ARTICLE

Molecular characterization of depression trait and state

Rammohan Shukla[1,2]✉, Dwight F. Newton[2,3], Akiko Sumitomo[2], Habil Zare[4], Robert Mccullumsmith[1,5], David A. Lewis[6], Toshifumi Tomoda[2] and Etienne Sibille[2,3,7]✉

Major depressive disorder (MDD) is a brain disorder often characterized by recurrent episode and remission phases. The molecular correlates of MDD have been investigated in case-control comparisons, but the biological alterations associated with illness trait (regardless of clinical phase) or current state (symptomatic and remitted phases) remain largely unknown, limiting targeted drug discovery. To characterize MDD trait- and state-dependent changes, in single or recurrent depressive episode or remission, we generated transcriptomic profiles of subgenual anterior cingulate cortex of postmortem subjects in first MDD episode ($n = 20$), in remission after a single episode ($n = 15$), in recurrent episode ($n = 20$), in remission after recurring episodes ($n = 15$) and control subject ($n = 20$). We analyzed the data at the gene, biological pathway, and cell-specific molecular levels, investigated putative causal events and therapeutic leads. MDD-trait was associated with genes involved in inflammation, immune activation, and reduced bioenergetics ($q < 0.05$) whereas MDD-states were associated with altered neuronal structure and reduced neurotransmission ($q < 0.05$). Cell-level deconvolution of transcriptomic data showed significant change in density of GABAergic interneurons positive for corticotropin-releasing hormone, somatostatin, or vasoactive-intestinal peptide ($p < 3 \times 10^{-3}$). A probabilistic Bayesian-network approach showed causal roles of immune-system-activation ($q < 8.67 \times 10^{-3}$), cytokine-response ($q < 4.79 \times 10^{-27}$) and oxidative-stress ($q < 2.05 \times 10^{-3}$) across MDD-phases. Gene-sets associated with these putative causal changes show inverse associations with the transcriptomic effects of dopaminergic and monoaminergic ligands. The study provides first insights into distinct cellular and molecular pathologies associated with trait- and state-MDD, on plasticity mechanisms linking the two pathologies, and on a method of drug discovery focused on putative disease-causing pathways.

Molecular Psychiatry; https://doi.org/10.1038/s41380-021-01347-z

INTRODUCTION

Major depressive disorder (MDD) is the world leading cause of years lost due to disability, with annual and lifetime prevalence of 6% and 18%, respectively [1]. Episodic phases of MDD are characterized by heterogeneous symptoms, including low mood, anhedonia, cognitive impairments, and physiological symptoms (i.e. activity and weight changes) [2]. For the majority of subjects, MDD follows a periodic trajectory of recurring depressive episodes of increasing severity, duration, and of progressive resistance to antidepressants, separated by gradually shortening and incomplete remission phases, leading to treatment-resistance and deteriorating functional fitness (Fig. 1A) [3]. This clinical trajectory suggests the presence of an underlying trait-like and/or progressive neuropathology, which might differ from that of episode or remission "states" of MDD.

The subgenual anterior cingulate cortex (sgACC) lies at the intersection of bottom-up sensory input and top-down cortical control and is involved in the integrated processing of emotions, including mood and reward [4]. Functional magnetic resonance imaging studies show that the activity of sgACC is increased in subjects with MDD, as well as in individuals with high neuroticism, fear of peer rejection, and healthy humans during experimentally induced sadness [5–7]. Successful antidepressant therapies reverse the hyperactivity of the sgACC, making it the target and suggested mediator of the therapeutic effects of deep brain stimulation [8, 9]. Magnetic resonance spectroscopy and transcranial magnetic stimulation studies suggest reduced gamma-aminobutyric acid (GABA) levels and cortical inhibition as a mechanism of sgACC dysregulation in MDD, which mostly normalize after successful treatment [10–13]. Large-scale transcriptomic studies in MDD postmortem sgACC samples demonstrate dysregulation in the cytoskeleton, rearrangement of neuronal processes, synaptic function, and presynaptic neurotransmission [14, 15] associated with GABA and glutamate receptor signaling. At the cellular level, reduced glial and increased neuronal densities were reported, associated with reduction in axon and dendrites [16, 17]. Reduced expression of markers for GABAergic interneurons targeting either the dendritic (e.g. somatostatin) or perisomatic (parvalbumin) compartments of pyramidal cells were reported, associated with reduced neurotrophic support [18–20].

To go beyond case-control studies and investigate molecular correlates of disease trait or state, we performed RNAseq-based transcriptome analysis of sgACC samples in postmortem cohort of

[1]Department of Neuroscience, University of Toledo, Toledo, OH 43614, USA. [2]Campbell Family Mental Health Research Institute of CAMH, Toronto, ON M5T 1R8, Canada. [3]Departments of Pharmacology and Toxicology, University of Toronto, Toronto, ON M5T1R8, Canada. [4]Department of Computer Science, Texas state University, San Marcos, TX 78666, USA. [5]Neurosciences Institute, ProMedica, Toledo, OH 43614, USA. [6]Department of Psychiatry, University of Pittsburgh, Pittsburgh, PA 15260, USA. [7]Departments of Psychiatry, University of Toronto, Toronto, ON M5T1R8, Canada. ✉email: Rammohan.Shukla@utoledo.edu; Etienne.Sibille@camh.ca

Received: 3 February 2021 Revised: 23 September 2021 Accepted: 4 October 2021
Published online: 22 October 2021

