Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Mouse spatial transcriptomics implicates the thalamus and cortex in autism and schizophrenia

David M. Young

davidyoung@llu.edu

University of California, San Francisco

Ruchira Sharma

University of California, San Francisco

Narjes Rohani

University of Oxford

Chimmi Dema

University of Oxford

Lindsay Liang

University of California, San Francisco

Bernie Devlin

University of Pittsburgh

Devanand S. Manoli

University of California, San Francisco

Stephan J. Sanders

University of Oxford

Research Article

Keywords: Autism spectrum disorder, schizophrenia, spatial transcriptomics, thalamus, cortex, gene enrichment

Posted Date: October 30th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11275405/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: Competing interest reported. S.J.S receives research funding from BioMarin Pharmaceutical. All other authors report no biomedical financial interests or potential conflicts of interest.

Abstract

Background

The past decade has seen tremendous progress in the identification of genes associated with complex neuropsychiatric disorders, including autism spectrum disorder (ASD) and schizophrenia. Expression patterns of these genes in single cell data strongly implicate excitatory and inhibitory neurons; however, there are limited data on the brain regions involved - a critical question for neurobiology. Spatial transcriptomics provide an opportunity to perform systematic multi-regional analyses to provide insights into this question.

Methods

We have generated a spatial transcriptomics dataset encompassing the diverse anatomical territories of the adult mouse brain sagittal midsection. We compare neuropsychiatric gene enrichment by applying Gene Fraction Enrichment Score (GFES), a novel statistical method that controls for differing neuronal proportions across regions.

Results

ASD-associated genes identified by exome sequencing were most enriched in the thalamus followed by the cortex. Schizophrenia genes from genome-wide association studies were also enriched in the thalamus, along with the hippocampus and cortex.

Limitations

Our study is limited to adult mouse brain and the sampled sections do not include all brain regions. Additionally, the analysis assumes that greater expression of genes associated with a disorder reflects greater involvement in the disorder, which might not necessarily hold true.

Conclusions

These findings add to the evidence that the thalamus plays a major role in neuropsychiatric disorders whilst supporting roles for the cortex and hippocampus. The results highlight shared and distinct patterns for pleiotropic brain disorders that could elucidate common underlying mechanisms and circuitry.

Background

Some brain-related disorders are caused by specific lesions in defined brain regions, such as the substantia nigra in Parkinson's Disease or the internal capsule in hemiplegic strokes. For many others, including autism spectrum disorder (ASD) and schizophrenia the brain regions involved remain poorly defined. Resolving the brain regions involved is of critical importance to identify molecular, cellular, or circuit level etiology and for translation, including biodistribution and neuromodulation.

In humans, neuroimaging offers a practical route to systematic analyses across multiple regions of the living brain, providing insights into structure, circuitry, and function. Analysis of large cohorts have identified subtle diffuse findings. These include structural changes, such as increased cortical thickness in ASD and decreased in schizophrenia [1–3] and functional changes, including disruption of the default mode network, often with decreased connectivity [3, 4]. The default mode network includes regions of the cortex and the hippocampus. Interpretation of neuroimaging data is complicated by small effect sizes and substantial variation among individuals and across studies.

Both ASD and schizophrenia are highly heritable, supporting a major role for genetic factors. Analyses of DNA from large cohorts of ASD and schizophrenia families, cases, and controls have yielded lists of genes associated with both disorders [5–10]. In both disorders these disorder-associated genes represent a foundation of causal factors from which to

investigate downstream neurobiology. Behavioral studies of animal models of individual genes have yielded mixed results, implicating most major brain regions [11]. In contrast, assessing the enrichment of these ASD and schizophrenia gene lists at the RNA level with transcriptomic data has clearly implicated excitatory and, to a lesser extent in ASD, inhibitory neurons in both disorders [9, 10, 12–14]. Most of the genes are expressed from the midfetal period onwards [14, 15]. Transcriptional changes can also be assessed in *postmortem* case/control analyses and have yielded altered expression of neuronal genes in the cortex from analyses of bulk tissue [16]. To date, other brain regions have not been assessed in sufficiently powered samples.

Ideally, we could identify brain regions where disorder-associated genes are commonly expressed. Spatial transcriptomics has emerged as a systematic method built upon the next-generation sequencing methods of scRNA-seq to profile the full transcriptome in a spatial context by simultaneously recording the position and identity of each RNA transcript [17]. Spatial transcriptomics approaches [17–19] have thus far been applied to a wide variety of brain disorders, including Alzheimer's disease [20, 21] to amyotrophic lateral sclerosis (ALS) [22]. By measuring across the transcriptome without prior selection of transcripts, spatial transcriptomics highlights regions enriched for disorder-associated genes, pointing to where they are more likely to be active and involved in disease when dysregulated.

Here, we present a multiregional, cross-disorder, spatial transcriptomics analysis of gene convergence in the adult mouse brain. We selected the sagittal midsection of the brain for its wide diversity of anatomical and functional territories and identified 29 clusters, which largely correspond to anatomical regions and sub-regions. To study the convergent expression of risk gene sets for polygenic neuropsychiatric disorders, we introduce a novel gene enrichment metric (GFES) to highlight spots expressing a preponderance of genes associated with each disorder whilst correcting for cell type proportion. Applying this to ASD and schizophrenia we observed enrichment in the murine thalamus and olfactory bulb, alongside the cortex and hippocampus.

Methods

Animals

The procedures of animal handling and tissue harvesting were approved by the Institutional Animal Care and Use Committee of University of California, San Francisco. All experiments were carried out on both male and female mice. The mouse strain (C57BL/6J) was obtained from The Jackson Laboratory and group housed in a UCSF barrier facility with a 12:12 hr light:dark cycle where food and water were available ad libitum.

Sample preparation

We performed whole brain dissections on 8-week-old (P56) male and female mice and flash froze the samples in isopentane. The frozen samples were then embedded in O.C.T. (cryostat-embedding compound, Tissue-Tek, Torrance, CA) over dry ice to avoid cracking and stored at -80°C. We used a cryostat to collect 2 lateral and 2 medial sagittal sections of the brain of 14µm thickness, using anatomical markers to identify midsection slices just lateral to the midline. The lateral sections were approximately 100–150µm apart from the medial sections. We mounted 4 such sections from 1 brain on a Visium Spatial Gene Expression Slide (10x Genomics, 1000187) and stored at -80°C until further processing.

Histology and sequencing

Spatial transcriptomics analysis was performed using the 10x Genomics Visium Spatial Gene Expression Assay. First, sections were placed on each of the four capture areas of the Visium Spatial Gene Expression slide. Each capture area contains a grid-like array of spots, where each spot contains primers to capture RNA lysed from permeabilized overlying tissue, including a unique spatial barcode of 12 bp to identify the spot's location, a unique molecular identifier of 12 bp to

distinguish each RNA transcript, and a poly(dT) sequence of 30 bp to capture mRNA for cDNA synthesis. After further cDNA amplification, the resulting libraries undergo Illumina sequencing, followed by alignment of transcript reads to the mouse genome using the 10x Spaceranger software package. Tissue fixation was performed in methanol at -20ºC for 30 minutes, then the slide was immediately stained with Hematoxylin and Eosin and imaged on an Automated Leica Aperio Versa 200 Slide Scanner at 20X magnification.

Next, permeabilization of tissue sections released the mRNA from cells, which were captured by primers on the Visium slide. Reverse transcription then added the spatial barcode, unique molecular identifier (UMI), and the partial Illumina sequencing Read 1 primer used for Visium. Following elution of samples from the Visium slide, one microliter of each sample was subjected to 25 cycles of qPCR. The optimal amplification cycles were determined using the Cq values at the exponential phase of the amplification plot, which is roughly 25% of the peak fluorescence. These values were used in the subsequent cDNA amplification reaction, followed by quality control on the Agilent Bioanalyzer. Ten microliters of cDNA were then used for fragmentation and library preparation, where dual indexes were added to the barcoded cDNA.

Quality control of the final libraries was completed on the Agilent 2100 Bioanalyzer using the High Sensitivity DNA Kit to determine the average library sizes, in addition to qPCR on the Applied Biosystems QuantStudio 5 Real-Time PCR System using the Roche KAPA Library Quantification Kit to determine the concentration of adapter-ligated libraries. Finally, libraries were pooled and sequenced on the NovaSeq 6000 S4 flow cell, paired-end 28x90 bp with 10 bp dual indexes.

Spatial transcriptomics data processing

We processed the Visium sequencing data using the 10x Spaceranger software. To prepare the histology images, we opened whole-slide images in QuPath, rotated images by 180º, downsampled them to a longest edge size of 2000 pixels, and saved them as TIFF files. In a Docker container with Spaceranger installed, we converted FASTQ files to BAMs and raw matrices through the $spaceran \geq rcount$ pipeline, using default settings and referencing the $refdata - \ \geq x - mm10 - 2020 - A$ mouse transcriptome reference from 10x.

After generating raw matrices, Seurat objects were created using default settings, which keeps only spots determined to be over tissue by Spaceranger's image processing. For quality control, we plotted gene counts by UMI counts for each sample, log-scaling each axis. To identify objective cutoffs for each type of count, we clustered points using DBSCAN (Density-Based Spatial Clustering and Application with Noise) as implemented by the $dbscan$ package. To find the tuning parameter epsilon, the maximum distance between two data points to be considered neighbors, we located the "elbow" position at 0.1 in a K nearest neighbor distance plot. For the minimum points parameter, we used a standard value of 3 based on 2 dimensions + 1. Clustering separated each sample into two dominant groups, with one large, dense cluster and a smaller, looser cluster. Clustering all spots from all samples together replicated this pattern and identified a threshold of about 100 UMIs and 100 genes between the clusters, which we took as the UMI and gene lower cutoffs along with a 30% mitochondrial upper cutoff for low-quality spots.

Once we filtered spots, we normalized batch effects by using the Seurat SCTransform data integration pipeline. After separately normalizing each sample batch (slides 1–2 and 3–4) by SCTransform, we performed anchor-based integration across batches using default settings except increasing the number of features to 3,000 for greater resolution.

Clustering and marker identification

We clustered spots using the Louvain-based shared nearest neighbor modularity optimization method implemented in Seurat, with default settings except increasing the number of dimensions to 30 when finding neighbors and reducing the resolution to 0.5. To visualize clusters, we plotted them after Uniform Manifold Approximation and Projection (UMAP) dimensional reduction, also generated with dimensions increased to 30.

