Article

Cell type specific transcriptomic differences in depression show similar patterns between males and females but implicate distinct cell types and genes

Malosree Maitra[1,2,3], Haruka Mitsuhashi[1,2,3], Reza Rahimian[1,3], Anjali Chawla[1,2,3], Jennie Yang[1,3], Laura M. Fiori[1,3], Maria Antonietta Davoli[1,3], Kelly Perlman[1,2,3], Zahia Aouabed[1,3], Deborah C. Mash[4], Matthew Suderman [5,6], Naguib Mechawar [1,2,3], Gustavo Turecki [1,2,3,7] ✉ & Corina Nagy [1,2,3,7] ✉

Major depressive disorder (MDD) is a common, heterogenous, and potentially serious psychiatric illness. Diverse brain cell types have been implicated in MDD etiology. Significant sexual differences exist in MDD clinical presentation and outcome, and recent evidence suggests different molecular bases for male and female MDD. We evaluated over 160,000 nuclei from 71 female and male donors, leveraging new and pre-existing single-nucleus RNA-sequencing data from the dorsolateral prefrontal cortex. Cell type specific transcriptome-wide threshold-free MDD-associated gene expression patterns were similar between the sexes, but significant differentially expressed genes (DEGs) diverged. Among 7 broad cell types and 41 clusters evaluated, microglia and parvalbumin interneurons contributed the most DEGs in females, while deep layer excitatory neurons, astrocytes, and oligodendrocyte precursors were the major contributors in males. Further, the Mic1 cluster with 38% of female DEGs and the ExN10_L46 cluster with 53% of male DEGs, stood out in the meta-analysis of both sexes.

Major depressive disorder (MDD) is a serious and potentially debilitating mental illness affecting 200–300 million people worldwide[1]. MDD is a leading cause of disability globally[1] and some prominent symptoms in patients with MDD include persistent low mood, decreased interest and/or pleasure, sleep and appetite disturbances, feelings of worthlessness, and suicidal thoughts[2]. A number of genetic variants have been identified which contribute to the heritability of MDD[3] and brain transcriptomic differences[4] are detected in this disease, but the molecular etiology of MDD is still only partially understood.

There are known dissimilarities in the epidemiology and pathophysiology of MDD between the sexes. Notably, it is twice as prevalent in women than men[5]. Symptomatology differs in that, women are more likely to have comorbid anxiety, so-called atypical depression, and recurrent episodes, while men are more likely to have comorbid substance use disorders and to die by suicide[6–8]. Sex-specific molecular profiles in MDD and corresponding animal models are often attributed to hormonal differences either during development or in adulthood, to the contributions of sex-chromosomes, or to inherent sex

[1]McGill Group for Suicide Studies, Douglas Institute, Verdun, QC, Canada. [2]Integrated Program in Neuroscience, McGill University, Montreal, QC, Canada. [3]Douglas Institute, Department of Psychiatry, McGill University, Montreal, QC, Canada. [4]Dr. Kiran C. Patel College of Allopathic Medicine, Nova Southeastern University, Ft. Lauderdale, FL, USA. [5]Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK. [6]MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK. [7]These authors jointly supervised this work: Gustavo Turecki, Corina Nagy. ✉e-mail: gustavo.turecki@mcgill.ca; corina.nagy@mcgill.ca

differences in the monoaminergic system or the hypothalamic-pituitary-adrenal axis (HPA), among other factors[6,9].

Recent studies in humans have attempted to address the gap in our knowledge of molecular sex differences in depression by examining MDD-associated sex-specific brain transcriptomic differences in human patients[7,10]. Using bioinformatic and meta-analysis approaches, combined with validation in animal models, these studies found that, overall, MDD-associated differences in brain transcriptomics are primarily sex-specific across brain regions, with very little overlap of differentially expressed genes (DEGs) and discordance in overall patterns of difference between the sexes.

Single-nucleus RNA-sequencing (snRNA-seq) can disentangle cell type specific transcriptomic contributions to complex neuropsychiatric conditions[11–15], and our recent snRNA-seq results[16] revealed disruptions in deep layer excitatory neurons and immature oligodendrocyte precursor cells (OPCs) in the prefrontal cortex (PFC) of males with MDD. Given the higher prevalence of MDD among women, the known sex-specific differences in MDD, and growing evidence that male and female MDD may be mediated by distinct brain molecular mechanisms, we conducted a study in a cohort of female individuals and applied an updated unified analysis pipeline to both the female and previously generated male cohorts. With a total of 71 individuals, 37 cases and 34 controls and over 160,000 single-nuclei profiled, our dataset represents the largest snRNA-seq study of the human brain in MDD to date. We found that the DEGs detected and the cell types with prominent differences were distinct in males and females. However, the overall patterns of MDD-associated gene expression difference within each cell type were consistent between the sexes. Whereas in males our analysis indicated a strong involvement of deep layer excitatory neurons, astrocytes, and OPCs—consistent with our previous report, in females we found a striking contribution of microglia and parvalbumin (PV) interneurons to MDD pathology.

Results

Profiling cells of the human dorsolateral prefrontal cortex (dlPFC)

snRNA-seq data was generated from the dlPFC for 20 female subjects with MDD and 18 neurotypical female controls (Fig. 1a, schematic; Table 1, demographic and sample characteristics; Supplementary Data 1, sequencing metrics) and combined with previously generated data from males[16]. After pre-processing with a unified pipeline (Methods: Sequencing, alignment, and generation of count matrices), we retained 160,711 high-quality nuclei with comparable contributions of sex (51% from females) and disease status (58% MDD). We used Harmony[17] to correct for covariates, including batch effects (Supplementary Fig. 1a–d), and applied the scclusteval[18] workflow to optimize the Seurat clustering parameters (Supplementary Fig. 2a–b) resulting in the identification of 41 nuclei clusters. Clusters mostly did not appear to be driven by batch, sex, brain bank, or subject (Supplementary Fig. 3a–e, g).

Of the 41 clusters, 40 could each be confidently annotated to one of 7 major brain cell types (Methods: Cluster annotation, Fig. 1b–d, Supplementary Fig. 4, Supplementary Data 2)—excitatory neurons (48% of nuclei), inhibitory neurons (18% of nuclei), oligodendrocytes (14% of nuclei), astrocytes (8% of nuclei), OPCs (5% of nuclei), endothelial cells (2.5% of nuclei), and microglia (2% of nuclei). The one unassigned cluster displayed a mixed expression profile of neuronal and glial marker genes (2% of nuclei).

We annotated 30 neuronal clusters, both excitatory (20 clusters) and inhibitory (10 clusters), using known subtype markers (Supplementary Fig. 5a, b). Excitatory neuronal clusters were annotated according to their layer of origin and inhibitory neuronal clusters according to their developmental origin, where applicable. For non-neuronal cells, we identified one microglial cluster, two clusters of astrocytes, and three clusters each of oligodendrocytes and OPCs.

Clusters annotated to the oligodendrocyte lineage (OL) were further characterized using pseudotime trajectory analysis (Methods: Pseudotime trajectory analysis; Supplementary Fig. 5c, d).

Using the gene expression patterns of our clusters and matching them to published clusters in several human brain datasets[19,20], we found close correspondence between observed cell types (Fig. 1e, Supplementary Fig. 6), further emphasizing that the quality of our data, clustering, and annotation are at par with other recent snRNA-seq datasets for the human brain. Besides the single cluster with mixed expression profile, two other clusters showed evidence of possible technical effects, ExN17 and ExN5 (Methods: Assessment of clustering quality).
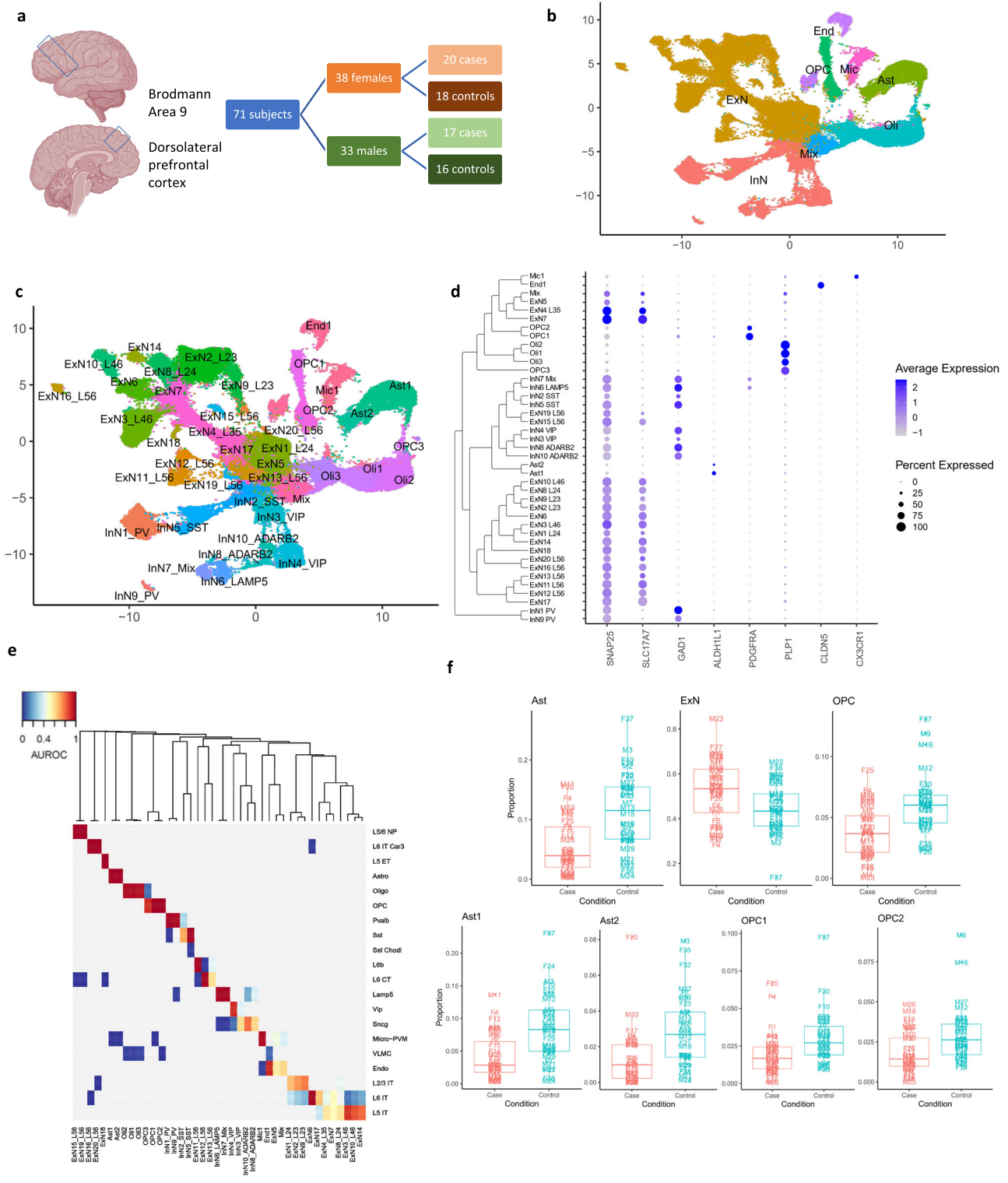
Cell types with altered proportions in MDD

We next examined whether proportions of nuclei in broad cell types and clusters differed between cases and controls. We observed that the proportions of nuclei per subject contributing to the broad astrocytic and OPC cell types were significantly decreased in cases compared to controls (two-sided Wilcoxon-test, FDR = 3.46×10^{-4}, Ast; FDR = 5.32×10^{-4}, OPC; Fig. 1f, Supplementary Data 3) and there were concomitant increases in excitatory neurons (FDR 0.0477). Similarly, there were reduced proportions of nuclei in both astrocytic clusters (Ast1, FDR 0.00188; Ast2, FDR 0.00291) and in two of three OPC clusters (OPC1, FDR 0.009799; OPC2, FDR 0.0168; Fig. 1f, Supplementary Data 3). The robustness of these differences was supported by sub-sampling analysis (Methods: Cell type proportions comparison). Splitting the male and female datasets revealed similar patterns as observed for the combined data (Supplementary Fig. 7). These results are similar to those found in analyses of other brain disorders[11,21], and indicate that there may be decreased proportion of astrocytes and OPCs may be reduced in MDD. Here the FDR refers to Benjamini and Hochberg correction.

Global cell type specific transcriptomic changes are largely concordant between the sexes

We next asked whether there are sex-specific differences in the gene expression patterns of individual cell types. To answer this question, we performed differential gene expression analysis comparing cases and controls in broad cell types and clusters, in males and females separately. In both males and females, we observed a high proportion of common DEGs between broad and cluster level analyses. However, consistent with previous studies showing distinct brain transcriptomic changes in males and females with MDD[10,22], few DEGs were common to both sexes (Fig. 2a). To compare overall patterns of depression-associated gene expression in males and females beyond those genes passing significance thresholds, we performed rank-rank hypergeometric overlap (RRHO) analysis[23] (Methods: Comparison of male and female results). Specifically, we used RRHO2 to compare the orderings of the genes induced by MDD association statistics in males compared to females. These orderings were generally moderately to strongly concordant between the sexes (Fig. 2b). Some evidence of discordance was visible only for Oli and OPC. There was a significant overlap between males and females in genes less expressed in MDD in Ast, ExN, and InN (warm colors in top right quadrant of RRHO plots) and an overlap in genes more expressed in MDD in Mic (warm colors in bottom left quadrant of RRHO plot).

