



OPEN

An integrative network analysis framework for identifying molecular functions in complex disorders examining major depressive disorder as a test case

Anup Mammen Oommen^{1,4}, Stephen Cunningham^{1,4}✉, Páraic S. O'Suilleabháin^{2,5}, Brian M. Hughes³ & Lokesh Joshi^{1,4}✉

In addition to the psychological depressive phenotype, major depressive disorder (MDD) patients are also associated with underlying immune dysregulation that correlates with metabolic syndrome prevalent in depressive patients. A robust integrative analysis of biological pathways underlying the dysregulated neural connectivity and systemic inflammatory response will provide implications in the development of effective strategies for the diagnosis, management and the alleviation of associated comorbidities. In the current study, focusing on MDD, we explored an integrative network analysis methodology to analyze transcriptomic data combined with the meta-analysis of biomarker data available throughout public databases and published scientific peer-reviewed articles. Detailed gene set enrichment analysis and complex protein–protein, gene regulatory and biochemical pathway analysis has been undertaken to identify the functional significance and potential biomarker utility of differentially regulated genes, proteins and metabolite markers. This integrative analysis method provides insights into the molecular mechanisms along with key glycosylation dysregulation underlying altered neutrophil-platelet activation and dysregulated neuronal survival maintenance and synaptic functioning. Highlighting the significant gap that exists in the current literature, the network analysis framework proposed reduces the impact of data gaps and permits the identification of key molecular signatures underlying complex disorders with multiple etiologies such as within MDD and presents multiple treatment options to address their molecular dysfunction.

Abbreviations

MDD	Major depressive disorder
ADD	Anti-depressant drugs
DEGs	Differentially expressed genes
GSEA	Gene set enrichment analysis
PTM	Post translational modification
GWAS	Genome-wide association studies

Depression and other common mental disorders constitute half of the leading causes of disability worldwide¹. It has recently been suggested that mental illness accounts for 32.4% of years lived with disability and 13.0% of disability-adjusted life-years and an estimated cost in excess of £2 billion and US\$ 40 billion every year for the employees in United Kingdom and United States respectively^{2,3}. Among different forms of mental disorders it is estimated that depression disorders (a mood disorder included in Diagnostic and Statistical Manual of Mental Disorders: DSM), is prevalent in ~4.4% of world population¹. Studies examining long-term mortality trajectories

¹Advanced Glycoscience Research Cluster (AGRC), National University of Ireland Galway, Galway, Ireland. ²Department of Psychology, University of Limerick, Limerick, Ireland. ³School of Psychology, National University of Ireland Galway, Galway, Ireland. ⁴Centre for Research in Medical Devices (CÚRAM), National University of Ireland Galway, Galway, Ireland. ⁵Health Research Institute, University of Limerick, Limerick, Ireland. ✉email: stephen.cunningham@nuigalway.ie; lokesh.joshi@nuigalway.ie

consider major depressive disorder (MDD), a major form of depressive disorder, as a critical mortality risk factor⁴. Epidemiological studies have demonstrated that the co-occurrence of diverse components of metabolic syndrome (high blood pressure, lipid abnormalities, hyperglycaemia and central obesity) with MDD, is associated with a fourfold increased risk of cardiovascular disease^{5–7}. Diverse factors attributed in the development of MDD include somatic illness, early life stress⁸, socioeconomic factors⁹, and genetic factors¹⁰. Depending on the subtype of MDD (e.g., melancholic, psychotic, catatonic, and atypical depression)¹¹ as well as diverse contributing factors, depressive symptoms are typically manifested as physical and emotional exhaustion, lack of interest or anhedonia, insomnia, impaired concentration, cynicism (depersonalization), and culminate in psychophysiological endpoints such as suicidal ideation and behaviour^{12,13}.

An appreciation of the complex relationship between the clinical symptoms of MDD and its underlying causal biomolecular network is crucial for disease diagnosis as well as clinical management of the disease. Network biology based approaches, coalescing diverse OMICs data with biomolecular interaction networks, have emerged as a powerful integrative and systems-level approach in understanding complex disease-disease and disease-biomolecular associations^{14–16}. When applied, such an analytic approach involves a dimensionality reduction of altered state of biomolecules to more comparable and interpretable set of biological pathways or molecular network information, providing significant insight into the biological mechanisms underlying a disease pathology^{17–20}. Such integrative network analysis approaches have significantly aided in identifying putative diagnostic biomarkers, therapeutic targets and pathophysiological mechanisms underlying complex diseases^{21–23}. Furthermore, the continued generation of large and publicly-available ‘omic’ data resources makes real the opportunity to examine complex disorders through a tiered systems biology approach, combining data collected across independent groups of individuals to identify novel biological parameters and potential causations, aiding to greater clinical interpretation and management. Complex neuropsychiatric diseases such as MDD, attributed by etiologic heterogeneity and a major comorbid risk factor for metabolic syndrome²⁴, can be examined and elucidated at a greater molecular level by such an integrative systems biology approach²⁵.

A number of studies had proposed comprehensive analysis framework at the systems level for characterizing functional biological pathways of candidate genes, proteins and metabolites in MDD^{26–31}. The findings and of these studies, focusing on identifying diagnostic markers and pathophysiological mechanisms associated with the symptoms of MDD revealed a bidirectional interaction between systemic stress, chronic low grade inflammatory response, altered neuronal pathways and synaptic signaling in the pathogenesis^{28–31}. Accumulating evidence suggest that the etiopathogenesis of MDD largely involve biochemical abnormalities in monoaminergic systems (serotonin, norepinephrine, and dopamine), glutamatergic neurotransmission, and structural abnormalities in the prefrontal projection systems^{32–34}. The majority of these pathway-centric enrichment analysis methods has largely been dependent on comprehensive sets of online databases, which differ significantly in terms of representation of biological pathways and statistical enrichment analysis conducted³⁵. Additionally, these databases often under represents or lack the relevance of post-translational modifications (PTM), a critical regulator of core biological pathways under health and disease conditions. For example, emerging data indicate alteration in glycosylation PTM, a critical modulator of neuronal functions and immune responses^{36–38}, to be significantly correlated with the inflammatory phenotype observed in MDD subjects. As examples of altered glycosylation, levels of the sialylated glycan structure Neu5Acα2-6GalNAc on plasma proteins³⁹, high branching plasma *N*-glycans with tri and tetra-sialylation in males and monogalactosylated *N*-glycome on IgG4 subclass in females⁴⁰, non-fucosylated biantennary glycans as well as α1-3-fucosylated triantennary glycans have all been shown to be significantly correlated with inflammatory cytokines and pathways in MDD subjects⁴¹. Moreover, majority of the systems biology approaches focusing primarily on transcriptomic or genomic data driven pathway enrichment analysis, often lack integration of non-genetic molecular measurements generated from the clinical samples. These data resources often available in the form of evidence based umbrella reviews; meta-analysis or systematic reviews provide a reliable source of molecular markers that are significantly associated with the disease phenotype.

In the current study, an integrative network analysis framework for the detailed biological process enrichment analysis to compile a comprehensive coverage of core biological processes associated with the MDD phenotype was undertaken. Diverse publicly available transcriptomic datasets associated with the MDD phenotype along with reported protein and metabolite markers, were subjected to detailed biological process enrichment analysis. In parallel and to compliment these datasets the biological role and significance of glycosylation PTM (lectins as well as glycoconjugate structures) was undertaken and integrated into the network analysis to assess correlation with the neurological function and inflammatory responses which are reported cellular dysfunctions in MDD subjects^{32–34,42–50}.

Methods

The design methodology explored for the integrative analysis encompassing transcriptomic assessment, reported bibliographic databases, open knowledge platforms and ‘omic’ databases is depicted in Fig. 1. The goal was to evaluate the functional significance and potential of this methodology in the identification of molecular function and biological pathway association of genes, proteins and metabolites for MDD, extending to the examination of the role of glycosylation PTM.

Gene expression data selection, bibliographic search and knowledge databases. For the transcriptomic analysis, public database Gene Expression Omnibus (GEO) (<https://www.ncbi.nlm.nih.gov/geo/>) was queried using the key word “Major Depressive Disorder”. Datasets included both blood and brain samples from MDD subjects and only those datasets with ≥ 3 replicates for both test and control samples were considered. Datasets without gene annotation were excluded from the analysis. For understanding the gene expression changes under treated conditions, data obtained from MDD subjects treated with anti-depressants drugs (ADD)

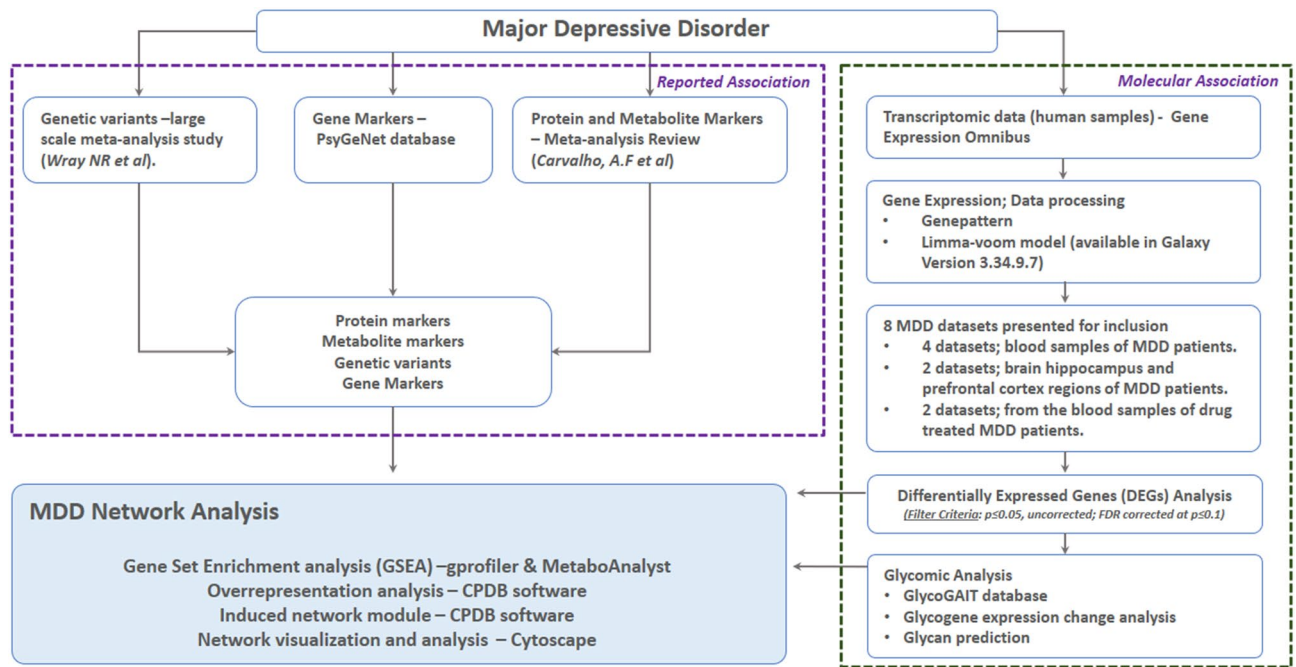


Figure 1. A flowchart depicting the sequential steps adopted in the current study for the integrative data analysis methodology. Various online software analysis tools as mentioned in the diagram were leveraged along with our proprietary GlycoGAIT database for detailed gene set enrichment analysis, complex protein–protein, gene regulatory and biochemical interaction pathway analysis.

