Comparative gene regulatory network analysis in Alzheimer's disease and major depressive disorder

Hanne Puype[1,2,3], Joke Deschildre[1,2,3], Vanessa Vermeirssen[1,2,3]

1 Lab for Computational Biology, Integromics and Gene Regulation (CBIGR), Cancer Research Institute Ghent (CRIG), Ghent, Belgium

2 Department of Biomedical Molecular Biology, Ghent University, Ghent, Belgium

3 Department of Biomolecular Medicine, Ghent University, Ghent, Belgium

Summary

Major depressive disorder and Alzheimer's disease are two prevalent and devastating disorders that still lack effective treatment. Despite distinct disease phenotypes, these two disorders are characterized by similar neuroinflammatory processes. However, it is not clear to what extent these processes are shared. To compare pathways and their regulators affected in Alzheimer's disease and depression, we adopted a systems biology approach. Consensus gene regulatory networks were built from RNA-seq data with GENIE3, CLR and Lemon-Tree. The two consensus networks were functionally characterized and compared to one another. Further, single-cell gene regulatory networks were constructed with SCENIC, to further characterize the regulatory programs active in specific cell types. In both disorders, aberrations in immune mechanisms and microglial activation were found, controlled by the same transcription factors (IKZF1, IRF8, NFATC2, RUNX1, SPI1, and TAL1). These transcription factors have been implicated in Alzheimer's disease before, while only one has been associated with depression hitherto. Moreover, in Alzheimer, dysfunctions in mitochondria, the proteasome and myelination were discovered. In depression, astrocyte function, GABA receptor signaling, and transport across the blood-brain barrier were aberrated. These disrupted pathways and transcription factors should be further investigated and might be used to develop novel therapies for neuroinflammatory disorders.

Keywords: gene regulatory networks, network inference, neurological disorders, brain, microglia, Alzheimer, depression

Introduction

Alzheimer's disease (AD) and major depression disorder (MDD) have a high burden on global health and still lack effective treatment. Recent insights indicate that several neurodegenerative and neuropsychiatric disorders are characterized by similar processes such as neuroinflammation[1]. However, despite shared pathophysiological mechanisms, these disorders ultimately display distinct phenotypes. Neuroinflammation is characterized by infiltrating leukocytes in the central nervous system (CNS) and activation of microglia[2], brain-residing macrophages that are crucial for the development and functioning of the brain[1]. On the other hand, microglia can release pro-inflammatory mediators, such as cytokines, causing neuronal and blood-brain barrier damage[3], and recruiting pro-inflammatory immune cells, exacerbating this neuroinflammation[3]. However, whether neuroinflammation is causative or consequential, and to what extent it plays a role in these disorders, is still not elucidated.

AD, which is mainly characterized by extracellular β-amyloid (Aβ) plaques and intracellular neurofibrillary tangles, is the most prevalent cause of dementia[4,5]. It is now clear that microglia also have a considerable influence on AD[6]. Normal functioning microglia protect against the development of the disease[6], as they are essential for the clearance of Aβ. Moreover, the majority of identified risk genes for AD are preferentially expressed in microglia[6]. Furthermore, activated microglia can induce a neurotoxic phenotype in astrocytes, leading to neurodegeneration[7]. Different neurodegenerative disorders are accompanied by psychiatric comorbidities. In AD, for instance, depression and anxiety are frequently seen, next to sleep disturbances[4,5].

MDD is a psychiatric disorder of which the pathophysiology is still largely unknown, and that is influenced by both genetic and environmental factors. The disease is characterized by decreased concentrations of the neurotransmitters serotonin, noradrenaline and dopamine in synaptic clefts[8]. In addition, the hypothalamic-pituitary-adrenal axis is defective[9]. Moreover, there are significant reductions in the number and density of astrocytes in several brain regions, and there is astrocyte hypotrophy[10]. There is also mounting evidence for the implication of the immune system in depression[8]. Cytokines such as IL-1, IL-6 and TNF-α are overexpressed and there are higher levels of T-cells and neutrophils in the CNS and periphery of MDD patients[8,11,12]. Activated microglia can lead to chronic inflammation. Consequently, a neuroinflammation hypothesis is rising. However, as depression is highly heterogeneous in different patients, so is neuroinflammation[11], with higher inflammatory levels correlating with treatment resistance or a more severe phenotype[12].

Network inference retrieves systems-wide insights into disease mechanisms from omics data. Gene regulatory networks (GRNs) map the directed, molecular interactions between

transcription factors (TFs) and target genes[13,14]. We and others have shown that integrating different omics data provides a more accurate, multi-modal view of gene regulation[13,14]. In this respect, adding TF-binding information to transcriptome data in network inference allows going beyond statistical association. There are several methods to infer GRNs, which are based on different underlying assumptions such as correlation, information theory, Boolean networks, Bayesian networks, regression, differential equations, and multi-omics integration methods[15,16]. Different network inference methods reveal complementary aspects of the underlying GRNs and a consensus solution provides a more robust and accurate representation of the true GRNs[13,15,17]. Researchers have already conducted network inference on several neuroinflammatory disorders[18,20–22]. The popular method WGCNA[23] builds co-expression networks, which consist of undirected interactions. Moreover, many researchers only infer networks on differentially expressed genes (DEGs), losing a substantial part of the data. One research group executed a system-level analysis of different neurodegenerative disorders[20]. Using microarray data from AD, Parkinson's disease, Huntington's disease, schizophrenia, amyotrophic lateral sclerosis, and multiple sclerosis patients and control samples, they extracted the DEGs and their TFs, and integrated validated protein-protein interactions. However, the number of samples they used was limited; around ten patients and ten control samples were used for every disease. In another study, researchers investigated schizophrenia, bipolar disorder, MDD, AD, and autism patients by reconstructing a GRN by using DNase-seq and TF binding sites data from the human brain and disease-specific DEGs from transcriptomic data[21].

In this study, we aimed to investigate system-wide common and distinct regulatory programs and pathways between the neuroinflammatory disorders AD and MDD. Therefore, we inferred GRNs using different computational methods. Using rank aggregation, a consensus network per disorder was constructed from these networks. The consensus networks of AD and MDD were compared to each other using different topology metrics and were functionally analyzed. In addition, single-cell GRNs were inferred by SCENIC using single-cell RNA-sequencing (scRNA-seq) data, allowing for the characterization of the cells that are implicated in the disorders. Neuroinflammation and activation of microglia were found in both AD and MDD, while specific disrupted pathways were discovered in one of the disorders. In AD, dysfunctions were discovered in mitochondria, the proteasome, and myelination by oligodendrocytes, while in MDD, dysfunctions were found in GABA neurotransmission, neurodevelopment, transport across the blood-brain barrier, and astrocyte functions. Furthermore, six common TFs were discovered playing an important role in microglial activation.

Methods

Retrieval and preprocessing of data

The following publicly available RNA-seq datasets were retrieved: GSE174367/syn22130832[24], GSE101521[25], and GSE80655[26]. We selected high-quality and biologically relevant datasets for our study as follows: human post-mortem brain samples, from the prefrontal cortex, with each dataset containing at least 20 diseased and 20 control samples, not enriched for certain cell types (e.g. for neurons or glial cells), and with a similar number of samples for each disorder.

R (version 4.0.3) was used with the *edgeR*[27] package for preprocessing. Raw counts were filtered to keep only genes with a counts per million (cpm) value greater than one in at least five samples, and were normalized by library size[28], and logarithmically transformed (prior count of one). Batch effects and outliers were detected with multidimensional scaling and hierarchical clustering. Before normalization, different datasets of the same disease were merged to make a compendium per disease, and a batch correction was executed after normalization, with the *removeBatchEffects* function from *edgeR*. Subsequently, protein-coding genes and highly variable genes were selected, respectively. Afterward, TFs with expression data that were left out due to filtering, were added again[29]. This ultimately resulted in 9724 and 9217 genes in the AD and MDD datasets, respectively. Lastly, scaling and centering by gene were performed. In the AD dataset, six outliers out of 95 samples were omitted: four AD samples and two controls. Additionally, in the first MDD dataset, one MDD outlier sample was deleted.

Bulk network inference

For all network inference methods, the TFs were defined as the TFs from Lovering et al.[29] (1455 TFs in total). We ran GENIE3[30] (R package *GENIE3*) with random forest and default parameters. CLR (*minet* package in R (v.4.0.3)[31]) was run with the Miller-Madow estimator and the equal frequency discretization. Edges were filtered to contain at least one TF. For Lemon-Tree[32] (in Java as a command-line program (v3.1.1), https://github.com/erbon7/lemon-tree), we used 100 permutations for the clustering and a minimum of ten genes per module. The three methods were run in a high-performance computing environment.

Consensus networks

From the different methods, an ensemble network was created by average rank aggregation using the *TopKLists*[33] R (v.4.1.3) package with the *Borda* function. The overlap in predicted TFs was calculated using the Jaccard similarity index, which was then used to group the genes into modules of coregulated genes through k-medoids clustering, using the R package *cluster*[34] with the *pam* function. The optimal number of clusters was calculated using the average silhouette width and Calinski-Harabasz index, with the *NbClust* package and *NbClust* function in R. However, in the *NbClust* function it was not possible to use k-medoids and the Jaccard similarity index, thus we approximated this by using the median and k-means methods and the

default dissimilarity metric on the scaled counts. As in all cases the optimal number of clusters was close to the number of clusters in the Lemon-Tree solution, i.e. 155 and 156 modules for AD and MDD respectively, we selected those numbers. There was for both the AD and the MDD networks one cluster with more than 1000 genes, and nine modules with more than 100 genes. The two large modules were briefly inspected and then omitted. They both contained genes of which the significant (adjusted p-value ≤ 0.05) functional enrichment terms indicated regulation of transcription and DNA binding, indicating the modules contained mostly TFs and co-factors.

TFs were assigned to each cluster as follows: TFs were ordered by the number of genes they regulate in the module, and they had to regulate at least half of the genes in the module. A maximum of ten TFs were chosen to be allocated to each module.

Visualization

Venn diagrams were made in R (v.4.1.3) with the package *VennDiagram*[35]. The modules from the consensus networks were visualized with Module Viewer[13] (https://github.com/thpar/ModuleViewer): a Java program in which the expression of different modules can be visualized, together with its TFs and annotation data. Network visualizations were created with the *igraph*[36] package in R or Cytoscape[37].

Functional characterization

The *enrichR*[38] package in R was used for functional enrichment analysis, with the databases Gene Ontology[39,40] (GO) Biological Process (2021), GO Molecular Function (2021), KEGG (2021)[41], Reactome (2016)[42] and WikiPathways Human (2021)[43]. Additionally, we used *enrichR* to look at the enrichment of target genes of certain TFs, using the libraries of ChEA (2016), ENCODE TF ChIP-seq (2015), TRANSFAC and JASPAR PWMs, and TRRUST Transcription Factors (2019). The terms were filtered to have an adjusted p-value of 0.05 or lower. Moreover, motif enrichment analysis was executed through The MEME Suite (Simple Enrichment Analysis SEA[44], default options) and RSAT (peak-motifs[45]). The upstream sequences of the genes were retrieved with RSAT[46] (retrieve sequence), with default options (sequence type: upstream -3000 to -1). The following databases were used to match the motifs in RSAT peak-motifs: JASPAR (core nonredundant vertebrates, 2020), HOMER (human TF motifs, 2016), cisBP (homo sapiens, 2019), HOCOMOCO (human TFs, 2017), and ENCODE (human TFs, 2018).

GWAS SNPs

Genome-wide association studies (GWAS) summary statistics were downloaded from the NHGRI-EBI GWAS Catalog[47] for Alzheimer disease (MONDO_0004975) and unipolar

depression (EFO_0003761) on 06/12/2022. The summary statistics were inspected for SNPs associated with genes contained in the prioritized modules. More specifically, the SNPs had to be located intragenically or a maximum of 3kb upstream of the gene's transcription start site. Additionally, it was checked whether there are SNPs enriched in certain modules of the networks of both disorders. Webgestalt (WEB-based GEne SeT AnaLysis Toolkit) [48] was utilized, with the method over-representation analysis, as input all the genes associated with SNPs (same conditions as above) for the respective disorder, with the custom database the modules and its genes of the disorder, and with the background all protein-coding genes. Multiple hypothesis testing was corrected for by the Benjamini-Hochberg method with FDR ≤ 0.05.

Network topology

Networks were compared with Venn diagrams and network characteristics (average degree, clustering coefficient, density, diameter, (directed) node betweenness, connected components, and the (directed) average path length), with the *VennDiagram* and *igraph* R (v.4.0.3) packages, respectively. The average degree was normalized for the number of nodes in the network. Additionally, two other distance measures were calculated. These were the Jaccard distance[49] and the Graphlet Correlation Distance (GCD), more specifically, GCD-11[50]. GCD-11 was executed by optimizing the Python scripts (adding brackets to the *print* function and changing *Queue* to *queue*) and using the orca.exe file from the authors. The networks were converted to the Leda format, with the *write_graph* function from the *igraph* package in R. GCD-11 is an undirected method. Hence, the 'duplicated' edges had to be omitted first (A-B versus B-A). In the first step, the Graphlet Degree Vector matrix is computed for each network. In the next step, the GCD is computed between the networks[50]. Additionally, heatmaps were made with the Graphlet Correlation Matrices, by using the output of the first step and calculating the Spearman correlation between the eleven non-redundant orbits[50].

Ground truth datasets

The networks were compared to ground truth regulatory interactions from OmniPath[51] and a frontal lobe network from Marbach et al.[52] (top 100 000 edges). From OmniPath, the gene regulatory interactions were retrieved, with confidence levels A-D. This consists of two main collections: DoRothEA and additional literature-curated interactions. DoRothEA[53] contains literature curated, ChIP-seq, gene expression-derived and TF binding site predicted data[51]. Marbach et al. have constructed 394 cell- or tissue-specific regulatory networks by integrating TF motifs with promoter and enhancer activity data[52]. The frontal lobe region network was chosen (from Synapse syn4956655), as the expression datasets are from the prefrontal cortex. These two GRNs from OmniPath and Marbach et al. were taken together and put into one

ground truth network, keeping only the unique edges and the TFs contained in the list by Lovering et al. (175 179 edges in total).