At the cluster level there was some evidence of discordance between the sexes, with 8 out of 34 clusters compared showing discordant patterns. This encompassed certain neuronal clusters, primarily excitatory neuronal, including ExN4_L35, ExN7, ExN12_L56, ExN13_L56, InN10_ADARB2 (Supplementary Fig. 8). Within the oligodendrocyte lineage, discordance is apparent for the Oli2, Oli3, and OPC1 clusters (Fig. 2c). Supplementary Data 4 summarizes the maximum -log10 p-values from RRHO2 analyses and the classification of the results into weak, moderate, strong categories or concordance and discordance—

with strongly concordant or discordant results providing the most convincing evidence for similarity or difference between the sexes. As can be seen from Supplementary Data 4, we can see moderate or strong evidence for concordance between males and females in not only neuronal cell types, but also glia (e.g., microglia and astrocytes).

Taken together we find that, although cell type specific statistically significant MDD-associated DEGs differ between the sexes, a threshold-free ranking approach to comparison shows considerable concordance between males and females for the majority of broad cell types and clusters.

We further assessed whether the similarities in cell type specific MDD-associated gene expression differences between males and females was likely to arise by chance using permutation analysis, which supported our conclusion that the similarities are not driven by chance (Methods: Permutation analysis, Supplementary Fig. 9e–j, Supplementary Data 4). For broad cell types, excluding the cluster annotated as having a mixed contribution of cell types, on average 91% of the time the real data yielded a higher correlation between male and female results than the permuted data. For clusters with concordant patterns, on average 90% of the time the real data yielded

Fig. 1 | Overview of cell types characterized in the dlPFC. a Schematic of study design. Diagrams depict the brain region of interest, Brodmann area 9, corresponding to the dlPFC. **b** UMAP plot colored by the broad cell types. **c** UMAP plot colored by the individual clusters identified and annotated. For UMAP plots, the *x* and *y*-axes represent the first and second UMAP co-ordinates respectively. **d** DotPlot depicting the expression of marker genes (*SNAP25* – neurons, *SLC17A7* – excitatory neurons, *GAD1* – inhibitory neurons, *ALDH1L1* – astrocytes, *PDGFRA* – oligodendrocyte precursor cells, *PLP1* – oligodendrocytes, *CLDN5* – endothelial cells, *CX3CR1* – microglia). The dendrogram next to the cluster names shows the relationship between the clusters by using the distance based on average expression of highly variable genes. **e** Best hits heatmap from MetaNeighbor showing the correspondence between the clusters in our dataset (columns) and the broad categories of cells identified in the Allen Brain Institute human motor cortex snRNA-seq dataset[20] (rows). **f** Boxplots showing the proportion of nuclei in each cluster for each subject split by cases and controls for the broad OPC, astrocyte, and excitatory neuron cell types and the Ast1, Ast2, OPC1, and OPC2 clusters (*n* = 37

cases, 34 controls, representing biologically independent samples for each cluster or broad cell type). The middle line is the median. The lower and upper hinges correspond to the 25th and 75th percentiles. Upper and lower whiskers extend from the upper or lower hinges to the largest or smallest value no further than 1.5 times the inter-quartile range from the hinge, where the inter-quartile range is the distance between the first and third quartiles. Points beyond the end of the whiskers are plotted individually. In Fig. 1c–e, excitatory neuronal cluster names contain approximate layer annotations and inhibitory neuronal cluster names contain MGE or CGE specific marker information as a suffix where applicable, as described in methods: Cluster annotation. For example, ExN10_L46 denotes a cluster of excitatory neurons with enrichment of marker genes from layer 4 to layer 6 of the cortex and InN1_PV denotes a cluster on inhibitory neurons with enrichment of the MGE specific marker PV. This convention is used throughout the paper. Brain diagram in 1a was created with BioRender.com. Source data are provided as a Source Data file.

Table 1 | Demographic and sample characteristics of cohorts

Group	Case (*n* = 37)		Control (*n* = 34)	
Sex	Female (*n* = 20)	Male (*n* = 17)	Female (*n* = 18)	Male (*n* = 16)
Age	45.10 ± 3.19 (0.92)	41.06 ± 4.66 (0.67)	47.89 ± 4.45 (0.92)	38.38 ± 4.58 (0.67)
PMI*	41.49 ± 3.07 (0.02)	41.69 ± 4.76 (0.10)	30.27 ± 4.73 (0.02)	32.02 ± 4.81 (0.10)
pH	6.58 ± 0.08 (0.06)	6.60 ± 0.07 (0.30)	6.34 ± 0.08 (0.06)	6.50 ± 0.06 (0.30)

Numbers in brackets are uncorrected *p*-values from Kruskal–Wallis test between the two conditions for the same sex. Numeric values in each cell represent the mean ± SEM. pH was unavailable for one female subject, F35.
*Significantly different between female cases and controls (*p*-value < 0.05).

a higher correlation between males and females than the permuted data. However, for clusters with evidence of discordance, the real correlation was higher than permuted correlation only 42% of the time.

Cell types with strongest MDD associations differ by sex

Next, we identified the cell types with the strongest evidence of dysregulation due to MDD in each sex. In males, our reanalysis indicated results consistent with those we reported previously, i.e., for broad cell types we identified the highest number of DEGs in astrocytes (90/151, 60%) and OPCs (54/151, 36%) (Fig. 3a, Supplementary Fig. 10a–d, Supplementary Data 5), whereas at the cluster level (Fig. 3b, Supplementary Figs. 10e–h, 11, Supplementary Data 5), the highest number of DEGs were found in a cluster of deep layer excitatory neurons—ExN10_L46 (238/447, 53%) and a cluster of astrocytes – Ast1 (98/447, 22%). A summary of the proportions of upregulated versus downregulated genes and unique DEGs versus DEGs shared across clusters is provided in Fig. 3e. Correlations between gene expression fold differences calculated in our reanalysis and our previous analysis are provided in Supplementary Fig. 10i–l (Methods: Differential expression analysis—Comparison of male differential expression results to previous results).

In females, for broad cell types, we detected a high number of DEGs in microglia only (74/85, 87%) (Fig. 3c, Supplementary Fig. 12a–d, Supplementary Data 6). The same analysis at the cluster level (Fig. 3d, Supplementary Fig. 12e–h, Supplementary Data 6) consistently showed the highest number of DEGs in the Mic1 (Fig. 3f; 68/180 DEGs, 38%) cluster with a large proportion (53/68, 78%) overlapping with the microglial DEGs at the broad level. We focused on cluster level results for follow up analyses (Methods: Differential expression analysis, for justification) and assessed the robustness of our microglial findings against misclassified or contaminating cells (Supplementary Fig. 12i).

The majority of microglial DEGs (47/68, 69%) were confirmed to be both transcribed and translated in microglia using a TRAP gene expression dataset in a lipopolysaccharide challenge mouse model[24].

In addition to microglia, several inhibitory neuronal clusters (Fig. 3g), including two *PVALB* expressing clusters—InN1_PV and InN9_PV as well as an *SST* expressing cluster—InN2_SST and an *ADARB2* expressing cluster—InN8_ADARB2 contained the majority of remaining DEGs. Our results thus pointed to dysregulation of microglia and inhibitory neurons, especially PV interneurons in females with MDD which further prompted us to explore the biological pathways within and possible interactions between these cell types which could be altered in MDD, as detailed below.

Further, our permutation analyses revealed that at the broad level the number of unique DEGs identified with the real data was higher than 93% of permutations for females and 97% of permutations for males (Supplementary Fig. 9a, b). At the cluster level, for males, the real number of unique DEGs was higher than the number of permuted unique DEGs 94% of the time (Supplementary Fig. 9c). The evidence from permutations was weaker at the cluster level for females with 60% of permutations revealing fewer unique DEGs than the real data (Supplementary Fig. 9d).

Meta-analysis reveals additive effects of depression-associated transcriptomic changes in males and females

To maximize statistical power to observe gene expression differences common to both males and females, we performed meta-analyses of the male and female data within each broad cell type and cluster. For broad cell types, the meta-analysis revealed upregulated genes in microglia and downregulated genes in astrocytes, with the majority of DEGs from the separate male and female analyses retained (Fig. 4, Supplementary Data 7). There were more DEGs in microglia (172 DEGs) than observed in the female dataset alone (74 DEGs), whereas there were fewer DEGs in astrocytes (53 DEGs) than identified in males alone (90 DEGs). 49/90 (54%) DEGs in the broad astrocytic cluster in males and 56/74 (76%) DEGs in the broad microglial cluster in females were recapitulated in the meta-analysis. There were 22 DEGs in OPCs in the meta-analysis, but the number was less than half compared to the independent analysis of the male dataset (54 DEGs) whereas for oligodendrocytes the number of DEGs was higher when the data were
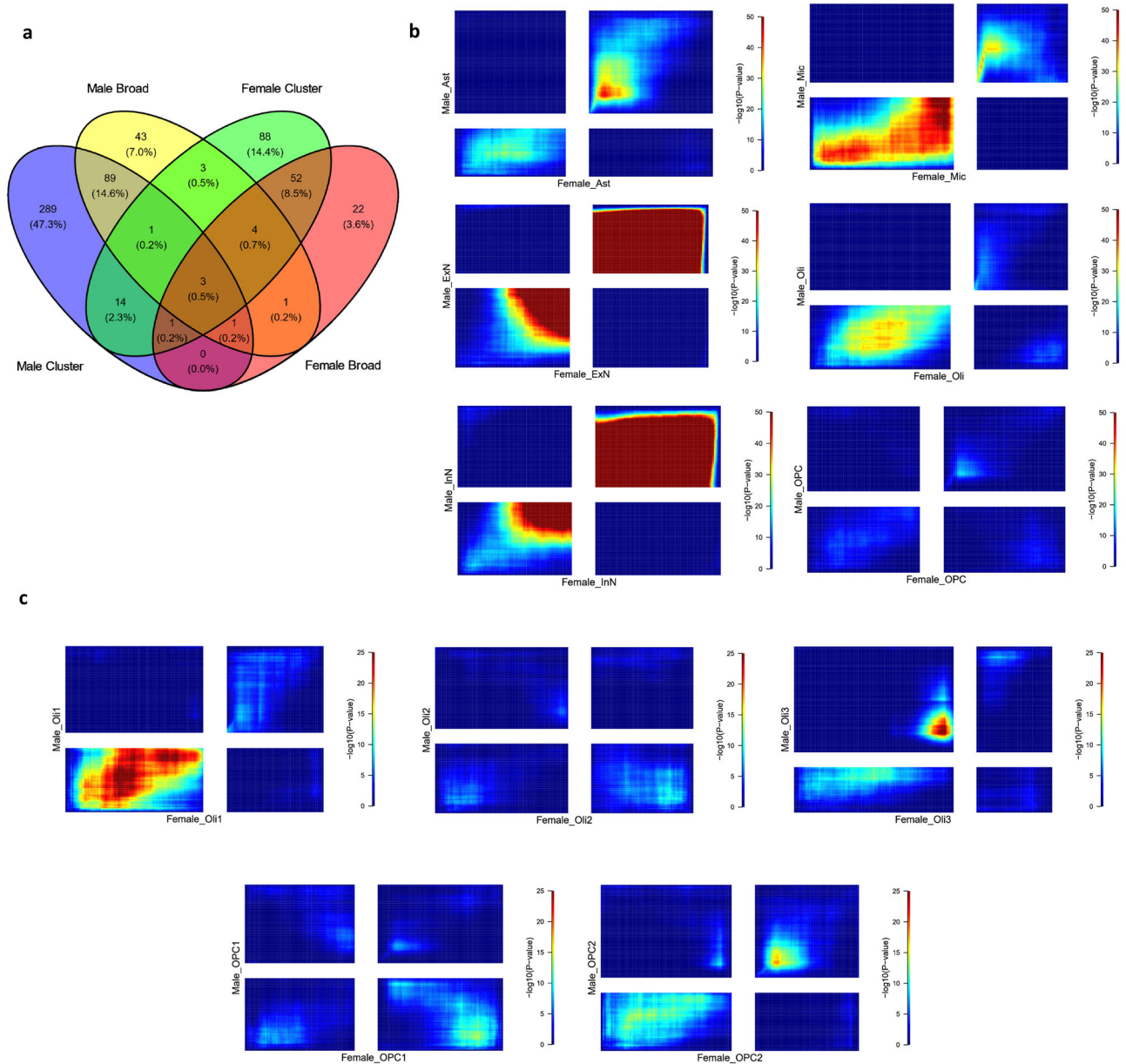
Fig. 2 | Overall comparison of cell type specific MDD-associated gene expression changes in males and females. a Venn diagram showing the overlap of DEGs between the male and female datasets at the broad cell type and cluster levels. **b** RRHO2 plots for correspondence between differential expression results for broad cell types in the female (*x*-axis) and male (*y*-axis) datasets. Warm colors in the bottom left and top right quadrants reflect overlap in genes with increased expression or decreased expression respectively, in cases versus controls between the male and female datasets. Warm colors in the top left and bottom right quadrants reflect overlaps in genes w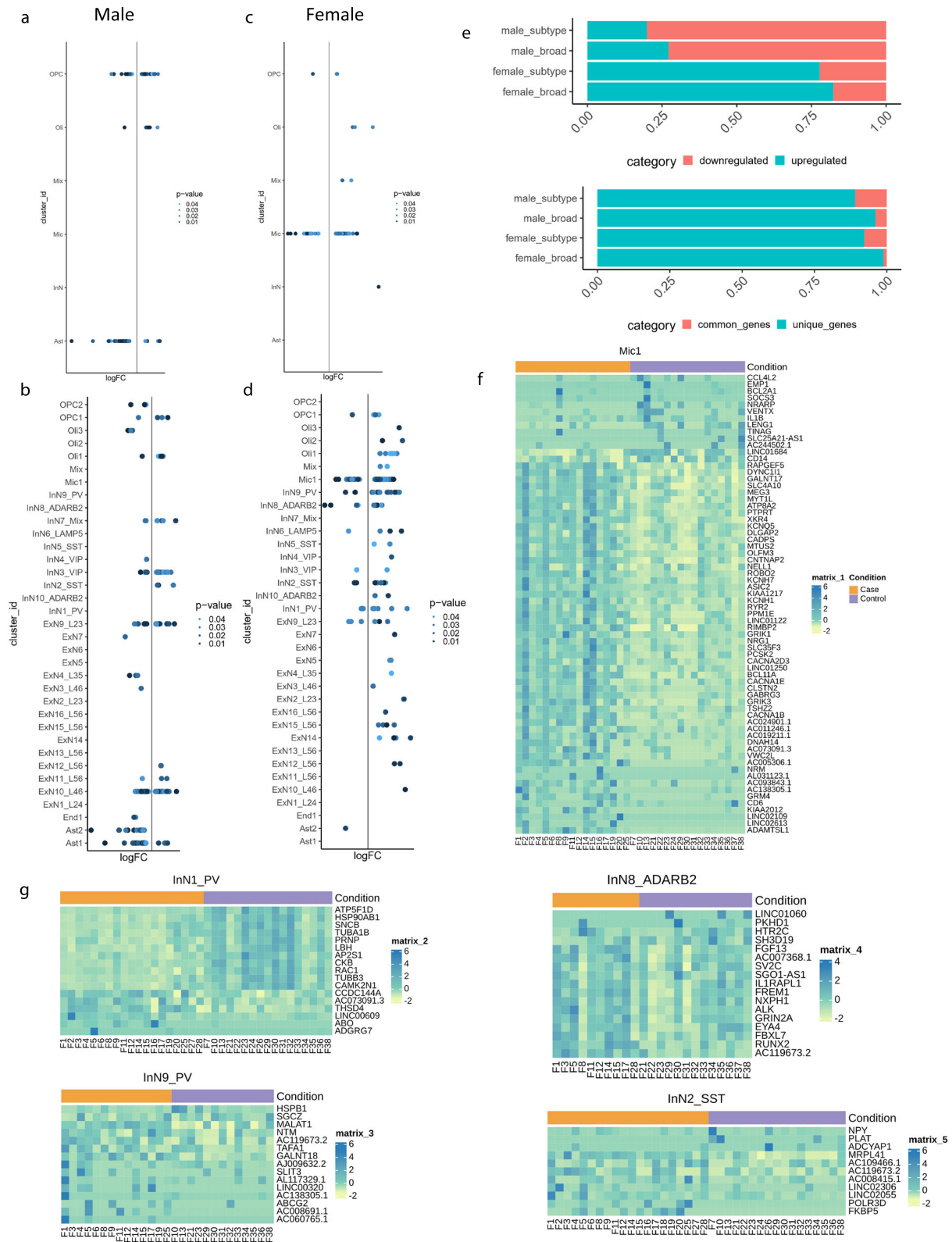ith the opposite direction of effects between the male and female datasets. For each dataset, genes were ranked according to the value of the log of the fold change multiplied by the negative base 10 logarithm of the uncorrected *p*-value from differential expression analysis. **c** RRHO2 plots similar to (**b**) but for oligodendrocyte lineage clusters. For RRHO2 plots comparing broad cell types the color scale maximum was set to a −log10(*p*-value) of 50, and for RRHO2 plots comparing clusters the color scale maximum was set to a −log10(*p*-value) of 25 for ease of comparison. RRHO2 uses one-sided hypergeometric tests, the p-values plotted here are uncorrected. Source data are provided as a Source Data file.