were also included (using the generic names of ADD, Suppl. Table 1). As the number of significant datasets was low, data generated in cell line and in vivo model system were incorporated. For data processing and gene annotation, online software tools GenePattern (<http://software.broadinstitute.org/cancer/software/genepattern/>)⁵¹ and Galaxy (<https://usegalaxy.org/>)⁵² were used throughout, following the procedures detailed in user manuals. In GenePattern, preprocessing of RNA-Seq data including normalization, missing value imputation, collapsing multiple probe set expression values into a single expression values were performed using the VoomNormalize (v2); ImputeMissingValues and CollapseDataset modules respectively. Differential expression of the preprocessed data files were then performed for individual datasets using the ComparativeMarkerSelection module⁵³. While, in Galaxy the limma-voom tool was utilized for the differential expression analysis accepting the default Trimmed Mean of M values (TMM) normalization method and applying filters to remove lowly expressed gene in each datasets selected for the analysis (we choose to retain genes if they are expressed at a counts per million above 0.5 in at least three samples). Significance of marker genes were calculated using default *p*-value adjustment methods in analysis⁵⁴. For few datasets normalized, log₂ transformed gene expression matrix files were used directly from the GEO databases, wherever available.

A systematic query of the PUBMED bibliographic database⁵⁵ using the search query [“major depressive disorder” (Title/Abstract)] AND [biomarker(Title/Abstract)] was performed. After filtering results using “review paper” and “full text” as parameters, a comprehensive umbrella review by Carvalho et al.⁵⁶ was identified reporting evidence based correlation of protein and metabolite markers for the 5 most prevalent and high burden major mental disorders. An Umbrella review systematically evaluates and collects information from multiple systematic reviews and meta-analyses on all outcomes of a given topic for which these have been performed, eliminating or greatly reducing reporting bias⁵⁷. From Carvalho et al., protein and metabolite markers reported with ‘strong correlation’ with the MDD phenotype was presented for integrative analysis.

Through PsyGeNET (Psychiatric disorders Gene association NETwork), genetic markers reported to be associated with MDD phenotype were compiled⁵⁸. Genetic variants from a published meta-analysis of 130,644 MDD cases⁵⁹ was also examined for detailed functional genomic data enrichment analysis to build a comprehensive pathway coverage of reported genetic and non-genetic markers associated with MDD.

Glycosylation process related gene pool. For the gene expression analysis of glycosylation process related genes, the GlycoGAIT database⁶⁰, an interactive web database developed within our research group was utilised. For this analysis, the GlycoGAIT database was enriched through manual capturing of enzymatic reactions for the glycosyltransferase and glycosidase enzymes, which were extracted from the BRENDA enzyme database (<https://www.brenda-enzymes.org/index.php>) and ExPASy bioinformatics resource portal (<https://www.expasy.org/>). For a number of interactions, where the reaction information is not available, interactions were curated manually from PUBMED sources. The reaction formats as well as the substrate and ligand short form representation are manually formatted to make it consistent throughout. Information on proteoglycan genes were extracted from the HGNC database (<https://www.genenames.org/>). For linking glycosylation reac-

tions to glycan structure predictions, well characterized human glycan epitope structures were also captured by manually selecting epitope structures from the GlycoEpitope database (https://www.glycoepitope.jp/epitopes/epitope_list). This list was further enriched with additional epitope structures reported to be present on immune cells from published literature sources^{61–64}.

Nomenclature mapping and gene family association. Prior to enrichment analysis, mapping of entity names to their respective standard nomenclatures were performed for metabolite entities using Chemical Translational Service (<http://cts.fiehnlab.ucdavis.edu/>) and for the protein and gene entities using the multi-symbol checker tool in HGNC database. This standard nomenclature mapping ensures no metabolite, protein or gene markers are missed during the enrichment analysis due to name mismatching. For detailed enrichment analysis, the dimensionality of the differentially expressed genes (DEGs) were reduced by generating smaller subset of more closely associated genes through mapping individual genes to respective gene families using the BioMart software available through HGNC web interface (<https://biomart.genenames.org/>). Top ranking gene family clusters were then identified from this list using the built-in statistical function in Excel “RANK function”.

Gene set enrichment analysis, cellular process mapping and network visualization. Functional enrichment analysis of the literature based protein markers was performed using the open-source software and web server—gProfiler⁶⁵. The HGNC gene symbol of the proteins were used as the query in the g:GOST functional profiling interface in gProfiler using Homo sapiens as the organism species, g:SCS threshold as the significance threshold and the user-defined *p*-value threshold as 0.05 in the advanced options (Suppl. Table 2a). Similarly, for the literature based metabolite markers enrichment analysis was performed using the web-based analytical pipeline for high-throughput metabolomics data analysis—MetaboAnalyst⁶⁶ and the Consensus-PathDB—(CPDB—<http://cpdb.molgen.mpg.de/>)⁶⁷. For the enrichment analysis, metabolite name mapping was performed manually to their respective KEGG, PUBCHEM and HMDB IDs^{68–70}. Due to greater coverage we used CPDB for subsequent analysis by inputting the KEGG IDs using the overrepresentation analysis for metabolite set function. The default list of 13 pathway-based databases were selected with a *p*-value cutoff 0.01.

For gene set enrichment analysis (GSEA), HGNC gene symbols of the DEGs were used as the query input in the enrichment analysis tools such as the ‘analysis’ feature available in the Enrichr tool⁷¹ or the ‘over-representation analysis’ feature available in CPDB⁶⁷. Besides, the induced network module analysis feature in CPDB was also used for identifying possible functional relationship between DEGs such as protein–protein, gene–protein and protein–metabolite reactions. The interaction network model generated for the DEGs is extracted as an XML-based file format using the CPDB visualization interface and ran through Cytoscape software (<http://www.cytoscape.org/>)⁷² for network visualization and analysis. The interaction edges are colored based on the nature of interaction and size of each node in the network is adjusted based on “neighborhood connectivity” wherein small size represents less connected entities and large size represents highly connected entities. Colours of the nodes are adjusted based on gene expression data.

Results

Functional enrichment analysis of markers identified from literature sources and knowledge databases. From the meta-analyses review published by Carvalho, et al.⁵⁶, 7 protein markers and 10 metabolite markers were identified that were significantly associated with MDD subjects. Functional enrichment analysis of these protein markers revealed cell proliferation, cell differentiation and growth factor signaling pathway as the major categories of biological pathways represented by the protein markers. Similarly, for the metabolite markers result obtained from the CPDB and MetaboAnalyst platform was primarily enriched in pathways associated with amino acid, nicotinamide, arachidonic acid and antioxidant metabolism (Suppl. Tables 2a, 2b, 2c; Suppl. Fig. 1). Biological processes identified from the enrichment analysis of the 62 genetic markers from the PsyGeNet database were broadly related to synaptic signaling, neurotransmitter metabolic process, neurotransmitter transport and neuron development (Suppl. Table 3a,b). Additional enrichment analysis of genetic variants identified from a recently published meta-analysis of 130,644 MDD cases⁵⁹ was performed. 67 gene were identified from this study, which either contain or are in linkage disequilibrium with a reported single nucleotide polymorphism. The GO biological process association of these gene markers were significantly enriched with the neuron generation and development processes (Suppl. Table 4a,b).

Differentially expressed genes. Using Comparative Marker Selection model (Genepattern software) for the microarray data and Limma-voom model (Galaxy Version 3.34.9.7) for the high throughput sequencing data^{73,74}, 8 out of 22 datasets were identified (Suppl. Table 5) that contained significantly differentially expressed genes (DEGs) based on the filter criteria set out ($p \leq 0.05$ uncorrected, FDR corrected at $p \leq 0.1$). The selected datasets from MDD subjects includes 4 datasets generated from the blood samples and 2 datasets from the brain hippocampus and prefrontal cortex regions. To compare the gene expression pattern under pathological and treated conditions, transcriptomics data under ADD treated conditions were also included in the analysis. However, owing to the current lack of datasets from the brain samples, only two datasets from the blood samples of ADD treated MDD subjects were included within the analysis undertaken. The DEGs identified after the filtering were further narrowed down by applying a cut-off value of ~ twofold up or down regulated genes. The resultant DEGs were pooled into three categories (Suppl. Table 6), based on the sample sources and types, which is as, detailed below:

- (i) Blood samples, untreated—228 upregulated genes and 198 down regulated genes.
- (ii) Brain samples, untreated—608 upregulated genes and 738 downregulated genes.

(iii) Drug treated, blood samples—215 upregulated genes and 129 downregulated genes.

Gene family association and GSEA of DEGs. DEGs from the blood and brain samples of MDD subjects and from the blood samples of ADD treated MDD subjects were mapped to their respective gene families (Suppl. Table 7a). Among these, the top ranking gene families were identified by applying a cut-off value of ≥ 5 . This method enabled to systematically narrow down the significant DEGs which covered ~14% to 22% of the genes across three sample categories which include (i) 46 upregulated and 42 downregulated genes from the blood samples (ii) 98 upregulated genes 115 downregulated genes from the brain samples (iii) 47 upregulated genes 27 downregulated genes from the drug treated blood samples. A limited number of significant glyco-genes were identified among the DEGs across the three data categories, with maximum representation from the brain samples (Suppl. Table 7b).