Single-cell network inference

scRNA-seq datasets from post-mortem human brain samples were from the same region as the bulk datasets, i.e. the prefrontal cortex. Both neurons and glial cells had to be included. Similarly as in the bulk datasets, control samples needed to be included and there had to be a similar number of samples (cells) for each disorder. The publicly available files (AD: GSE174367/syn22130832[24], MDD: GSE144136[54]) were preprocessed in R (v.4.1.3). The *Seurat*[55] package was used for quality control and preprocessing. Cells with too low or too high UMI (unique molecular identifier) counts were omitted. In addition, only protein-coding genes were selected. Subsequently, the dataset was logarithmically transformed (*NormalizeData* function) and the highly variable genes (8000 genes) were selected. Next, the TFs that had been omitted by the filtering, were added again. The cell annotations were taken from the original authors. In addition, SCENIC was executed to infer GRNs from the single-cell data[56]. It was run in Python (*pySCENIC*) with GRNBoost2, on a high-performance computing environment.

Results

Different population-level network inference methods retrieved different topologies, while consensus regulatory programs for AD and MDD were topologically similar

To compare regulatory circuits in AD and MDD, we first ran different GRN inference algorithms on bulk transcriptomic data. We utilized publicly available RNA-seq data from human post-mortem brain (prefrontal cortex) samples retrieved from GEO and Synapse: 46 healthy control and 43 patient samples with AD (GSE174367/syn22130832[24]) and 53 healthy control and 52 patient samples for depression (GSE101521[25] and GSE80655[26]) (Methods, Supplemental Table 1). The prefrontal cortex was the favored brain region, as a large number of studies sample from the prefrontal cortex, and it has been implicated in several psychiatric, neurodevelopmental and neurodegenerative disorders[21]. GENIE3[30], CLR[57], and Lemon-Tree[32] were used to infer networks (Methods). Two networks were made with every method, one with the AD dataset, and one with the MDD datasets. The top 100000 edges were selected from every method-specific network. These networks were then compared to each other, with about 10 000 edges shared by all three methods, for both disorders (Supplemental Figure S1). For an overview of the number of nodes of each network, see Supplemental Tables 2 and 3. To confirm that the different methods give complementary results, we looked at the overlap between the inferred interactions and 175179 ground truth regulatory interactions retrieved from the one hand of OmniPath, which includes, amongst other prior knowledge, known regulatory interactions, combined from different resources[51], and from the other hand of frontal

lobe GRNs inferred by integrating TF sequence motifs with promoter and enhancer activity data from the FANTOM5 project[52] (Methods). The overlap with the ground truth networks is low for all methods and both disorders (around 1.2%, Figure 1A), but there are more edges in common between every method separately and the ground truth (+/-600 edges) than between all three methods and the ground truth (134 edges for AD, 261 edges for MDD, Figure 1A). Topological measures indicate that these networks displayed both similarities and differences (Supplemental Table 4 & 5). There are three connected components in the GENIE3 MDD network, consisting of one main network and two separate interactions, while all the other networks consist of one connected component. Additionally, the clustering coefficient is 5.7 times lower in CLR networks, compared to GENIE3 and Lemon-Tree networks. Further, these networks were visualized with the igraph package in R (Supplemental Figure 2), revealing differences in overall topology. Overall, we observed that different GRN inference methods generated different network topologies.

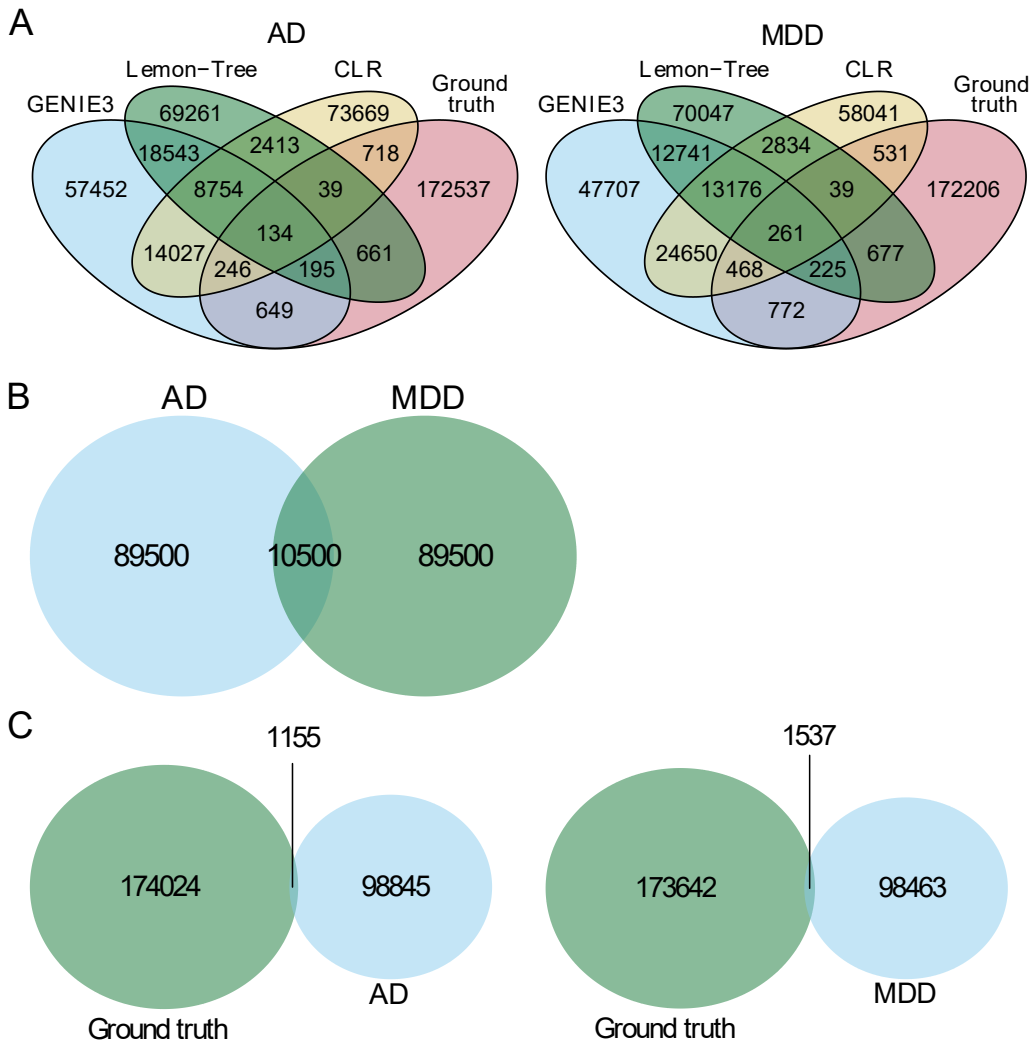


Figure 1. Different population-level network inference methods retrieved different topologies, while consensus regulatory programs for AD and MDD were topologically similar. A) Venn diagrams for the overlap between the three network inference methods (GENIE3, CLR and Lemon-Tree) and the curated regulatory interactions from OmniPath and Marbach et al. (frontal lobe GRNs). Left: overlap of the edges between the three networks for Alzheimer's disease (AD) and the ground truth. Right: overlap of the edges between the three networks for major depressive disorder (MDD) and the ground truth. B) Venn diagram of the edges of the two consensus networks of Alzheimer's disease (AD) and depression (MDD). There are 10500 edges in common. C) Venn diagrams for the overlap between the consensus networks of AD (left) and MDD (right) and the curated regulatory interactions.

Next, to compare the regulatory mechanisms between AD and MDD, we constructed consensus regulatory networks for both diseases, from the three different methods, through average rank aggregation. As such, more robust networks were retrieved. As before, the top 100 000 edges were retained. There were 1041 TFs and 9675 target genes generating a total of 9685 nodes in the consensus AD network (Supplemental Table 2). In the consensus MDD network, there were 1101 TFs, 9006 target genes, and a total of 9021 nodes (Supplemental Table 3). Of these, there were 7042 nodes in common between AD and MDD consensus networks. Topological characteristics (Supplemental Tables 4 & 5) indicated that both networks were similar to each other, and to the networks of the different network inference methods. The diameter of the consensus MDD network was one path longer than the diameter of the AD network. Both consensus networks consisted of one connected component. All topological characteristic values were higher for the MDD network, except for the directed average path length. There were 10500 common edges, which is about ten percent of the total edges (Figure 1B). To see where these common edges are situated in the two networks, histograms were made depicting the rank of these common edges in both networks (Supplemental Figure S3). Most edges are situated in the top ranks, thus with higher confidence. In the AD network, there are about 3500 shared edges in the top 10 000 ranks, while in the MDD network, about 2900 shared edges are in the top 10 000 ranks. The target genes (3331 genes) of these common edges are involved in the regulation of transcription, ion channel activity, synaptic transmission, cell migration, 'nervous system development', 'cellular response to cytokine stimulus', 'immune system', and related terms. Of the top 100 TFs of the consensus networks, there were 34 in common (Supplemental Table 6). Compared to the ground truth networks, the AD network had 1155 edges in common, while the MDD network had 1537 edges (Figure 1C). The two consensus networks were as well visualized with igraph (Supplemental Figure S2). In addition, the networks were compared with distance measures. Firstly, the Jaccard distance calculated between the two networks resulted in a value of 0.9446. The Jaccard distance gives a broad overview of the number of edges that are shared between two networks. The value indicates that there were few edges in common between the two networks (see also Figure 1B). Further, the Graphlet Correlation Distance (GCD)[50] was calculated at 1.063. For more information about this method, see Supplemental material.

Together, the topological measures indicated that the consensus networks of both disorders had a similar structure.

Functional annotation of the modules implicated disruption in immune, mitochondrial, and oligodendrocyte processes

To investigate the biological implications of the inferred GRNs, we disentangled the hairball structure of the GRNs by constructing modules, or clusters of coregulated genes. We hypothesized that these genes are part of the same or similar pathways, as they are regulated by the same TFs. The overlap in predicted TFs was calculated using the Jaccard similarity index, which was then used to group the genes into modules of coregulated genes through k-medoids clustering (Methods). The number of clusters to choose with k-medoids can be hard, especially for a large dataset. We chose to keep the number of clusters that were retrieved with Lemon-Tree, i.e. 155 for AD and 156 for MDD, as similar numbers were also retrieved when calculating the optimal number of clusters using the average silhouette width and Calinski-Harabasz index (Methods). In addition, TFs were assigned to the modules (Methods), where they had to be inferred as regulating at least half of the genes in the module, and with a maximum of ten TFs.

Functional enrichment analysis was performed on all modules using GO Biological Process, KEGG, Reactome, and WikiPathways annotations. Additionally, motif enrichment analysis using RSAT (peak-motifs) and The MEME Suit SEA (Simple Enrichment Analysis), and target gene enrichment of TFs with enrichR were executed (Methods). For both disorders, there were some modules enriched for the 'Alzheimer's disease' term from KEGG, indicating an etiological overlap between AD and MDD. We prioritized modules in two ways. Firstly, modules were searched for being functionally enriched for immune system-related terms (i.e. containing 'immune'/'inflammation' and terms related to cytokines, the complement system, innate and adaptive immune cells…), as both disorders are characterized by neuroinflammation. This resulted in ten modules for AD and twelve for MDD. Additionally, we looked for overlap in genes and TFs between these immune-related AD- and MDD-modules. Secondly, the modules were ordered according to the difference in expression of the module genes between control and disease phenotypes. Only significantly different (Wilcoxon Rank sum test, p-value < 0.05) modules were further inspected (15 for AD, 11 for MDD). Below, some modules that have interesting functionally enriched terms (i.e. more than five terms), and less than 80 genes are further elaborated upon. An overview of all modules can be found in the supplemental material.

In the first section, we go deeper into modules with shared pathways between AD and MDD. The 47 genes in AD-module 22 were functionally enriched for immune-related terms (Figure 2). Even though there was not a substantive difference between the expression in the

Alzheimer's patients and the controls, the control individuals tended to have a lower expression (see Figure 2A). There was an enrichment of several terms related to cytokines (IFN-γ, IL-2, IL-3, IL-6, IL-8, and TNF, Figure 2A). In addition, there were terms related to phagocytosis. TYROBP, which is an adaptor protein that associates with activating receptors of immune cells, has been implicated in AD before[58]. It associates with TREM2, and both are required for phagocytosis in microglia. However, both TYROBP and TREM2 were not part of the module. TREM2 is an important pattern-recognition receptor, involved in microglial phagocytosis, chemotaxis, survival, proliferation, and inflammatory response[6,59]. Microglia lacking TREM2 are not able to fulfill these functions, resulting in the exacerbation of AD. There were as well some terms related to T-cells. Classically activated microglia are activated by IFN-γ or LPS, and secrete the cytokines TNF, IL-1β, IL-2, IL-6, and IL-8, amongst others[60,61]. These microglia attract and activate T-cells, more specifically T_H1 and T_H17 T-cells. Phagocytosis is decreased in this subtype[61]. Higher levels of T-cell infiltration are seen in AD patients, however, their exact role remains unclear[62]. The TFs of this module were IKZF1, NFATC2 and RUNX1. All three genes are regulators of immune functions: NFATC2 is a regulator of cytokine release, and NFATC2 and RUNX1 are both involved in T-cell functioning[63,64]. Mutations in RUNX1 have been associated with increased risk for AD[65]. IKZF1 is involved in hematopoietic cell differentiation[66] and altered binding sites of IKZF1 have been implicated in risk genes for AD[67]. This has been seen for the *ABCA7* and *INPP5D* genes, of which the latter is also a part of this module. *INPP5D* encodes for the SHIP1 protein, which has been implicated in AD[62,67]. It acts as a negative regulator of myeloid cell proliferation/survival and chemotaxis, immune cells homeostasis, it regulates macrophage programming, phagocytosis and activation, and neutrophil migration[68]. Interestingly, this gene was found to be implicated in the activation of microglia in a mouse model of AD[69]. The inferred TFs *RUNX1* and *IKZF1* were also detected in the motif and/or TF enrichment analysis (Methods). Additionally, *SPI1*, *IRF8* and *TAL1* were found (see later in the text). The protein product of *SPI1* also binds *TYROBP, TREM2* and *INPP5D*[70–72]. *SPI1* has been implicated in microglial activation in AD before, particularly in antigen presentation and phagocytosis[73]. Several studies have found mutations that interrupt the binding of TAL1 in patients with AD[74]. Further, TAL1 was found to be implicated in MDD in two studies[75,76]. The enriched TFs and motifs of all modules can be found in the supplemental material.