To identify tissue specific gene markers, we first found differential markers using Seurat, comparing each cluster against all other clusters using default settings except increasing the minimum percentage threshold to 25% to find more broadly expressed genes. To refine this list, we took the markers with the greatest difference in percent expression within versus outside of the cluster and further filtered by the top log of the ratio of these percentages. Next, we applied the tau metric [23] as implemented in Kryuchko-Mostacci & Robinson-Rechavi [24] to each gene as a measure of tissue specificity to find the genes with the most focal, cluster-specific expression [24].

$$\tau = \frac{\sum_{i=1}^{n} (1 - \widehat{x_i})}{n-1}; \widehat{x_i} = \frac{x_i}{\underset{1 \leq i \leq n}{max} (x_i)}$$

Atlas alignment

To identify clusters anatomically, we mapped our samples into a common coordinate reference space using the Allen CCFv3 3D atlas to perform segmentation by registration. First, we manually identified the closest sagittal atlas slice corresponding to each of our sections. Next, we performed image registration to move each atlas section into alignment with the corresponding section. We used the Elastix image registration toolkit as implemented in our MagellanMapper software application [25]. After rescaling, rotating, and flipping the atlas section to approximate the sample orientation to aid registration initialization, we performed rigid registration using 1024 iterations of translation and 0-4096 iterations of affine transformations depending on the necessary cycles to align images visually. After rigid registration, we added deformable registration using 0-256 iterations of b-spline transformation with a voxel size of 150.

While these standard transformations typically aligned external borders well, internal structures often remained slightly misaligned, likely because of the multimodal registration [26]. To guide internal structure alignment in a semi-automated fashion, we identified clusters that clearly corresponded to large atlas regions, such as thalamic and hippocampal clusters. Using the corresponding-point b-spline transformation with 256 iterations and a voxel size of 30–50, we aligned the centers of each specified cluster-region pair. This atlas alignment allowed us to segment the samples using the atlas region annotations, where we assigned each spot to the corresponding atlas label at the spot's center pixel. To visualize the regions that mapped best to each cluster, we measured the proportion of spots in each cluster assigned to each atlas region and plotted these proportions as a heat map.

Cell type and region deconvolution

Each spot represents an amalgam of neighboring cells, potentially from diverse cell types. To estimate the proportion of cell types in each spot, we performed deconvolution using Robust Cell Type Decomposition (RCTD) from the spacexr package [27]. We ran each deconvolution in "full" doublet mode for no restrictions on cell type number and normalized the cell type proportions to a total of one per spot. To visualize these proportions, we plotted each cell type weighting on each sample to assess for expected locations. We also aggregated the mean weightings for each cell type within each cluster as a heat map. For region-specific datasets, such as the HypoMap hypothalamic atlas, we first subset our dataset to the corresponding region based on the registered anatomical delineations before performing the deconvolution.

Gene enrichment

We computed the gene enrichment fraction at each spot using the equation:

$$ASD_{i,j} = \frac{A_{i,j}}{T_{i,j}}$$

where $A_{i,j}$ = ASD gene UMIs at spot *i* in section *j*, $T_{i,j}$ = total gene UMIs at the same spot, making $ASD_{i,j}$ is the gene enrichment fraction.

To measure and compare the gene enrichment by cluster, we next performed an ANCOVA test as implemented in the R aov, where gene enrichment is the dependent variable, and cell type proportion is the categorical independent variable. Plots of the enrichment fraction versus neuronal proportion for each spot revealed heteroscedasticity — increased variance as neuronal fraction increased — which stabilized after applying a square root transformation of the enrichment values. We also observed a slight quadratic effect, which we modeled with a polynomial, resulting in the ANCOVA formula: `√ASD ~ neuron + neuron$^2$. $We \rightarrow oktheenrichmentofeachcluster \rightarrow betheestimatesasoutputomsummary.lm$`.

To identify statistical differences among regions, we could not use the estimates' p-values because the spots are spatially correlated with one another. Instead, we applied a bootstrapping resampling technique to measure 95% confidence intervals. For each cluster, we resampled its spots with replacement and reran the ANCOVA across all clusters, generating a new set of estimates. We repeated this process 1000 times, resulting in an estimate of the sampling distribution of ANCOVA estimates and their 95% confidence intervals.

To further validate the gene enrichment metric, we tested its ability to identify enrichment in simulations of gene expression. We simulated an RNA-seq dataset with 20,000 cells and 3,000 genes, where the number of UMIs for each gene in each cell was randomly taken from a uniform distribution. To simulate a baseline number of “expressed” genes, we randomly sampled 10% of all cell-gene pairs to have three times their original UMIs. To test whether the metric could detect differences in enriched clusters, we further expressed 10% of genes in two clusters, which demonstrated clear enrichment compared to the other clusters using the metric (Fig. S6a).

We also tested subtler enrichment by linearly increasing the number of expressed genes in the cells of three clusters. As expected, the metric showed a linear increase in enrichment across these clusters (Fig. S6b). To test non-linear enrichment, we increased the number of expressed genes exponentially across the whole simulated dataset, leading to an exponential enrichment curve using the metric, as anticipated (Fig. S6c).

In some cases, a single strongly expressed gene dominated the final enrichment results. To identify such genes, we inspected heatmaps of gene expression by cluster and compared results with and without these genes. We also compared enrichment on raw UMI counts versus normalized counts.

Geospatial interpolation through Kriging

To model gene enrichment spatially, we applied the Kriging geospatial interpolation technique [28] implemented in the $gstat$ package to predict values at higher resolution while accounting for covariates. Kriging models spatial autocorrelation, where points that are closer together are expected to be more similar. To model this autocorrelation, we first generated a semivariogram, which plots the variance of pairs of points by their distance, binned into similar distances. Next, we fit the variogram by empirically determining the sill (variance when the model starts to flatten), range (distance at this point), nugget (variance when distance is 0), and the mathematical model. With our fit variogram model, we used Ordinary Kriging to map it to a dense grid of points over each sample to predict the gene enrichment at each point using a weighted average of nearest neighboring points. To incorporate trends from the neuronal proportions, we applied Universal Kriging with neurons as the covariate, using the same formula as for the ANCOVA.

Results

Samples assayed from the sagittal midsection of mature mouse brain

We elected to assay sagittal midsections of the mature mouse brain (postnatal day 56 C57BL/6J; three female, one male) to enable multiple major brain regions to be assessed systematically, including cortex, hippocampus, basal ganglia, thalamus,

hypothalamus, and midbrain (Fig. 1a). Each 14μm section spanned the full extent of the brain dorsoventrally and from the posterior olfactory area to the midbrain anteroposteriorly, covering a substantial fraction of the brain's rostral-caudal and inferior-superior axes. Four sagittal sections were assessed from each mouse brain separated by about 100μm to capture a degree of medial-lateral variability (Fig. 1b-c).

Spatial gene expression was assayed using the 10x Visium platform. All four fresh frozen sections from each mouse were placed on a single Visium slide, which contains four capture areas of 6.5 x 6.5mm. Each capture area contained 4,992 spots of 55μm diameter arrayed in a grid-like pattern with 100μm between spot centers. Each spot captures RNA transcripts lysed from the overlying tissue, including a unique spatial barcode to identify the location of each transcript. An image of the tissue section was captured using a hematoxylin & eosin (H&E) stain and whole slide imaging microscope.

Anatomical alignment

Histological images were registered to the well-validated Allen Common Coordinate Framework version 3 (CCFv3) atlas [29] that represents an age-appropriate (P56) reference with 3D labels appropriate for all orthogonal axes, including the sagittal orientation. To identify the anatomical location of each slice on the medial-lateral axis, each microscopy image was compared to sagittal planes from the CCFv3 to identify the best match (Fig. 1c). Both rigid and deformable registration of each atlas plane to the corresponding tissue section was performed using the Elastix toolkit as implemented in our MagellanMapper software [25, 30], thereby assigning an anatomical structure to each spot (Fig. S1).

Transcriptomic quality control and clustering

Sequencing data were aligned to the mm10 mouse genome using STAR [31]. Data were processed using 10x Genomics' Space Ranger to yield tables of Unique Molecular Identifiers (UMIs) per feature per spot and analyzed by Seurat [32]. High quality spots were identified based on feature counts (≥ 100) and UMI counts (≥ 100) with thresholds determined by Density-Based Spatial Clustering and Application with Noise (DBSCAN) analysis [33] (Fig. S2a) and under 30% mitochondrial genes. A median of 3,274 high quality, tissue-covered spots were included per section (53,841 total), with a median of 16,541 UMIs per spot, and 4,974 features per spot (Fig. S2b). Seurat's SCTransform integration pathway was used to integrate and normalize counts across the 16 sections [32].

Applying Louvain clustering [34] to the 16 integrated sections identified 29 clusters (Fig. 2a), largely corresponding to anatomical structures, consistent with expectation [35] (Fig. 2b). Some major structures were subdivided into smaller clusters by layer (e.g., the cortex), by laterality (e.g., midline vs. medial clusters of L4-6 cortex and the thalamus), or by substructure (e.g., ventromedial hypothalamus). Three clusters were more dispersed (CTX-endo, TT-endo, BS-endo), with scattered spots enriched in hemoglobin genes probably representing spots with higher densities of vasculature.

Integrating anatomical registration and transcriptomic clusters

While anatomical registration aligned well to the outer tissue borders, some internal structures remained misaligned. Therefore, we coupled image registration with the transcriptomic data by leveraging 'anchor' regions with a clear match between clusters and anatomical structures, such as the thalamus. The centers of these anchor clusters and their corresponding atlas structures were linked to enable cluster-guided registration (Fig. S1). For each cluster, we plotted the proportion of spots assigned to each atlas-defined anatomical region as a heatmap (Fig. 2c). Most clusters matched a single anatomical region, highlighting the anatomical identity of each cluster.

Cluster marker identification

To further define and validate each cluster, we used Seurat to identify marker genes enriched in specific clusters (Fig. S3), refined the list to genes with large differences in percent expression across clusters, and applied a graded tissue-specificity metric, tau [23, 24]. Many of the resulting markers were substantially enriched in single clusters only (Fig. 3a), providing utility as anatomical markers, despite being less highly ranked based on mean differential fold change across clusters. For example, the marker with the most specificity to the hypothalamus was thyrotropin-releasing hormone, *Trh*, a known marker

of hypothalamic neurons, that had been ranked eighth based on differential fold change. Similarly, the highest ranked marker of the L4-6-mid cortex cluster was *Cort*, a gene expressed in cortical inhibitory neurons, whereas the highest differential marker was *Sst*, which happens to be structurally similar to *Cort* but also expressed in the ventromedial nucleus of the hypothalamus (VMH).

Each spatial spot assays the transcriptome of multiple overlaying cells; deconvolution can estimate the proportion of these cell types through comparison to the transcriptome of known cell types. We used established single-cell/nuclei RNA-seq datasets from specific brain regions (Fig. 3b) to perform deconvolution with RCTD [35, 39] to validate the anatomical assignment of clusters. The Zeisel *et al.* dataset [36] details the transcriptome of over half a million cells from contiguous regions dissected from the nervous system of adolescent mice (P20-30, CD-1/Swiss). Using cells from specific regions (e.g., cortex, thalamus, hippocampus, hypothalamus) as the input for deconvolution, we observed high concordance for spots assigned to the corresponding anatomical cluster (Fig. 3b). Thus, we demonstrate that our clusters are accurately assigned via image registration (Fig. 2), transcriptomic markers (Fig. 3a), and comparison to single cell data from dissected brain regions (Fig. 3b).