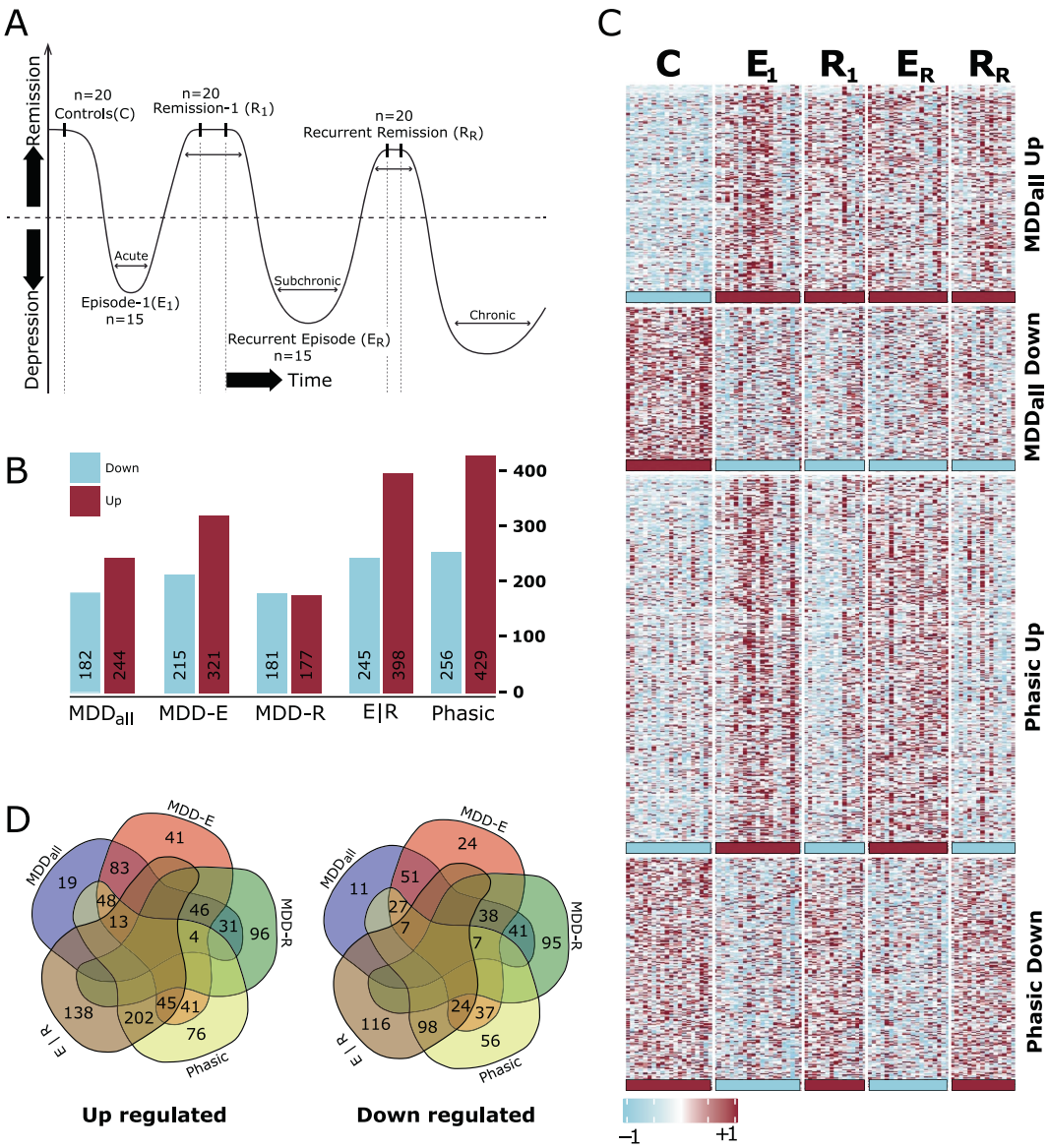

Fig. 1 RNAseq-based identification of trait and state-dependent phasic molecular changes in MDD. A Clinical evidence frequently shows recurrent episodes of MDD (valleys) of increasing severity, reduced therapeutic response and shorter remission periods (crests), **B** Numbers of differentially expressed genes associated with different contrasts ($p < 0.05$). **C** Heatmap of gene expression changes corresponding to the MDD-all (top two panels) and MDD-phasic (bottom two panels) contrasts. The colored bars at the bottom of each panel are visual summaries of gene effects within the section above. The MDD-all contrast shows consistent up or downregulation of genes across the various phases of MDD, compared to controls. The MDD-phasic contrast shows "waves" of gene changes that coincide with the episode and remission phases of MDD. Note that the "wave" pattern observed here is from cross-sectional, not longitudinal analysis. **D** Venn Diagram showing intersection of up- and downregulated genes associated with MDD-all, MDD-episode, MDD-remission, Episode/remission, and Phasic contrasts. Note the highest intersection between Episode/Remission and Phasic contrasts.

subjects in the first or recurrent phase of episodes or remission, and in healthy controls. We hypothesize that transcriptome profiling provides a broader snapshot of functional state and combined with ontological, systems biology, Bayesian network and perturbagen-induced transcriptome analyses, would enable the molecular characterization of disease trait, state, and chronicity, as well as putative causal biological changes and therapeutic targets (Supplementary Fig. 1).

MATERIAL AND METHODS

Human postmortem brain samples

Postmortem brain samples were collected during routine autopsies performed at the Allegheny County Medical Examiner's Office following procedures approved by the University of Pittsburgh Committee for Oversight of Research and Clinical Training Involving Decedents. Consensus DSM-IV diagnoses were made by an independent committee of experienced clinicians using information from structured interviews with family members, clinical records, toxicology results, and standardized psychological autopsies. A similar approach was employed to confirm the absence of a psychiatric diagnosis in comparison subjects. The MDD cohorts were carefully matched with controls to ensure that they did not differ in mean age, postmortem interval (PMI), brain pH, or RNA integrity number (RIN). Ninety samples including 20 control subjects, 20 subjects in first MDD episode (E_1), 15 in remission after a single episode (R_1), 20 in the recurrent episode (E_R), and 15 in remission after recurring episodes (R_R) (Supplementary Table 1). Samples comprising all six cortical layers were collected from coronal sections, as previously described [21]. Samples from the same cohort had been previously used in a large-scale proteomic study [22].

RNA extraction and sequencing library preparation

Total RNA was extracted from the sample homogenates using RNeasy Mini kit (Qiagen, Cat.No.74104) with in-column DNAse treatment using RNAase-Free DNase (Qiagen, Cat.No.79254). Sequencing libraries were prepared using SMARTer Stranded Total RNA-seq kit (Clontech Laboratories, Cat. No. 634876). All steps involved were performed according to the manufacturer's protocol.

Sequencing and data generation

Pooled libraries were sequenced in illumina HiSeq2500 to generate 2×100 paired end reads. On average ~41million paired end reads were generated (Supplementary Table 1, sequencing details), which were aligned to human reference genome GRCh38 provided by Ensembl using HISAT2 aligner. Count data were then generated for the reads aligned to exons and transcripts using the GenomicFeature and GenomicAlignments packages in R and gene model (GTF file) provided by Ensembl. On average, 9.7 million unique reads aligned to 21,705 genes considered for all downstream analysis, were obtained per sample.

Experimental contrasts and differential expression analysis

After removing low expressed genes (mean row sum ≤5 counts across all 90 samples), 21,684 genes were included in differential expression analysis (Supplementary Table 1). The experimental design allowed the following contrasts to be examined using the DESeq2 R package: (1) MDD-all: all MDD cohorts versus Controls, (2) MDD-episode: two MDD episode cohorts versus Controls ($E_n \mid C$), (3) MDD-remission: two remission cohorts versus controls, (4) Episode/Remission: two MDD episode cohorts versus two remission cohorts ($E_n \mid R_n$), and (5) MDD-phasic: capturing gene expression patterns coupled to the phasic changes across groups.