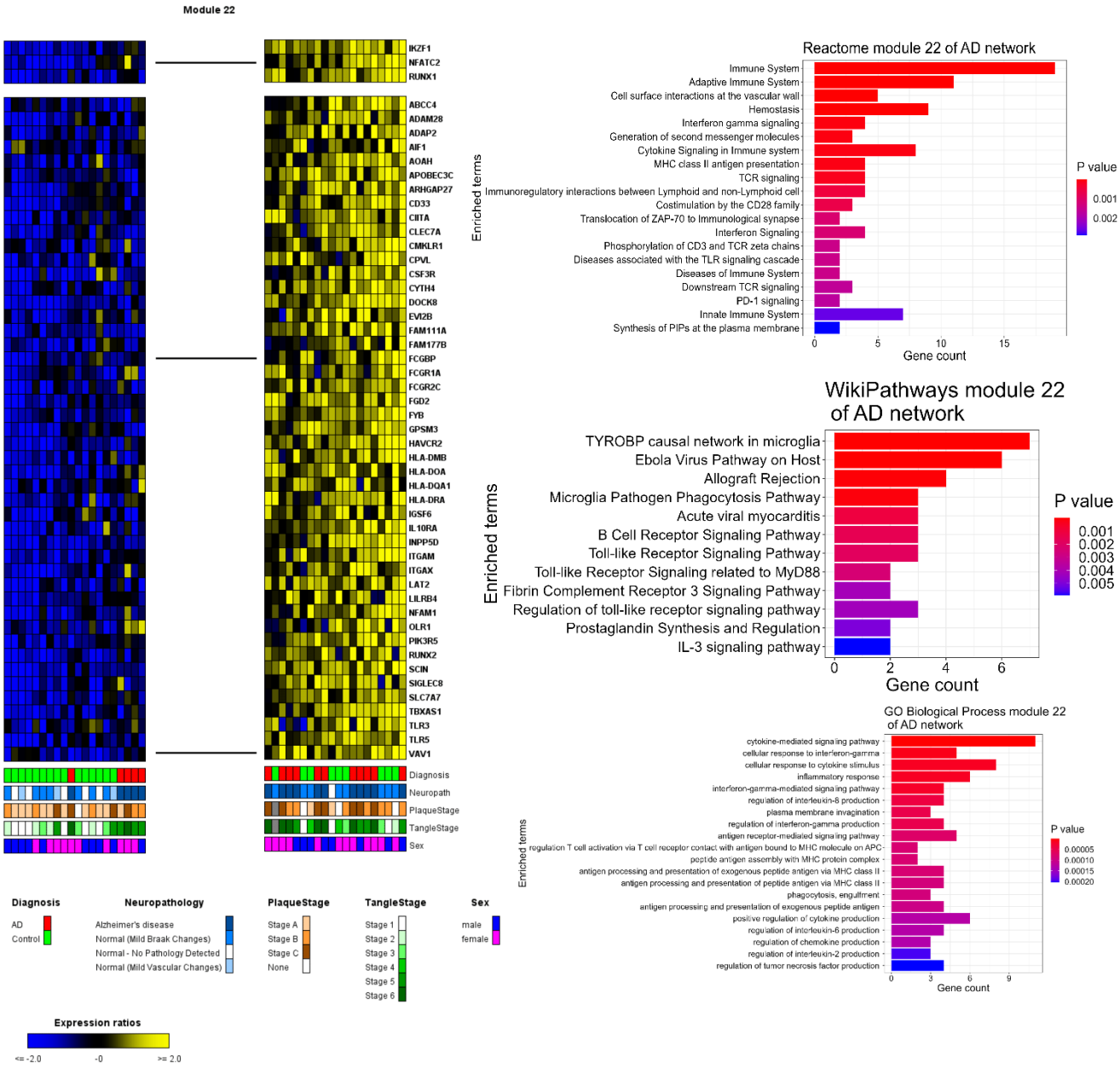


Figure 2. AD-module 22, containing 47 genes and regulated by IKZF1, NFATC2 and RUNX1 functions in cytokine signaling, antigen presentation and phagocytosis. The expression of the module was visualized with ModuleViewer, annotated with sample diagnosis, plaque and tangle stages, and sex (see legend). In addition, the functional enrichment plots from the databases GO Biological Process, Reactome and WikiPathways are pictured. The top twenty terms are depicted, ordered by increasing p-value. The number of genes from the module belonging to the terms is represented on the x-axis. See also Supplemental Figure S4.

Similarly, the 54 genes in MDD-module 24 were enriched for immune-related terms (Figure 3). However, here the expression of the depressive patients was lower than the expression of the controls (p-value 0.0473). There are six genes in common between this module and the AD-module 22: *ABCC4*, *CPVL*, *FCGR1A*, *GPSM3*, *IL10RA*, and *SLC7A7*. *ABCC4* and *FCGR1A* have been implicated in AD, while *CPVL* has been implicated in psychosis and AD[77]. CPVL is

a protease involved in the cleaving of phagocytosed particles in the lysosome, in an inflammatory protease cascade, and the pruning of peptides for antigen presentation[68]. Further, *IL10RA* has been associated with schizophrenia and multiple sclerosis[77]. In addition, IL10 has been associated with a worsening plaque load and reduced Aβ phagocytosis by microglia in mouse models of AD[78]. The TFs of this module are IKZF1, RUNX1 and TFEC. Hence, there are two TFs in common between this module and module 22 from the AD network. In addition, NFATC2 is part of the module. For this module, the inferred TFs IKZF1 and TFEC were also found in the motif enrichment analysis of the upstream promoter regions of the genes (Methods). Additionally, *IRF8*, *SPI1*, *NFATC2*, and *TAL1* were found in the motif and/or TF enrichment analysis. *SPI1* and NFATC2 were not inferred as a regulating TFs of this module, however, they are part of module 24. Interestingly, the protein product of *RUNX1* was found to bind *SPI1*[79]. Moreover, the protein product of *SPI1* was found to bind to the TFs *IRF8* and *RUNX1* in mice[71]. *IRF8* and *TAL1* were not inferred either, but come back in both this module and AD-module 22 as having enriched motifs in the promoter regions of the module genes. Looking at the functional enrichment analysis, there are some similar terms, and some differences between the two immune-related modules, so they do not entirely represent the same pathways. In contrast with the module from AD, the cytokine response (IFN-γ, IL-10) was not dominant here, but rather neutrophil activation stood out (Figure 3). This is in agreement with what has been reported in the literature: neutrophils and IFN-γ are increased in depression, while IL-10 is decreased[12]. IFN-γ is produced by T_H1 cells, while IL-10 is secreted by T_H2 and T_{reg} cells. The α-subunit of the IL-10 receptor (*IL10RA*) was one of the genes shared in the two modules. In addition, there were some enriched terms related to B- and T-cells. B-, T_H2 and T_{reg} cells are decreased, while T_H1 and T_H17 cells are increased in depression[12]. Both modules do have the 'TYROPB causal network in microglia' as the most enriched term, and 'Microglia Pathogen Phagocytosis Pathway' as the fourth and third term, respectively, from WikiPathways. As both modules contain immune-related terms and terms related to microglia, these modules are likely active in microglia and the majority of the reads could be coming from these cells. As the gene expression tends to be higher in Alzheimer's patients, there is more activation of microglia or there are more microglia in AD, compared to control individuals. On the other hand, as the expression of these genes tends to be lower in depressive patients, there is less activation or fewer microglia in this disease, compared to the controls. However, there are some terms with 'negative regulation of' in the MDD module (Figure 3), which might indicate that inhibitors of immune functions are downregulated in depression, resulting in a higher inflammatory response.

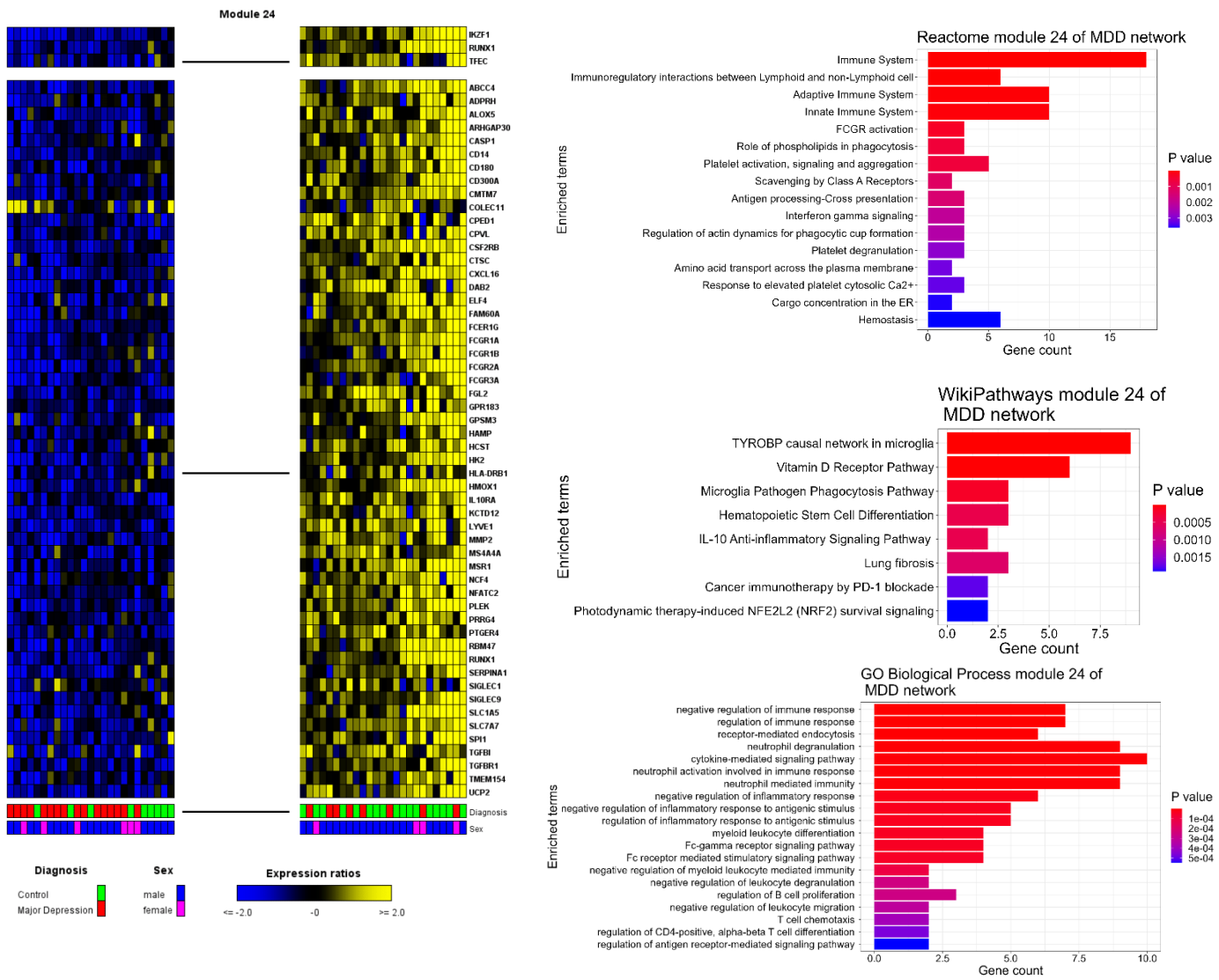


Figure 3. Module 24 from the MDD network, containing 54 genes and regulated by IKZF1, RUNX1 and TFEC, functions in neutrophil activation, cytokine release and phagocytosis. The expression of the module is visualized in Module Viewer, with the corresponding annotation data diagnosis and sex (see legend). In addition, the functional enrichment plots from the databases GO Biological Process, Reactome and WikiPathways are pictured. The top twenty terms are depicted, ordered by increasing p-value. The number of genes from the module belonging to the terms is represented on the x-axis. See also Supplemental Figure S5.

Module 39 (103 genes, Supplemental Figure S6) also had immune-related terms such as neutrophil activation, cytokine production (IL-6, TNF, IFN-γ) and B-cell activation. Additionally, one of the terms was 'amyloid-beta binding'. The genes with this term were *C1QA*, *CD74*, *FCGR2B*, *ITGB2*, *LILRB1*, *LILRB2*, *MSR1* and *TLR6*. C1QA is part of the complement system, the protein of *CD74* binds to class II MHC[80], FCGR2B is an Fc gamma receptor, ITGB2 is part of an integrin[68], LILRB1 and -2 are receptors for class I MHC antigens[81], MSR1 is a macrophage scavenger receptor and TLR6 is Toll-like receptor 6. Thus, these genes are involved in immune mechanisms as well. Similarly as in module 22, the expression of this

module seemed to be higher in AD patients (Supplemental figure S4), however, it is less pronounced. Module 39 had fourteen genes in common with module 24 from the MDD network. These were *ALOX5*, *CASP1* (caspase-1), *CD14*, *CD300A*, *CXCL16*, *FCGR1B*, *FCGR2A*, *FCGR3A*, *KCTD12*, *MSR1*, *PLEK* (pleckstrin), *RUNX1*, *SERPINA1* and *TGFBR1* (TGF-β receptor type-1). *ALOX5* encodes a lipoxygenase, which catalyzes the oxygenation of fatty acids. It is involved in several inflammatory processes, such as the biosynthesis of leukotrienes, dendritic cell migration, and CD40-induced production of reactive oxygen species[68]. CXCL16 is a chemokine and FCGR1B, -2A and -3A are immunoglobulin gamma Fc receptors. KCTD12 is a subunit of GABA-B receptors. MSR1 is again a macrophage scavenger receptor. *SERPINA1* encodes the α1-antitrypsin protein[68]. *ALOX5, CASP1, CD14, SERPINA1* and *TGFBR1* have been implicated with both MDD and AD before[77], while *FCGR2A, FCGR3A, MSR1, PLEK* and *RUNX1*[65] have been implicated with AD, and *KCTD12* with mood disorders[77]. Looking at the TFs of module 22 from the AD network, AD-module 39 had all three in common (*IKZF1*, *NFATC2* and *RUNX1*), plus *IRF8* (Supplemental figure S6). *IKZF1* and *RUNX1* were as well in common with the MDD-module 24. Module 39 had the inferred TFs with enriched target genes and motifs: *IKZF1*, *IRF8*, and *RUNX1*. *TFEC* and *SPI1* also had enriched TFs/ motifs, and had enriched motifs in the genes of AD-module 22 and MDD-module 24. They also have immunological functions. This indicates that this TF also steered these modules, but was not inferred as such.