Other single-cell/nuclei RNA-seq datasets provide greater resolution within anatomical structures. Considering over 23,000 cells from the neocortex of P53-59 C57BL/6J mice in the Tasic et al. 2016 dataset [37], we observed concordance with excitatory neurons in specific cortical layers matching our anatomical clusters (Fig. 3b). Similarly, integration of 17 datasets to yield a murine hypothalamic atlas for 384,000 cells (HypoMap) [38], enabled specific nuclei of the hypothalamus to be identified in our spatial transcriptomic data (Fig. 3b).

Spatial gene enrichment corrected for cell type proportion

Having registered the spatial images and defined and validated spatial clusters, we assessed gene enrichment within spatial clusters to yield insights into brain regions related to ASD and schizophrenia. Prior analyses have demonstrated that genes associated with brain phenotypes are enriched in neurons [9, 10, 14]. Since the proportion of neurons varies across brain regions (Fig. S4), we must correct for the neuronal percentage per spot to assess regional enrichment. Therefore, we used the Zeisel *et al.* whole-brain dataset [36] to deconvolve each spot to estimate neuronal fraction. As expected, spots in the cortex and hippocampus had high neuronal fractions, while spots in the corpus callosum had low neuronal fractions.

Based on the logic that a gene needs to be expressed to contribute to a disorder, gene enrichment studies often use over-representation analysis to assess whether a defined list of genes are expressed more than expected in a specific cell type [40]. Each gene is defined as expressed or not expressed, based on exceeding a specific threshold, e.g., present in 25% of cells in a cell-type cluster [10]. The fraction of genes on the list that are expressed in the cluster is compared to the total number of genes assessed to calculate a p-value (e.g., Fisher's exact test). This approach benefits from simplicity; however, the inclusion of covariates, such as neuronal proportion, is complicated, varying the threshold defining 'expressed' can alter results, and the total number of expressed genes varies between clusters, confounding interpretation. Of note, the use of a threshold of cells with counts to define gene expression suggests that single cell gene enrichment analysis is measuring whether a gene is highly expressed, rather than an absolute measure of expressed or not expressed. This approach is embraced in Functional Class Storing (FCS) method [40] that uses quantitative expression to assess enrichment, typically generating a score per gene.

To apply gene enrichment analysis to spatial transcriptomic data, we assessed the proportion of total UMIs that aligned to genes in the list per spot, yielding a graded, self-normalized enrichment metric. To compare gene enrichment across clusters, we applied ANCOVA to the per spot metrics, treating clusters as the independent, categorical variable and per spot neuronal proportion as a covariate. This ANCOVA approach calculates regression estimates per cluster that can be compared between clusters. Because each spot is likely to be next to a spot in the same cluster, each spot is not independent, therefore the ANCOVA p-values are not informative. To overcome this limitation, we used a bootstrap procedure in which we assign a subset of spots randomly to each cluster and use ANCOVA to yield regression estimates. With multiple bootstrapping iterations we can define a 95% confidence interval per cluster, enabling significant differences

between clusters to be evaluated (Fig. S5). We have termed this graded, self-normalizing, covariate compatible, threshold-free, cell/spot-level metric 'Gene Fraction Enrichment Score' (GFES). Through simulation, we demonstrate that GFES can accurately and quantitatively distinguish different patterns of enrichment across transcriptomic clusters (Fig. S6).

Autism genes are enriched in the thalamus, olfactory bulb, and cortex

We next considered the enrichment for genes associated with neuropsychiatric disorders, starting with 72 ASD-associated genes from large-scale exome analyses (false discovery rate (FDR) ≤ 0.001 equivalent to genome-wide significance, Fig. S7, S8, Table S1) [9]. Of note, considering a larger list of 255 ASD-associated genes at genome-wide significance (FDR ≤ 0.1) yielded similar patterns (Fig. S9).

Across the 29 clusters, corrected for neuronal fraction [10], all clusters showed some degree of enrichment (Table S2), reflecting the high expression levels of ASD-associated genes throughout the brain [10, 41]. Eight clusters were especially enriched for ASD genes, mapping to the thalamus, main olfactory bulb (MOB), and cortex (L2-3, L4-6-med) (Fig. 4). The highest GFES metric for ASD genes (0.89, 95% CI:0.88–0.90; Fig. 4a) was observed in the medial thalamus (TH-med) with the four other thalamic clusters (TH-endo, PF, TH-antdor, TH-mid) also in the top seven ASD enriched clusters by GFES. The neighboring, but functionally and anatomically distinct, epithalamus (EPI) showed relatively low enrichment (0.70, 95% CI:0.69–0.72; Fig. 4a). After thalamus, the main olfactory bulb was the most enriched, followed by two cortical clusters (L2-3, L4-6-med) with no significant difference between upper and lower cortical layers. Two other cortical layers (CTX-endo, L4-6-middle) showed relatively low enrichment.

The observed thalamic enrichment is especially marked for nine ASD-associated genes (Fig. 4c). The clearest thalamic pattern was for *Tcf7l2* (Fig. 4d), a transcription factor that is more highly expressed in neurons of the thalamus than the cortex in both mice and humans (Fig. S10 [42]). Despite the strong *Tcf7l2* signal, thalamic enrichment of ASD-associated genes persists with its removal from the analysis (Fig. S9). Neuronal expression of *Nrxn1*, a pre-synaptic cell adhesion protein, and *Ctnnb1*, the Wnt signaling β-catenin protein show clear thalamic enrichment over other captured brain regions (Fig. 4d). Other ASD-associated genes are enriched in both the thalamus and cortex, such as *Stxbp1* which regulates the SNARE complex for neurotransmitter release.

Schizophrenia genes are enriched in the olfactory bulb, cortex, hippocampus, and thalamus

Exome-based analyses identify ten genes associated with schizophrenia through rare protein-coding variation [7], but this is too few genes for reliable analysis of gene enrichment. We therefore used genes identified from common variation in a recent large-scale GWAS study, using the 435 reported high-confidence protein-coding genes in a "broader fine-mapped set" (Table S1) [8]. Like the ASD-associated genes, these are enriched for neuronal markers as shown previously [8] and replicated in our dataset (Fig. S7), so we applied the same correction for neuronal proportion as in ASD. Assessing spatial enrichment identified similar brain regions in schizophrenia and ASD (R^2 = 0.37, p = 0.0005; see Fig. S11), with the main olfactory bulb (MOB), cortex (L2-3, L4-6-med), and thalamus (TH-med) being ranked highly by GFES in both disorders (Fig. 5a). Applying the Kriging geospatial model to both schizophrenia and ASD enrichment generated similar landscapes of gene enrichment (Fig. 5b, Fig. S11), although with a higher degree of enrichment in the dentate gyrus (DG) and Cornu Ammonis (CA) of the hippocampus in schizophrenia than ASD.

As with ASD, a minority of the 435 genes contribute strongly to the regional enrichment patterns (Fig. 5c). Multiple schizophrenia fine mapped genes are enriched in the hippocampus and cortex, including neurogranin, *NRGN*/*Nrgn*, the tyrosine kinase *PTK2B*/*Ptk2b*, the ribosomal protein *RPL13/Rpl13*, and the transcription factor *EGR1/Egr1* (Fig. 5d).

However, within these four examples, the finemap posterior probability (FPP) values are low for *NRGN*, *PTK2B*, and *RPL13* (Table 1), raising the possibility that other genes might better reflect the causal genetic effects of the locus [8]. Of note, within the *NRGN* locus, another gene (*VSIG2*, a membrane bound immune regulator) has a higher FPP (0.38) than *NRGN* with

the same index SNP [8], raising the possibility that *NRGN* does not contribute to schizophrenia liability. Removing *Nrgn*, the most prominent hippocampal gene, reorders the most enriched clusters, ranking the MOB and thalamus clusters at the top (Fig. 5e).

Focusing on the spatial dynamic genes with the higher fine-mapping estimates of the index SNP (Table 1) also supports the enrichment in the thalamus, cortex, and hippocampus. *ATP2A2*/*Atp2a2* is a calcium transporter associated with the endoplasmic reticulum and sarcoplasm and is enriched in the thalamus (TH-med). *NGEF/Ngef* is a neuronal guanine nucleotide exchange factor enriched in the cortex. *GPM6A/Gpm6a*, a neuronal membrane glycoprotein [43], enriched in both the dentate gyrus and thalamus and *GABBR2/Gabbr2* is a type B GABA receptor enriched in the thalamus (Fig. 5d).

Table 1
Selected fine mapped GWAS loci for spatially dynamic genes from Trubetskoy *et al.*[8]

Index SNP	Fine mapped SNP	Allele frequency	GWAS odds ratio	GWAS p-value	Finemap posterior probability	Human Gene	Human Ensembl ID	Mouse gene
rs4766428	rs4766428	0.46	0.93	$2x10^{-17}$	0.994	*ATP2A2*	ENSG00000174437	*Atp2a2*
rs778371	rs778371	0.26	0.93	$8x10^{-17}$	0.546	*NGEF*	ENSG00000066248	*Ngef*
rs12498839	rs12498839	0.07	0.89	$1x10^{-13}$	0.330	*GPM6A*	ENSG00000150625	*Gpm6a*
rs7915131	rs7915131	0.43	1.04	$4x10^{-8}$	0.213	*ZNF365*	ENSG00000138311	*Zfp365*
rs1121296	rs1121296	0.41	1.05	$3x10^{-9}$	0.202	*EGR1*	ENSG00000172260	*Egr1*
rs41533650	rs41533650	0.20	0.93	$8x10^{-14}$	0.169	*GPM6A*	ENSG00000150625	*Gpm6a*
rs10985811	rs10985811	0.21	0.95	$3x10^{-8}$	0.111	*GABBR2*	ENSG00000136928	*Gabbr2*
rs9926049	rs9926049	0.27	0.95	$3x10^{-10}$	0.083	*GRIN2A*	ENSG00000183454	*Grin2a*
rs2022265	rs2022265	0.47	1.05	$3x10^{-10}$	0.063	*SNAP91*	ENSG00000065609	*Snap91*
rs7938083	rs7938083	0.29	0.95	$2x10^{-9}$	0.041	*TMX2*	ENSG00000213593	*Tmx2*
rs1940171	rs55661361	0.32	1.06	$1x10^{-12}$	0.041	*NRGN*	ENSG00000154146	*Nrgn*
rs11076631	rs12932628	0.41	0.96	$4x10^{-7}$	0.034	*RPL13*	ENSG00000167526	*Rpl13*
rs73229090	rs2271920	0.37	0.96	$1x10^{-6}$	0.033	*PTK2B*	ENSG00000120899	*Ptk2b*
rs3752827	rs3752827	0.33	1.05	$5x10^{-10}$	0.012	*YWHAE*	ENSG00000108953	*Ywhae*

Returning to the exome-derived list of ten genes, two have dynamic spatial profiles (Fig. 5d, Fig. S12). *GRIN2A*/*Grin2a* is enriched in the hippocampus, cortex, and thalamus and *CACNA1G*/*Cacna1g* is enriched in the thalamus (Fig. 5d). The gene *NRXN1*, enriched in the thalamus (Fig. 4d) is also robustly associated with both ASD and schizophrenia, mostly from copy number variants [44].

