval, 23.5 hours [range, 8–32 hours]). To understand heterogeneity across the primate amygdala, tissue samples were collected from the corresponding areas of the dorsal amygdala Ce region and the ventral portions of the La region (vLa) in both primate species. Rhesus tissue was obtained from the California National Primate Research Center. The right hemisphere of the rhesus macaque was immediately cut into eight coronal sections and flash frozen with liquid nitrogen. All study procedures were performed in accordance with the guidelines set forth by the University of California Davis Institutional Animal Care and Use Committee (IACUC). Human tissue was obtained from the Brain Endowment for Mental Health (University of California, Davis). Alternating hemispheres across cases are either fixed or coronally blocked into 1-cm-thick slabs and flash frozen according to our established protocols (8, 28). All experiments were approved by the Institutional Review Board and the IACUC at the University of California, Davis. Additional information on tissue collection and how the animals were handled prior to experimentation can be found in a previous study that utilized the same rhesus cohort (29).

Samples and Data Collection

Fresh frozen tissue blocks from one hemisphere per case containing amygdala were subsequently flash frozen in OCT before being sectioned on a cryostat in order to bring the block face to the same rostrocaudal level across all cases (Figure 1A,C,G). Sections representing the block surface were stained with the Acetylcholinesterase Rapid Staining Kit (MBL International, no. 8450) according to manufacturer's recommendations for unfixed sections to confirm rostrocaudal depth and to locate boundaries of individual amygdala subnuclei (Figure 1C,G).

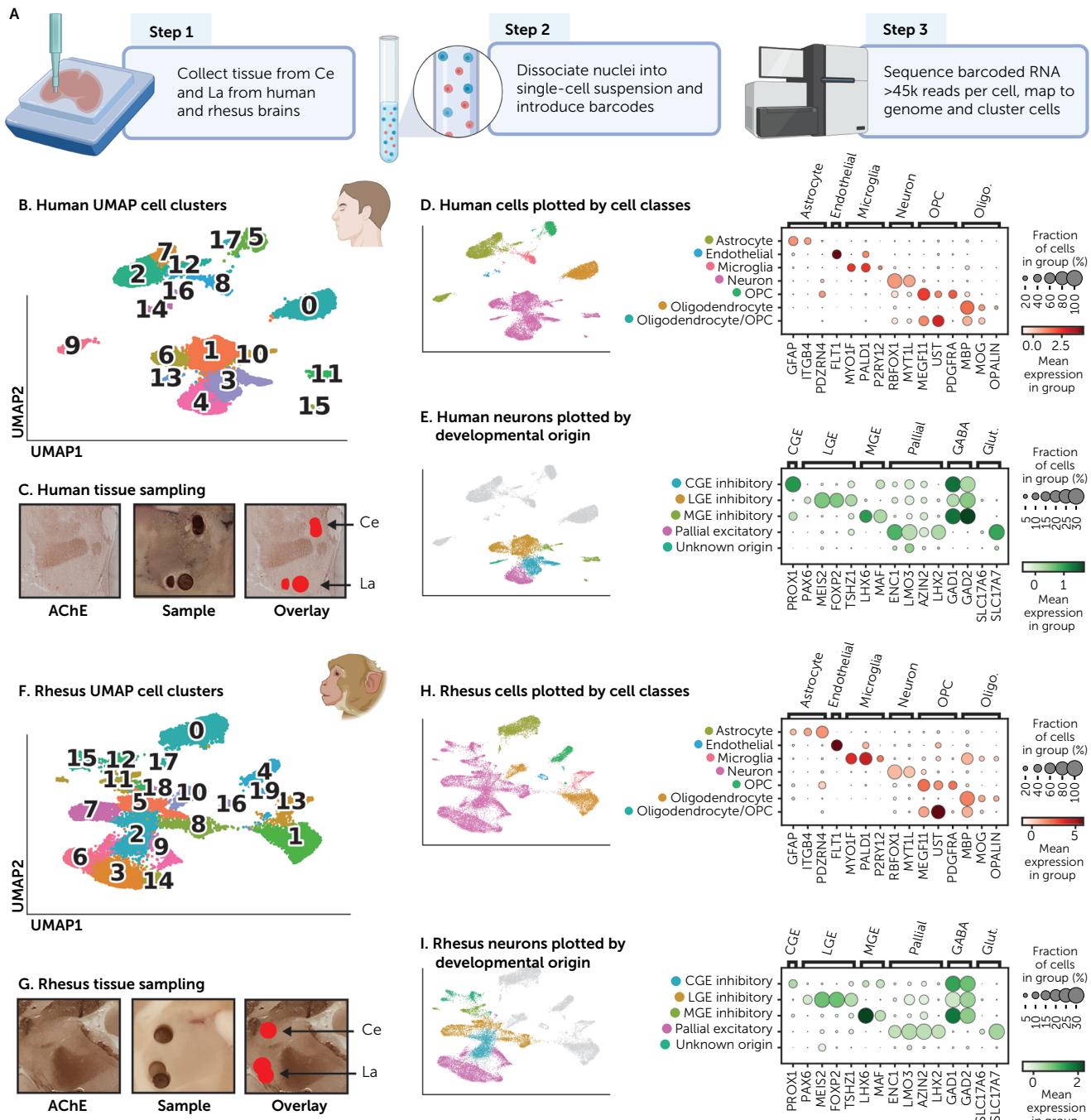
Under the guidance of expert primate neuroanatomists (C.M.S., A.S.F., J.L.F.), 1.5-mm tissue biopsy punches at a depth of 2 to 3 mm were used to retrieve samples from Ce and vLa, including embedded ITCs. One punch was performed per region per case, unless a shallow first punch was taken. If

a shallow first punch was performed, a second overlapping punch was taken if deemed appropriate by our expert primate neuroanatomists. Samples were placed in a Biomasher II Micro Tissue Homogenizer (Kimble) and immediately processed for nucleus preparations (Figure 1A). Single-nucleus suspensions were made following minor modifications to the 10x Genomics demonstrated protocol "Nuclei Isolation From Cell Suspensions and Tissue for Single Cell RNA Sequencing" (CG000124 revE). Briefly, 1 mL of lysis buffer (10 mM Tris-HCl pH7.4, 10 mM NaCl, 3 mM MgCl₂, 0.1% Nonidet P40 substitute, in nuclease-free water) was added to each sample, homogenized 20×, and incubated on ice for 5 minutes. Samples were centrifuged at 850 RCF for 5 minutes at 4°C. The resulting pellet was rinsed three times by resuspending in 1.5% BSA in PBS and centrifuged at 850 RCF for 5 minutes at 4°C. Finally, the pellet was resuspended in 250 μL 1.5% BSA in PBS and passed through a 40-μm Flowmi cell strainer (Bel-Art, H13680-0040).

Single-nucleus suspensions were fixed according to the Nuclei Fixation Kit (SplitBio, SB1003) specifications. An aliquot of each fixed sample was counted in duplicate on a Countess II FL automated cell counter using ethidium homodimer-1 (Fisher E1169) to positively label nuclei. Samples were diluted to 520 nuclei/μL, optimal sample loading concentration for library preparation, and slowly frozen as aliquots using a Mr. Frosty Freezing Container (Thermo Scientific, 5100-0001). All barcoding and library preparation was performed by the UC Davis Genomics Core using the Single Cell Whole Transcriptome Kit (W01030, V1.3.1, Parse Biosciences) with strict adherence to the manufacturer's instructions. This kit uses a combinatory barcoding process to uniquely identify individual nuclei in lieu of specialized cell separation technology. Briefly, cells are distributed across 48 wells to undergo round 1 barcoding using an *in situ* reverse transcription reaction with barcoded primers unique to those wells. Samples are then pooled before undergoing further separation and labeling across round 2 (96 wells) and round 3 (96 wells), where additional barcodes are added through *in situ* ligation reactions. Finally, samples are divided into eight sub-libraries and lysed. This combinatorial barcode approach permits each cell to receive four barcodes, with the potential for 3,538,944 possible barcode combinations.

Libraries were prepared by the UC Davis Genomics Core and sequenced on a NovaSeq S4 300 using 150 bp pair-ended reads (PE150), which targeted a total of 13,149 nuclei for Ce and 8,245 nuclei for vLa across all three rhesus subjects, and 6,973 nuclei for Ce and 9,289 nuclei for vLa across all three human subjects, resulting in a total of 67,312 reads for rhesus and 52,378 reads for human, across Ce and La.

Processing Pipeline for Single-Nucleus Sequencing of the La and Ce Amygdala From Rhesus and Human
Sequenced reads from human and rhesus FASTQ files were aligned to the GRCH38 and RheMac10 genome assemblies,

FIGURE 1. snRNA-seq reveals variation in cell classes and neurons from varied developmental lineages across primate species^a

^a Panel A presents an overview of our tissue processing pipeline used to process amygdala samples from human and rhesus macaque. Data-driven Leiden clustering of cells across regions revealed a number of unique cell clusters in human (panel B) and rhesus monkey (panel F) samples punched from Ce and vLa based on subnuclei localization on corresponding acetylcholinesterase (AChE) stained sections from the amygdala (panel C, human; panel G, rhesus monkey). Cross-region clusters showed expected cell classes as seen in UMAP projections, where each cluster is colored by cell class, and dot plots to show marker gene expression across humans (panel D) and rhesus macaques (panel H). Clusters also revealed excitatory and inhibitory neurons derived from varied developmental lineages (CGE, LGE, MGE, pallium), shown in UMAP projections where each cluster is colored by developmental lineage (human, panel E; rhesus, panel I). "Unknown Origin" is used to indicate that these cells do not express a large proportion of the specific marker genes shown. Glut. = glutamate; Oligo. = oligodendrocyte. (Portions of the figure were created with BioRender.com.)

respectively. Because *Macaca mulatta* did not undergo the same population bottleneck that *Homo sapiens* went through, there is significantly more genetic heterogeneity across the rhesus macaque genome, as compared to humans (30). This

increased genetic heterogeneity has made it difficult to fully annotate the rhesus genome. Publicly available reference genomes are often lacking key gene annotations, which can make accurate cluster identification and cross-species

analysis difficult. To overcome this challenge, a reference genome was built using annotations created by CAT (31). CAT uses the annotation of evolutionary close relatives to the rhesus macaque in order to build a reference that identifies genes that are homologous across species. Human samples were annotated using the latest Ensembl annotation (release 106). Alignment and count matrix generation were performed using the Parse Biosciences splitpipe tool (version 0.9.3). Preprocessing, visualization, and clustering were performed on the generated count matrix using Scanpy (version 1.8.2) (32) and anndata (version 0.7.8). Samples were QC filtered to exclude cells that had a high percentage of mitochondrial and ribosomal genes. The cutoff for mitochondrial gene proportion in this study was 0.4%, and for ribosomal genes, 0.15%. This resulted in one human and one macaque sample being excluded, leaving three human and three macaque samples. Cell counts were normalized by read depth, while excluding highly variable genes (i.e., top 20), and log transformed. Cells and genes were QC filtered to exclude low-expressing cells (i.e., <200 genes), rare genes (i.e., present in <10 cells), and cells that likely reflect more than one cell (i.e., >10,000 genes) (for all functions used in this study, see the GitHub repository at <https://github.com/asfox>).

Cell Class and Developmental Origin Cluster Identification Strategy

In order to perform cluster identification, 100 principal components were identified across the top 6,000 most variable genes for use in clustering analyses. Batch correction was performed using batch-balanced k-nearest neighbors (BBKNN) (version 1.5.1) (33) across 100 principal components. Depending on the analysis, different subsets of cells were clustered. Clustering was performed using Leiden clustering (default resolution of 1) and visualized using UMAP (spread of 2) on the preprocessed and batch-corrected datasets to identify and interpret cell clusters.

In this report, we distinguish between major cell classes (neurons, astrocytes, oligodendrocytes, oligodendrocyte precursor cells [OPCs], microglia, endothelial cells), cell types (i.e., subtypes of cells that we aim to identify within these major cell classes), and cell clusters (our data-derived estimates of cell types). Major cell classes were identified using established marker genes for neurons (*RBFOX1*, *RBFOX3*) (34), microglia (*MYO1F*, *PALDI*) (35), astrocytes (*GFAP*, *ITGB4*) (35), oligodendrocytes (*MBP*) (35, 36), oligodendrocyte progenitor cells (*PDGFRA*, *UST*) (35, 37), and endothelial cells (*FLT1*) (35, 38). Neurons were further subdivided into cell types using marker genes for glutamatergic (*SLC17A6*, *SLC17A7*, *SLC17A8*) and GABAergic (*GAD1*, *GAD2*) cells (35, 39). Clusters were labeled as belonging to a particular class if >60% expressed the marker of interest. Markers for developmental lineages were used to further classify neurons as coming from the excitatory pallidal lineage (*EMC1*, *LMO3*, *AZIN2*, *LHX2*) (17, 19, 40, 41) and inhibitory developmental lineages (ganglionic eminence,

CGE:*PROX1*; *MGE*:*LHX6*, *MAF*; *LGE*:*PAX6*, *MEIS2*, *FOXP2*, *TSHZ1*) (42–47).

Identifying Unique Marker Genes for Cluster Identification

Ideal naming of cell clusters would rely on unique marker genes that could be validated using histopathological methods and could guide reverse translation using cell type-specific functional assays in animal models. Existing tools fall short of this goal by focusing on genes that are most highly expressed in specific cell clusters, as compared to other clusters, which can inhibit translation. Thus, we searched for genes that were unique markers of each cell cluster to guide interpretation and assist in cluster naming. The total number of cells that contained each gene (i.e., present or not) was computed for each Leiden cluster. Unique counts were computed by calculating the difference between the percentage of cells in a cluster that expressed the gene, compared to the maximum percentage of cells expressing that gene in another cluster. For a gene to be categorized as being “unique,” it had to meet two criteria. First, the gene had to be expressed in at least 18% of cells within the cluster. Second, the percentage of cells that expressed the gene had to be approximately threefold greater than in any other cluster, with many genes showing significantly higher in cluster expression. As discussed in the results, many cell clusters did not have any genes that meet our criteria for uniqueness. To name these cell clusters, a gene within the top 0.1 percentile of mean normalized expression for that cluster was used (i.e., based on differential expression with other clusters). Cell clusters failing to express any unique genes are denoted with an asterisk.