Analytical details of the contrasts

The curve fitting for MDD-phasic was performed in two steps in R. For the first step, the command "pData$MDD-Phasic = sin(−pi/2 + 2* pi/2* pData $group)" adds a vector named MDD-phasic to the sample phenotype data. The vector is used to model the sinusoidal nature of the cohort. Here -pi/2 and pi/2 captures the down and up phase, respectively. The constant 2 ensures that the model searches for alternate down and up pattern. "pData" denotes phenotype data table and "group" denotes the cohort. Within different group the order is preserved as $C \rightarrow E_1 \rightarrow R_1 \rightarrow E_R \rightarrow R_R$. The second step involved using the MDD-phasic column in the phenotype data as contrast (like other contrast) for finding genes associated with it. For all contrasts, we regressed out the effect of top three covariates explaining highest variability in our data (Supplementary Fig. 2, Age, Sex, and postmortem interval (PMI)). This was implemented using a Likelihood ratio test employing a full design of ~Age + Sex + PMI + Contrast and a reduced design of ~Age + Sex + PMI. A *p*-value threshold of 0.05 was used.

Pathway enrichment analysis

Biological pathways affected in different contrasts were determined using gene set enrichment analysis (GSEA) [23] (Supplementary Table 1). The 21,684 genes ranked by Wald statistic were tested against the three Gene Ontologies (GO): Biological Process (GOBP), Molecular Function (GOMF) and Cellular Component (GOCC). Updated gene set (pathway) lists were obtained from the Bader lab (http://download.baderlab.org/EM_Genesets/). To compare the effect of MDD pathology across different contrasts, the normalized enrichment score of significant pathways (*P* value < 0.05, *q* value < 0.25) was used to generate heatmaps.

A focused analysis of a priori functional themes was performed to better identify the character of the biological changes in the enrichment results. Specifically, we selected significantly enriched pathways which either (1) contained the name of the functional theme based on a text-based query or (2) were identified as child terms (nested pathways) of pre-selected parent terms which represent the theme.

Hypergeometric analysis

As a cross validation of our study, we looked for overlap between gene sets associated with different contrasts (*p* value < 0.05) and lists of differentially expressed genes provided by the authors in the following studies: Pantazatos et al. [24], Ramaker et al. [25], Scifo et al. [22], Labonte et al. [26], and Ding et al. [27]. The significance of the overlap was assessed using hypergeometric distribution implemented by Geneoverlap package in R [28]. Given two gene-lists, a significance test for their overlap can be described using hypergeometric distribution performed with a genomic background representing universe of known genes (21196 gene, default used by the package). The null-hypothesis represents an odd-ratio < 1 whereas the alternate hypothesis represents an odd-ration > 1. A similar analysis was performed to assess enrichment of different cell-type clusters in the WGCNA modules (described later in the Bayesian Network analysis).

Quantitative polymerase chain reaction (qPCR)

Differentially expressed (DE) genes belonging to a leading-edge subset (core set of transcripts that accounts for the enrichment signal [23]) in our enrichment analysis were used to validate the differential expression results using qPCR (Supplementary Fig. 3). Top five samples representing either control or diseased state were selected based on their expression profile. Total RNA (same as used for generating the sequencing libraries) were reverse transcribed to cDNA using PrimeScript RT Master Mix (TaKaRa). cDNA, primers, and TB Green Premix ExTaq (TIi RNaseH Plus) (TaKaRa) were mixed in 96-well PCR plate, and qPCR was performed in triplicate using CFX96 Real-Time System (Bio-Rad). Results were normalized to GAPDH internal control. Primers used are detailed below.

Primer name	Sequence (5′≥3′)
RHOB-F	TCGGAGCTAAGATGGTGTTAT
RHOB-R	GGTCAAGTCCTGTTCATGC≥
CRYM-F	CAAGGAGAGGTACGGGTCT
CRYM-R	CCAGGGTGACTGTGATGA
ARID5B-F	GAACCCGCACAAACCTAC
ARID5B-R	TGCCTAAGACCTGGAACTG
CNTN2-F	AACACTGCCAACCTGACC
CNTN2-R	CACCAAGACAAGCCCTTT
DDX25-F	AGGGATTGATGTGAAGCAG
DDX25-R	CTATGCGGTGGAGGTAGG
STX1A-F	ACGACCAGTGAGGAGCTG
STX1A-R	GCTTCGAGATGCTGGAGT
VEGFA-F	CCTTGCTGCTCTACCTCCAC
VEGFA-R	ATGATTCTGCCCTCCTCCTT

Cellular deconvolution analysis

To estimate the cell type proportion and identify putative disease-associated cellular differences, we adapted the deconvolution analysis described by Baron et. al. [29] and implemented in bscqsc package in R. The analysis consists of the following five steps.

A) Identifying cell-type-specific marker-genes: Using the single nucleus data set for Anterior cingulate cortex available form Allen brain atlas, we first identified cluster of cell types and markers specific to each cluster (Supplementary Table 2) using SEURAT package in R for single cell analysis [30]. In order to segregate cell clusters based on subtle differences in expression, the resolution parameter was set to 1.2

B) Building the reference basis matrix of marker-genes: This matrix contains the highly discriminatory marker-gene expression of each cell-type cluster averaged across all the cells of a given cell type cluster. For a given cluster, we considered a gene specific to a cluster only when it showed ≥3-fold difference in expression when compared to its expression in other cell clusters.

C) Estimating proportions: The resulting reference matrix was used to estimate cell proportion in different cohorts and contrast using Support Vector Regression implemented by CIBERSORT package in R [31]. For MDD-all, MDD-episode, MDD-remission, and Episode/Remission, the Wilcoxon test was used to estimate the significance of altered cell-type proportion (Supplementary Table 2). For MDD-phasic as detailed above, linear model was fitted across all cohort (with order preserved as $C \rightarrow E_1 \rightarrow R_1 \rightarrow E_R \rightarrow R_R$) and F-test was used to estimate the significance of altered cell-type proportion (Supplementary Table 2).

D) Adjusting the bulk tissue gene expression differences for density: In order to know the influence of a cell-type, we have to look for the overall changes with and without it. In other words, are the disease-related

changes affect by the cell type showing significant change in density? To estimate this possibility, we need to regress out the influence of changes in cell type ratios and compare it with the results obtained without regressing it (i.e., the results from original differential expression analysis). If the findings (in terms of significantly up and down genes) persist after regression, it indicates that the change is associated with cell-type whose ratio was regressed. Based on this principal, we statistically regress out the effect of a cell types of cluster (DTI-cluster) which showed the significant change in proportion in Episode/Remission contrast. This was implemented by expanding the design-model used for finding differential expression between episode and remission cohort by incorporating the estimated cell-type proportion (from step C) in the design matrix.

E) Finding cell-type specific differential expression: As described previously [32] and implemented in the csSAM (a package in R) we use the two separate differential expression analyses, (i.e. with and without regressing the cell proportion) to fit a linear model that estimates interaction term between cell proportion and contrast to find cell type cluster specific differential expression.