Discussion

The contribution of specific brain regions to most human traits and disorders remains undefined, especially for neuropsychiatric disorders. Here, we investigate this question in novel spatial transcriptomic data from sagittal midsections

of the adult mouse brain.

Site-specific transcriptomic data largely clustered by known anatomical regions (Fig. 2), such as the thalamus or cortex, with a few clusters defined by the intersection of anatomical regions and high proportions of endothelial marker genes (e.g., CTX-endo, TH-endo). To associate neuropsychiatric traits with these clusters, we developed GFES – the Gene Fraction Enrichment Score – to assess per spot gene enrichment controlling for covariates, such as cell type proportion. Because multiple prior analyses demonstrate that genes associated with neuropsychiatric traits are enriched in neurons [9, 10, 45], we used GFES to control for neuronal proportion focusing on the degree of gene enrichment by region rather than regional neuronal density.

Integrating hundreds of genes associated with ASD and schizophrenia with mouse spatial transcriptomic data using GFES, we assessed regional enrichment corrected for neuronal density. No single brain region emerged unequivocally in either disorder, instead multiple large brain regions were enriched. This fits with the observation that symptoms of ASD and schizophrenia are not a frequent consequence of individual brain lesions [46–48]. This may indicate that these disorders are not mediated by any individual brain region acting in isolation, but rather a circuit level phenomenon acting via neuronal impairment across multiple regions.

Nonetheless, some brain regions were substantially more enriched for ASD and/or schizophrenia genes than others. Regional enrichment was somewhat correlated between the two disorders (R^2 = 0.37, Fig. S12) and similar regions ranked highly (thalamus, cortex, olfactory bulb) despite the ASD analysis using genes defined by exome sequencing and schizophrenia using GWAS and fine mapping. While clinically distinct, there is also epidemiological and genetic evidence of etiological overlap, including a five-fold increase schizophrenia relative risk in autistic individuals than in controls [49, 50] and shared genetic loci [7, 9].

The cortical enrichment in both conditions is consistent with the wealth of evidence linking this human-enlarged brain region to neuropsychiatric traits including transcriptomics, structural and functional neuroimaging, and the known role in cognition, behavior, and language [3, 15, 16, 51–55]. However, because the cortex is often analyzed in isolation, it is hard to contextualize this evidence. While our data support cortical involvement through the high enrichment values, in neither disorder is cortex the region most enriched for causal genetic factors.

The hippocampus has been assessed extensively as a model of synaptic plasticity [56], which, in turn, has been implicated in ASD and schizophrenia [57, 58]. However, there is less evidence directly linking the hippocampus to these traits [54, 55, 59]. The hippocampus has a well-established role in memory [60, 61] and an emerging role in social behavior [62]. Our data support a role of the hippocampus in schizophrenia (Fig. 5), but this is less clear in ASD (Fig. 4).

Surprisingly, the strongest enrichment in ASD was observed in the medial thalamus (Fig. 4), a region also ranked fifth in schizophrenia (Fig. 5). Enrichment in the middle or anterior-dorsal thalamus was less pronounced; the limited availability of additional thalamic transcriptomic datasets prevents further resolution. Changes in the shape of the mediodorsal thalamus (included in the medial thalamus) and pulvinar thalamus (primate specific) are reported in ASD using structural MRI [63]. With functional MRI, overconnectivity between the cortex and thalamus have been described in both ASD and schizophrenia [3, 64–66] and a similar observation is observed in organoid models of the ASD- and schizophrenia-associated 22q11.2 microdeletion [67]. However, other analyses identify reduced thalamocortical connectivity in schizophrenia [68]. Our transcriptomic data are not able to define specific circuits; however, the enrichment in both the thalamus and cortex could be consistent with a role of thalamocortical circuitry. This thalamic enrichment for ASD-associated genes is consistent with analysis of human prenatal single cell data across development [69] and human prenatal spatial transcriptomic data [70].

The clearest thalamic-enriched ASD-associated gene was for *Tcf7l2*. In the human thalamus, *TCF7L2* is expressed in most cell types but especially excitatory neurons [71], whereas in the human cortex expression is primarily in progenitor cells and glia [72, 73] but not in mature neurons. *TCF7L2* is robustly associated with ASD (FDR = 0.0004 [9]) and neurodevelopmental delay in exome-sequenced cohorts (FDR = 4×10^{-11}) [9, 74] and characterized as clinical syndrome in a case series of 11

individuals [75]. *CTNNB1* follows a similar, but less dramatic expression pattern in human and mouse data [42, 71, 72] and similarly robustly associated with ASD (ASD FDR = 8 x 10^{-5}, NDD FDR = 0). Of note, the proteins encoded by *TCF7L2* (TCF7L2, previously TCF-4) and *CTNNB1* (β-catenin) interact in the canonical Wnt signaling pathway [76]. β-catenin binds to TCF7L2 leading to transcriptional activation of Wnt target genes [77], so haploinsufficiency of either gene would be expected suppress Wnt target genes. To assess if the Wnt pathway is involved in ASD, we considered 254 genes in the Wnt signaling pathway, as defined by PANTHER Pathways (P00057) for gene ontology (GO:0016055). We observed a trend towards enrichment across the 72 most robust ASD-associated genes (OR: 3.31, 95% CI:0.88-Inf, p = 0.07, Fisher's exact test) that is significant with the more liberal 255 ASD-associated gene list (OR: 4.22, 95% CI:2.43-Inf, p = 3.3x10^{-5}, Fisher's exact test). To date, a thalamic Wnt signaling pathway has not been defined.

We also observed substantial enrichment of the main olfactory bulb (MOB), which is the most enriched region in schizophrenia (Fig. 5) and third most enriched in ASD (Fig. 4). Olfaction is critical to social behavior and interaction in rodents, but it is hard to interpret this finding in humans where olfaction plays less of a role and the MOB is dramatically smaller [78]. Observed MOB enrichment may be indicative of the importance of sensory processing in neuropsychiatric traits and social behavior. Of note, assessment of postmortem gene expression in the cortex of individuals with ASD identified the most substantial dysregulation in sensory cortex (V1, visual cortex), rather than the prefrontal cortex [52].

Limitations

We note several limitations to our analyses. First, assessments were based on the adult mouse, which may not capture regions that are transiently enriched in earlier developmental windows or human-specific expression patterns; notably analysis of human prenatal thalamus observes similar enrichment for ASD-genes [69, 70]. Future assessments would be enhanced by longitudinal data and integrated analysis with human data [45, 71, 79]. Second, while spatial transcriptomics enables systematic cross-regional analyses, the sagittal slices we selected exclude some major brain structures, such as the cerebellum, amygdala, and some hypothalamic nuclei. There is a need for more comprehensive multi-regional transcriptomic data. Third, we make the implicit assumption that high expression of genes associated with a trait indicates a causal role. This is logical for truly specific genes (e.g., *Tcf7l2* in the thalamus) but it is less evident for moderate quantitative differences; the sensitivity of GFES to highly specific genes is a strength in this context. Caution about the relationship between high expression and causality is warranted in the field at large. Analyses including transcriptomic assessment of conditional heterozygous knockdown models of ASD/schizophrenia associated genes complemented by evidence of downstream disruption, e.g., protein levels, electrophysiology, and calcium imaging, will be necessary to validate these findings.

Conclusions

Systematic analysis of multiregional spatial transcriptomic data supports the role of the thalamus and cortex in ASD and schizophrenia and also implicates the sensory structures and the hippocampus. Together, these shared and distinct patterns in the brain provide further insight into the underlying neurobiology and circuitry of these two disorders that may contribute to their complexity.

Abbreviations

ALS
Amyotrophic lateral sclerosis
ASD
Autism spectrum disorder
CCFv3
Allen Common Coordinate Framework version 3
DBSCAN

Density-Based Spatial Clustering of Applications with Noise
EPI
Epithalamus
FCS
Functional Class Storing
FDR
False discovery rate
FPP
Fine-mapping posterior probability
GFES
Gene Fraction Enrichment Score
GWAS
Genome-wide association study
MOB
Main olfactory bulb
RCTD
Robust Cell Type Decomposition
UMAP
Uniform Manifold Approximation and Projection
UMI
Unique molecular identifier

Declarations

Acknowledgments

We thank Mylinh Bernardi and Horng-Ru Lin of the Gladstone Genomics Core for their assistance with Visium spatial gene expression assay for fresh frozen tissue.

Authors' contributions

D.Y and S.J.S conceptualized and designed the study; B.D, D.S.M, and S.J.S secured funding for the project; R.S performed the experiments and generated the spatial transcriptomic data; D.Y and L.L processed the data; D.Y, B.D, and S.J.S performed the data and statistical analyses; D.Y and B.D developed the GFES methodology; D.Y, N.R, C.D, BD, S.J.S interpreted the results; D.Y and S.J.S prepared the figures; D.Y and S.J.S wrote the original draft of the manuscript; D.Y, N.R, C.D, B.D, D.S.M, and S.J.S reviewed and revised the manuscript; C.D organized the tables and figures and reviewed the references; all authors read and approved the final manuscript.

Funding

Sequencing was performed at the UCSF CAT, supported by UCSF PBBR, RRP IMIA, and NIH 1S10OD028511-01 grants. This work was supported by a grant from the Simons Foundation Autism Research Initiative Sex Differences Collaboration (SFARI #736613 to S.J.S.), the National Institute of Mental Health (grant numbers R01MH125516 to S.J.S, R01MH129751 to S.J.S.), and HDRUK: Health Data Research UK QQ2 Molecules to Health Records Driver Programme to S.J.S.

Competing interests

S.J.S receives research funding from BioMarin Pharmaceutical. All other authors report no biomedical financial interests or potential conflicts of interest.

Ethics approval and consent to participate

Animal handling and tissue collection procedures were approved by the Institutional Animal Care and Use Committee at the University of California, San Francisco.

Consent for publication

Not applicable.

Preregistration

Not applicable.

Software versions

We used the following software versions: Space Ranger 2.0, QuPath 0.3.2, Seurat 4.1, dbscan 1.1, Elastix 5.0, MagellanMapper 1.6, spacexr 2.0, gstat 2.1, R 4.2, Python 3.8, Java 17.

Data availability

The data will be released upon publication. The raw, unprocessed spatial transcriptomics data will be available on GEOXXX as FASTQ files. The processed data will be available as an RDS Seurat object file, also on GEOXXX.

Code availability

The source code including the enrichment pipeline will be released upon publication on GitHub at XXX.