meta-analyzed (21 versus 7 DEGs in the male dataset alone). The decrease in number of MDD-associated DEGs in OPCs when combining the male and female cohorts indicates that gene expression differences in OPCs in MDD are dissimilar between the sexes. This agrees with the discordance of depression-related transcriptomic changes between sexes in OPCs in our RRHO2 analysis.

At the cluster level, we found that upregulated DEGs in Mic1 and downregulated DEGs in ExN10_L46 stood out as the top findings in the meta-analysis (Fig. 4, Supplementary Data 7). Once again, we found more microglial DEGs (128 DEGs) via the meta-analysis compared to

the female data alone (68 DEGs) and more DEGs in ExN10_L46 (254 DEGs) than with the male data alone (238 DEGs).

Given the overall between-sex concordance in MDD-associated gene expression changes detected in RRHO2, it is not surprising that clusters with prominent differential expression from the individual cohorts also stood out in the meta-analysis. Taken together these results further support that the global patterns of change in gene expression within cell type are generally consistent between males and females, especially for excitatory neurons and microglia, with a few notable exceptions such as OPCs.

a. Male

c. Female

b.

d.

e.

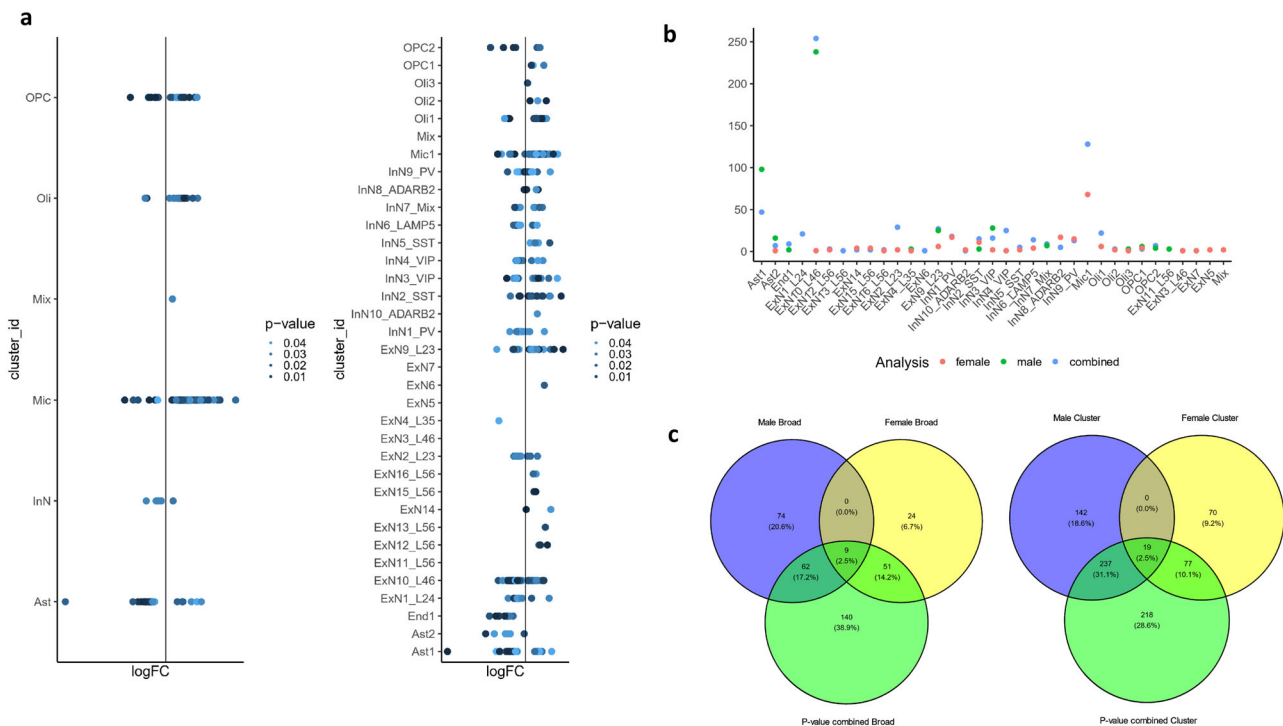f. Mic1

g. InN1_PV

InN9_PV

InN8_ADARB2

InN2_SST

Female cell type specific DEGs are enriched for previous MDD-linked genes

The relevance of the DEGs we have identified to psychiatric disorders was evaluated by referring to the PsyGeNET[25] text-mining database. Compared to other disorders, depressive disorders had the most gene-disease associations with the female cell type specific DEGs (>60; Fig. 5a). The next largest number of gene-disease associations

was for schizophrenia (<40). Statistically, the overlap of all DEGs at the cluster level with disease-associated genes in PsyGeNET was significant only for two disease categories, Depressive disorders (hypergeometric test, p = 0.0378) and Alcohol use disorders (hypergeometric test, p = 0.0141). Further, for the top 5 clusters with highest numbers of DEGs in the female cluster-level analysis, gene-disease associations for depression and related disorders in

Fig. 3 | Cell type specific differential gene expression in males and females with MDD. a, b Distribution of differentially expressed genes in (**a**) broad cell types and (**b**) clusters with increased and decreased expression in male cases compared to controls. **c, d** Distribution of differentially expressed genes in (**c**) broad cell types and (**d**) clusters with increased and decreased expression in female cases compared to controls. For **a**–**d**, points are colored by the corrected *p*-value for differential expression, and upregulated genes are plotted to the right of the midline while downregulated genes are plotted to the left. **e** Barplots showing proportions of up and downregulated genes and unique and shared genes. For males, the majority of DEGs were decreased in expression in cases compared to controls both at the broad (110/151, 73%) and cluster levels (358/447, 80%) and most DEGs were cell type specific both at the broad (145/151, 96% unique DEGs) and cluster (398/447, 89%

unique DEGs) level. For females, the majority of DEGs were upregulated both at the broad (70/85, 82%) and the cluster level (140/180, 78%) and most DEGs were cell type specific both at the broad (84/85, 99% unique DEGs) and cluster (166/180, 92% unique DEGs) level. **f, g** Heatmaps showing the pseudobulk expression of differentially expressed genes in top clusters with highest number of DEGs in the female cluster level analysis—**f** microglia, **g** inhibitory neuronal clusters. For **f, g**, the plotted values are pseudobulk CPMs (counts per million) calculated with edgeR and muscat and scaled per row (by gene). For all heatmaps (**f, g**), the annotation bar at the top is colored orange for cases and purple for controls, and rows and columns are not clustered. Statistical testing corresponding to Fig. 3a–d were performed with the edgeR (glmQLFit, glmQLFtest), FDR (Benjamini & Hochberg) corrected p-values are plotted. Source data are provided as a Source Data file.

Fig. 4 | p-value combination meta-analysis results. a, b Distribution of DEGs across the (left) broad cell types and (right) clusters after *p*-value combination meta-analysis. **b** Numbers of DEGs (*y*-axis) in each cluster for the male analysis, female analysis, and meta-analysis. **c, d** Overlap of meta-analysis DEGs with the

individual analyses of the male and female datasets for (**c**) broad cell types and (**d**) clusters. The statistical test performed is Fisher combination of *p*-values as implemented in metaRNAseq, the test is one-sided, and the *p*-values were Benjamini Hochberg corrected. Source data are provided as a Source Data file.

PsyGeNET were identified for several DEGs (Fig. 5b, c). Therefore, our cell type specific DEG findings in females recapitulated previously reported gene-disease associations.

Disease-relevant biological pathways revealed by cell type specific transcriptomic changes in females with MDD

To explore the underlying pathways associated with the cell type specific transcriptomic changes in females with MDD, we performed pre-ranked gene set enrichment analysis (GSEA; Methods: Pre-ranked gene set enrichment analysis). Female microglia from cases showed significant negative enrichment scores for inflammation-related Reactome pathway gene sets including "Interferon Gamma signaling", "Interleukin 4 and Interleukin 13 signaling", "Interleukin 10 signaling", and "TNFR2 non-canonical NF-KB pathway" (Fig. 5d, Supplementary Data 8). "Neuronal system" gene sets were positively enriched with contributions from "Voltage-gated potassium channels", "Class C/3 metabotropic glutamate/pheromone receptors", and "Neurexins and neuroligins" among others (Fig. 5d). Interestingly both

pro- and anti-inflammatory immune signaling pathway gene sets were downregulated which may indicate that MDD-associated dysregulation of gene expression in microglia involves more than just a microglial inflammatory response.

Further, both PV interneuron clusters showed a negative enrichment of heat shock factor 1 (HSF1) related terms—"HSF1 activation" in InN9_PV and "HSF1 dependent transactivation" in InN1_PV. Moreover, both clusters showed an enrichment of the gene sets "Cellular response to external stimuli" and "Metabolism of RNA". The InN1_PV cluster showed further enrichment of immune gene sets such as "Innate immune system", "Adaptive immune system", and "Cytokine signaling in immune system" and interestingly in the context of sex differences in depression, "ESR mediated signaling", pertaining to the estrogen receptor.

