Supplementary Materials

Genetic analysis of the human microglia transcriptome across brain regions, aging and disease pathologies

Authors

Katia de Paiva Lopes[1,2,3,4,*], Gijsje J. L. Snijders[5,6,*], Jack Humphrey[1,2,3,4,*], Amanda Allan[1,2,3,4], Marjolein Sneeboer[7], Elisa Navarro[1,2,3,4], Brian M. Schilder[1,2,3,4], Ricardo A. Vialle[1,2,3,4], Madison Parks[1,2,3,4], Roy Missall[5], Welmoed van Zuiden[5], Frederieke Gigase[5,6], Raphael Kübler[5], Amber Berdenis van Berlekom[7], Emily M. Hicks[1,2,3,4], Chotima Böttcher[8], Josef Priller[8], René S. Kahn[5,6], Lot D. de Witte[5,6,*,#], Towfique Raj[1,2,3,4,*,#]

Affiliations

1. Nash Family Department of Neuroscience & Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, NY, United States of America
2. Ronald M. Loeb Center for Alzheimer's Disease, Icahn School of Medicine at Mount Sinai, New York, NY, United States of America
3. Department of Genetics and Genomic Sciences & Icahn Institute for Data Science and Genomic Technology, Icahn School of Medicine at Mount Sinai, New York, NY, United States of America
4. Estelle and Daniel Maggin Department of Neurology, Icahn School of Medicine at Mount Sinai, New York, NY, United States of America
5. Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, United States of America
6. Mental Illness Research, Education and Clinical Center (MIRECC), James J Peters VA Medical Center, New York City, NY, United States of America
7. Department of Translational Neuroscience, University Medical Center Utrecht Brain Center, Utrecht University, 3584 CG Utrecht, The Netherlands
8. Department of Neuropsychiatry and Laboratory of Molecular Psychiatry, Charité-Universitätsmedizin Berlin, 10117 Berlin, Germany.

* These authors contributed equally

# these authors share correspondence:

Lot D. de Witte: lotje.dewitte@mssm.edu

Towfique Raj: towfique.raj@mssm.edu

Supplementary Note

Supplementary Methods

Microglia isolation protocol

Brain tissue was stored in Hibernate media (Gibco) at 4 °C upon processing within 24 hours after autopsy. Brain tissue was first mechanically dissociated through a metal sieve in a glucose- potassium-sodium buffer (GKN-BSA; 8.0 g/L NaCl, 0.4 g/L KCl, 1.77 g/L Na2HPO4.2H2O, 0.69 g/L NaH2PO4.H2O, 2.0 g/L D-(1)-glucose, 0.3% bovine serum albumin (BSA, Merck, Darmstadt, Germany); pH 7.4) and supplemented with collagenase Type I (3700 units/mL; Worthington, USA) and DNase I (200 µg/mL; Roche, Switzerland) or 2% of Trypsin (Invitrogen) at 37 °C for 30 min or 60 min while shaking. The suspension was put over a 100 µM cell strainer and washed with GKN-BSA buffer in the centrifuge (1800 rpm, slow brake, 4 °C, 10 min) before the pellet was resuspended in 20 mL GKN-BSA buffer. 10 mL of Percoll (Merck, Darmstadt, Germany) was added dropwise and the tissue homogenate was centrifuged at 4000 rpm (fast acceleration, slow brake at 4 °C, 30 min). The middle layer was collected and washed with GKN-BSA buffer, followed by resuspension and centrifuging in a magnetic-activated cell sorting (MACS) buffer (PBS, 1% heat-inactivated fetal cow serum (FCS), 2 mM EDTA; 1500 rpm, 10 °C, 10 min). Microglia were positively selected with CD11b-conjugated magnetic microbeads (Miltenyi Biotec, Germany) according to the manufacturer's protocol. Microglia were stored in RLT buffer + 1% 2-Mercaptoethanol or lysed in TRIzol reagent (Invitrogen, USA). RNA was isolated using RNeasy Mini kit (Qiagen) adding the DNase I optional step or as described in detail before[1]. Library preparation was performed at Genewiz using the Ultra-low input system which uses Poly-A selection. CD11b is also present on perivascular macrophages in the CNS. However, we have previously shown by mass cytometry that the percentage of macrophages ($CD206^{high}$) was low[2]. In a subsample of our cohort we checked the percentage of (infiltrating) monocytes by measuring migration inhibitory factor related protein 14 (MRP14+) cells (n = 56 samples derived from 19 donors). The percentage of (infiltrating) monocytes was 0.04 % on average. Moreover, we showed in a subsample (*n* = 91 samples derived from 30 donors) of our cohort that the mean percentage of microglia ($P2Y12^{+}CD64^{+}$) cells is 95%, ranging from 56.6% to 99.9%. The percentage of P2Y12-CD64+ is 3.2%, suggesting low levels of infiltration of macrophages/monocytes (see **Figure S2**).

We performed this study using a validated protocol for post-mortem microglia isolation[1–4], which includes the assessment of DNA-intercalator 7-amino-actinomycin D (7-AAD) staining to detect non-viable cells with flow cytometry. After each isolation, we performed trypan blue staining of the microglial cells to check the amount of death cells with microscopy. The resulting cell viability was between 70% and 98%.

DNA isolation

In short, a small piece of brain tissue was cut and placed in a 96-deepwell plate (Thermo Fisher) on dry ice. A mastermix containing binding enhancer, PBS and proteinase K (Qiagen) was added to each well. The solution was incubated overnight at 65 degrees Celsius. The KingFisher™ (KF) Duo system was applied with a 12 pin magnet head which enabled processing of 12 samples per run using microtiter 96 deepwell plates (Thermo Fisher). Prior to the extraction process, the deepwell plates were filled with the following reagents: wash buffer (Qiagen), 80% Ethanol, elution buffer (Qiagen) and tip combs (Qiagen). For extraction, 440 µl of DNA binding buffer beads (Qiagen) were added to the sample and mixed by pipetting. After that, automated extraction was executed and completed within 30 minutes using the protocol provided by the manufacturer. The extracted DNA was eluted in 200 µl elution buffer and subsequently transferred to 1.5 ml tubes for storage.

Differential expression analysis

Differential age-by-region gene expression was measured by fitting linear mixed models using DREAM (considering repeated donor measures). Analogously as in the standard region differential analysis, expression data were first normalized and transformed using voomWithDreamWeights. Four models were fitted, setting each brain region as a reference. The model included sex, donor ID, cause of death, the first 4 ancestry MDS values (C1-4), % mRNA bases, median insert size, and % ribosomal bases, plus an interaction term with age and region. Next, we extracted the interaction coefficients for all pairwise region comparisons and selected genes with FDR < 0.05.

Differential sex-related analysis was performed with the following model: sex, donor ID, age, region, cause of death, the first 4 ancestry MDS values (C1-4), % mRNA bases, median insert size, and % ribosomal bases, with the difference that we have excluded the genes from the chromosomes X and Y. The effect of diagnosis was analyzed with DREAM in microglial samples from subjects with dementia (n = 15 samples derived from 9 donors; **Supplementary Table 10)**, Parkinson's disease (PD; n = 18 samples derived from 12 donors, **Supplementary Table 11)**, major depressive disorder (MDD; n = 74 samples derived from 21 donors, **Supplementary Table 12**; bipolar disorder or schizophrenia (BD/SCZ; n = 37 samples derived from 14 donors, and controls (n = 96 samples derived from 38 donors, **Supplementary Table 13)**. The final model included sex, donor ID, age, region, cause of death, the first 4 ancestry MDS values (C1-4), % mRNA bases, median insert size and % ribosomal bases.

Differential Transcript Usage

Transcript expression was estimated in each sample using RSEM with the GENCODE v30 transcript reference. Lowly expressed transcripts were removed with the threshold transcript counts per million > 1 in at least 30% of all samples. This retained 49,694 transcripts from 10,818 genes. Differential transcript usage was tested simultaneously between each region and with donor age using satuRn[5], a fast method for computing differential transcript usage. No current differential transcript usage tool can model random effects so we were unable to account for shared donors, but otherwise the same technical covariates were used as in the differential expression modelling. Pairwise comparisons between each pair of regions were extracted using the limma::makeContrasts() function. To correct for test statistic inflation due to correlation across transcripts and donors, we employed a more stringent empirical FDR correction[6], but used a more liberal FDR threshold of 0.1 to increase our discovery set of transcripts. Transcripts were considered differentially used between regions at a |log odds ratio| > 1 and an empirical FDR < 0.1. For aging, transcripts were filtered at an FDR < 0.1 but no effect size cutoffs were used.

Single-cell sequencing

To generate single-cell data we isolated CD11b+ microglial cells from the medial frontal gyrus and hippocampus of one additional donor (MG-22). We used the instructions of the Single-Cell 3' Reagent Kits v2 (10 x Genomics) to construct scRNAseq barcoded libraries. In short, we loaded ~10,000 microglial cells from both regions separately into a slot of Chromium Chip. GEMs were incubated in a thermal cycler for the generation of barcoded cDNA. The cDNA was fragmented after amplification and processed for sequencing by ligating adapters and sample indices. The libraries were sequenced on an Illumina Hi-Seq system. The average sequencing depth was 2197.959 raw reads per cell, and the average was 1108.241 UMIs per cell. We analyzed 8589 cells for the two brain regions combined. Low-quality cells with > 10% mitochondrial gene content were removed. Duplicate cells were filtered by removing cells with greater than 10000 transcripts. Genes not detected in at least 3 cells were removed. After quality control, we analyzed 5990 single cells in total. We used the

standard Seurat workflow (Seurat v. 3.2.3) to normalize (NormalizeData; default parameters), scale (ScaleData; on all genes regressing out total UMI count and mitochondrial percentage per cell) and performed dimensionality reduction using principal component analysis with the top 2000 highly variable genes (RunPCA). The first 13 principal components were used for PCA-Louvain clustering (FindNeighbors and FindClusters). To gain sufficient detail to detect small subpopulations within one donor, the cluster resolution was set to 0.1. We used UMAP, a non-linear dimensionality reduction, to visualize the data (RunUMAP with the first 13 PCs) and produce feature plots in low dimensional space (DimPlot and FeaturePlot). Cluster-enriched genes were identified using logistic regression (FindAllMarkers function with default thresholds and only.pos = TRUE) and were used to manually annotate clusters by cell type.

IFN-y and LPS stimulated microglia

Microglia were isolated from the MFG of 6 unaffected controls, cultured and stimulated as described before[1,2]. After overnight incubation, microglia medium was supplemented with 100 ng/mL LPS from Escherichia coli 0111:B4 (Merck, Germany) or with 50 ng/ml interferon gamma INF-γ for 6 hours. Microglial RNA was isolated using the TRIzol method and cDNA libraries were generated using the SMART-Seq v4 Ultra Low Input RNA Kit for Sequencing Components (Takara) according to the manufacturer's protocol. The libraries were sequenced as 150 bp paired-end reads with an average read depth of 42 million (range 16-98M) read pairs on an Illumina HiSeq 2500. FASTQ files were processed through RAPiD (v19.09.1). Two outliers (one unstimulated sample and one LPS-exposed sample) were detected with the hclust function in R and excluded from further analysis. Differential expression was tested using DESeq2[7], testing all genes with greater than 1 read count per million in at least 30% of samples. The effect of condition was calculated separately for LPS and IFNy, controlling for donor variation. Since donor highly correlated with all other covariates measured, no other variables were included in the design. Genes with an adjusted P-value of < 0.05 were considered significant.

Colocalization and fine-mapping

For the MiGA eQTLs and sQTLs we used the METASOFT random effects meta-analysis across the four regions for colocalization analyses. We used the coloc package[8] to test whether SNPs from different disease GWAS colocalized with expression and splicing QTLs from microglia, monocytes and brain (dorsolateral prefrontal cortex). For each genome-wide significant locus in a GWAS we extracted the nominal summary statistics of association for all SNPs within 1 megabase either upstream/downstream of the top lead SNP (2Mb-wide region total). In each QTL dataset we then extracted all nominal associations for all SNP-gene pairs within that range and tested for colocalization between the GWAS locus and each gene. We used thresholds of posterior probability H4 (PP4) ≥ 0.5 for suggestive, ≥ 0.7 for moderate and ≥ 0.9 for strong colocalization, respectively. We restricted our colocalizations to GWAS SNP - eQTL SNP pairs where the distance between their respective top SNPs was ≤ 500kb or the two lead SNPs were in moderate linkage disequilibrium (r2 > 0.1), taken from the 1000 Genomes (Phase 3) European populations using the LDLinkR package[9]).

For splicing QTLs we followed the same approach but collapsed junctions to return only the highest PP4 value for each gene in each locus. For presenting results across diseases, we merged overlapping loci from the four Alzheimer's disease studies together, presenting results with the highest PP4 value for each gene. Due to the smaller window of association (100kb from the center of the intron excision cluster) we restricted reported colocalizations to cases where the GWAS SNP and the top sQTL SNP were either within 100kb of each other or in moderate linkage disequilibrium ($r^2 > 0.1$).

We used echolocator[10] to perform statistical and functional fine-mapping of each GWAS locus with a suggestive colocalization PP4 > 0.5. echolocatoR combines the output of multiple fine-mapping tools to identify high-confidence putative causal SNPs. The full 2Mb window was fine-mapped in each locus to better take into account more widespread LD architectures. SNPs with minor allele frequency (MAF) < 5% were removed as we were primarily focused in identifying common risk factors. We fine-mapped using ABF[11], FINEMAP[12], SUSIE[13], and POLYFUN + SUSIE[14], with 1000 Genomes (Phase 3) European samples as our LD reference. For tools that permitted, we set the maximum number of causal SNPs per locus to 5. Each tool produces a 95% credible set of SNPs, which can be understood as SNPs with a posterior probability > 95% of being causal for the given phenotype. echolocatoR then defines SNPs that are present in multiple tools' credible sets as "consensus SNPs" with a higher level of confidence in their causality. Importantly, we define the lead GWAS SNP as the lead variant listed in the original summary statistics, which may or may be not prioritized by fine-mapping as a credible or consensus SNP.