In addition to the immune-related modules above, two modules had an overlap in genes and TFs, and had a significantly differential expression between cases and controls: module 153 from AD (Supplemental Figure S7) and module 110 from MDD (Supplemental Figure S8). Module 153 of the AD network contained terms related to apoptosis, 'regulation of inflammatory response', and 'cellular senescence'. Patients had a higher expression of the genes compared to control individuals (p-value 0.0258). The TFs were *CEBPD*, *LEF1*, *MAFF*, *NEUROD6*, and *RREB1*. *CEBPD* and *NEUROD6* were as well found in the motif and TF enrichment analysis. Module 110 contains genes involved in the unfolded protein response and response to cellular stress, 'regulation of angiogenesis', and also in cytokine signaling (IL-17, TNF, IL-18, TGF-β, IL-3). Surprisingly, the expression was lower in MDD patients compared to control individuals (p-value 0.0419). The TFs of this module are *ATF3, CSRNP1, FOS, FOSB, HES1, JUN, JUNB, KLF4*, and *MAFF*. *ATF3*, *FOS*, *JUN*, *JUNB*, and *KLF4* had motifs and TFs enriched. FOS and JUN can dimerize to form the AP-1 complex, which supports higher cognitive functions[82,83]. These modules have two genes in common: *RGS1* and *TNFAIP3*. RGS1 regulates G protein-coupled receptor signaling cascades[84], while TNFAIP3 is involved in NFκB signaling and has both ubiquitin ligase and deubiquitinase activity[85]. Additionally, they had the TF *MAFF* in common.

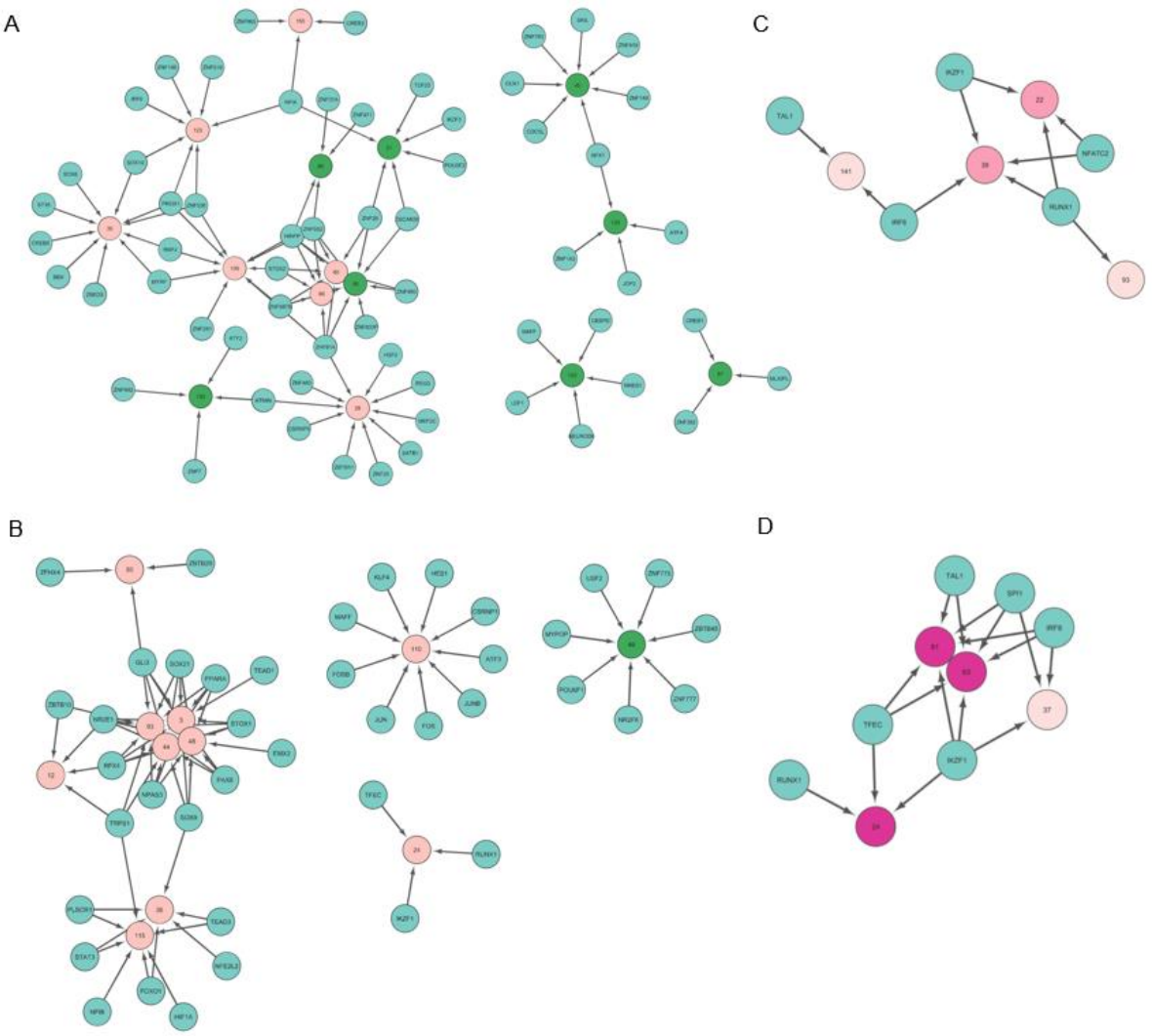


Figure 4. Networks of AD and MDD share regulators and targets. A) Modules with significant difference in expression between Alzheimer's (AD) patients and control individuals, and their regulators. B) Modules with significant difference in expression between depressive patients (MDD) and control individuals, and their regulators. C) Overview of the modules regulated by the TFs IKZF1, IRF8, NFATC2, RUNX1, SPI1, TAL1 and TFEC in the AD network. D) Overview of the modules regulated by the TFs IKZF1, IRF8, NFATC2, RUNX1, SPI1, TAL1 and TFEC in the MDD network. TFs are in blue, in A-B, the upregulated modules are in green, downregulated in pink. In C-D, the colors are according to the rank of the modules. The darker the pink, the more difference between diseased patients and normal controls.

In the next paragraphs, modules will be described with a significant gene expression difference between diseased patients and healthy controls, and that contain pathways specific to the two disorders. The networks of these modules and their TFs can be found in Figure 4a and 4b, and the Module Viewer visualizations of the modules in the supplemental figures. According to the functional enrichment analysis, module 26 (Supplemental Figure S9) of AD was involved in cholesterol biosynthesis and intracellular protein transport, and module 60 (Supplemental Figure S10) of the AD network was involved in mitochondrial pathways, such as mitochondrial protein import and the citric acid cycle, and metabolism of proteins. There were seven genes

of module 60 enriched in the 'Alzheimer's disease' term of the KEGG database (*PSMB7, PSMC6, MAPT, CYC1, COX6A1, GAPDH, RTN4*). PSMB7 and PSMC6 are part of the proteasome, and *MAPT* encodes the tau protein. CYC1 is part of the Cytochrome C family, COX6A1 is involved in the mitochondrial respiratory chain, while GADPH is involved in glycolysis[68]. Lastly, RTN4 is a axonal outgrowth inhibitor[86]. The TFs of module 26 were *ATMIN, CSRNP3, HSF2, MEF2C, PEG3, SATB1, ZBTB11, ZNF25, ZNF384* and *ZNF814*. The TFs of module 60 were *HINFP, STOX2, ZNF26, ZNF480, ZNF552, ZNF587B* and *ZNF814*. Module 35 from the AD network (Supplemental Figure S11) had terms related to membrane lipid biosynthesis, lipid metabolism, actin reorganization, cell movement and the term 'Oligodendrocyte specification and differentiation, leading to myelin components for CNS', which indicated that this module contained genes involved in myelination by oligodendrocytes. The expression of the AD patients was lower compared to the controls (p-value 0.0483), indicating a dysfunction in myelination by oligodendrocytes in AD. The TFs of this module *were BBX, CREB5, MYRF, PROX1, RBPJ, SOX8, SOX10, ST18, ZBED3* and *ZNF536*. Module 66 (Supplemental Figure S12) contained genes with functions related to 'GTP binding', 'pathways of neurodegeneration' (*COX8A, TUBA1A, TUBB, CHMP2B, NDUFC1, RAC1, TUBB4A*, and *COX6B1*), 'Alzheimer disease' (*COX8A, TUBA1A, TUBB, NDUFC1, TUBB4A, COX6B1*) and metabolism. The genes in this module had an overall lower expression in AD patients compared to controls (p-value 0.0359). The TFs of this module were *HINFP, STOX2, ZNF552, ZNF587B* and *ZNF814*. Looking at the motif analysis for these AD-modules, the following TFs were retrieved that were as well inferred: for module 26 *MEFC2*; for module 35 *RBPJ* and *SOX10*; for module 60 *HINFP* and for module 66 *HINFP*.

Module 93 of the MDD network (Supplemental Figure S13) had terms related to oligodendrocytes, 'regulation of nervous system development', GABA receptor signaling, 'mBDNF and proBDNF regulation of GABA neurotransmission', and kinase activity. The TFs of this module are *HINFP*, *STOX2, ZNF26, ZNF480, ZNF552, ZNF587B* and *ZNF814*. It is seen that the balance between proBDNF and BDNF is disturbed in MDD and that the levels of GABA are decreased in MDD patients[8]. The expression of this module was indeed decreased in patients compared to controls (p-value 0.0087). Module 115 (Supplemental Figure S14) of the MDD network contains genes related to 'inflammatory response', neutrophil activation, cytokine response, and the complement system, but also to 'glucose homeostasis' and 'regulation of lipid storage'. MDD patients have a lower expression (p-value 0.0370). *FOXO1, HIF1A, NFIB, PLSCR1, TEAD3, TRPS1* and *STAT3* are the inferred TFs. Module 36 (Supplemental Figure S15) had terms related to cytokine response and (negative regulation of) apoptosis. The expression is lower in patients compared to controls (p-value 0.0419). There are three genes (*CASP7* (caspase-7), *TP53*, *TNFRSF1A* (Tumor necrosis factor receptor

superfamily member 1A)) in the KEGG 'Alzheimer's disease' term. The TFs of this module are *FOXO1, NFE2L2, PLSCR1, TEAD3, SOX9* and *STAT3*. Module 44 (Supplemental Figure S16) from the MDD network is involved in 'regulation of macrophage activation', 'transport across blood-brain barrier', and 'aminoglycan biosynthetic process'. The expression was lower in MDD patients compared to controls (p-value 0.0079). The TFs are *GLI3, NPAS3, NR2E1, PAX6, PPARA, RFX4, SOX9, SOX21, STOX1* and *TRPS1*. MDD-module 49 (Supplemental Figure S17) contained genes related to 'regulation of neuron death', secretion, and protein transport. The expression was higher in MDD patients (p-value 0.0452). *MYPOP, NR2F6, POU6F1, USF2, ZBTB45, ZNF775* and *ZNF777* are the TFs of this module. For MDD-modules 36 and 115, only *STAT3* was retrieved as having enriched target genes. For the motif analysis, the promoter regions of the genes of module 36 had motifs enriched for *FOXO1*, *STAT*, and *SOX9*; for module 115, *FOXO1*; for module 93, *GLI3* and *PAX6*.

Additionally, GWAS summary statistics from the GWAS Catalog (Methods) were downloaded to see whether SNPs associated with AD or MDD were present intragenic or upstream (max 3kb) of genes in the prioritized modules. A full overview of these SNPs can be seen in Supplemental Table 9. Indeed, intragenic SNPs associated with AD were found in module genes *ABI3, ADAMTS4, BLNK, C12orf65, CD33, CTSB, DISC1, DLG2, ENOPH1, FRMD5, HAVCR2, HLA-DQA1, HS3ST5, INPP5D, ITGAX, LAMP1, MAPT, MS4A6A, MSR1, MTMR2, NKAIN2, PDE4B, PDE8A, PHF21B, PLCG2, RAC1, SESTD1, SORL1, TBXAS1, TNXB*, and *WDFY4*. SNPs upstream of *CD33* and *MS4A6A* were found to be associated with AD. For MDD, intragenic SNPs were found in *ACSS3, FAT1, GPC5, MCC, NTRK2, OSMR, PTAR1*, and *TLR4*, next to a SNP upstream of *ANXA1*. However, the prioritized modules were not overrepresented for genes associated with SNPs of the neuroinflammatory disorder (Methods). There were three modules from the AD network significantly (FDR ≤ 0.05) enriched for genes associated with SNPs: modules 2, 41 and 100. Module 2 contained genes involved in cell migration, 'blood vessel morphogenesis', cell differentiation, 'transforming growth factor beta receptor signaling pathway', 'complement and coagulation cascades' and related terms. Module 41 contained genes related to voltage-gated potassium channel activity. Lastly, in module 100 the genes were involved in 'potassium ion transport' and cation channel activity, cell-cell adhesion, synaptic transmission and some related terms. However, none of these modules showed a difference in gene expression between patients and controls. For MDD, there were no modules enriched for genes associated with SNPs.

Overall, for both AD and MDD, dysregulation of immune functions was found, but with only a partial overlap in pathways. Additionally, in AD, we discovered mitochondrial, proteasome and oligodendrocyte myelination dysfunctions. On the other hand, in MDD, dysfunctions in GABA

neurotransmission, oligodendrocyte function, transport across the blood-brain barrier and glucose homeostasis were detected.

