

Supplementary Information

Contains:

Supplementary Materials and Methods, and References

Supplementary Figures and Legends

Materials and Methods

1 Cohort details

1.1 Bulk RNA sequencing cohort

The cohort used for the bulk RNA sequencing of this study consisted of 286 tissue samples from three frontal cortex subregions (dACC, dlPFC, mPFC) of which 138 unique donor IDs were shared across regions. The diagnosis breakdown was 190 samples with primary or secondary diagnosis of MDD and 96 neurotypical controls. Tissue acquisition and processing, diagnostic assessment protocols, ethical considerations, oversight, and governance of the LIBD PTSD brain collection have been previously described¹. Donation sites included the Office of the Chief Medical Examiner of the State of Maryland, Western Michigan University Homer Stryker MD School of Medicine, the Department of Pathology, University of North Dakota School of Medicine and Health Sciences, and the County of Santa Clara Medical Examiner-Coroner Office, and Gift of Life Michigan. Audiotaped and witnessed informed consent was obtained from the legal next-of-kin for every case according to the MD Department of Health IRB (protocol #12-24), and WCG IRB (protocol #20111080). The LIBD Autopsy telephone screening was done at time of donation with the legal next-of-kin and consisted of 39 items about the donor's medical, social, psychiatric, substance use, and treatment history. Retrospective clinical diagnostic reviews were conducted for every brain donor to include data from: autopsy reports, toxicology testing, forensic investigations, neuropathological examinations, telephone screening, and psychiatric/substance abuse treatment record reviews and/or supplemental family informant interviews (when possible and needed). Sex was determined based on autopsy, next of kin interviews, medical records, and genotype, which showed no inconsistencies. The variable was considered as a proxy of biological sex. All data were compiled and summarized in a detailed psychiatric narrative summary and was reviewed independently by two board-certified psychiatrists to determine lifetime psychiatric diagnoses according to Diagnostic and Statistical Manual of Mental Disorders (DSM-5). Non-psychiatric control donors were excluded if they had acute drug and alcohol intoxication/use at time of death. For every subject, we had access to extensive trauma-related information including age and type of trauma exposure, age of onset of psychiatric diagnosis, treatment history including hospitalizations, and all DSM-5 criteria and Research Domain Criteria relevant to their diagnoses, especially with regards to PTSD and MDD. All case subjects had primary or secondary diagnosis of MDD and approximately half of them were exposed to varying amounts of childhood trauma. Demographic information can be found in Supplementary Table 1A.

1.2 Single cell sequencing cohorts

Demographic information for the sample used for the single cell RNA sequencing (snRNA-seq) portion of this study can be found in Supplementary Table 6A.

1.2.1 McLean Cohort – dlPFC

A total of 46 samples were collected from the study by Chatzinakos et al.² to create the McLean Cohort. The cohort was comprised of 16 neurotypical controls and 30 samples with primary or secondary diagnosis of MDD. In this study, we used the raw FASTQ files provided by the above publication.

1.2.2 McLean Cohort – mPFC

Postmortem human brain tissue was obtained from the Lieber Institute for Brain Development (LIBD) repository, per section 1.1. A total of 49 subjects were selected and were split into 16 neurotypical controls and 33 samples with secondary diagnosis of MDD.

1.2.3 McGill Cohort - dlPFC

A total of 70 samples were collected from the studies by Nagy et al.³, and Maitra et al.⁴, to create the McGill Cohort. The cohort was comprised of 33 neurotypical controls and 37 samples with primary diagnosis of MDD. In this study we used the raw FASTQ files provided by the above publications.

1.2.4. MPI Cohort - oFC

A total of 39 samples were collected from the study by Fröhlich et al.⁵, to create the MPI cohort. The subset of the cohort we used comprised of 33 neurotypical controls and 6 samples with primary diagnosis of MDD. In this study we used the publicly available processed Seurat object from the above publication.

2 Bulk tissue processing and data generation

2.1 Data generation

All bulk RNA extractions and sequencing runs were performed at LIBD. RNA was extracted from postmortem homogenate brain tissue of gray matter dissected from the dACC, dlPFC and mPFC. Total RNA was extracted using the RNeasy kit (Qiagen) from approximately 100mg homogenate tissue for each region. RNA-seq was performed using Ribosomal RNA depletion library preparation with strand-specificity, using the TruSeq RNA Sample Preparation v2 kit. Library construction was performed using cDNA fragments from the total RNA that undergo end repair processing by T4 DNA polymerases, T4 polynucleotide kinases, and Klenow DNA polymerases. This process adds a single adenosine and then ligates cDNA fragments to paired-end adaptors using T4 DNA ligase. An index was incorporated into these Illumina adaptors to allow samples to be multiplexed in one lane of the sequencer flow cell. 150 bp paired-end libraries were run on the HiSeq 3000. The Illumina Real Time Analysis module was used to perform image analysis and base calling and the BCL Converter (CASAVA v1.8.2) was used to generate sequence reads. The sequencing depth was 40-60 million paired-end reads, yielding 80-120 million reads per sample.

2.2 Quality control and data preprocessing

The RNAseq-pipeline⁶ was used to preprocess the first batch of sequencing reads. The later batches of data were processed with SPEAQeasy, the continuation of the RNAseq-pipeline project. Both workflows align the sequencing reads to the human genome sequence GRCh38 using the HISAT2 aligner^{7,8}. The pipeline produced SummarizedExperiment R objects⁹ with quantification of all genomic features as annotated by GENCODE release 25. The pipeline also provided coverage data for unannotated expressed regions and all exon-exon junctions found in the RNA-seq read alignments. Custom R code¹⁰ was used to merge data across the batches processed with the RNAseq-pipeline and SPEAQeasy, forming a single set of SummarizedExperiment R objects for downstream analysis.

2.3 RNA data normalization

For the region-specific analysis, gene expression data were split per cortical subregion; mPFC, dlPFC and dACC. These brain region-specific datasets were then normalized separately. Within each subregion, gene count data were first filtered to remove lowly expressed features using the `filterByExpr()` function from the `edgeR`¹¹ package. For the across-region analysis, each of the subregion count objects were merged and analyzed as a single gene count dataset, and the filtering of the lowly expressed features was performed as described above. The diagnosis for each subject (i.e., NTC or MDD) was supplied as the group argument to this function while default values were used for the remaining arguments. Normalization factors were then calculated using the `calcNormFactors()` function from `edgeR` via the weighted trimmed mean of M-values (TMM) method. Lastly, the `voom()` function from the `limma`¹² package was used to normalize the expression data and transform the counts to log₂-counts per million for DE analysis.

2.4 Genetic/Ancestry PC calculation

Genotypes from three cohorts were subjected to standard single nucleotide polymorphism (SNP) quality control prior to imputation. The quality-controlled raw genotype data were phased with Eagle v2.4 and imputed using TOPMED imputation server¹³ with the TOPMED-r3 reference panel. Post-imputation, SNPs were filtered to retain those with INFO score > 0.8, and minor allele frequency (MAF) > 0.01. Imputed VCFs from all cohorts used in this study were merged on common SNPs using BCFtools¹⁴. Ancestry PCs were computed with the standardized global ancestry pipeline (https://github.com/nievergeltlab/global_ancestry). The pipeline uses SNPweights¹⁵ applied to 10,000 curated ancestry-informative markers anchored to a reference panel of 2,911 individuals from 71 populations across six continental groups ($k = 6$)¹⁶. Using the processed genotype dataset, we estimated ancestry PCs, which were included as covariates in downstream analyses to account for population stratification. Genotype data were unavailable for 2 samples from the McLean cohort; these samples were excluded from ancestry-informed analyses.

2.5 Cell type proportion calculation

Deconvolution was performed with the `ReferenceBasedDecomposition` function from the R package `BisqueRNA` version 1.0.5¹⁷, using the `use.overlap = FALSE` option. We used the single cell objects generated from each frontal cortex subregion (Section 7.6) as reference; we deconvolved the dlPFC with the dlPFC snRNA-Seq, the mPFC with the mPFC snRNA-seq and for the dACC we used the sum of the dACC-deconvolved with dlPFC and dACC-deconvolved with mPFC. The seven cell-types considered in the tissue deconvolution were excitatory neurons, inhibitory neurons, oligodendrocytes, OPC, astrocytes, immune cells and perivascular cells. Marker genes were selected by first filtering for genes common between the bulk data and the reference data, then calculating the ratio of the mean expression of each gene in the target cell-type over the highest mean expression of that gene in a non-target cell-type. The 25 genes with the highest ratios for each cell-type were selected as markers.

2.6 Variance Partition Analysis

To quantify the relative contributions of biological and technical sources of variation to genome-wide transcriptomic profiles, we applied `variancePartition` to the normalized bulk RNA-seq expression data. The analysis was performed on voom-normalized data. A linear mixed-effects model was specified as:

$\sim (1|gDx) + Age + (1|Sex) + PMI + RIN + mitoMapped + Ex + Inh + Microglia + Astro + Oligo + OPC + (1|Region) + (1|BrNum) + ancestryPC1 + ancestryPC2$

For each gene, the proportion of expression variance explained by each covariate was estimated using the iterative fitting procedure implemented in variancePartition. Variance fractions were then summarized across all expressed genes and visualized as violin plots ranked by mean variance explained per covariate.

2.7 Differential Expression Analysis

2.7.1 Region-Specific Differential Expression Analysis

Region-specific differential gene expression (DGE) analyses were performed using the limma package. For each region, we fit a linear model including [diagnosis \(gDx\)](#), [age at death](#), [sex](#), [postmortem interval \(PMI\)](#), [RNA integrity number \(RIN\)](#), [reads mapped to mitochondrial genes](#), [ancestry principal components \(anPC1-2\)](#), and [deconvoluted cell type proportions \(Ex, Inh, Oligo, Astro, OPC, Immune\)](#). Genes were considered differentially expressed at a false discovery rate (FDR) of 0.05 using the Benjamini-Hochberg method for multiple testing correction.

2.7.2 Cross-Region Differential Expression Analysis

To account for repeated measures across regions from the same donor, we used the dream function from the variancePartition package¹⁸. The model, [named FC-wide](#), included fixed effects for [diagnosis \(gDx\)](#), [age at death](#), [sex](#), [cortical subregion](#), [postmortem interval \(PMI\)](#), [RNA integrity number \(RIN\)](#), [reads mapped to mitochondrial genes](#), [ancestry principal components \(anPC1-2\)](#), and [deconvoluted cell type proportions \(Ex, Inh, Oligo, Astro, OPC, Immune\)](#), along with random intercept for donor identity (BrNum). Multiple testing correction was applied using the Benjamini-Hochberg method, with an FDR threshold of 0.05.