Cell-type specific promoter-enhancer data

We downloaded processed cell-type specific promoter and enhancer data[15] using the echolocatoR package[10]. Briefly, fluorescent activated nuclear sorting (FANS) was performed on post-mortem human brains to isolate PU.1+ microglia, NEUN+ neurons, OLIG2+ oligodendrocytes, and NEUN- LHX2+ astrocyte nuclei. Chromatin immunoprecipitation sequencing (ChIP-seq) was performed for the histone modifications H3K27ac and H3K4me3, which identify activated chromatin regions and promoters, respectively. Regions were defined by Nott and colleagues as active promoters when a H3K4me3 peak overlapped H3K27ac within 2000bp of the nearest transcription start site, whereas active enhancers were defined as H3K27ac peaks that did not overlap any H3K4me3 peaks. In addition, they performed proximity ligation-assisted ChIP-seq (PLAC-seq[16]), which identifies long-range connections between H3K4me3-positive promoter regions and other genomic regions. We downloaded coordinates for ChIP-seq and PLAC-seq in each cell type and overlapped our SNP sets in each colocalized locus using GenomicRanges in R[17].

Predicting transcription factor binding motif disruption

We used motifbreakR[18], a package that computes the similarity of a genomic sequence to a range of transcription factor motifs and calculates the potential disruption to each motif caused by a SNP with a score and P-value. We used as input 426 human transcription factor motifs from the HOCOMOCO database[19] as provided by motifbreakR. We ran the package on all lead QTL, lead GWAS and fine-mapped SNPs found at the USP6NL locus (using the AD GWAS from Jansen et al.) and the P2RY12 locus from the Nalls et al PD GWAS. Motifbreakr returns metrics for each tested motif on how much each allele of a SNP disrupts or enhances predicted binding. Each comparison then undergoes a background correction to compute p-values. Full results for both loci are in Table S23.

Visualization

All plots were created using ggplot2[20] in R (version 3.6.0), with ggrepel[21], ggfortify[22], patchwork[23], and ggbio[24] for additional layers of visualization.

Supplementary Results

Pure microglial samples from post-mortem brain tissue

We collected 314 human microglial samples from 115 donors, of which 46 donors were non-neurological disease controls (**Supplementary Figure 1**). Microglia were isolated from fresh post-mortem tissue using CD11b-beads[1–3]. Mass cytometry (CyTOF) proteomic and scRNA-seq analyses found an average of 95% of cells were positive for the microglia-specific markers P2Y12 and TMEM119 (**Supplementary Figure 2A and 3)**. The other cells are mainly myeloid in origin, expressing monocyte marker genes (CD64+/P2Y12- and MRP+ in CyTOF; MRC1/LYV1 in scRNA-seq). All microglial samples expressed known microglia-specific genes at high levels (**Supplementary Fig. 2C**). After rigorous quality control, 255 microglial samples from 100 different donors and four different regions were included (**Supplementary Fig. 1 and 4; Supplementary Table 1).** The four regions comprise two cortical regions: the medial frontal gyrus (MFG) and superior temporal gyrus (STG); and two subcortical regions: the thalamus (THA) and subventricular zone (SVZ).

Neurological disease loci regulate microglia gene expression

We combined the results of four different statistical and functional fine-mapping approaches to create a set of fine-mapped variants at each disease locus (see Methods, **Supplementary Table 22**). We defined variants prioritized by at least one fine-mapping tool at a posterior probability ≥ 95% of being causal for the phenotype as “credible set SNPs”, and variants prioritized by at least two tools as “consensus SNPs”. In addition, because coloc does not take linkage disequilibrium (LD) structure into account, which can confound results due to non-independence between SNPs, we calculated the LD between each lead QTL variant and the set of fine-mapped SNPs at each locus using 1000 Genomes (phase 3) European reference samples (**Supplementary Fig. 26**). In GWAS loci with multiple colocalized genes this can be used to prioritize the most likely candidate. In the CD19 locus in PD, this approach gave additional weight to suggesting an eQTL in *SPNS1* as the likely causal gene, due to high LD between the *SPNS1* lead QTL SNP and the GWAS SNP (r^2*=0.78)*, compared to *TUFM* and *SULT1A2* (**Figure S27**). With the expanded set of SNPs at each locus, we then overlapped each SNP with sets of defined promoter and enhancer regions in microglia, neurons, oligodendrocytes, and astrocytes[15].

References

1. Melief, J. *et al.* Characterizing primary human microglia: A comparative study with myeloid subsets and culture models. *Glia* **64**, 1857–1868 (2016).
2. Böttcher, C. *et al.* Human microglia regional heterogeneity and phenotypes determined by multiplexed single-cell mass cytometry. *Nat. Neurosci.* **22**, 78–90 (2019).
3. Sneeboer, M. A. M. *et al.* Microglia in post-mortem brain tissue of patients with bipolar disorder are not immune activated. *Transl. Psychiatry* **9**, 153 (2019).
4. Melief, J. *et al.* Phenotyping primary human microglia: tight regulation of LPS responsiveness. *Glia* **60**, 1506–1517 (2012).
5. Gilis, J., Vitting-Seerup, K., Van den Berge, K. & Clement, L. satuRn: Scalable analysis of differential transcript usage for bulk and single-cell RNA-sequencing applications. *F1000Res.* **10**, 374 (2021).
6. Efron, B. Large-scale simultaneous hypothesis testing. *J. Am. Stat. Assoc.* **99**, 96–104 (2004).
7. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
8. Giambartolomei, C. *et al.* Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet.* **10**, e1004383 (2014).
9. Myers, T. A., Chanock, S. J. & Machiela, M. J. LDlinkR: An R package for rapidly calculating linkage disequilibrium statistics in diverse populations. *Front. Genet.* **11**, 157 (2020).
10. Schilder, B. M., Humphrey, J. & Raj, T. echolocatoR: an automated end-to-end statistical and functional genomic fine-mapping pipeline. *Bioinformatics* (2021) doi:10.1093/bioinformatics/btab658.
11. Wakefield, J. A Bayesian measure of the probability of false discovery in genetic epidemiology studies. *Am. J. Hum. Genet.* **81**, 208–227 (2007).
12. Benner, C. *et al.* FINEMAP: efficient variable selection using summary data from genome-wide association studies. *Bioinformatics* **32**, 1493–1501 (2016).
13. Wang, G., Sarkar, A., Carbonetto, P. & Stephens, M. A simple new approach to variable selection in regression, with application to genetic fine mapping. *J. R. Stat. Soc. Series B Stat. Methodol.* **82**, 1273–1300 (2020).
14. Weissbrod, O. *et al.* Functionally informed fine-mapping and polygenic localization of complex trait heritability. *Nat. Genet.* **52**, 1355–1363 (2020).
15. Nott, A. *et al.* Brain cell type-specific enhancer-promoter interactome maps and disease-risk association. *Science* **366**, 1134–1139 (2019).
16. Fang, R. *et al.* Mapping of long-range chromatin interactions by proximity ligation-assisted ChIP-seq. *Cell Res.* **26**, 1345–1348 (2016).
17. Lawrence, M. *et al.* Software for computing and annotating genomic ranges. *PLoS Comput. Biol.* **9**, e1003118 (2013).
18. Coetzee, S. G., Coetzee, G. A. & Hazelett, D. J. motifbreakR: an R/Bioconductor package for predicting variant effects at transcription factor binding sites. *Bioinformatics* **31**, 3847–3849 (2015).
19. Kulakovskiy, I. V. *et al.* HOCOMOCO: towards a complete collection of transcription factor binding models for human and mouse via large-scale ChIP-Seq analysis. *Nucleic Acids Res.* **46**, D252–D259 (2018).
20. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*. (Springer-Verlag New York, 2009).
21. Slowikowski, K. *ggrepel: Repel overlapping text labels away from each other*. (Github).
22. Tang, Y., Horikoshi, M. & Li, W. Ggfortify: Unified interface to visualize statistical results of popular R packages. *R J.* **8**, 474 (2016).
23. Pedersen, T. L. The Composer of Plots [R package patchwork version 1.1.1]. (2020).
24. Yin, T., Cook, D. & Lawrence, M. ggbio: an R package for extending the grammar of graphics for genomic data. *Genome Biol.* **13**, R77 (2012).
25. Hoffman, G. E. & Schadt, E. E. variancePartition: interpreting drivers of variation in complex gene expression studies. *BMC Bioinformatics* **17**, 483 (2016).
26. Peters, M. J. *et al.* The transcriptional landscape of age in human peripheral blood. *Nat. Commun.* **6**, 8570 (2015).
27. Navarro, E. *et al.* Dysregulation of mitochondrial and proteolysosomal genes in Parkinson's disease myeloid cells. *Nat Aging* **1**, 850–863 (2021).

List of figures

List of tables

Supplementary Tables 1-23 are available as a separate file (in .xls format). Full nominal and permuted eQTL and sQTL summary statistics per brain region are available from Zenodo at https://doi.org/10.5281/zenodo.4118605 (eQTL) and https://doi.org/10.5281/zenodo.4118403 (sQTL). Results for eQTL and sQTL meta-analysis and colocalization with diseases are available from Zenodo at https://doi.org/10.5281/zenodo.4118676. Allelic information for all QTL analyses is available from Zenodo at https://doi.org/10.5281/zenodo.4301005 .

Supplementary Table 1: Metadata for all microglia samples (n = 255).

Supplementary Table 2: Full summary statistics from microglia differential gene expression between male and female.

Supplementary Table 3: Full summary statistics from microglia differential gene expression between MFG and SVZ.

Supplementary Table 4: Full summary statistics from microglia differential gene expression between SVZ and STG.

Supplementary Table 5: Full summary statistics from microglia differential gene expression between THA and MFG.

Supplementary Table 6: Full summary statistics from microglia differential gene expression between THA and STG.

Supplementary Table 7: Full summary statistics from microglia differential gene expression between THA and SVZ.

Supplementary Table 8: Full summary statistics from microglia differential gene expression between MFG and STG.

Supplementary Table 9: Microglia K-means clusters and splicing genes pathway analysis and upstream regulator results and replication with external dataset (Van der Poel et al. 2019).

Supplementary Table 10: Full summary statistics from microglia differential gene expression between Dementia and controls.

Supplementary Table 11: Full summary statistics from microglia differential expression between Parkinson's disease and controls.

Supplementary Table 12: Full summary statistics from microglia differential expression between Major Depressive Disorder and controls.

Supplementary Table 13: Full summary statistics from microglia differential gene expression between Schizophrenia/Bipolar Disorder and controls.

Supplementary Table 14: Full summary statistics from age-related analysis (Dream software).

Supplementary Table 15: Microglia pathway analysis results for the significant age-related genes (up and down regulated at FDR < 0.05) and age-related splicing genes and replication with external datasets.

Supplementary Table 16: Full summary statistics from the differential gene expression analysis using an interaction-term between age and region.

Supplementary Table 17: Number of eGenes, sGenes and sClusters per region.

Supplementary Table 18: Meta-analysis mashR significant results (lfsr < 0.05) for eQTLs.

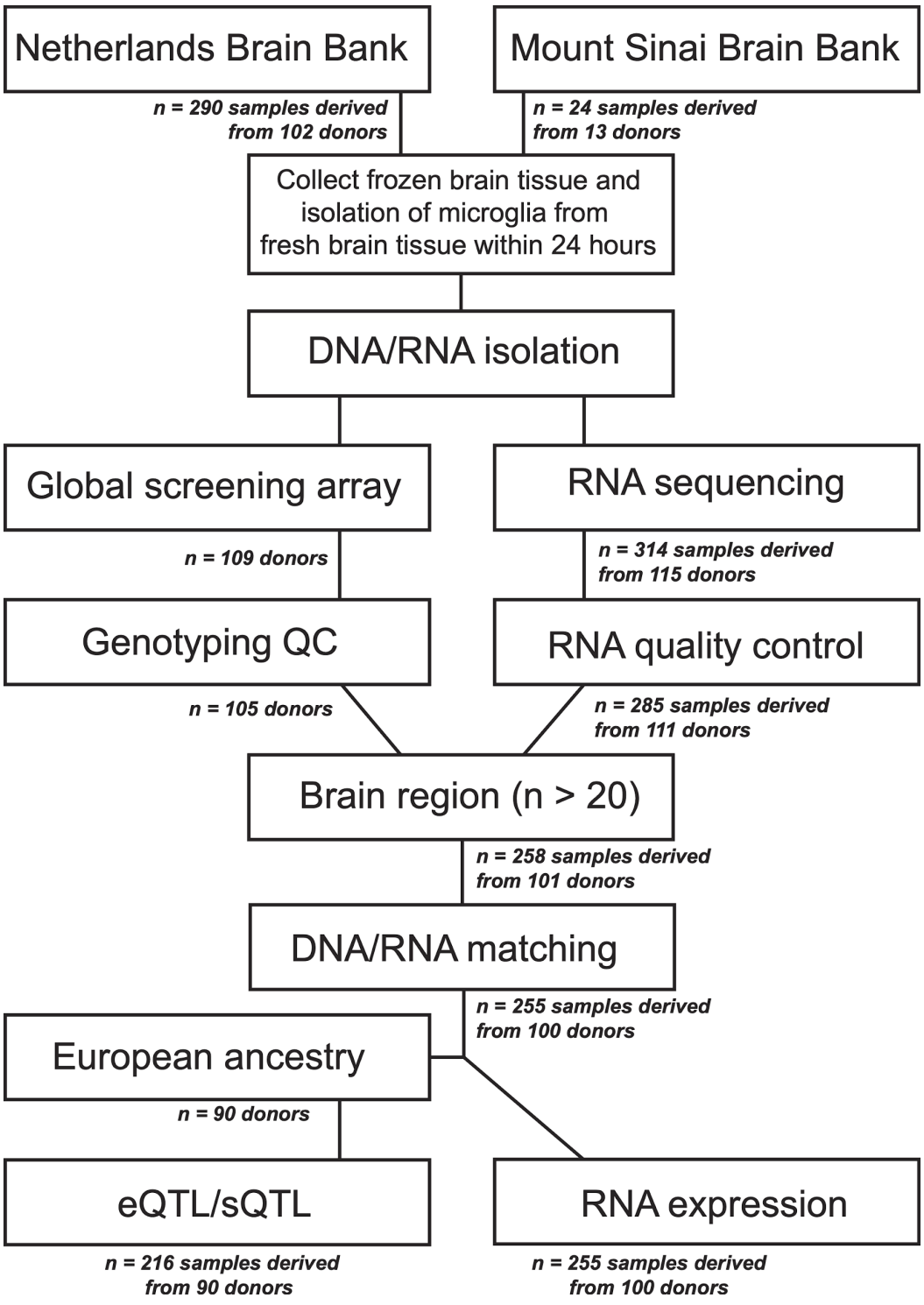
Supplementary Table 19: Meta-analysis mashR significant results (lfsr < 0.05) for sQTLs.