Glycosylation process related gene pool. GlycoGAIT database, compiled together from multiple online databases, captures 564 proteins which represents the unique list of well characterized unique list of lectins, enzymes, transporters and other proteins involved in glycosylation PTM along with a list of 38 genes categorized as proteoglycans (Suppl. Tables 8–10), collectively referred to as glyco-genes hereafter. Information on 331 glycosylation reactions that represent 53 glycosylation reaction patterns in the network format (Suppl. Tables 11, 12) are also incorporated. In addition to glycosylation reactions, the GlycoGAIT database also captured 194 well-characterized glycan epitope structures (Suppl. Table 13).

Enrichment analysis of DEGs—blood samples of MDD subjects. GSEA and the GO biological process overlap analysis of the top ranking clustered gene families (cut off value ≤ 5) for the blood sample DEGs primarily showed enrichment of immune signaling pathways such as cellular response to cytokine stimulus, positive regulation of immune response, platelet activation, neutrophil mediated immunity, markers of myeloid–lymphoid maturation, complement–coagulation cascades and glycosaminoglycan metabolism. Key cellular processes identified for the glyco-gene subset involve mucin type O-Glycan biosynthesis and chondroitin sulfate/dermatan sulfate metabolism (Suppl. Table 14). Significant gaps in process association was observed in various pathway databases for the glyco-gene subsets.

Molecular mechanisms underpinning the GSEA results and the over-representation analysis results were studied by using the induced network module available in CPDB⁶⁷. The interconnected network revealed a strong inflammatory signaling pathway comprising of immunomodulatory receptors, guanine exchange factors, adaptor proteins and non-receptor tyrosine kinases (Fig. 2). Downregulated genes, identified to be linked with this inflammatory signaling cascade, comprised of primarily markers of myeloid and lymphoid maturation that belongs to cluster of differentiation family^{75–79}; few anti-inflammatory proteins such as PILRA (paired immunoglobulin like type 2 receptor alpha)⁸⁰, SLAMF7 (SLAM family member-7)⁸¹ and LILRA3 (leukocyte immunoglobulin like receptor-A3)⁸². Other category identified from this group consist of proteins involved in cell adhesion and migration such as PLAUR (plasminogen activator, urokinase receptor)⁸³, CLEC4E (C-type lectin domain family 4 member E)⁸⁴ and ICAM1 (intercellular adhesion molecule-1)⁸⁵.

Protein–protein interaction network captured the association of glyco-gene DEGs majorly with the anti-coagulant protein PLAUR⁸⁶ and the Integrin signaling cascade (Fig. 2). Binary interaction network between the upregulated glyco-gene GALNT14 and the proteoglycan VCAN (versican) with the SELPLG (selectin P ligand)—a critical protein involved in leukocyte trafficking⁸⁷, might indicate enhanced signaling pathways for leukocyte adhesion. Tyrosine sulfation of leukocyte adhesion molecules, chemokines and chemokine receptors has been implicated in promoting atherosclerosis by enhanced recruitment of mononuclear cells^{88,89} and is captured in the network model as a protein–protein interaction between TPST1 (tyrosylprotein sulfotransferase-1) and the chemokine receptor CCR2. One of the most striking observations in the downregulated glyco-gene subsets was the decreased expression of enzymes such as DSE (dermatan sulfate epimerase), CHST12 (carbohydrate sulfotransferase-12) and UST (uronyl 2-sulfotransferase) involved in chondroitin/dermatan sulfate biosynthesis, the two immunomodulatory glycosaminoglycans^{90–94}.

Significant gap in enrichment analysis databases were identified for a number of glyco-genes in the blood sample DEGs. These glyco-genes, which are reported to be critical regulators of inflammatory responses, involve OGA (O-GlcNAcase) and B4GALT5 (β -1,4-galactosyltransferase-5)^{95–99}. Other few important glyco-genes identified in this category are lectins involved in the plasma protein clearance, immune cell adhesion and anti-coagulant properties such as C-type lectin receptors ASGR1 (asialoglycoprotein receptor-1), DGCR2 (DiGeorge syndrome critical region gene-2) and thrombomodulin (THBD)^{100,101}.

Enrichment analysis of DEGs—brain samples of MDD subjects. Carrier and channel proteins involved in ion coupled cell transport mechanisms, growth factor signaling pathways as well as protein coding genes involved in neurotransmission and synapse formation were some of the top scoring enriched processes identified from the GSEA and the GO biological process analysis. Similarly, glucose, sialic acid, glycosphingolipid and glycosaminoglycan metabolism, glycosylphosphatidylinositol (GPI) anchor and O-glycan biosynthesis were the top scoring enriched processes for the glyco-gene subsets. Few biological processes associated with ion homeostasis, chemokine mediated signaling pathway, cellular response to cytokine stimulus, ubiquitination, RNA metabolic and biosynthetic processes were also identified from these DEGs (Suppl. Table 15).

A unique list of 243 DEGs were selected for the induced network model based analysis which extracted a highly interconnected transcriptional network with the ubiquitination machinery, neurotransmitter and ion homeostasis transporters as well as adhesion molecules (Fig. 3). It may be interesting to evaluate whether the altered transcriptional signatures and its regulatory interaction network with the transporter machinery might

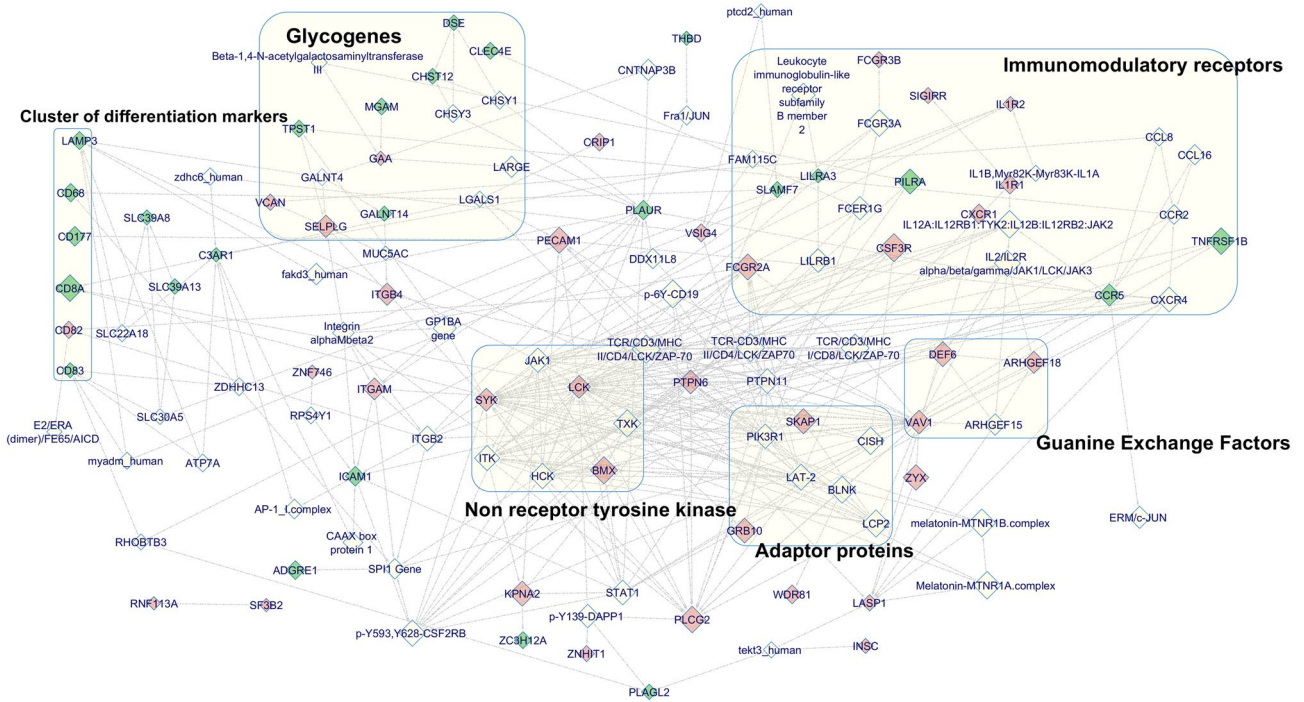


Figure 2. Network visualization of the integrated network module of both upregulated and downregulated genes identified from the blood samples of MDD patients. The interaction network is generated using the CPDB induced network modules by inputting the gene subsets from the top ranking gene family clusters along with the glycogenes. The compact subnetwork is created by applying a z-score threshold of 20 in the CPDB user interface. Exported network model from the CPDB is processed in the Cytoscape visualization tool using GeneMANIA Force Directed Layout and the entities are manually aligned. Colours of the nodes are adjusted based on gene expression data wherein wine red colour represents upregulated genes and dark green represents the downregulated genes. Major network modules with maximum representation of the DEGs are highlighted in the graph using a light yellow background.

explain the decreased norepinephrine, GABA and increased levels of glutamate neurotransmitters observed in MDD subjects^{102–105}. Moreover, DEGs representing the Fibroblast growth factor (FGF) signaling pathway, dysregulation of which was associated with the anxiolytic and antidepressant function with severely depressed humans¹⁰⁶, were found to be connected with the ubiquitination and the solute carrier family proteins in the network. Similarly, the upregulated expression pattern of neuropeptide GAL (galanin) among the DEGs was consistent with a previous report¹⁰⁷ and might indicate dysregulation in pathways involved in mood regulation.