Cell Cluster Naming Strategy

Final cell cluster names have the following format: [Region or Species][Leiden Cluster Number][Cell Class][Developmental Origin (if Neuron)][Unique Marker or Overexpressed Gene]. Cells named based on overexpressed genes, rather than a unique marker, are indicated by an asterisk. To facilitate clarity and brevity in text and figures, cell cluster names include abbreviations as follows: Astrocyte, *Ast.*; Neuron, *Neu.*; Oligodendrocyte, *Oligo.*; Microglia, *Micro.*; Unknown, *Unkn.*; Endothelial, *Endo.*; Pallial, *Pal.*

Testing Differences in Cell Type Proportion

To test for significant differences in the proportion of cells in a particular cluster, we used single-cell composition analysis. This analysis was performed in separate analyses across regions and species using the Python package scCODA (version 0.1.7) (33). scCODA allows for the identification of significant composition changes within single-cell sequencing data. This is done using a Bayesian model, which attempts to account for the effect of one cell’s proportion influencing the observed proportion of other cells, thus controlling for false positives. The model structure is based on the Dirichlet-multinomial model. In the model, each cell

cluster is described by a covariate in a log-linear linkage. For the analysis, a default normal prior was used for intercepts. The covariate for each cell cluster was performed via a spike-and-slab prior. scCODA was applied to cross-region (Ce vs. vLa) and cross-species (human vs. rhesus) analyses. Significance was determined using Hamiltonian Monte Carlo sampling to estimate credible effects using false discovery rate (FDR) correction, and log-fold change was reported.

Differential Expression Between Cell Clusters

All differential expression analyses were performed in Python using the Scanpy (version 1.8.2) implementation of a Wilcoxon rank-sum test and FDR correction. The non-parametric Wilcoxon test was used to minimize the influence of outliers. To maximize our power to detect differences that differentiate neuronal cell clusters from each other, this analysis was performed on neuron clusters only. Neuronal clusters were first identified using established marker genes for neurons (*RBFOX1* and *RBFOX3*). Comparisons were performed for each gene between each cluster and all other clusters, in a one-against-all fashion. Significance was determined using FDR correction ($q < 0.05$). Fold change was calculated by Scanpy as $\text{foldchanges} = (\text{self.expm1_func[mean_group]} + 1e-9) / (\text{self.expm1_func[mean_rest]} + 1e-9)$. Any gene that showed a fold change of NA was filled with zero.

Differential Expression Between Inhibitory and Excitatory Neurons Between Regions

To test for differences within GABAergic and glutamatergic neurons between amygdala subnuclei (Ce and La), we performed differential expression analyses using Wilcoxon rank-sum test and FDR correction implemented in Scanpy (version 1.8.2). Inhibitory and excitatory cells were identified as expressing markers for GABAergic (*GAD1*, *GAD2*) and glutamatergic (*SLC17A6*, *SLC17A7*, *SLC17A8*) neurons. Any cells that simultaneously expressed both excitatory and inhibitory markers were excluded. Within each species, we performed differential expression analysis between Ce and La separately for GABAergic and glutamatergic cells to identify transcripts that were significantly (FDR-corrected $q < 0.05$) up- or downregulated in Ce compared to La, that is, higher-expressed in Ce (up) or higher-expressed in La (down). Follow-up analyses were performed using rank-rank hypergeometric tests to examine the overlap in significantly up- and downregulated genes between humans and rhesus monkeys; this produced a 2×2 overlap matrix (monkey: inhibitory/excitatory/down vs. human: inhibitory/excitatory). Adjusted p values for each differential expression analysis were log-scored and signed by the direction of the effect. Because this analysis was being performed across species, genes expressed in only a single species were excluded. In addition, because many genes were not differentially expressed between Ce and vLa, p values equal to 1 were excluded (without this step, the results remain consistent, but there is a large inflection point at $p=1$). The

remaining p values were ranked. We then took 100 steps through the rank for each species and performed a hypergeometric test at each pair of ranks to see if the overlap between gene lists was greater than would be expected by chance. The p values for the hypergeometric test were then log-scaled for visualization. This analysis was performed four times to obtain a 2×2 matrix of rank-rank hypergeometric maps for Ce versus vLa for rhesus versus human and GABAergic versus glutamatergic. The resulting data are visualized such that each point in each rank-rank hypergeometric overlap image is a single hypergeometric test comparing the overlap of the genes above the selected rank threshold for both humans (x-axis) and rhesus monkeys (y-axis).

Co-Clustering Comparisons With HUGO Gene Lists

To test the extent to which marker genes and Leiden clusters could recapitulate other cellular functions, we compared them to clusters created from other functionally relevant genes from the HUGO database (e.g., ion channels, which are relevant to electrophysiological properties). This analysis was performed on rhesus and human cells separately. To do this, we first created subsets of data containing only genes from each cluster, and renormalized and clustered these subsets of data. This produced a set of clusters for each gene set. We then computed co-clustering matrices for each gene set, which resulted in a series of number-of-cells-by-number-of-cells matrices. We then calculated the percentage of cells that co-clustered together compared to the total number of cells that were in both co-clustering matrices for both marker and Leiden gene sets as compared to each of the HUGO gene sets.

MetaNeighbor to Compare Across Species and to Other Datasets

We next sought to test the replicability of relevant findings in relation to previously published snRNA-seq analyses of the rhesus and human amygdala (48, 49). Datasets were downloaded from the Neuroscience Multi-omic Archive (identifier nemo:dat-rtmm5q2) and the Gene Expression Omnibus database (50) (identifier GSE195445). We searched for corresponding cell types in existing datasets using Meta-Neighbor (51). More specifically, we assessed cell cluster replicability using pyNM (<https://github.com/gillislabs/pyMN>; downloaded on March 22, 2024) by pretraining a reference taxonomy based on our cell clusters (using pymn.trainModel) and searching for corresponding clusters in the two rhesus amygdala and one human amygdala reference cell atlases (using pymn.MetaNeighborUS). This analysis takes the correlations between all pairs of cells to build a network where each node is a cell and edges between nodes represent their similarity. To obtain a stringent mapping between our cell clusters and the reference cell types, we used one-versus-best area under the receiver operating characteristic curve (AUROC). AUROC calculates the probability that a classifier correctly predicts that a cell of type X outranks a cell not of type

X within the test dataset based on similarity to the labeled data in the training dataset(s).

Permutation Tests to Identify Disease-Relevant Cell Clusters

A permutation test was performed to identify disease-relevant cell clusters using Python. This test was performed on disease-related gene lists from curated gene lists for consensus agreed-upon genes implicated in ASD (52–54). Because consensus gene lists are not available for anxiety and depressive disorders, we focused on published genome-wide gene-association studies (GWGAs) of schizophrenia (54–57), neuroticism (58), anxiety disorders (59), and depressive disorders (60). We included neuroticism because of the relative risk for anxiety and depressive disorders, as well as the high-quality GWGAS data available (i.e., based on >500k individuals). GWGAs aggregate across multiple SNPs to determine the likelihood of an individual gene being involved. GWGAs reflect a combination of individual SNPs that alter the expression of that transcript, transcript splicing, secondary mRNA structure, probability of methylation, and nonsynonymous variation that impacts the translated protein structure, as well as numerous other ways that individual SNPs may exert their effects. Thus, there is no assumption that individual SNPs are associated with transcript levels, but rather this approach provides a reasonable estimate of what genes might be involved in a particular disorder that does not rely on making an assumption about the mechanism of a particular SNP. Specifically, we examined genes from previously published GWGAs of anxiety disorders (59), neuroticism (58), and depressive disorders (60), as well as current consensus predictions of genes implicated in schizophrenia (54–57) and ASD (56, 61, 62).

All permutation-based analyses were performed in human cells from cross-species clustering of the Ce and vLa. Permutation analyses were performed across human cells to identify cell cluster enrichment for each list of genes in two ways. First, we summed the number of genes expressed in each disease-relevant list for each cell cluster and compared it to 1,000 permutations of randomly selected genes. Because clusters with fewer cells or fewer unique genes can bias permutation analyses, we excluded *Ce 0: Neu_JLS1** and *La 3: Neu_Pal_EGR4**, which had fewer unique genes, and we excluded clusters with <50 cells. Significance was determined based on the proportion of random gene sets with summed reads greater than in the disease-relevant gene set (*enrichment*). Next, a relative enrichment permutation was computed by summing the number of genes from each disease-relevant gene list in each cell type minus the summed reads in other cell types, and comparing it to 1,000 permutations of randomly selected genes. In this second analysis, significance was determined based on the proportion of random gene lists where the *difference* in summed genes was higher than in the disease-relevant gene set (*relative enrichment*). The first analysis provides information about enrichment of disease-relevant genes, which may be shared

between cell clusters, and the second provides a relative enrichment score to identify specific cell clusters that are comparatively enriched relative to other cell clusters.

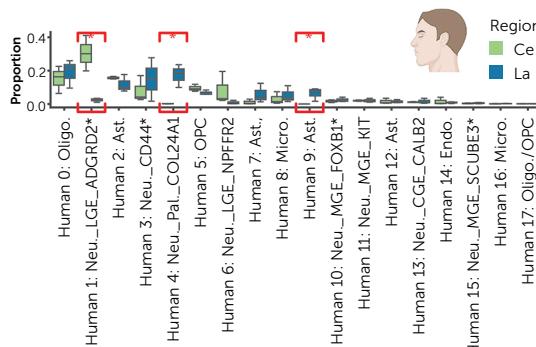
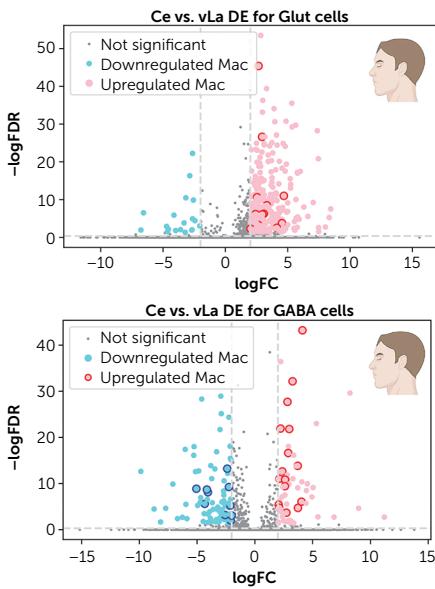
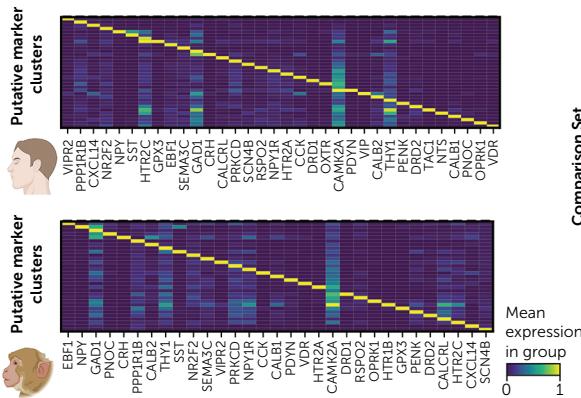
RNAscope in Situ Hybridization to Confirm Coexpression of ITC Markers

Verification of cluster *Ce 3: Neu_LGE_FOXP2* was performed in the same rhesus individuals using single-molecule fluorescence in situ hybridization (smFISH) via RNAscope HiPlex v2 on fixed 14- μ m-thick coronal sections from three control cases, following the manufacturer's instructions. Sections were baked for 30 minutes at 60°C, further fixed in 4% paraformaldehyde for 30 minutes, and then dehydrated through 5-minute increments of graded ethanol (50%, 70%, 100%, 100%). Sections were processed using the RNAscope HiPlex12 Reagent Kit (488, 550, 650) v2 (ACD Bio, 324419), including epitope retrieval and protease III treatment for 15 minutes at 40°C, as directed for fixed tissue. Probes for all six targets—*GAD1*, *NPFFR2*, *TSHZ1*, *GAD2*, *FOXP2*, and *SLC17A7* (see Table S2 in the online supplement)—were hybridized and then amplified en masse. Fluorescent tags (tails) were then applied to the first three probe targets, and sections were counterstained with DAPI and coverslipped with ProLong gold mountant (Invitrogen, P36930). Images were collected using a Zeiss LSM 880 confocal microscope. An initial 5 \times DAPI scan of the entire section was collected (by A.S.F., C.M.S., and K.L.H.) to ensure the accurate anatomical location of Ce. Subsequently, 63 \times scans were collected within the well-defined Ce region for one case, with detailed visual inspection of the other two to confirm results. Tails T1–T3 were removed following application of the manufacturer's cleaving solution, fluorescent tails T4–T6 applied, and then sections were re-coverslipped. New scans for T4–T6 covering all regions captured in round 1 were collected, permitting individual nuclei to be analyzed for all six probe targets. Finally, T4–T6 probes were cleaved, and the slides were reimaged to collect background signals. Not including DAPI, all detectable signals remaining following this final cleavage step were assumed to be endogenous autofluorescence. Image registration and background subtraction across rounds was accomplished using RNAscope HiPlex Image Registration software (ACD Bio, version 2.0.1). Following image alignment, autofluorescence specific to each individual wavelength was subtracted from the probe images. Final image analysis and figure creation were performed using the FISH module in the HALO software program (Indica Labs).