Supplementary Table 20: Region-specific eQTLs according to mashR (lfsr < 0.05 and effect size at least 2-fold larger in one tissue than in any other)

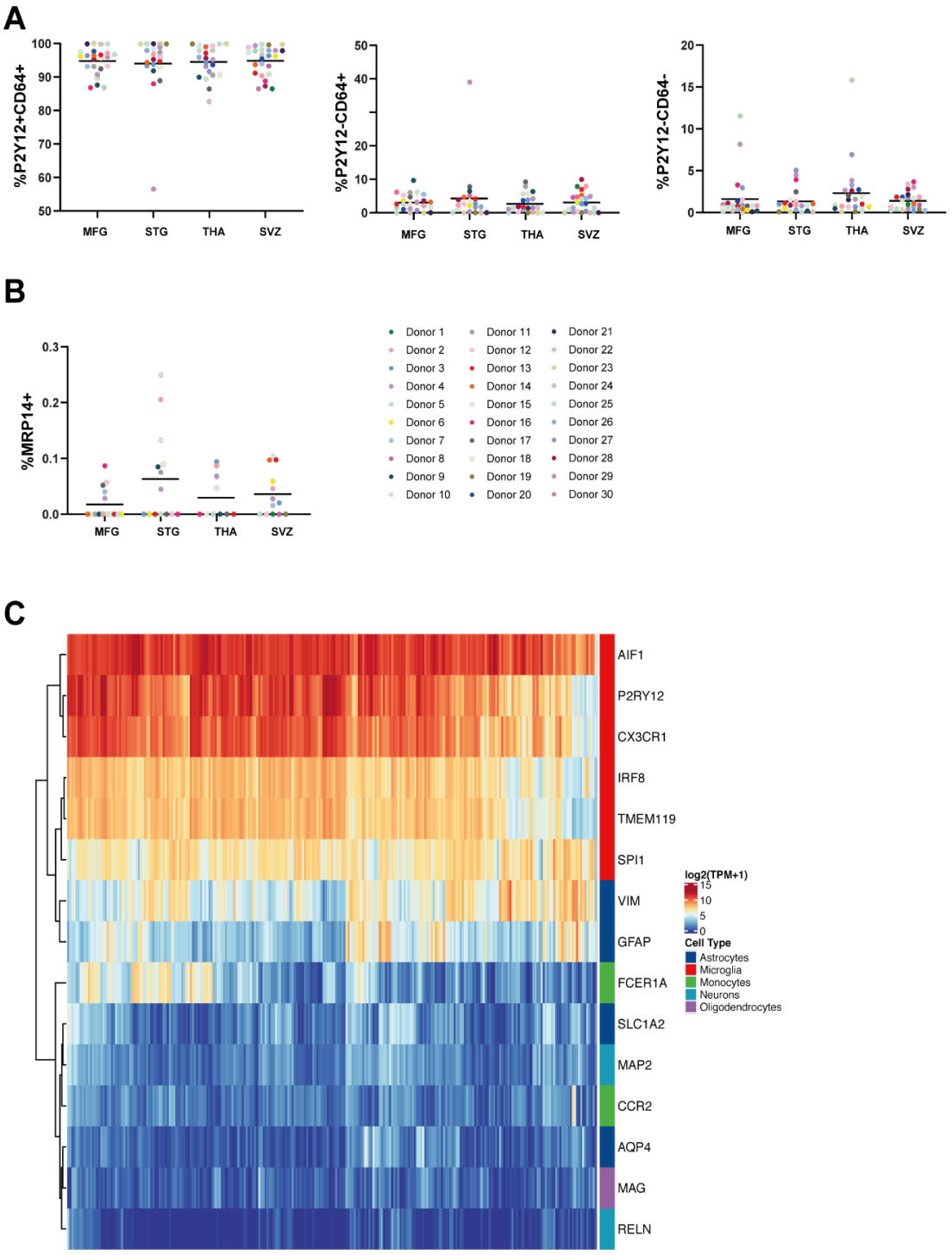
Supplementary Table 21: COLOC results at PP4 > 0.5.

Supplementary Table 22: Fine-mapping results for the 79 loci that are colocalized with MiGA eQTLs.

Supplementary Table 23: Result of the Motif breaker analysis for *USP6NL* and *P2RY12* genes.

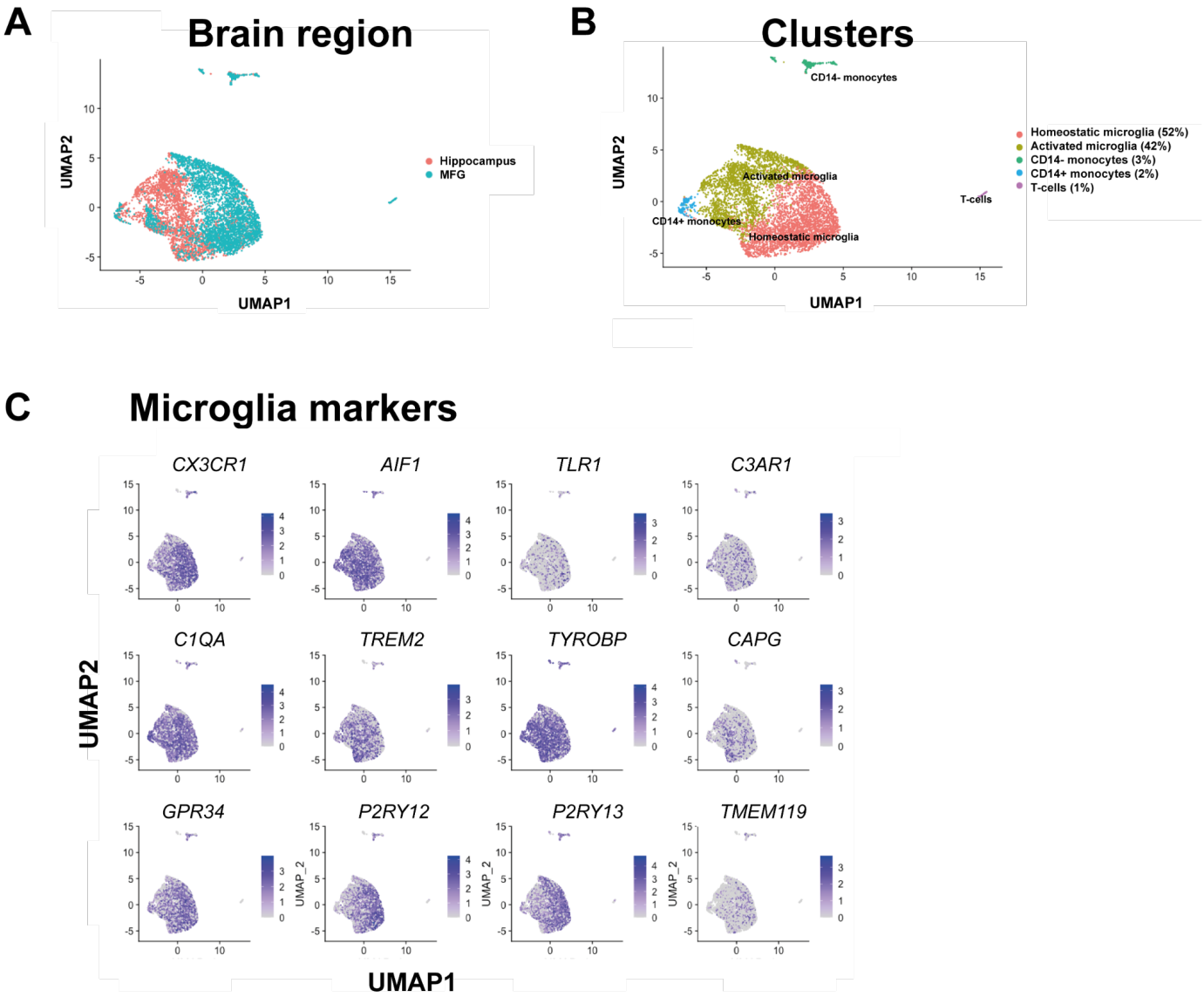


Supplementary Figure 1: Flowchart of quality control.



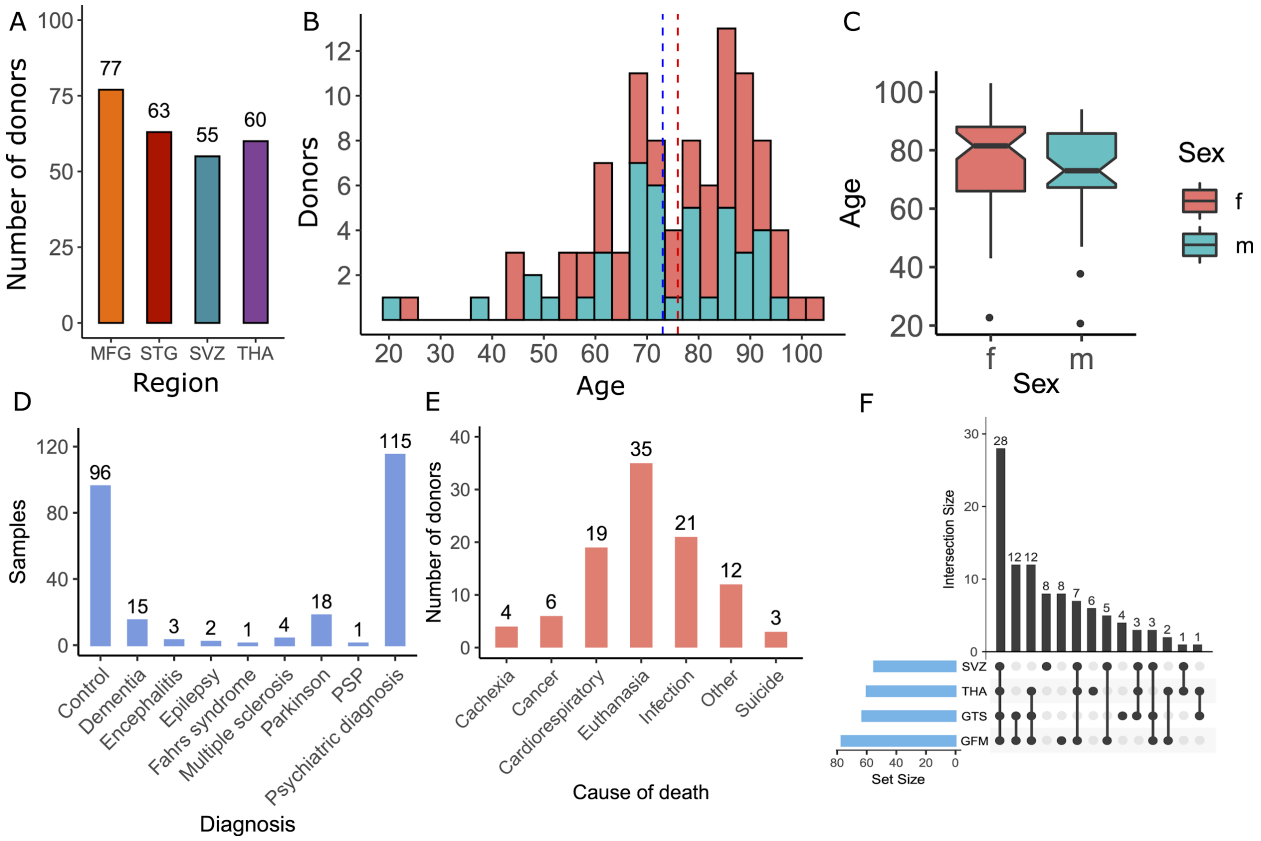
Supplementary Figure 2: Purity of microglia samples.

A) CyTOF analysis of CD11b-isolated cells on a subsample of *n* = 30 donors. Cells are gated for P2Y12+CD64+ (microglia); P2Y12-/CD64+ (macrophages) and P2Y12-/CD64- (non-myeloid cells) B) CD11b isolated cells are gated for MPR14+ (monocytes) in subsample of n = 19 donors with CyTOF analysis C) Expression levels (TPM+1 $\log_2$ scale) of cell markers for 255 samples. Blue colors indicate low expression and red colors indicate high gene expression.