Focused analysis of the glycogene DEGs enabled us to identify few of them to be associated with the transcriptional hub machinery (Fig. 3). These genes are primarily involved in the complex N-glycan synthesis, heparan sulfate biosynthesis, ganglioside metabolism, GPI-anchor biosynthesis processes and polysialic acid modification of neuronal cell adhesion molecules¹⁰⁸. The decreased expression of ST8SIA3 (ST8 α -N-acetyl-neuraminidase α -2,8-sialyltransferase-3) involved in the synthesis of c-series ganglioside and HEXB (hexosaminidase) enzyme involved in regulating GM2 ganglioside content, might indicate altered neuronal repair¹⁰⁹ synaptic plasticity and neurogenesis^{110,111}. Interestingly, in the network model, HEXB gene was associated with the hypoxia regulated hexokinase isoform HK2 and glucokinase enzyme GCK, which are critical regulators of glucose metabolism¹¹⁰ and hence might indicate a role of adaptive mechanism to support brain energy metabolism. Similarly, downregulated expression pattern of gene transcripts involved in the GPI-anchor biosynthesis pathway: PIGN (phosphatidylinositol glycan anchor biosynthesis class N) and PIGF (phosphatidylinositol glycan anchor biosynthesis class F) might indicate altered cell surface expression of various proteins that play vital role in neuronal differentiation, synapse development as well as axon guidance^{110,112–114}.

Enrichment analysis of DEGs—blood samples of ADD treated MDD subjects. GSEA and GO biological process categorization of DEGs from the blood samples of ADD treated MDD subjects highlighted phagosome formation, endosomal/lysosomal pathways, positive regulation of cell death and immune responses, antigen processing and presentation, as some of the top scoring enriched pathways. While, for the downregulated gene subsets this include spliceosome, histone modification, mRNA processing, organic cyclic compound or cellular nitrogen compound metabolic process and DNA damage. Similar analysis for the glycogene DEGs yielded sialic acid metabolism, O-linked glycosylation and glycosaminoglycan metabolism as the major enriched cellular processes (Suppl. Table 16).

DEGs from the top ranking gene family clusters were subjected to induced network model based analysis by seeding 83 unique gene list integrated from the upregulated and downregulated gene lists, including the

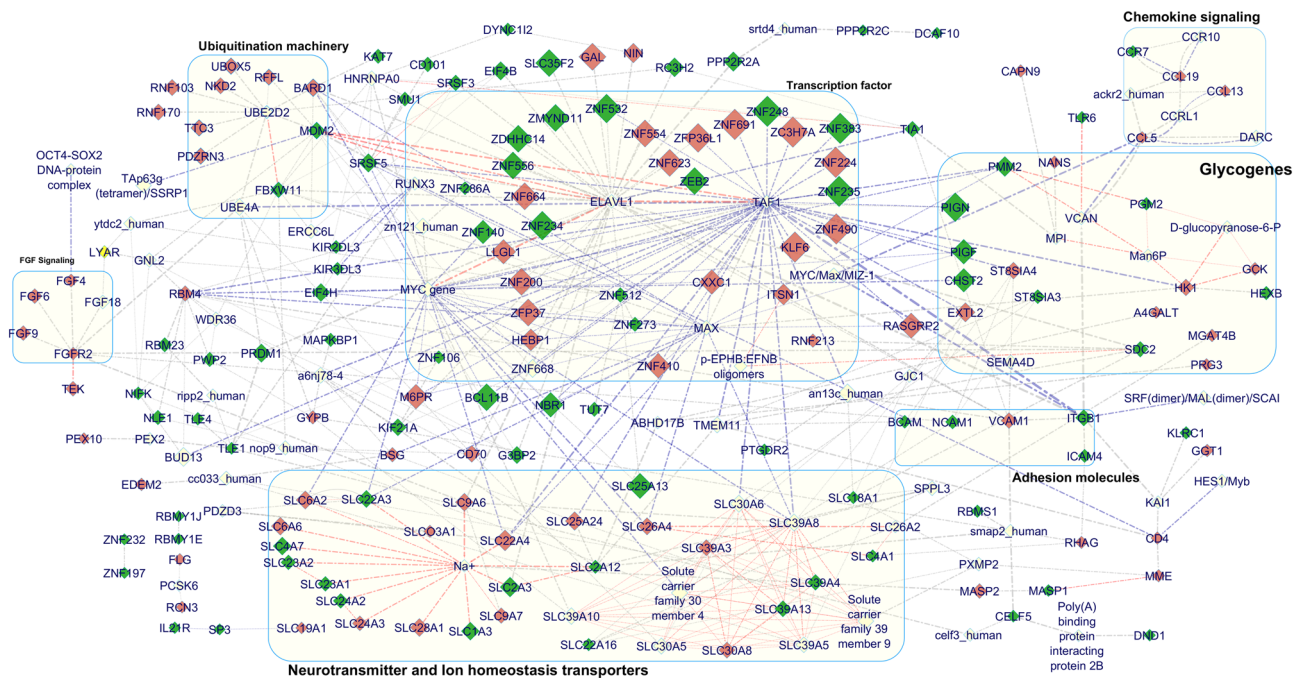


Figure 3. Integrated network model generated for both upregulated and downregulated genes identified from the brain samples of MDD patients using the CPDB induced network modules. The current network model comprises of 218 nodes out of the 243 input genes from the top ranking gene family clusters along with the glycogenes. The compact subnetwork is created by applying a z-score threshold of 20 in the CPDB user interface. Exported network model from the CPDB is manually aligned in the GeneMANIA Force Directed Layout using the Cytoscape visualization tool. Colours of the nodes are adjusted based on gene expression data wherein wine red colour represents upregulated genes and dark green represents the downregulated genes. The DEGs belonging to network modules with maximum representation is highlighted in the graph using light yellow background.

glycogenes. Induced network model based analysis using CPDB generated protein–protein interaction networks linking chromatin remodeling and spliceosome machineries with the proteins involved in antigen presentation and immunomodulatory receptor signaling pathways (Fig. 4). Additional literature search on the entities of this seed network suggests a change in the innate immune response pathway^{115–119} as well as co-stimulation of antigen activated lymphocytes¹²⁰ under drug treated condition, which need to be studied in detail. Decreased expression pattern of the linker histones associated with chromatin modification (Fig. 4) suggest that epigenetic modifications are involved in gene expression reprogramming¹²¹ under drug treated conditions in the peripheral blood mononuclear cells.

Interestingly, among the glycogenes two major glycosyltransferases FUT7 (fucosyltransferase-7) and ST3GAL4 (ST3 beta-galactoside alpha-2,3-sialyltransferase-4) were found to be upregulated and was linked with the TLR and HLD-DRA protein entities (Fig. 4). Another important gene which was found to be upregulated and code for the poly-N-acetyllactosamine synthesizing enzyme, B3GNT8 (UDP-GlcNAc:betaGal-β-1,3-N-acetylglucosaminyltransferase-8), was not found in the induced network model. Due to lack of sufficient information in the induced network model, detailed analysis of the enzymatic reaction and possible glycan structure formation for these enzymes were performed using the GlycoGAT reaction database. From the reaction table these three enzymes were mapped to 19, 25, 32 and 45 reaction codes (Suppl. Table 12). Manual search for the maximum combination of these reaction codes in the epitope table, lead to the identification of possible glycan epitope structures and were mapped to Lewis^a- Lewis^x (Le^a-Le^x), 3'-sialyl Le^a-Le^x (3'sLe^a-Le^x), 3's-Di-Le^a, 3's-Di-Le^x, O-Fucose Glycan /EGF Repeat, Sialyl 6-Sulfo LacNAc, Sialyl 6-Sulfo Le^x, sLe^a, sLe^x, sLe^{x-1}, VIM-2 (CD65, CDw65, sLe^{x-1}) and Myeloglycan structures (Suppl. Table 13). The sialylated Lewis-type blood group antigens as well as the myeloglycan or the polylactosaminolipid structures are known to modulate immune function in both myeloid and lymphoid cells^{122–125}.

Discussion

Diverse integrative network analysis approaches have been adopted in the past to identify the aberrant pathway networks underlying complex diseases by leveraging multiple “-omic” studies¹²⁶. Similar integrative analysis is required to understand the biological implications underlying the phenotypical and physiological expression of complex disorders such as MDD. In addition to the depressive phenotype, MDD subjects are also associated with underlying immune dysregulation, which has been shown to correlate with the metabolic syndrome prevalent in MDD subjects presenting a complex biomarker network¹²⁷.

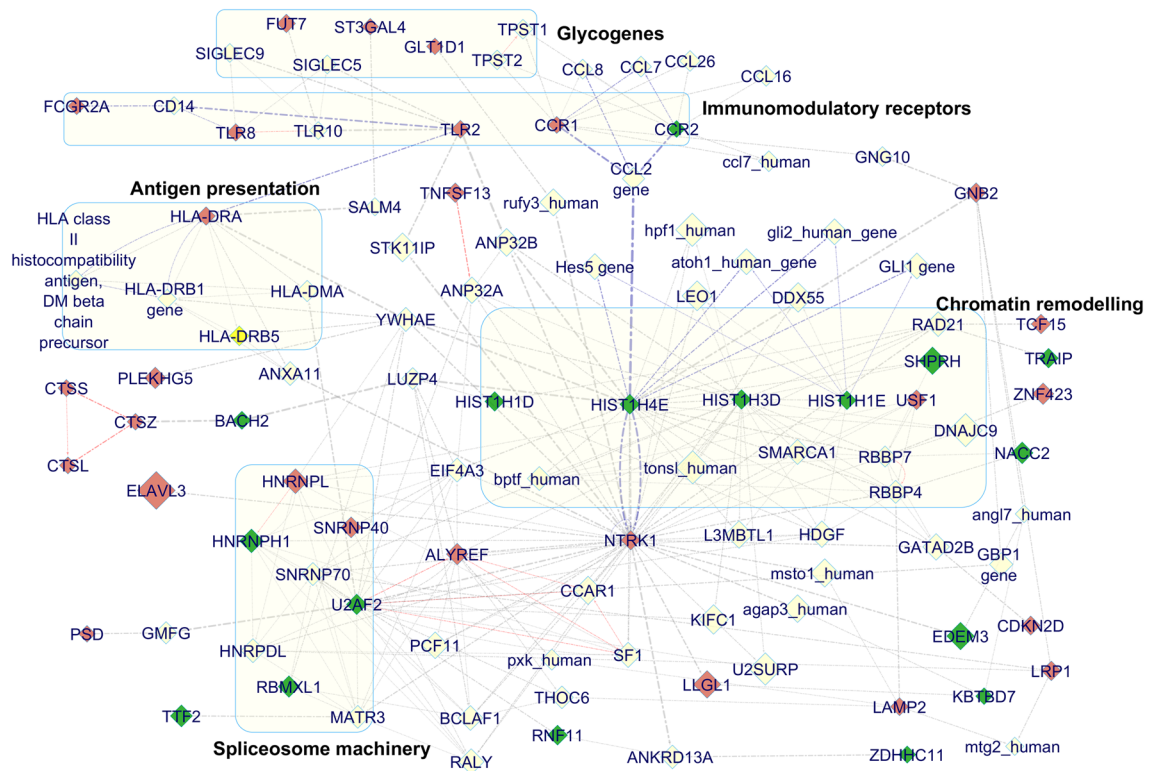


Figure 4. Integrated network model of 82 DEGs from the top ranking gene family clusters identified from the blood samples of anti-depressant drug treated MDD patients using the CPDB induced network modules. The compact network model is created by applying a z-score threshold of 20 in the CPDB user interface. GeneMANIA Force Directed Layout in the Cytoscape visualization tool was used for manually aligning and analyzing the exported network from CPDB. Colours of the nodes are adjusted based on gene expression data wherein wine red colour represents upregulated genes and dark green represents the downregulated genes. The core network modules with maximum DEG representation is highlighted in the graph using light yellow background.