References

1. van Rooij D, Anagnostou E, Arango C, Auzias G, Behrmann M, Busatto GF, et al. Cortical and Subcortical Brain Morphometry Differences Between Patients With Autism Spectrum Disorder and Healthy Individuals Across the Lifespan: Results From the ENIGMA ASD Working Group. Am J Psychiatry. 2018;175:359–69. https://doi.org/10.1176/appi.ajp.2017.17010100.
2. Bedford SA, Park MTM, Devenyi GA, Tullo S, Germann J, Patel R, et al. Large-scale analyses of the relationship between sex, age and intelligence quotient heterogeneity and cortical morphometry in autism spectrum disorder. Mol Psychiatry. 2020;25:614–28. https://doi.org/10.1038/s41380-019-0420-6.
3. Moreau CA, Raznahan A, Bellec P, Chakravarty M, Thompson PM, Jacquemont S. Dissecting autism and schizophrenia through neuroimaging genomics. Brain. 2021;144:1943–57. https://doi.org/10.1093/brain/awab096.
4. Parellada M, Andreu-Bernabeu Á, Burdeus M, San José Cáceres A, Urbiola E, Carpenter LL, et al. In Search of Biomarkers to Guide Interventions in Autism Spectrum Disorder: A Systematic Review. Am J Psychiatry. 2023;180:23–40. https://doi.org/10.1176/appi.ajp.21100992.
5. Choi L, An J-Y. Genetic architecture of autism spectrum disorder: Lessons from large-scale genomic studies. Neurosci Biobehav Rev. 2021;128:244–57. https://doi.org/10.1016/j.neubiorev.2021.06.028.
6. Sanders SJ, He X, Willsey AJ, Ercan-Sencicek AG, Samocha KE, Cicek AE, et al. Insights into Autism Spectrum Disorder Genomic Architecture and Biology from 71 Risk Loci. Neuron. 2015;87:1215–33. https://doi.org/10.1016/j.neuron.2015.09.016.
7. Singh T, Poterba T, Curtis D, Akil H, Al Eissa M, Barchas JD, et al. Rare coding variants in ten genes confer substantial risk for schizophrenia. Nature. 2022;604:509–16. https://doi.org/10.1038/s41586-022-04556-w.
8. Trubetskoy V, Pardiñas AF, Qi T, Panagiotaropoulou G, Awasthi S, Bigdeli TB, et al. Mapping genomic loci implicates genes and synaptic biology in schizophrenia. Nature. 2022;604:502–8. https://doi.org/10.1038/s41586-022-04434-5.

9. Fu JM, Satterstrom FK, Peng M, Brand H, Collins RL, Dong S, et al. Rare coding variation provides insight into the genetic architecture and phenotypic context of autism. Nat Genet. 2022;54:1320–31. https://doi.org/10.1038/s41588-022-01104-0.
10. Satterstrom FK, Kosmicki JA, Wang J, Breen MS, De Rubeis S, An J-Y, et al. Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. Cell. 2020;180:568–e58423. https://doi.org/10.1016/j.cell.2019.12.036.
11. Garrido D, Beretta S, Grabrucker S, Bauer HF, Bayer D, Sala C, et al. Shank2/3 double knockout-based screening of cortical subregions links the retrosplenial area to the loss of social memory in autism spectrum disorders. Mol Psychiatry. 2022. https://doi.org/10.1038/s41380-022-01756-8.
12. Elmentaite R, Domínguez Conde C, Yang L, Teichmann SA. Single-cell atlases: shared and tissue-specific cell types across human organs. Nat Rev Genet. 2022. https://doi.org/10.1038/s41576-022-00449-w.
13. Piwecka M, Rajewsky N, Rybak-Wolf A. Single-cell and spatial transcriptomics: deciphering brain complexity in health and disease. Nat Rev Neurol. 2023;1–17. https://doi.org/10.1038/s41582-023-00809-y.
14. Li M, Santpere G, Kawasawa YI, Evgrafov OV, Gulden FO, Pochareddy S, et al. Integrative functional genomic analysis of human brain development and neuropsychiatric risks. Science. 2018;362. https://doi.org/10.1126/science.aat7615.
15. Willsey AJ, Sanders SJ, Li M, Dong S, Tebbenkamp AT, Muhle RA, et al. Coexpression Networks Implicate Human Midfetal Deep Cortical Projection Neurons in the Pathogenesis of Autism. Cell. 2013;155:997–1007. https://doi.org/10.1016/j.cell.2013.10.020.
16. Gandal MJ, Zhang P, Hadjimichael E, Walker RL, Chen C, Liu S, et al. Transcriptome-wide isoform-level dysregulation in ASD, schizophrenia, and bipolar disorder. Science. 2018;362:eaat8127. https://doi.org/10.1126/science.aat8127.
17. Maynard KR, Collado-Torres L, Weber LM, Uytingco C, Barry BK, Williams SR, et al. Transcriptome-scale spatial gene expression in the human dorsolateral prefrontal cortex. Nat Neurosci. 2021;24:425–36. https://doi.org/10.1038/s41593-020-00787-0.
18. Ståhl PL, Salmén F, Vickovic S, Lundmark A, Navarro JF, Magnusson J, et al. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. Science. 2016;353:78–82. https://doi.org/10.1126/science.aaf2403.
19. Ortiz C, Navarro JF, Jurek A, Märtin A, Lundeberg J, Meletis K. Molecular atlas of the adult mouse brain. Sci Adv. 2020;6:eabb3446. https://doi.org/10.1126/sciadv.abb3446.
20. Chen W-T, Lu A, Craessaerts K, Pavie B, Sala Frigerio C, Corthout N, et al. Spatial Transcriptomics and In Situ Sequencing to Study Alzheimer's Disease. Cell. 2020;182:976–e99119. https://doi.org/10.1016/j.cell.2020.06.038.
21. Chen S, Chang Y, Li L, Acosta D, Li Y, Guo Q, et al. Spatially resolved transcriptomics reveals genes associated with the vulnerability of middle temporal gyrus in Alzheimer's disease. Acta Neuropathol Commun. 2022;10:188. https://doi.org/10.1186/s40478-022-01494-6.
22. Maniatis S, Äijö T, Vickovic S, Braine C, Kang K, Mollbrink A, et al. Spatiotemporal dynamics of molecular pathology in amyotrophic lateral sclerosis. Science. 2019;364:89–93. https://doi.org/10.1126/science.aav9776.
23. Yanai I, Benjamin H, Shmoish M, Chalifa-Caspi V, Shklar M, Ophir R, et al. Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. Bioinformatics. 2005;21:650–9. https://doi.org/10.1093/bioinformatics/bti042.
24. Kryuchkova-Mostacci N, Robinson-Rechavi M. A benchmark of gene expression tissue-specificity metrics. Brief Bioinform. 2017;18:205–14. https://doi.org/10.1093/bib/bbw008.
25. Young DM, Fazel Darbandi S, Schwartz G, Bonzell Z, Yuruk D, Nojima M, et al. Constructing and optimizing 3D atlases from 2D data with application to the developing mouse brain. eLife. 2021;10. https://doi.org/10.7554/eLife.61408.
26. Goubran M, Leuze C, Hsueh B, Aswendt M, Ye L, Tian Q, et al. Multimodal image registration and connectivity analysis for integration of connectomic data from microscopy to MRI. Nat Commun. 2019;10:5504. https://doi.org/10.1038/s41467-019-13374-0.

27. Cable DM, Murray E, Zou LS, Goeva A, Macosko EZ, Chen F, et al. Robust decomposition of cell type mixtures in spatial transcriptomics. Nat Biotechnol. 2022;40:517–26. https://doi.org/10.1038/s41587-021-00830-w.
28. Krige DG. A statistical approach to some basic mine valuation problems on the Witwatersrand. J South Afr Inst Min Metall. South Afr Inst Min Metall. 1951;52:119–39. https://doi.org/10.10520/AJA0038223X_4792.
29. Wang Q, Ding S-L, Li Y, Royall J, Feng D, Lesnar P, et al. The Allen Mouse Brain Common Coordinate Framework: A 3D Reference Atlas. Cell. 2020;181:936–e95320. https://doi.org/10.1016/j.cell.2020.04.007.
30. Klein S, Staring M, Murphy K, Viergever MA, Pluim JPW. elastix: a toolbox for intensity-based medical image registration. IEEE Trans Med Imaging. 2010;29:196–205. https://doi.org/10.1109/TMI.2009.2035616.
31. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinforma Oxf Engl. 2013;29:15–21. https://doi.org/10.1093/bioinformatics/bts635.
32. Hao Y, Hao S, Andersen-Nissen E, Mauck WM, Zheng S, Butler A, et al. Integrated analysis of multimodal single-cell data. Cell. 2021;184:3573–e358729. https://doi.org/10.1016/j.cell.2021.04.048.
33. Ester M, Kriegel H-P, Xu X, Sander J. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. Portland, Oregon: AAAI; 1996. pp. 226–31.
34. Blondel VD, Guillaume J-L, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. J Stat Mech Theory Exp. 2008;2008:P10008. https://doi.org/10.1088/1742-5468/2008/10/P10008.
35. Chen J, Liu W, Luo T, Yu Z, Jiang M, Wen J, et al. A comprehensive comparison on cell-type composition inference for spatial transcriptomics data. Brief Bioinform. 2022;bbac245. https://doi.org/10.1093/bib/bbac245.
36. Zeisel A, Hochgerner H, Lönnerberg P, Johnsson A, Memic F, van der Zwan J, et al. Molecular Architecture of the Mouse Nervous System. Cell. 2018;174:999–e101422. https://doi.org/10.1016/j.cell.2018.06.021.
37. Tasic B, Menon V, Nguyen TN, Kim TK, Jarsky T, Yao Z, et al. Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. Nat Neurosci. 2016;19:335–46. https://doi.org/10.1038/nn.4216.
38. Steuernagel L, Lam BYH, Klemm P, Dowsett GKC, Bauder CA, Tadross JA, et al. HypoMap—a unified single-cell gene expression atlas of the murine hypothalamus. Nat Metab. 2022;4:1402–19. https://doi.org/10.1038/s42255-022-00657-y.
39. Li B, Zhang W, Guo C, Xu H, Li L, Fang M, et al. Benchmarking spatial and single-cell transcriptomics integration methods for transcript distribution prediction and cell type deconvolution. Nat Methods. 2022. https://doi.org/10.1038/s41592-022-01480-9.
40. Maleki F, Ovens K, Hogan DJ, Kusalik AJ. Gene Set Analysis: Challenges, Opportunities, and Future Research. Front Genet. 2020;11:654. https://doi.org/10.3389/fgene.2020.00654.
41. Fazel Darbandi S, An J-Y, Lim K, Page NF, Liang L, Young DM, et al. Five autism-associated transcriptional regulators target shared loci proximal to brain-expressed genes. Cell Rep. 2024;43:114329. https://doi.org/10.1016/j.celrep.2024.114329.
42. Kang HJ, Kawasawa YI, Cheng F, Zhu Y, Xu X, Li M, et al. Spatio-temporal transcriptome of the human brain. Nature. 2011;478:483–9. https://doi.org/10.1038/nature10523.
43. Gutiérrez Fuster R, León A, Aparicio GI, Brizuela Sotelo F, Scorticati C. Combined additive effects of neuronal membrane glycoprotein GPM6a and the intercellular cell adhesion molecule ICAM5 on neuronal morphogenesis. J Neurochem. 2025;169:e16231. https://doi.org/10.1111/jnc.16231.
44. Shanta O, Klein M, Sacks M, MacDonald JR, Maihofer A, Ahangari M, et al. A cross-disorder analysis of CNVs finds novel loci and dose-dependent relationships of genes to psychiatric traits. MedRxiv Prepr Serv Health Sci. 2025. https://doi.org/10.1101/2025.07.11.25331310. 2025.07.11.25331310.
45. Werling DM, Pochareddy S, Choi J, An J-Y, Sheppard B, Peng M, et al. Whole-Genome and RNA Sequencing Reveal Variation and Transcriptomic Coordination in the Developing Human Prefrontal Cortex. Cell Rep. 2020;31:107489. https://doi.org/10.1016/j.celrep.2020.03.053.