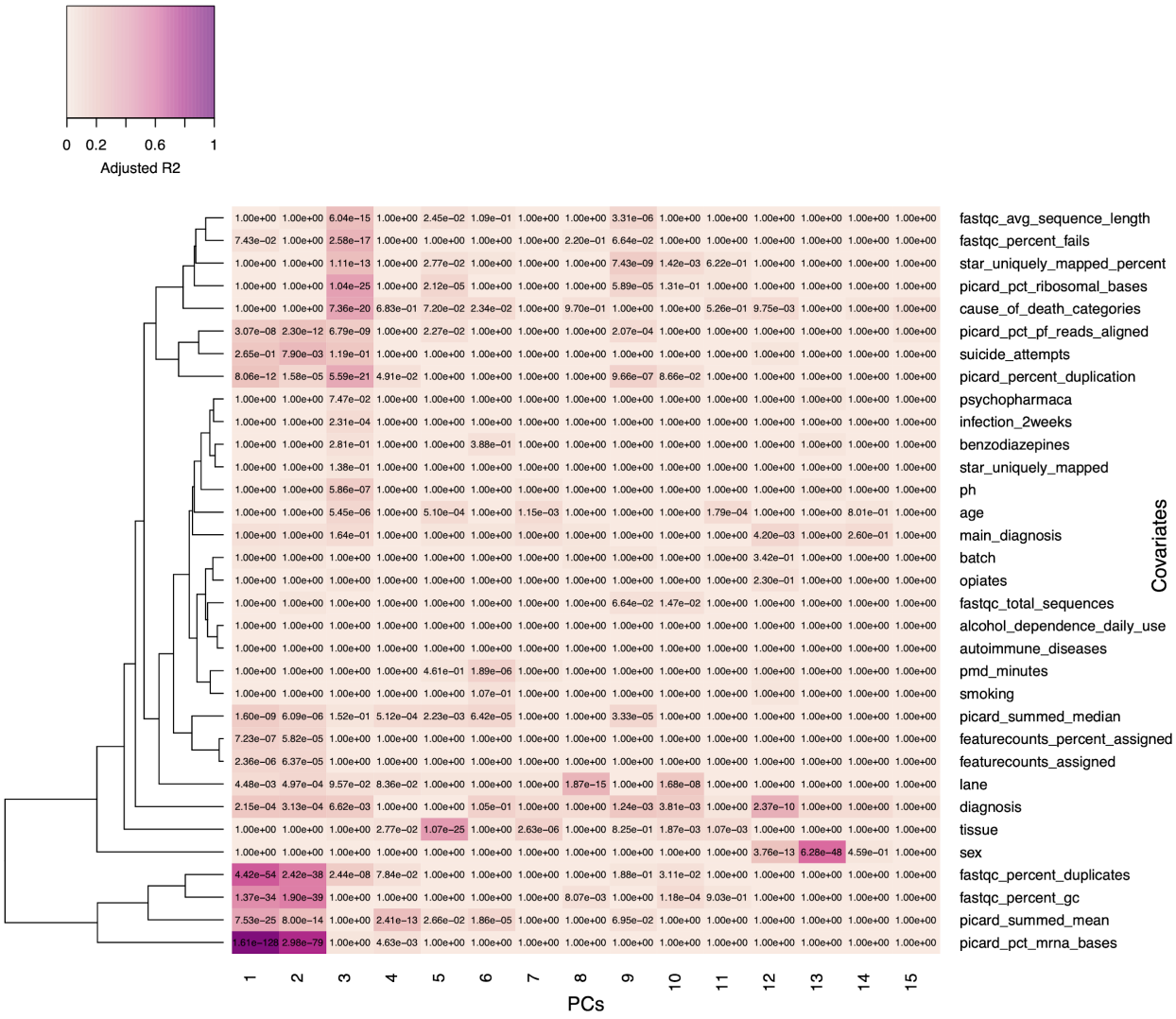
Supplementary Figure 3: Single-cell expression profiling of CD11b isolated microglia from individual donor using 10x genomics.

A) UMAP visualization and unsupervised clustering of CD11b isolated microglia from Parkinson's disease donor from medial frontal gyrus (MFG) and hippocampus. B) UMAP visualization and unsupervised clustering identified five subsets of microglia clusters. C)Average expression of microglia specific genes.

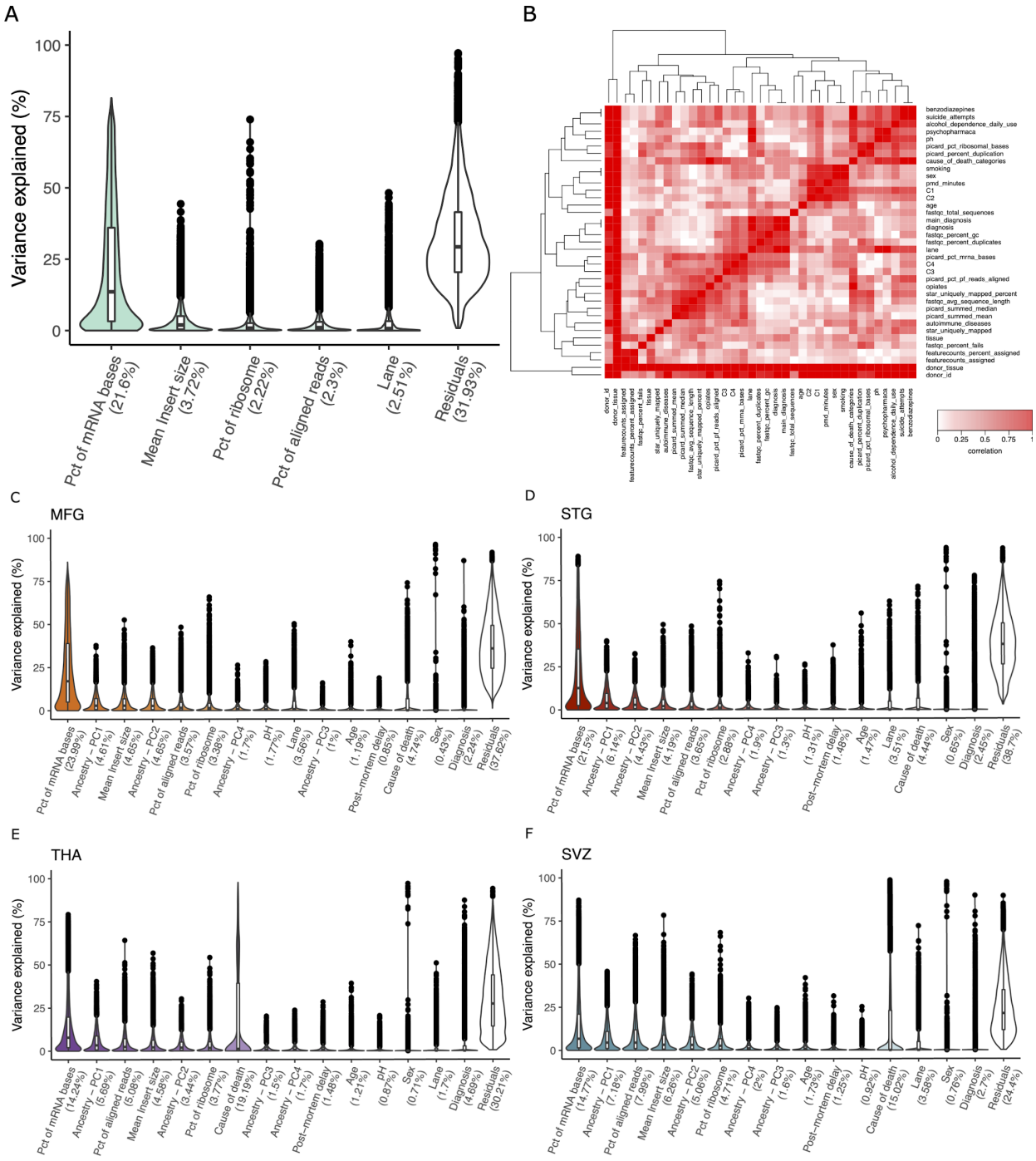


Supplementary Figure 4: Overview of the data.

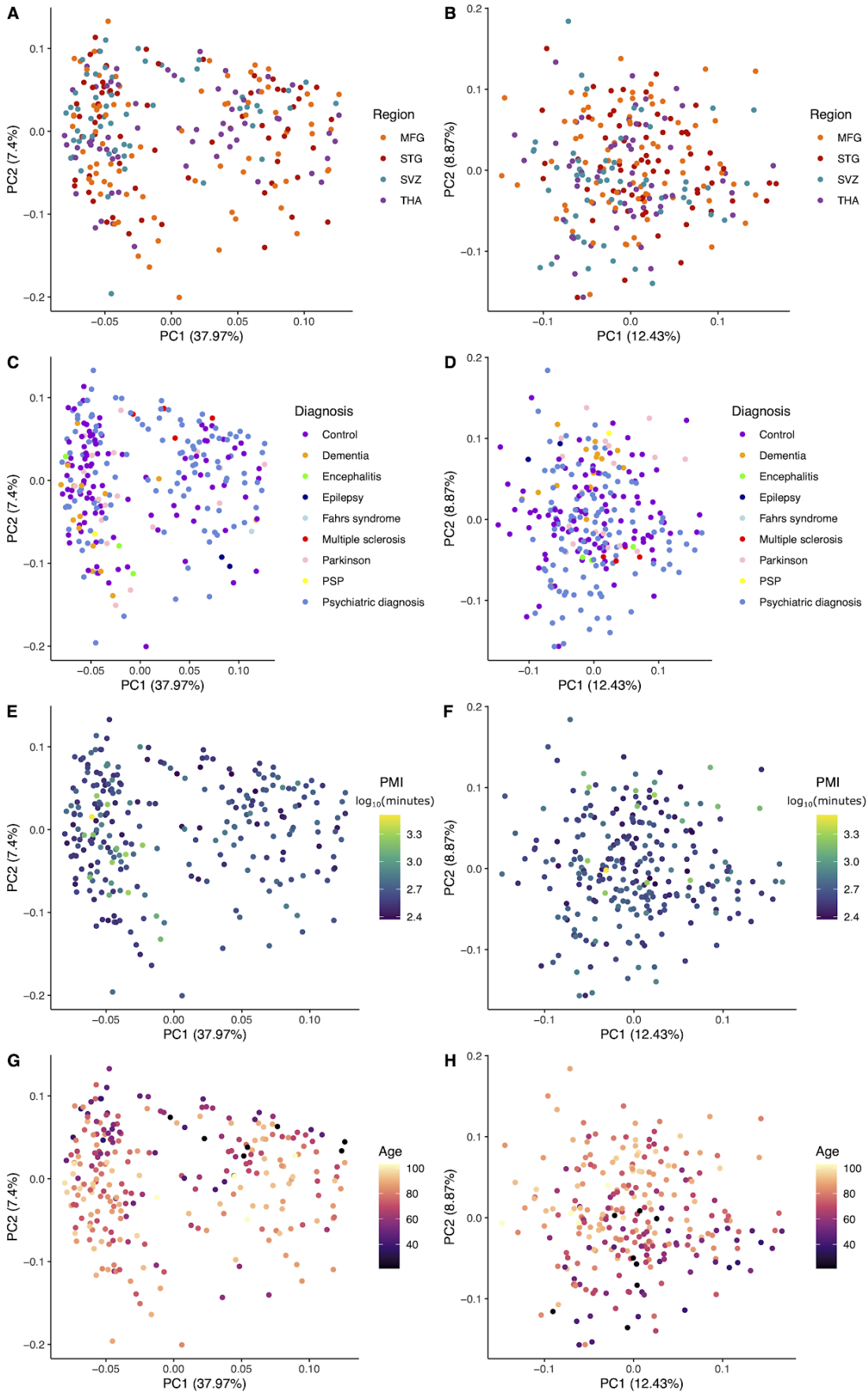
A) Number of donors by brain region. B-C) Age range of the 100 donors in this study. Blue indicates the distribution of male donors and red indicates the distribution of female donors. Dashed lines for mean age. D) Frequency of diagnosis. E) Frequency of cause of death. F) The number of brain regions in this study. Each donor donated one up to four brain regions.

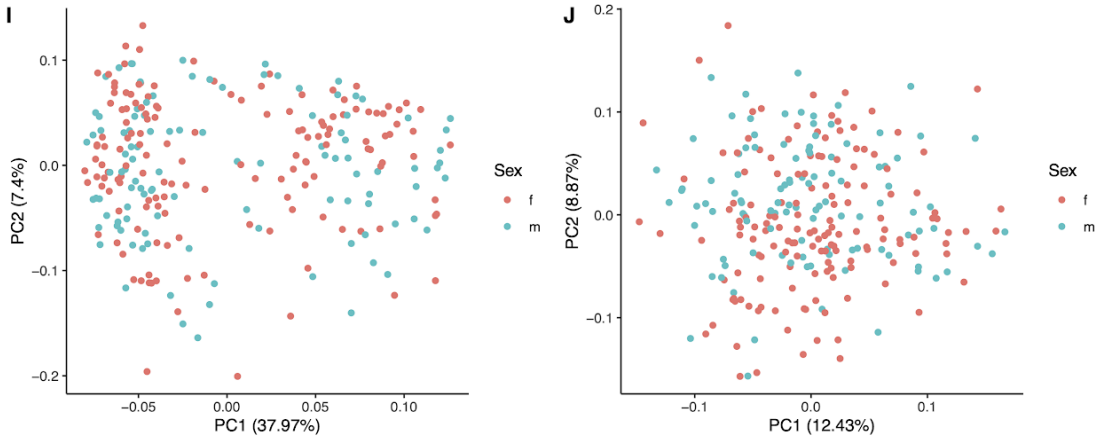


Supplementary Figure 5: Sources of variation in the gene expression data.
Linear regression between the first 15 Principal Components (PCs) and the covariates. Colors correspond to the adjusted R-squared. The *P*-values are from the Linear regression, two-sided and Bonferroni adjusted. “Tissue” refers to which brain region the microglia were isolated from.