The undertaking to compile all available data for network and systems biology studies for a single disorder is not a trivial task, limited by analysis and computational techniques currently available, and by the complexity of biological interaction and network. The potential to introduce bias, impacting observations and reported findings exists across all integrated and systems biology networks being subjected to numerous independent datasets, subject to different classification and as is the current case with MDD small-sample datasets. Limitations suffered in common across such studies include, (i) restrictions on the basis of inclusion through the use of keywords and terms; (ii) limited datasets and their size and classification; (iii) sampling and confounding factors; (iv) limiting analysis to top ranked findings, potential missing key molecular interactions as network noise, analysis threshold; (v) subject to level of detail and biological functions assigned in databases.

A number of knowledge databases focusing on MDD as well as molecular meta-analysis data are currently available^{26,27,128–131}. There are however, several challenges in the integration of omics data within a network. These extended to but not limited to include experimental and inherent biological noise, differences among experimental platforms, detection bias, and unclear molecular cascades and mechanisms. A common trend, reporting the significant gap in correlation of both genetic and non-genetic biomarkers with the MDD phenotype also exists^{132,133}. Commencing this study, a recent umbrella review was selected to give an evidence-based correlation of both proteins and metabolite markers for MDD⁵⁶. Functional enrichment analysis of the MDD associated peripheral markers from this study revealed enrichment of cellular processes associated with inflammatory response, endothelial dysfunction and neurological functions suggesting a close interrelationship between these biological events in MDD condition, which might drive these patients to higher risk of depressive symptoms and metabolic syndrome^{134,135}.

From the blood samples of MDD subjects, gene expression patterns indicated heightened inflammatory signaling pathways and decreased myeloid and lymphoid maturation as well as anti-inflammatory pathways. Results summarized in this study are consistent with the recent findings of increased neutrophil to lymphocyte ratio, platelet volume and activity in MDD subjects^{5,77,136,137}. Further to, downregulated gene transcripts identified from the blood samples mapped to complement—coagulation cascades indirectly indicating reduced anti-coagulant activity in MDD subjects. A recent proteomic investigation of plasma based biomarkers in MDD subjects also reported increased expression of proteins involved in coagulation cascades¹³⁸ which might lead to prothrombotic platelet phenotype as demonstrated in a study by Lopez-Vilchez et al.¹³. These findings coincide

with the inflammation theory of depression, which suggests a neuro-inflammatory circuit that underlies the development of depression^{139–141}.

Gene expression pattern from the brain samples revealed dysregulated neuroprotective mechanisms. The altered gene expression markers of mRNA splicing, translational machinery and ubiquitination machinery, from the brain samples, might indicate an enhanced cellular stress level and altered gene transcription reprogramming. Current analysis, revealed several gene expression markers that correlate with altered levels of neurotransmitters and dysregulated neuronal ion homeostasis which might indicate defective neuron myelination and synaptic signaling^{10,103–106,142} in MDD subjects. Moreover, the genetic findings from a recent GWAS indirectly supports our methodology as the results obtained here in the current analysis were majorly associated with cytokine and immune responses in addition to neuronal development and morphogenesis, showing strong parallel with the GWAS⁵⁹.

In order to understand the impact on gene expression patterns by ADD treatment, relevant gene expression data were queried from the public databases. In this regard, one of the major limitations is in the lack of availability of sufficient brain sample data from ADD treated MDD subjects. DEGs identified from the blood samples of different ADD treated MDD subjects highlight primarily an increase in innate immune response pathway with co-stimulation of antigen-activated lymphocytes. Whether these pathways drive specifically the anti-inflammatory signaling or augment the innate immune response is not clear but this possibility has been proposed and it needs more detailed research to support this hypothesis. A recent systematic review which studied the impact of ADD on innate and adaptive immunity by Bournazos et al., reported high heterogeneity in results and small number of comparable studies¹⁴³. Considering the limited availability of drug treated samples, studies that are more detailed are required to delineate the molecular mechanisms underlying the impact of ADD on immune response pathways, which is considered as one of the factors that underlie treatment-resistant depression in MDD subjects¹⁴⁴.

Within the GSEA and induced network analysis, DEGs related to glycoconjugates were analysed. Glycosylation PTM as well as glycoconjugate structures are known to play crucial role in modulating central and peripheral nervous system functions and immune responses^{37,145,146}. Results obtained from this study revealed significant gap in databases that support GSEA for glycosylation process related reaction networks. With the aid of GlycoGAT database and manual literature based analysis, meaningful interpretation for the altered glycoconjugate expression was attempted. From the MDD blood samples transcriptional reprogramming of glycoconjugates that favor pro-inflammatory response in lymphocytes and monocytes and downregulation of glycosaminoglycans, which have anti-inflammatory and anti-coagulant properties, were identified. The gene expression pattern from the brain samples indicate increased expression of glycoconjugates involved in synthesis of sialylated structures which are a prominent feature of brain glycoproteins such as those observed on cell adhesion molecules, ion channels and many ligand-activated receptors in neuronal and glial cells^{38,147}. Decreased expression of enzymes involved in GPI pathway and in regulation of GM2 ganglioside content might indicate altered brain glycolipid structures, which are crucial for synaptic plasticity and neurogenesis^{110,111}. One of the most interesting aspects of the glycoconjugate expression data analysis was the prediction of sialylated Lewis-type blood group antigens as well as the myeloglycan or the polylectosaminolipid structure formation in blood monocytes and leukocytes under ADD treatment condition. Hence, it will be interesting to validate how these specific glycan structures fits with the anti-inflammatory claim of these drugs^{148–151} which can potentially developed as diagnostic markers.

Despite significant advancement in the field of 'omics' data analysis and the increasing accessibility to open repositories, no central guideline or platform for an integrative analysis framework has yet been proposed for standardization. For the continued development and acceptance of validity of integrative analysis, standardization at multiple levels is essential. Unlike the analysis of a single dataset, the power of integrative analysis framework stems from the fact that multiple independent data sources can be analyzed to explore trends, similarities or variations in the behavioral patterns of molecular functions and biological processes, which can be correlated with a disease phenotype, outcome and treatment.

The methodology presented here provides an exemplary framework to systematically integrate diverse 'omics' data in an open, inclusive and unbiased manner. Researchers, without the need for advanced computational demands can implement and apply the current approach to available datasets requiring minimal computational resources. Presenting a methodological approach, which could be utilized and adapted for exploiting existing datasets without requiring field expertise using publically open and available resources and repositories. Presenting the inclusion of integrative analysis to become a more routine undertaking and building a mass of analysis which can be refined for standardization and proposing validation albeit through replicated integrative studies within single conditions. Using such an approach the study provides novel insights into the molecular mechanisms underlying altered immune response as well as dysregulated neuronal survival maintenance and synaptic functioning in MDD subjects. Furthermore, for the first time altered glyco-genome expression profile from the blood and brain samples of MDD subjects were explored using GlycoGAT database which might explain the relevance of glycoconjugate structures in modulating the immune responses and neuronal functions.

Received: 1 December 2020; Accepted: 14 April 2021

Published online: 06 May 2021

References

1. Depression and Other Common Mental Disorders: Global Health Estimates. Geneva: World Health Organization; 2017. Licence: CC BY-NC-SA 3.0 IGO.
2. WHO Library Cataloguing-in-Publication Data investing in mental health: evidence for action. World Health Organization. ISBN 978 92 4 156461 8 (NLM classification: WM 30).