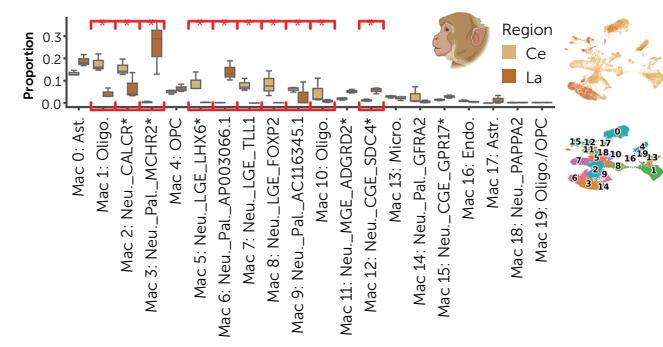
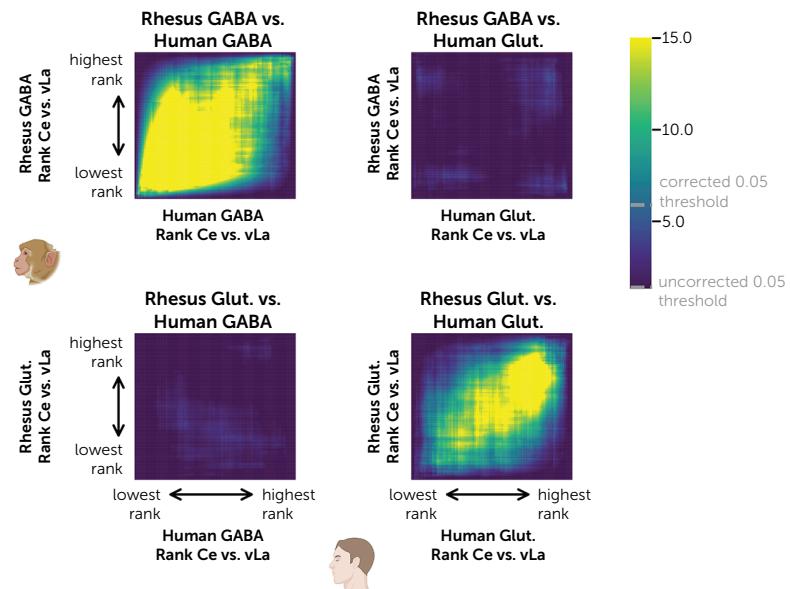
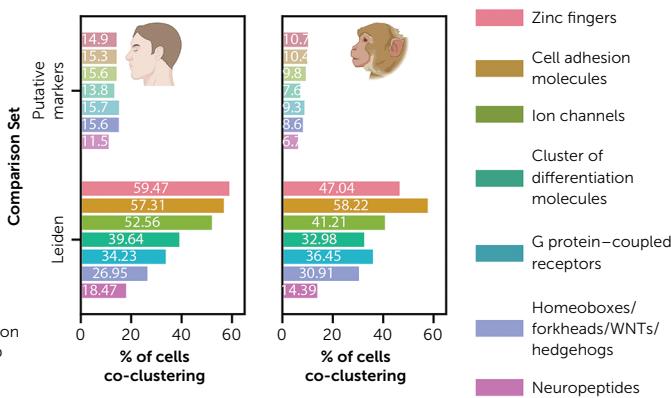
RESULTS

Characterizing Cell Types Within Ce and La

We dissociated cells and sequenced RNA from cells in dorsal and ventral regions of rhesus monkey (N=3) and human (N=3) amygdala (see Figures S1 and S2 in the online supplement). Samples include Ce and vLa, as well as ITCs,

FIGURE 2. Cross-region comparisons of Ce and vLa in humans and rhesus monkeys^a**A. Human: Differences in cell proportions between regions****C. GABAergic and glutamatergic cells express different genes in Ce and vLa****E. Histopathological marker genes are expressed in different cells**

^a Bar plots showing differences between the proportion of cells coming from the Ce and vLa in humans (panel A) and rhesus macaques (panel B) revealed many cells to be more likely to be found in one region than another (highlighted in red with an asterisk; the plot whiskers denote 1.5 times the interquartile range). Volcano plots show differential expression of transcripts between human Ce and vLa glutamatergic neurons (panel C, top) and GABAergic neurons (panel C, bottom). Ce upregulated genes are shown in pink, while Ce downregulated genes, that is, those that are upregulated in vLa, are shown in blue. Genes that were also significant in rhesus macaques are denoted by larger dots outlined in red or blue. In panel D, rank-rank hypergeometric tests reveal cross-species overlap between differentially expressed genes in glutamatergic neurons and GABAergic neurons. Reclustering of proposed marker genes illustrates that our data are consistent with previous histopathological studies showing that these proposed markers are expressed in different cells in both humans (panel E, top) and rhesus macaques (panel E, bottom). However, clustering based on proposed

B. Rhesus: Differences in cell proportions between regions**D. Genes differentially expressed between regions in GABAergic and glutamatergic cells are consistent across species****F. Histopathological marker genes and cell clusters as predictors of variation in the expression of functionally relevant genes**

which are distributed throughout the amygdala. The results identified 19 distinct cell clusters in rhesus and 17 distinct cell clusters in humans, including multiple distinct neuronal cell clusters (12 rhesus and eight human) (Figure 1). Cell clusters were classified into distinct cell classes based on established marker genes (e.g., *GFAP* for astrocytes) (Figure 1D and H). Neuronal cell clusters were further classified using marker genes for excitatory/inhibitory cells and developmental lineage (i.e., pallidal, CGE, MGE, LGE) (Figure 1E and I). These data demonstrate that the amygdala contains multiple cell classes, including neurons of varied developmental origin, but do not represent a comprehensive taxonomy of amygdala cell types.

Differences in Cell Class Composition Between Nuclei

To further explore the differences between regions, we focused on neurons, as they account for the majority of functionally relevant cellular diversity. Because our interest is focused on identifying putative cell types that can be the focus of translational research, we aimed to name cells based on unique marker genes when possible, and on highly expressed genes when no unique genes could be identified (see Tables S3 and S4 in the online supplement). The results demonstrated multiple neuron cell clusters that contained uniquely expressed genes, which included previously identified cell types, including a cluster of Ce neurons that expressed *PENK* (*Ce 9: Neu..LGE_PENK*), a rhesus cluster that expressed *CALCR* or *LHX6* (*Mac 2: Neu..CALCR* and *Mac 2: Neu..LGE_LHX6*), and human neurons that expressed *CALB2* (*Human 13: Neu..CGE_CALB2*). All of these neuronal cell types have been found to be expressed in previous amygdala snRNA-seq studies (39, 49, 52, 63, 64). In particular, a recent cross-species study of whole amygdala by Yu et al. (48) included primate samples (*CALCR* and *PENK*). Notably, in the data presented here, even with relaxed thresholds (i.e., >18% of cells expressing a potential marker), not all cells expressed unique marker genes. Similar results were found when comparing splits in hierarchical cluster dendograms (data not shown).

To characterize differences in cell type composition between Ce and vLa, we tested the differences in the proportion of cells belonging to each cell cluster between subnuclei, while accounting for the number of cells from other cell clusters. The results demonstrated that three human and 10 rhesus cell clusters show differences in cell proportion between subnuclei (Figure 2A,B; see also Table S5 in the online supplement), including both neuronal and nonneuronal clusters in both species. Consistent with the known differences between subnuclei, the largest differences between nuclei were in inhibitory and excitatory neurons, which were enriched in Ce and vLa, respectively.

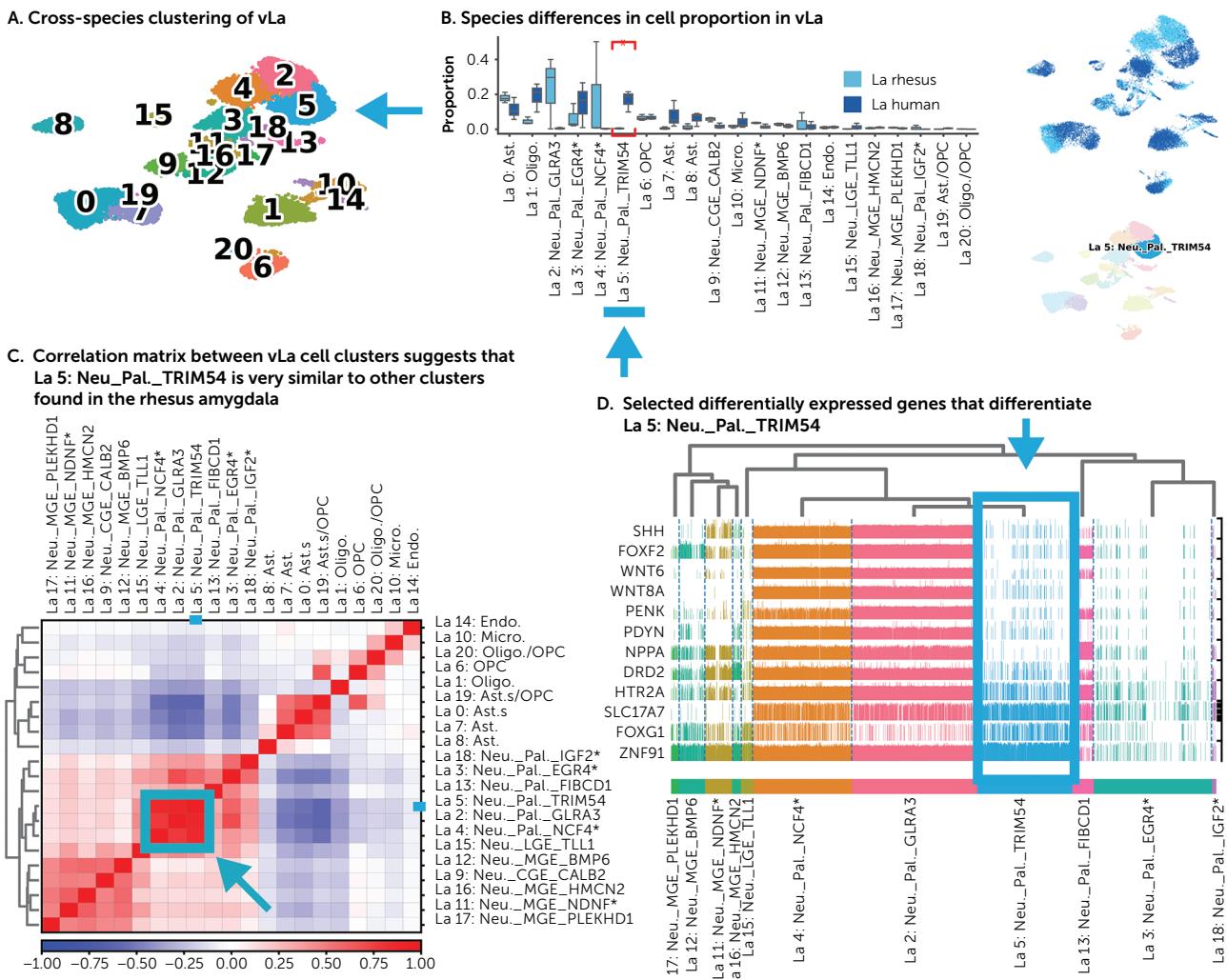
Examining Cross-Region Heterogeneity Within Glutamatergic and GABAergic Neurons

To determine the extent to which this excitatory/inhibitory characterization was sufficient to distinguish between subnuclei, we examined the molecular differences between glutamatergic and GABAergic cells between Ce and vLa across cell clusters. Differential expression analyses were performed within excitatory and inhibitory cells between Ce and vLa, separately for rhesus and human samples. The results demonstrated robust differences in the transcriptomes of excitatory and inhibitory neurons between Ce and vLa in both species (FDR $q < 0.05$; rhesus: 5,863 genes in inhibitory cells and 1,557 genes in excitatory cells; human: 517 genes in inhibitory cells and 707 genes in excitatory cells) (Figure 2C; see also Table S6 in the online supplement). To test the extent to which these differences were reflective of cross-species differences between regions, we performed rank-rank hypergeometric tests at multiple nominal p thresholds. This analysis revealed significant cross-species consistency in the Ce versus vLa differences in excitatory and inhibitory neurons, demonstrating that these differences are robust to species and technical variation (p values < 0.05 , Šidák corrected) (Figure 2D). Post hoc inspection of the data revealed that inhibitory vLa and Ce cells were more likely to express MGE and LGE markers (20), respectively, highlighting heterogeneity of developmental origins as contributing to regional differences.

Relationships Between Data-Derived Cell Clusters and Previously Published Proposed Cell Type Markers

Notably, our data-derived neuronal cell clusters did not clearly support proposed neuronal cell type markers from histopathological studies (*SST*, *CRH*, *PRKCD*, etc.). To determine whether our data were consistent with previous work, we reclustered data based on proposed marker genes for specific cell types in the Ce and La (see Table S7 in the online supplement) (16, 48, 52, 65, 66). The results were consistent with previous work in rodent and in situ studies, demonstrating proposed marker genes to be expressed in different cells (Figure 2E). We next sought to assess the utility of these proposed marker genes as cell type indicators. Because the implicit benefit of a cell type marker is that it can inform us about other aspects of the cell, we examined the percentage of marker- and data-derived clusters that co-clustered with functionally relevant HUGO gene set clusters (e.g., ion channels, which are relevant to electrophysiological properties; see Table S8 in the online supplement). The results demonstrated significantly better correspondence between our data-derived clusters and functionally relevant HUGO gene set clusters, as compared to proposed marker-based clusters (Mann-Whitney U test, $p < 0.001$) (Figure 2F). For example, human amygdala

marker genes, as compared to our data-driven clustering, revealed proposed marker genes to be less likely to co-cluster with other functionally relevant HUGO gene set clusters in both humans (panel F, left) and rhesus macaques (panel F, right). Oligo.=Oligodendrocyte; Ast.=Astrocyte; Neu.=Neuron; Pal.=Pallium; Endo.=Endothelial; Micro.=Microglia. (Portions of the figure were created with BioRender.com.)

FIGURE 3. Cross-species comparisons of amygdala cell clusters^a

^a Cross-species UMAP clustering of vLa (panel A) revealed a number of evolutionarily conserved primate amygdala cell clusters in each region. In panel B, bar plots showing the proportion of cells in each region identify one vLa cell cluster, *La 5: Neu_Pal_TRIM54*, that is more likely to be found in humans compared to rhesus monkeys. In panel C, a correlation matrix shows that transcript expression in the human-enriched vLa neuron cluster, *La 5: Neu_Pal_TRIM54*, is very similar to *La 2: Neu_Pal_GLRA3* and *La 4: Neu_Pal_NCF4**. In panel D, track plots highlight a number of genes that differentiate *La 5: Neu_Pal_TRIM54* from other similar clusters. Blue arrows indicate *La 5: Neu_Pal_TRIM54* across analyses. Ast.=Astrocyte; Oligo.=Oligodendrocyte; Neu.=Neuron; Pal.=Pallium; Micro.=Microglia; Endo.=Endothelial.

cells from HUGO ion-channel cell clusters co-clustered with our data-derived clusters 52.8% of the time, as compared to 16% of the time for marker-based clusters. These results reinforce the limitations of histopathologically defined cell markers.