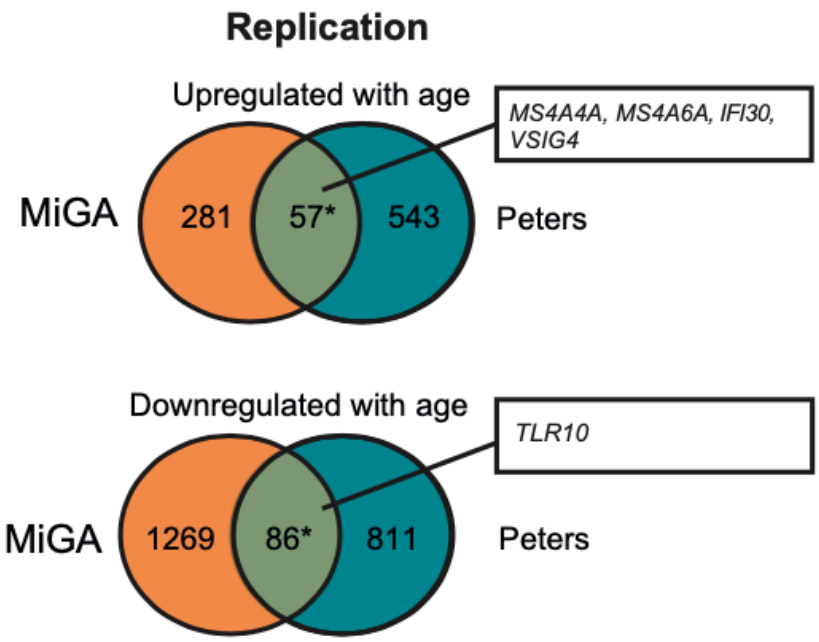


Supplementary Figure 6: Main sources of expression variation and correlation of covariates.
A) Percentage of variance explained by technical factors for 255 samples for each gene, variancePartition package[25]. Numbers between parentheses indicate mean values. Data are presented as percentage (%) of total variance explained. Box plots show median, box spans first to third quartiles, and whiskers extend 1.5 times the interquartile range (IQR) from the box. B) Pairwise canonical correlation between the covariates. Red indicates high correlation; white indicates low correlation. C-F) Percentage of variance explained by technical and biological covariates by each brain region. Data are presented as a percentage (%) of total variance explained. Box plots show median, box spans first to third quartiles, and whiskers extend 1.5 times the interquartile range (IQR) from the box.

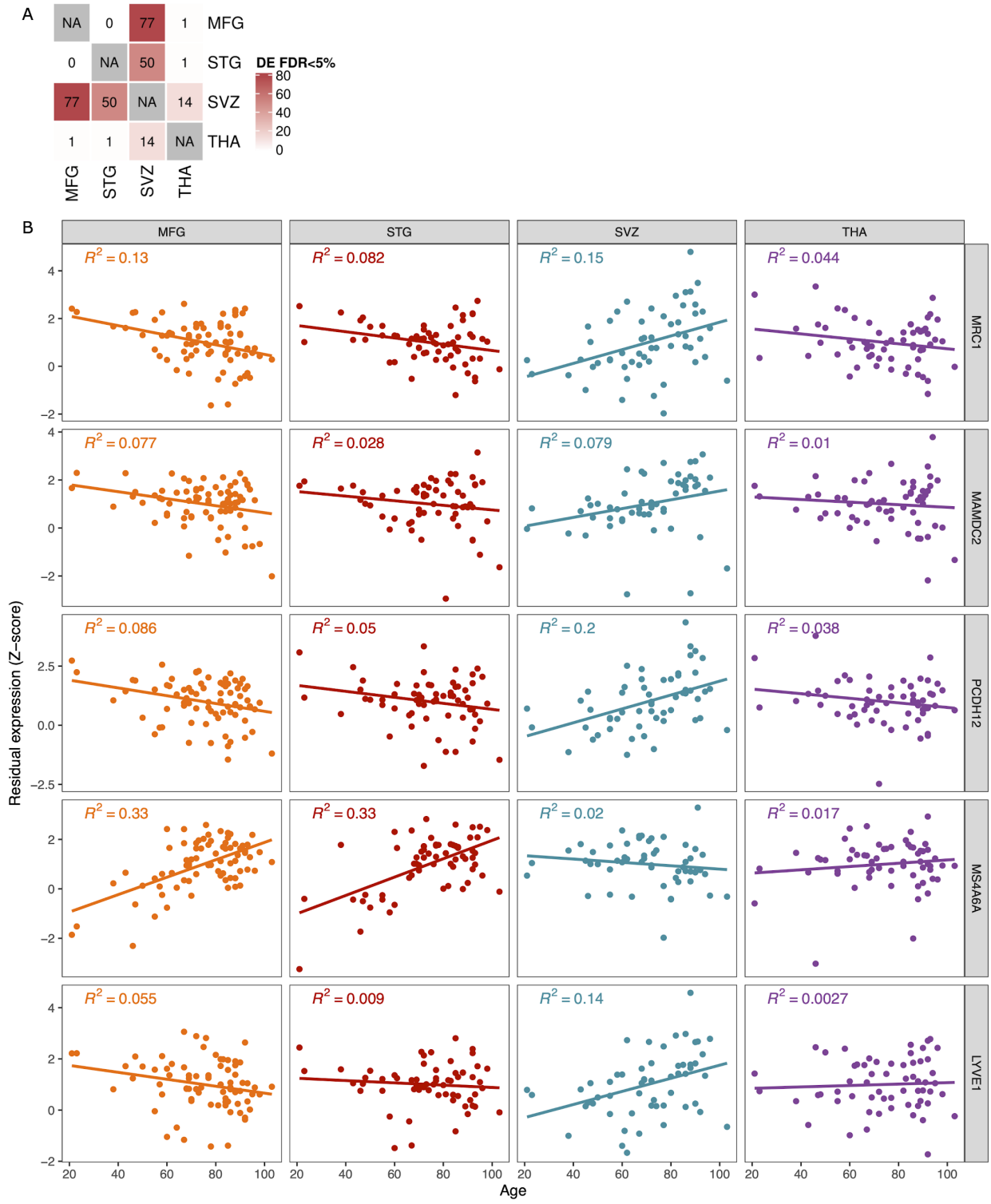

A
B
C
D
E
F
G
H
PC1 (37.97%)
PC2 (7.4%)
PC1 (12.43%)
PC2 (8.87%)
Region
MFG
STG
SVZ
THA
Diagnosis
Control
Dementia
Encephalitis
Epilepsy
Fahrs syndrome
Multiple sclerosis
Parkinson
PSP
Psychiatric diagnosis
PMI
log10(minutes)
Age



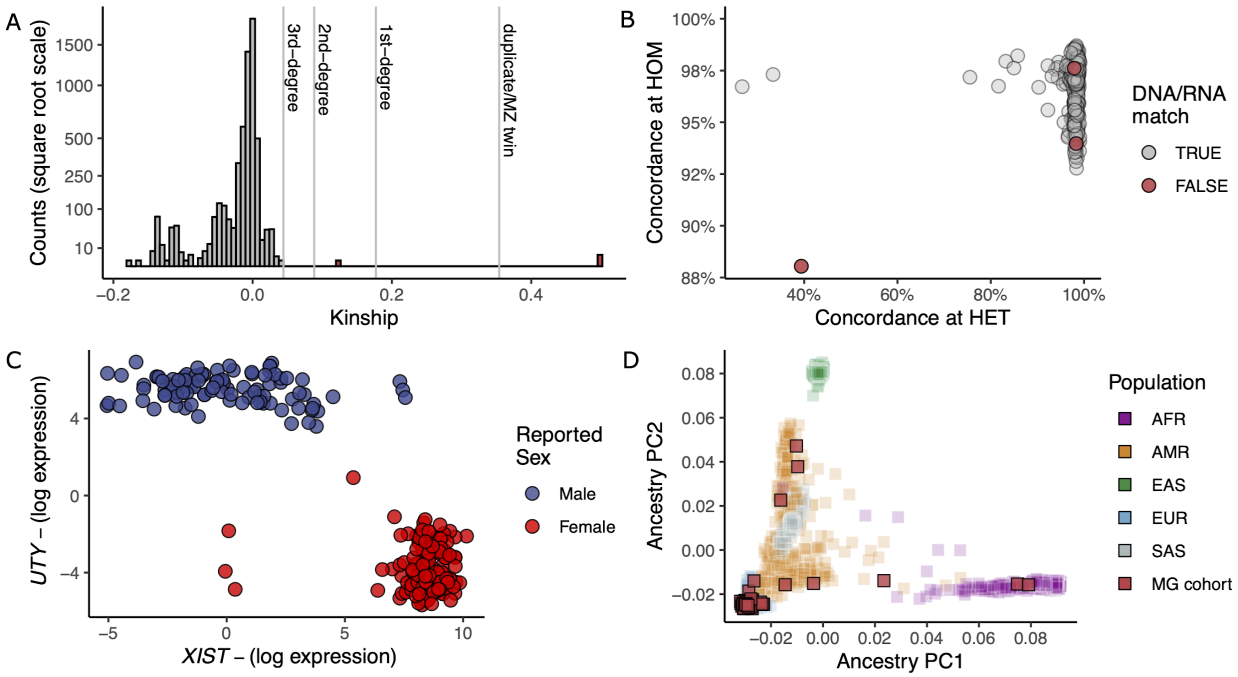
Supplementary Figure 7: Principal component analysis (PCAs) and data adjustment.
Pair of PCA before (left A,C,E,G) and after (right B,D,F,H) correction by regressing out technical confounders, colored by A-B) region; C-D) diagnosis; E-F) post-mortem interval (minutes); G-H) age; I-J) sex. The plots show voom-TMM normalized expression for all samples.



Supplementary Figure 8: Replication of MiGA age-related genes with peripheral human blood. Upregulated in age between MiGA and Peters et al[26]. (OR 7.14; 95% CI = 5.21-9.66, P<1e-16). Downregulated in age between MiGA and Peters (OR = 1.49; 95% CI = 1.17-1.88; P = 6e-4).

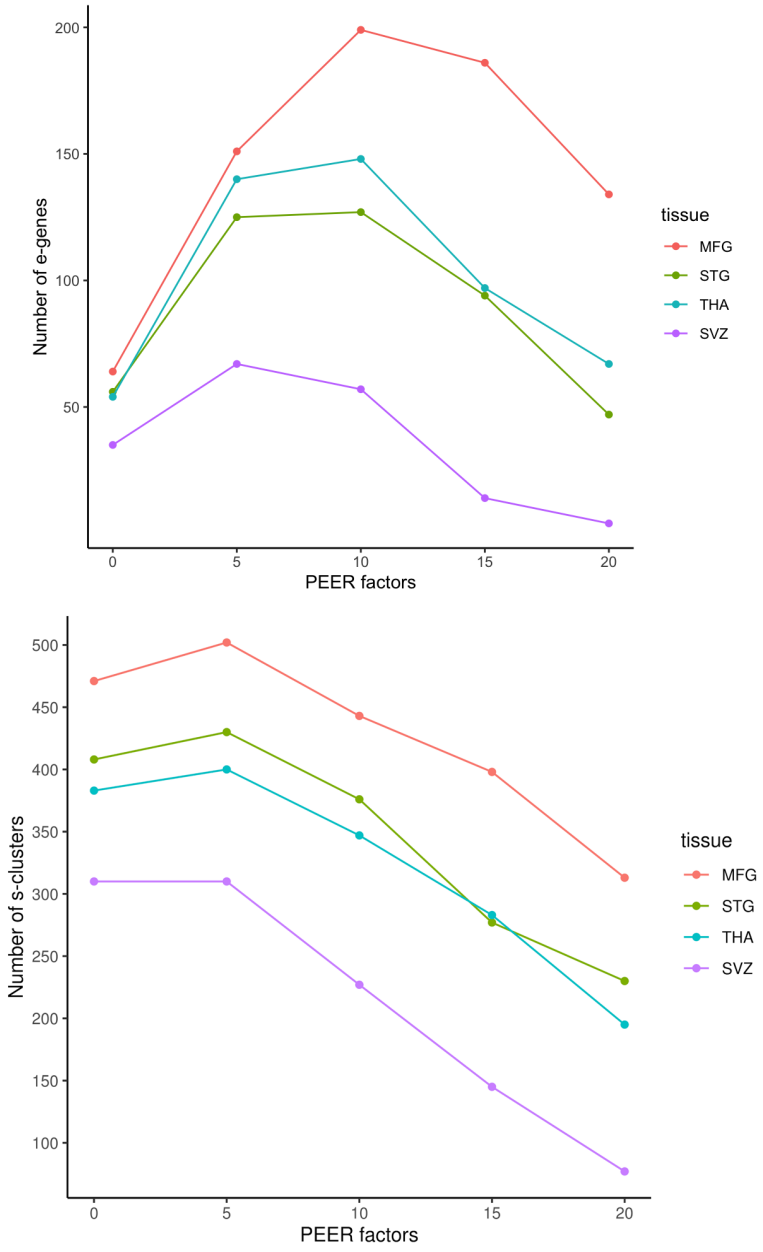


Supplementary Figure 9: Differential expression by age interaction with brain regions.
A) Number of pairwise DE genes (FDR<0.05) using a linear model with interaction between age and region. B) Top five genes prioritized by the DE analysis to show flipped effects in direction of gene expression per age and region.