46. Samanta D. An Updated Review of Tuberous Sclerosis Complex-Associated Autism Spectrum Disorder. Pediatr Neurol. 2020;109:4–11. https://doi.org/10.1016/j.pediatrneurol.2020.03.008.
47. Hamner T, Shih E, Ichord R, Krivitzky L. Children with perinatal stroke are at increased risk for autism spectrum disorder: Prevalence and co-occurring conditions within a clinically followed sample. Clin Neuropsychol. 2022;36:981–92. https://doi.org/10.1080/13854046.2021.1955150.
48. Saponaro S, Giuliano A, Bellotti R, Lombardi A, Tangaro S, Oliva P, et al. Multi-site harmonization of MRI data uncovers machine-learning discrimination capability in barely separable populations: An example from the ABIDE dataset. NeuroImage Clin. 2022;35:103082. https://doi.org/10.1016/j.nicl.2022.103082.
49. Jutla A, Foss-Feig J, Veenstra-VanderWeele J. Autism spectrum disorder and schizophrenia: an updated conceptual review. Autism Res Off J Int Soc Autism Res. 2022;15:384–412. https://doi.org/10.1002/aur.2659.
50. Lai M-C, Kassee C, Besney R, Bonato S, Hull L, Mandy W, et al. Prevalence of co-occurring mental health diagnoses in the autism population: a systematic review and meta-analysis. Lancet Psychiatry. 2019;6:819–29. https://doi.org/10.1016/S2215-0366(19)30289-5.
51. Batiuk MY, Tyler T, Dragicevic K, Mei S, Rydbirk R, Petukhov V, et al. Upper cortical layer-driven network impairment in schizophrenia. Sci Adv. 2022;8:eabn8367. https://doi.org/10.1126/sciadv.abn8367.
52. Gandal MJ, Haney JR, Wamsley B, Yap CX, Parhami S, Emani PS, et al. Broad transcriptomic dysregulation occurs across the cerebral cortex in ASD. Nature. 2022;611:532–9. https://doi.org/10.1038/s41586-022-05377-7.
53. Wamsley B, Bicks L, Cheng Y, Kawaguchi R, Quintero D, Margolis M, et al. Molecular cascades and cell type-specific signatures in ASD revealed by single-cell genomics. Science. 2024;384:eadh2602. https://doi.org/10.1126/science.adh2602.
54. Dear R, Wagstyl K, Seidlitz J, Markello RD, Arnatkevičiūtė A, Anderson KM, et al. Cortical gene expression architecture links healthy neurodevelopment to the imaging, transcriptomics and genetics of autism and schizophrenia. Nat Neurosci. 2024;27:1075–86. https://doi.org/10.1038/s41593-024-01624-4.
55. Rexrode LE, Hartley J, Showmaker KC, Challagundla L, Vandewege MW, Martin BE, et al. Molecular profiling of the hippocampus of children with autism spectrum disorder. Mol Psychiatry. 2024;29:1968–79. https://doi.org/10.1038/s41380-024-02441-8.
56. Citri A, Malenka RC. Synaptic plasticity: multiple forms, functions, and mechanisms. Neuropsychopharmacol Off Publ Am Coll Neuropsychopharmacol. 2008;33:18–41. https://doi.org/10.1038/sj.npp.1301559.
57. Bourgeron T. From the genetic architecture to synaptic plasticity in autism spectrum disorder. Nat Rev Neurosci. 2015;16:551–63. https://doi.org/10.1038/nrn3992.
58. Stephan KE, Baldeweg T, Friston KJ. Synaptic plasticity and dysconnection in schizophrenia. Biol Psychiatry. 2006;59:929–39. https://doi.org/10.1016/j.biopsych.2005.10.005.
59. Banker SM, Gu X, Schiller D, Foss-Feig JH. Hippocampal contributions to social and cognitive deficits in autism spectrum disorder. Trends Neurosci. 2021;44:793–807. https://doi.org/10.1016/j.tins.2021.08.005.
60. Konkel A, Warren DE, Duff MC, Tranel DN, Cohen NJ. Hippocampal amnesia impairs all manner of relational memory. Front Hum Neurosci. 2008;2:15. https://doi.org/10.3389/neuro.09.015.2008.
61. Olton DS, Becker JT, Handelmann GE. Hippocampus, space, and memory. Behav Brain Sci. United Kingdom: Cambridge University Press; 1979;2:313–65. https://doi.org/10.1017/S0140525X00062713
62. Schafer M, Schiller D, Navigating Social Space. Neuron. 2018;100:476–89. https://doi.org/10.1016/j.neuron.2018.10.006.
63. Schuetze M, Park MTM, Cho IY, MacMaster FP, Chakravarty MM, Bray SL. Morphological Alterations in the Thalamus, Striatum, and Pallidum in Autism Spectrum Disorder. Neuropsychopharmacol Off Publ Am Coll Neuropsychopharmacol. 2016;41:2627–37. https://doi.org/10.1038/npp.2016.64.

64. Ayub R, Sun KL, Flores RE, Lam VT, Jo B, Saggar M, et al. Thalamocortical connectivity is associated with autism symptoms in high-functioning adults with autism and typically developing adults. Transl Psychiatry. 2021;11:1–9. https://doi.org/10.1038/s41398-021-01221-0.
65. Woodward ND, Giraldo-Chica M, Rogers B, Cascio CJ. Thalamocortical dysconnectivity in autism spectrum disorder: An analysis of the Autism Brain Imaging Data Exchange. Biol Psychiatry Cogn Neurosci Neuroimaging. 2017;2:76–84. https://doi.org/10.1016/j.bpsc.2016.09.002.
66. Cho KIK, Shenton ME, Kubicki M, Jung WH, Lee TY, Yun J-Y, et al. Altered Thalamo-Cortical White Matter Connectivity: Probabilistic Tractography Study in Clinical-High Risk for Psychosis and First-Episode Psychosis. Schizophr Bull. 2016;42:723–31. https://doi.org/10.1093/schbul/sbv169.
67. Shin D, Kim CN, Ross J, Hennick KM, Wu S-R, Paranjape N, et al. Thalamocortical organoids enable in vitro modeling of 22q11.2 microdeletion associated with neuropsychiatric disorders. Cell Stem Cell. 2024;31:421–e4328. https://doi.org/10.1016/j.stem.2024.01.010.
68. Sheffield JM, Huang AS, Rogers BP, Giraldo-Chica M, Landman BA, Blackford JU, et al. Thalamocortical Anatomical Connectivity in Schizophrenia and Psychotic Bipolar Disorder. Schizophr Bull. 2020;46:1062–71. https://doi.org/10.1093/schbul/sbaa022.
69. Wang X, Li Y, Young DM, Ljungdahl A, Dema C, Rohani N et al. Spatiotemporal analysis of autism gene enrichment implicates cortex, thalamus, and hypothalamus. 2026. doi.org/10.64898/2026.05.14.724487
70. Aivazidis A, Memi F, Rademaker K, Koko M, Roberts K, Trinh A et al. A spatial transcriptomic atlas of autism-associated genes identifies convergence in the developing human thalamus. bioRxiv. 2025. https://doi.org/10.1101/2025.11.05.685843
71. Kim CN, Shin D, Wang A, Nowakowski TJ. Spatiotemporal molecular dynamics of the developing human thalamus. Science. 2023;382:eadf9941. https://doi.org/10.1126/science.adf9941.
72. Nano PR, Fazzari E, Azizad D, Martija A, Nguyen CV, Wang S, et al. Integrated analysis of molecular atlases unveils modules driving developmental cell subtype specification in the human cortex. Nat Neurosci. 2025;28:949–63. https://doi.org/10.1038/s41593-025-01933-2.
73. Szewczyk LM, Lipiec MA, Liszewska E, Meyza K, Urban-Ciecko J, Kondrakiewicz L, et al. Astrocytic β-catenin signaling via TCF7L2 regulates synapse development and social behavior. Mol Psychiatry. 2024;29:57–73. https://doi.org/10.1038/s41380-023-02281-y.
74. Kaplanis J, Samocha KE, Wiel L, Zhang Z, Arvai KJ, Eberhardt RY, et al. Evidence for 28 genetic disorders discovered by combining healthcare and research data. Nat Deciphering Dev Disorders Study. 2020;586:757–62. https://doi.org/10.1038/s41586-020-2832-5.
75. Dias C, Pfundt R, Kleefstra T, Shuurs-Hoeijmakers J, Boon EMJ, van Hagen JM, et al. De novo variants in TCF7L2 are associated with a syndromic neurodevelopmental disorder. Am J Med Genet A. 2021;185:2384–90. https://doi.org/10.1002/ajmg.a.62254.
76. Graham TA, Ferkey DM, Mao F, Kimelman D, Xu W. Tcf4 can specifically recognize beta-catenin using alternative conformations. Nat Struct Biol. 2001;8:1048–52. https://doi.org/10.1038/nsb718.
77. Bem J, Brożko N, Chakraborty C, Lipiec MA, Koziński K, Nagalski A, et al. Wnt/β-catenin signaling in brain development and mental disorders: keeping TCF7L2 in mind. FEBS Lett. 2019;593:1654–74. https://doi.org/10.1002/1873-3468.13502.
78. Hartig R, Wolf D, Schmeisser MJ, Kelsch W. Genetic influences of autism candidate genes on circuit wiring and olfactory decoding. Cell Tissue Res. 2021;383:581–95. https://doi.org/10.1007/s00441-020-03390-8.
79. Bakken TE, Jorstad NL, Hu Q, Lake BB, Tian W, Kalmbach BE, et al. Comparative cellular analysis of motor cortex in human, marmoset and mouse. Nature. 2021;598:111–9. https://doi.org/10.1038/s41586-021-03465-8.