Bayesian network analysis

To find probabilistic causal associations of disease states with gene co-expression modules representing different biological themes we used weighted gene coexpression network analysis (WGCNA) [33] and Pigengene [34] package in R. First, we used consensus WGCNA to generate co-expression modules common across samples and compared them between control and disease states. For each cohort, count data normalized based on size factor (using DESeq2 package in R [35]) was used to create a matrix of pairwise Pearson correlations between genes, which was then transformed to a signed adjacency matrix using power $\beta = 12$. To calculate the interconnectedness between genes, which defines "modules", we derived the topological overlap, which gives a biologically meaningful measurement of similarity between two genes, based on their co-expression relationship with all other genes [36]. We then identified consensus modules between the control and the diseased cohorts using blockwiseConsensusModules function [37] with consensusQuantile setting set to 0.50. We obtained 22-consensus modules with numeric label 1 to 22 and color labels for all. Label 0 and color grey was assigned to genes not assigned to any module and was removed from further analysis. The identified modules were functionally characterized using all three Gene Ontologies and top pathways with Bonferroni corrected p-value < 0.05 were used to label the modules.

The first principal component (the eigengene) summarizes a given module by accounting for the maximum variability of all its constituent genes. It is used as a proxy for the overall gene expression of this module and to identify mechanisms associated with disease states.

Using the Pigengene R package [34], module eigengenes were used as random variables to train a Bayesian network that modeled the probabilistic dependencies between all modules. In order to establish the probabilistic dependencies (which in the DAG network is represented by the direction) an otherwise correlational network must be anchored with an external data (called causal anchors or genetic priors, as explained by Zhang et al. [38]. For example, in many instances SNPs data available with the transcriptomic data is used as anchor [38]. In absence of such data and to simplify our inference, we have included the experimental groups (i.e., labels of the cohort) as a categorical variable called Disease node. Specifically, the Disease node is an observed variable that can take one of the following values: Control (C), first Episode (E_1), first Remission (R_1), Recurrent Episode (E_R), and Recurrent Remission (R_R). There can be many possible directional associations between the disease node and other nodes, (which in our case are gene modules identified using WGCNA). The best direction however is selected based on average of 1000 permuted networks using all nodes. As the consensus is drawn from all disease cohorts the network essentially models MDD, regardless of trait and state.

Connectivity map (Cmap) analysis

Cmap contains expression profiles induced by ~19,000 different small molecules (perturbagens) in ~77 different cell lines [39]. Genes in module 10 were submitted as query in the Cmap API (https://clue.io/query). The output file (cs_nlx476251.gct) containing the raw weighted connectivity scores for all molecules, or perturbagens, tested in different experimental conditions and cell types was used for further analysis. A connectivity scores summarizes the similarity between the query and signature profile of a perturbagen based on Kolmogorov-Smirnov enrichment statistics [23]. To compare an observed connectivity score to all others in the database we calculate the percentile score, referred as "tau". The weighted connectivity score ranges from −1 to +1 and accordingly the tau ranged from −100 to +100. Negative and positive scores indicate dissimilarity and similarity between query and perturbagen signatures, respectively. In this study we restricted our search space to all perturbagen tested in three neuronal cell lines (NEU, NEC.KCL, NPC) available in the Cmap database [39]. As described before [39], we summarized the multiple experimental conditions under which a given perturbagen was tested in a given cell line separately for each cancer and neuronal cell lines based on the largest effect size magnitude between 33rd and 67th percentile. To select the top perturbagen candidate, we used the rank-ordered row sum of tau score [39].

Drug target identification

Using the available two-dimensional (2D) structures (http://lincsportal.ccs.miami.edu/SmallMolecules/catalog) for top drugs associated with Module M10 and M16 molecules we searched for protein target and target class (protein family to which the target belongs) using SwissTargetPrediction, a web tool based working on structure function relationship of drugs and biomolecule [40]. Note that M16 was neither downstream nor directly upstream of disease the node and thus served as reference against which the drugs associated with M10, the one directly connected to disease node, was accessed. A drug can potentially dock to many protein targets of different target class with different probabilities. We reported the target class with highest docking probabilities.

RESULTS

MDD trait and state are characterized by distinct biological changes in sgACC

RNAseq-based expression profiles from sgACC samples were obtained from one control (C) and four cohorts of MDD subjects during a first depressive episode (E_1), remission after first episode (R_1), recurrent episode (E_R) or remission after recurrent episodes (R_R) (Supplementary Table 1). Differential expression and biological pathway analyses were performed along the following group contrasts: MDD-all, comparing control samples to all other cohorts and representing depression-trait; MDD-episode, comparing control samples with those in depressive state; MDD-remission, comparing control samples with those in remission state; Episode/remission, a direct comparison of depressive and remission states, and MDD-phasic, capturing the oscillating changes between depression and remission (Fig. 1B and methods). Events captured by MDD-all were considered MDD-trait pathologies while those captured by the last four contrasts were considered as MDD-state pathologies.

Overall, upregulated genes were more common than downregulated genes across contrasts (Fig. 1B, Supplementary Table 1). The MDD-all contrast showed consistent up- or downregulation patterns of gene expression regardless of phases, compared to controls (Fig. 1C). The MDD-phasic contrast revealed robust patterns of gene expression that cross-sectionally oscillated with episodes and remission phases ("waves-pattern" in Fig. 1C). The MDD-phasic gene-set overlapped substantially with the Episode/Remission contrasts (Fig. 1D), which captured the range of the "wave-pattern". Aside from genes shared with MDD-all (Fig. 1D), minimal overlap was observed between genes identified in the MDD-episode and MDD-remission contrasts.

Upregulated pathways (p value < 0.01; FDR < 0.25) were similarly more common than downregulated pathways (Fig. 2, Supplementary Table 1). For the MDD-all contrast, multiple pathways associated with inflammation, immune-system, angiogenesis, and vascular growth factors were upregulated. Other upregulated pathways included transcription associated with stress response and protein modification processes (JNK-cascade, MAPK-activity, auto-ubiquitination; Supplementary Table 1). Few downregulated pathways were observed and associated with mitochondrial-function, bioenergetics, and neuropeptide-signaling.