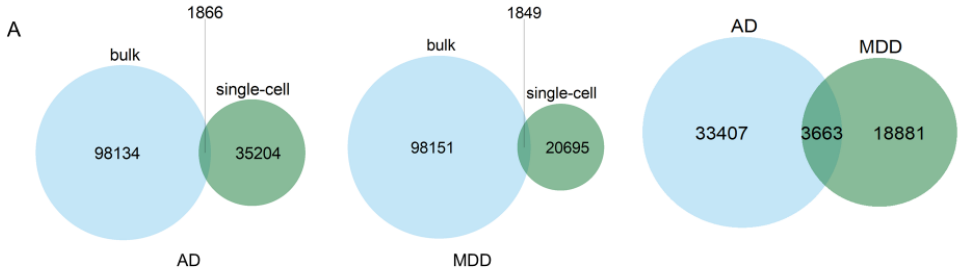
Further characterization with single-cell analysis implicates microglia, astrocytes and oligodendrocytes in Alzheimer's disease and depression

To identify in which cell types the pathways observed in the population-level analysis are manifested, we turned to single-cell network inference with scRNA-seq data. Single-nucleus RNA-seq datasets were as well retrieved from the prefrontal cortex. For depression, the dataset GSE144136 was found[54], where all samples were male. For Alzheimer's, the dataset GSE174367/syn22130832 was utilized[24] (Supplemental Table 1). The raw counts were preprocessed with the *Seurat* package in R (Methods). We used SCENIC to infer single-cell GRNs. SCENIC[56,87] infers regulons, together with the activity scores of the different regulons in specific cell types. A regulon signifies a TF together with its target genes. Firstly, GRNBoost is used to infer GRNs, which are then further refined to direct target genes by using TF motif data, from RcisTarget, reducing the number of false positives. The single-cell networks of MDD (22 544 edges, 4814 nodes) had fewer edges and nodes than the single-cell networks of AD (37 070 edges, 6360 nodes). There were 359 and 265 regulating TFs inferred in the AD and MDD networks, respectively. An overview of the number of nodes, target genes and TFs of the single-cell networks can be seen in Supplemental Table 9. The networks were visualized with the igraph package in Supplemental Figure 2. The single-cell MDD networks consisted of two connected components. Network characteristics (Supplemental Table 10) were similar to the bulk network characteristics, except for the (directed) node betweenness, which was lower in the single-cell networks.

The single-cell networks were compared to the population-level consensus networks. Firstly, the edges were compared between the networks of the same disease with Venn diagrams and the Jaccard index. The two disorders have a similar number of edges in common (Figure 5A). The Jaccard distance of the population-level and single-cell networks was 0.9862 and 0.9847 for AD and MDD, respectively, indicating that there are large differences between the inferred edges of the population-level networks and the single-cell networks. Furthermore, the number of TFs shared between the bulk and single-cell networks for the same disease was 258 TFs in the AD networks, and 238 TFs in the MDD networks. The single-cell networks were as well compared to each other through the number of overlapping edges (Figure 5a) and the Jaccard index. There are 3663 edges in common, which is about ten percent of the total edges of the networks. The Jaccard distance was 0.9345. Thus, the Jaccard distance between the two single-cell networks was smaller than the Jaccard distances between the population-level and single-cell networks of the same disorder. This could be explained by the fact that the single-cell networks were retrieved by a different method than the consensus population-level

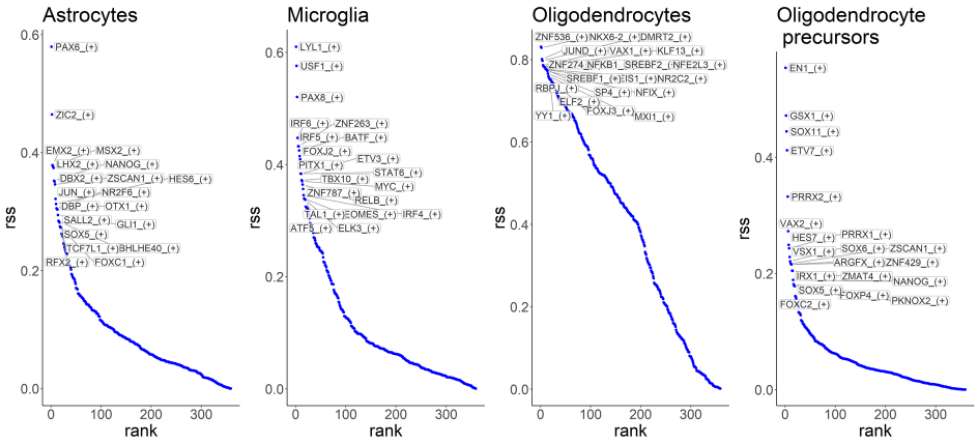
networks. Moreover, the number of regulators and edges was substantially less in the single-cell networks compared to the population-level networks. SCENIC makes use of RcisTarget, which cannot predict regulons for a TF with an unknown motif and which identifies direct binding of the TFs to its target genes. In comparison, the Jaccard distance of the population-level networks of AD and MDD was 0.9446.

A

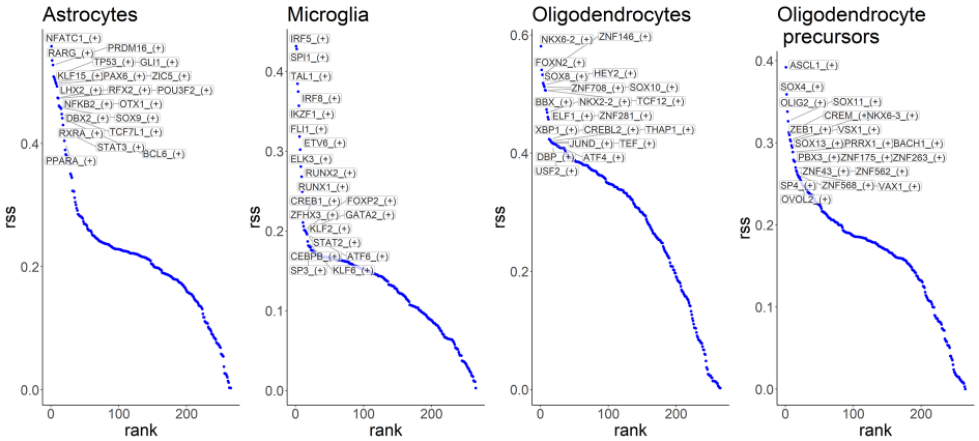

1866
bulk
single-cell
98134
35204
AD
1849
bulk
single-cell
98151
20695
MDD
AD
MDD
33407
3663
18881

1

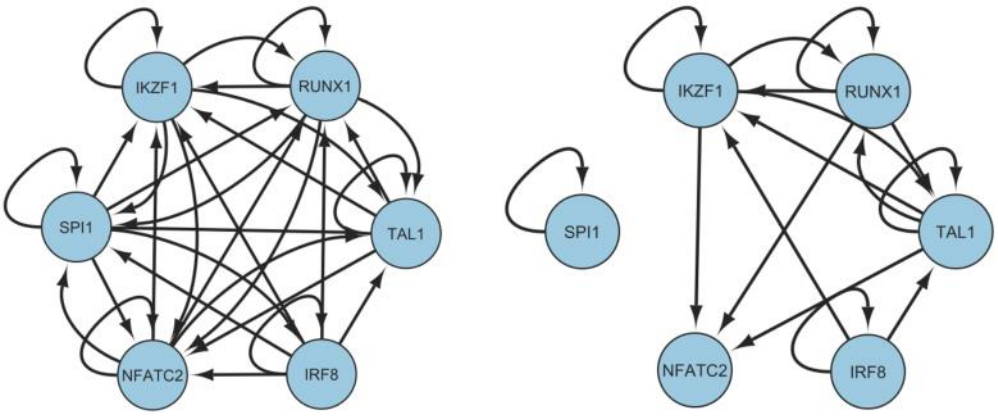
B


Astrocytes
Microglia
Oligodendrocytes
Oligodendrocyte precursors
rss
rank

C


Astrocytes
Microglia
Oligodendrocytes
Oligodendrocyte precursors
rss
rank

D


IKZF1
RUNX1
SPI1
TAL1
NFATC2
IRF8
AD
SPI1
IKZF1
RUNX1
TAL1
NFATC2
IRF8
MDD

3

Figure 5. Further characterization with single-cell analysis implicates microglia, astrocytes and oligodendrocytes in Alzheimer's disease and depression. A) Venn diagrams for the comparison between the edges of the consensus bulk networks and the single-cell networks of Alzheimer's disease (AD) and major depressive disorder (MDD). Right: Venn diagram of the edges of the single-cell networks of Alzheimer's disease (AD) and major depressive disorder (MDD). There are 3663 edges in common. B) Regulon specificity score (RSS) plots from the single-cell Alzheimer's networks (AD) for each glial cell type. The top 20 regulons are highlighted on the plots. The RSS is depicted on the y-axis, the rank of the regulons is depicted on the x-axis. C) RSS plots from the single-cell depressive networks (MDD) for each glial cell type. The top 20 regulons are highlighted on the plots. The RSS is depicted on the y-axis, the rank of the regulons is depicted on the x-axis. D) Microglial core gene regulatory networks with six TFs. The single-cell Alzheimer's disease (AD) network is depicted on the left, and the single-cell depression (MDD) network is depicted on the right.

To further investigate which regulons are active in which cell types, we looked at their activity and specificity. Table 1 represents an overview of the highly active regulons in each cell type (i.e. regulatory programs highly active in that cell type), for both the AD and MDD networks. Heatmaps with the activity scores of the regulons for each cell type can be found in the supplemental material. Of the regulons inferred in both networks, a large part that is highly active in a certain cell type of one of the two networks, is also active in the same cell type in the other network (red in Table 1).

Table 1. Overview of the highly active regulons (depicted as the TF) in different cell types of the brain in single cell networks of Alzheimer's disease and major depressive disorder. Transcription factors in red are also active in the other network, in the same cell type.

Cell type	*Alzheimer's disease*	*Major depressive disorder*
Excitatory neurons	AHR, HLF, MEF2C, ZEB1, ZMAT4	FOXP1, NFAT5, PBX1, RFX3, TEAD4, ZMAT4, ZNF282, ZNF699
Inhibitory neurons	HLF, MEF2C, PKNOX2, ZEB1, ZMAT4	DLX1, DLX2, DLX5, FOXN3, HIVEP3, LHX6, MAFB, NR2F2, TCF4, ZMAT4, ZNF282
Microglia	ELF1, ELK3, ETS2, ETV6, FLI1, IKZF1, IRF8, MAF, NFATC2, RUNX1, SPI1, STAT6	FOXN3, IRF8, RUNX1
Oligodendrocytes	MXI1, NFIX, SREBF2, ZNF536	FOXN2, SOX10
Oligodendrocyte precursor cells	PRRX1, PRRX2, SOX6, VSX1, ZEB1, ZNF227	FOXN3, PBX3, PRRX1, SOX13, SOX4, ZEB1
Astrocytes	FOXO1, RFX2, SOX5, TCF7L1, TCF7L2	FOXO1, PAX6, RARG, RXRA, SOX2, SOX9, TCF7L2

In addition, the regulon specificity score was calculated for the four glial cell types (astrocytes, microglia, oligodendrocytes, and oligodendrocyte-precursor cells). This score indicates to what extent regulons are specific to a certain cell type (the higher the score, the more specific)[87].

The top 20 regulons for these cell types are indicated in Figure 5B and 5C. There were three top regulons shared between the astrocytes and oligodendrocyte precursor cells in the single-cell AD networks: *NANOG*, *SOX5* and *ZSCAN1*. The regulons of *DBX2*, *GLI1*, *LHX2*, *OTX1*, *PAX6*, *RFX2,* and *TCF7L1* were in the top 20 regulons of the astrocytes for both single-cell networks. The regulons of *ELK3*, *IRF5* and *TAL1* were most active in microglia in both single-cell networks. *NKX6-2* and *JUND* were shared top regulons of the oligodendrocytes. Lastly, the regulons of *SOX11, PRRX1*, and *VSX1* were shared top regulons in the oligodendrocyte precursor cells.

We also investigated in which cell types the TFs of the prioritized modules of the population-level network inference were active. The regulons of *IKZF1, IRF8, NFATC2, RUNX1, SPI1,* and *TAL1* were highly active in the microglia in the single-cell AD networks, and the regulons of *IKZF1, IRF8, RUNX1, SPI1,* and *TAL1* were most active in the microglia in the single-cell MDD networks, which confirms the hypothesis that the immune-related bulk consensus modules (module 22 and 39 from AD and 24 from MDD) represented regulatory programs active in microglia. The regulon of *TFEC* is not present in both single-cell networks, and the regulon of *NFATC2* is not present in the single-cell MDD networks. The overlap between the module genes and target genes in the regulons is high for the AD regulons (± 69% of module genes found in regulons), but lower for the MDD regulons (± 27% of module genes found in regulons). This could be explained by the fact that there were fewer genes in the regulons of the single-cell MDD networks.

As the regulons of *IKZF1*, *IRF8*, *NFATC2*, and *RUNX1* were highly active in microglia in the single-cell data of AD and MDD, and there was a difference in the regulatory programs in Alzheimer's and depression in the population-level networks, these were further investigated. The TFs SPI1 and TAL1, which were not inferred but came back in the motif enrichment, were also further inspected in the single-cell networks. A core GRN could be retrieved with only these six TFs, highly regulating one another in both neuroinflammatory disorders (except for *SPI1* in MDD) with more interconnecting edges in the AD than in the MDD network (Figure 5D). IRF8 only regulated itself and the other TFs, but was not regulated by any of the other TFs, except for SPI1 in the AD network. In addition, in the MDD network, NFATC2 was only regulated by other TFs, but did not regulate any target genes, as this gene was not inferred as a regulating TF in the single-cell MDD networks. As can be seen in the figure, the TFs also regulated themselves, which are direct interactions supported by TF binding information from SCENIC. Thus, these TFs have the binding sites of their protein product in the upstream regulatory region of their genes. There were as well some feed-forward loops present.