Thus, our GSEA of the female microglia and PV interneuron differential expression results revealed dysregulated Reactome pathway gene sets which are functionally relevant in these cell types and plausibly associated with sex differences.
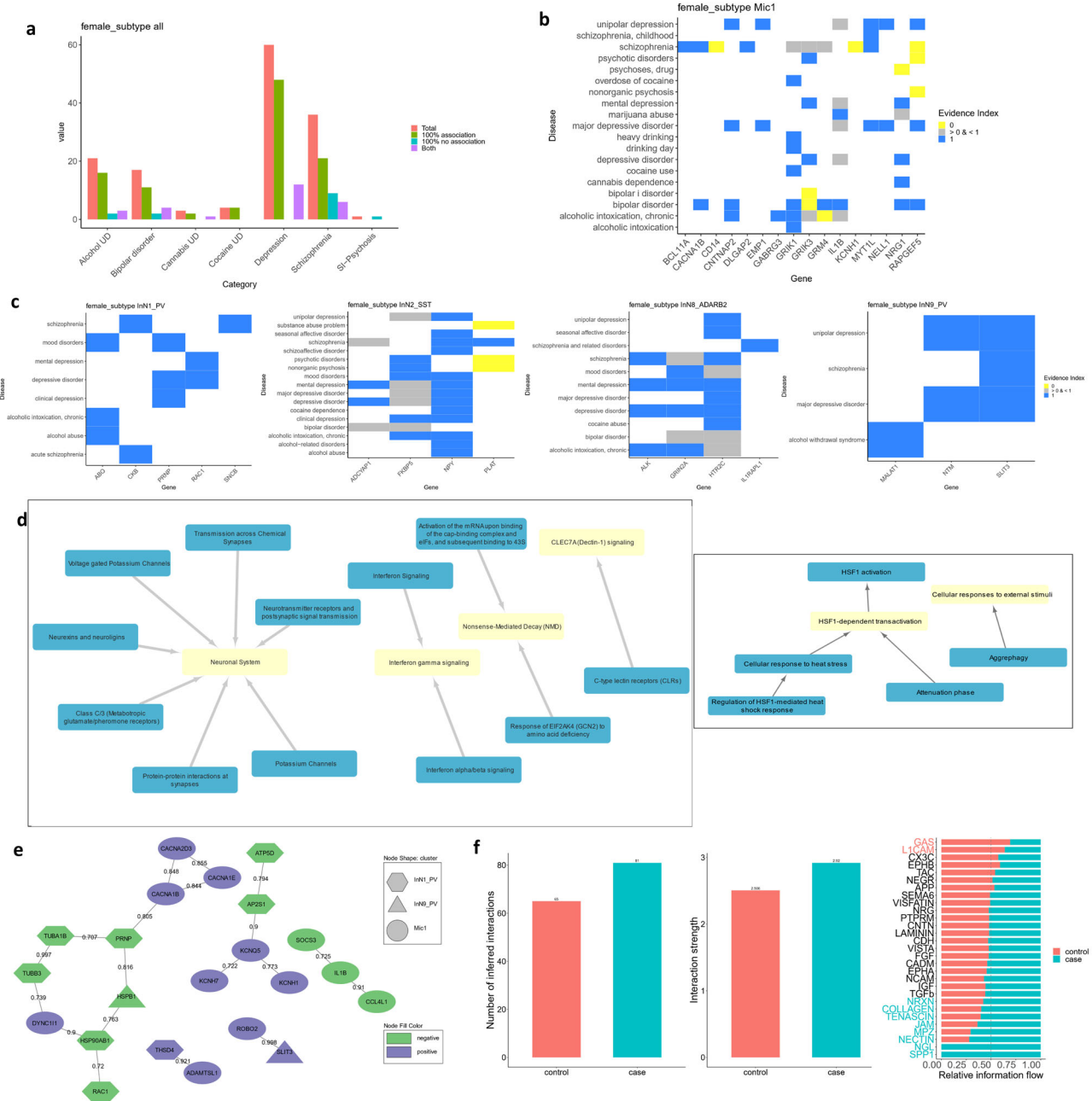
Fig. 5 | Characterization of cell type specific DEGs in females with MDD.
a PsyGeNET literature reported gene-disease association bar plot for all DEGs in the female cluster level analysis. The *y*-axis shows the number of gene-disease associations. "100% association" indicates all evidence is in support, "100% no association" indicates the opposite, while "Both" indicates mixed support. **b**, **c** Gene-disease association heatmaps for 5 clusters with the highest numbers of DEGs in females: **b** microglia, **c** inhibitory neuron clusters. Evidence index of 1 indicates that all literature supports the association, while 0 indicates that there is no support for the association. Values in between indicate partial support. **d** Networks showing the relationship between main gene sets (yellow) and all gene sets (blue) with enrichment in pre-ranked GSEA with Reactome pathways in Mic1 (left) and InN9_PV (right)

in females. Controlling for the overlap between gene sets, the main gene sets are independently enriched. **e** STRING network showing DEGs in female microglia and PV interneurons whose protein products have reported interactions. The shape of the node represents the cluster in which the DEG was detected, and the color represents the direction of fold change in cases compared to controls. The numbers on the edges represent the confidence scores for the interactions. **f** (left) Bar plots showing the number and strength of ligand-receptor communications within and between PV interneurons and microglia in cases and controls. (right) Relative strength of communication in different signaling pathways for cases and controls. Source data are provided as a Source Data file.

Assessing the relationship between microglia and PV interneuron dysregulation in females with MDD using protein–protein interaction assessment

To further assess the functional relevance of striking gene expression differences in microglia and PV interneurons in females with MDD, we examined whether the protein products of DEGs in these clusters belonged to interacting networks. STRING[26] protein-protein

interaction (PPI) analysis (Methods: STRING analysis) revealed links between the protein products of several DEGs in the microglia and the PV interneurons. We focused on the top two interactions, based on the STRING confidence score. These interactions were between protein products of DEGs coming from microglia and PV interneurons and with the same direction of change (Fig. 5e). The *ROBO2* gene, which encodes a canonical cell migration guidance receptor[27],

was increased in microglia whereas one of its corresponding ligands, *SLIT3*[27] was increased in expression in the InN9_PV cluster. In addition, *ADAMSTL1* and *THSD4* (also known as *ADAMTSL6*), two members of the ADAMTS-like family of proteins, which have extracellular matrix (ECM) binding properties[28], were upregulated in microglia and in the InN1_PV cluster respectively. The PPI network analysis results point to the intriguing possibility that changes in communication between microglia and PV interneurons through the ECM and cell surface molecules contribute to depression-associated brain pathology in females.

Assessing the relationship between microglia and PV interneuron dysregulation in females with MDD using ligand-receptor interaction assessment

Building upon the indications from PPI assessment we explored the possible changes in ligand-receptor expression in microglia and PV interneurons between female cases and controls with CellChat[29] (Methods: CellChat analysis). CellChat identified more interacting ligand-receptor pairs and estimated increased communication strength overall within and between and within microglia and PV interneurons in cases compared to controls (Fig. 5f). CellChat further identified several signaling pathways (groups of related ligand-receptor pairs) with decreased (top pathway: GAS) and increased (top pathway: SPP1) communication in cases compared to controls (Fig. 5f). Within these top signaling pathways, we specifically identified a probable increase in SPP1 to integrin communication and decrease in GAS6-MERTK communication from microglia to PV interneurons and vice versa, respectively (Supplementary Fig. 13).

WGCNA confirms MDD dysregulated pathways in female microglia and PV interneurons

Next, we performed weighted-gene co-expression network analysis (WGCNA) using the pseudobulk gene expression profiles to identify correlated modules of genes associated with MDD in microglia and PV interneurons in females.

In microglia, 8 modules out of 44 had a significant correlation with case-control status (p-value < 0.05; Fig. 6a). Further, the MEturquoise module which is positively correlated with MDD-status (correlation 0.627, $p = 7.26 \times 10^{-5}$) showed a significant overlap ($p = 5 \times 10^{-56}$; Methods: Weighted gene co-expression network analysis) with upregulated DEGs in microglia in female cases (Fig. 6b). MEturquoise also showed an enrichment of Reactome pathway gene sets related to ion channels, neurotransmitter receptors, and the neuronal system (Fig. 6c) similar to gene sets found upregulated in microglia in female cases by GSEA.

In PV interneurons (including nuclei in the InN1_PV and InN9_PV clusters), 16 of 55 modules were significantly associated with case-control status ($p < 0.05$, Fig. 6d). In addition, downregulated DEGs from InN1_PV and InN9_PV significantly overlapped ($p = 3.24 \times 10^{-7}$) with the genes from the MEturquoise module (Fig. 6e). The MEturquoise module which is negatively associated with MDD (correlation -0.582, $p = 0.00016$), had over-representation of 489 Reactome pathway gene sets. Of these, 30 pathways overlapped with the main downregulated pathways previously identified with GSEA in InN1_PV or InN9_PV (Fig. 6f). The overlapping pathways included "HSF1 activation", "HSF1 dependent transactivation", and "ESR mediated signaling". Further, upregulated DEGs in InN1_PV and InN9_PV significantly overlapped with the genes of two modules which had a positive association with MDD-status: MEred (correlation 0.568, $p = 0.0002$) and MEgreenyellow (correlation 0.426, $p = 0.0085$).

Overall, the female microglia and PV interneuron WGCNA results further support our MDD-associated DEG and Reactome Pathway findings in these clusters.

Discussion

Cell type specificity of depression-associated transcriptomic changes

There is a sizable body of postmortem literature describing differences from cellular morphology to proteomic and transcriptomic profiles in individuals with depression. Classic cytological experiments from the turn of the century identified abnormalities in morphology and distribution of cell types, but also put into question cell number, size, and neuropil density, particularly for neurons and astrocytes[30–34]. Transcriptomic studies have to some extent implicated all broad cell types[35–37]. Results from our present and previous study confirm this implication of multiple cell types including excitatory and inhibitory neurons, astrocytes, OPCs, and microglia.

Our study highlights potential cell type specific transcriptomic targets for treatment and intervention in MDD. Given that different cell types appear to be implicated in MDD in males and females, approaches to treatment may need to be different as well. At the very least, our findings strengthen the evidence in support of including female subjects in pre-clinical and clinical research, which had been historically neglected and continues to be neglected in biomedical research.

Sex-specificity of depression-associated transcriptomic differences

Only recently have postmortem transcriptomic studies of MDD begun to incorporate sex as a biological factor[7,10]. These studies, which analyzed bulk tissue samples, reported distinct gene expression differences in males and females with very limited overlap of DEGs. Our findings are consistent with these studies in that the cell types with most prominent differences in gene expression—and the DEGs within these cell types—were quite separate for males and females (Fig. 2a). However, in contrast, we found that within each cluster and broad cell type the threshold-free patterns of MDD-associated difference in gene expression were highly concordant between the sexes in most cases, except for most oligodendrocyte lineage clusters (Fig. 2b, c). This overall agreement between the sexes was confirmed by a meta-analysis of the male and female data (Fig. 4a–c). Possibly, by using single-nucleus methodology, our results provided better resolution for threshold free analyses.

Notably, recent reviews on the sex specificity of transcriptomic differences in MDD[22,38] suggest that in females with MDD there is reduced microglial activation and increased synaptic connectivity while the opposite is true for males. This theory is supported by the downregulation we observed in MDD females in microglial inflammatory pathways such as interferon and NF-KB signaling.

Immune response is innately different across sexes leading to inflammatory responses that vary with age and sex resulting in a bias in susceptibility to the development of diseases from autoimmune to infections to cancer[39]. Microglia, the resident immune cell of the brain, showed the most significant difference in gene expression compared to control subjects specifically in females and not males (Fig. 3a–d). It has been hypothesized that these differences result from distinct starting points between sexes[6]. Notably, many microglial immune functions are mediated by gonadal hormones including transcriptional regulators such as suppressor of cytokine signaling 3, *SOCS3*[40], which is downregulated in our data and influences the expression of other cytokines.

Microglial contributions to MDD in females

The salient implication of microglia specifically in females is consistent with differences in the distribution, structure, transcriptome, and proteome of microglia between the sexes, in both health and disease[41–43]. Furthermore, the number and phenotype of microglia differ by sex in the rodent brain[44–46], and several recent rodent studies demonstrated sex-specificity of microglial response to stress in various brain regions[47–49]. These studies describe changes in genes involved in
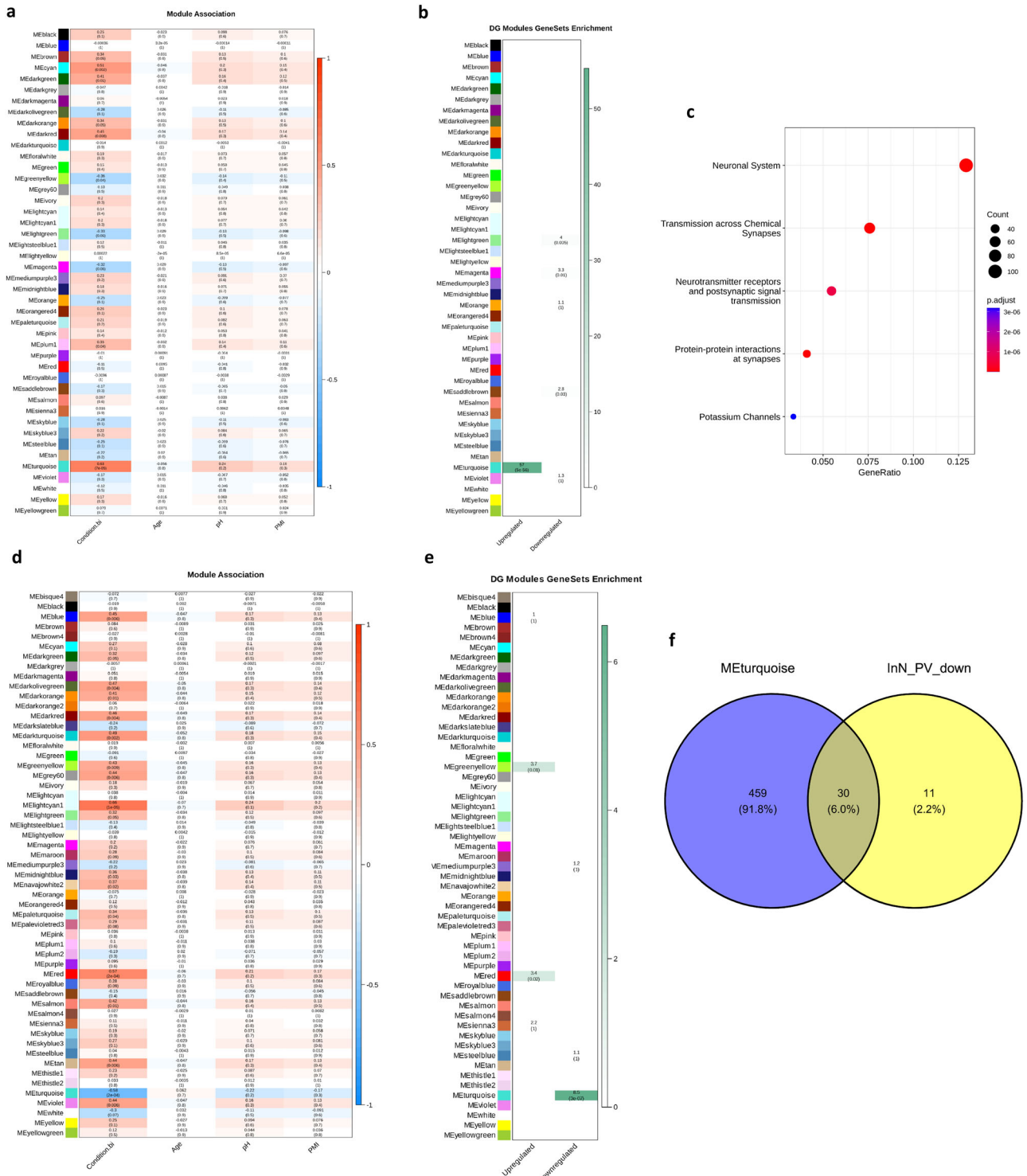
Fig. 6 | WGCNA results for microglia and PV interneurons in females. a Heatmap showing the correlation and associated *p*-value, in parentheses, of Mic1 WGCNA module eigengenes with case-control status and covariates (age, pH, PMI). **b** Heatmap showing the test-statistic and FDR corrected *p*-value, in parentheses, for one-sided Fisher tests of overlap between the Mic1 WGCNA module member genes and DEGs in females in Mic1. **c** Top Reactome pathway gene sets over-represented in Mic1 WGCNA, in the MEturqouise module using one-tailed hypergeometric testing. Uncorrected *p*-values are plotted. **d** Heatmap showing the correlation and associated *p*-value, in parentheses, of InN_PV WGCNA module eigengenes with case-control status and covariates.