2.7.3 Secondary Analyses

2.7.3.1 Sub-analyses

To assess the robustness of MDD-associated transcriptional alterations across clinically relevant subgroups, we performed a series of stratified differential gene expression (DGE) analyses using the same cross-region modeling framework described above, while restricting the case group according to predefined clinical criteria. Specifically, analyses were performed for: (i) individuals with a primary diagnosis of MDD and no comorbid PTSD (MDD-only; 96 NTC vs 95 cases), (ii) individuals with MDD and comorbid PTSD (PTSD-comorbid; 96 NTC vs 95 cases), (iii) female-only (46 F NTC vs 91 F cases) and male-only (50 M NTC vs 99 M cases) MDD cases, and (iv) a low-substance (polypharmacy or opioid) MDD subgroup (96 NTC vs 90 cases). For each analysis, the relevant subgroup was compared against the same control cohort, and all covariates and random effects included in the primary cross-region model were retained. Concordance of transcriptional signatures across subgroup analyses was evaluated by comparing differential expression effect sizes and gene-level significance patterns with those observed in the primary MDD analysis.

2.7.3.2 Sensitivity analyses

To evaluate the potential influence of childhood trauma history, and substance use on the observed MDD-associated transcriptional signatures, we performed a series of sensitivity analyses using the cross-region modeling framework. Childhood trauma (CT) exposure was defined from manual review of physician narrative reports and modeled as an ordinal variable, ranging from 0 for neurotypical controls, 1 for individuals with diagnosis of MDD and no reported CT (n = 103), and

2 for individuals with a diagnosis of MDD and reported CT (n = 87). Additional analyses incorporated measures of substance use severity derived from clinical records, including polysubstance severity (0 = no mention, 1 = intoxication, 2 = abuse, 3 = dependence) and opioid severity (0 = no mention, 1 = intoxication, 2 = abuse, 3 = dependence). Differential expression results from each sensitivity model were compared with the primary MDD analysis to determine whether observed transcriptional signatures were robust to variation in trauma exposure and substance-use related factors.

2.8 Correlation Analyses

Pairwise correlations were calculated between log fold changes (logFC) derived from different models using the spearman method. We assessed concordance at multiple stringency levels, including genome-wide logFC correlations and analyses restricted to genes passing one-sided or two-sided FDR thresholds. These comparisons were used to assess the consistency of transcriptional signatures across models and phenotypes.

3 Summary-data Mendelian Randomization (SMR) Analysis

To identify putative causal genes underlying depression, we performed Summary-data-based Mendelian Randomization (SMR) analysis using the web-based SMR tool (<https://yanglab.westlake.edu.cn/smr-portal/>)¹⁹. We utilized the latest genome-wide association study (GWAS) summary statistics for depression released by the Psychiatric Genomics Consortium²⁰, which includes over 412k individuals diagnosed with depression. As the expression quantitative trait loci (eQTL) reference panel, we used the *BrainMeta* eQTL catalogue, which compiles high-quality cis-eQTL data from bulk and single-nucleus RNA-sequencing datasets across multiple human brain regions (n = 2,865).

Analyses were restricted to autosomal probes with at least one significant cis-eQTL ($p < 5 \times 10^{-8}$). SMR testing was performed using the default parameters. To distinguish pleiotropic effects from linkage, we applied the HEIDI (Heterogeneity in Dependent Instruments) test, excluding associations with HEIDI p-values < 0.05 , consistent with prior recommendations. Associations passing both SMR significance (FDR < 0.05 , Benjamini-Hochberg corrected) and HEIDI filtering were considered evidence of a shared genetic signal between gene expression and MDD risk, consistent with a shared causal variant model.

4 Transcriptome-Wide Association Studies (TWAS)

TWAS was performed using JEPEGMIX3_P (Gene-level Joint Analysis of Functional SNPs in Cosmopolitan Cohorts)^{21,22} with default settings. Gene expression weights²³ were integrated with GWAS summary statistics²⁰ to conduct transcriptome-wide gene-trait association analyses. JEPEGMIX3_P aggregates SNP-level GWAS signals into gene-level association statistics while accounting for linkage disequilibrium using an LD reference panel. Significant gene-trait associations (GTAs) were identified after false discovery rate (FDR) correction, using an FDR-adjusted $P < 0.05$ threshold.

4.1 Pairwise enrichment analysis of significant genes across methods

To quantify the degree of overlap between genes identified by DGE, SMR, and TWAS, we performed pairwise enrichment analyses using Fisher's exact tests and complementary permutation testing.

For each pairwise comparison, the background gene universe was defined as the intersection of genes that were testable in both analyses being compared, thereby accounting for differences in gene coverage across methods. Significant gene sets were defined using the significance thresholds established for each respective analysis. One-sided Fisher's exact tests (alternative = "greater") were used to test whether the observed overlap exceeded that expected by chance. Expected overlap was calculated as the product of the sizes of the two significant gene sets divided by the size of the pairwise gene universe, and fold enrichment was computed as the ratio of observed to expected overlap. Odds ratios and associated P values were obtained from Fisher's exact test.

To independently validate enrichment, empirical permutation testing was performed for each comparison. One significant gene set was held fixed while a random gene set of identical size was sampled without replacement from the corresponding pairwise background universe. This procedure was repeated 10,000 times to generate a null distribution of overlap sizes. Empirical one-sided P values were calculated as the proportion of permutations producing overlaps greater than or equal to the observed overlap. Benjamini–Hochberg correction was applied to Fisher's exact test P values across the three pairwise comparisons.

5 MAGMA enrichment analysis

To evaluate the genetic relevance of transcriptional alterations identified in bulk cortical tissue, we tested whether differentially expressed genes were enriched for common variant associations from various psychiatric GWAS, namely major depression²⁰, PTSD²⁴, schizophrenia²⁵, and bipolar disorder²⁶. Differentially expressed genes were identified separately in each cortical region (mPFC, dlPFC, and dACC) and in the FC-wide analysis. Genes passing a false discovery rate threshold of 5% (FDR < 0.05) were separated into upregulated and downregulated sets. Gene-set enrichment analyses were performed using MAGMA²⁷. SNP-to-gene mapping was performed using MAGMA's default window of 10kb and 1.5kb downstream of the transcription start and end sites based on NCBI 37 build. The gene location file and reference linkage disequilibrium structure were based on the 1000 Genomes Phase 3 European panel. Gene-level p-values were computed using the SNP-wise mean model. Competitive gene-set analyses were then performed to determine whether genes differentially expressed in cortical tissue showed stronger GWAS associations than expected relative to other genes in the genome. Multiple testing correction across gene sets was performed per disorder, using FDR and results were considered significant at an FDR-adjusted p value < 0.05.

6 Gene Set Enrichment and Spatial Annotation Analyses

6.1 Gene Ontology and Chromosomal Location Enrichment

Gene set enrichment analysis (GSEA) was performed using the fgsea package²⁸ on gene-level ranked lists from each differential expression analysis. Genes were ranked by the [t-statistic](#), and enrichment was assessed against the Gene Ontology Biological Process (GOBP; MSigDB category C5, subcategory BP) and positional (MSigDB category C1) gene sets from the Molecular Signatures Database. For the chromosomal enrichment analysis as assessed by the positional gene

sets, a more permissive threshold of $FDR < 0.1$ was used to account for the lower resolution and smaller gene set sizes of positional annotations. All other enrichments were interpreted at $FDR < 0.05$.

To aid biological interpretation of the GOBP enrichment results, we clustered all significant gene sets based on semantic similarity using the simplifyEnrichment package²⁹. We applied the Louvain clustering algorithm to identify clusters of functionally related GO terms. To determine the most central and representative “parent” nodes within each cluster, we computed spread centrality scores using the CePa package³⁰. This allowed us to annotate each cluster with a biologically meaningful summary label for improved interpretability.

6.2 Spatial registration of the DEGs in cortical layers

We assessed the enrichment of the MDD differentially expressed genes (DEGs; 5% FDR) from the FC-wide model and each region specific model of the main analysis into cortical layer marker genes as described in the spatiaLIBD package³¹. We used a broad clustering resolution shown to recapitulate the leptomeninges, the classic histological cortical layers 1-6 (L1-6) and white matter. This approach allowed us to evaluate whether stress-related transcriptional changes were preferentially localized to specific layers of the cortical column, offering spatial context for the observed gene expression patterns. To validate this approach we also generated reference gene sets from our cell type markers from the snRNA-seq analysis (Section 7.6).

7 Multi-Omics Factor Analysis (MOFA)

We ran MOFA as implemented in the R package MOFA2³² for multi-omic, multi-region integration of the shared samples across the three PFC subregions as well as proteomic and methylomic data from the mPFC extracted from Daskalakis et al.³³ ($N = 138$ samples, $N_{\text{views}} = 5$). Due to the unequal number of features across the omics layers, we retained features with expression variance at or above the 60th percentile across samples, resulting in a range of 4,702-9,164 features for each omic/region. The MOFA algorithm was run to generate 30 factors over slow convergence. We tested each factor for association with demographic covariates using a simple linear modeling strategy--treating factors as the dependent outcome variable, with diagnosis (NTCs or MDD), sex, psychosis and lifetime smoking variables incorporated as categorical predictors and age, methylation age, PMI, RIN, mitoMapped and ancestry PCs as numerical predictors (excluding potential interaction terms for simplicity) – using drop-one F tests to assess significance.

8 Single-Nucleus RNA-Sequencing Processing and Analysis

8.1 mPFC tissue processing and data generation

8.1.1 Brain tissue processing

Tissue acquisition and processing have been described previously³³. For the current study, microdissections of the mPFC were collected and stored at -80°C until processing. Approximately 70 mg of tissue were used for nuclear isolation and snRNA-seq.