Similarities and Differences Across Primate Species

To understand the relative distribution of cell clusters in the rhesus and human amygdala, we performed cross-species comparisons. Because Ce and La cells are molecularly and cellularly distinct, we performed separate cross-species clustering for Ce and vLa, and tested for differences in cell cluster proportion between species while accounting for the proportion of other cell clusters (Figure 3A–D; see also Figure S3A,B and Tables S4 and S5 in the online supplement). The results demonstrated that the majority of cell clusters

were not significantly different in the proportion of observed cells between primate species, suggesting that the composition of the amygdala is largely conserved across species (see Supplemental Results in the online supplement for additional detail).

Exploration of *La 5: Neu_Pal_TRIM54* revealed it to be similar to other vLa clusters, *La 2: Neu_Pal_GLRA3* and *La 4: Neu_Pal_NCF4** (Figure 3C), which tended to have more rhesus samples (though there were no significant differences in the concentration of these cell clusters and they were present at low levels in human samples). Interestingly, *La 5: Neu_Pal_TRIM54* was primarily differentiated from *La 2: Neu_Pal_GLRA3* and *La 4: Neu_Pal_NCF4** based on genes that it did not express. This could not be explained by cell quality metrics, such as the number of reads per cell or the average number of genes per cell. Genes largely absent from *La 5:*

Neu_Pal_TRIM5, relative to *La 2: Neu_GLRA3* and *La 5: Neu_Pal_TRIM5*, included genes that encode developmentally relevant molecules, neuropeptides, and neurotransmitter receptors (Figure 3D). Because these differences could also reflect differences in the spatial distribution of cell types across species, we searched for similar cell types in previously published whole-amamygdala analyses from rhesus macaque and human data. The results identified a putatively homologous amygdala cell type in human (cluster 9 from human amygdala in Yu et al. [48]; one-vs-best AUROC: 0.59), but not rhesus, further supporting *La 5: Neu_Pal_TRIM5* as a human-enriched vLa cell cluster that may reflect a developmental deviation from a vLa cell type observed in the rhesus macaque.

Relevance of Specific Cell Clusters to Psychopathology Based on Expression of GWAS-Identified Genes

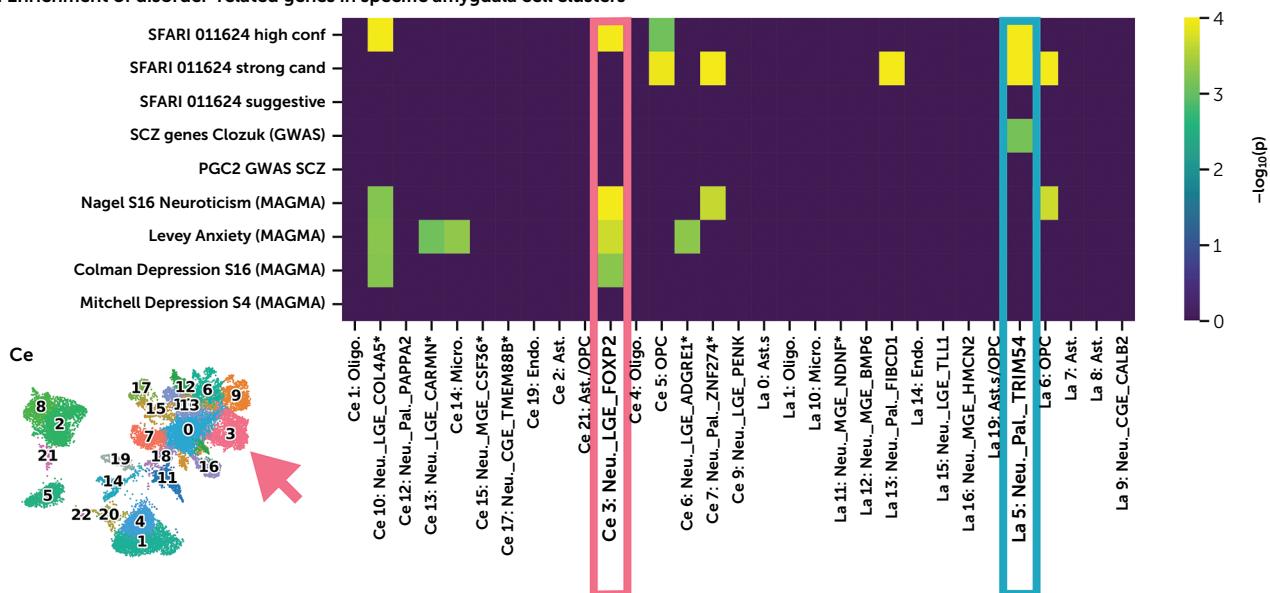
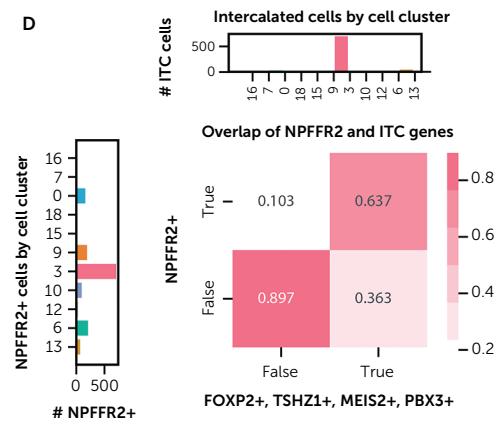
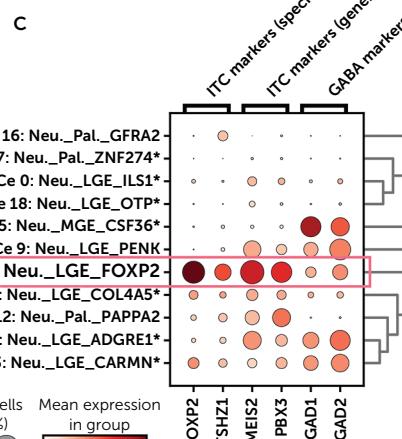
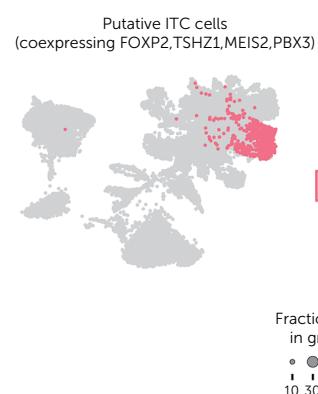
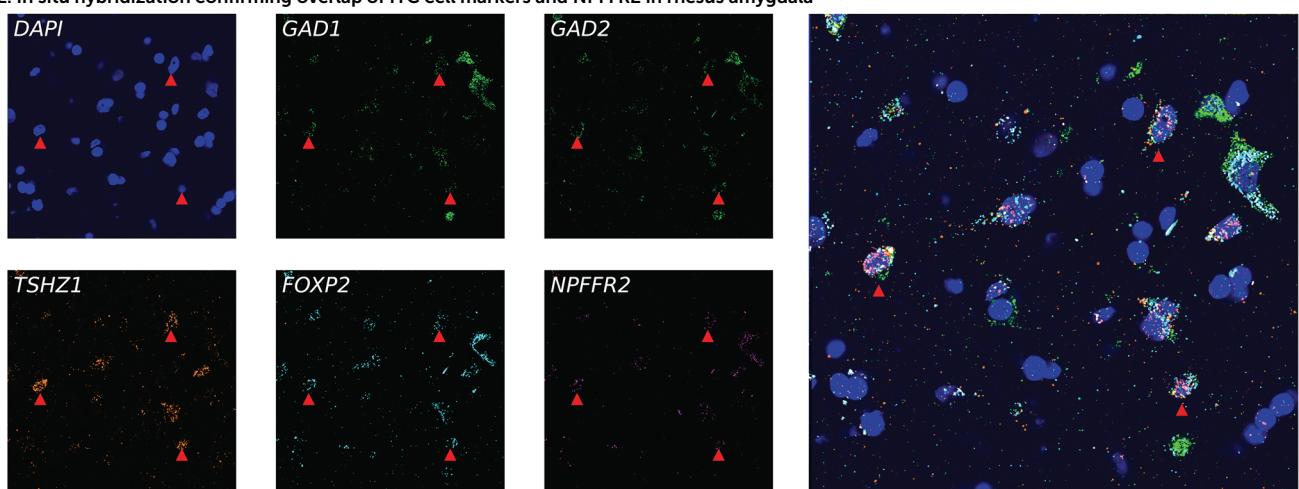
Having identified cross-species cell clusters in different amygdala subnuclei, we sought to determine the potential relevance of these cell clusters to various psychiatric and neurodevelopmental disorders. We performed permutation tests to determine the extent to which transcripts from genes identified in genome-wide association studies (GWASs) were more likely to aggregate in specific cell clusters in the human amygdala. The results demonstrated that multiple cell clusters were enriched for genes across multiple forms of psychopathology and neurodevelopmental disorders (see Figure S4A and Table S9 in the online supplement). Critically, there was relatively low overlap between gene lists, suggesting that cell clusters identified in multiple disorders are not strictly the result of overlapping gene lists (see Figure S4B in the online supplement). Consistent with other studies, neuronal clusters tended to express psychopathology-relevant genes (49), raising the possibility that these results reflect alterations in nonspecific neuronal genes that are expressed in multiple neuronal cell clusters. To identify cell clusters that were uniquely enriched for psychopathology-related genes, we tested the relative enrichment in each cluster compared to other clusters. The results revealed a subset of clusters that are relatively enriched for disease-relevant genes (Figure 4A; see also Table S9 in the online supplement). Significant relative enrichment for ASD-related genes was found for *La 5: Neu_Pal_TRIM54*, the human-enriched vLa cell cluster described above (Figure 3A–D). This suggests that there may be a human-enriched cell type that is particularly vulnerable to the genetic alterations associated with ASD.

In addition, relative enrichment of neuroticism-, anxiety-, and depression-related genes was found for *Ce 3: Neu_LGE_FOXP2*. This cluster expressed high levels of *FOXP2* (Figure 4B), which has been proposed to be a marker of ITCs in rodents (67). Further exploration revealed that markers of ITCs in rodents (*FOXP2* and *TSHZ1*) (67, 68), as well as general markers of the developmental lineage giving rise to rodent ITCs (*MEIS2* and *PBX3*), were highly expressed in *Ce 3: Neu_LGE_FOXP2* (Figure 4B). To confirm coexpression

and demonstrate that these cells could be found in the intercalated region, we performed RNAscope *in situ* hybridization of ITC markers *GAD*, *FOXP2*, and *TSHZ1* along the Ce border (Figure 4E; see also Figure S5 in the online supplement). The results demonstrated that these genes are coexpressed in individual cells in the intercalated region along the Ce border, suggesting that they are indeed ITCs. In addition, we identified a small set of intercalated cells in vLa (data not shown), which also showed this pattern of gene expression, consistent with ITCs being distributed throughout the amygdala. Moreover, comparisons with two extant snRNA-seq datasets from the human and monkey amygdala revealed this cell cluster to correspond to previously identified clusters across primates, including other putatively ITC clusters (see Supplemental Results in the online supplement) (48, 49). Together, these results strongly suggest that *Ce 3: Neu_LGE_FOXP2* consists of cells that are similar to ITCs in rodents. Further exploration of *Ce 3: Neu_LGE_FOXP2* enriched genes revealed a G protein-coupled receptor for neuropeptide FF (*NPFFR2*; log-fold change=22, $p<0.001$; see Table S4 in the online supplement). *NPFFR2* was more likely to be expressed in intercalated cells, compared to other cells (i.e., those that expressed *FOXP2+*, *TSHZ1+*, *MEIS2+*, and *PBX3+*; Fisher's exact test, $p<0.001$) (Figure 4D). Specifically, *NPFFR2* is expressed in 65.5% (419/640) of intercalated cells within *Ce 3: Neu_LGE_FOXP2*. This was broadly consistent across all cells that expressed these intercalated marker genes in Ce, as across all clusters 64% of intercalated cells express *NPFFR2*, as compared to 10% of non-intercalated cells (Figure 4D). RNAscope *in situ* hybridization further confirmed coexpression of *NPFFR2* in ITCs (Figure 4E). Together, these data suggest that ITCs may be particularly relevant to understanding anxiety and depressive disorders, and demonstrate the utility of snRNA-seq for identifying potential treatment targets that should be explored in future research.

DISCUSSION

The amygdala has been implicated as a potential source of pathophysiology in multiple psychiatric and neurodevelopmental disorders, including anxiety disorders (6, 69, 70), depressive disorders (71), ASD (8, 9, 69, 72–74), and psychosis spectrum disorders (75–77). Collectively, these disorders represent a major source of human suffering. Developing treatment targets will require a translational research approach that incorporates a deeper consideration of the specific cells and molecular processes that are impacted in these disorders. Our findings are largely consistent with previous snRNA-seq studies on the primate amygdala and complement these studies by focusing on region-specific cell clusters (previous studies used whole-amamygdala dissections) and analyses that are focused on supporting translational psychiatry. In short, the monkey and human snRNA-seq data presented here underscore the need to consider the amygdala subnuclei as distinct and provide

FIGURE 4. Disorder-relevant enrichment identifies intercalated cells as related to anxiety disorders^a**A. Enrichment of disorder-related genes in specific amygdala cell clusters****B. UMAP location and expression of putative ITC cells enriched in Ce 3: Neu_LGE_FOXP2****E. In situ hybridization confirming overlap of ITC cell markers and NPFFR2 in rhesus amygdala**

valuable insights that can guide translational research into the role of amygdala cell types in psychiatric and neurodevelopmental disorders.

Ce and vLa Are Distinct at the Level of Cells and Molecules

Anatomical studies have identified meaningful differences between nuclei in terms of their composition and connectivity. The data outlined above underscore these findings and reaffirm that different amygdala subnuclei should not be lumped together for the identification of biomarkers and treatment targets. Previous single-cell/nucleus RNA-seq studies have relied on whole-amygdala dissections (48, 63, 64, 78), making the spatial distribution of these cells across amygdala subnuclei unclear and limiting the conclusions that could be drawn. Importantly, although the Ce and vLa are primarily comprised of GABAergic and glutamatergic neurons, respectively, our data further show that there is substantial molecular heterogeneity between the GABAergic/glutamatergic neurons across amygdala subnuclei. These data make it clear that distinctions that rely solely on inhibitory/excitatory neurons are insufficient for characterizing different amygdala subnuclei. Theories of subnucleus-specific function are emerging and will be critical for the next generation of neuroscience-focused psychiatry. In addition, these findings have clear implications for functional neuroimaging studies that aim to identify biomarkers for psychiatric disease. In particular, the cellular and transcriptomic heterogeneity within Ce and vLa suggests that the same genetic and environmental factors are likely to manifest as distinct biomarkers across different amygdala subnuclei. This may be a contributing factor to large-scale efforts to use neuroimaging to identify amygdala activation as a biomarker for psychiatric disease (e.g., 10, 11). Critically, these regions are spatially distinct, and analyses/reanalyses can be used to examine unique biomarkers in each Ce and vLa separately.