3. Vigo, D., Thornicroft, G. & Atun, R. Estimating the true global burden of mental illness. *Lancet Psychiatry* **3**(2), 171–178 (2016).
4. O'Suilleabhain, P. S. & Hughes, B. M. Neuroticism predicts all-cause mortality over 19-years: the moderating effects on functional status, and the angiotensin-converting enzyme. *J. Psychosom. Res.* **110**, 32–37 (2018).
5. Marazziti, D., Rutigliano, G., Baroni, S., Landi, P. & Dell'Osso, L. Metabolic syndrome and major depression. *CNS Spectr.* **19**(4), 293–304 (2014).
6. Vancampfort, D. *et al.* Metabolic syndrome and metabolic abnormalities in patients with major depressive disorder: a meta-analysis of prevalences and moderating variables. *Psychol. Med.* **44**(10), 2017–2028 (2014).
7. Foley, D. L. *et al.* Major depression and the metabolic syndrome. *Twin Res. Hum. Genet.* **13**(4), 347–358 (2010).
8. Saleh, A. *et al.* Effects of early life stress on depression, cognitive performance and brain morphology. *Psychol. Med.* **47**(1), 171–181 (2017).
9. Freeman, A. *et al.* The role of socio-economic status in depression: results from the COURAGE (aging survey in Europe). *BMC Public Health* **16**(1), 1098 (2016).
10. Lohoff, F. W. Overview of the genetics of major depressive disorder. *Curr. Psychiatry Rep.* **12**(6), 539–546 (2010).
11. van Loo, H. M., de Jonge, P., Romeijn, J. W., Kessler, R. C. & Schoevers, R. A. Data-driven subtypes of major depressive disorder: a systematic review. *BMC Med.* **10**, 156 (2012).
12. Kennedy, S. H. Core symptoms of major depressive disorder: relevance to diagnosis and treatment. *Dialogues Clin. Neurosci.* **10**(3), 271–277 (2008).
13. Yang, Y. *et al.* The extrinsic coagulation pathway: a biomarker for suicidal behavior in major depressive disorder. *Sci. Rep.* **6**, 32882 (2016).
14. Naylor, S. & Chen, J. Y. Unraveling human complexity and disease with systems biology and personalized medicine. *Pers. Med.* **7**(3), 275–289 (2010).
15. Furlong, L. I. Human diseases through the lens of network biology. *Trends Genet.* **29**(3), 150–159 (2013).
16. Xiong, M., Feghali-Bostwick, C. A., Arnett, F. C. & Zhou, X. A systems biology approach to genetic studies of complex diseases. *FEBS Lett.* **579**(24), 5325–5332 (2005).
17. Knox, S. S. From “omics” to complex disease: a systems biology approach to gene-environment interactions in cancer. *Cancer Cell Int.* **10**, 11 (2010).
18. Cantini, L. *et al.* Benchmarking joint multi-omics dimensionality reduction approaches for the study of cancer. *Nat. Commun.* **12**(1), 124. <https://doi.org/10.1038/s41467-020-20430-7> (2021).
19. Reimand, J. *et al.* Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. *Nat. Protoc.* **14**(2), 482–517 (2019).
20. Schadt, E. E. Molecular networks as sensors and drivers of common human diseases. *Nature* **461**(7261), 218–223 (2009).
21. Wang, I. M. *et al.* Systems biology approach for new target and biomarker identification. *Curr. Top. Microbiol. Immunol.* **363**, 169–199 (2013).
22. Barabási, A. L., Gulbahce, N. & Loscalzo, J. Network medicine: a network-based approach to human disease. *Nat. Rev. Genet.* **12**(1), 56–68 (2011).
23. Conte, F. *et al.* A paradigm shift in medicine: a comprehensive review of network-based approaches. *Biochim. Biophys. Acta Gene Regul. Mech.* **1863**(6), 194416 (2020).
24. Nousen, E. K., Franco, J. G. & Sullivan, E. L. Unraveling the mechanisms responsible for the comorbidity between metabolic syndrome and mental health disorders. *Neuroendocrinology* **98**(4), 254–266 (2013).
25. Alawieh, A. *et al.* Systems biology, bioinformatics, and biomarkers in neuropsychiatry. *Front. Neurosci.* **6**, 187 (2012).
26. Pu, J., Liu, Y., Zhang, H., Tian, L., Gui, S., Yu, Y., Chen, X., Chen, Y., Yang, L., Ran, Y., Zhong, X., Xu, S., Song, X., Liu, L., Zheng, P., Wang, H. & Xie, P. An integrated meta-analysis of peripheral blood metabolites and biological functions in major depressive disorder. *Mol. Psychiatry* (2020).
27. Guo, L. *et al.* MK4MDD: a multi-level knowledge base and analysis platform for major depressive disorder. *PLoS ONE* **7**(10), e46335 (2012).
28. Jia, P., Kao, C. F., Kuo, P. H. & Zhao, Z. A comprehensive network and pathway analysis of candidate genes in major depressive disorder. *BMC Syst Biol.* **5**(Suppl 3), S12 (2011).
29. Fan, T., Hu, Y., Xin, J., Zhao, M. & Wang, J. Analyzing the genes and pathways related to major depressive disorder via a systems biology approach. *Brain Behav.* **10**(2), e01502 (2020).
30. Martins-de-Souza, D. *et al.* Identification of proteomic signatures associated with depression and psychotic depression in post-mortem brains from major depression patients. *Transl. Psychiatry.* **2**(3), e87 (2012).
31. Gottschalk, M. G., Wesseling, H., Guest, P. C. & Bahn, S. Proteomic enrichment analysis of psychotic and affective disorders reveals common signatures in presynaptic glutamatergic signaling and energy metabolism. *Int. J. Neuropsychopharmacol.* **18**(2), pyu019 (2014).
32. Hashimoto, K. Emerging role of glutamate in the pathophysiology of major depressive disorder. *Brain Res. Rev.* **61**(2), 105–123 (2009).
33. Nicu, M. J., Ionescu, D. F., Mathews, D. C., Richards, E. M. & Zarate, C. A. Jr. Second messenger/signal transduction pathways in major mood disorders: moving from membrane to mechanism of action, part I: major depressive disorder. *CNS Spectr.* **18**(5), 231–241 (2013).
34. Rajkowska, G. Histopathology of the prefrontal cortex in major depression: what does it tell us about dysfunctional monoaminergic circuits?. *Prog. Brain Res.* **126**, 397–412 (2000).
35. Mubeen, S. *et al.* The impact of pathway database choice on statistical enrichment analysis and predictive modeling. *Front. Genet.* **10**, 1203. <https://doi.org/10.3389/fgene.2019.01203> (2019) (Erratum in: *Front. Genet.* **2020**;11:436).
36. Freeze, H. H., Eklund, E. A., Ng, B. G. & Patterson, M. C. Neurological aspects of human glycosylation disorders. *Annu. Rev. Neurosci.* **38**, 105–125 (2015).
37. Rudd, P. M., Elliott, T., Cresswell, P., Wilson, I. A. & Dwek, R. A. Glycosylation and the immune system. *Science* **291**(5512), 2370–2376 (2001).
38. Wolfert, M. A. & Boons, G. J. Adaptive immune activation: glycosylation does matter. *Nat. Chem. Biol.* **9**(12), 776–784 (2013).
39. Yamagata, H. *et al.* Altered plasma protein glycosylation in a mouse model of depression and in patients with major depression. *J. Affect. Disord.* **233**, 79–85 (2018).
40. Park, D. I. *et al.* Blood plasma/IgG N-glycome biosignatures associated with major depressive disorder symptom severity and the antidepressant response. *Sci. Rep.* **8**(1), 179 (2018).
41. Boeck, C. *et al.* Alterations of the serum N-glycan profile in female patients with Major Depressive Disorder. *J. Affect. Disord.* **234**, 139–147 (2018).
42. Strawbridge, R., Young, A. H. & Cleare, A. J. Biomarkers for depression: recent insights, current challenges and future prospects. *Neuropsychiatr. Dis. Treat.* **13**, 1245–1262 (2017).
43. Miller, A. H. & Raison, C. L. The role of inflammation in depression: from evolutionary imperative to modern treatment target. *Nat. Rev. Immunol.* **16**(1), 22–34 (2016).
44. Anisman, H. & Merali, Z. Cytokines, stress and depressive illness: brain-immune interactions. *Ann. Med.* **35**(1), 2–11 (2003).
45. Hayley, S., Merali, Z. & Anisman, H. Stress and cytokine-elicited neuroendocrine and neurotransmitter sensitization: implications for depressive illness. *Stress.* **6**(1), 19–32 (2003).