Figures

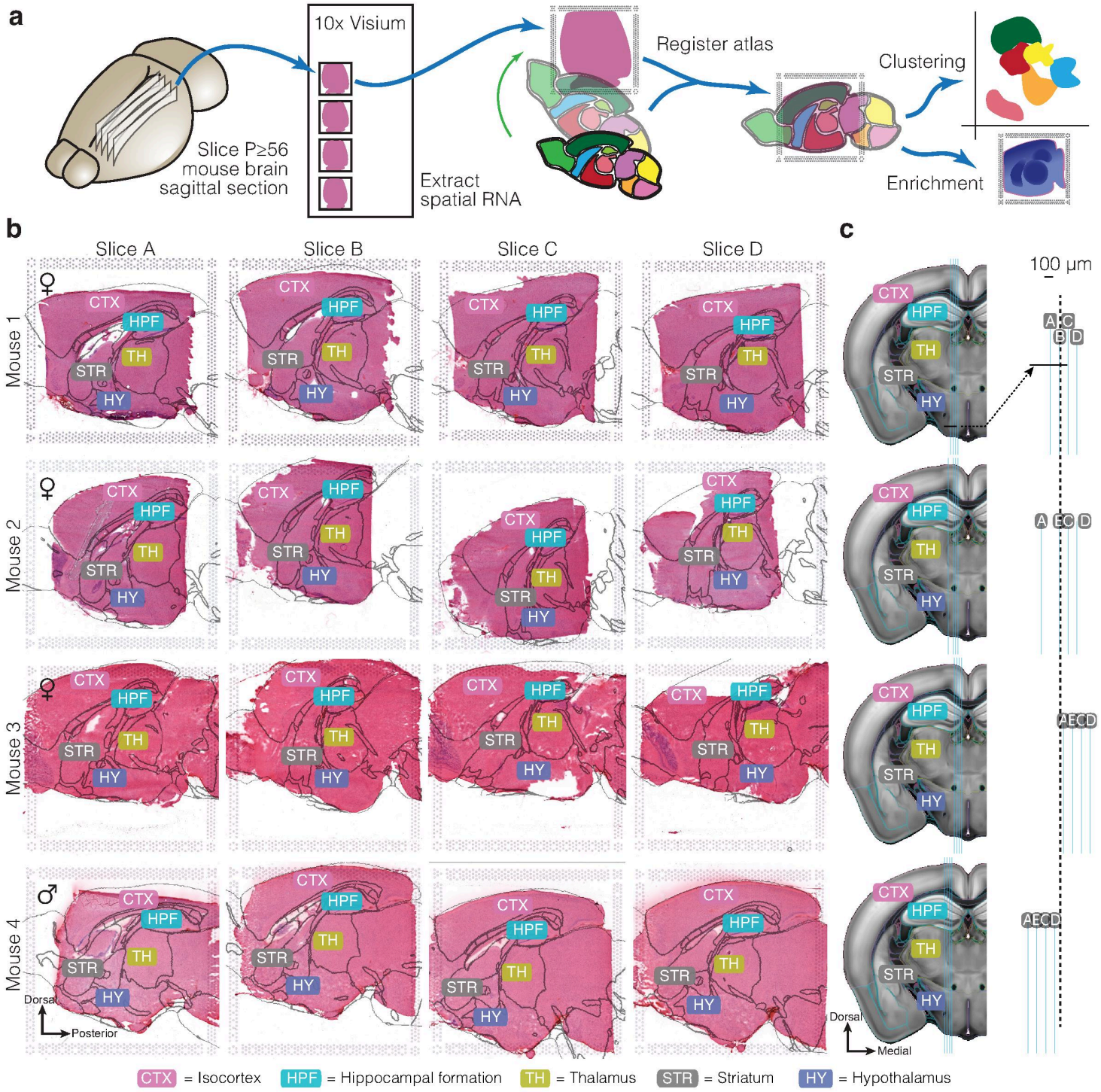


Figure 1

Experimental overview and anatomical orientation. a) Serial sagittal adult mouse brain sections were collected onto 10x Visium capture areas, which extracted RNA onto spatially barcoded spots. Tissue sections were segmented by atlas registration to identify clusters and enriched regions. **b)** Tissue samples from 16 sections in four mouse brains (rows) by serial slice (column) from lateral (left) to medial (right). Registered anatomical boundaries are overlaid and labelled. **c)** Corresponding slice positions represented in the coronal plane (left) and magnified for relative positioning (right).

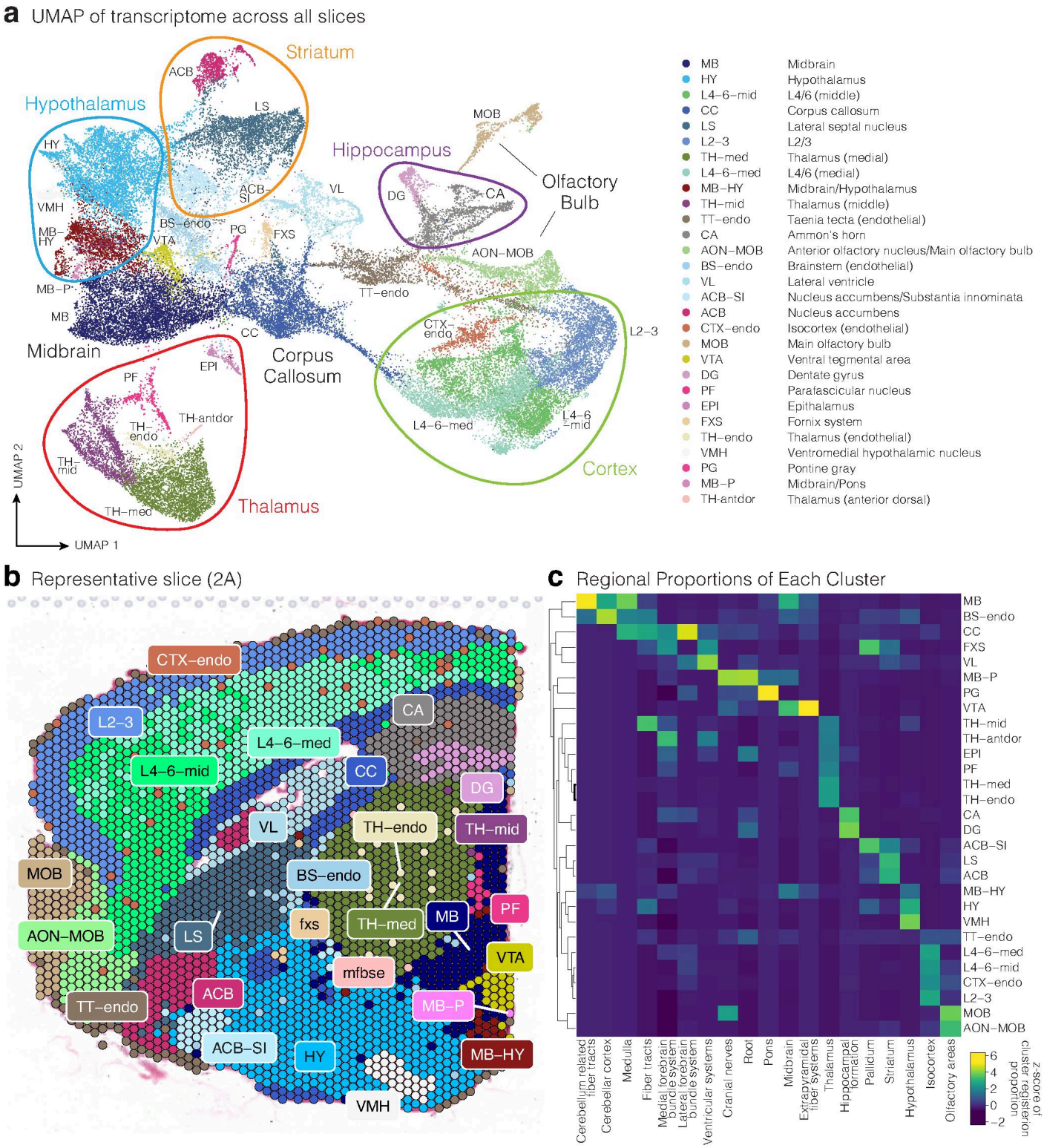


Figure 2

Cluster identification by anatomical alignment. a) UMAP of transcriptomic spots with spots colored by cluster identity. Clusters related to similar anatomical regions are circled and/or labelled. **b)** Spatial clustering shown on a representative slice (Slice 4A, see Fig. 1b). Most clusters aggregated into anatomically distinct groups, though some clusters span multiple regions (e.g., MB-HY) or are dispersed within a region (e.g., TH-endo). **c)** Heatmap of clusters (y-axis) by registered CCFv3 atlas region showing the standardized z-scores of the proportion of each cluster occupied by CCFv3 registered atlas regions (x-axis). Most clusters predominantly match only one region.