case-control status and covariates. **e** Heatmap showing the test-statistic and FDR corrected *p*-value, in parentheses, for one-sided Fisher tests of overlap between the InN_PV WGCNA module member genes and DEGs in females the InN1_PV or InN9_PV clusters. **f** Venn diagram showing the overlap of Reactome pathway gene sets enriched in the InN_PV WGCNA module MEturquoise (associated negatively with case status) and downregulated via GSEA in cases within InN1_PV or InN9_PV. For 6a, 6d the statistical test performed was a Pearson correlation as implemented in the WGCNA package and *p*-values are one-sided and uncorrected. Source data are provided as a Source Data file.

cellular stress and immune function with brain-region and sex-specific variation. This is roughly analogous to the female-specific pathway dysregulation we observed in microglia and PV interneurons in MDD (Fig. 5d, Supplementary Data 8).

Most studies examining peripheral markers report increased inflammation in MDD[50], but studies in brain tissue have reported increases[50], decreases[51], or changes in both directions[52] in the expression of pro-inflammatory molecules. Moreover, several depression-linked genetic variants in pro-inflammatory genes, including in *IL1B*, *TNFA*, and *CRP*, are associated with decreased expression[53]. Recently the concept of a pro-inflammatory versus anti-inflammatory state of microglia has been challenged[54]. In the brain, amidst close interactions with multiple cell types, microglia adopt more diverse states with varying levels of pro- and anti-inflammatory markers, and this is being underscored by single-cell data[54]. Our results reflect altered microglial transcription in MDD females versus controls, with pro-inflammatory (interferon and NF-KB signaling) and anti-inflammatory (IL4, IL13, and IL10 signaling) pathways simultaneously downregulated (Supplementary Data 8).

We observed evidence that further "neuronal" pathways—including neurotransmitter signaling and ion channels—were upregulated in female MDD microglia, in both differential expression and WGCNA results (Supplementary Data 8, Fig. 6c). Microglia have long been known to express neurotransmitter receptors and ion channels. Mounting evidence suggests these canonically "neuronal" gene products regulate microglial activity[55–58], and our results suggest that changes in their expression may contribute to MDD pathophysiology, at least in females.

PV interneuron and microglia crosstalk in females with MDD

Together with striking changes in microglial gene expression, we observed dysregulation in PV interneurons. PV interneurons, among other interneuron subtypes, are implicated in stress and depression with evidence for sex-specific changes[59,60]. Most PV interneurons are encapsulated by ECM structures called perineuronal nets (PNNs) which help protect them from cellular stress, and microglia are known to regulate PNNs[61]. Oxidative and cellular stress relate to PV neuron and PNN deficits in animal models[62] and cellular stress may be part of the molecular pathology in MDD[63].

We found evidence of dysregulated cellular stress pathways, such as heat-shock factor activation, in PV interneurons in MDD females via differential expression analysis and WGCNA (Supplementary Data 8, Fig. 6f). Moreover, both analyses pointed to dysregulation of estrogen receptor mediated signaling. The expression of many genes is regulated by the ligand-bound estrogen receptor and difference in estrogen levels are known to contribute to differences in brain physiology between the sexes[41].

Beyond effects in individual cell types, our results imply potentially impaired communication between PV interneurons and microglia in females with MDD (Fig. 5e, f). Microglial synaptic regulation involves migration of microglia towards specific neurons and in glioblastomas this migration can be regulated by SLIT-ROBO signaling[27]. The *SLIT3* gene and its corresponding receptor gene *ROBO2* were upregulated in PV interneurons and microglia respectively in females with MDD. Of note, genetic variation in *SLIT3* has been associated with depression[64].

We observed that the ECM-binding protein genes *ADAMTSL1* and *THSD4* were upregulated in microglia and PV interneurons, respectively, in MDD females. These recently characterized ADAMTS-like proteins lack the enzymatic domains through which ADAMTSs break PNN components, but they have been proposed to protect these components from degradation by mimicking ADAMTS binding[28,65]. We therefore conjecture that microglial migration cued by PV interneurons, followed by concerted alterations of the ECM by these two cell types stabilize PNNs in females with MDD. A recent study—

including males and females—reported increased PNN number in the PFC of MDD subjects who experienced early life adversity[66], and our molecular findings might underlie one sex-specific mechanism for PNN alterations in MDD.

Our preliminary assessment also points to downregulation of PV interneuron to microglia signaling via GAS6-MERTK and upregulation of SPP1 to integrin signaling in the opposite direction in females with MDD. Together, MERTK and GAS6 promote homeostasis and neuronal survival and they are disrupted in several nervous system disorders[67]. On the other hand, microglial osteopontin (SPP1), promotes remyelination in multiple sclerosis and is neuroprotective near infarcts in stroke but in Alzheimer's disease it is part of the "disease-associated microglia" signature[68]. The role of these signaling molecules in depression, if any, are yet to be determined.

Limitations

This study has limitations that should be considered. We could not directly compare male and female cell type specific transcriptomes or assess the interaction of sex and disease status given that we are using data from two sex-specific datasets. Thus, the implication of different cell types in MDD between males and females could be partly attributable to differences in methodology (such as library preparation chemistry, tissue collection approach, or nuclei isolation protocol, among other factors) for generating the two datasets. However, we attempted to mitigate this by applying a unified pre-processing pipeline and joint definition of cell types. Our findings are consistent with previous evidence for sex-specific mechanisms for depression etiology in animal models and human studies[6,9,69,70].

Our permutation analysis indicated our DEGs at the cluster level for females may not be as robust as for the male cluster level analysis and the broad analysis for both sexes. However, our main findings in females at the cluster level are in microglia, and 78% of microglial DEGs in the female cluster analysis are also present in the female broad analysis, results that were robust, according to the permutation approach. Further our DEGs from the female cluster level analysis were supported by our WGCNA results, partially mitigating the concern that the DEGs can be an artefact of the differential expression analysis strategy.

Although our study included data from over 160,000 nuclei, the number of subjects was small relative to the large number of genes tested for associations with MDD. The relatively small number of subjects included in this study limits our statistical power to detect cell type specific disease-relevant genes and pathways. Further, our results may not be generalizable to all populations and this work will need to be extended with larger sample sizes from diverse populations. However, the number of subjects included in our study compares favorably to most published snRNA-seq studies of neuropsychiatric conditions to date, which have included anywhere between 11 and 48 subjects[11,12,15,71,72].

We did not identify a separate sub-population of disease-associated microglia as observed in some neurological disorders[73]. This may partly be due to the lack of cytoplasmic transcripts in snRNA-seq limiting the information about microglial states[74]. Nevertheless, a recent study highlighted similarities between cellular and nuclear microglia RNA-seq data from mouse and human—fresh and frozen— CNS samples[75]. Nuclear microglia transcriptomes are a reliable proxy for cellular transcriptomes and are less affected by cell isolation-based transcriptional artifacts[75]. We were able to detect inflammatory pathway dysregulation in female microglia despite the limitations.

Our CellChat and STRING results are speculative. We cannot draw conclusions about the proximity of microglia to PV interneurons or the presence of PNNs, as snRNA-seq involves dissociation of the tissue with loss of spatial and structural information. Neither can we conclude that protein expression is changed for our DEGs. Future studies using

spatial transcriptomic techniques coupled with immunohistochemistry may better answer these questions.

Lastly, a few clusters may be of lower quality (biased by batch or according to quality parameters, and inconsistency with other datasets or with cluster-enriched genes; ExN17, ExN5, and Mix). However, given that these clusters did not contribute substantially to our differential expression results, their impact on our main conclusions is likely to be limited.

Outcomes

We provide a cell type and sex-specific assessment of transcriptomic changes in the dlPFC in MDD using snRNA-seq. Our dataset represents a rich resource which will stimulate further fruitful investigations of sex- and cell type specific molecular pathways in depression. While most transcriptomic changes in males with MDD are observed in deep layer excitatory neurons, astrocytes, and OPCs, in females the changes are concentrated in microglia and PV interneurons. Although major dysregulated cell types and genes are distinct for each sex, within broad cell types and clusters the patterns of transcriptomic differences in MDD are primarily concordant between males and females. Finally, preliminary evidence hints that in females with MDD, impaired communication between microglia and PV interneurons may be an important feature of MDD molecular pathology.

Methods

Male snRNA-seq dataset

We used published snRNA-seq data from a cohort of male subjects with or without MDD[16]. We started with the raw FASTQ files available through GEO (GSE144136) and reprocessed the data, dropping two runs from one subject (number 25) with low quality results based on the previous analysis. All male samples for the study had been obtained from the Douglas Bell-Canada Brain Bank.

Post-mortem brain samples in the female cohort

This study was approved by the Douglas Institute IRB. Human post-mortem dlPFC tissue was obtained from the Douglas- Bell-Canada Brain Bank (www.douglasbrainbank.ca, all female case samples and eight female control samples) and from the University of Miami Miller School of Medicine Brain Endowment Bank (https://med.miami.edu/programs/brain-endowment-bank, ten female control samples). Informed consent from next of kin was obtained for each individual included in this study. Frozen histological grade samples of gray and white matter were dissected from the dlPFC (Brodmann Area 9) by expert neuroanatomists and stored at −80 °C. Psychological autopsies were performed using proxy-based interviews complemented by medical charts, as previously described[76]. A summary of sample demographic characteristics is provided in Table 1. All cases included in this study died while affected by MDD or unspecified depressive disorder, whereas controls were neurotypical individuals who died suddenly without prolonged agonal periods and did not have evidence of axis I disorders. The post-mortem interval (PMI) represents the delay between an individual's death and collection and processing of the brain. One female case subject and three female control subjects were Hispanic, two female control subjects were African American, and race information was missing for one female case. All other female subjects were Caucasian, as were all subjects in the male cohort.

Nuclei extraction, single-nuclei capture, and library preparation for female cohort

Nuclei were extracted from coronal cryosections or tissue shavings across the cortical layers and white matter, weighing between 40 and 65 mg, obtained using a cryostat at −20 °C with thickness set to 100 microns. Nuclei were extracted as previously described[77]. Two versions of the iodixanol gradient were used—a weaker gradient using 17.5% and 15% (w/v) concentrations of iodixanol (batches 3F, 7F, 2F) and a

stronger gradient using the 29% and 25% (w/v) concentrations of iodixanol (batches 6F, 8F, 12F), as previously published[78], and we found the stronger gradient to perform better. Nuclei were resuspended in wash buffer and stained using Hoescht 33342 (1:2000). 10 uL of nuclei were loaded onto EVE cell counting slides (MBI) and imaged using an Olympus VS120 Slide Scanner (10X magnification) and counted using the QuPath[79] software (version 0.2.0) with the "Watershed cell detection" functionality.

We used the 10x Genomics Chromium controller for single-cell gene expression to isolate single nuclei for downstream RNA library preparation with 10x Genomics Chromium Single Cell 3' reagents. For samples processed with version 2 of the Chromium chemistry (Supplementary Data 1), we followed the protocols as outlined by the user guide (CG00052_SingleCell3_ReagentKitv2UserGuide_RevB; latest version at https://bit.ly/3dUNOLZ), whereas for sample processed with version 3 of the Chromium chemistry (Supplementary Data 1) we followed the protocols as outlined by the user guide (CG000204_ChromiumNextGEMSingleCell3_v3.1_Rev_D, https://assets.ctfassets.net/an68im79xiti/1eX2FPdpeCgnCJtw4fj9Hx/7cb84edaa9eca04b607f9193162994de/CG000204_ChromiumNextGEMSingleCell3_v3.1_Rev_D.pdf). The catalog numbers for the 10X Genomics single-cell RNA-seq kits for the v2 chemistry and v3 chemistry were 120237 and 1000121, respectively. The only modification was for loading concentration, which we increased by 30% as we assessed the capture of nuclei to be slightly less efficient than cell encapsulation. Nuclei were loaded to capture 3000 per sample, but because of a systematic error in counting the actual number of nuclei captured per sample was variable (Supplementary Data 1).