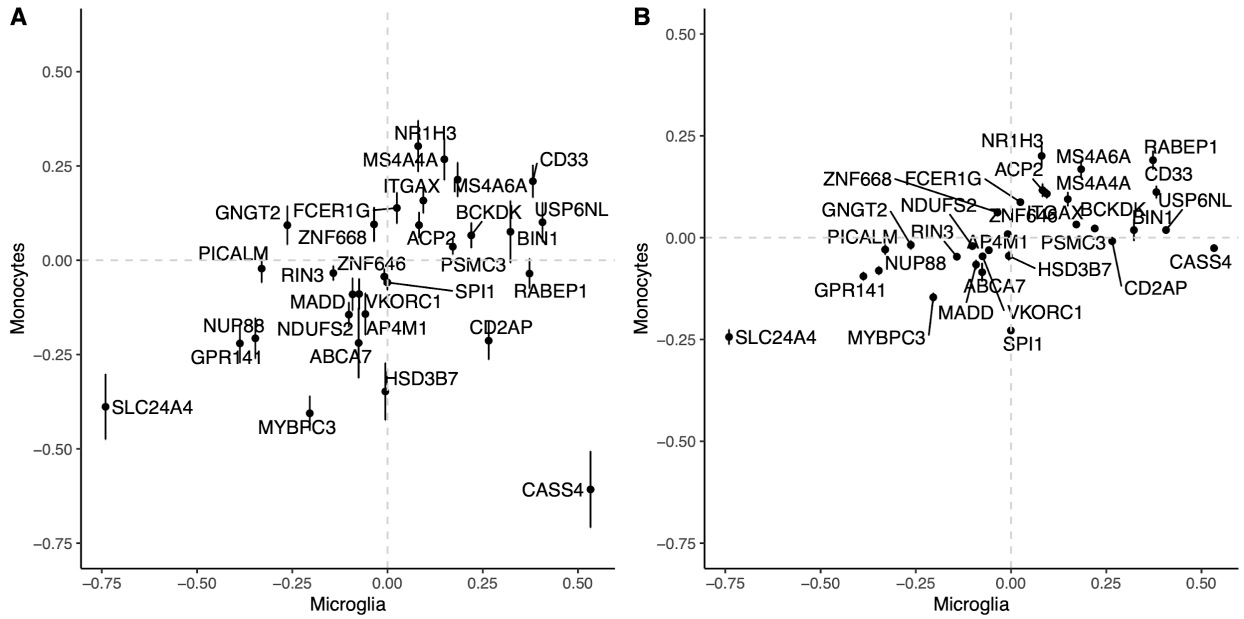
Supplementary Figure 10: Genotyping QC.

A) Distribution of number of samples per estimated kinship measured using KING. Estimated kinship coefficient ranges of >0.354, [0.177, 0.354], [0.0884, 0.177] and [0.0442, 0.0884] corresponds to duplicate or monozygotic twins, 1st-degree, 2nd-degree, and 3rd-degree relationships respectively. B) DNA-RNA sample matching using QTLtools-mbv measured by the percentage of concordance at heterozygous and homozygous genotypes. Samples in red failed to match ids between each data. C) Log scaled expression (*voom*) of sex chromosome genes (*UTY* and *XIST*) for each donor colored by reported sex (blue = male, red = female). D) First two ancestry principal components of MiGA samples (red squares) on top of 1000 Genome individuals colored by major populations. African [AFR], Admixed American [AMR], East Asian [EAS], European [EUR], South Asian [SAS].



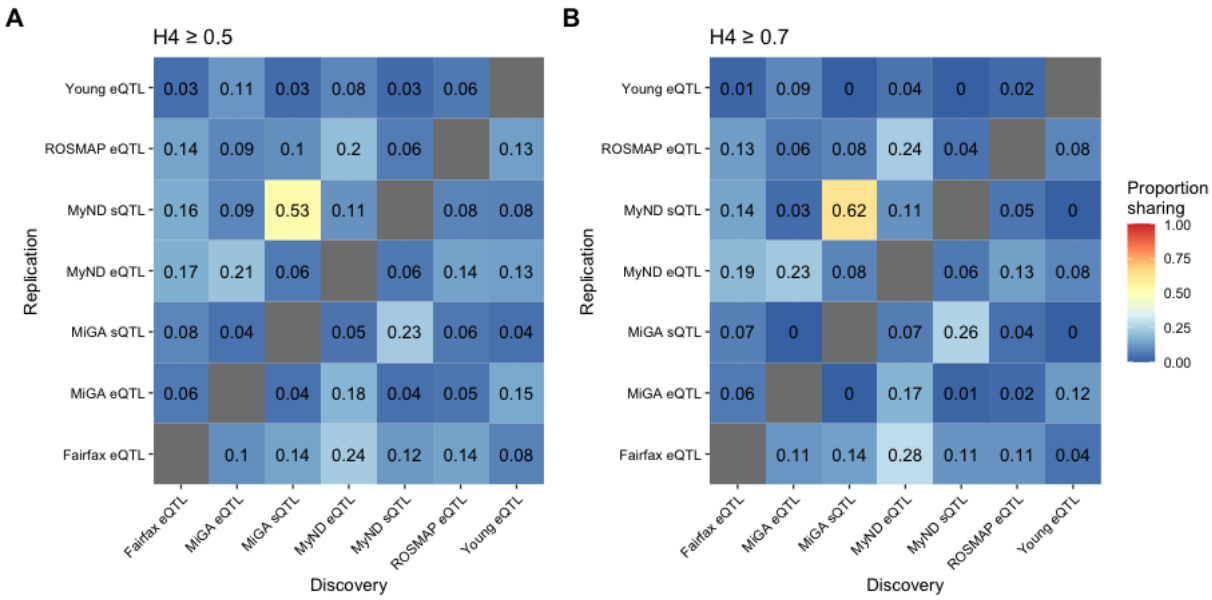
Supplementary Figure 11: Probabilistic Estimation of Expression Residuals (PEER) correction for non-genetic factors in eQTL and sQTL analyses.

To account for hidden effects in gene expression data, such as technical artifacts, we used PEER. After adjusting for age and sex, a number of PEER factors (from 0 to 20) was tested to maximize eQTLs per gene (top) and splicing QTLs per intron cluster (bottom) discovery at q-val < 0.05. Based on this result, 5 (SVZ) and 10 (MFG, STG, and THA) PEER factors were selected for eQTLs and 0 (SVZ) and 5 (MFG, STG, and THA) PEER factors for sQTLs. Factors were regressed out of the respective gene expression or splicing matrix before running the association tests.



Supplementary Figure 12: eQTL effect sizes of the AD-associated genes.

A) x axis shows the eQTL betas from microglia dataset from this study (MiGA), and y axis shows the eQTL betas of monocytes from MyND dataset[27]). B) x axis shows the eQTL betas from microglia (MiGA), and y axis shows the eQTL betas of monocytes from Fairfax et al. Vertical lines indicate standard error.



Supplementary Figure 13: Pairwise sharing of colocalized genes.

Locus-Gene combinations across all 5 diseases that colocalized at PP4 >= 0.5 (**A**) or >= 0.7 (**B**) were compared between pairs of QTL datasets. For example, 53% of genes that colocalized PP4 > 0.5 in MiGA sQTLs also colocalized in MyND sQTLs, whereas only 23% of the MyND sQTLs were found in the MiGA sQTLs.


Schizophrenia
GWAS: Ripke et al, 2014; PP4 ≥ 0.5
dataset Microglia Monocytes Brain type eQTL sQTL
ACVR2A 0.06 0.58 0.06 chr2_149429178_D
TMCO6 0.04 0.27 0.08 0.12 0.8 0.73
PCDHGA9 0.06 0.51
PCDHA13 0.59
NDUFA2 0.06 0.07 0.71 0.09 0.02 0 chr5_140143664_I
IK 0.13 0.07 0.46 0.22 0.73 0.38
DND1 0.89
SNAP91 0.06 0.77 chr6_84280274_D
MAD1L1 0.42 0.05 0.05 0.07 0.54 0.08 chr7_2025096_I
CLCN3 0.07 0.05 0.32 0.92 rs10520163
SNX19 0.73 0.17 0.02 0.99 0.97 0 0.46 rs10791097
NT5C2 0.07 0.65 0.23 0.05 0.96 0.76 0.01
BORCS7-ASMT 0.94 rs11191419
BORCS7 0.05 0.11 0 0.94
CUL3 0.07 0.93 0.38 0.35 0.05 0.89 rs11685299
ZNF804A 0.05 0.05 0 0.97 rs11693094
SETD6 0.06 0.07 0.33 0.68 0.23 rs12325245
CNOT1 0.09 0.97 0.1 0.15 0 0
TMEM219 0.13 0.44 0.13 0.1 0.53
TAOK2 0.06 0.05 0.27 0.85
SMG1P5 0 0.86 0 0.98
SEPT1 0.62 0.15 0.07 0
MAPK3 0.08 0.06 0.65 0.07 0 0.94
KCTD13 0.1 0.16 0.05 0.14 0.75 0.27 rs12691307
INO80E 0.06 0.07 0.07 0.12 0.58 0.4
GDPD3 0.69 0.49 0.21
ASPHD1 0.35 0.92
AC120114.1 0.4 0.07 0.5
AC106782.4 0.55
AC106782.3 0.55
STAT6 0.1 0.09 0.12 0.05 0.86 0 0.01 rs12826178
CDK4 0.66 0.05 0.05 0.04
ZFYVE21 0.03 0.09 0.06 0.24 0.06 0.72
PPP1R13B 0.08 0.59 rs12887734
MARK3 0.04 0.12 0.05 0 0.68 0 0
IMMP2L 0.93 0.12 0.4 0.97 0.57 rs13240464
IFRD1 0.82
KCNQ5 0.55 rs1339227
AC074131.1 0.58 rs16867576
RGS6 0.05 0.06 0.09 0.09 0.05 0.91 rs2332700


NEK4 0.08 0.05 0.52 0.04 0.51
MUSTN1 0.06 0.06 0.47 0.98
MIR135A1 0.54
GNL3 0.08 0.97 0.07 0.12 0.98 0.87 0
rs2535627
MPHOSPH9 0.06 0.69 0.07 0.05 0.75 0.07
ARL6IP4 0.04 0.76 0.05 0.99 0.01 0.04
rs2851447
ZNF506 0 0.51 0.02 0.07 0
MEF2B 0.05 0.5 0.06
MAU2 0.05 0.23 0.07 0.16 0.61 0.07
LPAR2 0.92 0.06 0 0 0.07
rs2905426
CISD2 0.11 0.05 0.57
rs35518360
KDM3B 0.06 0.09 0.39 0.84 0.04
FAM53C 0.26 0.05 0.85 0.16 0.04 0
rs3849046
TSNARE1 0.06 0.48 0.05 0.01 0.72 0
rs4129585
TBC1D5 0.06 0.08 0.94
rs4330281
SGSM2 0.05 0.19 0.09 0.92 0.03 0
rs4523957
MAN2A2 0.08 0.05 0.04 0.22 0.74 0.06
FURIN 0.06 0.15 0.07 0.41 0.25 0.61 0.98
FES 0.03 0.45 0.16 0.59 0.35 0.99 0.14
rs4702
TCTN1 0.06 0.05 0.04 0.79
HVCN1 0.2 0.78 0.03
rs4766428
IVD 0.04 0.96 0.05 0.01 0.07 0 0
GPR176 0.07 0.23 0.57
rs56205728
IRF3 0.8 0.05 0.04 0.44 0.02 0
rs56873913
CYP2D8P 0.05 0.05 0.7
CYP2D7 0.06 0.38 0.92
CYP2D6 0.07 0.38 0.92 0.04 1
rs6002655
ACTR5 0.07 0.06 0.06 0.17 0.09 0.76 0.07
rs6065094
HECW2 0.5 0.04
ANKRD44-IT1 0.3 0.52
ANKRD44 0.52 0.05 0.05
AC010746.1 0.78
rs6434928
SRPK2 0.19 0.06 0.3 0.62 0.07
rs6466055
CSE1L 0.06 0.65
rs7267348
STAG1 0.1 0.1 0.61 0 0.08
PCCB 0.93 0.07 0.95 0.22 0.7
rs7432375
TRANK1 0.06 0.59 0.26 0.97
LRRFIP2 0.18 0.55 0.07 0 0.99 0 0
rs75968099
SDCCAG8 0.22 0.67 0.68 0 0.77 0 0.67
rs77149735
CTSH 0.66 0.07 0 0 0.01
rs8042374
SLC7A6 0.31 0.08 0.15 0 0.12 0 0.68
RRAD 0.06 0.61
PRMT7 0.08 0.12 0 0.06 0 0.98
rs8044995

Supplementary Figure 14: Full colocalization results in schizophrenia.
Colocalization PP4 displayed for each GWAS locus (right text) and gene (left text) for each QTL dataset. An empty value means no QTL was present for testing for that gene in that dataset.