8.1.2 Nuclear isolation and library preparation

Nuclei were isolated from frozen tissue blocks (~70 mg) as previously described³³. Briefly, tissue samples were Dounce homogenized in a homogenization buffer (0.25 M sucrose, 25 mM KCl, 5

mM MgCl₂, 20 mM Tricine-KOH pH 7.8, 1 mM DTT, 0.15 mM spermine, 0.5 mM spermidine, protease inhibitors, 0.04% BSA, 5% IGEPAL) and filtered through a 40 µm strainer. Samples were then mixed with an equal amount of working solution (50% iodixanol OptiPrep density gradient medium, 150 mM KCl, 30 mM MgCl₂, 120 mM Tricine-KOH, pH 7.8), loaded on top of a 30%/40% OptiPrep density gradient and spun at 10,000 x g at 4°C for 5 min. Samples were recovered from the OptiPrep interphase and resuspended in a PBS solution with 0.04% BSA to the optimal concentration recommended by 10x Genomics (700 to 1200 cells/ul). All steps were performed on ice and all buffers were supplemented with 1 U/µL RNAsin (Promega N2611), 1 U/µL Superscript (Thermo Fisher AM2696) and 0.04% BSA. Samples were processed in sets of 6 resulting in 9 batches total which were balanced for group and sex. Nuclei suspensions were counted in duplicates on a Luna II Automated Cell Counter to determine sample concentrations. Prior to Gel Bead-in-emulsion (GEM) generation, a multiplexing strategy was executed wherein individual samples from each batch were pooled together. This pooling was performed by combining an equivalent number of nuclei from each donor, targeting an expected recovery of 8,000 nuclei per subject. snRNA-seq libraries were subsequently generated using the 10x Genomics Chromium platform and the Single Cell 3' Reagent Kit v3.1 (Protocol Revision D). All downstream steps, including GEM generation and library construction, were performed in strict accordance with the manufacturer's standard protocols.

8.1.3 Sequencing

Following library preparation, sample quantity was assessed using a 2100 Bioanalyzer HS DNA chip (Agilent) and Qubit dsDNA HS Assay kit (Invitrogen™, Q32852). Samples were pooled at a final concentration of 10 nM and sequenced using NovaSeq 25B flow cells aiming for an average of 45k reads per nucleus. Sequencing was executed in a paired-end, dual-index format with a standard 10x Genomics Chromium v3.1 configuration consisting of 28 cycles for Read 1 (capturing the 16 bp cell barcode and 12 bp UMI), 10 cycles for the i7 sample index, 10 cycles for the i5 sample index, and 90 cycles for Read 2 (capturing the cDNA insert).

8.2 Raw Data Processing

Raw FASTQ files from snRNA-seq libraries from the McLean (mPFC and dlPFC) and the McGill cohort (dlPFC) were processed with Cellranger (v7.1.0) against the GRCh38 (2020-A) reference genome using default parameters. Raw gene-barcode matrices were imported into Seurat for downstream analysis. The MPI cohort dataset was used as a working Seurat object without further modifications⁵.

8.3 Genotype-based donor demultiplexing

Donor identities of the multiplexed McLean-mPFC cohort were assigned using a genotype-guided demultiplexing approach based on natural genetic variation. Cellranger processed alignments and filtered cellular barcodes were processed with cellsnr-lite³⁴(v1.2.3) to quantify allele-specific expression at known polymorphic sites. A donor genotype reference VCF generated from matched genomic DNA samples corresponding to each sequencing library was supplied to cellsnr-lite to restrict analysis to known donor variants. SNP quantification was performed using default counting procedures while requiring a minimum minor allele frequency of 0.05, a minimum read count of 10 per variant, a minimum aligned read length of 30 bp, and a minimum mapping quality score of 30. Unique molecular identifiers (UMIs) were used for allele counting.

Donor assignment was then performed using vireo³⁵ (v0.5.9), a Bayesian demultiplexing framework that leverages donor genotype information to probabilistically assign individual nuclei

to donors. Vireo was run using genotype calls from the matched donor reference panel and was executed both with ambient RNA estimation enabled and under a no-doublet model to evaluate assignment robustness. Posterior donor assignment probabilities and doublet classifications generated by vireo were used for downstream quality control and donor identification analyses.

8.4 Quality Control

Quality control was performed independently for each dIPFC and mPFC dataset.

8.4.1 mPFC dataset

Nuclei that were not assigned to any donor following Vireo demultiplexing were excluded from subsequent QC steps. Doublets were detected and excluded using scDbtFinder³⁶. Nuclei with less than 500 detected genes or more than 25,000 detected genes were excluded to remove low-complexity profiles and potential residual multiplets. Additionally, nuclei with greater than 5% of reads mapping to mitochondrial genes were removed. Thresholds were selected based on the empirical distributions of quality-control metrics across all nuclei, corresponding approximately to the 5th percentile for the lower cutoff and the 99th percentile for the upper and mitochondrial-content cutoffs. Samples with less than 1,000 assigned nuclei post demultiplexing and QC were dropped providing us a working dataset of 49 samples. Remaining nuclei passing all quality-control criteria were retained for downstream analyses.

8.4.2 dIPFC datasets

Quality control was performed independently for each dIPFC dataset. Nuclei were retained if they expressed more than 500 genes, with total UMI counts less than 30,000, and with <5% mitochondrial gene content. Doublets were excluded using scrublet³⁷ with the appropriate thresholds. Thresholds were selected based on the empirical distributions of quality-control metrics across all nuclei, corresponding approximately to the 5th percentile for the lower cutoff and the 99th percentile for the upper and mitochondrial-content cutoffs. Nuclei passing all quality-control criteria were retained for downstream analyses.

8.5 Data Integration and Clustering

Following QC, datasets were merged into a single Seurat object. Gene expression was log-normalized, and 3,000 highly variable features were selected using the “vst” method. To reduce data dimensionality and preserve structure for large-scale integration, cells were sketched based on their leverage score, selecting 80,000 representative cells for modeling.

Dimensionality reduction was performed via PCA, and the top 100 principal components were used for dataset integration using Harmony³⁸, regressing out technical and biological confounders including tissue, sequencing batch (orBatch), sample ID, and library version. Clustering was performed on the 100 Harmony embeddings using Louvain community detection at a resolution of 0.1, following neighbor graph construction with the FindNeighbors function.

8.6 Cell Type Annotation and Subclustering

Clusters were annotated based on canonical markers for major brain cell types, including excitatory neurons, inhibitory neurons, astrocytes, oligodendrocytes, oligodendrocyte precursor cells (OPCs), microglia, and perivascular cells.

Each major cell class was extracted and processed separately. The raw Seurat objects from each brain region, namely dIPFC, mPFC and oFC, were partitioned by cortical subregion. Each dataset was normalized using NormalizeData and highly variable genes were identified with

FindVariableFeatures. We regressed out technical covariates including percent.mt and nCount during scaling with ScaleData, followed by PCA dimensionality reduction (with appropriate number of PCs according to the size of the cluster, ranging from 10 to 50 PCs). Batch correction and integration were performed independently per region using the Harmony algorithm, regressing out donor effect and sex in oFC, donor effect, brain bank batch, chemistry version and sex in dlPFC and donor effect, and sex in mPFC. UMAP embeddings and shared nearest-neighbor graphs were computed using Harmony-reduced dimensions, followed by Louvain clustering at the optimal resolution. Preliminary cell-type labels were assigned via runAzimuth³⁹ where nuclei were annotated to their respective subclasses, and any nuclei belonging to irrelevant cell types were discarded. The oFC, mPFC and dlPFC were then merged and the combined object was re-normalized and scaled as above. Dimensionality reduction was performed via PCA (30-100 PCs), followed by Harmony-based batch correction across donors, sex and cortical subregion. UMAP projection and clustering were performed on Harmony-corrected embeddings and Azimuth annotations were re-applied post-integration for consistency. Clusters were annotated on the bases of unsupervised clustering and reference-based subclass predictions. Additional resolution of fine-grained subclusters was achieved using FindSubCluster whenever applicable. Marker genes for each subcluster were identified and used to validate the cell subclass assignment and further attribute functional annotation to each subcluster via overrepresentation analysis on GOBP terms. After independently performing initial annotation and filtering by cell type, we compiled a list of high-confidence nuclei for the downstream analysis, which now included 863,511 nuclei, clustered in 7 coarse classes, 11 biologically major classes and 29 subclasses.

8.7 Cell Type Composition Analysis

To assess differences in cellular composition across diagnostic groups, sex and brain regions, we utilized the *crumblr*⁴⁰ framework, which is designed for compositional analysis of cell-type abundance data. For each sample, the number of nuclei belonging to each of the 11 cell classes was quantified and aggregated into a sample-by-cell-type count matrix. Sample-level metadata included diagnosis, primary diagnosis, sex, age, brain region, sequencing batch, dataset version, and genetic ancestry principal components (anPC1 and anPC2). Raw cell-type counts were transformed using a centered log-ratio (CLR) transformation with a pseudocount of 0.5 to account for compositional constraints and zero counts. Variance attributable to biological and technical covariates was estimated using *variancePartition*⁴¹. The proportion of variance explained by diagnosis, sex, age, brain region, sequencing batch, dataset version, donor identity, and genetic ancestry was quantified using linear mixed-effects models. Differential cell-type abundance analyses were performed using the *dream*⁴² framework implemented in *variancePartition*. Separate mixed-effects models were fitted to evaluate associations with MDD diagnosis, primary diagnosis of MDD or PTSD, sex, and frontal cortex subregion. Models included relevant biological and technical covariates, including age, sex, sequencing batch, dataset version, brain region, and ancestry principal components, with donor identity modeled as a random effect where applicable. Statistical significance was assessed using empirical Bayes moderation, and Bonferroni correction was applied across tested cell classes to account for multiple comparisons.

8.8 Cell-type enrichment of GWAS genes

8.8.1 Stratified linkage disequilibrium score regression.

To evaluate whether common genetic risk for MDD and PTSD was enriched within cell-type-specific transcriptional programs, we performed stratified linkage disequilibrium score regression

(S-LDSC) ⁴³. Cell-type-specific markers (top 2,000 most significantly upregulated in each cell class) were derived from each of the 11 major cell classes to define cell-type-specific annotations. Gene-level annotations were converted to SNP-level annotations by assigning variants located within ± 10 kb of each annotated gene. Chromosome-specific annotation files were generated using the European reference panel from the 1000 Genomes Project Phase 3 dataset. Linkage disequilibrium (LD) scores were calculated for each annotation using LDSC with HapMap3 SNPs and a 1-cM LD window. Partitioned heritability analyses were performed using stratified LDSC while jointly modeling the baselineLD v2.3 annotation set to account for known functional genomic features. MDD GWAS ²⁰ and PTSD GWAS ²⁴ summary statistics were munged and analyzed using the LDSC framework, and enrichment statistics, regression coefficients, and associated standard errors were estimated for each cell-type annotation. Statistical significance was assessed using false discovery rate (FDR) correction across tested cell classes.