Results for the MDD-episode and MDD-all contrasts were highly similar (Fig. 2, Columns 1-2) for upregulated pathways related to

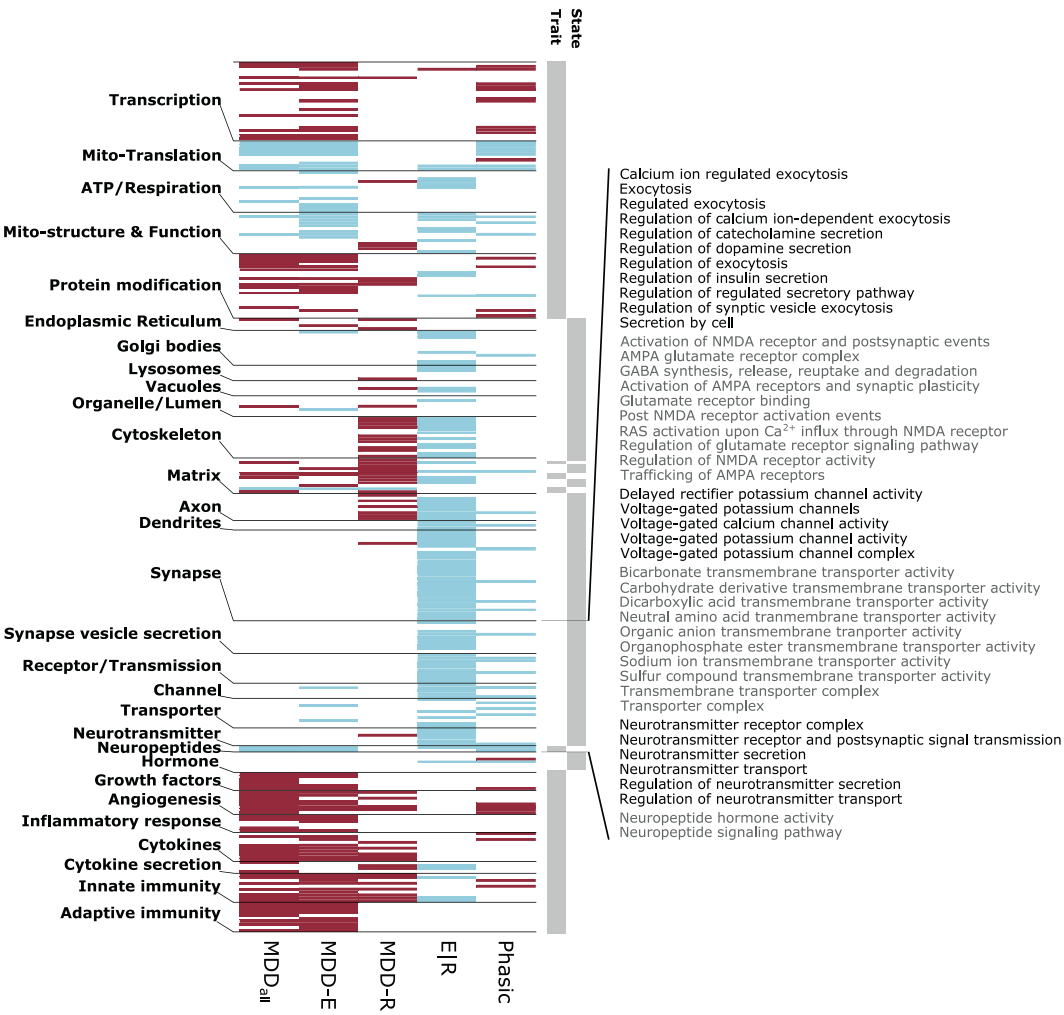

Fig. 2 Profiles of biological pathway affected in the various MDD group contrasts. Altered pathways associated with different biological themes (left) per contrast. The blue and orange colors represent up- and downregulated pathways, respectively. The vertical columns to the right indicate broad association of trait or state pathologies with biological themes. Trait and State were mostly defined by the first contrast (MDD-ALL), then by consistency across related biological pathways. For instance, See Supplementary Table 1 for details and further biological theme re-grouping.

immune-system, inflammatory-response, translation related to stress response and aspects of protein-modifications, and down-regulated pathways related to bioenergetic functions. For neurotransmission, pathways associated with Voltage-gated-potassium-channels and *transporter* activities were downregulated, in addition to neuropeptide signaling pathways.

The MDD-Remission contrast, similar to the MDD-all and MDD-episode contrasts, was associated with upregulated inflammation- and immune-related pathways, although limited to innate versus both adaptive- and innate-immunity in the MDD-episode contrast (Fig. 2, Column-3). Additional changes included multiple upregulated pathways related to cytoskeleton, axonal and extracellular matrix, suggesting a structural and cellular reorganization during remission.

Finally, the direct Episode/remission contrast (Fig. 2, column-4) was associated with downregulated (during episode) bioenergetic-related pathways (*ATP, cellular-respiration*, and *Mitochondrion-related* pathways), consistent with findings from the Episode- or Remission-contrasts. Surprisingly, we observed an additional large downregulation during episodes of multiple pathways related to the structure and function of all neuronal compartments (cytoskeleton, organelles, matrix, axon, dendrites, synapse, channel, receptor/transmission dendrite, vesicle-secretion, extracellular-matrix), multiple aspects of synaptic function, neurotransmission mediated by glutamate, GABA, dopamine, catecholamines, and multiple membrane transporter functions involving potassium- and calcium-gated voltage channels. This suggests multiple moderate state-dependent changes affecting multiple aspects of neuronal signaling that are only detectable by direct comparison between remission and episode phases and missed when traditionally comparing MDD subjects to controls. Changes matching the oscillatory nature of the disease trajectory (MDD-phasic) identified similar functional themes as in the Episode/Remission contrast (Fig. 2, Column-5), although to lower degrees.

Results were validated using three approaches. First, we compared our results to RNAseq data from three independent MDD versus control studies [24–26]. This analysis revealed significant overlap (Hypergeometric-overlap test) of differentially expressed gene lists from Pantazatos et al. with MDD-episode (Up: $p = 0.05$), MDD-phasic (Up: $p = 0.046$) and trend-level overlap of MDD-all (Up: $p = 0.062$); Ramaker et al. with MDD-episode (Up: $p = 0.020$), Episode/remission (Up: $p = 0.002$, Down: $p = 0.042$), MDD-phasic (Up: $p = 0.00014$) and Labonte et al. with Episode/remission (BA25: Up: $p = 4.78^{-06}$, Down: $p = 0.00033$), MDD-phasic (Up: $p = 8.70^{-04}$, Down: $p = 0.035$). Second, comparison to mass