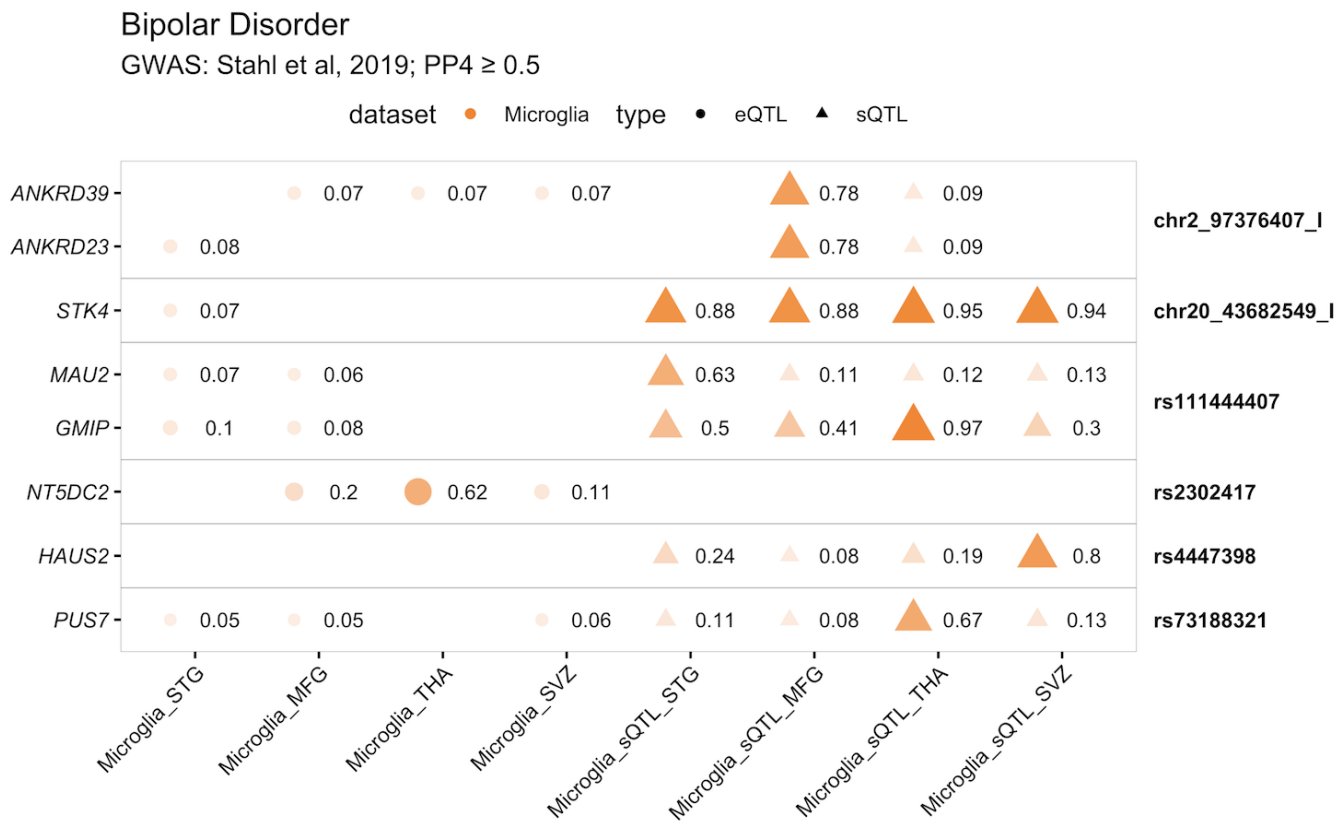

Schizophrenia
GWAS: Ripke et al, 2014; PP4 ≥ 0.5
dataset Microglia type eQTL sQTL
FAM83F 0.07 0.1 0.58 chr22_39987017_D
SNX19 0.16 0.3 0.34 0.55 0.37 rs10791097
NYAP2 0.69 rs11685299
CNOT1 0.08 0.06 0.86 0.12 0.14 0.22 rs12325245
SMG1P5 0.04 0.07 0.08 0.36 0.65 0.32 0.18 rs12691307
DOC2A 0.08 0.08 0.13 0.16 0.55
NT5DC2 0.14 0.6 0.11 rs2535627
MPHOSPH9 0.06 0.74 0.9 0.12 0.28 rs2851447
ARL6IP4 0.09 0.06 0.06 0.06 0.04 0.24 0.25 0.72
MAU2 0.06 0.07 0.56 0.1 0.12 0.11 rs2905426
CDC23 0.55 rs3849046
TBC1D5 0.07 0.06 0.22 0.22 0.57 rs4330281
RAP1GAP2 0.65 0.07 rs4523957
UNC45A 0.05 0.06 0.11 0.68 0.09 0.11 rs4702
FES 0.07 0.05 0.86 0.95 0.13 0.3
PUS7 0.09 0.06 0.09 0.67 rs6466055
SAG 0.56 0.07 0.06 0.07 rs6704768
ATG16L1 0.53
CLDN18 0.6 rs7432375
MLH1 0.06 0.06 0.51 0.13
ITGA9-AS1 0.63 rs75968099
GOLGA4 0.06 0.08 0.07 0.13 0.66
SDCCAG8 0.14 0.45 0.09 0.6 0.46 0.38 0.19 rs77149735
SLC9A5 0.11 0.61 0.07 0.07 rs8044995
XPNPEP3 0.06 0.08 0.06 0.99 0.95 0.9 0.99 rs9607782
ACO2 0.06 0.07 0.08 0.57
Microglia_STG
Microglia_MFG
Microglia_THA
Microglia_SVZ
Microglia_sQTL_STG
Microglia_sQTL_MFG
Microglia_sQTL_THA
Microglia_sQTL_SVZ

Supplementary Figure 15: Colocalization results for each regional microglia dataset in schizophrenia.

Colocalization PP4 displayed for each GWAS locus (right text) and gene (left text) for each QTL dataset. An empty value means no QTL was present for testing for that gene in that dataset.


Bipolar Disorder
GWAS: Stahl et al, 2019; PP4 ≥ 0.5
dataset Microglia Monocytes Brain type eQTL sQTL
WDR73 0.06 0.09 0.14 0.96 0.99
SLC28A1 0.06 0.74
AC048382.5 0.05 0.06 0.99
chr15_85357857_I
LMAN2L 0.06 0.08 0.01 0.1 0.2 0.97
CNNM4 0.06 0.08 0.35 0.73 0.09
chr2_97376407_I
STK4 0.98 0.05 0.06 0.96 0 0
DBNDD2 0.08 0.61 0.05 0.07
chr20_43682549_I
MRPS33 0.07 0.62 0.07 0.05 0.09 0 0.02
chr7_140700006_I
CCS 0.06 0.95 0.02 0.96
AP002748.3 0.09 0.05 0.89
rs10896090
MAU2 0.06 0.08 0.06 0.04 0.56 0.04
GMIP 1 0.06 0.87 0 0.04
rs111444407
AC091132.2 0.62
rs112114764
ZCCHC2 0.05 0.06 0 0.13 0.04 0.98
rs11557713
FGF3 0.52
rs12575685
PPM1M 0.43 0.06 0.05 0.61 0.06 0
NEK4 0.08 0.05 0.52 0.04 0.89
MUSTN1 0.06 0.06 0.36 0.86 0.16
rs2302417
MIR135A1 0.54
GNL3 0.1 0.97 0.06 0.18 0.98 0.64 0.01
CD47 0.21 0.12 0.04 0.9 0.76 0.36
rs3804640
LRRC57 0.1 0.89 0.34 0.05 0.07 0.03 0
AC018362.3 0.89 0.07
AC018362.1 0.89 0.19
rs4447398
LRRFIP2 0.1 0.55 0.11 0 0.99 0 0
rs9834970
MiGA eQTL
MiGA sQTL
Young eQTL
MyND eQTL
MyND sQTL
Fairfax eQTL
ROSMAP eQTL

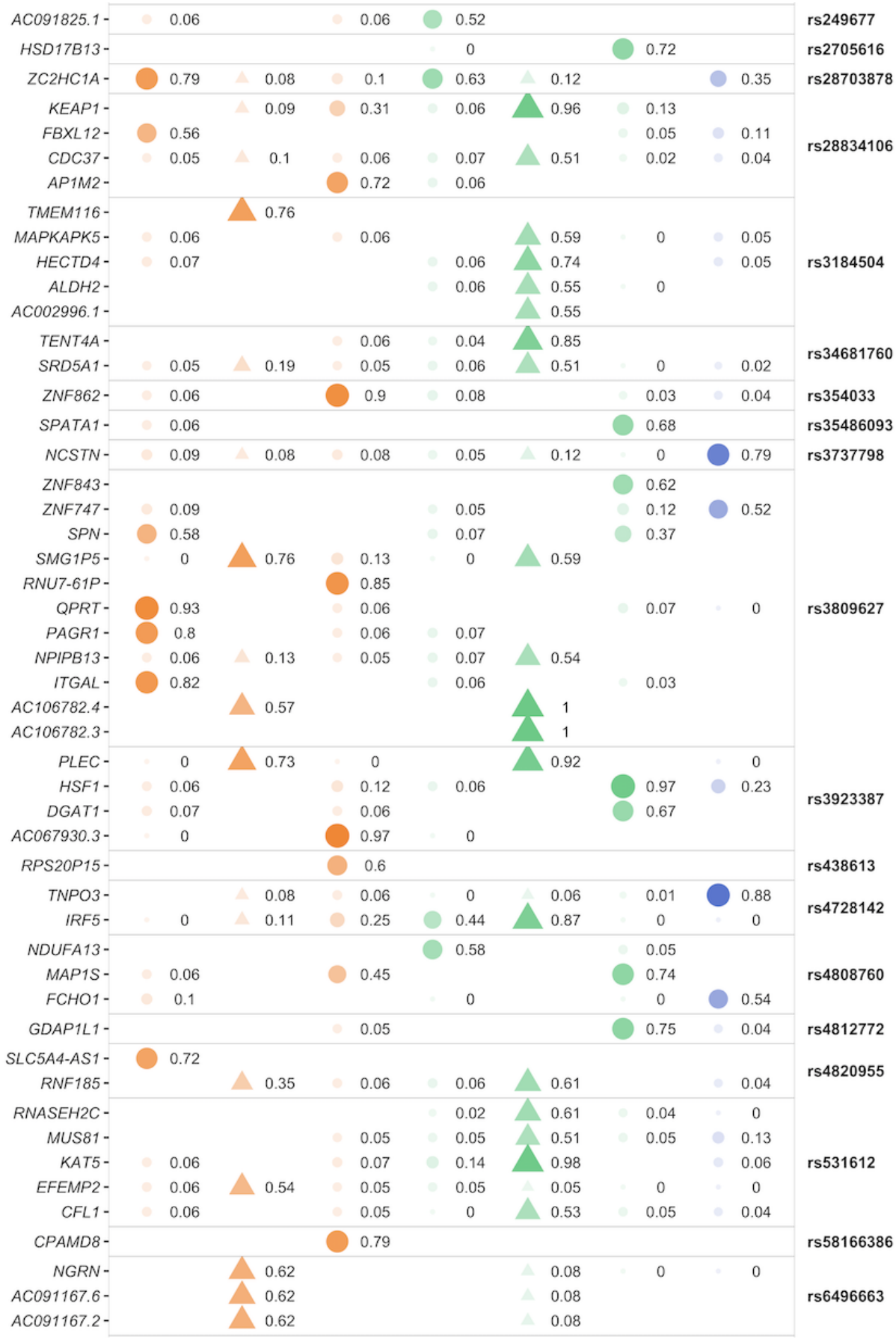
Supplementary Figure 16: Full colocalization results in bipolar disorder.
Colocalization PP4 displayed for each GWAS locus (right text) and gene (left text) for each QTL dataset. An empty value means no QTL was present for testing for that gene in that dataset.

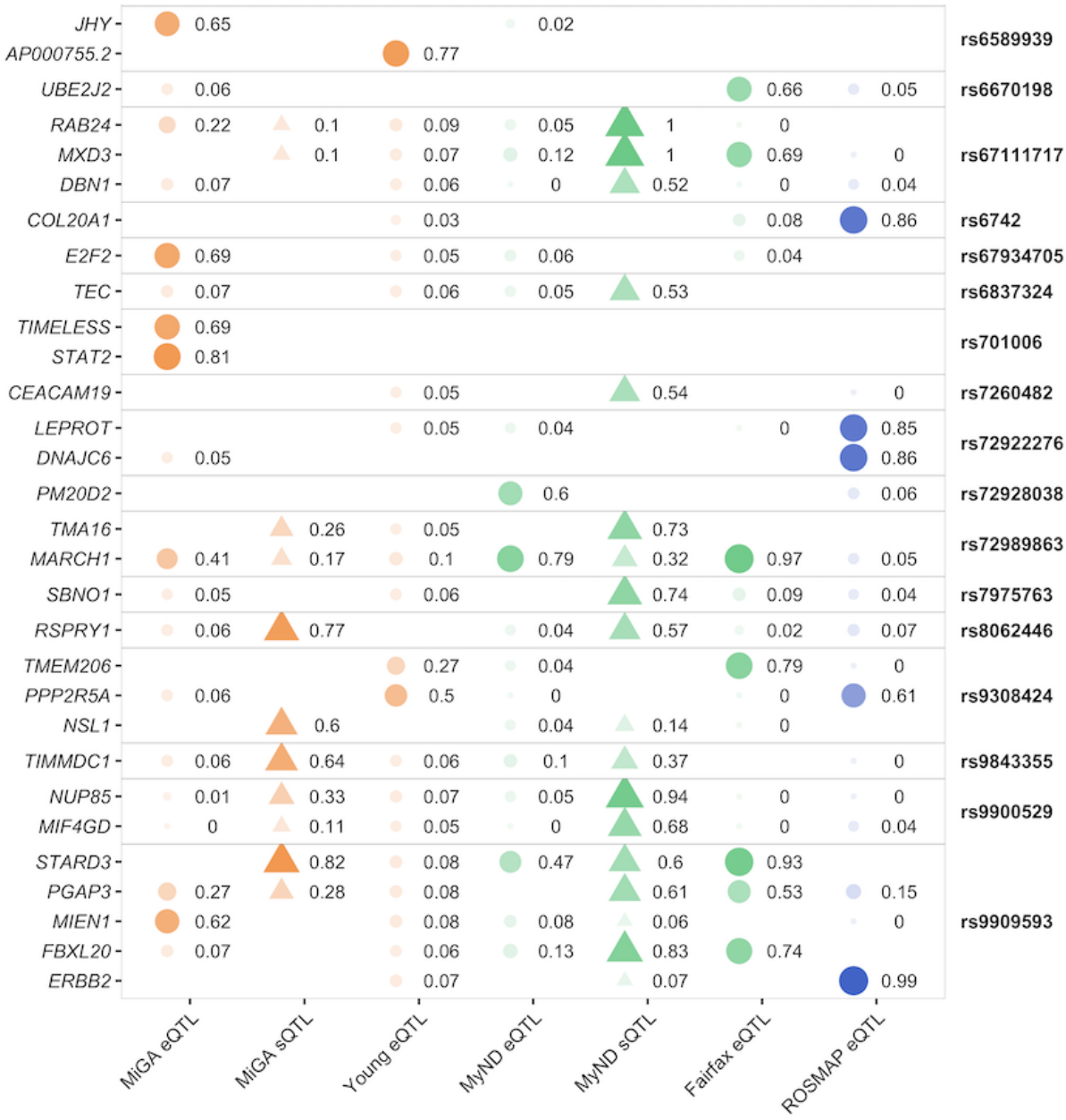


Supplementary Figure 17: Colocalization results for each regional microglia dataset in bipolar disorder.