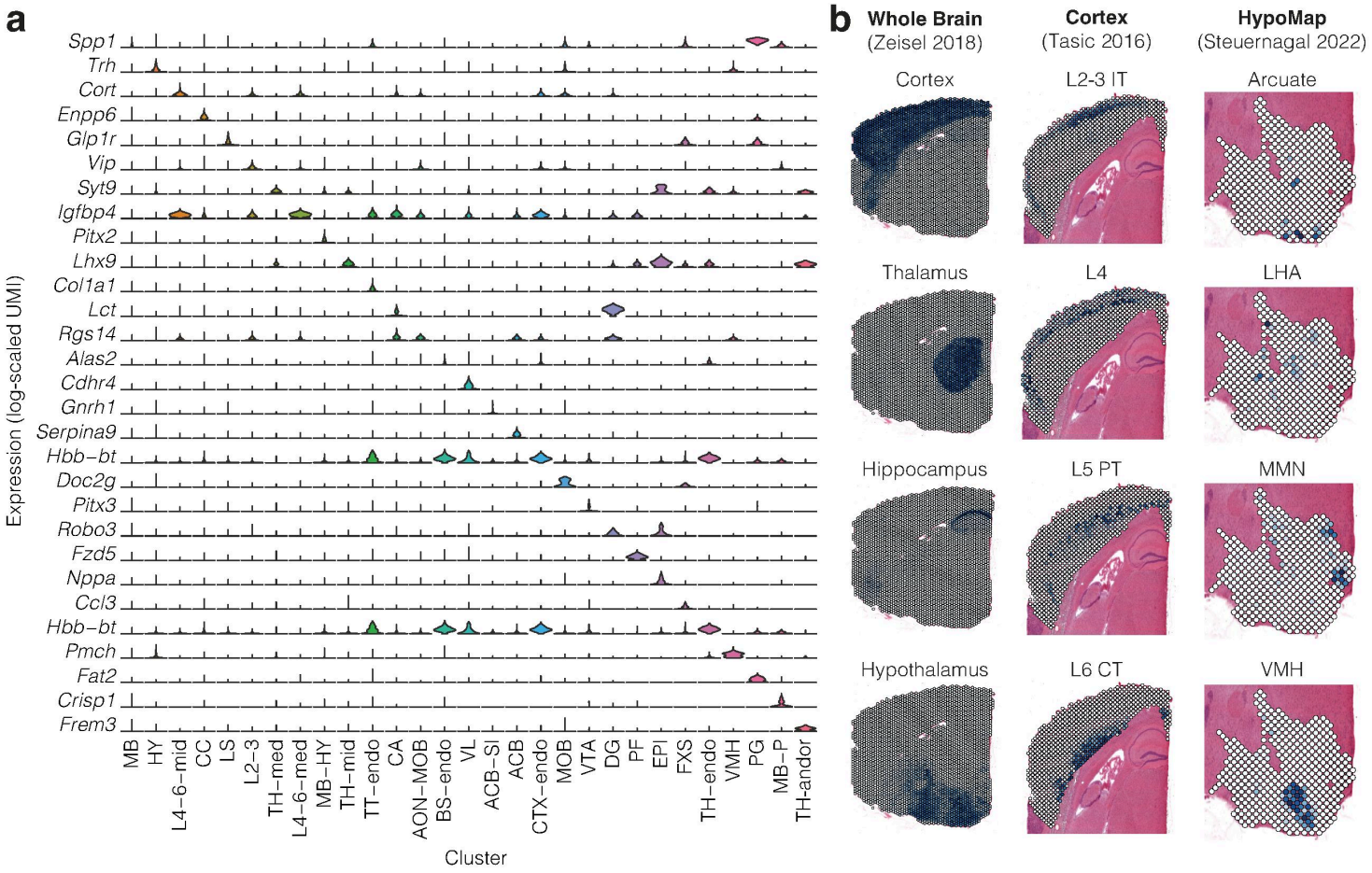


Figure 3

Transcriptomic clusters delineated by tissue-specific markers and deconvolution. a) Violin plots show expression levels by cluster of tissue-specific genes identified by cluster-specific differential expression and high tau values. **b)** Cluster anatomical assignments compared to external reference datasets of the whole brain [36], cortex [37], and hypothalamus [38]. For each anatomical spot, Seurat prediction scores are represented by shade, with dark blue representing higher prediction scores, normalized to each spatial slice.

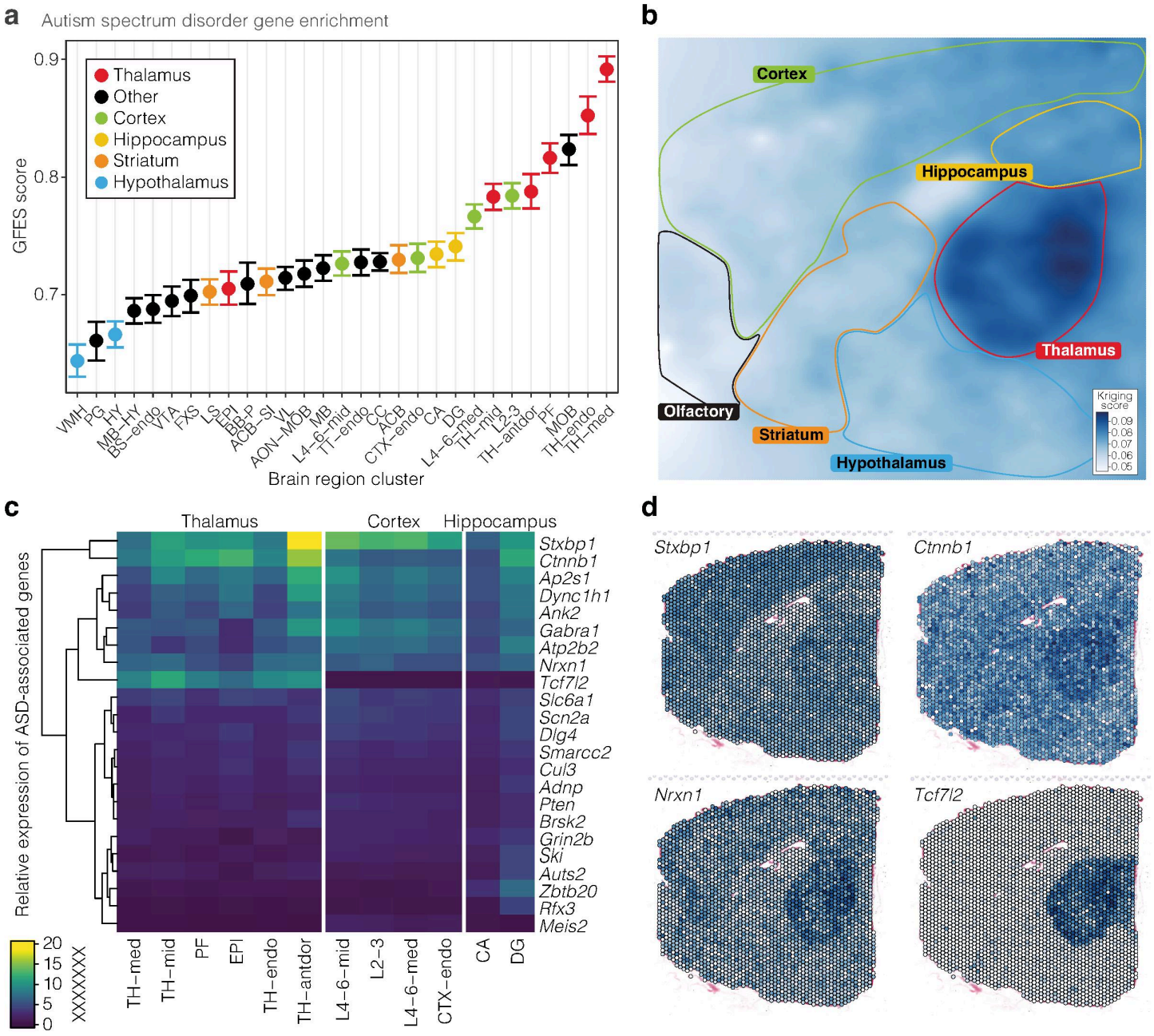


Figure 4

Enrichment of ASD-associated genes in the thalamus. a) GFES score based on 72 ASD-associated genes by regional cluster corrected for neuronal proportion. Error bars show 95% confidence intervals calculated by bootstrapping. **b)** Kriging plot showing spatial enrichment across all ASD genes assessed. **c)** Heatmap of gene expression across clusters for the most highly expressed genes. **d)** Spatial gene expression plots of specific ASD-associated genes.

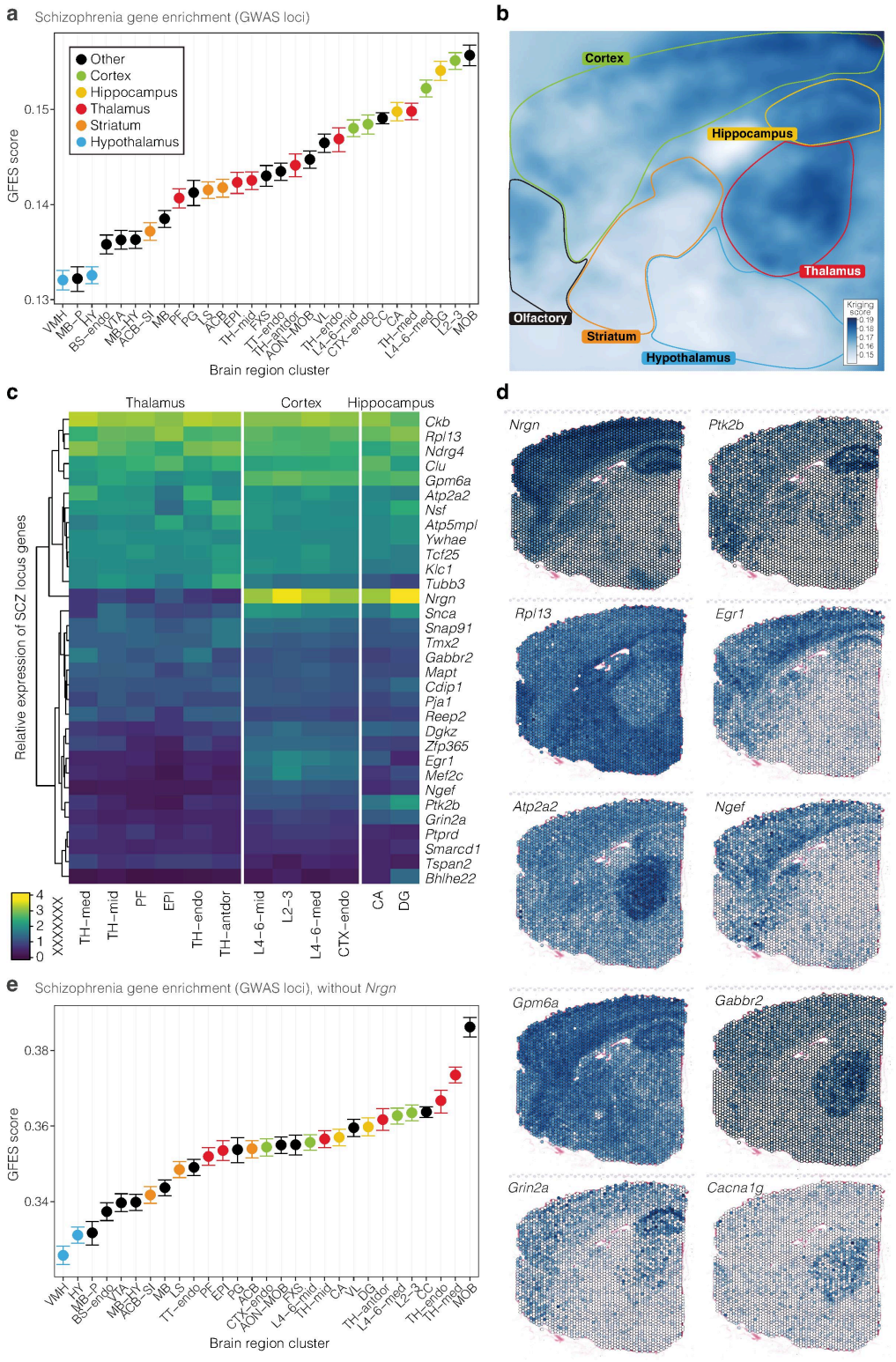


Figure 5

Enrichment of schizophrenia-associated genes in the olfactory bulb, cortex, thalamus, and hippocampus. a) GFES score based on 435 fine-mapped schizophrenia genes from GWAS by regional cluster corrected for neuronal proportion. Error bars show 95% confidence intervals calculated by bootstrapping. **b)** Kriging plot showing spatial enrichment across all schizophrenia genes assessed. **c)** Heatmap of gene expression across clusters for the most highly expressed genes. **d)**Spatial gene expression plots of key schizophrenia risk genes. **e)**Analysis in panel 'a' is repeated excluding the gene *Nrgn*.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementalFigures.pdf
- TableS1.xlsx
- TableS2.xlsx