Individual Transcript Markers Are Not Unique to Particular Cell Types

In the amygdala, individual marker genes are likely insufficient to understand cellular heterogeneity. Studies in rodents have revealed distinct cells that express particular marker genes that are functionally and, in some cases, structurally distinct (12, 13, 79, 80). Critically, these marker genes appear to cluster into unique cell clusters in rodent

transcriptomic studies. Historically, translational studies have relied on verifying that putative markers of distinct cell types are expressed in different primate cells using histopathological methods. Although the existence of distinct cells expressing distinct marker genes is consistent with our snRNA-seq data (Figure 2E), we did not find previously identified individual marker genes to be uniquely expressed in particular data-derived cell clusters in primates. Rather, consistent with previous work in the primate amygdala, proposed marker genes tended to be expressed across a number of cell clusters (48). Compared to data-derived cell clusters, putative marker genes were less likely to overlap with clusters defined solely based on genes associated with other functionally relevant aspects of neuronal function, for example, ion channel genes (Figure 2F). This illustrative analysis highlights how individual histopathological markers are unlikely to be sufficient to understand cellular heterogeneity in the amygdala. This is particularly important for studies of the amygdala, the cells of which, as compared to cortical cell types, are not obviously differentiated based on location (e.g., layer 2 vs. layer 5/6) or morphology (e.g., projection cells vs. interneurons). Although useful for elucidating amygdala function in rodents, our data question the utility of individual marker genes for translational studies.

Further complicating translational efforts, many cell clusters did not express *unique* marker genes. An ideal marker gene would be expressed in most cells in a particular cell type and minimally expressed in other cell types, to allow for complementary investigation of the same cell populations using *in situ* hybridization, immunohistochemistry, and viral vector-mediated cell type–specific gene manipulation (e.g., using Cre lines or enhancers). Consistent with previous work, these data reinforce the fact that individual marker genes are insufficient to understand cell type heterogeneity (81–83). Increasing the number of samples, cells, and reads and extending these data to include females will provide additional precision and reliability, and is likely to identify additional heterogeneity (80). It remains possible that additional data could identify *unique* marker genes for cell clusters as the parcellation becomes more refined and the transcripts within each cell are more fully sequenced. However, it is becoming increasingly clear that markers based on individual peptides or receptors will be insufficient for guiding translational work and that researchers will need to incorporate multiple markers, drawing heavily on genes

^a In panel A, results of permutation analyses show relative enrichment of disorder-relevant genes in specific amygdala cross-species Ce and vLa cell clusters. To be conservative, only reads from human cells were used to compute relative overexpression. Significant clusters ($p < 0.05$) are colored based on $-\log_{10}(p)$. The results included La 5: Neu...Pal...TRIM54, the human-enriched cluster highlighted in Figure 3, which was relatively enriched for ASD-related genes (blue box). In addition, results included Ce 3: Neu...LGE_FOXP2, which was relatively enriched for genes related to neuroticism, anxiety disorder, and depressive disorder (salmon-colored box and arrow). Ce 3: Neu...LGE_FOXP2 is shown as expression per cell overlaid on a UMAP projection (panel A, bottom left). This cell cluster was enriched for markers of ITCs, shown by the location of putative intercalated cells that coexpress FOXP2, TSHZ1, MEIS2, and PBX3 in the Ce UMAP projection (panel B) and a dot plot (panel C). A 2x2 overlap plot between intercalated cells and NPFFR2 reveals that this G protein receptor is enriched in intercalated cells (panel D; Fisher's exact test, $p < 0.001$), with inset bar plots showing the number of cells in each cluster that express markers for intercalated cells (panel D, top) and NPFFR2 (panel D, left). The presence of cells coexpressing ITC markers was confirmed using RNAscope *in situ* hybridization along the ventral edge of Ce (panel E; see also Figure S5 in the online supplement). Oligo.=Oligodendrocyte; Neu.=Neuron; Pal.=Pallium; Micro.=Microglia; Ast.=Astrocyte; Endo.=Endothelial.

that are involved in development and differentiation, to perform cross-modality translational science (83). Until such markers are identified, researchers utilize multiple markers when possible, and acknowledge that they are likely studying heterogeneous populations.

The Composition of the Ce and vLa Across Primates Is Largely Conserved

Comparisons between the proportion of cells in each species revealed most cell clusters to be consistent across species (39 of 44 cell clusters). Thus, for most cell clusters, the rhesus monkey can provide an excellent model organism for studying cells and molecules that are relevant to the human amygdala.

A notable exception is a cluster of vLa excitatory neurons (*La 5: Neu_Pal_TRIM54*) that was significantly enriched in humans across multiple primate amygdala datasets (see Supplemental Results in the online supplement for additional clusters that differed in their concentration between species). Although this cell cluster had increased expression of numerous genes, including *CRHBP*, *TRIM54*, and the KRAB zinc-finger gene *ZNF91*, which has undergone a series of structural changes in recent primate evolution (84), there were no genes that were uniquely expressed in *La 5: Neu_Pal_TRIM54* compared to similar cell types. The similarity of this cluster to other clusters impedes direct *in situ* validation of these species differences, although the presence of unique genes suggests that this may be possible with a multigene panel. Thus, we cannot rule out the possibility that these cells are present in other species, and/or not truly a distinct cell type, reflecting more nuanced differences, such as those induced by activity-dependent regulation. Nevertheless, these human-specific cell clusters represent a potentially critical insight into the human amygdala, and further work should be done to validate our findings and identify the developmental origin of these cells. Although *La 5: Neu_Pal_TRIM54* retained many similarities to other cell clusters, it was differentiated by a notable sparsity of many transcripts, including the transcript for sonic hedgehog protein (*SHH*), which is integral to the development and differentiation of neuronal cell types. The existence of human-specific cell clusters represents a major challenge for translational research, but by understanding the factors that give rise to these human-specific cells, we have the opportunity to “humanize” animal cells and/or develop *in vitro* models (e.g., induce human pluripotent cells into specific disorder-relevant neuronal cell types).

Specific Cell Types Express Disorder-Related Genes

Ultimately, we anticipate that targeting specific molecules in specific cell types will be critical for optimal treatment and prevention of psychiatric and neurodevelopmental disorders. When examining the enrichment of disease-relevant genes in each cell cluster independently, we found that many neuronal clusters were enriched for genes identified by human GWASs of amygdala-relevant disorders. These data

are consistent with data from other groups (49), and follow from the observation that many GWASs in psychiatric and neurodevelopmental disorders identify numerous genes that are pan-neuronal (e.g., *RBFOX1* in ASD, schizophrenia, depression, and neuroticism). It is possible that these genes represent a disorder-independent cross-cell-type vulnerability factor, but additional mechanistic studies are needed to test this hypothesis.

We identified clusters that were enriched for disease-relevant genes as compared to other clusters, which are most relevant to the development of disorder-specific treatment targets. With respect to ASD-related genes, we identified enrichment within the human-specific cluster *La 5: Neu_Pal_TRIM54* described above. This raises the possibility of a recently evolved cell type that underlies an increased vulnerability to neurodevelopmental disorders. Although more work will be needed to fully understand the implications of this finding, these data suggest a developmental shift in the composition of a particular cell type that provides a unique neurodevelopmental vulnerability for ASD.

Though a potentially powerful tool, examining over-expression of disease-relevant transcripts is limited by our burgeoning understanding of disease-relevant genes. As a result, examinations of disease-relevant gene expression are suggestive, rather than unequivocal. This becomes more important when examining uncurated, data-derived gene lists derived from GWAS analyses. GWAS results contain both false positives and false negatives, and the factors that mediate SNP-phenotype associations remain poorly understood. We attempted to overcome this limitation by focusing on multiple GWGAs for each disorder, each representing a previous best guess for what genes underlie phenotypic differences based on a combination of individual polymorphisms while taking into account linkage disequilibrium. However, these gene lists are incomplete, and the absence of a gene on this list does not mean that a gene does not contribute to the phenotype. Nevertheless, this approach can provide suggestive information that can help prioritize further study of specific cell types to confidently implicate them in psychiatric and/or neurodevelopmental disorders.

With this in mind, we found *Ce 3: Neu_LGE_FOXP2* to be uniquely enriched for genes implicated in neuroticism, anxiety disorders, and depressive disorders. This cell cluster coexpressed genetic markers for ITCs, which are distributed throughout the amygdala (85). We validated coexpression of ITC genes in the intercalated region between Ce and Ba (Figure 4D and E). ITCs receive input from multiple amygdala subnuclei (e.g., Ba and La), as well as frontal and dopaminergic inputs (67, 86). Preclinical rodent work suggests that ITCs are critical for extinction learning, through feed-forward inhibition of Ce and Ba/La (86). The data presented here support preclinical rodent research implicating the ITCs as critical for switching between states of fear/anxiety (86, 87). Although rodent work demonstrates that different ITC clusters can differentially impact threat responses (86), the location of these cell populations in

primates is not well understood, and there are no well-established gene markers to differentiate clusters (although see reference 85). Excitingly, this cell cluster is enriched for expression of the transcript for neuropeptide FF receptor 2 (*NPFFR2*). *NPFFR2* is coexpressed with markers of ITCs in snRNA-seq and RNA-scope data, suggesting that this receptor may be preferentially expressed in the cell lineage that gives rise to ITCs. Initial studies in rodents suggest *NPFFR2* to be relevant for anxiety- and depressive-like behaviors and highlight the therapeutic potential of *NPFFR2* ligands (88–90). Together, these data suggest that ITCs and *NPFFR2* are excellent targets for reverse translation, and may ultimately prove to be relevant treatment targets for anxiety and depressive disorders.

CONCLUSIONS

Psychiatric phenotypes manifest through various pathways, involving multiple brain regions, cells, and molecules. Consequently, there are likely countless mechanisms that could lead to the development of any given clinical phenotype. Technical limitations of preclinical and clinical research tools preclude a careful consideration of cellular and molecular heterogeneity in the human brain. Here, we show how snRNA-seq data can help facilitate cross-species translational psychiatry by elucidating cellular and molecular heterogeneity across the amygdala. Specifically, 1) we demonstrate that broad definitions of region (i.e., “the” amygdala) and/or neuron class (i.e., excitatory vs. inhibitory) are likely to be insufficient for identifying molecules and cell types that can be targeted for optimal treatment development; 2) we reinforce the growing understanding that traditional histopathological markers are likely insufficient to guide cross-species translation; 3) we show how snRNA-seq data can be leveraged in combination with human genetic association studies to refine and prioritize specific cell types for studies in animal disease models; and 4) we show how single-cell transcriptomic data might be used to identify and prioritize receptors that can serve as potential treatment targets. Future research will have to integrate these insights and overcome these challenges with other techniques, including protein expression and cell type-specific manipulations to uncover the functional role of amygdala cell types. Together, these findings offer valuable insights that can help focus basic science targets to better understand human psychopathology and, ultimately, guide the development of new treatments for people who are suffering.

AUTHOR AND ARTICLE INFORMATION

Department of Psychology (Kamboj, Fox), California National Primate Research Center (Kamboj, Bauman, Fox), and MIND Institute (Carlson, Ander, Hanson, Bauman, Schumann), University of California, Davis; Department of Psychiatry and Behavioral Sciences (Carlson, Hanson, Schumann), Department of Neurology (Ander), and Department of Physiology and Membrane Biology (Murray, Bauman), School of Medicine, University of California, Davis; Department of Neuroscience and

Department of Psychiatry, School of Medicine and Dentistry, University of Rochester, Rochester, NY (Fudge).

Send correspondence to Dr. Fox (dfox@ucdavis.edu) and Dr. Schumann (cschumann@ucdavis.edu).

Development of the animal model was supported by a grant from the Simons Foundation to the late Dr. Paul Patterson (SFARI 9900060). This research was supported in part by NIMH grant R21MH119650 (to Drs. Schumann and Bauman), the UC Davis Conte Center (NIMH grant P50MH106438-6618), the California National Primate Research Center (NIH grant P51-OD011107, to Dr. Fox as core scientist), and the UC Davis MIND Institute Intellectual and Developmental Disabilities Research Center (National Institute of Child Health and Human Development grant P50HD103526).

The authors thank everyone in the Fox and Schumann Labs, the staff at the UC Davis High Performance Computing Core, and the staff at the California National Primate Research Center, in particular Amanda Chu. The authors also thank Alex Pollen and Matthew Schmitz for providing guidance and support in mapping cross-species RNA-seq data, and Avni Shah, Aaron Davis, and Alicja Omanska for technical assistance in tissue processing.

The authors report no financial relationships with commercial interests.

All code for processing and analyses is available on GitHub, at <https://github.com/asfox>. The data for this report have been made publicly available using the Gene Expression Omnibus database with the following id: GSE262905.

Received July 31, 2023; revision received April 16, 2024; accepted May 29, 2024.