For the other modules, there was only a small overlap in genes between the regulons of the TFs and the modules steered by these TFs. Thus, to see whether the regulons and modules converge on the pathway level, functional enrichment analysis was executed for these regulons. Module 35 from the AD network, with terms involving lipids and oligodendrocytes, had TFs (*BBX, SOX10, SOX8, RBPJ, ZNF536*) of which the regulons were most active in oligodendrocytes and oligodendrocyte precursor cells. The AD regulons of *SOX8* and *SOX10* had genes enriched for 'oligodendrocyte differentiation', myelination, 'amoeboidal-type cell migration', and 'oligodendrocyte specification and differentiation, leading to myelin components for CNS'. Module 153 from the AD network, with terms involving apoptosis and inflammatory response, was steered by TFs of which the regulons were most active in microglia (*CEBPD*), endothelial cells (*CEBPD* AD, *LEF1* MDD), and excitatory neurons (*CEBPD*). The AD regulon of CEBPD contained genes involved in cytokine response, apoptosis, regulation of cell cycle, T_H cell differentiation, and 'cellular senescence', amongst others. The AD regulon of *LEF1* also contained genes involved in cytokine response.

Module 36 of the MDD network, containing terms related to cytokine response and apoptosis, had TFs most active in astrocytes (*FOXO1, SOX9, STAT3* MDD) and microglia (*STAT3* AD). The MDD regulon of *FOXO1* had genes enriched in 'positive regulation of apoptotic process', regulation of cell differentiation, and 'cellular response to cytokine stimulus', amongst others. The AD regulon of *STAT3* contained terms related to apoptosis and cytokine response as well. Module 44, containing terms related to transport across the blood-brain barrier, had TFs of which the regulons were most active in astrocytes (*PAX6, SOX9, SOX21, PPARA*). The *PAX6* regulon of the single-cell MDD network contained genes related to the 'maintenance of blood-brain barrier', and macrophage chemotaxis, amongst others. The *SOX9* regulon was also enriched for the term 'transport across blood-brain barrier'. In modules 36 and 44, the expression was lower in the patients. Astrocyte dysfunction has been linked with depression previously, and with reductions in the number and density of astrocytes as well[10]. Module 93 from the MDD network had some TFs (*PAX6, PPARA, SOX21*) of which the regulons were most active in astrocytes in the single-cell networks, while the module had functional enrichment terms related to oligodendrocytes, 'regulation of nervous system development', GABA receptor signaling, and BDNF. The MDD regulons of *PAX6* and *SOX21* both contained genes involved in nervous system development, while the PPARA regulon had the 'regulation of lipid transport' term. MDD-module 110, involved in cytokine signaling and unfolded protein response, contained TFs most active in astrocytes in the AD single-cell network (*FOS, JUN*), and endothelial cells in the MDD single-cell networks (*FOS, JUNB, KLF4*), as well as inhibitory neurons (*ATF3*), and mixed cell types (*JUN, JUNB*). The MDD regulons of *FOS*, *JUN*, and *JUNB* contained genes involved in 'cellular response to cytokine stimulus', 'response to

unfolded protein', apoptosis, and 'cellular responses to stress', while the *KLF4* regulon was involved in cytokine response and angiogenesis. Lastly, module 115 from the MDD network, containing terms related to inflammatory response and 'glucose homeostasis' and 'regulation of lipid storage', had TFs of which the regulons were most active in astrocytes (*FOXO1, STAT3* MDD), microglia (*STAT3* AD), and inhibitory neurons (*NFIB*). The MDD *FOXO1* and AD *STAT3* regulons contained genes related to cytokine response, amongst others. The MDD *NFIB* regulon contained terms related to brain development and chemotaxis. In conclusion, by inferring single-cell GRNs, we were able to implicate microglia in both disorders, oligodendrocytes in AD and astrocytes in MDD.

Discussion

Our purpose was to unravel common and distinct regulatory programs of neuroinflammation and other processes in AD and MDD. Finding a common pathway that could be targeted in different disorders, would possibly shine a new light on the treatment of neuroinflammatory disorders. The field of GRN mapping in neuroinflammatory disorders has grown largely over the last few years. However, limitations remain in only using one network inference method[17], using only DEGs[99], or using bulk data to infer the networks. In the first part of this work, population-level GRNs were inferred with different methods (GENIE3, CLR and Lemon-Tree). There was only a small overlap in the results retrieved from the different methods, highlighting the different underlying assumptions and algorithms. Next, consensus networks were made for each disorder, from which modules were retrieved. Given the small overlap in the different network inference methods and the significant overlap with the ground truth for each network inferred by a specific method, we also observed here that building consensus networks across algorithms provide a more robust view of the true underlying network[13,17]. The modules were functionally analyzed with GO, KEGG, Reactome, and WikiPathways. We prioritized several modules based on a significant difference in gene expression between patients and controls, a functional enrichment related to immune mechanisms, and the number of genes. Three of these modules contained overlapping immune-related terms and TFs (AD-modules 22 and 39, and MDD-module 24). The gene expression of module 22 was increased in AD patients compared to controls, while it was decreased in MDD patients (module 24) compared to controls, indicating a distinct disruption of these pathways in the two disorders. This can also be explained by the small overlap in genes (± 12% overlap) between these modules.

Further, networks were inferred with the single-cell GRN inference method SCENIC[56] on scRNA-seq data. As such, cell-type-specific regulons were retrieved for both AD and MDD. The main drawback of bulk RNA-seq is the fact that all reads are averaged over the cells, and that the results are influenced by the number of cells of each cell type present in the samples. Additionally, SCENIC uses TF-binding information to refine direct regulatory interactions. Most

of the TFs inferred in the population-level immune-related modules had regulons that were highly active in microglia, in both AD and depression, and it was seen that these TFs regulate each other. Thus, a large part of the genes in the immune-related bulk modules could be coming from reads in microglia. It is already known that microglia play a pivotal role in AD. As mentioned above, the modules of the AD network had a higher gene expression in patients compared to controls, indicating activation of microglia in AD. In a mouse model of AD, NFATC2 was found to be implicated in the activation of microglia[69]. The researchers have crossed AβPP/PS1 mice with NFATc2$^{-/-}$ mice, which resulted in mice with diminished cytokine levels, reduced microgliosis and astrogliosis, but with no effect on plaque load, compared to AβPP/PS1 mice. In addition, IRF8 contributes to microglial activation by regulating microglial immune responses and chemotaxis[88]. The expression of IRF8 was found to be increased in the brains and microglia of an AD mouse model[88]. Moreover, Aβ was discovered to promote the expression of IRF8, while overexpression of IRF8 aggravated microglial activation.

On the other hand, in module 24 of the MDD network, there were some terms with negative regulation, which makes it hard to determine which cells and/or pathways are decreased or increased here. Hence, this needs further research. Phagocytosis was recurring in both modules. Despite TYROBP not being a part of modules 22 and 24 in the population-level networks, it was regulated by the TFs IKZF1, IRF8, RUNX1, and SPI1 in the AD single-cell network and by IKZF1 and IRF8 in the single-cell MDD network. TYROBP plays an important role in signal transduction in microglia and was found to be significantly upregulated in the brain of patients with AD[89]. It is important for the phagocytic activity of microglia, together with TREM2, which is a well-known risk gene for AD. In addition, TYROBP can also suppress cytokine production and secretion. Thus, TYROBP deficiency could lead to aberrant phagocytosis and increased cytokine release, which the AD brain tries to counteract by increasing the expression levels of TYROBP.

In addition to the immune-related processes, other disturbed pathways were found in the AD network, such as mitochondrial pathways and the proteasome system. Additionally, myelination by oligodendrocytes seemed to be dysregulated in AD. Myelin damage has been implicated before in AD, and it is not known whether the damage and death of oligodendrocytes is a result of the Aβ plaques and neurofibrillary tangles, or that it is more a primary event[90]. Some studies found evidence of myelin damage before the occurrence of plaques and tangles in animal models, which might indicate that demyelination is a driving force in AD[90]. Moreover, it was found that myelin damage happens mainly in the proximity of Aβ plaques. On the other hand, GABA (γ-amino butyric acid) signaling and BDNF (brain-derived neurotrophic factor) were found to be involved in MDD and have been implicated before in the disease, next to dysregulation of nervous system development. Lastly, cytokine response and astrocytes were

implicated, and enrichment for genes involved in apoptosis was found, indicating there is higher apoptosis of astrocytes in depression. A reduction in the number of astrocytes can be caused by chronic stress[10]. Several studies support that there are significant reductions in the number and density of astrocytes in several brain regions.

Pearl and colleagues have reconstructed a GRN for the human brain by integrating brain-specific DNase footprinting and TF-gene co-expression[21]. Then they retrieved the DEGs from RNA-seq and microarray data from schizophrenia, bipolar disorder, MDD, AD, and autism patients and controls and identified TFs whose target genes were enriched in these DEGs. Interestingly, of the top 100 TFs of the AD and MDD consensus networks, MEF2C and TFEB were two common TFs, which they found as well as hub TFs in AD[21]. MEFC2 was the second top TF in our AD network, indicating this TF had a broad function in the network. MEF2C plays an essential role in hippocampal-dependent learning and memory, and is crucial for normal neuronal development and electrical activity[91]. It is as well involved in some immunological functions[92], and has been linked to AD before[93]. The regulon of MEF2C was highly active in both the excitatory and inhibitory neurons in the single-cell network of AD. In addition, in this paper from Pearl et al., microglia-enriched TF networks were upregulated in AD[21]. In another study, ZEB1 and MAFB were identified as candidate TFs of AD-specific cis-regulatory elements[94]. Here, we found ZEB1 to be highly active in excitatory neurons, inhibitory neurons and oligodendrocyte-precursor cells in the single-cell networks. They also found MEF2C to be highly involved in regulatory interactions.

Interestingly, a recent study investigated the shared genetic etiology between AD and MDD using GWAS datasets[95]. The researchers found a moderate level of polygenic overlap between the two disorders. Noteworthy, they found a stronger overlap if the SNPs in the APOE region were excluded from the analysis, because of a very strong association between SNPs in this region and AD, with a change in p-value of up to 227 orders of magnitude, compared to the other SNPs[95]. An enrichment of SNPs was found on chromosome eleven, which were linked to expression regulation in myeloid cells, such as microglia. The associated SNPs were mapped to 40 most proximal genes when studying enrichment for AD, given MDD GWAS association, of which nine genes (*BIN1, CELF1, CR1, FERMT2, MS4A6A, PICALM, PTK2B, SORL1*, and *SPI1*) were already known as AD risk genes[95]. These nine genes are involved in two major pathways; immune response (*CELF1, CR1, MS4A6A,* and *SPI1*) and regulation of endocytosis (*BIN1, FERMT2, PICALM, PTK2B*, and *SORL1*). This indicates that indeed MDD and AD both have an immunological compound, and that there is some genetic overlap in immune aberrations. *SPI1* was part of module 24 from the consensus MDD network, its motifs were enriched in the genes of the three immune-related modules, and its regulon was highly active in microglia in the single-cell AD network. This gene was as well prioritized in the single-

cell analysis of AD in the paper of the original data for AD[24]. In another paper, they also found *SPI1* to be implicated in AD, where it was regulated by *NFKB1* and its expression strongly correlated with AD/Braak Progression in the hippocampus[96]. Similarly, in a new GWAS of AD, the most significantly enriched gene sets were related to tau and amyloid, lipids, endocytosis, and immunity[97]. The endocytosis pathway might be related to microglial phagocytosis, which is impaired in AD[97]. In addition, they found several genes implicated in the TNF-α signaling pathway.

To predict cell-type disease genes and GRNs of neurodegenerative and neuropsychiatric diseases, scGRNom (single-cell Gene Regulatory Network prediction from multi-omics) was developed [98]. They used chromatin interaction and TF binding site data to infer a reference GRN linking TFs, enhancers and promotors of the target genes. The final GRN is predicted by TF-target gene expression relationships inferred through elastic net regression[98]. To determine the cell-type-specific disease genes and regulatory elements, SNPs associated with the disease are integrated with the GRNs. The researchers applied the pipeline to SCZ and AD and found that enhancers in cell-type-specific GRNs are enriched with GWAS SNPs. They used healthy brain multi-omics data with disease-specific SNPs. For AD, they also found cellular response to stress and immune response activation in microglia[98].

There are some limitations of the used methods. Firstly, only TFs were used as regulators, without taking into account chromatin modifiers, co-factors, or regulatory RNA molecules such as lncRNA and miRNA. In addition, only transcriptomics data was used to infer the population-level networks, without additional information, resulting in possibly indirect interactions. However, little multi-omics data is available from the same individuals[99]. SCENIC uses additional data from RcisTarget, resulting in the inference of direct interactions. Further, k-medoids clustering and Lemon-Tree can only assign genes to one module. In reality, however, genes mostly belong to different pathways. Next, by using publicly available data, there is less control over the samples that are used. For instance, it is not always known if the patients took medication. Additionally, batch effects had to be accounted for regarding the depressive samples. Moreover, the ground truth regulatory interactions were not condition-specific. Further, overall, research is still focused on individuals of European descent, resulting in a limited genetic diversity, and the findings cannot be extrapolated to the whole human population. Lastly, post-mortem brain samples are prone to the degradation of RNA molecules[100], and these samples only represent one snapshot of the patient's lifetime[100].

In conclusion, neuroinflammation and microglia activation were found to be dysregulated in both AD and MDD. In addition, several pathways and cell types were found to be specifically dysregulated in one of the disorders. Regarding the activation of microglia, different TFs were

found that were implicated in both AD and MDD. Of particular interest were the TFs *IKZF1, IRF8, NFATC2, RUNX1, SPI1*, and *TAL1*. All of these TFs have been implicated with AD, while only *TAL1* has been implicated with MDD before. In addition, *INPP5D, TYROBP* and *TREM2* have been linked with the activation of microglia in AD, while *TYROBP* and *TREM2* have been linked with MDD. Moreover, *TFEB* and *MEF2C* were two hub TFs in AD and MDD that were also found to be hub TFs in AD before. All these genes and pathways need to be further investigated, especially in MDD, and can be prioritized when developing novel therapies.

Author contributions

Contributions according to the CRediT system (https://www.ucl.ac.uk/library/research-support/open-access/credit-taxonomy, HP = Hanne Puype, JD = Joke Deschildre, VV = Vanessa Vermeirssen): Conceptualization: VV, Data curation: HP, Formal analysis: HP, Investigation: HP, Methodology: HP, VV, Project administration: VV, Software: HP, JD, Supervision: VV, Validation: HP, Visualization: HP, Writing—original draft: HP, VV, Writing—rewriting and editing HP, JD, VV. All authors read and approved the final manuscript.