8.8.2 MAGMA

We performed cell type-specific gene-set analyses using MAGMA as in section 4. Cell class and subclass specific markers as calculated above were used and gene-set enrichment analysis were performed using MAGMA's gene property analysis mode. Multiple testing correction across gene sets was performed using FDR and results were considered significant at an FDR-adjusted p value < 0.05 .

8.8.3 Combining enrichment analyses

To integrate evidence across complementary enrichment methods, MAGMA and S-LDSC p values for each cell type were combined using Fisher's combined probability test (metap, v1.13). The resulting combined p values were Bonferroni corrected across all tested cell types.

8.9 Differential Gene Expression Analysis

Differential gene expression analysis was performed using linear models applied to pseudobulked expression profiles. For cell types lacking highly distinct substructure, i.e., OPCs, T cells, macrophages and endothelial cells, we used fixed-effect linear models. For all other cell types with pronounced heterogeneity, comprising multiple transcriptionally distinct subclasses (excitatory and inhibitory neurons, astrocytes, oligodendrocytes, microglia, mural cells, and fibroblasts), we applied linear mixed-effects models with a random intercept by donor to account for the nested data structure.

Pseudobulk expression profiles were generated by aggregating raw RNA counts across cells from the same individual and cluster, separately for each cell class and subclass. These summed counts were converted to counts per million (cpm), and lowly expressed genes were filtered either using the filterByExpr function or by applying manual thresholds tailored to each subpopulation depending on downstream analysis. To account for differences in library size across samples, normalization factors were computed using the trimmed mean of M-values (TMM) method. These factors were incorporated into the voom transformation (limma package), which models the mean-variance relationship of the log-transformed counts and generates observation-level precision weights. The resulting voom-normalized expression values (log₂ scale) were used for linear modeling.

For cell classes containing multiple subclasses or clusters, we performed analyses using linear mixed-effects models. Pseudobulk expression profiles were generated separately for each constituent subclass or cluster within a parent cell class, and these profiles were jointly analyzed within a repeated-measures framework. Models included fixed effects for diagnosis, subclass or

cluster identity, age, sex, frontal cortex subregion, chemistry version, and ancestry PCs. Donor identity was modeled as a random intercept to account for the non-independence of multiple subclass- or cluster-level observations originating from the same individual. This framework enabled estimation of an overall MDD-associated transcriptional effect across all constituent subclasses or clusters while accounting for differences in cellular composition and representation within each parent cell class. Gene expression counts were analyzed using the dream framework with empirical Bayes moderation. Multiple testing correction was performed using the Benjamini–Hochberg procedure, and genes with an FDR < 0.05 were considered significant.

For OPCs, T cells, macrophages and endothelial cells, differential expression was performed using a similar mixed model framework accounting for the random effect of the donor identity but without including any subclass information, since these cell types were not meaningfully subclustered.

Similarly to the bulk analysis, we conducted sub-analyses for: (i) individuals with a primary diagnosis of MDD and no comorbid PTSD (MDD-only; 98 NTC vs 58 cases), (ii) individuals with MDD and comorbid PTSD (PTSD-comorbid; 98 NTC vs 45 cases), and (iii) female-only (44 F NTC vs 49 F cases) and male-only (54 M NTC vs 54 M cases) MDD cases. For each analysis, the relevant subgroup was compared against the same control cohort, and all covariates and random effects included in the primary model per cell type were retained. Further, we performed a sub-analysis accounting for childhood trauma (CT) exposure, wherever available, i.e., the McLean and MPI cohorts (n = 28), using the same approach as the primary model per cell type but using the CT exposure as predictor.

8.10 Single-cell SMR

To identify genes whose genetically regulated expression was associated with MDD, we performed SMR analyses – as in section 3 -- using cell type-specific expression cis-quantitative trait locus (eQTL) reference datasets⁴⁴ from 983 individuals and MDD GWAS summary statistics²⁰. To distinguish pleiotropic associations from linkage disequilibrium between distinct causal variants, we additionally applied the heterogeneity in dependent instruments (HEIDI) test. Genes were considered significant SMR candidates if they passed false-discovery-rate (FDR) correction (FDR < 0.05) and showed no evidence of heterogeneity (PHEIDI > 0.05).

8.10.1 Pairwise enrichment analysis of significant genes across methods

Pairwise enrichment and permutation analysis for evaluating the degree of overlap between single cell DGE analysis and scSMR was performed as in Section 4.1

8.11 Comparison of differential expression signatures across (sub-)analyses

To evaluate the consistency of disease-associated transcriptional signatures across the main analysis and the sub-analyses, we compared genome-wide differential expression results patterns. For each cell type, genome-wide differential expression signatures were represented by gene-level log₂FC from the corresponding differential expression models. Spearman rank correlations were calculated between differential expression signatures for each pairwise separately for each cell type.

8.12 Correlation of bulk and single-nucleus differential expression signatures

To assess the concordance between bulk tissue and cell type-specific transcriptional alterations, we compared differential expression signatures from the main bulk RNA-seq analysis (MDD) with pseudobulk differential expression results generated for each major cell class. Differential

expression signatures were derived from partially overlapping cohorts, with approximately 60 individuals represented in both datasets. Genes detected in both analyses were retained for subsequent comparisons. For each comparison, we quantified directional differential expression (DDE) scores as the signed significance of differential expression, calculated as $-\log_{10}P \times \text{sign}(\log_2FC)$. Spearman correlations were computed between bulk and cell type DDE scores across all shared genes to evaluate global concordance of disease-associated transcriptional changes. Rank-rank hypergeometric overlap (RRHO2)⁴⁵ analysis was performed using signed DDE scores to identify significant overlap between bulk and cell type transcriptional signatures without imposing arbitrary significance thresholds. Genes were ranked according to their DDE scores in each dataset and overlap enrichment across ranked gene lists was visualized using RRHO2.

8.13 Gene set enrichment analysis

GSEA was performed using the fgsea package on gene-level ranked lists from the differential expression analysis per cell type. Genes were ranked by the t-statistic, and enrichment was assessed against the GO:BP. Enrichments were interpreted positive when the normalized enrichment score was positive, and vice versa, at an FDR < 0.05 level.

8.14 CellChat analysis

To infer and compare cell-cell communication between control and MDD, we used the CellChat⁴⁶ package. [The addition of the new samples allowed us in parallel to investigate the cell-cell communication differences in each primary diagnosis, i.e., MDD-only and PTSD-comorbid as well.](#) We subset the normalized counts into control and each diagnostic group and processed them in parallel as independent CellChat objects. For each of the datasets, we used the preprocessing functions ‘identifyOverExpressedGenes’, ‘identifyOverExpressedInteractions’, and ‘projectData’ with standard parameters. We utilized all available ligand-receptor interactions from CellChatDB. Communication probabilities between cell types were computed using the computeCommunProb() function with the triMean method and adjusting for population size. Low-confidence interactions were filtered out using a minimum cell count threshold (min.cells = 10). Communication probabilities were also inferred at the pathway level with computeCommunProbPathway(), and signaling networks were aggregated using aggregateNet(). Finally, the function ‘netAnalysis_computeCentrality’ was applied to identify dominant senders and receivers in the intercellular communication networks created. At this point, we combined the control and respective diagnosis group CellChat objects for downstream comparison analysis. Differential number and weight of interactions were calculated. To detect altered signaling pathways with distinct topological and quantitative features, we applied rankNet, which ranks signaling networks based on interaction count or strength.

We quantified total cell-cell communication by extracting interaction count matrices from CellChat for Control and MDD/MDD-only/PTSD-comorbid conditions. Global differences in total communication were first evaluated using a Chi-squared test on the raw interaction count matrices. To assess statistical significance of observed differences in total signaling, as well as group-specific differences (e.g., neuron-, glia-, immune-, and perivascular-involved interactions), we performed permutation tests (n = 10,000). Interaction matrices were flattened to long format, and interactions involving each cell group were identified based on sender or receiver identity. Observed differences in total interaction counts between conditions were compared to a null distribution generated by randomly permuting condition labels. Two-tailed empirical p-values

were calculated as the proportion of permuted differences greater than or equal to the absolute observed difference.

References

1. Mighdoll, M. I. *et al.* Implementation and clinical characteristics of a posttraumatic stress disorder brain collection. *J Neurosci Res* **96**, 16–20 (2018).
2. Chatzinakos, C. *et al.* Single-Nucleus Transcriptome Profiling of Dorsolateral Prefrontal Cortex: Mechanistic Roles for Neuronal Gene Expression, Including the 17q21.31 Locus, in PTSD Stress Response. *AJP appli.ajp.20220478* (2023) doi:10.1176/appi.ajp.20220478.
3. Nagy, C. *et al.* Single-nucleus transcriptomics of the prefrontal cortex in major depressive disorder implicates oligodendrocyte precursor cells and excitatory neurons. *Nature Neuroscience* 1–11 (2020) doi:10.1038/s41593-020-0621-y.
4. Maitra, M. *et al.* Cell type specific transcriptomic differences in depression show similar patterns between males and females but implicate distinct cell types and genes. *Nat Commun* **14**, 2912 (2023).
5. Fröhlich, A. S. *et al.* Single-nucleus transcriptomic profiling of human orbitofrontal cortex reveals convergent effects of aging and psychiatric disease. *Nat Neurosci* **27**, 2021–2032 (2024).
6. Huber, W. *et al.* Orchestrating high-throughput genomic analysis with Bioconductor. *Nat Methods* **12**, 115–121 (2015).
7. Zhang, Y., Park, C., Bennett, C., Thornton, M. & Kim, D. Rapid and accurate alignment of nucleotide conversion sequencing reads with HISAT-3N. *Genome Res.* **31**, 1290–1295 (2021).
8. Eagles, N. J. *et al.* SPEAQeasy: a scalable pipeline for expression analysis and quantification for R/bioconductor-powered RNA-seq analyses. *BMC Bioinformatics* **22**, 224 (2021).