46. Qiu, Y., Peng, Y. & Wang, J. Immunoregulatory role of neurotransmitters. *Adv. Neuroimmunol.* **6**(3), 223–231 (1996).
47. Heim, C., Newport, D. J., Mletzko, T., Miller, A. H. & Nemeroff, C. B. The link between childhood trauma and depression: insights from HPA axis studies in humans. *Psychoneuroendocrinology* **33**(6), 693–710 (2008).
48. Jeon, S. W. & Kim, Y. K. The role of neuroinflammation and neurovascular dysfunction in major depressive disorder. *J. Inflamm. Res.* **11**, 179–192 (2018).
49. Jeon, S. W. & Kim, Y. K. Neuroinflammation and cytokine abnormality in major depression: Cause or consequence in that illness? *World J. Psychiatry.* **6**(3), 283–293 (2016).
50. Bob, P. *et al.* Dissociative symptoms and neuroendocrine dysregulation in depression. *Med. Sci. Monit.* **14**(10), CR499–CR504 (2008).
51. Kuehn, H., Liberzon, A., Reich, M. & Mesirov, J. P. Using GenePattern for gene expression analysis. *Curr. Protoc. Bioinform.* Chapter 7:Unit 7.12 (2008).
52. Blankenberg, D. & Hillman-Jackson, J. Analysis of next-generation sequencing data using Galaxy. *Methods Mol. Biol.* **1150**, 21–43 (2014).
53. Gould, J., Getz, G., Monti, S., Reich, M. & Mesirov, J. P. Comparative gene marker selection suite. *Bioinformatics* **22**(15), 1924–1925 (2006).
54. Batut B, Hiltmann S, Bagnacani A, Baker D, Bhardwaj V, Blank C, Bretaudeau A, Brillet-Guéguen L, Čech M, Chilton J, Clements D, Doppelt-Azeroual O, Erxleben A, Freeberg MA, Gladman S, Hoogstrate Y, Hotz HR, Houwaart T, Jagtap P, Larivière D, Le Corguillé G, Manke T, Mareuil F, Ramirez F, Ryan D, Sigloch FC, Soranzo N, Wolff J, Videm P, Wolfien M, Wubuli A, Yusuf D; Galaxy Training Network, Taylor J, Backofen R, Nekrutenko A, Grüning B. Community-driven data analysis training for biology. *Cell Syst.* **6**(6), 752–758.e1 (2018).
55. <https://pubmed.ncbi.nlm.nih.gov/>.
56. Carvalho, A. F. *et al.* Evidence-based umbrella review of 162 peripheral biomarkers for major mental disorders. *Transl. Psychiatry.* **10**(1), 152 (2020).
57. Papatheodorou, S. Umbrella reviews: what they are and why we need them. *Eur. J. Epidemiol.* **34**, 543–546 (2019).
58. Gutiérrez-Sacristán, A. *et al.* PsyGeNET: a knowledge platform on psychiatric disorders and their genes. *Bioinformatics* **31**(18), 3075–3077 (2015).
59. Wray, N. R., Ripke, S., Mattheisen, M., Trzaskowski, M., Byrne, E. M., Abdellaoui, A. *et al.* Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat. Genet.* **50**(5), 668–681 (2018).
60. Oommen, A. M. *et al.* GlycoGAIT: a web database to browse glycogenes and lectins under gastric inflammatory diseases. *J. Theor. Biol.* **406**, 93–98 (2016).
61. Stroud, M. R. *et al.* Myeloglobyan, a series of E-selectin-binding polylectosaminolipids found in normal human leukocytes and myelocytic leukemia HL60 cells. *Biochem. Biophys. Res. Commun.* **209**(3), 777–787 (1995).
62. Babu, P. *et al.* Structural characterisation of neutrophil glycans by ultra sensitive mass spectrometric glycomics methodology. *Glycoconj. J.* **26**(8), 975–986 (2009).
63. Antonopoulos, A., North, S. J., Haslam, S. M. & Dell, A. Glycosylation of mouse and human immune cells: insights emerging from N-glycomics analyses. *Biochem. Soc. Trans.* **39**(5), 1334–1340 (2011).
64. Blixt, O. *et al.* Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. *Proc. Natl. Acad. Sci. U. S. A.* **101**(49), 17033–17038 (2004).
65. Raudvere, U. *et al.* g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.* **47**(W1), W191–W198 (2019).
66. Chong, J. & Xia, J. Using MetaboAnalyst 4.0 for metabolomics data analysis, interpretation, and integration with other omics data. *Methods Mol. Biol.* **2104**, 337–360 (2020).
67. Kamburov, A., Wierling, C., Lehrach, H. & Herwig, R. ConsensusPathDB—a database for integrating human functional interaction networks. *Nucleic Acids Res.* **37**(Database issue), D623–D628 (2009).
68. Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**(1), 27–30 (2000).
69. Wang, Y. *et al.* PubChem BioAssay: 2017 update. *Nucleic Acids Res.* **45**(D1), D955–D963 (2017).
70. Wishart, D. S. *et al.* HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Res.* **46**(D1), D608–D617 (2018).
71. Kuleshov, M. V. *et al.* Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **44**(W1), W90–W97 (2016).
72. Su, G., Morris, J. H., Demchak, B. & Bader, G. D. Biological network exploration with Cytoscape 3. *Curr. Protoc. Bioinform.* **47**, 8–13 (2014).
73. Smyth, G. K. Limma: linear models for microarray data. In *Bioinformatics and Computational Biology Solutions Using R and Bioconductor* (eds Gentleman, R. *et al.*) 397–420 (Springer, 2005).
74. Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. Voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **15**, R29 (2014).
75. Bai, M. *et al.* CD177 modulates human neutrophil migration through activation-mediated integrin and chemoreceptor regulation. *Blood* **130**(19), 2092–2100 (2017).
76. Malaguarnera, L., Marsullo, A., Zorena, K., Musumeci, G. & Di Rosa, M. Vitamin D(3) regulates LAMP3 expression in monocyte derived dendritic cells. *Cell Immunol.* **311**, 13–21 (2017).
77. Demir, S. *et al.* Neutrophil-lymphocyte ratio in patients with major depressive disorder undergoing no pharmacological therapy. *Neuropsychiatr. Dis. Treat.* **11**, 2253–2258 (2015).
78. Iqbal, A. J. *et al.* Human CD68 promoter GFP transgenic mice allow analysis of monocyte to macrophage differentiation in vivo. *Blood* **124**(15), e33–44 (2014).
79. Elsayh, K. I., Zahran, A. M., Lotfy Mohamad, I. & Aly, S. S. Dendritic cells in childhood sepsis. *J. Crit. Care.* **28**(5), 881.e7–13 (2013).
80. Banerjee, A. *et al.* Modulation of paired immunoglobulin-like type 2 receptor signaling alters the host response to *Staphylococcus aureus*-induced pneumonia. *Infect. Immun.* **78**(3), 1353–1363 (2010).
81. Kim, J. R., Horton, N. C., Mathew, S. O. & Mathew, P. A. CS1 (SLAMF7) inhibits production of proinflammatory cytokines by activated monocytes. *Inflamm. Res.* **62**(8), 765–772 (2013).
82. An, H. *et al.* Soluble LILRA3, a potential natural antiinflammatory protein, is increased in patients with rheumatoid arthritis and is tightly regulated by interleukin 10, tumor necrosis factor- α , and interferon- γ . *J. Rheumatol.* **37**(8), 1596–1606 (2010).
83. Madsen, C. D. & Sidenius, N. The interaction between urokinase receptor and vitronectin in cell adhesion and signalling. *Eur. J. Cell Biol.* **87**(8–9), 617–629 (2008).
84. Sharma, A., Steichen, A. L., Jondle, C. N., Mishra, B. B. & Sharma, J. Protective role of Mincle in bacterial pneumonia by regulation of neutrophil mediated phagocytosis and extracellular trap formation. *J. Infect. Dis.* **209**(11), 1837–1846 (2014).
85. Katakai, T., Habiro, K. & Kinashi, T. Dendritic cells regulate high-speed interstitial T cell migration in the lymph node via LFA-1/ICAM-1. *J. Immunol.* **191**(3), 1188–1199 (2013).
86. Mahmood, N., Mihalciou, C. & Rabbani, S. A. Multifaceted role of the urokinase-type plasminogen activator (uPA) and its receptor (upar): diagnostic, prognostic, and therapeutic applications. *Front. Oncol.* **8**, 24 (2018).

87. Sperandio, M., Gleissner, C. A. & Ley, K. Glycosylation in immune cell trafficking. *Immunol. Rev.* **230**(1), 97–113 (2009).
88. Koltsova, E. & Ley, K. Tyrosine sulfation of leukocyte adhesion molecules and chemokine receptors promotes atherosclerosis. *Arterioscler. Thromb. Vasc. Biol.* **29**(11), 1709–1711 (2009).
89. Ouyang, Y. B., Lane, W. S. & Moore, K. L. Tyrosylprotein sulfotransferase: purification and molecular cloning of an enzyme that catalyzes tyrosine O-sulfation, a common posttranslational modification of eukaryotic proteins. *Proc. Natl. Acad. Sci. U. S. A.* **95**(6), 2896–2901 (1998).
90. Miyamoto, K. *et al.* Chondroitin 6-O-sulfate ameliorates experimental autoimmune encephalomyelitis. *Glycobiology* **24**(5), 469–475 (2014).
91. Shao, C. *et al.* Comparative glycomics of leukocyte glycosaminoglycans. *FEBS J.* **280**(10), 2447–2461 (2013).
92. Iovu, M., Dumais, G. & du Souich, P. Anti-inflammatory activity of chondroitin sulfate. *Osteoarthr. Cartil.* **16**(Suppl 3), S14–S18 (2008).
93. Trowbridge, J. M. & Gallo, R. L. Dermatan sulfate: new functions from an old glycosaminoglycan. *Glycobiology* **12**(9), 117R–R125 (2002).
94. Belmiro, C. L. *et al.* Dermatan sulfate reduces monocyte chemoattractant protein 1 and TGF- β production, as well as macrophage recruitment and myofibroblast accumulation in mice with unilateral ureteral obstruction. *Braz. J. Med. Biol. Res.* **44**(7), 624–633 (2011).
95. Baudoin, L. & Issad, T. O-GlcNAcylation and inflammation: a vast territory to explore. *Front. Endocrinol. (Lausanne)*. **5**, 235 (2015).
96. Yang, Y. R. *et al.* Elevated O-GlcNAcylation promotes colonic inflammation and tumorigenesis by modulating NF- κ B signaling. *Oncotarget* **6**(14), 12529–12542 (2015).
97. Li, S. F. *et al.* Downregulation of β 1,4-galactosyltransferase 5 improves insulin resistance by promoting adipocyte commitment and reducing inflammation. *Cell Death Dis.* **9**(2), 196 (2018).
98. Iwabuchi, K. & Nagaoka, I. Lactosylceramide-enriched glycosphingolipid signaling domain mediates superoxide generation from human neutrophils. *Blood* **100**(4), 1454–1464 (2002).
99. Zhang, L. *et al.* The Immunological regulation roles of porcine β -1, 4 galactosyltransferase V (B4GALT5) in PRRSV infection. *Front. Cell Infect. Microbiol.* **8**, 48 (2018).
100. Okamoto, T., Tanigami, H., Suzuki, K. & Shimaoka, M. Thrombomodulin: a bifunctional modulator of inflammation and coagulation in sepsis. *Crit. Care Res. Pract.* **2012**, 614545 (2012).
101. Harris, R. L., van den Berg, C. W. & Bowen, D. J. ASGR1 and ASGR2, the genes that encode the asialoglycoprotein receptor (Ashwell receptor), are expressed in peripheral blood monocytes and show interindividual differences in transcript profile. *Mol Biol Int.* **2012**, 283974 (2012).
102. Krishnan, V. & Nestler, E. J. The molecular neurobiology of depression. *Nature* **455**(7215), 894–902 (2008).
103. Luscher, B., Shen, Q. & Sahir, N. The GABAergic deficit hypothesis of major depressive disorder. *Mol. Psychiatry*. **16**(4), 383–406 (2011).
104. Colvin, R. A., Davis, N., Nipper, R. W. & Carter, P. A. Zinc transport in the brain: routes of zinc influx and efflux in neurons. *J. Nutr.* **130**(5S Suppl), 1484S–S1487 (2000).
105. Gess, B., Röhr, D. & Young, P. Ascorbic acid and sodium-dependent vitamin C transporters in the peripheral nervous system: from basic science to clinical trials. *Antioxid. Redox Signal.* **19**(17), 2105–2114 (2013).
106. Turner, C. A., Eren-Koçak, E., Inui, E. G., Watson, S. J. & Akil, H. Dysregulated fibroblast growth factor (FGF) signaling in neurological and psychiatric disorders. *Semin. Cell Dev. Biol.* **53**, 136–143 (2016).
107. Barde, S. *et al.* Alterations in the neuropeptide galanin system in major depressive disorder involve levels of transcripts, methylation, and peptide. *Proc. Natl. Acad. Sci. U. S. A.* **113**(52), E8472–E8481 (2016).
108. Mühlhoff, M., Rollenhagen, M., Werneburg, S., Gerardy-Schahn, R. & Hildebrandt, H. Polysialic acid: versatile modification of NCAM, SynCAM 1 and neuropilin-2. *Neurochem. Res.* **38**(6), 1134–1143 (2013).
109. Yu, R. K., Tsai, Y. T., Ariga, T. & Yanagisawa, M. Structures, biosynthesis, and functions of gangliosides—an overview. *J. Oleo Sci.* **60**(10), 537–544 (2011).
110. Um, J. W. & Ko, J. Neural glycosylphosphatidylinositol-anchored proteins in synaptic specification. *Trends Cell Biol.* **27**(12), 931–945 (2017).
111. Hooper, A. W. M., Alamilla, J. F., Venier, R. E., Gillespie, D. C. & Igdoura, S. A. Neuronal pentraxin 1 depletion delays neurodegeneration and extends life in Sandhoff disease mice. *Hum. Mol. Genet.* **26**(4), 661–673 (2017).
112. Matas-Rico, E., van Veen, M. & Moolenaar, W. H. Neuronal differentiation through GPI-anchor cleavage. *Cell Cycle* **16**(5), 388–389 (2017).
113. Gennarini, G. *et al.* The role of Gpi-anchored axonal glycoproteins in neural development and neurological disorders. *Mol. Cell Neurosci.* **81**, 49–63 (2017).
114. Paulick, M. G. & Bertozzi, C. R. The glycosylphosphatidylinositol anchor: a complex membrane-anchoring structure for proteins. *Biochemistry* **47**(27), 6991–7000 (2008).
115. Mergenthaler, P., Lindauer, U., Dienel, G. A. & Meisel, A. Sugar for the brain: the role of glucose in physiological and pathological brain function. *Trends Neurosci.* **36**(10), 587–597 (2013).
116. Söderberg, M., Raffalli-Mathieu, F. & Lang, M. A. Inflammation modulates the interaction of heterogeneous nuclear ribonucleoprotein (hnRNP) I/polypyrimidine tract binding protein and hnRNP L with the 3' untranslated region of the murine inducible nitric-oxide synthase mRNA. *Mol. Pharmacol.* **62**(2), 423–431 (2002).
117. Krieg, A. M. & Vollmer, J. Toll-like receptors 7, 8, and 9: linking innate immunity to autoimmunity. *Immunol. Rev.* **220**, 251–269 (2007).
118. Oliveira-Nascimento, L., Massari, P. & Wetzler, L. M. The role of TLR2 in infection and immunity. *Front. Immunol.* **3**, 79 (2012).
119. Cribbs, D. H. *et al.* Extensive innate immune gene activation accompanies brain aging, increasing vulnerability to cognitive decline and neurodegeneration: a microarray study. *J. Neuroinflamm.* **9**, 179 (2012).
120. Vega, J. A., García-Suárez, O., Hannestad, J., Pérez-Pérez, M. & Germanà, A. Neurotrophins and the immune system. *J. Anat.* **203**(1), 1–19 (2003).
121. Dillon, S. R., Gross, J. A., Ansell, S. M. & Novak, A. J. An APRIL to remember: novel TNF ligands as therapeutic targets. *Nat. Rev. Drug Discov.* **5**(3), 235–246 (2006).
122. Wang, G. *et al.* PAK1 regulates RUFY3-mediated gastric cancer cell migration and invasion. *Cell Death Dis.* **6**, e1682 (2015).
123. Döring, Y. *et al.* Deficiency of the sialyltransferase St3Gal4 reduces Ccl5-mediated myeloid cell recruitment and arrest: short communication. *Circ. Res.* **114**(6), 976–981 (2014).
124. Sperandio, M. The expanding role of α 2-3 sialylation for leukocyte trafficking in vivo. *Ann. N. Y. Acad. Sci.* **1253**, 201–205 (2012).
125. Bengtson, P., Lundblad, A., Larson, G. & Pählsson, P. Polymorphonuclear leukocytes from individuals carrying the G329A mutation in the alpha 1,3-fucosyltransferase VII gene (FUT7) roll on E- and P-selectins. *J. Immunol.* **169**(7), 3940–3946 (2002).
126. Maccarrone, G. *et al.* Psychiatric patient stratification using biosignatures based on cerebrospinal fluid protein expression clusters. *J. Psychiatr. Res.* **47**(11), 1572–1580 (2013).
127. Chen, L., Xuan, J., Gu, J., Wang, Y., Zhang, Z., Wang, T. L. *et al.* Integrative network analysis to identify aberrant pathway networks in ovarian cancer. *Pac. Symp. Biocomput.* 31–42 (2012).
128. Carvalho, A. F. *et al.* Bias in peripheral depression biomarkers. *Psychother. Psychosom.* **85**(2), 81–90 (2016).