Sequencing, alignment, and generation of count matrices

The majority of samples in the female cohort (36) were sequenced using the Illumina NovaSeq 6000 but two samples were sequenced using BGI DNB-seq technology. Sequencing metrics are provided in Supplementary Data 1. All samples from the male cohort were realigned. Alignment was performed and count matrices were generated with Cell Ranger version 5.0.1 against the GRCh38 reference available on the 10X Genomics website (refdata-gex-GRCh38-2020-A, https://support.10xgenomics.com/single-cell-gene-expression/software/release-notes/build). We ran the "cellranger count" command using the "−include-introns" option and all other options set to default.

An initial 174,178 nuclei were obtained with Cell Ranger default cell filtering. The median value of mean reads per cell was 71,279, the average mapping rate to the transcriptome was 68.8%, the average fraction of reads in cells was 71%, and the average sequencing saturation was 78.5% (Supplementary Data 1). There was higher intronic mapping rate (Kruskal−Wallis test p-value 0.0029) and a lower exonic mapping rate (Kruskal−Wallis test p-value 0.0048) for cases compared to controls, but no significant differences in any other sequencing quality control metrics (Supplementary Data 1).

The filtered gene barcode matrices were individually loaded into R[80] (versions 4.0.2 and 4.1.2) for downstream analysis and processed with Seurat[81] (4.0.3.9000 and 4.0.5). Percentage of reads from mitochondrially encoded genes were calculated before filtering, added as metadata, and used as a quality control parameter for nuclei filtering, after which the mitochondrial genes were removed for downstream analysis. The parameters for filtering were as follows:

Male cohort: nCount_RNA < 35000, nFeature_RNA > 350, percent.mt < 10

Female cohort v2 chemistry: nCount_RNA < 25000, nFeature_RNA > 250, percent.mt < 10

Female cohort v3 chemistry: nCount_RNA < 120000, nFeature_RNA > 350, percent.mt < 10

After filtering, we obtained 79,058 nuclei in the male cohort (43,347 from cases, 35,711 from controls) and 81,653 nuclei in the

female cohort (49,926 from cases, 31,727 from controls). In the female cohort, after filtering, the median across samples of the median number of UMIs per cell and the median the number of genes per cell were 2758.5 and 1711.5 respectively (Supplementary Data 1). In the males, the corresponding numbers were 2530.5 and 1638.25 respectively (Supplementary Data 1).

Dimensionality reduction and data integration

We performed SCTransform on each Seurat object individually and used the SelectIntegrationFeatures function to set the variable genes for downstream analysis. We scaled each cell to 10,000 counts and ran log normalization. We regressed out nCount_RNA and percent.mt from the counts to get scaled gene expression values for variable genes, which was used as input for calculating 100 PCA components. We corrected PCA components with Harmony[17] to account for batch, chemistry, and sample specific effects. This helped align the datasets as seen in the UMAP projections produced before and after correction (Supplementary Fig. 1a–d). All UMAPs in figures were created using Seurat.

Clustering

We tested of a range of combinations of clustering parameters for the Seurat package (FindClusters function) using the scclusteval[18] subsampling (80% of all cells, 100 times) and stability comparison workflow using Jaccard indices, with some customization. With each subsampling, PCA and Harmony were recalculated. The parameters tested were: k-param: 20, 30; resolution: 0.1, 0.3, 0.5, 0.7, 0.9, 1.1, 1.3, 1.5; number of Harmony corrected PCs to use: 70, 80.

We then set a threshold for the minimum stability with a chooseR-like[82] approach based on the bootstrapped medians of the median Jaccard index across all the clusters and all the parameter sets tested. We selected parameters that maximized the number of clusters while passing the threshold of cluster stability: 70 Harmony corrected PCA components, a k-nearest neighbors' parameter of 30, and a resolution of 0.7 (Supplementary Fig. 2a, b). Repeating the Harmony correction with a seed set followed by clustering with the optimal parameters produced 41 clusters. Final UMAPs were produced using all 100 Harmony corrected PCA components and all calculation parameters set to default.

Cluster annotation

Genes enriched in clusters were calculated using the wilcoxauc function from presto[83] with default parameters, and filtered with the following criteria: padj < 0.05, logFC > log(1.5), pct_in-pct_out > 10. For annotation, the following known cell type marker genes were assessed in the cluster enriched genes:

Macrophage/microglia: *SPI1, MRC1, TMEM119, CX3CR1*; Endothelial: *CLDN5, VTN, VIM*; Astrocytes: *GLUL, SOX9, AQP4, GJA1, NDRG2, GFAP, ALDH1A1, ALDH1L1*; OPCs: *PDGFRA, PCDH15, OLIG2, OLIG1*; Oligodendrocytes: *PLP1, MAG, MOG, MOBP, MBP*; Neurons: *SNAP25, RBFOX3*; Excitatory neurons: *SATB2, SLC17A7, SLC17A6*; Inhibitory neurons: *GAD1, GAD2, SLC32A1*, Inhibitory neuronal subtypes: *VIP, PVALB, SST, ADARB2, LHX6, LAMP5, PAX6*.

In addition, the expression of cell type specific genes from BRETIGEA[84] were assessed using the Seurat AddModuleScore function (Supplementary Fig. 4).

Twenty clusters of excitatory cells were identified (Supplementary Fig. 5a) including four superficial cortical layer neuronal clusters (ExN1_L24, ExN2_L23, ExN8_L24, ExN9_L23), ten deep cortical layer neuronal clusters (ExN3_L46, ExN4_L35, ExN10_L46, ExN11_L56, ExN12_L56, ExN13_L56, ExN15_L56, ExN16_L56, ExN19_L56, ExN20_L56) and six excitatory neuronal clusters without an obvious pattern of cortical layer specific marker expression (ExN5, ExN6, ExN7, ExN14, ExN17, ExN18). The layer annotations of excitatory neuronal clusters were supported by assessment of enrichment for genes known to be

specific to the different layers of the cortex using spatial transcriptomics results from Maynard et al.[85] (data in Supplementary Table 4 of the cited publication).

We identified 10 inhibitory clusters (Supplementary Data 2, Supplementary Fig. 5b), that can broadly be divided into cells likely derived from the medial ganglionic eminence (MGE; InN1_PV, InN9_PV, InN2_SST, InN5_SST) based on *LHX6, SST*, or *PVALB* enrichment, or the caudal ganglionic eminence (CGE; InN3_VIP, InN4_VIP, InN6_LAMP5, InN8_ADARB2, InN10_ADARB2) based on *ADARB2* enrichment. The InN2_SST cluster was enriched for *SST* and *GAD1* expression but had no *LHX6* enrichment. The InN8_ADARB2 (also referred to interchangeably as InN8_Mix) cluster also showed enrichment for *SST*. One inhibitory neuron cluster with enrichment for both *ADARB2* and *LHX6* (InN7_Mix), which has been previously reported[86].

Assessment of clustering quality

Contribution of batches, groups, brain banks, and subjects was relatively uniform across clusters (Supplementary Fig. 3a–d, g). Endothelial, microglial, and oligodendrocyte lineage cells showed a higher percentage of contribution from the females compared to the males, possibly due a different dissection strategy used for the two cohorts such that for the female cohort more white matter tissue was included in the nuclei extractions. All but one cluster (number 34, later annotated as ExN17) had contributions from both the male and female cohorts and one cluster was primarily composed of cells from the female cohort (number 11, ExN5). ExN17 also showed exceptionally high numbers of UMIs detected per nucleus (Supplementary Fig. 3f). Moreover, one cluster (number 17, which was later annotated as showing a mixed expression profile—Mix) had relatively high percentage of mitochondrial reads (Supplementary Fig. 3f). These clusters are likely driven by technical effects rather than representing biologically driven cell subtypes or cell states, but they only represented < 6% of our data.

Comparison to other datasets

MetaNeighbor. We used MetaNeighbor[87] to compare the clusters in our dataset to several published datasets[16,19,20]. For the Song et al., 2020 data we used the h5_a[88], h7[89], h10[90], and h14[86] datasets which contain adult human cortical cells or nuclei and were reprocessed by the authors. We used our own dataset as a reference to train the model, for consistency of comparisons across the datasets and limited the analysis to the same variable genes we used for PCA and clustering. MetaNeighbor best hits plots are shown in Fig. 1e and Supplementary Fig. 6.

Spatial label transfer. We used Seurat to transfer the labels for layer annotation from a spatial transcriptomics dataset[85] to our dataset (Supplementary Fig. 5a). Each tissue section of the spatial transcriptomics data was treated separately and one section each from two different subjects were assessed (data shown for one subject and section—151673). Both the spatial and snRNA-seq data were preprocessed with SCTransform and transfer anchors were identified using the "canonical correlation analysis" option before transferring the labels.

Pseudotime trajectory analysis

We used "slingshot"[91] to build a pseudotime trajectory with our OL nuclei (Supplementary Fig. 5c). We built the pseudotime trajectory with the male and female datasets combined. OL nuclei were subset and UMAP was rerun using the following parameters: dims = 1:10, min.dist = 0.1, spread = 5, n.neighbors = 100, chosen to capture the global patterns in the data. Slingshot was run with the resulting UMAP as input and using the following parameters: extend = "n", start.clus = "OPC2", end.clus = "Oli3", stretch = 0.1, thresh = 0.3, once again chosen to capture the broad patterns in the data. The start and end

clusters were chosen based on their position in the UMAP, and cluster labels were provided. The oligodendrocyte lineage (OL) clusters were arranged from OPC2 at one end of the pseudotime trajectory, followed by OPC1 and OPC3, a small cluster possibly corresponding to committed oligodendrocyte precursors (COPs). At the other end of the pseudotime trajectory Oli2, Oli1, and Oli3 were placed sequentially and could represent the order of oligodendrocyte clusters from myelinating to mature states.

We fit the expression of genes along pseudotime by splitting the data for males and females before using tradeSeq[92] (Supplementary Fig. 5d). We ran fitGAM on the UMI counts for each gene, with age, PMI, pH, and batch as covariates, with conditions set to case and control status, and nknots of 5, based on evaluation of a range of nknot values. The fitted expression of OL marker genes was visualized using the plotSmoothers function. The pseudotime trajectory analysis was performed following the vignette available here: https://kstreet13.github.io/bioc2020trajectories/articles/workshopTrajectories.html.

Cell type proportions comparison

The percentage of nuclei in each cluster and each broad cell type for each sample was calculated and compared between cases and controls with Wilcoxon tests using rstatix[93]. To further mitigate the effect of outliers we obtained p-values for the Wilcoxon test using bootstrapping with 10000 replicates (R package boot;[94] Supplementary Data 3) which supported the initial results. Lastly, we also examined the distribution of p-values (Supplementary Data 3) from the Wilcoxon test rerun after randomly sub-sampling 70% of the nuclei 100 times similar to a previous study[11] and confirmed the pattern of changes in proportion preserved after sub-sampling (Supplementary Data 3). Wilcoxon tests were also repeated for the male and female datasets separately (Supplementary Fig. 7) as described above. All boxplots in Fig. 1 and Supplementary Fig. 7 are made using the geom_boxplot function from ggplot2, and the detailed description of the boxplot elements can be found, in the documentation for the function which is linked here: https://ggplot2.tidyverse.org/reference/geom_boxplot.html.

Differential expression analysis

We performed pseudobulk differential gene expression analysis using muscat[95] and edgeR[96] at the broad cell type and cluster levels in males and females separately. Pseudobulk expression profiles were obtained by summing the raw UMI counts for each gene for each sample within the broad cell type or cluster. Only one run of the male sample 24 was included in these analyses (M24_2 excluded). In addition, subjects were only included if they had a minimum of 10 cells in the broad cell type or 5 cells minimum in the cluster. The covariates included age, pH, PMI, and batch and muscat's default gene and sample filtering were disabled. Further, internal checks within muscat excluded clusters where the number of samples with sufficient cells was not enough, given the model being used. DEGs were selected using an FDR (Benjamini & Hochberg) adjusted local (within cluster or broad cell type) p-value < 0.05 and logFC > log2(1.1) and non-zero expression value in at least 3 samples. The isOutlier function (nmads 5, log "FALSE") from scater[97] was used to flag potential outliers on the CPMs from edgeR, as an additional assessment for genes that were called as differentially expressed. Flagged outliers were not removed from analysis. For one female subject with missing pH, F35, the average pH across all female subjects was substituted.

Since the only difference between the broad and cluster level microglial results lies in the exclusion criteria for subjects based on number of cells contributed the input data and outcomes were similar between these analyses, and we focused on the cluster level results in females for follow-up analyses.

All Venn diagrams to show overlap of differentially expressed genes were made with ggvenn (version 0.1.9). Heatmaps for differentially expressed genes were made with ComplexHeatmap (version 2.10.0).

Comparison of male differential expression results to previous results.

For the male differential expression results, using linear regression we compared the log fold changes per gene for top clusters with highest numbers of DEGs from the current analysis with the per gene estimates for similar clusters with high numbers of DEGs in our previous analysis[16]. Considering only the top 1000 genes in common ranked by the p-values in the current analysis, we found moderate positive relationships with R-squared values in the 0.13–0.32 range (Supplementary Fig. 10i–l). Considering that the analysis approaches were quite distinct at every upstream and downstream step, these results support a similar pattern of changes in gene expression in the male data as we had previously reported.

Sub-clustering of microglia for differential expression analysis in females.