9. Morgan, M., Obenchain, V., Hester, J. & Boutej, H. Summarized Experiment: A container (S4 class) for matrix-like assays. *Bioconductor* (2024).
10. Eagles, Nick, Collado-Torres, Leonardo & Huuki-Myers, Louise. LieberInstitute/joelPTSD_year2: v0_methods_with_documentation. (2022).
11. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
12. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research* **43**, e47 (2015).
13. Das, S. *et al.* Next-generation genotype imputation service and methods. *Nat Genet* **48**, 1284–1287 (2016).
14. Danecek, P. *et al.* Twelve years of SAMtools and BCFtools. *Gigascience* **10**, giab008 (2021).
15. Chen, C.-Y. *et al.* Improved ancestry inference using weights from external reference panels. *Bioinformatics* **29**, 1399–1406 (2013).
16. Kouakou, M. R. *et al.* Sites of active gene regulation in the prenatal frontal cortex and their role in neuropsychiatric disorders. *Am J Med Genet B Neuropsychiatr Genet* **186**, 376–388 (2021).
17. Jew, B. *et al.* Accurate estimation of cell composition in bulk expression through robust integration of single-cell information. *Nat Commun* **11**, 1971 (2020).
18. Hoffman, G. E. & Schadt, E. E. variancePartition: interpreting drivers of variation in complex gene expression studies. *BMC Bioinformatics* **17**, 483 (2016).

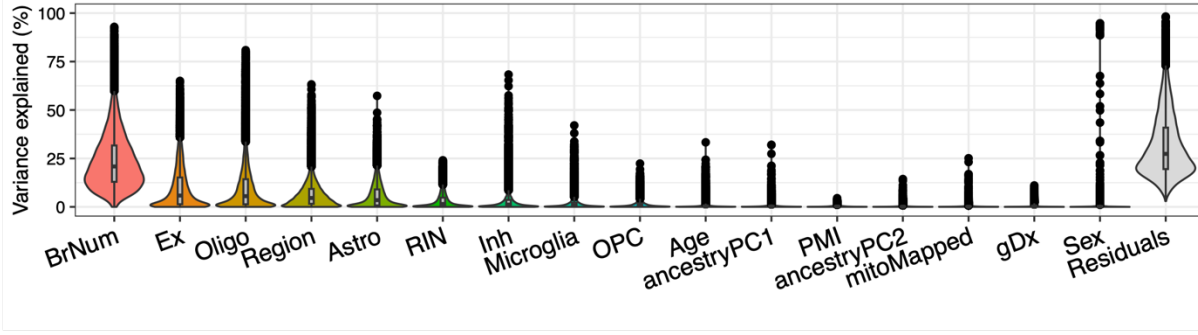
19. Qi, T. *et al.* Genetic control of RNA splicing and its distinct role in complex trait variation. *Nat Genet* **54**, 1355–1363 (2022).
20. Adams, M. J. *et al.* Trans-ancestry genome-wide study of depression identifies 697 associations implicating cell types and pharmacotherapies. *Cell* **188**, 640-652.e9 (2025).
21. Chatzinakos, C. *et al.* TWAS pathway method greatly enhances the number of leads for uncovering the molecular underpinnings of psychiatric disorders. *Am J Med Genet B Neuropsychiatr Genet* **183**, 454–463 (2020).
22. Chatzinakos, C. *et al.* JEPEGMIX2: improved gene-level joint analysis of eQTLs in cosmopolitan cohorts. *Bioinformatics* **34**, 286–288 (2018).
23. Bhattacharya, A. *et al.* Isoform-level transcriptome-wide association uncovers genetic risk mechanisms for neuropsychiatric disorders in the human brain. *Nat Genet* **55**, 2117–2128 (2023).
24. Nievergelt, C. M. *et al.* Genome-wide association analyses identify 95 risk loci and provide insights into the neurobiology of posttraumatic stress disorder. *Nat Genet* **56**, 792–808 (2024).
25. Trubetskoy, V. *et al.* Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature* **604**, 502–508 (2022).
26. O’Connell, K. S. *et al.* Genomics yields biological and phenotypic insights into bipolar disorder. *Nature* **639**, 968–975 (2025).
27. Leeuw, C. A. de, Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: Generalized Gene-Set Analysis of GWAS Data. *PLOS Computational Biology* **11**, e1004219 (2015).
28. Korotkevich, G. *et al.* Fast gene set enrichment analysis. 060012 Preprint at <https://doi.org/10.1101/060012> (2021).

29. Gu, Z. & Hübschmann, D. Simplify enrichment: A bioconductor package for clustering and visualizing functional enrichment results. *Genomics, Proteomics & Bioinformatics* <https://doi.org/10.1016/j.gpb.2022.04.008> (2022) doi:10.1016/j.gpb.2022.04.008.
30. Gu, Z. & Wang, J. CePa: an R package for finding significant pathways weighted by multiple network centralities. *Bioinformatics* **29**, 658–660 (2013).
31. Pardo, B. *et al.* spatialLIBD: an R/Bioconductor package to visualize spatially-resolved transcriptomics data. *BMC Genomics* **23**, 434 (2022).
32. Argelaguet, R. *et al.* Multi-Omics Factor Analysis—a framework for unsupervised integration of multi-omics data sets. *Molecular Systems Biology* **14**, e8124 (2018).
33. Daskalakis, N. P. *et al.* Systems biology dissection of PTSD and MDD across brain regions, cell types, and blood. *Science* **384**, eadh3707 (2024).
34. Huang, X. & Huang, Y. Cellsnp-lite: an efficient tool for genotyping single cells. *Bioinformatics* **37**, 4569–4571 (2021).
35. Huang, Y., McCarthy, D. J. & Stegle, O. Vireo: Bayesian demultiplexing of pooled single-cell RNA-seq data without genotype reference. *Genome Biol* **20**, 273 (2019).
36. Germain, P.-L., Lun, A., Meixide, C. G., Macnair, W. & Robinson, M. D. Doublet identification in single-cell sequencing data using *scDblFinder*. Preprint at <https://doi.org/10.12688/f1000research.73600.2> (2022).
37. Wolock, S. L., Lopez, R. & Klein, A. M. Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. *Cell Syst* **8**, 281-291.e9 (2019).
38. Korsunsky, I. *et al.* Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods* **16**, 1289–1296 (2019).

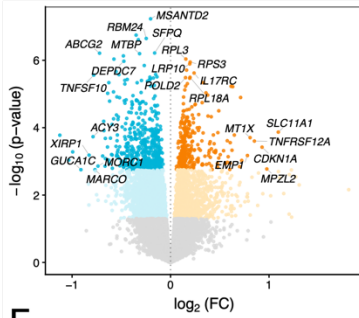
39. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e29 (2021).
40. Hoffman, G. E. & Roussos, P. Fast, flexible analysis of differences in cellular composition with crumblr. 2025.01.29.635498 Preprint at <https://doi.org/10.1101/2025.01.29.635498> (2025).
41. Hoffman, G. E. & Schadt, E. E. variancePartition: interpreting drivers of variation in complex gene expression studies. *BMC Bioinformatics* **17**, 483 (2016).
42. Hoffman, G. E. & Roussos, P. Dream: powerful differential expression analysis for repeated measures designs. *Bioinformatics* **37**, 192–201 (2021).
43. Finucane, H. K. *et al.* Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat Genet* **47**, 1228–1235 (2015).
44. Jang, B. *et al.* A meta-analysis of single-nucleus expression quantitative trait loci linking genetic risk to brain disorders. *Nat Genet* **58**, 737–747 (2026).
45. Cahill, K. M., Huo, Z., Tseng, G. C., Logan, R. W. & Seney, M. L. Improved identification of concordant and discordant gene expression signatures using an updated rank-rank hypergeometric overlap approach. *Sci Rep* **8**, 9588 (2018).
46. Jin, S., Plikus, M. V. & Nie, Q. CellChat for systematic analysis of cell–cell communication from single-cell transcriptomics. *Nat Protoc* **20**, 180–219 (2025).