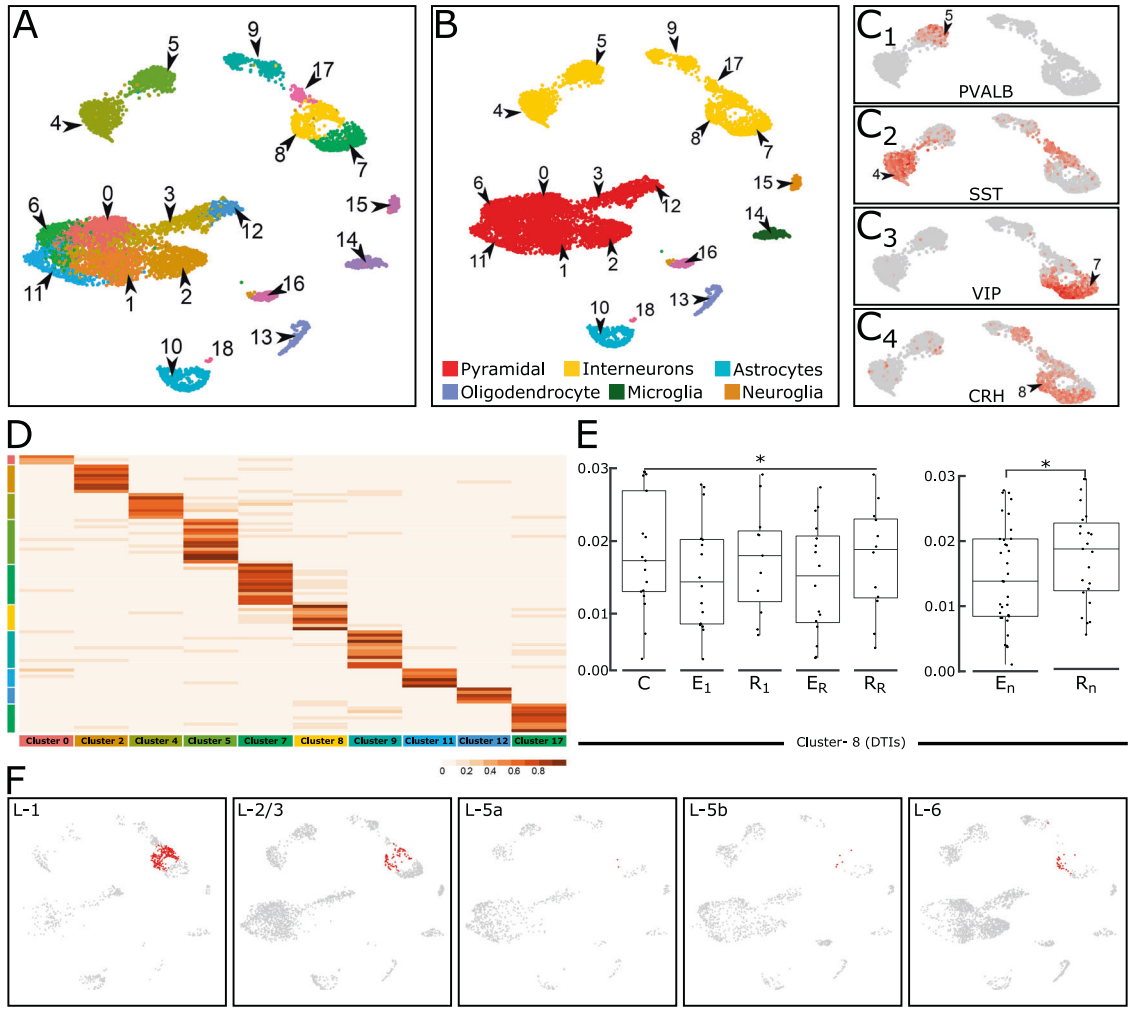

Fig. 3 Cell-type deconvolution of gray matter RNAseq reveals gene expression changes in synchrony with MDD phases for *CRH*-, *SST*- and *VIP*-expressing GABAergic interneurons. A Clustering of human anterior cingulate cortex single nuclei RNAseq data from Allen Brain Atlas. **B** Clusters were characterized globally using known markers of pyramidal neurons (*SLC17A7*), interneurons (*GAD1* and *GAD2*), astrocytes (*GFAP*), oligodendrocytes (*OPALIN*), microglia (*CX3CR1*) and neuroglia (*PDGFRA*). **C** Clusters were characterizing locally using the top enriched expressed genes (upregulated) by cluster. C_1) Cluster 5 enriched with *PVALB*. C_2) Cluster 4 enriched with *SST*. C_3) Cluster 7 enriched with VIP. C_4) Cluster 8 enriched with *CRH*. Note that some clusters have overlapping expression of different markers. For instance, cluster 8 (*CRH*) is also enriched in *SST* and *VIP*. **D** Heatmap of highly discriminative gene, i.e., markers used as reference for estimating relative cell densities in (**E**). Rows represent genes and columns represent the cluster identified in (**A**). Only representative clusters are shown here. **E** Density differences of cluster 8 cell types in Episode/remission contrast (right) and phasic contrast (left). The *y*-axis shows the relative densities of cluster 8 cell types while the x-axis represents different cohort used in the study, matching the expected 1–2% cell-type representation for the associated markers. Wilcoxon test and *F*-test were used to estimate the significance for Episode/remission and phasic contrast, respectively. *<0.05. Please see Supplementary Table 2 for changes in cellular densities across all cell-type cluster F) Layer specific distribution of cluster 8 neurons (yellow in **A**, red in **F**). Note the high enrichment of cluster 8 neurons in layers 1 and 2/3.

spectrometry-based proteomics results in the same cohort [22] showed significant correlation ($n = 3000$ genes/proteins; $R = 0.34$; $p = 9.34 \times 10^{-94}$) and overlaps between protein and gene changes for the MDD-episode ($p = 0.02$), Episode/Remission ($p = 0.043$), and trend-level overlap for the MDD-remission ($p = 0.062$) contrasts. Third, technical validation by qPCR showed highly concordant results for leading-edge genes (core set of differentially expressed transcripts that accounts for the enrichment signal) in key affected pathways within different contrasts ($p < 2.2 \times 10^{-16}$; Supplementary Fig. 3A, B).

Together, these results uncovered two parallel pathological entities: a depressive trait pathology associated with inflammation, immune-system activation and decreased bioenergetics, and a depressive-state pathology matching clinical phases and associated with reduced cell structure and neurotransmission during episodes.