Declaration of interest

The authors declare no competing interests.

References

1. Pape, K., Tamouza, R., Leboyer, M., and Zipp, F. (2019). Immunoneuropsychiatry — novel perspectives on brain disorders. Nature Reviews Neurology *15*, 317–328. 10.1038/s41582-019-0174-4.

2. Schwartz, M., and Deczkowska, A. (2016). Neurological Disease as a Failure of Brain-Immune Crosstalk: The Multiple Faces of Neuroinflammation. Trends Immunol *37*, 668–679. 10.1016/j.it.2016.08.001.

3. Chew, G., and Petretto, E. (2019). Transcriptional Networks of Microglia in Alzheimer's Disease and Insights into Pathogenesis. Genes (Basel) *10*. 10.3390/genes10100798.

4. Soria Lopez, J.A., González, H.M., and Léger, G.C. (2019). Alzheimer's disease. Handb Clin Neurol *167*, 231–255. 10.1016/b978-0-12-804766-8.00013-3.

5. Lane, C.A., Hardy, J., and Schott, J.M. (2018). Alzheimer's disease. Eur J Neurol *25*, 59–70. 10.1111/ene.13439.

6. Hansen, D.V., Hanson, J.E., and Sheng, M. (2018). Microglia in Alzheimer's disease. J Cell Biol *217*, 459–472. 10.1083/jcb.201709069.

7. Liddelow, S.A., Guttenplan, K.A., Clarke, L.E., Bennett, F.C., Bohlen, C.J., Schirmer, L., Bennett, M.L., Münch, A.E., Chung, W.-S., Peterson, T.C., et al. (2017). Neurotoxic reactive astrocytes are induced by activated microglia. Nature *541*, 481–487. 10.1038/nature21029.

8. Pitsillou, E., Bresnehan, S.M., Kagarakis, E.A., Wijoyo, S.J., Liang, J., Hung, A., and Karagiannis, T.C. (2020). The cellular and molecular basis of major depressive disorder: towards a unified model for understanding clinical depression. Mol Biol Rep *47*, 753–770. 10.1007/s11033-019-05129-3.

9. Boku, S., Nakagawa, S., Toda, H., and Hishimoto, A. (2018). Neural basis of major depressive disorder: Beyond monoamine hypothesis. Psychiatry Clin Neurosci *72*, 3–12. 10.1111/pcn.12604.

10. Wang, Q., Jie, W., Liu, J.H., Yang, J.M., and Gao, T.M. (2017). An astroglial basis of major depressive disorder? An overview. Glia *65*, 1227–1250. 10.1002/glia.23143.

11. Liu, C.S., Adibfar, A., Herrmann, N., Gallagher, D., and Lanctôt, K.L. (2017). Evidence for Inflammation-Associated Depression. In Inflammation-Associated Depression: Evidence, Mechanisms and Implications Current Topics in Behavioral Neurosciences., R. Dantzer and L. Capuron, eds. (Springer International Publishing), pp. 3–30. 10.1007/7854_2016_2.

12. Drevets, W.C., Wittenberg, G.M., Bullmore, E.T., and Manji, H.K. (2022). Immune targets for therapeutic development in depression: towards precision medicine. Nat Rev Drug Discov, 1–21. 10.1038/s41573-021-00368-1.

13. Vermeirssen, V., De Clercq, I., Van Parys, T., Van Breusegem, F., and Van de Peer, Y. (2014). Arabidopsis Ensemble Reverse-Engineered Gene Regulatory Network Discloses Interconnected Transcription Factors in Oxidative Stress. The Plant Cell *26*, 4656. 10.1105/tpc.114.131417.

14. Loers, J.U., and Vermeirssen, V. (2022). SUBATOMIC: a SUbgraph BAsed mulTi-OMIcs clustering framework to analyze integrated multi-edge networks. BMC Bioinformatics *23*, 363. 10.1186/s12859-022-04908-3.

15. Saint-Antoine, M.M., and Singh, A. (2020). Network inference in systems biology: recent developments, challenges, and applications. Current Opinion in Biotechnology *63*, 89–98. https://doi.org/10.1016/j.copbio.2019.12.002.

16. Banf, M., and Rhee, S.Y. (2017). Computational inference of gene regulatory networks: Approaches, limitations and opportunities. Biochimica et Biophysica Acta - Gene Regulatory Mechanisms *1860*, 41–52. 10.1016/j.bbagrm.2016.09.003.

17. Marbach, D., Costello, J.C., Küffner, R., Vega, N.M., Prill, R.J., Camacho, D.M., Allison, K.R., Kellis, M., Collins, J.J., and Stolovitzky, G. (2012). Wisdom of crowds for robust gene network inference. Nat Methods *9*, 796–804. 10.1038/nmeth.2016.

18. Gandal, M.J., Haney, J.R., Parikshak, N.N., Leppa, V., Ramaswami, G., Hartl, C., Schork, A.J., Appadurai, V., Buil, A., Werge, T.M., et al. (2018). Shared molecular neuropathology across major psychiatric disorders parallels polygenic overlap. Science *359*, 693–697. 10.1126/science.aad6469.

19. Sadeghi, I., Gispert, J.D., Palumbo, E., Muñoz-Aguirre, M., Wucher, V., D'Argenio, V., Santpere, G., Navarro, A., Guigo, R., and Vilor-Tejedor, N. (2021). Brain transcriptomic profiling reveals common alterations across neurodegenerative and psychiatric disorders. bioRxiv, 2021.08.16.456345. 10.1101/2021.08.16.456345.

20. Godini, R., Fallahi, H., and Ebrahimie, E. (2019). A comparative system-level analysis of the neurodegenerative diseases. Journal of Cellular Physiology *234*, 5215–5229. 10.1002/jcp.27330.

21. Pearl, J.R., Colantuoni, C., Bergey, D.E., Funk, C.C., Shannon, P., Basu, B., Casella, A.M., Oshone, R.T., Hood, L., Price, N.D., et al. (2019). Genome-Scale Transcriptional Regulatory Network Models of Psychiatric and Neurodegenerative Disorders. Cell Syst *8*, 122-135.e7. 10.1016/j.cels.2019.01.002.

22. Sadeghi, I., Gispert, J.D., Palumbo, E., Muñoz-Aguirre, M., Wucher, V., D'Argenio, V., Santpere, G., Navarro, A., Guigo, R., and Vilor-Tejedor, N. (2022). Brain transcriptomic profiling reveals common alterations across neurodegenerative and psychiatric disorders. Computational and Structural Biotechnology Journal *20*, 4549–4561. 10.1016/j.csbj.2022.08.037.

23. Langfelder, P., and Horvath, S. (2007). Eigengene networks for studying the relationships between co-expression modules. BMC Syst Biol *1*, 54. 10.1186/1752-0509-1-54.

24. Morabito, S., Miyoshi, E., Michael, N., Shahin, S., Martini, A.C., Head, E., Silva, J., Leavy, K., Perez-Rosendahl, M., and Swarup, V. (2021). Single-nucleus chromatin accessibility and transcriptomic characterization of Alzheimer's disease. Nature genetics *53*, 1143–1155. 10.1038/s41588-021-00894-z.

25. Pantazatos, S.P., Huang, Y.-Y., Rosoklija, G.B., Dwork, A.J., Arango, V., and Mann, J.J. (2017). Whole-transcriptome brain expression and exon-usage profiling in major depression and suicide: evidence for altered glial, endothelial and ATPase activity. Molecular Psychiatry *22*, 760–773. 10.1038/mp.2016.130.

26. Ramaker, R.C., Bowling, K.M., Lasseigne, B.N., Hagenauer, M.H., Hardigan, A.A., Davis, N.S., Gertz, J., Cartagena, P.M., Walsh, D.M., Vawter, M.P., et al. (2017). Post-mortem molecular profiling of three psychiatric disorders. Genome Medicine *9*, 72. 10.1186/s13073-017-0458-5.

27. Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139–140. 10.1093/bioinformatics/btp616.

28. Robinson, M.D., and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biology *11*, R25. 10.1186/gb-2010-11-3-r25.

29. Lovering, R.C., Gaudet, P., Acencio, M.L., Ignatchenko, A., Jolma, A., Fornes, O., Kuiper, M., Kulakovskiy, I.V., Lægreid, A., Martin, M.J., et al. (2021). A GO catalogue of human DNA-binding transcription factors. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms, 194765. 10.1016/j.bbagrm.2021.194765.

30. Huynh-Thu, V.A., Irrthum, A., Wehenkel, L., and Geurts, P. (2010). Inferring Regulatory Networks from Expression Data Using Tree-Based Methods. PLoS One *5*, e12776. 10.1371/journal.pone.0012776.

31. Meyer, P.E., Lafitte, F., and Bontempi, G. (2008). minet: A R/Bioconductor Package for Inferring Large Transcriptional Networks Using Mutual Information. BMC Bioinformatics *9*, 461. 10.1186/1471-2105-9-461.

32. Bonnet, E., Calzone, L., and Michoel, T. (2015). Integrative Multi-omics Module Network Inference with Lemon-Tree. PLOS Computational Biology *11*, e1003983. 10.1371/journal.pcbi.1003983.

33. Schimek, M.G., Budinská, E., Kugler, K.G., Švendová, V., Ding, J., and Lin, S. (2015). TopKLists: a comprehensive R package for statistical inference, stochastic aggregation, and visualization of multiple omics ranked lists. Stat Appl Genet Mol Biol *14*, 311–316. 10.1515/sagmb-2014-0093.

34. Mächler, M., Rousseeuw, P., Struyf, A., Hubert, M., and Hornik, K. (2022). Cluster: Cluster Analysis Basics and Extensions.

35. Chen, H., and Boutros, P.C. (2011). VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. BMC Bioinformatics *12*, 35. 10.1186/1471-2105-12-35.

36. Csardi, G., and Nepusz, T. (2005). The Igraph Software Package for Complex Network Research. InterJournal *Complex Systems*, 1695.

37. Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res *13*, 2498–2504. 10.1101/gr.1239303.

38. Kuleshov, M.V., Jones, M.R., Rouillard, A.D., Fernandez, N.F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S.L., Jagodnik, K.M., Lachmann, A., et al. (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res *44*, W90-97. 10.1093/nar/gkw377.

39. Gene Ontology Consortium (2021). The Gene Ontology resource: enriching a GOld mine. Nucleic Acids Res *49*, D325–D334. 10.1093/nar/gkaa1113.

40. Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet *25*, 25–29. 10.1038/75556.

41. Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., Katayama, T., Kawashima, S., Okuda, S., Tokimatsu, T., et al. (2008). KEGG for linking genomes to life and the environment. Nucleic Acids Research *36*, D480–D484. 10.1093/nar/gkm882.

42. Jassal, B., Matthews, L., Viteri, G., Gong, C., Lorente, P., Fabregat, A., Sidiropoulos, K., Cook, J., Gillespie, M., Haw, R., et al. (2020). The reactome pathway knowledgebase. Nucleic Acids Res *48*, D498–D503. 10.1093/nar/gkz1031.

43. Martens, M., Ammar, A., Riutta, A., Waagmeester, A., Slenter, D.N., Hanspers, K., A. Miller, R., Digles, D., Lopes, E.N., Ehrhart, F., et al. (2021). WikiPathways: connecting communities. Nucleic Acids Research *49*, D613–D621. 10.1093/nar/gkaa1024.

44. Bailey, T.L., and Grant, C.E. (2021). SEA: Simple Enrichment Analysis of motifs. 2021.08.23.457422. 10.1101/2021.08.23.457422.

45. Thomas-Chollier, M., Herrmann, C., Defrance, M., Sand, O., Thieffry, D., and van Helden, J. (2012). RSAT peak-motifs: motif analysis in full-size ChIP-seq datasets. Nucleic Acids Research *40*, e31. 10.1093/nar/gkr1104.

46. Nguyen, N.T.T., Contreras-Moreira, B., Castro-Mondragon, J.A., Santana-Garcia, W., Ossio, R., Robles-Espinoza, C.D., Bahin, M., Collombet, S., Vincens, P., Thieffry, D., et al. (2018). RSAT 2018: regulatory sequence analysis tools 20th anniversary. Nucleic Acids Research *46*, W209–W214. 10.1093/nar/gky317.

47. Buniello, A., MacArthur, J.A.L., Cerezo, M., Harris, L.W., Hayhurst, J., Malangone, C., McMahon, A., Morales, J., Mountjoy, E., Sollis, E., et al. (2019). The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. Nucleic Acids Res *47*, D1005–D1012. 10.1093/nar/gky1120.

48. Liao, Y., Wang, J., Jaehnig, E.J., Shi, Z., and Zhang, B. (2019). WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. Nucleic Acids Res *47*, W199–W205. 10.1093/nar/gkz401.

49. Tantardini, M., Ieva, F., Tajoli, L., and Piccardi, C. (2019). Comparing methods for comparing networks. Scientific reports *9*, 17557. 10.1038/s41598-019-53708-y.

50. Yaveroğlu, Ö.N., Malod-Dognin, N., Davis, D., Levnajic, Z., Janjic, V., Karapandza, R., Stojmirovic, A., and Pržulj, N. (2014). Revealing the Hidden Language of Complex Networks. Scientific reports *4*, 4547. 10.1038/srep04547.

51. Türei, D., Korcsmáros, T., and Saez-Rodriguez, J. (2016). OmniPath: guidelines and gateway for literature-curated signaling pathway resources. Nat Methods *13*, 966–967. 10.1038/nmeth.4077.

52. Marbach, D., Lamparter, D., Quon, G., Kellis, M., Kutalik, Z., and Bergmann, S. (2016). Tissue-specific regulatory circuits reveal variable modular perturbations across complex diseases. Nat Methods *13*, 366–370. 10.1038/nmeth.3799.

53. Garcia-Alonso, L., Holland, C.H., Ibrahim, M.M., Turei, D., and Saez-Rodriguez, J. (2019). Benchmark and integration of resources for the estimation of human transcription factor activities. Genome Res. *29*, 1363–1375. 10.1101/gr.240663.118.