Supplementary Figures

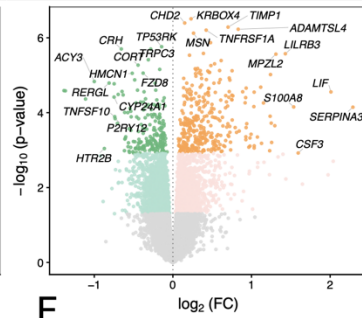
A



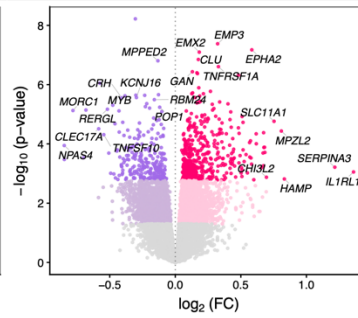
B



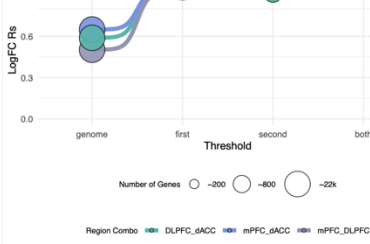
C



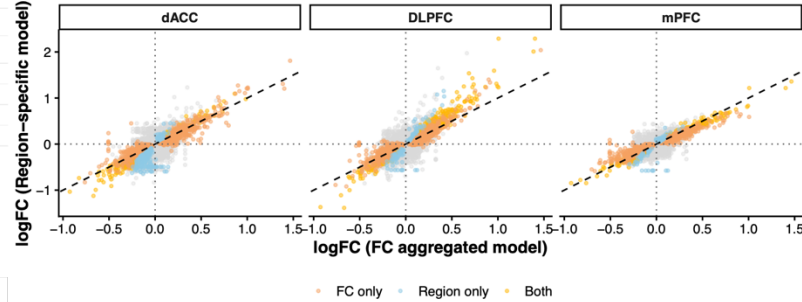
D



E

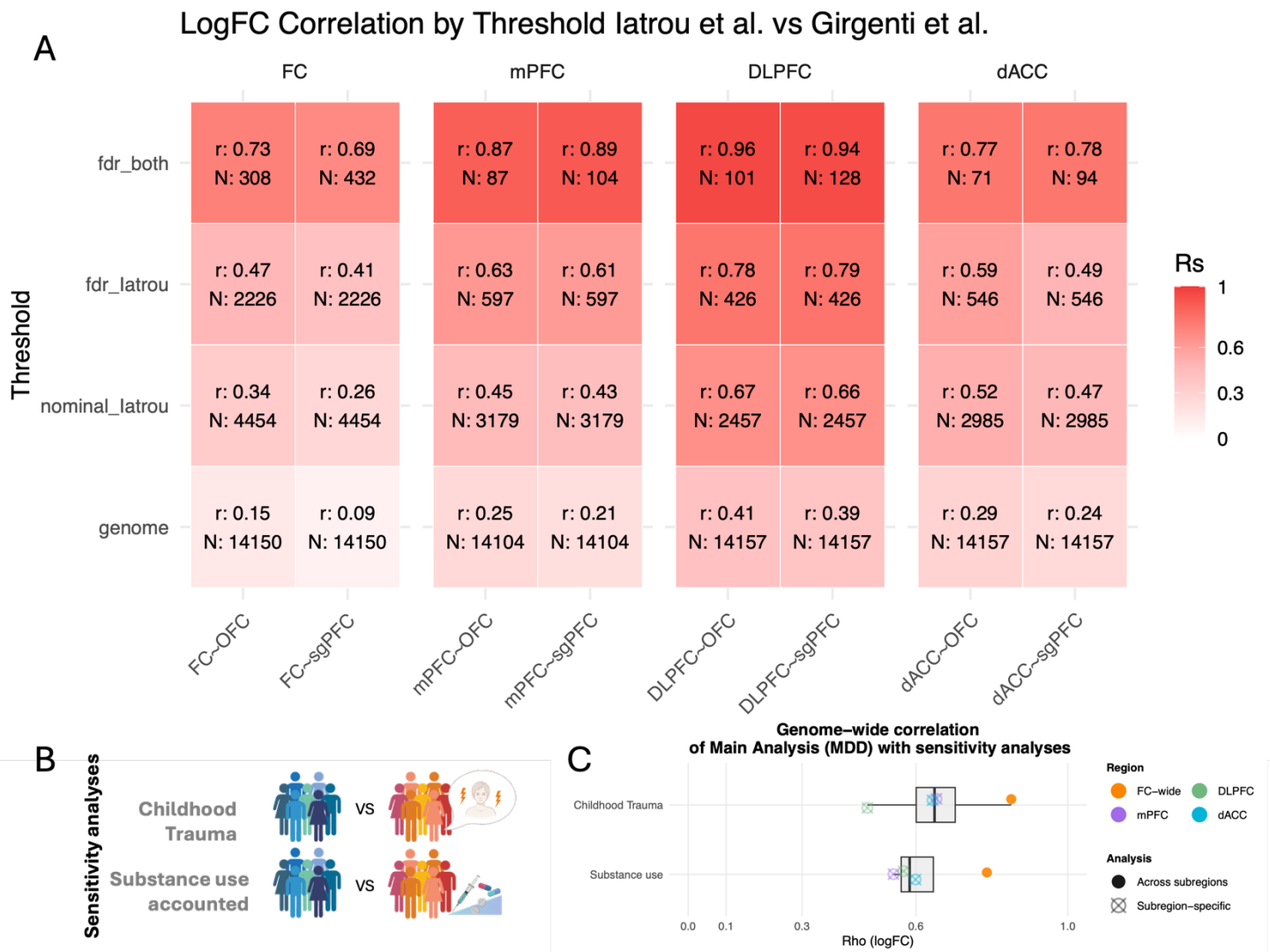


F

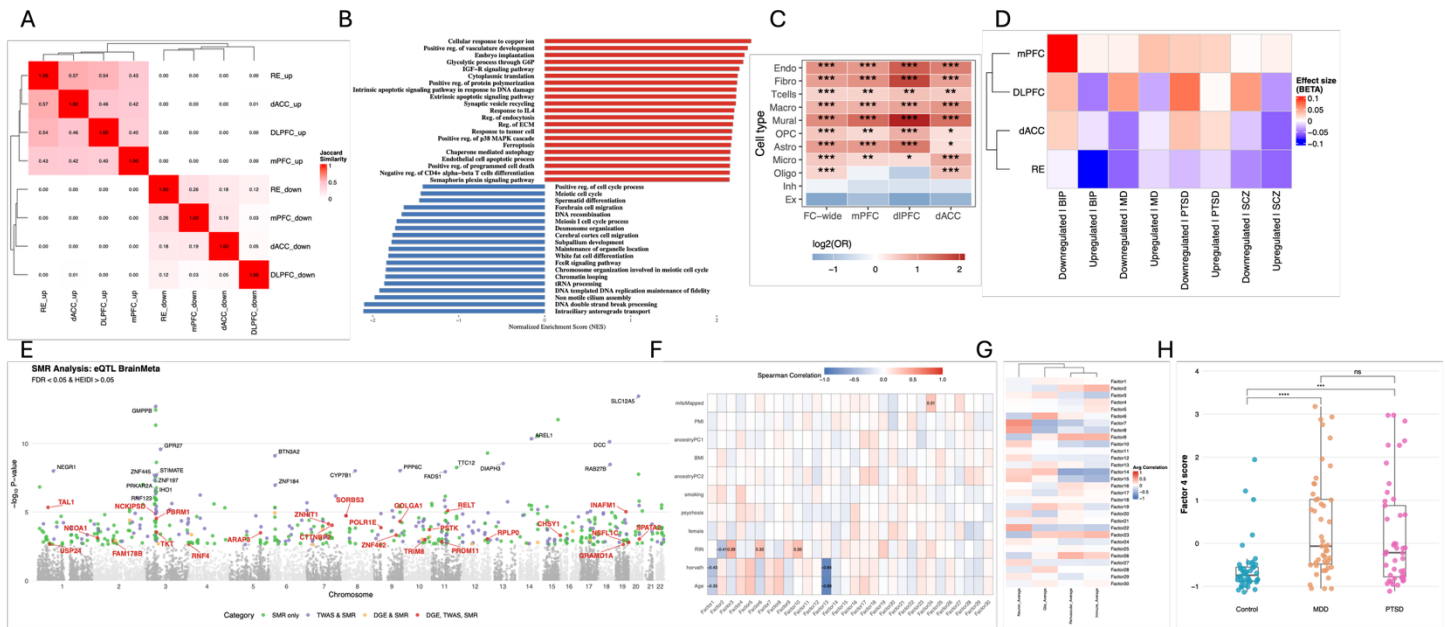


Supplementary Figure 1. Regional decomposition of MDD-associated transcriptional signatures.

(A) Variance partition analysis quantifying the proportion of gene expression variance explained by each covariate included in the FC-aggregated linear mixed model (diagnosis, sex, age, region, PMI, RIN, mitochondrial read fraction, ancestry PCs, and inferred cell-type proportions), with donor modeled as a random effect. Each point represents one gene; the bars show the median variance explained across genes per covariate. Donor identity accounted for the largest proportion of variance (average 23.8%), while region explained an average of 6.6%, confirming that inter-individual variability dominates the transcriptome and that the FC-aggregated framework effectively partitions this structure. (B–D) Volcano plots of differential gene expression from the mPFC- (B), dLPFC- (C), and dACC-specific (D) linear models. Each point represents one gene; the x-axis shows $\log_2$ fold change and the y-axis shows $-\log_{10}(\text{FDR-adjusted p-value})$. Genes passing $\text{FDR} < 0.05$ are highlighted (upregulated: red; downregulated: blue); the top 10 most significant genes in each direction are labeled. (E) Bump plot displaying Spearman correlations between each pair of region-specific analyses at three levels of gene selection: genome-wide (all expressed genes), FDR-significant genes from the FC-aggregated model, and genes reaching FDR significance in both region-specific models being compared. Point size reflects the number of shared DEGs contributing to each correlation estimate. Genome-wide correlations between individual subregions were moderate ($\rho = 0.50\text{--}0.65$), increasing substantially when restricted to FDR-significant genes, with near-identical effect size rankings among replicated genes ($\rho = 0.98\text{--}0.99$), indicating that the most robust MDD-associated signals are highly conserved across subregions. (F) Scatterplots of $\log_2$ fold change values for genes detected in both the FC-aggregated model and each region-specific analysis (mPFC, dLPFC, dACC). Points are colored by significance category: FC-aggregated significant only (yellow), region-specific significant only (blue), and significant in both (orange). The predominance of orange points along the diagonal confirms strong directional concordance among robustly dysregulated genes, while yellow and blue points reflect genes whose detection is driven by the increased power of the aggregated framework or region-specific biological variation, respectively.