RNA-seq data deconvolution identified *SST*-, *VIP*- and *CRH*-expressing interneurons as affected by phasic state-dependent changes in MDD

The present data were derived from gray matter samples where relative cell densities and associated biological signals are masked. To address this limitation, we used layer-specific single nucleus RNAseq data from ACC (from Allen-Brain-Atlas) and statistical methods to deconvolute it. We identified 19 cell-type clusters (Fig. 3A), characterized in two stages: First, globally, using pyramidal-neurons, interneurons, astrocytes, oligodendrocytes, microglia, and neuroglia markers (Fig. 3B), and second, locally, based on the genes most enriched in each cluster (Fig. 3C), for instance, the GABAergic interneuron markers *PVALB*, *SST*, *VIP* and *CRH* as enriched genes in clusters 5, 4, 7 and 8 respectively. Next, we identified discriminatory gene-sets for these 19 cell-type clusters (Fig. 3D, Supplementary Table 2) and used them to

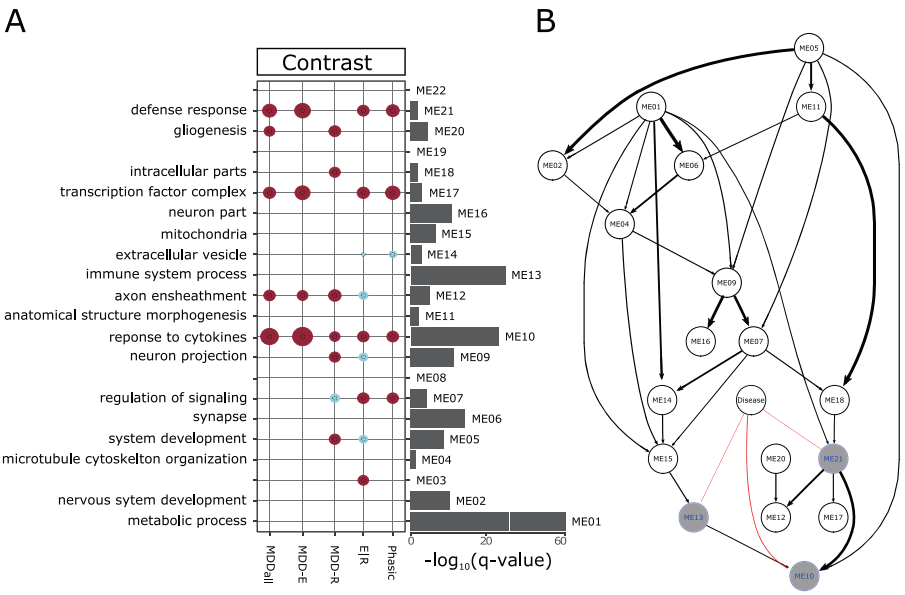

Fig. 4 Prioritizing putative causal gene modules in MDD using Bayesian network. A Consensus gene modules were identified using the combined RNAseq data across diseases phases and characterized using GO enrichment. GO terms highly enriched for each module are shown on the left, and −log10(*p*-values) (FDR < 0.05) of enrichment tests for the GO term are shown on the right. The middle panels show the enrichment of modules in the MDD contrasts (From Fig. 1B). The blue and yellow circles show enrichment in up- and downregulated genes respectively for each module. The size of circles is proportional to the −log10 of *p* values associated the enrichment. The black dotted circles represent the scale = −log10(*p*-value = 0.05). Note that module ME03, ME08, ME19 and ME22 were not associated with any GO terms and were removed from further analysis. **B** Directed acyclic graph (DAG) obtained after fitting Bayesian network to the eigengenes of the characterized consensus modules, which are shown as nodes. The edges represent the dependencies between nodes from early to later (referred as parent to child). The thickness of the edges is proportional to the number of times, in percentage, the edge was detected in 1000 permutations used to generate the Bayesian-network. The connections of the "Disease" node to its child nodes are shown in red.

estimate the relative cell-type densities in the cohort datasets [31]. Results show that cluster-8, which co-expressed *SST*, *VIP* and *CRH* genes usually expressed in dendrite targeting interneurons (henceforth: DTI-cluster) displayed lower marker gene expression in episode compared to remission phases ($p < 0.042$, Fig. 3E, right panel) and that these changes followed a phasic trajectory matching the illness states ($p < 0.016$, Fig. 3E, left panel). DTI-cluster also corresponds to the "neuropeptide-signaling" gene group previously identified in the MDD-All, MDD-episode and MDD-Phasic contrasts (Fig. 2). An analysis of layer-specificity further showed enrichment of DTI-cluster genes in cortical layers 1 and 2/3 (Fig. 3F). Validating our results, we observed significant overlap (HGA, $q < 0.03$) of DTI-cluster with a list of 566 genes obtained from meta-analysis of MDD across corticolimbic regions, including the sgACC [27].

Next, we investigated the putative impact of altered DTI-cluster cell types on overall biological changes. Focusing on Episode/remission (Fig. 3E, right), the most significant contrast, we compared gene expression differences with and without regressing out the variability associated with the cell type density differences of the DTI-cluster. In total, 721 genes showed lower *p* values after regression (Supplementary Table 2) and were associated with the following biological pathways ($q < 0.05$) downregulated in episodes: voltage-gated-potassium-channel-complex, postsynaptic-membrane, asymmetric-synapse, axon-ensheathment, and regulation-of-cell-morphogenesis (Supplementary Table 2). Finally, to assess DTI-cluster-specific differential expression, we included the interaction between DTI-cluster cell densities and the episode and remission variables in a statistical linear model [41] and identified 75 upregulated and 175 downregulated (Supplementary Table 2, *p* value < 0.01) genes (in episode) corresponding to upregulated chemokine-binding
($q < 0.04$) and downregulated *nf-KappaB* ($q < 0.01$) pathways.

Together, inferring cell type origin of gene expression revealed state-dependent phasic changes affecting upper layers *SST*-, *VIP*- and *CRH*-expressing interneurons, synchronized with reduced neuronal signaling and related to inflammation and immune-related events.

Investigating causality using Bayesian gene network analysis

While group contrasts and cell-based analyses provided biological and cellular insight, they did not inform on causal links. For this, we summarized the expression profiles across all five cohorts into modules of co-expressed genes using consensus weighted gene co-expression network analysis [33] and used the module eigengenes to construct a Bayesian-network modeling probabilistic dependency between modules [34, 42]. Twenty-two consensus modules were identified, containing 31 to 4494 genes. Eighteen modules were enriched for at least one gene-ontology functional category ($q < 0.05$; Fig. 4A, Supplementary Table 3) and used to create the DAG (Fig. 4B). The direction of edges and probabilistic dependencies were first inspected for biological and theoretical consistency. For instance, module M05, the source of all modules, is enriched in pathways associated with system-development and connected to module M02 and M011, enriched in pathways associated with nervous-system-development and anatomical-structure-morphogenesis, respectively. Similarly, the relationship between the immune-system and cytokines [43] was captured by the high probability parent-child association between module M21, enriched with immune-response, and module M10, enriched with cytokine response pathways.