REFERENCES

1. Baron-Cohen S, Ring HA, Bullmore ET, et al: The amygdala theory of autism. *Neurosci Biobehav Rev* 2000; 24:355–364
2. Aleman A, Kahn RS: Strange feelings: do amygdala abnormalities dysregulate the emotional brain in schizophrenia? *Prog Neurobiol* 2005; 77:283–298
3. Amaral DG, Schumann CM, Nordahl CW: Neuroanatomy of autism. *Trends Neurosci* 2008; 31:137–145
4. Schumann CM, Bauman MD, Amaral DG: Abnormal structure or function of the amygdala is a common component of neurodevelopmental disorders. *Neuropsychologia* 2011; 49:745–759
5. Fox AS, Kalin NH: A translational neuroscience approach to understanding the development of social anxiety disorder and its pathophysiology. *Am J Psychiatry* 2014; 171:1162–1173
6. Craske MG, Stein MB, Eley TC, et al: Anxiety disorders. *Nat Rev Dis Primers* 2017; 3:17024
7. Bobilev AM, Perez JM, Tamminga CA: Molecular alterations in the medial temporal lobe in schizophrenia. *Schizophr Res* 2020; 217: 71–85
8. Avino TA, Barger N, Vargas MV, et al: Neuron numbers increase in the human amygdala from birth to adulthood, but not in autism. *Proc Natl Acad Sci U S A* 2018; 115:3710–3715
9. Weir RK, Bauman MD, Jacobs B, et al: Protracted dendritic growth in the typically developing human amygdala and increased spine density in young ASD brains. *J Comp Neurol* 2018; 526:262–274
10. Grogan SE, Fox AS, Shackman AJ: The amygdala and depression: a sober reconsideration (editorial). *Am J Psychiatry* 2022; 179: 454–457
11. Tamm S, Harmer CJ, Schiel J, et al: No association between amygdala responses to negative faces and depressive symptoms: cross-sectional data from 28,638 individuals in the UK Biobank cohort. *Am J Psychiatry* 2022; 179:509–513
12. Janak PH, Tye KM: From circuits to behaviour in the amygdala. *Nature* 2015; 517:284–292
13. Fadok JP, Markovic M, Toyote P, et al: New perspectives on central amygdala function. *Curr Opin Neurobiol* 2018; 49:141–147

14. Johnston JB: Further contributions to the study of the evolution of the forebrain. *J Comp Neurol* 1923; 35:337–481
15. Swanson LW, Petrovich GD: What is the amygdala? *Trends Neurosci* 1998; 21:323–331
16. Beyeler A, Dabrowska J: Neuronal diversity of the amygdala and the bed nucleus of the stria terminalis, in *Handbook of Behavioral Neuroscience*. Edited by Urban JH , Rosenkranz JA, London, Elsevier, 2020, pp 63–100. <https://www.sciencedirect.com/science/article/pii/B9780128151341000039>
17. Garcia-Calero E, Martínez-de-la-Torre M, Puelles L: A radial histogenetic model of the mouse pallial amygdala. *Brain Struct Funct* 2020; 225:1921–1956
18. Aerts T, Seuntjens E: Novel perspectives on the development of the amygdala in rodents. *Front Neuroanat* 2021; 15:786679
19. Medina L, Abellán A, Desfilis E: Evolving views on the pallium. *Brain Behav Evol* 2022; 96:181–199
20. Medina L, Abellán A, Morales L, et al: Evolution and development of amygdala subdivisions: pallial, subpallial, and beyond. *Brain Behav Evol* 2023; 98:1–21
21. Namburi P, Beyeler A, Yorozu S, et al: A circuit mechanism for differentiating positive and negative associations. *Nature* 2015; 520: 675–678
22. Barger N, Stefanacci L, Schumann CM, et al: Neuronal populations in the basolateral nuclei of the amygdala are differentially increased in humans compared with apes: a stereological study. *J Comp Neurol* 2012; 520:3035–3054
23. Schumann CM, Scott JA, Lee A, et al: Amygdala growth from youth to adulthood in the macaque monkey. *J Comp Neurol* 2019; 527: 3034–3045
24. Lew CH, Hanson KL, Groeniger KM, et al: Serotonergic innervation of the human amygdala and evolutionary implications. *Am J Phys Anthropol* 2019; 170:351–360
25. Perelman P, Johnson WE, Roos C, et al: A molecular phylogeny of living primates. *PLoS Genet* 2011; 7:e1001342
26. Kalin NH, Shelton SE: Nonhuman primate models to study anxiety, emotion regulation, and psychopathology. *Ann N Y Acad Sci* 2003; 1008:189–200
27. Phillips KA, Bales KL, Capitanio JP, et al: Why primate models matter. *Am J Primatol* 2014; 76:801–827
28. Schumann CM, Amaral DG: Stereological estimation of the number of neurons in the human amygdaloid complex. *J Comp Neurol* 2005; 491:320–329
29. Page NF, Gandal MJ, Estes ML, et al: Alterations in retrotransposition, synaptic connectivity, and myelination implicated by transcriptomic changes following maternal immune activation in nonhuman primates. *Biol Psychiatry* 2021; 89:896–910
30. Warren WC, Harris RA, Haukness M, et al: Sequence diversity analyses of an improved rhesus macaque genome enhance its biomedical utility. *Science* 2020; 370:eabc6617
31. Fiddes IT, Armstrong J, Diekhans M, et al: Comparative Annotation Toolkit (CAT): simultaneous clade and personal genome annotation. *Genome Res* 2018; 28:1029–1038
32. Wolf FA, Angerer P, Theis FJ: SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol* 2018; 19:15
33. Büttner M, Ostner J, Müller CL, et al: scCODA is a Bayesian model for compositional single-cell data analysis. *Nat Commun* 2021; 12: 6876
34. Darnell RB: RNA protein interaction in neurons. *Annu Rev Neurosci* 2013; 36:243–270
35. Karlsson M, Zhang C, Méar L, et al: A single-cell type transcriptomics map of human tissues. *Sci Adv* 2021; 7:eaabh2169
36. Zecević N, Andjelković A, Matthieu JM, et al: Myelin basic protein immunoreactivity in the human embryonic CNS. *Brain Res Dev Brain Res* 1998; 105:97–108
37. Beiter RM, Rivet-Noor C, Merchak AR, et al: Evidence for oligodendrocyte progenitor cell heterogeneity in the adult mouse brain. *Sci Rep* 2022; 12:12921
38. Sawano A, Iwai S, Sakurai Y, et al: Flt-1, vascular endothelial growth factor receptor 1, is a novel cell surface marker for the lineage of monocyte-macrophages in humans. *Blood* 2001; 97:785–791
39. Hochgerner H, Singh S, Tibi M, et al: Neuronal types in the mouse amygdala and their transcriptional response to fear conditioning. *Nat Neurosci* 2023; 26:2237–2249
40. Cucas LA, Miyoshi G, Carney RS, et al: Emx1-lineage progenitors differentially contribute to neural diversity in the striatum and amygdala. *J Neurosci* 2009; 29:15933–15946
41. Abellán A, Desfilis E, Medina L: Combinatorial expression of Lef1, Lhx2, Lhx5, Lhx9, Lmo3, Lmo4, and Prox1 helps to identify comparable subdivisions in the developing hippocampal formation of mouse and chicken. *Front Neuroanat* 2014; 8:59
42. Miyoshi G, Young A, Petros T, et al: Prox1 regulates the subtype-specific development of caudal ganglionic eminence-derived GABAergic cortical interneurons. *J Neurosci* 2015; 35:12869–12889
43. Asgarian Z, Oliveira MG, Stryjewska A, et al: MTG8 interacts with LHX6 to specify cortical interneuron subtype identity. *Nat Commun* 2022; 13:5217
44. Alifragis P, Liapi A, Parnavelas JG: Lhx6 regulates the migration of cortical interneurons from the ventral telencephalon but does not specify their GABA phenotype. *J Neurosci* 2004; 24:5643–5648
45. Pai EL, Vogt D, Clemente-Perez A, et al: Mafb and c-Maf have prenatal compensatory and postnatal antagonistic roles in cortical interneuron fate and function. *Cell Rep* 2019; 26:1157–1173.e5
46. Kroll TT, O’Leary DD: Ventralized dorsal telencephalic progenitors in Pax6 mutant mice generate GABA interneurons of a lateral ganglionic eminence fate. *Proc Natl Acad Sci U S A* 2005; 102: 7374–7379
47. Schmitz MT, Sandoval K, Chen CP, et al: The development and evolution of inhibitory neurons in primate cerebrum. *Nature* 2022; 603:871–877
48. Yu B, Zhang Q, Lin L, et al: Molecular and cellular evolution of the amygdala across species analyzed by single-nucleus transcriptome profiling. *Cell Discov* 2023; 9:19
49. Chiou KL, Huang X, Bohnen MO, et al: A single-cell multi-omic atlas spanning the adult rhesus macaque brain. *Science* 2023; 9:eadh1914
50. Edgar R, Domrachev M, Lash AE: Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* 2002; 30:207–210
51. Crow M, Paul A, Ballouz S, et al: Characterizing the replicability of cell types defined by single cell RNA-sequencing data using MetaNeighbor. *Nat Commun* 2018; 9:884
52. Wang Y, Krabbe S, Eddison M, et al: Multimodal mapping of cell types and projections in the central nucleus of the amygdala. *ELife* 2023; 12:e84262
53. Mariani J, Coppola G, Zhang P, et al: FOXG1-dependent dysregulation of GABA/glutamate neuron differentiation in autism spectrum disorders. *Cell* 2015; 162:375–390
54. Schizophrenia Working Group of the Psychiatric Genomics Consortium: Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 2014; 511:421–427
55. Farrell MS, Werge T, Sklar P, et al: Evaluating historical candidate genes for schizophrenia. *Mol Psychiatry* 2015; 20:555–562
56. Gandal MJ, Zhang P, Hadjimichael E, et al: Transcriptome-wide isoform-level dysregulation in ASD, schizophrenia, and bipolar disorder. *Science* 2018; 362:eaat8127
57. Pardiñas AF, Holmans P, Pocklington AJ, et al: Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nat Genet* 2018; 50: 381–389
58. Nagel M, Jansen PR, Stringer S, et al: Meta-analysis of genome-wide association studies for neuroticism in 449,484 individuals identifies novel genetic loci and pathways. *Nat Genet* 2018; 50:920–927
59. Levey DF, Gelernter J, Polimanti R, et al: Reproducible genetic risk loci for anxiety: results from ~200,000 participants in the Million Veteran Program. *Am J Psychiatry* 2020; 177:223–233

60. Coleman JRI, Gaspar HA, Bryois J, et al: The genetics of the mood disorder spectrum: genome-wide association analyses of more than 185,000 cases and 439,000 controls. *Biol Psychiatry* 2020; 88:169–184
61. Sorrells SF, Paredes MF, Velmeshov D, et al: Immature excitatory neurons develop during adolescence in the human amygdala. *Nat Commun* 2019; 10:2748
62. Zhang P, Omanska A, Ander BP, et al: Neuron-specific transcriptomic signatures indicate neuroinflammation and altered neuronal activity in ASD temporal cortex. *Proc Natl Acad Sci U S A* 2023; 120:e2206758120
63. Tran MN, Maynard KR, Spangler A, et al: Single-nucleus transcriptome analysis reveals cell-type-specific molecular signatures across reward circuitry in the human brain. *Neuron* 2021; 109: 3088–3103.e5
64. Siletti K, Hodge R, Mossi Albiach A, et al: Transcriptomic diversity of cell types across the adult human brain. *Science* 2023; 382:eadd7046
65. Dilly GA, Kittleman CW, Kerr TM, et al: Cell-type specific changes in PKC-delta neurons of the central amygdala during alcohol withdrawal. *Transl Psychiatry* 2022; 12:289
66. O'Leary TP, Kendrick RM, Bristow BN, et al: Neuronal cell types, projections, and spatial organization of the central amygdala. *iScience* 2022; 25:105497
67. Aksoy-Aksel A, Gall A, Seewald A, et al: Midbrain dopaminergic inputs gate amygdala intercalated cell clusters by distinct and cooperative mechanisms in male mice. *ELife* 2021; 10:e63708
68. Kuerbitz J, Arnett M, Ehrman S, et al: Loss of intercalated cells (ITCs) in the mouse amygdala of Tshz1 mutants correlates with fear, depression, and social interaction phenotypes. *J Neurosci* 2018; 38: 1160–1177
69. MacMillan S, Szczesko PR, Moore GJ, et al: Increased amygdala: hippocampal volume ratios associated with severity of anxiety in pediatric major depression. *J Child Adolesc Psychopharmacol* 2003; 13:65–73
70. Roy AK, Fudge JL, Kelly C, et al: Intrinsic functional connectivity of amygdala-based networks in adolescent generalized anxiety disorder. *J Am Acad Child Adolesc Psychiatry* 2013; 52:290–299.e2
71. Bielau H, Trübner K, Krell D, et al: Volume deficits of subcortical nuclei in mood disorders: a postmortem study. *Eur Arch Psychiatry Clin Neurosci* 2005; 255:401–412
72. Morgan JT, Barger N, Amaral DG, et al: Stereological study of amygdala glial populations in adolescents and adults with autism spectrum disorder. *PLoS ONE* 2014; 9:e110356
73. Andrews DS, Aksman L, Kerns CM, et al: Association of amygdala development with different forms of anxiety in autism spectrum disorder. *Biol Psychiatry* 2022; 91:977–987
74. Herrington JD, Maddox BB, Kerns CM, et al: Amygdala volume differences in autism spectrum disorder are related to anxiety. *J Autism Dev Disord* 2017; 47:3682–3691
75. Kreczmanski P, Heinsen H, Mantua V, et al: Volume, neuron density and total neuron number in five subcortical regions in schizophrenia. *Brain* 2007; 130:678–692
76. Williams MR, Pattini S, Pearce RK, et al: Basolateral but not corticomedial amygdala shows neuroarchitectural changes in schizophrenia. *J Neurosci Res* 2016; 94:544–547
77. Pantazopoulos H, Wiseman JT, Markota M, et al: Decreased numbers of somatostatin-expressing neurons in the amygdala of subjects with bipolar disorder or schizophrenia: relationship to circadian rhythms. *Biol Psychiatry* 2017; 81:536–547
78. Zhang L, Cheng Y, Wu S, et al: Molecular taxonomy of the primate amygdala via single-nucleus RNA sequencing analysis. *Sci Bull (Beijing)* 2021; 66:1379–1383
79. Haubensak W, Kunwar PS, Cai H, et al: Genetic dissection of an amygdala microcircuit that gates conditioned fear. *Nature* 2010; 468:270–276
80. BRAIN Initiative Cell Census Network (BICCN): A multimodal cell census and atlas of the mammalian primary motor cortex. *Nature* 2021; 598:86–102
81. Fischer S, Gillis J: How many markers are needed to robustly determine a cell's type? *iScience* 2021; 24:103292
82. Dumitrescu B, Villar S, Mixon DG, et al: Optimal marker gene selection for cell type discrimination in single cell analyses. *Nat Commun* 2021; 12:1186
83. Krienen FM, Levandowski KM, Zaniewski H, et al: A marmoset brain cell census reveals persistent influence of developmental origin on neurons. *BioRxiv*, October 19, 2022. <https://www.biorxiv.org/content/10.1101/2022.10.18.512442v1>
84. Jacobs FM, Greenberg D, Nguyen N, et al: An evolutionary arms race between KRAB zinc-finger genes ZNF91/93 and SVA/L1 retrotransposons. *Nature* 2014; 516:242–245
85. Zikopoulos B, John YJ, García-Cabezas MÁ, et al: The intercalated nuclear complex of the primate amygdala. *Neuroscience* 2016; 330: 267–290
86. Asede D, Doddapaneni D, Bolton MM: Amygdala intercalated cells: gate keepers and conveyors of internal state to the circuits of emotion. *J Neurosci* 2022; 42:9098–9109
87. Hagihara KM, Bukalo O, Zeller M, et al: Intercalated amygdala clusters orchestrate a switch in fear state. *Nature* 2021; 594: 403–407
88. Lin YT, Huang YL, Tsai SC, et al: Ablation of NPFFR2 in mice reduces response to single prolonged stress model. *Cells* 2020; 9: 2479
89. Yu Z, Lin YT, Chen JC: Knockout of NPFFR2 prevents LPS-induced depressive-like responses in mice. *Int J Mol Sci* 2021; 22:7611
90. Nguyen T, Marusich J, Li JX, et al: Neuropeptide FF and its receptors: therapeutic applications and ligand development. *J Med Chem* 2020; 63:12387–12402