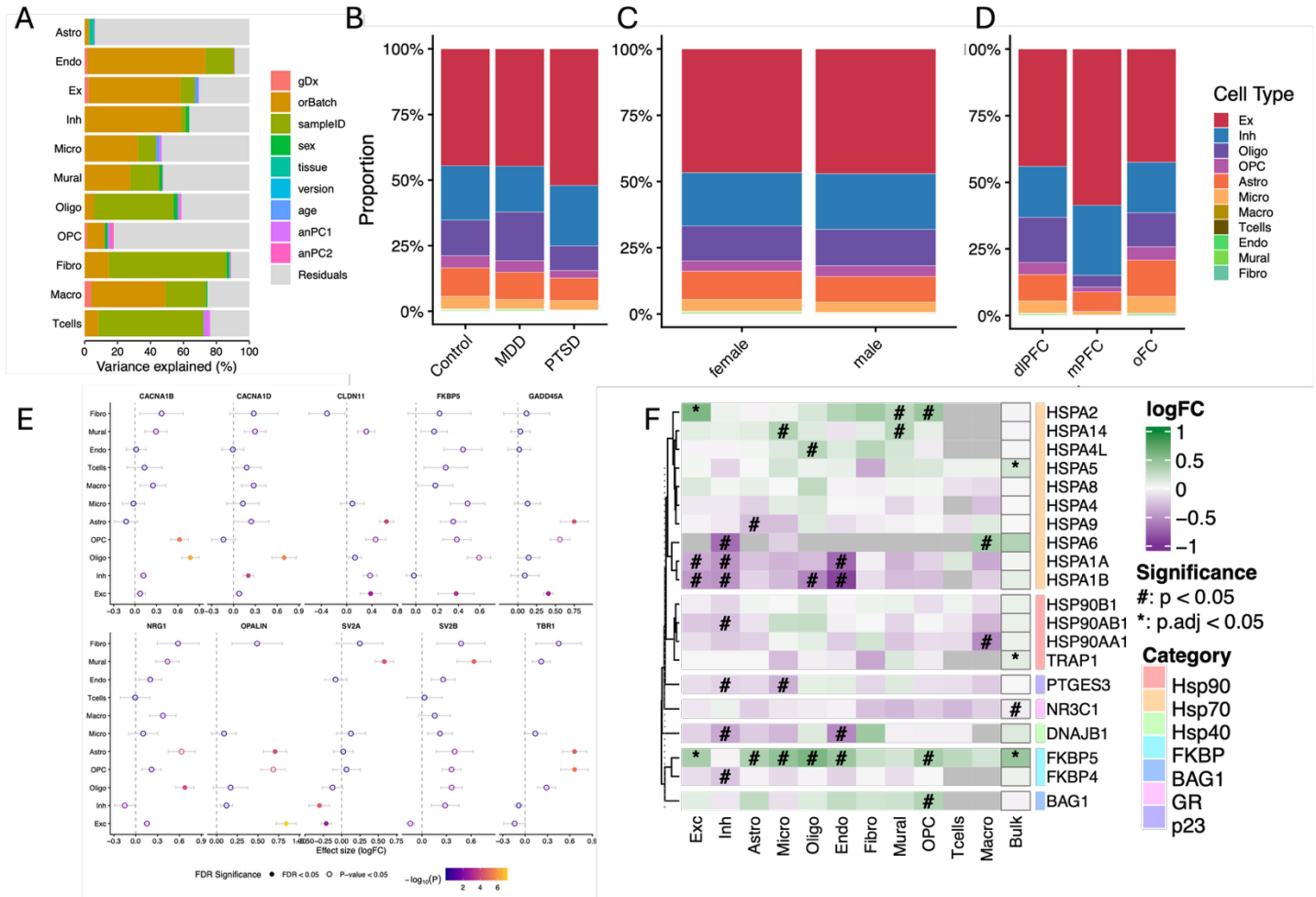


Supplementary Figure 2. External replication and sensitivity analyses of MDD-associated transcriptional signatures. (A) Heatmap displaying pairwise Spearman correlations of genome-wide differential expression effect sizes between the MDD-associated transcriptional profiles from Girgenti et al. (2021) — comprising the orbitofrontal cortex (oFC) and subgenual PFC (sgPFC) — and the region-specific and FC-aggregated models from the current study. Correlations are shown at the genome-wide, nominal level in this study and restricted to FDR-significant genes in either this study or both. (B) Schematic overview of the sensitivity analyses conducted, illustrating the two analytical frameworks applied: (i) inclusion of reported childhood trauma (CT) exposure as a predictor variable, and (ii) inclusion of the degree of substance use/abuse/dependence as an additional covariate in the FC-aggregated and region-specific models. (C) Boxplots showing genome-wide Spearman correlations of differential expression effect sizes between each sensitivity analysis — CT and substance use-adjusted — and the primary FC-aggregated model, across the FC-aggregated and three region-specific models. Each boxplot displays the median (vertical line), interquartile range (box), and whiskers extending to $1.5 \times$ the IQR. Both sensitivity analyses produced highly concordant signatures relative to the primary model (CT: $\rho \approx 0.85$; substance use: $\rho \approx 0.78$), confirming that the core MDD-associated transcriptional signal is not driven by childhood trauma exposure or comorbid substance use.



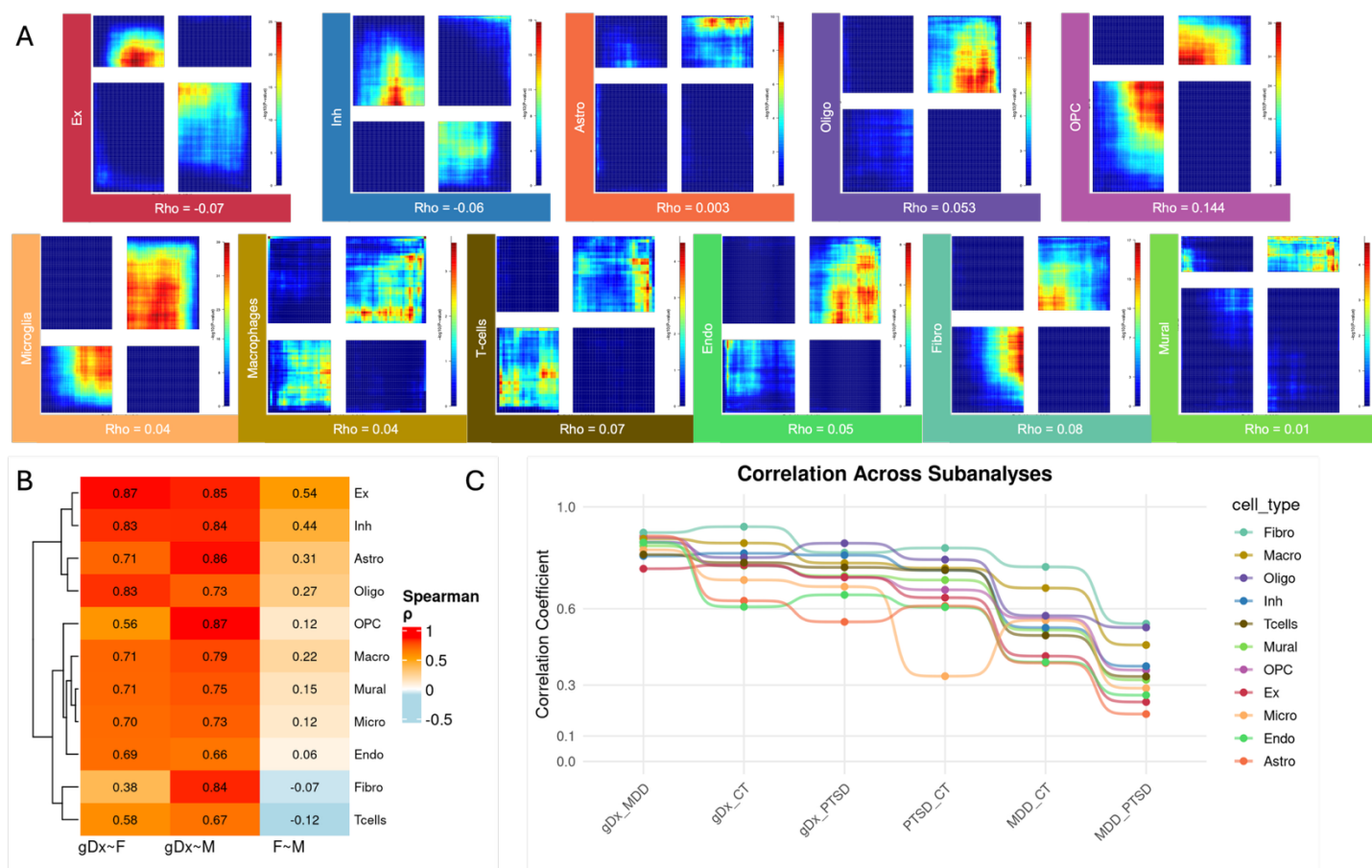
Supplementary Figure 3. Extended pathway, spatial, genomic, and multi-omic analyses.

(A) Heatmap of pairwise Jaccard similarity indices comparing the sets of significantly enriched GO:BP pathways (upregulated and downregulated separately) between the FC-aggregated model and each region-specific analysis (mPFC, dlPFC, dACC). Scale represents Jaccard similarity (0–1), with higher values indicating greater overlap in enriched pathway sets. Upregulated pathways showed stronger cross-regional concordance than downregulated pathways, consistent with the asymmetry observed in NES correlations. (B) Barplot displaying the top 20 upregulated and top 20 downregulated GO:BP pathways from the FC-aggregated analysis, ranked by NES. Color reflects enrichment direction (red: positive; blue: negative) and point size or bar length reflects $-\log_{10}(\text{FDR-adjusted } p\text{-value})$. (C) Heatmap showing enrichment of cell-type-specific marker gene sets (from the snRNA-seq atlas generated in this study) among MDD-associated DEGs from the FC-aggregated and each region-specific bulk analysis, serving as orthogonal validation of the spatial registration results. Color reflects $\log_2$ odds ratio (blue: depletion; red: enrichment). Significance is indicated by asterisks (* FDR < 0.05; ** FDR < 0.01; *** FDR < 0.001). Non-neuronal cell-type markers were consistently enriched and neuronal markers depleted across all analyses, mirroring the leptomeningeal and white matter spatial enrichment patterns. (D) Heatmap of MAGMA gene-set enrichment results for upregulated and downregulated DEGs from the FC-aggregated and each region-specific analysis, tested against GWAS summary statistics for major depression (MD), PTSD, schizophrenia (SCZ), and bipolar disorder (BIP). Color reflects standardized effect size (blue: negative; red: positive). No associations survived correction for multiple testing, consistent with emerging evidence that disease-state transcriptional changes and genetic risk architecture reflect partially distinct biological processes. (E) Manhattan plot displaying SMR results across all tested genes. Points are colored by analytical category: green, genes passing both SMR (FDR < 0.05) and HEIDI ($p > 0.05$) filters; purple, genes identified in both TWAS and SMR; orange, genes overlapping between SMR eGenes and bulk DEGs; red, genes identified across all three approaches (DEG, TWAS, and SMR). (F) Heatmap of Spearman correlations between the 30 MOFA latent factor scores and demographic and clinical variables. Color reflects Spearman ρ (blue: -1 ; red: $+1$). Correlations with absolute $\rho > 0.3$ are labeled with their exact value. (G) Heatmap of Spearman correlations between the 30 MOFA latent factor scores and deconvolved cell-type proportion estimates derived from the snRNA-seq atlas. Color reflects Spearman ρ (blue: -1 ; red: $+1$). (H) Boxplot of MOFA Factor 4 scores stratified by diagnostic group: neurotypical controls, MDD-only, and PTSD-comorbid. Each boxplot displays the median (horizontal line), interquartile range (box), and whiskers extending to $1.5 \times$ the IQR; individual samples are overlaid as points. Pairwise comparisons are indicated by significance brackets (** $p_{\text{adj}} < 0.01$; *** $p_{\text{adj}} < 0.001$; ns: not significant). Factor 4 scores were significantly elevated in both MDD-only ($p_{\text{adj}} = 2.2 \times 10^{-4}$) and PTSD-comorbid ($p_{\text{adj}} = 4.5 \times 10^{-4}$) groups relative to controls, with no significant difference between the two diagnostic subgroups ($p_{\text{adj}} = 0.24$).



Supplementary Figure 4. Cell-type composition, covariate structure, and cross-cell-type transcriptional signatures.