Colocalization PP4 displayed for each GWAS locus (right text) and gene (left text) for each QTL dataset. An empty value means no QTL was present for testing for that gene in that dataset.


Multiple Sclerosis
GWAS: IMSGC 2019; PP4 ≥ 0.5
dataset Microglia Monocytes Brain type eQTL sQTL
TMEM54 0.71 0.13 0.05
MTMR9LP 0.54 0.08 0.08 0.05 0
chr1:32738415
UBAC2 0.09 0.68 0.15 0.04 0.08 0 0.07
chr13:100026952
CIITA 0.64 0.06 0.05
chr16:11213951
Z99496.1 0.05 0.68 0.1
chr6:119215402
ATXN1 0.56
chr6:14691215
DARS 0.79 0.03 0.02 0.09 0
rs10191360
KLHL11 0.05 0.66
COASY 0.06 0.36 0.05 0.65 0.06
ATP6V0A1 0.28 0.05 0.71 0.03
rs1026916
GPR108 0.04 0.26 0.04 0.04 0.51 0 0
rs1077667
CCDC7 0.54
rs1087056
MRPL45P2 0.2 0.46 0.24 0.07 0.64
KPNB1 0.07 0.05 0.32 0.15 0.65 0.43
EFCAB13 0.12 0.75 0.1 0 0.9
CDC27 0.09 0.76 0.06
rs11079784
HHEX 0.65 0.05 0.07
rs1112718
STIP1 0.29 0.29 0.79
FERMT3 0.07 0.09 0.05 0.11 0.52 0.06 0.02
rs11231749
RPL37 0.31 0.04 0.54
rs11749040
AC016747.1 0 0.05 0.01 0.54
rs1177228
EVI5 0.41 0.25 0.1 0.9 0.08 0.03
rs11809700
CCDC84 0.05 0.25 0.06 0.05 0.75 0 0.01
CBL 0.06 0.11 0.83 0 0.08
rs12365699
MPND 0.05 0.63 0.72 0.51 0.07
rs12971909
LBH 0.07 0.05 0.05 0.52 0.04
rs13414105
PLXNB2 0.86 0.05 0.13 0 0.05
rs140522
PPP1R15A 0.62 0.05 0.04
CD37 0.63 0.07 0.05 0.67 0.01 0.02
rs1465697
SEC24C 0.08 0.16 0.08 0.58 0.08
ECD 0.06 0.69
AGAP5 0.06 0.59
rs17741873
JADE2 0.15 0.06 0.05 0.92
rs2084007
TBC1D3P1-DHX40P1 0.09 0.92
DHX40P1 0 0.92
DHX40 0.08 0.07 0.99 0.08 0.03
rs2150879
PTPRJ 0.65
NUP160 0 0.87 0.17
CELF1 0.07 0.13 0.05 0.65
ARFGAP2 0.08 0.62 0.14 0.14 0.04 0.04
rs2269434


AC091825.1 0.06 0.06 0.52 rs249677
HSD17B13 0 0.72 rs2705616
ZC2HC1A 0.79 0.08 0.1 0.63 0.12 0.35 rs28703878
KEAP1 0.09 0.31 0.06 0.96 0.13
FBXL12 0.56 0.05 0.11
CDC37 0.05 0.1 0.06 0.07 0.51 0.02 0.04 rs28834106
AP1M2 0.72 0.06
TMEM116 0.76
MAPKAPK5 0.06 0.06 0.59 0 0.05
HECTD4 0.07 0.06 0.74 0.05 rs3184504
ALDH2 0.06 0.55 0
AC002996.1 0.55
TENT4A 0.06 0.04 0.85
SRD5A1 0.05 0.19 0.05 0.06 0.51 0 0.02 rs34681760
ZNF862 0.06 0.9 0.08 0.03 0.04 rs354033
SPATA1 0.06 0.68 rs35486093
NCSTN 0.09 0.08 0.08 0.05 0.12 0 0.79 rs3737798
ZNF843 0.62
ZNF747 0.09 0.05 0.12 0.52
SPN 0.58 0.07 0.37
SMG1P5 0 0.76 0.13 0 0.59
RNU7-61P 0.85
QPRT 0.93 0.06 0.07 0 rs3809627
PAGR1 0.8 0.06 0.07
NPIPB13 0.06 0.13 0.05 0.07 0.54
ITGAL 0.82 0.06 0.03
AC106782.4 0.57 1
AC106782.3 1
PLEC 0 0.73 0 0.92 0
HSF1 0.06 0.12 0.06 0.97 0.23 rs3923387
DGAT1 0.07 0.06 0.67
AC067930.3 0 0.97 0
RPS20P15 0.6 rs438613
TNPO3 0.08 0.06 0 0.06 0.01 0.88 rs4728142
IRF5 0 0.11 0.25 0.44 0.87 0 0
NDUFA13 0.58 0.05
MAP1S 0.06 0.45 0.74 rs4808760
FCHO1 0.1 0 0 0.54
GDAP1L1 0.05 0.75 0.04 rs4812772
SLC5A4-AS1 0.72 rs4820955
RNF185 0.35 0.06 0.06 0.61 0.04
RNASEH2C 0.02 0.61 0.04 0
MUS81 0.05 0.05 0.51 0.05 0.13
KAT5 0.06 0.07 0.14 0.98 0.06 rs531612
EFEMP2 0.06 0.54 0.05 0.05 0.05 0 0
CFL1 0.06 0.05 0 0.53 0.05 0.04
CPAMD8 0.79 rs58166386
NGRN 0.62 0.08 0 0
AC091167.6 0.62 0.08 rs6496663
AC091167.2 0.62 0.08

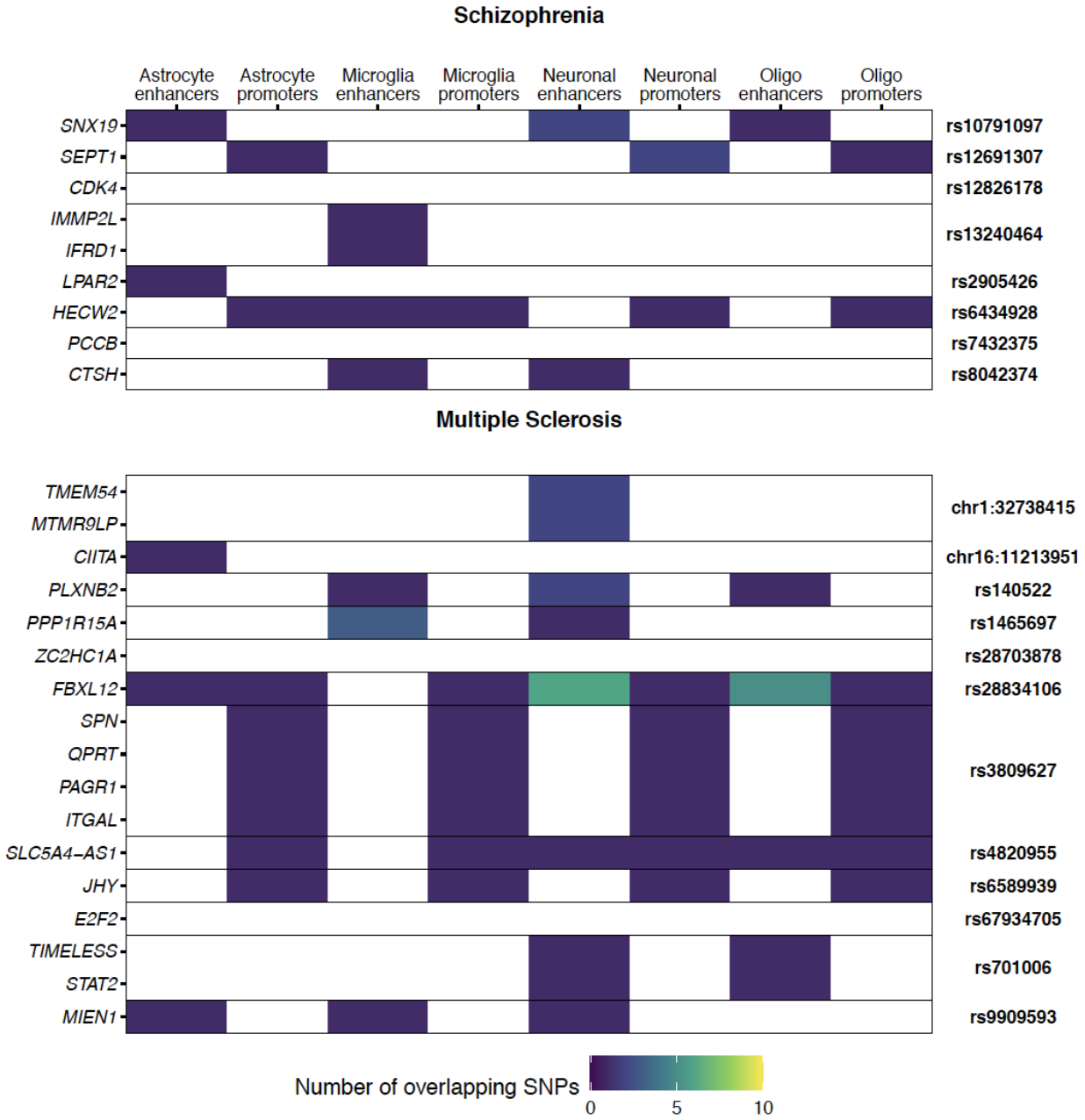


Supplementary Figure 18: Full colocalization results in multiple sclerosis.
Colocalization PP4 displayed for each GWAS locus (right text) and gene (left text) for each QTL dataset. An empty value means no QTL was present for testing for that gene in that dataset.


Multiple Sclerosis
GWAS: IMSGC 2019; PP4 ≥ 0.5
dataset Microglia type eQTL sQTL
EFCAB13 0.13 0.08 0.08 0.07 0.95 0 0.97 0.54
AC068234.1 0.64
rs11079784
C2orf74 0 0 0 0.04 0.07 0.75 0.1
AC016747.4 0.07 0.75 0.07 0.1
rs1177228
YPEL5 0.07 0.06 0.05 0.06 0.09 0.08 0.53 0.11
rs13414105
CD37 0.06 0.66 0.29 0.36 0.09
rs1465697
TNFRSF1A 0.1 0.06 0.09 0.08 0.8 0.11
rs1800693
PPA2 0.08 0.06 0.11 0.45 0.71 0.17 0.15
rs2726479
ZC2HC1A 0.16 0.64 0.24 0.08 0.1
rs28703878
TMEM219 0.07 0.06 0.06 0.5 0.08 0.09
SMG1P5 0.03 0.07 0.07 0.07 0.13 0.69 0.23 0.16
rs3809627
AC009065.4 0.62 0.07
rs405343
TMEM87B 0.06 0.05 0.07 0.84 0.21 0.1
rs57116599
ASAP3 0.82 0.1 0.06 0.06
rs67934705
OS9 0.09 0.06 0.34 0.77 0.19 0.09
rs701006
SH3BP1 0.06 0.59 0.07
rs760517
NUP85 0.08 0.11 0.26 0.59 0.24 0.12
GGA3 0.05 0.09 0 0 0.54 0.42 0.66 0.24
rs9900529
Microglia_STG
Microglia_MFG
Microglia_THA
Microglia_SVZ
Microglia_sQTL_STG
Microglia_sQTL_MFG
Microglia_sQTL_THA
Microglia_sQTL_SVZ

Supplementary Figure 19: Colocalization results for each regional microglia dataset in multiple sclerosis.

Colocalization PP4 displayed for each GWAS locus (right text) and gene (left text) for each QTL dataset. An empty value means no QTL was present for testing for that gene in that dataset.



Supplementary Figure 20: Overlap of colocalized microglia eQTLs with epigenomic features in SCZ and MS.

Regions defined by Nott et al (2019) as cell-type specific promoters and enhancers were overlapped with SNP sets for each colocalizing microglia QTL - GWAS locus. SNP sets consisted of the lead GWAS SNP, the lead QTL SNP and any fine-mapped consensus or credible SNPs. Results are summarized here by the number of SNPs in the set that overlap with a particular feature type.