A subset of microglia clustered next to oligodendrocytes in the UMAP, which could reflect misclassified cells, doublets, or even immune oligodendroglia[98,99] or white matter microglia[100]. To determine the robustness of our microglial results to the presence of this subset of cells, we sub-clustered the microglial cluster. We found variable features within the microglial population, reran PCA and Harmony, and optimized clustering parameters (resolution 0.01, other parameters default) using silhouette scores. We excluded any subclusters which expressed oligodendrocyte lineage markers (*PLP1* and *ZFPM2*). Then we reran differential expression analysis on female microglia using the same parameters as initially used and found that new per gene logFCs showed a strong positive association (linear regression) with the initial results (Supplementary Fig. 12i). Given that sub-setting microglia to most confident nuclei did not substantially alter the results and for downstream analysis, we proceeded with the DEGs obtained using the full microglial cluster.

Comparison of male and female results

Rank-rank hypergeometric overlap. We performed a threshold free, rank-rank hypergeometric overlap (RRHO) analysis with RRHO2[23]. Within each cluster or broad cell type, genes were scored using the product of the logFC and the negative log base 10 uncorrected p-value from differential expression analysis in the male and female datasets separately. The scored gene lists were provided to RRHO2_initialize function (method "hyper" and log10.ind "TRUE") and the results were plotted using the RRHO2_heatmap function.

Meta-analysis by Fisher combination of *p*-values. To meta-analyze the male and female differential expression results per broad cell type and per cluster, we used Fisher combination of p-values as implemented in the metaRNASeq[101] R package on the uncorrected p-values after filtering out genes detected in <3 samples. We also used an FDR (Benjamini & Hochberg) adjusted p-value threshold of 0.05 for genes to be considered significantly changed in the meta-analysis and removed any genes with opposite direction of change between the two datasets.

Permutation analysis. We permuted the cases versus control labels 100 times, within each batch, within the male and female datasets separately, and re-ran our differential expression analysis to obtain a distribution of the number of cell type specific unique DEGs (counting once any DEGs repeated across multiple clusters or cell types), at the broad and cluster level, for males and females, with randomly permuted groups (Supplementary Fig. 9a–d). We also calculated the Spearman correlation between the differential expression gene scores (log fold change multiplied by the negative log base 10 uncorrected p-value, as used in RRHO2 analysis) between

male and female datasets per cell type and cluster. We plotted the distribution of correlation coefficients obtained between the male and female datasets using permuted case versus control labels for broad cell types (Supplementary Fig. 9e–j). Further, we assessed for each cluster what percentage of Spearman correlation coefficients calculated using permuted results were less than the Spearman correlation coefficient observed with the real labels (Supplementary Data 4), with a higher percentage representing a correlation that is less likely to appear by chance with random case versus control labels.

Functional interpretation of female differential expression results

Pre-ranked gene set enrichment analysis (GSEA). For the microglial (Mic1) and PV interneuron (InN1_PV, InN9_PV) differential expression results we individually performed pre-ranked Gene Set Enrichment Analysis[102] with FGSEA[103] using the same ranking metric as used for RRHO2 (product of log fold change and the negative log base 10 of the uncorrected p-value). We evaluated the Reactome pathway[104] gene sets obtained from msigdbr[105]. The following parameters were used for the fgsea function: eps = 0.0, minSize = 15, maxSize = 1000 and any pathways with Benjamini-Hochberg adjusted p-value < 0.1 were considered to be significant. Finally, we ran collapsedPathways with pval.threshold = 0.01 to get the main pathways for each cluster.

PsyGeNET analysis. With the list of DEGs from the female dataset across all clusters, we ran enrichedPD from psygenet2r[106] with database = "ALL" and other parameters set to default to find the psychiatric disorders for which our DEGs showed an enrichment. Next, we ran psygenetGene with database= "ALL" and other parameters set to default, and created a geneAttrPlot for the evidence index for all DEGs from all clusters in females to summarize the links between our DEGs and psychiatric disorders reported in PsyGeNET[25]. In addition, we similarly ran psygenetGene, individually on the DEGs from Mic1, InN1_PV, InN2_SST, InN9_PV, InN8_ADARB2, and plotted the corresponding gene-disease association heatmaps with plot type = "GDA heatmap".

STRING analysis. We used STRING DB[26] (version 11.5) to assess the relationships between the protein products of our DEGs in female microglia and PV interneurons. The entire list of DEGs from these clusters (Mic1, InN1_PV, and InN9_PV) were provided as input and the confidence level was set to high (interaction score > 0.7). We then exported the network to Cytoscape (3.9.1), colored genes by direction of change in expression, shaped DEG nodes based on their cluster of origin, and labeled the edges with the confidence scores for the interactions.

CellChat analysis. We subset the relevant nuclei from females in Mic1, InN1_PV, and InN9_PV and performed CellChat[29] analysis. We relabeled all PV interneuron nuclei as InN_PV. For cases and controls independently, we sequentially ran identifyOverExpressedGenes and identifyOverExpressedInteractions with lenient default parameters to find the ligand-receptor gene combinations overexpressed in these cell types. Next, we ran computeCommunProb (with nboot = 1000) followed by computeCommunProbPathway, netAnalysis_computeCentrality, and aggregateNet with default parameters to find the ligand-receptor pathways present. Lastly, we merged the case and control objects and ran computeNetSimilarityPairwise with type "functional". Finally, we used the compareInteractions, rankNet, and netVisual_bubble to visualize the results. We used the following vignette for CellChat analysis: https://github.com/sqjin/CellChat/blob/master/tutorial/Comparison_analysis_of_multiple_datasets.html.

Weighted gene co-expression network analysis (WGCNA)
Weighted gene co-expression network analysis (WGCNA) was performed to identify co-expression modules using the snRNA-seq expression data[107]. First, the aggregated expression for each female sample in microglia and PV interneuron clusters (InN1_PV and InN9_PV combined) was calculated by summing the counts per gene across all nuclei. We excluded subjects that did not have at least 5 microglial nuclei or 5 PV interneuron nuclei (InN1_PV and InN9_PV combined). To account for known external sample traits, the counts were corrected for age, pH, PMI, and batch (same as covariates used for differential gene expression analysis) using limma[108]. In addition, lowly expressed genes with total counts of below 5 were removed. A soft thresholding power of 10 and 12, respectively, with a minimum module size of 30 genes, were used for network construction and module detection for microglia and PV interneurons. Each module was correlated with the phenotype (healthy control vs MDD), and significance was determined using a p-value < 0.05.

To further characterize modules correlated with MDD, Fisher tests for overlap were performed to calculate the over-representation of DEGs as described previously[109]. In addition, the functional annotation of modules was determined using Reactome Pathway gene set over-representation analysis provided by clusterprofiler[110].

Reporting summary
Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability
Raw sequencing data (FASTQ files) for the female cohort generated in this study is available on GEO (accession number: GSE213982) along with the processed gene-barcode matrix and metadata including both male and female cohorts. The raw sequencing data for the male cohort are also available on GEO (accession number: GSE144136). Additionally, the processed data from this study are available on the UCSC Cell Browser for easy visualization: https://dlpfc-mdd.cells.ucsc.edu. Source data for all figures in this paper are provided on Zenodo: https://doi.org/10.5281/zenodo.7884086. The reference genome version used is available on the 10X Genomics website (refdata-gex-GRCh38-2020-A, https://support.10xgenomics.com/single-cell-gene-expression/software/release-notes/build). Allen Brain Institute motor cortex data used for MetaNeighbor comparison are available for download here: https://portal.brain-map.org/atlases-and-data/rnaseq/human-m1-10x. STAB reprocessed data from published snRNA-seq and scRNA-seq datasets used for MetaNeighbor comparison is available here: https://mai.fudan.edu.cn/stab/help/. The spatial transcriptomics data used here for label transfer is available using the spatialLIBD (version 1.6.0) R package and through the AWS download links provided here: https://github.com/LieberInstitute/spatialLIBD.

Code availability
All scripts used to analyze data are provided in a Github repository (https://github.com/MgssGroup/snRNASeq_public/). The DOI for the version of the scripts used is: https://zenodo.org/badge/latestdoi/634913347.

References

1. James, S. L. et al. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* **392**, 1789–1858 (2018).
2. American Psychiatric, A. *Diagnostic and Statistical Manual Of Mental Disorders: DSM-5-TR*. (American Psychiatric Association Publishing, Washington, DC, 2022).

3. Howard, D. M. et al. Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nat. Neurosci.* **22**, 343–352 (2019).

4. Gandal, M. J. et al. Shared molecular neuropathology across major psychiatric disorders parallels polygenic overlap. *Science* **359**, 693–697 (2018).

5. Kessler, R. C. & Bromet, E. J. The epidemiology of depression across cultures. *Annu. Rev. Public Health* **34**, 119–138 (2013).

6. Seney, M. L. & Sibille, E. Sex differences in mood disorders: perspectives from humans and rodent models. *Biol. Sex. Differ.* **5**, 17 (2014).

7. Seney, M. L. et al. Opposite molecular signatures of depression in men and women. *Biol. Psychiatry* **84**, 18–27 (2018).

8. Bachmann, S. Epidemiology of suicide and the psychiatric perspective. *Int. J. Environ. Res. Public Health* **15**, 1425 (2018).

9. Labaka, A., Goñi-Balentziaga, O., Lebeña, A. & Pérez-Tejada, J. Biological sex differences in depression: a systematic review. *Biol. Res Nurs.* **20**, 383–392 (2018).

10. Labonté, B. et al. Sex-specific transcriptional signatures in human depression. *Nat. Med.* **23**, 1102–1111 (2017).

11. Morabito, S. et al. Single-nucleus chromatin accessibility and transcriptomic characterization of Alzheimer's disease. *Nat. Genet.* **53**, 1143–1155 (2021).

12. Smajić, S. et al. Single-cell sequencing of human midbrain reveals glial activation and a Parkinson-specific neuronal state. *Brain* **145**, 964–978 (2022).

13. Reiner, B. C. et al. Single-nuclei transcriptomics of schizophrenia prefrontal cortex primarily implicates neuronal subtypes. Preprint at *bioRxiv* https://doi.org/10.1101/2020.07.29.227355 (2020).

14. Ruzicka, W. B. et al. Single-cell dissection of schizophrenia reveals neurodevelopmental-synaptic axis and transcriptional resilience. Preprint at *medRxiv* https://doi.org/10.1101/2020.11.06.20225342 (2020).

15. Velmeshev, D. et al. Single-cell genomics identifies cell type–specific molecular changes in autism. *Science* **364**, 685 (2019).

16. Nagy, C. et al. Single-nucleus transcriptomics of the prefrontal cortex in major depressive disorder implicates oligodendrocyte precursor cells and excitatory neurons. *Nat. Neurosci.* **23**, 771–781 (2020).

17. Korsunsky, I. et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).

18. Tang, M. et al. Evaluating single-cell cluster stability using the Jaccard similarity index. *Bioinformatics* **37**, 2212–2214 (2021).

19. Song, L. et al. STAB: a spatio-temporal cell atlas of the human brain. *Nucleic Acids Res.* **49**, D1029–d1037 (2021).

20. Bakken, T. E. et al. Comparative cellular analysis of motor cortex in human, marmoset and mouse. *Nature* **598**, 111–119 (2021).

21. Batiuk M.Y. et al. Upper cortical layer-driven network impairment in schizophrenia. *Sci Adv.* **8**, eabn8367 (2022).

22. Seney, M. L., Glausier, J. & Sibille, E. Large-scale transcriptomics studies provide insight into sex differences in depression. *Biol. Psychiatry* **91**, 14–24 (2022).

23. 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).

24. Boutej, H. et al. Diverging mRNA and protein networks in activated microglia reveal SRSF3 suppresses translation of highly upregulated innate immune transcripts. *Cell Rep.* **21**, 3220–3233 (2017).

25. Gutiérrez-Sacristán A. et al. Text mining and expert curation to develop a database on psychiatric diseases and their genes. *Database (Oxford)* **2017**, bax043 (2017).

26. Szklarczyk, D. et al. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res.* **49**, D605–d612 (2021).

27. Geraldo, L. H. et al. SLIT2/ROBO signaling in tumor-associated microglia and macrophages drives glioblastoma immunosuppression and vascular dysmorphia. *J. Clin. Investig.* **131**, e141083 (2021).

28. Porter, S., Clark, I. M., Kevorkian, L. & Edwards, D. R. The ADAMTS metalloproteinases. *Biochem. J.* **386**, 15–27 (2005).

29. Jin, S. et al. Inference and analysis of cell-cell communication using CellChat. *Nat. Commun.* **12**, 1088 (2021).

30. Stockmeier, C. A. & Rajkowska, G. Cellular abnormalities in depression: evidence from postmortem brain tissue. *Dialogues Clin. Neurosci.* **6**, 185–197 (2004).

31. Rajkowska, G. et al. Morphometric evidence for neuronal and glial prefrontal cell pathology in major depression. *Biol. Psychiatry* **45**, 1085–1098 (1999).

32. Cotter, D., Mackay, D., Landau, S., Kerwin, R. & Everall, I. Reduced glial cell density and neuronal size in the anterior cingulate cortex in major depressive disorder. *Arch. Gen. Psychiatry* **58**, 545–553 (2001).

33. Cotter, D. et al. Reduced neuronal size and glial cell density in area 9 of the dorsolateral prefrontal cortex in subjects with major depressive disorder. *Cereb. Cortex* **12**, 386–394 (2002).

34. Ongür, D., Drevets, W. C. & Price, J. L. Glial reduction in the subgenual prefrontal cortex in mood disorders. *Proc. Natl Acad. Sci. USA* **95**, 13290–13295 (1998).

35. Sequeira, A. et al. Global brain gene expression analysis links glutamatergic and GABAergic alterations to suicide and major depression. *PLoS ONE* **4**, e6585 (2009).