54. Nagy, C., Maitra, M., Tanti, A., Suderman, M., Théroux, J.F., Davoli, M.A., Perlman, K., Yerko, V., Wang, Y.C., Tripathy, S.J., et al. (2020). Single-nucleus transcriptomics of the prefrontal cortex in major depressive disorder implicates oligodendrocyte precursor cells and excitatory neurons. Nat Neurosci *23*, 771–781. 10.1038/s41593-020-0621-y.

55. Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. Cell *184*, 3573-3587.e29. 10.1016/j.cell.2021.04.048.

56. Aibar, S., González-Blas, C.B., Moerman, T., Huynh-Thu, V.A., Imrichova, H., Hulselmans, G., Rambow, F., Marine, J.C., Geurts, P., Aerts, J., et al. (2017). SCENIC: single-cell regulatory network inference and clustering. Nat Methods *14*, 1083–1086. 10.1038/nmeth.4463.

57. Faith, J.J., Hayete, B., Thaden, J.T., Mogno, I., Wierzbowski, J., Cottarel, G., Kasif, S., Collins, J.J., and Gardner, T.S. (2007). Large-Scale Mapping and Validation of Escherichia coli Transcriptional Regulation from a Compendium of Expression Profiles. PLOS Biology *5*, e8. 10.1371/journal.pbio.0050008.

58. Haure-Mirande, J.-V., Audrain, M., Ehrlich, M.E., and Gandy, S. (2022). Microglial TYROBP/DAP12 in Alzheimer's disease: Transduction of physiological and pathological signals across TREM2. Molecular Neurodegeneration *17*, 55. 10.1186/s13024-022-00552-w.

59. Troutman, T.D., Kofman, E., and Glass, C.K. (2021). Exploiting dynamic enhancer landscapes to decode macrophage and microglia phenotypes in health and disease. Mol Cell *81*, 3888–3903. 10.1016/j.molcel.2021.08.004.

60. Goldmann, T., and Prinz, M. (2013). Role of Microglia in CNS Autoimmunity. Clin Dev Immunol *2013*, 208093. 10.1155/2013/208093.

61. Martinez, F.O., and Gordon, S. (2014). The M1 and M2 paradigm of macrophage activation: time for reassessment. F1000Prime Rep *6*, 13. 10.12703/P6-13.

62. Jorfi, M., Maaser-Hecker, A., and Tanzi, R.E. (2023). The neuroimmune axis of Alzheimer's disease. Genome Medicine *15*, 6. 10.1186/s13073-023-01155-w.

63. Bettelli, E., Dastrange, M., and Oukka, M. (2005). Foxp3 interacts with nuclear factor of activated T cells and NF-kappa B to repress cytokine gene expression and effector functions of T helper cells. Proc Natl Acad Sci U S A *102*, 5138–5143. 10.1073/pnas.0501675102.

64. Ono, M., Yaguchi, H., Ohkura, N., Kitabayashi, I., Nagamura, Y., Nomura, T., Miyachi, Y., Tsukada, T., and Sakaguchi, S. (2007). Foxp3 controls regulatory T-cell function by interacting with AML1/Runx1. Nature *446*, 685–689. 10.1038/nature05673.

65. Kimura, R., Kamino, K., Yamamoto, M., Nuripa, A., Kida, T., Kazui, H., Hashimoto, R., Tanaka, T., Kudo, T., Yamagata, H., et al. (2007). The DYRK1A gene, encoded in chromosome 21 Down syndrome critical region, bridges between β-amyloid production and tau phosphorylation in Alzheimer disease. Human Molecular Genetics *16*, 15–23. 10.1093/hmg/ddl437.

66. Dijon, M., Bardin, F., Murati, A., Batoz, M., Chabannon, C., and Tonnelle, C. (2008). The role of Ikaros in human erythroid differentiation. Blood *111*, 1138–1146. 10.1182/blood-2007-07-098202.

67. Rosenthal, S.L., Barmada, M.M., Wang, X., Demirci, F.Y., and Kamboh, M.I. (2014). Connecting the Dots: Potential of Data Integration to Identify Regulatory SNPs in Late-Onset Alzheimer's Disease GWAS Findings. PLOS ONE *9*, e95152. 10.1371/journal.pone.0095152.

68. The UniProt Consortium (2021). UniProt: the universal protein knowledgebase in 2021. Nucleic Acids Research *49*, D480–D489. 10.1093/nar/gkaa1100.

69. Manocha, G.D., Ghatak, A., Puig, K.L., Kraner, S.D., Norris, C.M., and Combs, C.K. (2017). NFATc2 Modulates Microglial Activation in the AβPP/PS1 Mouse Model of Alzheimer's Disease. J Alzheimers Dis *58*, 775–787. 10.3233/JAD-151203.

70. Huang, K., Marcora, E., Pimenova, A.A., Di Narzo, A.F., Kapoor, M., Jin, S.C., Harari, O., Bertelsen, S., Fairfax, B.P., Czajkowski, J., et al. (2017). A common haplotype lowers PU.1 expression in myeloid cells and delays onset of Alzheimer's disease. Nat Neurosci *20*, 1052–1061. 10.1038/nn.4587.

71. Satoh, J.-I., Asahina, N., Kitano, S., and Kino, Y. (2014). A Comprehensive Profile of ChIP-Seq-Based PU.1/Spi1 Target Genes in Microglia. Gene Regul Syst Bio *8*, 127–139. 10.4137/GRSB.S19711.

72. Efthymiou, A.G., and Goate, A.M. (2017). Late onset Alzheimer's disease genetics implicates microglial pathways in disease risk. Mol Neurodegener *12*, 43. 10.1186/s13024-017-0184-x.

73. Rustenhoven, J., Smith, A.M., Smyth, L.C., Jansson, D., Scotter, E.L., Swanson, M.E.V., Aalderink, M., Coppieters, N., Narayan, P., Handley, R., et al. (2018). PU.1 regulates Alzheimer's disease-associated genes in primary human microglia. Molecular Neurodegeneration *13*, 44. 10.1186/s13024-018-0277-1.

74. Schwartzentruber, J., Cooper, S., Liu, J.Z., Barrio-Hernandez, I., Bello, E., Kumasaka, N., Young, A.M.H., Franklin, R.J.M., Johnson, T., Estrada, K., et al. (2021). Genome-wide meta-analysis, fine-mapping and integrative prioritization implicate new Alzheimer's disease risk genes. Nature Genetics *53*, 392–402. 10.1038/s41588-020-00776-w.

75. Samaan, Z., Lee, Y.K., Gerstein, H.C., Engert, J.C., Bosch, J., Mohan, V., Diaz, R., Yusuf, S., Anand, S.S., Meyre, D., et al. (2015). Obesity genes and risk of major depressive disorder in a multiethnic population: a cross-sectional study. The Journal of Clinical Psychiatry *76*, e1611-1618. 10.4088/JCP.14m09720.

76. Rasheed, M., Asghar, R., Firdoos, S., Ahmad, N., Nazir, A., Ullah, K.M., Li, N., Zhuang, F., Chen, Z., and Deng, Y. (2022). A Systematic Review of Circulatory microRNAs in Major Depressive Disorder: Potential Biomarkers for Disease Prognosis. International Journal of Molecular Sciences *23*, 1294. 10.3390/ijms23031294.

77. Piñero, J., Ramírez-Anguita, J.M., Saüch-Pitarch, J., Ronzano, F., Centeno, E., Sanz, F., and Furlong, L.I. (2020). The DisGeNET knowledge platform for disease genomics: 2019 update. Nucleic Acids Research *48*, D845–D855. 10.1093/nar/gkz1021.

78. Lobo-Silva, D., Carriche, G.M., Castro, A.G., Roque, S., and Saraiva, M. (2016). Balancing the immune response in the brain: IL-10 and its regulation. Journal of Neuroinflammation *13*, 297. 10.1186/s12974-016-0763-8.

79. Huang, G., Zhang, P., Hirai, H., Elf, S., Yan, X., Chen, Z., Koschmieder, S., Okuno, Y., Dayaram, T., Growney, J.D., et al. (2008). PU.1 is a major downstream target of AML1 (RUNX1) in adult mouse hematopoiesis. Nat Genet *40*, 51–60. 10.1038/ng.2007.7.

80. Riberdy, J.M., Newcomb, J.R., Surman, M.J., Barbosa, J.A., and Cresswell, P. (1992). HLA-DR molecules from an antigen-processing mutant cell line are associated with invariant chain peptides. Nature *360*, 474–477. 10.1038/360474a0.

81. Shiroishi, M., Kuroki, K., Ose, T., Rasubala, L., Shiratori, I., Arase, H., Tsumoto, K., Kumagai, I., Kohda, D., and Maenaka, K. (2006). Efficient leukocyte Ig-like receptor signaling and crystal structure of disulfide-linked HLA-G dimer. J Biol Chem *281*, 10439–10447. 10.1074/jbc.M512305200.

82. Herring, C.A., Simmons, R.K., Freytag, S., Poppe, D., Moffet, J.J.D., Pflueger, J., Buckberry, S., Vargas-Landin, D.B., Clément, O., Echeverría, E.G., et al. (2022). Human prefrontal cortex gene regulatory dynamics from gestation to adulthood at single-cell resolution. Cell *185*, 4428-4447.e28. 10.1016/j.cell.2022.09.039.

83. Shaulian, E., and Karin, M. (2002). AP-1 as a regulator of cell life and death. Nat Cell Biol *4*, E131-136. 10.1038/ncb0502-e131.

84. Denecke, B., Meyerdierks, A., and Böttger, E.C. (1999). RGS1 is expressed in monocytes and acts as a GTPase-activating protein for G-protein-coupled chemoattractant receptors. J Biol Chem *274*, 26860–26868. 10.1074/jbc.274.38.26860.

85. Wertz, I.E., O'Rourke, K.M., Zhou, H., Eby, M., Aravind, L., Seshagiri, S., Wu, P., Wiesmann, C., Baker, R., Boone, D.L., et al. (2004). De-ubiquitination and ubiquitin ligase domains of A20 downregulate NF-kappaB signalling. Nature *430*, 694–699. 10.1038/nature02794.

86. GrandPré, T., Nakamura, F., Vartanian, T., and Strittmatter, S.M. (2000). Identification of the Nogo inhibitor of axon regeneration as a Reticulon protein. Nature *403*, 439–444. 10.1038/35000226.

87. Van de Sande, B., Flerin, C., Davie, K., De Waegeneer, M., Hulselmans, G., Aibar, S., Seurinck, R., Saelens, W., Cannoodt, R., Rouchon, Q., et al. (2020). A scalable SCENIC workflow for single-cell gene regulatory network analysis. Nat Protoc *15*, 2247–2276. 10.1038/s41596-020-0336-2.

88. Zeng, Q., Man, R., Luo, Y., Zeng, L., Zhong, Y., Lu, B., and Wang, X. (2017). IRF-8 is Involved in Amyloid-β1-40 (Aβ1-40)-induced Microglial Activation: a New Implication in Alzheimer's Disease. Journal of Molecular Neuroscience *63*, 159–164. 10.1007/s12031-017-0966-1.

89. Ma, J., Jiang, T., Tan, L., and Yu, J.-T. (2015). TYROBP in Alzheimer's disease. Mol Neurobiol *51*, 820–826. 10.1007/s12035-014-8811-9.

90. Papuć, E., and Rejdak, K. (2018). The role of myelin damage in Alzheimer's disease pathology. Arch Med Sci *16*, 345–351. 10.5114/aoms.2018.76863.

91. Akhtar, M.W., Kim, M.-S., Adachi, M., Morris, M.J., Qi, X., Richardson, J.A., Bassel-Duby, R., Olson, E.N., Kavalali, E.T., and Monteggia, L.M. (2012). In Vivo Analysis of MEF2 Transcription Factors in Synapse Regulation and Neuronal Survival. PLOS ONE *7*, e34863. 10.1371/journal.pone.0034863.

92. Han, J., Jiang, Y., Li, Z., Kravchenko, V.V., and Ulevitch, R.J. (1997). Activation of the transcription factor MEF2C by the MAP kinase p38 in inflammation. Nature *386*, 296–299. 10.1038/386296a0.

93. Kim, J.H. (2018). Genetics of Alzheimer's Disease. Dement Neurocogn Disord *17*, 131–136. 10.12779/dnd.2018.17.4.131.

94. Anderson, A.G., Rogers, B.B., Loupe, J.M., Rodriguez-Nunez, I., Roberts, S.C., White, L.M., Brazell, J.N., Bunney, W.E., Bunney, B.G., Watson, S.J., et al. (2023). Single nucleus multiomics identifies ZEB1 and MAFB as candidate regulators of Alzheimer's disease-specific cis-regulatory elements. Cell Genomics *0*. 10.1016/j.xgen.2023.100263.

95. Lutz, M.W., Sprague, D., Barrera, J., and Chiba-Falek, O. (2020). Shared genetic etiology underlying Alzheimer's disease and major depressive disorder. Translational Psychiatry *10*. 10.1038/s41398-020-0769-y.

96. Khullar, S., and Wang, D. (2023). Predicting brain-regional gene regulatory networks from multi-omics for Alzheimer's disease phenotypes and Covid-19 severity. Human Molecular Genetics, ddad009. 10.1093/hmg/ddad009.

97. Bellenguez, C., Küçükali, F., Jansen, I.E., Kleineidam, L., Moreno-Grau, S., Amin, N., Naj, A.C., Campos-Martin, R., Grenier-Boley, B., Andrade, V., et al. (2022). New insights into the genetic etiology of Alzheimer's disease and related dementias. Nat Genet, 1–25. 10.1038/s41588-022-01024-z.

98. Jin, T., Rehani, P., Ying, M., Huang, J., Liu, S., Roussos, P., and Wang, D. (2021). scGRNom: a computational pipeline of integrative multi-omics analyses for predicting cell-type disease genes and regulatory networks. Genome Med *13*, 95. 10.1186/s13073-021-00908-9.

99. Yan, J., Risacher, S.L., Shen, L., and Saykin, A.J. (2018). Network approaches to systems biology analysis of complex disease: integrative methods for multi-omics data. Briefings in bioinformatics *19*, 1370–1381. 10.1093/bib/bbx066.

100. Dong, X., Liu, C., and Dozmorov, M. (2021). Review of multi-omics data resources and integrative analysis for human brain disorders. Brief Funct Genomics *20*, 223–234. 10.1093/bfgp/elab024.