Data supplement for Kamboj et al., Translational Insights From Cell Type Variation Across Amygdala Subnuclei in Rhesus Monkeys and Humans. Am J Psychiatry
 (10.1176/appi.ajp.20230602)

TABLE S1. Individual Study Subjects

Case ID	Sex	Age	Cause of Death	PMI (hrs)	Brain Weight (g)	Number of Nuclei Isolated	
						Hemisphere	
Human							
BEMH0048	M	15	Drowning	31	1330	Right	34,400 (LAT), 220,500 (Ce)
BEMH0033	M	17	Suicide	32	1465	Right	81,200 (LAT), 43,700 (Ce)
BEMH0037	M	19	Suicide	23	1610	Left	79,500 (LAT), 42,500 (Ce)
Macaca mulatta							
40295	M	3.5-4				Right	78,900 (LAT), 93,100 (Ce)
40335	M	3.5-4				Right	70,100 (LAT), 89,100 (Ce)
40478	M	3.5-4				Right	67,700 (LAT), 72,400 (Ce)

TABLE S2. RNAscope HiPlex Probes

Round	Probe	Cat. No (ACD Bio)	Wavelength
Round 1	Mmu-GAD1-T1	501811-T1	488nm
	Mmu-NPFFR2-T2	1286711-T2	550nm
	Mmu-TSHZ1-T3	801141-T3	650nm
Round 2	Mmu-GAD2-T4	461701-T4	488nm
	Mmu-FOXP2-T5	1107421-T5	550nm
	Mmu-SLC17A7-T6	492791-T6	650nm

FIGURE S1. Cross region scatter and violin plots for key quality control metrics. Scatter plot axes are total counts against number of genes by count, where color represents mitochondrial percentage. Violin plots draw attention to the number of genes by counts, total counts, percent mitochondrial expression and percent ribosomal expression for each sample.

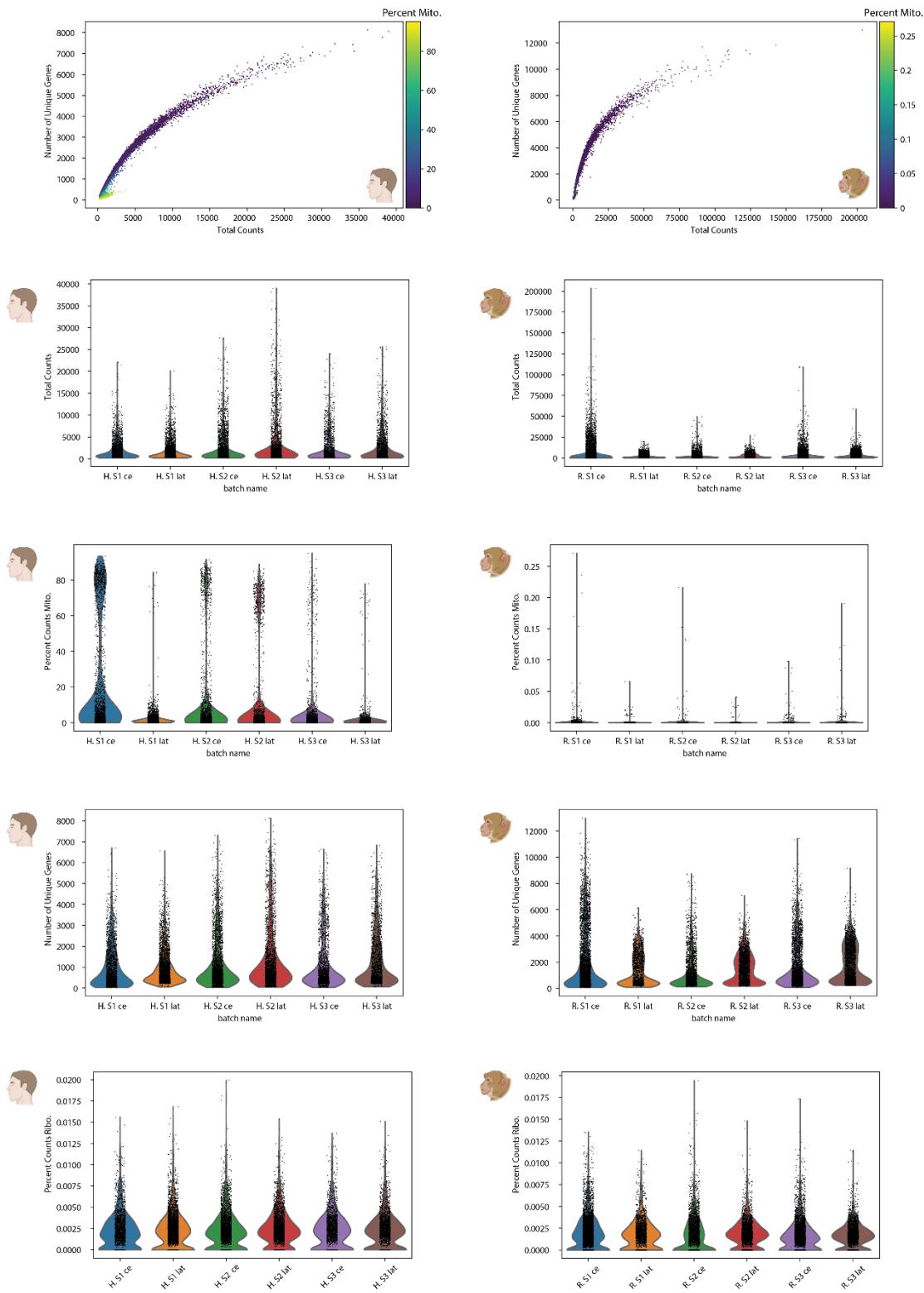


FIGURE S2. Cross-region UMAP images from Figure 1 with each cell colored by subject for human (left) and rhesus monkey (right).

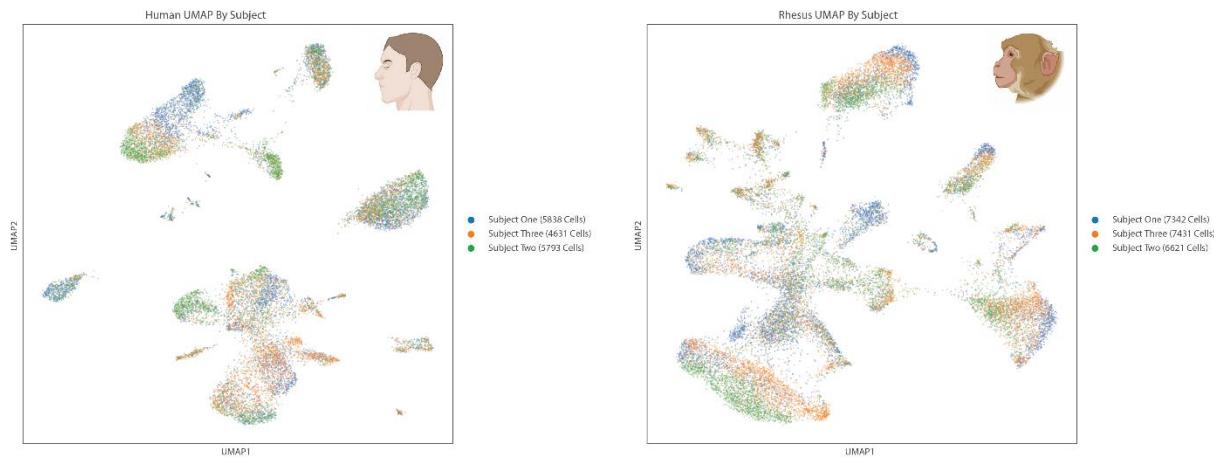


Table S3 and Table S4 are in separate supplemental Excel files.

TABLE S5. Differences in Cell Type Proportion

Rhesus	Human		Central amygdala		Lateral amygdala	
Cell Type	Effect (Log ₂)	Cell Type	Effect (Log ₂)	Cell Type	Effect (Log ₂)	Cell Type
Mac 0: Ast.	0.385061	Human 0: Oligo.	0.13439	Ce 0: Neu._ILS1*	-0.27732	La 0: Ast.
Mac 1: Oligo.	-1.93724	Human 1: Neu._LGE_ADGRD2*	-3.0689	Ce 1: Oligo.	2.537023	La 1: Oligo.
Mac 2: Neu._CALCR*	-1.36506	Human 2: Ast.	0.13439	Ce 2: Ast.	-0.27732	La 2: Neu._Pal._GLRA3
Mac 3: Neu._Pal._MCHR2*	4.971202	Human 3: Neu._CD44*	0.13439	Ce 3: Neu._LGE_FOXP2	-0.27732	La 3: Neu._Pal._EGR4*
Mac 4: OPC	0.385061	Human 4: Neu._Pal._COL24A1	4.030008	Ce 4: Oligo.	-0.27732	La 4: Neu._Pal._NCF4*
Mac 5: Neu._LGE_LHX6*	-3.32869	Human 5: OPC	0.13439	Ce 5: OPC	-0.27732	La 5: Neu._Pal._TRIM54
Mac 6: Neu._Pal._AP003066.1	4.595821	Human 6: Neu._LGE_NPFFR2	0.13439	Ce 6: Neu._LGE_ADGRE1*	-0.27732	La 6: OPC
Mac 7: Neu._LGE_TLL1	-4.41031	Human 7: Ast.	0.13439	Ce 7: Neu._Pal._ZNF274*	-0.27732	La 7: Ast.
Mac 8: Neu._LGE_FOXP2	-3.54756	Human 8: Micro.	0.13439	Ce 8: Ast.	2.165189	La 8: Ast.
Mac 9: Neu._Pal._AC116345.1	-2.39786	Human 9: Ast.	3.07756	Ce 9: Neu._LGE_PENK	1.876308	La 9: Neu._CGE_CALB2
Mac 10: Oligo.	-1.35173	Human 10: Neu._MGE_FOXB1*	0.13439	Ce 10: Neu._LGE_COL4A5*	-0.27732	La 10: Micro.
Mac 11: Neu._MGE_ADGRD2*	0.385061	Human 11: Neu._MGE_KIT	0.13439	Ce 11: Oligo.	1.928809	La 11: Neu._MGE_NDNF*
Mac 12: Neu._CGE_SDC4*	1.917378	Human 12: Ast.	0.13439	Ce 12: Neu._Pal._PAPPA2	-0.27732	La 12: Neu._MGE_BMP6
Mac 13: Micro.	0.385061	Human 13: Neu._CGE_CALB2	0.13439	Ce 13: Neu._LGE_CARMN*	-0.27732	La 13: Neu._Pal._FIBCD1
Mac 14: Neu._Pal._GFRA2	0.385061	Human 14: Endo.	0.13439	Ce 14: Micro.	-0.27732	La 14: Endo.
Mac 15: Neu._CGE_GPR17*	0.385061	Human 15: Neu._MGE_SCUBE3*	0.13439	Ce 15: Neu._MGE_CSF36*	-0.27732	La 15: Neu._LGE_TLL1
Mac 16: Endo.	0.385061	Human 16: Micro.	0.13439	Ce 16: Neu._Pal._GFRA2	-0.27732	La 16: Neu._MGE_HMCN2
Mac 17: Astr.	0.385061	Human 17: Oligo./OPC	0.13439	Ce 17: Neu._CGE_TMEM88B*	-0.27732	La 17: Neu._MGE_PLEKHD1
Mac 18: Neu._PAPPA2	0.385061			Ce 18: Neu._OTP*	-0.27732	La 18: Neu._Pal._IGF2*
Mac 19: Oligo./OPC	0.385061			Ce 19: Endo.	-0.27732	La 19: Ast./OPC
				Ce 20: Micro./Oligo.	-0.27732	La 20: Oligo./OPC
				Ce 21: Ast./OPC	-0.27732	
				Ce 22: Oligo./OPC	-0.27732	

Table S6 is in a separate supplemental Excel file.