129. Köhler, C. A. *et al.* Peripheral cytokine and chemokine alterations in depression: a meta-analysis of 82 studies. *Acta Psychiatr. Scand.* **135**(5), 373–387 (2017).
130. Pu, J. *et al.* MENDA: a comprehensive curated resource of metabolic characterization in depression. *Brief. Bioinform.* **21**(4), 1455–1464 (2020).
131. Cheng, L. *et al.* MetSigDis: a manually curated resource for the metabolic signatures of diseases. *Brief. Bioinform.* **20**(1), 203–209 (2019).
132. Kennis, M. *et al.* Prospective biomarkers of major depressive disorder: a systematic review and meta-analysis. *Mol. Psychiatry.* **25**(2), 321–338 (2020).
133. Menezes, I. C., von Werne, B. C., Lacchini, R. & Juruena, M. F. Genetic biomarkers for differential diagnosis of major depressive disorder and bipolar disorder: a systematic and critical review. *Behav. Brain Res.* **357–358**, 29–38 (2019).
134. Dunbar, J. A. *et al.* Depression: an important comorbidity with metabolic syndrome in a general population. *Diabetes Care* **31**(12), 2368–2373 (2008).
135. Halaris, A. Inflammation, heart disease, and depression. *Curr. Psychiatry Rep.* **15**(10), 400 (2013).
136. Ziegelstein, R. C., Parakh, K., Sakhuja, A. & Bhat, U. Platelet function in patients with major depression. *Intern. Med. J.* **39**(1), 38–43 (2009).
137. Musselman, D. L. *et al.* Platelet activation and secretion in patients with major depression, thoracic aortic atherosclerosis, or renal dialysis treatment. *Depress Anxiety.* **15**(3), 91–101 (2002).
138. Ataoglu, A. & Canan, F. Mean platelet volume in patients with major depression: effect of escitalopram treatment. *J. Clin. Psychopharmacol.* **29**(4), 368–371 (2009).
139. Lopez-Vilchez, I. *et al.* Prothrombotic platelet phenotype in major depression: downregulation by antidepressant treatment. *J. Affect. Disord.* **159**, 39–45 (2014).
140. Haapakoski, R., Ebmeier, K. P., Alenius, H. & Kivimäki, M. Innate and adaptive immunity in the development of depression: an update on current knowledge and technological advances. *Prog. Neuropsychopharmacol. Biol. Psychiatry.* **66**, 63–72 (2016).
141. Hughes, M. M., Connor, T. J. & Harkin A. Stress-related immune markers in depression: implications for treatment. *Int. J. Neuropsychopharmacol.* **19**(6), pyw001 (2016).
142. Chen, D., Yu, S. P. & Wei, L. Ion channels in regulation of neuronal regenerative activities. *Transl. Stroke Res.* **5**(1), 156–162 (2014).
143. Bournazos, S., Wang, T. T. & Ravetch, J. V. The role and function of Fcγ receptors on myeloid cells. *Microbiol. Spectr.* **4**(6), <https://doi.org/10.1128/microbiolspec.MCHD-0045-2016>(2016).
144. Eyre, H. A., Lavretsky, H., Kartika, J., Qassim, A. & Baune, B. T. Modulatory effects of antidepressant classes on the innate and adaptive immune system in depression. *Pharmacopsychiatry* **49**(3), 85–96 (2016).
145. Strawbridge, R. *et al.* Inflammation and clinical response to treatment in depression: a meta-analysis. *Eur. Neuropsychopharmacol.* **25**(10), 1532–1543 (2015).
146. Freeze, H. H., Eklund, E. A., Ng, B. G. & Patterson, M. C. Neurology of inherited glycosylation disorders. *Lancet Neurol.* **11**(5), 453–466 (2012).
147. Schnaar, R. L., Gerardy-Schahn, R. & Hildebrandt, H. Sialic acids in the brain: gangliosides and polysialic acid in nervous system development, stability, disease, and regeneration. *Physiol. Rev.* **94**(2), 461–518 (2014).
148. Postigo, A. A., Marazuela, M., Sánchez-Madrid, F. & de Landázuri, M. O. B lymphocyte binding to E- and P-selectins is mediated through the de novo expression of carbohydrates on in vitro and in vivo activated human B cells. *J. Clin. Invest.* **94**(4), 1585–1596 (1994).
149. Buffone, A. J. *et al.* Silencing α1,3-fucosyltransferases in human leukocytes reveals a role for FUT9 enzyme during E-selectin-mediated cell adhesion. *J. Biol. Chem.* **288**(3), 1620–1633 (2013).
150. Nazimek, K. *et al.* The role of macrophages in anti-inflammatory activity of antidepressant drugs. *Immunobiology* **222**(6), 823–830 (2017).
151. Curzytek, K. *et al.* Inhibitory effect of antidepressant drugs on contact hypersensitivity reaction is connected with their suppressive effect on NKT and CD8(+) T cells but not on TCR delta T cells. *Int. Immunopharmacol.* **28**(2), 1091–1096 (2015).

Acknowledgements

We thank Dr. Gopalakrishna Kallapura, Leucine Rich Bio Private Limited, Bengaluru, India for his valuable guidance in using the online software tools used for gene expression analysis.

Author contributions

A.M.O. analysed the data and prepared initial manuscript draft; A.M.O., S.C. and L.J. contributed to drafting, structuring, preparation, review and final editing manuscript for submission. P.S. and B.H. contributed to review and final edition.

Funding

AMO, SC, POS, BH, LJ received funding from Irish Research Council New Horizons (REPRO/2016/72). In addition, AMO, SC, LJ received funding support from ‘CURAM’ funded under Science Foundation Ireland (SFI) and the European Regional Development Fund (Grant Number 13/RC/2073) and ‘GEMMA’ funded under European Union’s Horizon 2020 research and innovation programme, grant agreement No 825033. SC also received support from Higher Education Authority, Ireland.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-89040-7>.

Correspondence and requests for materials should be addressed to S.C. or L.J.

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

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021