36. Tripp, A., Kota, R. S., Lewis, D. A. & Sibille, E. Reduced somatostatin in subgenual anterior cingulate cortex in major depression. *Neurobiol. Dis.* **42**, 116–124 (2011).

37. Lutz, P. E. et al. Association of a history of child abuse with impaired myelination in the anterior cingulate cortex: convergent epigenetic, transcriptional, and morphological evidence. *Am. J. Psychiatry* **174**, 1185–1194 (2017).

38. Bollinger, J. L. Uncovering microglial pathways driving sex-specific neurobiological effects in stress and depression. *Brain Behav., Immun. - Health* **16**, 100320 (2021).

39. Klein, S. L. & Flanagan, K. L. Sex differences in immune responses. *Nat. Rev. Immunol.* **16**, 626–638 (2016).

40. Shi, M. et al. Estrogen receptor-regulated SOCS3 modulation via JAK2/STAT3 pathway is involved in BPF-induced M1 polarization of macrophages. *Toxicology* **433-434**, 152404 (2020).

41. Rahimian, R., Cordeau, P. & Kriz, J. Brain response to injuries: when microglia go sexist. *Neuroscience* **405**, 14–23 (2019).

42. Rahimian, R., Belliveau, C., Chen, R. & Mechawar, N. Microglial inflammatory-metabolic pathways and their potential therapeutic implication in major depressive disorder. *Front. Psychiatry* **13**, 871997 (2022).

43. Guneykaya, D. et al. Transcriptional and translational differences of microglia from male and female brains. *Cell Rep.* **24**, 2773–2783.e2776 (2018).

44. Caplan, H. W., Cox, C. S. & Bedi, S. S. Do microglia play a role in sex differences in TBI? *J. Neurosci. Res.* **95**, 509–517 (2017).

45. Schwarz, J. M. & Bilbo, S. D. Sex, glia, and development: interactions in health and disease. *Horm. Behav.* **62**, 243–253 (2012).

46. Lenz, K. M., Nugent, B. M., Haliyur, R. & McCarthy, M. M. Microglia are essential to masculinization of brain and behavior. *J. Neurosci.* **33**, 2761–2772 (2013).

47. Bollinger, J. L., Bergeon Burns, C. M. & Wellman, C. L. Differential effects of stress on microglial cell activation in male and female medial prefrontal cortex. *Brain Behav. Immun.* **52**, 88–97 (2016).

48. Woodburn, S. C., Bollinger, J. L. & Wohleb, E. S. Synaptic and behavioral effects of chronic stress are linked to dynamic and sex-

specific changes in microglia function and astrocyte dystrophy. *Neurobiol. Stress* **14**, 100312 (2021).

49. Bollinger, J. L., Collins, K. E., Patel, R. & Wellman, C. L. Behavioral stress alters corticolimbic microglia in a sex- and brain region-specific manner. *PLoS ONE* **12**, e0187631 (2017).

50. Wang, Q., Roy, B., Turecki, G., Shelton, R. C. & Dwivedi, Y. Role of complex epigenetic switching in tumor necrosis factor-α upregulation in the prefrontal cortex of suicide subjects. *Am. J. Psychiatry* **175**, 262–274 (2018).

51. Morrison, F. G. et al. Reduced interleukin 1A gene expression in the dorsolateral prefrontal cortex of individuals with PTSD and depression. *Neurosci. Lett.* **692**, 204–209 (2019).

52. Mahajan, G. J. et al. Altered neuro-inflammatory gene expression in hippocampus in major depressive disorder. *Prog. Neuro-Psychopharmacol. Biol. Psychiatry* **82**, 177–186 (2018).

53. Barnes, J., Mondelli, V. & Pariante, C. M. Genetic Contributions of Inflammation to Depression. *Neuropsychopharmacology* **42**, 81–98 (2017).

54. Ransohoff, R. M. A polarizing question: do M1 and M2 microglia exist? *Nat. Neurosci.* **19**, 987–991 (2016).

55. Liu, H., Leak, R. K. & Hu, X. Neurotransmitter receptors on microglia. *Stroke Vasc. Neurol.* **1**, 52–58 (2016).

56. Eder, C. Ion channels in microglia (brain macrophages). *Am. J. Physiol.-Cell Physiol.* **275**, C327–C342 (1998).

57. Stolero, N. & Frenkel, D. The dialog between neurons and microglia in Alzheimer's disease: the neurotransmitters view. *J. Neurochem.* **158**, 1412–1424 (2021).

58. Bhandage, A. K. & Barragan, A. GABAergic signaling by cells of the immune system: more the rule than the exception. *Cell. Mol. Life Sci.* **78**, 5667–5679 (2021).

59. Wellman, C. L. et al. Sex differences in risk and resilience: stress effects on the neural substrates of emotion and motivation. *J. Neurosci.* **38**, 9423–9432 (2018).

60. Pesarico, A. P. et al. Chronic stress modulates interneuronal plasticity: effects on PSA-NCAM and perineuronal nets in cortical and extracortical regions. *Front. Cell. Neurosci.* **13**, 197 (2019).

61. Gibel-Russo, R., Benacom, D. & Di Nardo, A. A. Non-cell-autonomous factors implicated in parvalbumin interneuron maturation and critical periods. *Front. Neural Circuits* **16**, 875873 (2022).

62. Steullet, P. et al. Oxidative stress-driven parvalbumin interneuron impairment as a common mechanism in models of schizophrenia. *Mol. Psychiatry* **22**, 936–943 (2017).

63. Kowalczyk, M. et al. Cellular response to unfolded proteins in depression. *Life* **11**, 1376 (2021).

64. Glessner, J. T. et al. Duplication of the SLIT3 locus on 5q35.1 predisposes to major depressive disorder. *PLoS ONE* **5**, e15463 (2010).

65. Hirohata, S. et al. Punctin, a Novel ADAMTS-like molecule, ADAMTSL-1, in extracellular matrix. *J. Biol. Chem.* **277**, 12182–12189 (2002).

66. Tanti, A. et al. Child abuse associates with increased recruitment of perineuronal nets in the ventromedial prefrontal cortex: a possible implication of oligodendrocyte progenitor cells. *Mol. Psychiatry* **27**, 1552–1561 (2022).

67. Burstyn-Cohen, T. & Hochberg, A. TAM signaling in the nervous system. *Brain Plasticity* **7**, 33–46 (2021).

68. Rosmus, D.-D., Lange, C., Ludwig, F., Ajami, B. & Wieghofer, P. The role of osteopontin in microglia biology: current concepts and future perspectives. *Biomedicines* **10**, 840 (2022).

69. Brivio, E., Lopez, J. P. & Chen, A. Sex differences: Transcriptional signatures of stress exposure in male and female brains. *Genes Brain Behav.* **19**, e12643 (2020).

70. Blokland, G. A. M. et al. Sex-dependent shared and nonshared genetic architecture across mood and psychotic disorders. *Biol. Psychiatry* **91**, 102–117 (2022).

71. Lau, S.-F., Cao, H., Fu, A. K. Y. & Ip, N. Y. Single-nucleus transcriptome analysis reveals dysregulation of angiogenic endothelial cells and neuroprotective glia in Alzheimer's disease. *Proc. Natl Acad. Sci. USA* **117**, 25800–25809 (2020).

72. Mathys, H. et al. Single-cell transcriptomic analysis of Alzheimer's disease. *Nature* **570**, 332–337 (2019).

73. Chen, Y. & Colonna, M. Microglia in Alzheimer's disease at single-cell level. Are there common patterns in humans and mice? *J. Exp. Med.* **218**, e20202717 (2021).

74. Thrupp, N. et al. Single-nucleus RNA-Seq is not suitable for detection of microglial activation genes in humans. *Cell Rep.* **32**, 108189 (2020).

75. Gerrits, E., Heng, Y., Boddeke, E. & Eggen, B. J. L. Transcriptional profiling of microglia; current state of the art and future perspectives. *Glia* **68**, 740–755 (2020).

76. Dumais, A. et al. Risk factors for suicide completion in major depression: a case-control study of impulsive and aggressive behaviors in men. *Am. J. Psychiatry* **162**, 2116–2124 (2005).

77. Maitra, M. et al. Extraction of nuclei from archived postmortem tissues for single-nucleus sequencing applications. *Nat. Protoc.* **16**, 2788–2801 (2021).

78. Kriaucionis, S. & Heintz, N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science* **324**, 929–930 (2009).

79. Bankhead, P. et al. QuPath: Open source software for digital pathology image analysis. *Sci. Rep.* **7**, 16878 (2017).

80. R Core Team. R: A Language and Environment for Statistical Computing. https://cran.r-project.org/ (2021).

81. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e3529 (2021).

82. Patterson-Cross, R. B., Levine, A. J. & Menon, V. Selecting single cell clustering parameter values using subsampling-based robustness metrics. *BMC Bioinforma.* **22**, 39 (2021).

83. Korsunsky, I., Nathan, A., Millard, N. & Raychaudhuri, S. Presto scales Wilcoxon and auROC analyses to millions of observations. Preprint at *bioRxiv* https://doi.org/10.1101/653253 (2019).

84. McKenzie, A. T. et al. Brain cell type specific gene expression and co-expression network architectures. *Sci. Rep.* **8**, 8868 (2018).

85. Maynard, K. R. et al. Transcriptome-scale spatial gene expression in the human dorsolateral prefrontal cortex. *Nat. Neurosci.* **24**, 425–436 (2021).

86. Hodge, R. D. et al. Conserved cell types with divergent features in human versus mouse cortex. *Nature* **573**, 61–68 (2019).

87. Fischer, S., Crow, M., Harris, B. D. & Gillis, J. Scaling up reproducible research for single-cell transcriptomics using Meta-Neighbor. *Nat. Protoc.* **16**, 4031–4067 (2021).

88. Li, M. et al. Integrative functional genomic analysis of human brain development and neuropsychiatric risks. *Science* **362**, eaat7615 (2018).

89. Lake, B. B. et al. Integrative single-cell analysis of transcriptional and epigenetic states in the human adult brain. *Nat. Biotechnol.* **36**, 70–80 (2018).

90. Habib, N. et al. Massively parallel single-nucleus RNA-seq with DroNc-seq. *Nat. Methods* **14**, 955 (2017).

91. Street, K. et al. Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* **19**, 477 (2018).

92. Van den Berge, K. et al. Trajectory-based differential expression analysis for single-cell sequencing data. *Nat. Commun.* **11**, 1201 (2020).

93. Kassambara, A. rstatix: Pipe-friendly framework for basic statistical tests. https://cran.r-project.org/package=rstatix (2021).

94. Canty, A. & Ripley, B. D. boot: Bootstrap R (S-Plus) Functions. https://cran.r-project.org/package=boot (2021).

95. Crowell, H. L. et al. muscat detects subpopulation-specific state transitions from multi-sample multi-condition single-cell transcriptomics data. *Nat. Commun.* **11**, 6077 (2020).

96. 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).

97. McCarthy, D. J., Campbell, K. R., Lun, A. T. L. & Wills, Q. F. Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. *Bioinformatics* **33**, 1179–1186 (2017).

98. Jäkel, S. et al. Altered human oligodendrocyte heterogeneity in multiple sclerosis. *Nature* **566**, 543–547 (2019).

99. Kokkosis, A. G., Madeira, M. M., Mullahy, M. R. & Tsirka, S. E. Chronic stress disrupts the homeostasis and progeny progression of oligodendroglial lineage cells, associating immune oligodendrocytes with prefrontal cortex hypomyelination. *Mol. Psychiatry* **27**, 2833–2848 (2022).

100. Rahimian, R., Perlman, K., Canonne, C. & Mechawar, N. Targeting microglia-oligodendrocyte crosstalk in neurodegenerative and psychiatric disorders. *Drug Discov. Today* **27**, 2562–2573 (2022).

101. Rau, A., Marot, G. & Jaffrézic, F. Differential meta-analysis of RNA-seq data from multiple studies. *BMC Bioinforma.* **15**, 91 (2014).

102. Subramanian, A. et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).

103. Korotkevich, G. et al. Fast gene set enrichment analysis. Preprint at *bioRxiv* https://doi.org/10.1101/060012 (2021).

104. Gillespie, M. et al. The reactome pathway knowledgebase 2022. *Nucleic Acids Res.* **50**, D687–D692 (2021).

105. Dolgalev, I. msigdbr: MSigDB Gene Sets for Multiple Organisms in a Tidy Data Format. https://cran.r-project.org/package=msigdbr (2021).

106. Gutiérrez-Sacristán A. et al. psygenet2r: a R/Bioconductor package for the analysis of psychiatric disease genes. *Bioinformatics* **33**, 4004–4006 (2017).

107. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinforma.* **9**, 559 (2008).

108. Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47–e47 (2015).

109. Perez, J. M. et al. Hippocampal subfield transcriptome analysis in schizophrenia psychosis. *Mol. Psychiatry* **26**, 2577–2589 (2021).

110. Wu, T. et al. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innovation* **2**, 100141 (2021).

Acknowledgements

Author contributions

M.M. processed samples, performed data analysis, and wrote the manuscript. H.M., A.C., and Z.A. supported data analysis. R.R. and K.Y. interpreted data and reviewed the manuscript. J.Y., L.M.F., and M.A.D. processed samples. D.C.M. provided samples. M.S. provided bioinformatic expertise and reviewed the manuscript. N.M. contributed to sample procurement and reviewed the manuscript. G.T. obtained funding. G.T. and C.N. jointly supervised the study design, implementation, and manuscript preparation. C.N. additionally performed snRNA-seq experiments.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-023-38530-5.

Correspondence and requests for materials should be addressed to Gustavo Turecki or Corina Nagy.

Peer review information *Nature Communications* thanks Carmine Pariante and the other, anonymous, reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.