TABLE S7. Proposed Marker Genes for Specific Cell Types

Marker Genes (non-exhaustive references)

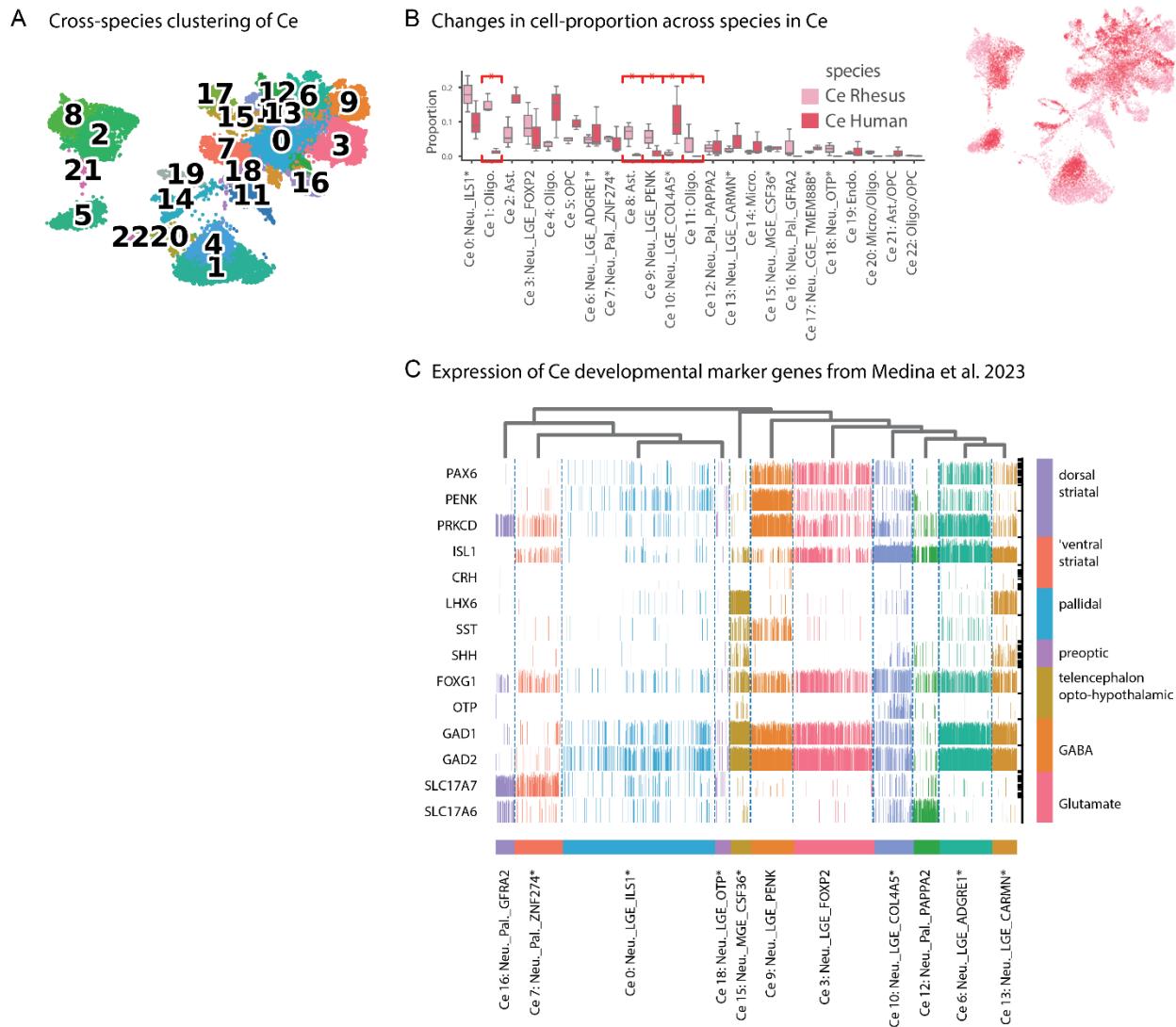
VIPR2 (1)
PPP1R1B (2)
CXCL14 (3)
NR2F2 (3)
NPY (4)
SST (3)
HTR2C (1)
GPX3 (1)
EBF1 (1)
SEMA3C (1)
GAD1 (3)
CRH (3)
CALCRL (1)
PRKCD (3)
RSP02 (5)
HTR2A (1)
CCK (3)
DRD1 (1)
OXTR (3)
CAMK2A (2)
PDYN (1)
VIPR2 (1)
CALB2 (5)
PENK (1)
DRD2 (3)
TAC1 (1)
NTS (3)
CALB1 (2)
PNOC (1)
OPRK1 (1)
VDR (1)

1. Wang Y, Krabbe S, Eddison M, et al.: Multimodal mapping of cell types and projections in the central nucleus of the amygdala. *eLife* 2023; 12:e84262
2. O'Leary TP, Sullivan KE, Wang L, et al.: Extensive and spatially variable within-cell-type heterogeneity across the basolateral amygdala. *eLife* 2020; 9:e59003
3. O'Leary TP, Kendrick RM, Bristow BN, et al.: Neuronal cell types, projections, and spatial organization of the central amygdala. *iScience* 2022; 25:105497
4. Beyeler A, Dabrowska J: Neuronal diversity of the amygdala and the bed nucleus of the stria terminalis. *Handb Behav Neurosci* 2020; 26:63–100
5. Dilly GA, Kittleman CW, Kerr TM, et al.: Cell-type specific changes in PKC-delta neurons of the central amygdala during alcohol withdrawal. *Transl Psychiatry* 2022; 12:1–10

TABLE S8. Functionally Relevant HUGO Gene Sets

HUGO Category	Group Numbers	Group Link
Ion_Channels	177	https://www.genenames.org/data/genegroup/#!/group/177
Cell_Adhesion_Molecules	16	https://www.genenames.org/data/genegroup/#!/group/16
Cell_Adhesion_Molecules	589	https://www.genenames.org/data/genegroup/#!/group/589
Cell_Adhesion_Molecules	597	https://www.genenames.org/data/genegroup/#!/group/597
Cell_Adhesion_Molecules	945	https://www.genenames.org/data/genegroup/#!/group/945
Cell_Adhesion_Molecules	1582	https://www.genenames.org/data/genegroup/#!/group/1582
Cell_Adhesion_Molecules	1583	https://www.genenames.org/data/genegroup/#!/group/1583
Zinc_Finger	26	https://www.genenames.org/data/genegroup/#!/group/26
Zinc_Finger	26_sub	https://www.genenames.org/data/genegroup/#!/group/26
Cluster_of_DF_Molecules	471	https://www.genenames.org/data/genegroup/#!/group/471
G-Protein_Coupled_Receptor	139	https://www.genenames.org/data/genegroup/#!/group/139
Homeoboxes_Forkheads_WNT_Hedgehog	360	https://www.genenames.org/data/genegroup/#!/group/360
Homeoboxes_Forkheads_WNT_Hedgehog	508	https://www.genenames.org/data/genegroup/#!/group/508
Homeoboxes_Forkheads_WNT_Hedgehog	522	https://www.genenames.org/data/genegroup/#!/group/522
Homeoboxes_Forkheads_WNT_Hedgehog	526	https://www.genenames.org/data/genegroup/#!/group/526
Homeoboxes_Forkheads_WNT_Hedgehog	527	https://www.genenames.org/data/genegroup/#!/group/527
Homeoboxes_Forkheads_WNT_Hedgehog	529	https://www.genenames.org/data/genegroup/#!/group/529
Homeoboxes_Forkheads_WNT_Hedgehog	1373	https://www.genenames.org/data/genegroup/#!/group/1373
Neuropeptides	542	https://www.genenames.org/data/genegroup/#!/group/542

FIGURE S3. Cross-species comparisons of amygdala cell-clusters. Cross-species clustering of Ce (A) revealed a number of evolutionarily conserved primate amygdala cell-types in each region. Barplots showing the proportion of cells in each region, identify a subset of cells that are more likely to be found in one species compared to another for both Ce (B). Track plots highlight a number of developmentally relevant from Medina et al. 2023 genes that differentiate Ce cell-clusters (C).



Supplemental Results

Similarities and differences across primate species: Additional Clusters

Of the 5 cell-clusters with significantly different proportions across species, we identified 1 Ce neuron cell-cluster with greater proportion in rhesus (Ce 9: *Neu._LGE_PENK*, Log₂ fold change 1.88) and 1 Ce cluster with greater proportion in humans (Ce 10: *Neu._LGE_COL4A5**; Log₂ fold change 2.58). Additionally, we identified 3 non-neuronal clusters that were enriched in rhesus (Ce 1: Oligo; Log₂ fold change 2.54, Ce 8: Ast.; Log₂ fold change 2.17, Ce 11: Oligo.; Log₂ fold change 1.93). When exploring proportional differences between the rhesus and human, vLa 1 neuron cell cluster was found to be enriched in Humans (La 5: *Neu._Pal._TRIM54*; Log₂ fold change 2.41).

Exploration of the composition of rhesus enriched cell-cluster Ce 8: *Neu._LGE_PENK* revealed expression of PAX6 and PDYN, suggesting it is from the dorsal striatal embryonic division (20). In contrast, Ce 10: *Neu._LGE_COLA45** expresses the ventral striatal embryonic division marker *ISL1*. Together, these data suggest a relative enrichment of dorsal and ventral striatal embryonic division cells in rhesus monkeys and humans, respectively. These data raise the possibility that differences in cell-clusters between species may reflect changes in the development of cell-types that result in different overall patterns of gene expression, but more data is required to make strong inferences.

Similarities and differences across primate species: Comparisons with previously published amygdala datasets

To examine the replicability of observed cross-species differences, which could be due to differences in punch locations, we compared our results to two previously published primate amygdala studies, Yu et al. (49) which included human and rhesus whole-amygdala samples, and Chiou et al. (50) which included rhesus amygdala samples. We used MetaNeighbor (pyNM) to create cell-type predictions based on our data, and searched for cells in other datasets that best fit this pattern. Using this approach, we found that La 5: *Neu._Pal._TRIM54*, which we identified as being specific to humans, had a corresponding cell-type in the human data from Yu et al., (Yu-Human Cluster 9 one-vs-best AUROC: 0.59) but had no corresponding cell-type in rhesus amygdala from Yu et al. or Choi et al.

In the Ce, we identified 5 clusters as having different proportions between humans and rhesus monkeys. We found strong support for Ce 8: Ast., which had corresponding clusters in the 2 rhesus but not the human datasets (Ce 8: Ast.: Yu-Rhesus Astrocyte Cluster 0 one-vs-best AUROC: 0.85; Noah-Rhesus astrocytes one-vs-best AUROC: 0.70); partial support for Ce 11: Oligo. (Ce 11: Oligo.: Yu-Rhesus Ext. Cluster 18 one-vs-best AUROC: 0.50; Noah-Rhesus glutamatergic neurons 37 one-vs-best AUROC: 0.48), which corresponded to excitatory clusters; partial support for Ce 1: Oligo. which had a corresponding cluster in 1/2 rhesus datasets but not the human dataset (Ce 1: Oligo.: Noah-Rhesus oligodendrocytes one-vs-best AUROC: 0.43); but did not find support for the species specificity of Ce 9: *Neu._LGE_PENK* or Ce 10: *Neu._LGE_COL4A5**, which did not have a corresponding cluster in either dataset.

Because of our focus on the "intercalated" cell cluster we performed the same MetaNeighbor the same analyses to ensure that Ce 3: *Neu._LGE_FOXP2* had corresponding clusters in other datasets. Results demonstrated a cluster similar to Ce 3: *Neu._LGE_FOXP2* in

all three datasets (Yu-Human Cluster 11 one-vs-best AUROC: 0.66; Yu-Rhesus Cluster 28 one-vs-best AUROC: 0.55; Noah-Rhesus GABAergic neurons 18 one-vs-best AUROC: 0.75).

Notably, *Ce 3: Neu_LGE_FOXP2* was most similar to rhesus and human clusters that identified as intercalated cells by Yu et al., by examining makers derived from mouse. Interestingly *Ce 3: Neu_LGE_FOXP2 did highly express markers of either DRD1 or SEMA3C which Yu et al., found to be markers of specific ITC subtypes.*

Taken together, these data support our interpretation of disease-relevant clusters, showing that *Ce 3: Neu_LGE_FOXP2* is conserved across primates, and *La 5: Neu_Pal_TRIM54* is enriched in humans.

FIGURE S4. Enrichment results for permutation analyses showing enrichment of disorder-relevant genes in many amygdala cell-clusters (A). Significant ($p < .05$) cell-clusters enriched for a particular disorder-gene-list are colored by $-\log_{10}(p\text{-value})$. Rows indicate the disorder-relevant gene-list. Columns indicate cell-clusters. Cell-clusters are named based on: [Region or species][Leiden Cluster Number][Cell Class][Developmental Origin (if Neuron)][Unique Marker or Overexpressed Gene]. Although many neuron clusters were significant, this is unlikely due to overlap between gene lists, as the percent overlap matrix between genes reported in different disorder-related gene-lists reveal relatively little overlap (B).

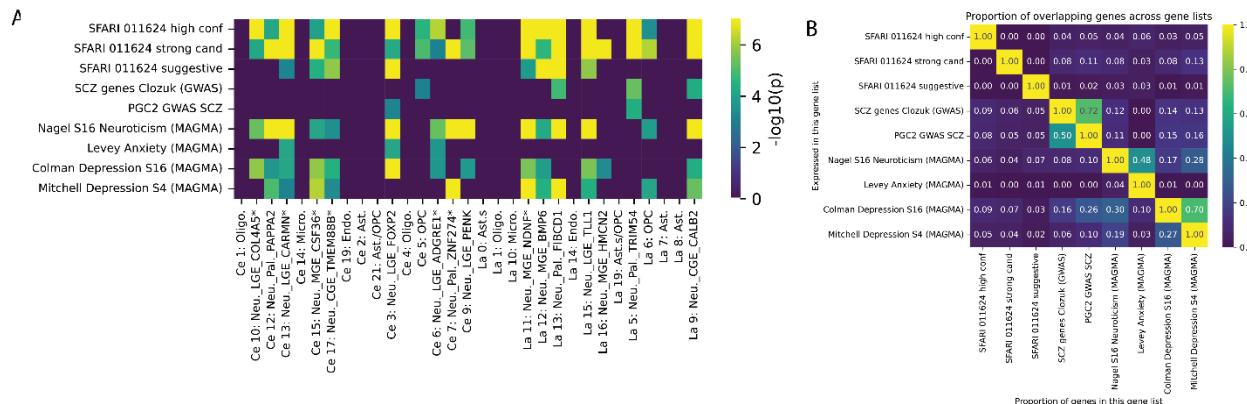


Table S9 is in a separate supplemental Excel file.

FIGURE S5. RNAscope in situ hybridization was used to confirm the presence of ITC markers along the ventral border of Ce. Ce was first localized using AChE on the adjacent slice (A; Ce outlined) and a 5X scan of the amygdala for DAPI (B) to select the region to be scanned at 20X (dashed- red box). Individual images for DAPI, GAD1, GAD2, VGLUT, TSHZ1, FOXP2, and NPFFR2, as well as the combined image for the target region are shown (C). The region selected for 63X image shown in Figure 4 is outlined (white box).