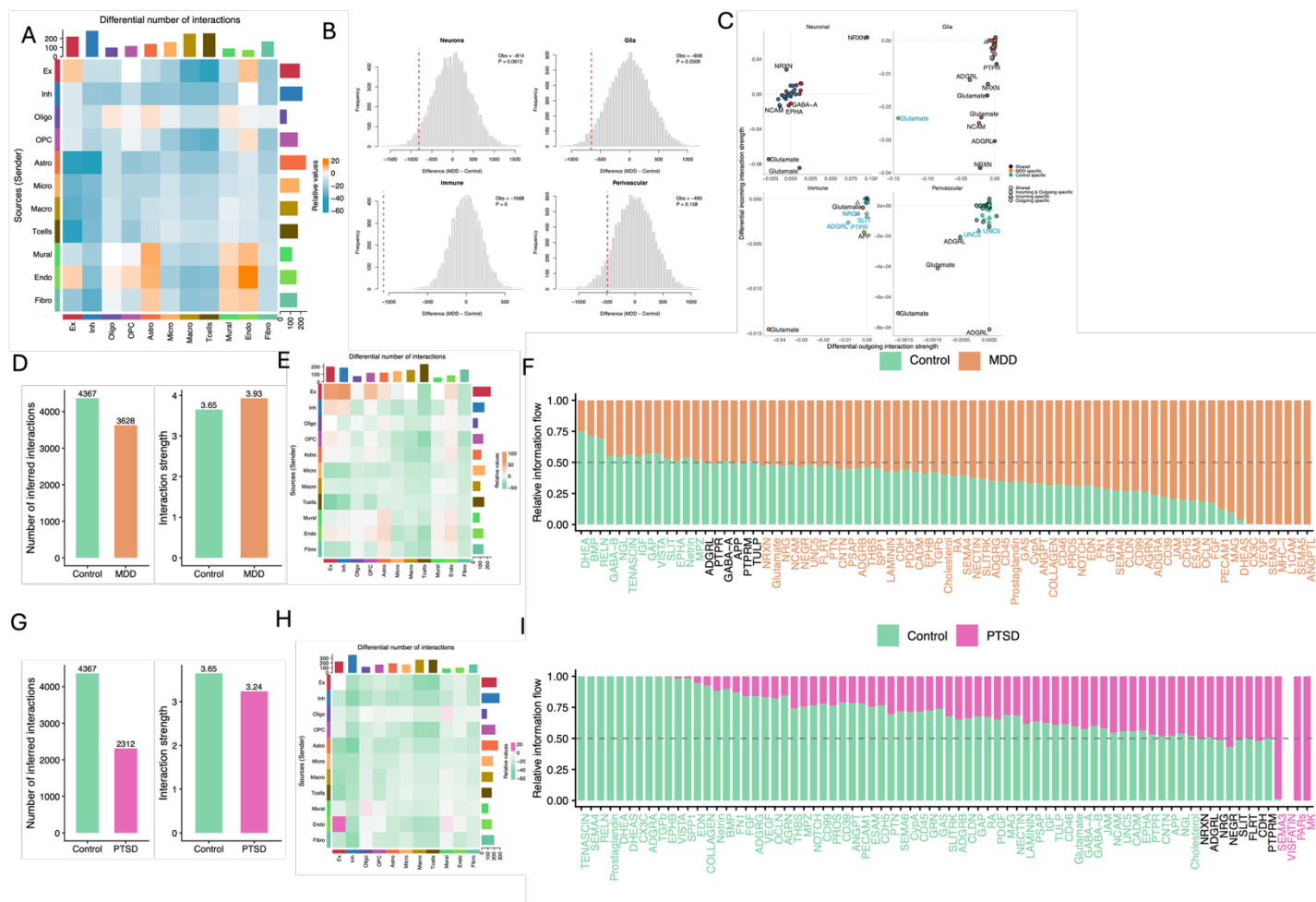
(A) Stacked bar plot of cell-type proportion variance explained by each covariate included in the pseudobulk linear mixed model, estimated separately for each major cell type using crumbly. Each bar represents one cell type; color-coded segments reflect the proportion of variance attributable to diagnosis, sex, age, ancestry PCs, study batch, chemistry version, tissue subregion, and donor (random effect). (B–D) Stacked bar plots showing cell-type proportion estimates for each major cell class across diagnostic subgroups — neurotypical controls, MDD-only, and PTSD-comorbid (B); biological sex — female and male (C); and cortical subregion — dlPFC, mPFC, and oFC (D). A modest reduction in macrophages was observed in MDD-only cases ($T = -2.91$, Bonferroni-adjusted $p = 0.045$); astrocytes were significantly more abundant in the mPFC relative to other regions ($T = 4.06$, Bonferroni-adjusted $p = 7.8 \times 10^{-4}$) after correcting for covariates. Sex and PTSD-comorbid effects did not survive multiple testing correction. (E) Forest plot of genes showing MDD-associated differential expression across multiple cell types, prioritized based on the number of FDR-significant cell-type populations, number of nominally associated populations, cross-cell-type directional consistency, and effect size magnitude. Each row represents one gene; points show log₂ fold change estimates per cell type, with error bars reflecting standard error. Point color reflects -log₁₀(FDR-adjusted p-value); filled points indicate FDR significance (FDR < 0.05); open points indicate nominal significance. Genes with consistent cross-cell-type dysregulation include synaptic and calcium signaling mediators (*NRG1*, *SV2A*, *SV2B*, *CACNA1B*, *CACNA1D*), myelination-associated factors (*CLDN11*, *OPALIN*), and stress- and activity-responsive regulators (*GADD45A*, *FKBP5*). (F) Heatmap of normalized expression of glucocorticoid receptor (GR) complex chaperone and co-chaperone genes across major cell types. Rows represent individual genes grouped by functional chaperone category (GR-stabilizing: HSP90, p23; functional folding: HSP40, HSP70; co-regulators: FKBP5, FKBP4; receptor: NR3C1); columns represent cell types and the bulk tissue analysis. Color reflects log₂ fold change in MDD relative to controls (green: upregulated; purple: downregulated). Statistical significance is indicated by symbol (# nominal $p < 0.05$; * FDR < 0.05). Coordinated downregulation of GR-stabilizing and folding chaperones alongside *FKBP5* upregulation and *NR3C1* downregulation was observed across multiple cell types, suggesting broad impairment of GR complex assembly and nuclear translocation in MDD.



Supplementary Figure 5. Cross-modality concordance and subgroup specificity of MDD transcriptional signatures at cell-type resolution.

(A) Rank-Rank Hypergeometric Overlap (RRHO) plots comparing bulk differential expression results (x-axis) with pseudobulk differential expression results for each major cell type (y-axis). Genes are independently ranked by their differential expression z-score in each analysis; color reflects $-\log_{10}(\text{hypergeometric p-value})$ for the overlap between the top or bottom of each ranked list at a sliding threshold. Concordant upregulation (bulk $\uparrow$ / cell type $\uparrow$) appears in the top-right quadrant; concordant downregulation (bulk $\downarrow$ / cell type $\downarrow$) in the bottom-left; discordant signals occupy the top-left (bulk $\uparrow$ / cell type $\downarrow$) and bottom-right (bulk $\downarrow$ / cell type $\uparrow$) quadrants. Hot colors (yellow-red) in the concordant quadrants indicate significant overlap beyond chance. Beneath each RRHO plot, the Spearman ρ of the genome-wide z-score correlation between the bulk and cell-type analysis is reported, providing a summary measure of overall transcriptional concordance.

(B) Heatmap of Spearman correlations of genome-wide differential expression effect sizes between sex-stratified and primary pseudobulk analyses, shown separately for each major cell type (rows). Three comparisons are displayed as columns: primary MDD analysis (gDx) versus female-specific MDD analysis (gDx vs. F); primary MDD analysis versus male-specific MDD analysis (gDx vs. M); and female-specific versus male-specific analysis (F vs. M). Color reflects Spearman ρ . **(C)** Bump plot displaying Spearman correlations of genome-wide differential expression effect sizes between pairwise subgroup comparisons across major cell types. Each line represents one cell type (color-coded); the x-axis shows the six pairwise comparisons: primary MDD (gDx) vs. MDD-only (gDx_MDD), primary MDD vs. childhood trauma-exposed (gDx_CT), primary MDD vs. PTSD-comorbid (gDx_PTSD), PTSD-comorbid vs. childhood trauma (PTSD vs. CT), MDD-only vs. childhood trauma (MDD vs. CT), and MDD-only vs. PTSD-comorbid (MDD vs. PTSD). The y-axis shows Spearman ρ .



Supplementary Figure 6. Diagnosis-stratified cell-cell communication remodeling in MDD and PTSD-comorbid cases. (A) Heatmap of the differential number of predicted significant ligand-receptor interactions between MDD and neurotypical controls (MDD minus control), organized by sender (ligand-expressing; rows) and receiver (receptor-expressing; columns) cell type. Color reflects the magnitude and direction of change (orange: more interactions in MDD; blue: fewer interactions in MDD). (B) Density plots showing the distribution of interaction number differences (MDD minus control) aggregated at the level of major cell class groups, derived from 10,000 permutations of diagnostic labels. The red vertical line indicates the observed interaction number difference for each group. Groups whose observed value falls outside the permuted null distribution ($p < 0.05$) are considered to show significant remodeling of communication architecture in MDD. (C) Scatterplot of pathway-level communication strength for incoming and outgoing signals per major cell group. Each point represents one signaling pathway in one cell group; the x-axis shows incoming signal strength and the y-axis shows outgoing signal strength, or vice versa as applicable. Point color reflects cell type identity. Point shape denotes the directionality category of the pathway: square = predominantly incoming; triangle = predominantly outgoing; diamond = both incoming and outgoing; circle = shared across conditions. Label color indicates condition specificity: blue = control-enriched pathway activity; orange = MDD-enriched pathway activity. (D) Bar plot of the total number of inferred interactions and aggregate interaction strength (MDD-only minus control) across major cell-type sender-receiver pairs in the MDD-only subgroup. (E) Heatmap of the differential number of predicted significant ligand-receptor interactions between MDD-only cases and neurotypical controls, formatted as in (A). Color reflects MDD-only minus control interaction counts (orange: more interactions in MDD-only; green: fewer). (F) Stacked bar plots of relative information flow for neurotypical controls and MDD-only subgroup across all inferred signaling pathways. Each bar represents one signaling pathway; color indicates the condition showing significantly greater pathway activity: green for control-enriched, orange for MDD-only-enriched pathways. (G) Bar plot of the total number of inferred interactions and aggregate interaction strength (PTSD-comorbid minus control) across major cell-type sender-receiver pairs, formatted as in (D). (H) Heatmap of the differential number of predicted significant ligand-receptor interactions between PTSD-comorbid cases and neurotypical controls, formatted as in (A) and (E). Color reflects PTSD-comorbid minus control interaction counts. (I) Stacked bar plots of relative information flow for controls and the PTSD-comorbid subgroup across all inferred signaling pathways. Each bar represents one signaling pathway; green for control-enriched, and pink for PTSD-comorbid-enriched pathways.