A "disease-node" capturing the disease phases across cohorts (See methods) was also included in the DAG. This node was linked to three DAG modules: module M13, enriched with innate-immune-response, module M10, enriched with cytokine-response, and response-to-stress, and module M21, associated with defense–response and stress–response mostly associated with oxidative stress, as suggested by the child module M17, enriched with response-to-oxidative-stress and cellular-response-to-stress (Supplementary Table 3). M10 had the highest probabilistic

association to the disease module and is a child to disease-associated M13 and M21, making it the endpoint of the graph and the most disease-associated module. This is supported by the enrichment of M10 with upregulated pathways in all seven contrasts (Fig. 4A). M10 was enriched in astrocyte (cluster-10, $p < 0.02$) and microglia (cluster-14, $p < 5.3 \times 10^{-9}$) markers, M13 in *SST*-expressing interneuron (cluster-04, $p < 0.03$) and microglia (cluster-14, $p < 5.55 \times 10^{-127}$) markers, whereas M21 was not enriched in cell markers. As an independent validation, we looked for hypergeometric-overlap with cell-specific gene-sets from a previous study [44]. Module M10 was enriched in microglia ($p < 0.01$), module M13 in microglia ($p < 2.23 \times 10^{-27}$) and module M21 in endothelial-cells ($p < 0.03$). Overlap of module with lamina-specific gene-sets [45] suggest the enrichment of all three modules in layer-1 (M10: $p < 1.2 \times 10^{-14}$; M13: $p < 2.29 \times 10^{-27}$; M21: $p < 0.03$), which coincides with the significant enrichment of disease associated DTI-cluster cells in superficial layer-1 (Fig. 3F).

Together, the Bayesian-network organized gene-expression from the control and MDD cohorts into a coherent and directional graph. The direct association of the MDD module with three DAG modules suggests that immune and cellular stress functions are proximal and potentially causal to the trajectory and clinical manifestation of the illness, and cell-enrichment analyses implicate superficial layer inhibitory-neurons, glial, and immune-cells as potential mediators.

The MDD gene module identified in the Bayesian network is associated with known antidepressant-related pathways and novel drug targets

Assuming that the M10 MDD-associated module may have a causal role in the illness, we looked for drugs that mimicked or antagonized its gene-expression profile. For this, we probed the M10 associated gene-set against the connectivity-map, a database of transcriptomic responses to multiple molecules in diverse cell types.

A total of 33 molecules were identified as antagonizing (Fig. 5) and 50 molecules as mimicking (Fig. 5) the M10 gene-expression profiles in neuronal cell lines, potentially reflecting therapeutic or MDD-inducing effects, respectively [46]. Among the M10-antagonizing molecules, three were either dopamine, its precursor (glutamyl-dopamine) or its structural variant (6-nitrodopamine), 15 (45.45%) had G-coupled protein receptors (GPCRs) as their target class, of which five were dopaminergic, four were serotonergic, and two targeted the serotonin and/or dopamine transporters. The other major class of drug target includes enzyme (6/33), electrochemical transporter (4/33) and nuclear receptors (3/33). Notably, seven of the identified drugs are either clinically used to treat neuropsychiatric disorders or identified as a potential drug in human postmortem studies (i.e. spermidine) [47].

Among the MDD-mimicking molecules (Fig. 5), 12 (24%) were of GPCRs target class, of which 6 were serotonergic and three were dopaminergic receptors (mostly with opposing effects of compounds antagonizing the M10 expression profile), and one each for opioid, melatonin, and adrenergic receptors. The other major classes were epigenetic regulators (9/50, i.e., eraser, HDAC inhibitors) and protease (5/50).

In contrast, using the M16 gene module, which is part of the DAG, but not enriched in any MDD contrasts (Fig. 4A), only seven and nine molecules were identified as having antagonizing or mimicking effects and the target class associated with antagonizing molecules belonged to epigenetic regulators (3/7), nuclear receptor (2/7), kinase (1/7) and voltage-gated channel (1/7).

DISCUSSION

The waxing-and-waning nature of MDD (Fig. 1A) across depression and remission states is well studied in sgACC; but the molecular correlates of MDD in this brain region are limited to case-control setting, and molecular changes between and across states are largely unknown. Here, using transcriptome profiles from human postmortem sgACC subjects in depressive or remission phases of MDD, and from control subjects, we uncovered two parallel pathological entities (Fig. 1B): First, a stable MDD-state pathology associated with inflammation, immune-activation, and reduced bioenergetics, and second, a dynamic state-dependent pathology (matching the clinical phases) associated with neuronal structure and function, affecting both fast (GABA, glutamate) and slow (catecholamines, neuropeptides) neurotransmitters, sets of changes that are downregulated during episodes and not observed or reversed during remission phases. Further in-silico analyses provided perspective on cellular specificity and putative causal processes: First, a cellular deconvolution analysis associated *CRH*-, *SST*- and *VIP*-expressing dendrite targeting interneurons with phasic MDD changes; Second, a probabilistic Bayesian-network analysis suggested causal roles for immune-system activation, cytokine-response, and oxidative-stress across MDD phases. In support of this, the putative causal pathways were antagonized by drugs associated with clinical response using a drug-induced transcriptome database; thus, providing a new approach for novel therapeutic-target discovery.

Dissociating trait and state pathologies of MDD

The MDD-trait and -state pathologies (Fig. 2) correspond to distinct biological pathways. The MDD-trait pathology shows robust links to inflammation and immune activation consistent with prior reports [48]. These events often recruit extensive signal transduction pathway protein modifications, such as proteins phosphorylation in the MAPK pathway [49], and increased demand on local blood supply, mediated by increased angiogenesis, which are both also associated here with MDD-trait pathology. These findings are consistent with literature on MDD for inflammation [50], immune-system activation [51], MAPK-activation [49], but less so for angiogenesis [52].

"Trait" was defined based on statistical association with the MDD-All contrast, but some of the identified trait phenotypes displayed various degrees of significance in the separate contrasts. For instance, pathways related to innate immunity were upregulated in all three contrasts, but changes in adaptive immunity and inflammatory response were limited to the MDD-All and MDD-episode, suggesting state-dependent immune features. Reduced bioenergetics was identified as trait pathology, represented by decreased expression of genes implicated in mitochondrial structure and function, and in cellular energy production. This was observed in the MDD-All and MDD-Episode contrasts, but not the MDD-remission, suggesting a partial reversal during remission, as supported by significant results in the Episode/remission and phasic contrast analyses. Reduced bioenergetics is consistent with prior research in MDD [53] and associated with inflammation and immune system activation [54, 55], hence linking the two MDD trait-related main findings from this study.

The MDD-state pathology, as defined by all but MDD-all contrast, was associated with the classical neurotransmission changes reported in depression [56], negatively affecting excitatory glutamatergic and inhibitory GABAergic systems, voltage-gated channel activities, and structural elements of neurotransmission (axons, dendrites, vesicles), consistent with studies showing reduced pyramidal neuron dendrites and sizes [57, 58]. The fact that these results are only observed in Episode/Remission contrast suggests moderate state-dependent changes affecting neuronal signaling that are more difficult to identify in case-control studies. In contrast, neuronal structural components (cytoskeleton, matrix, axons) were restored during remission phases (i.e., significant in the MDD-remission contrast). These results show that the often-reported neurotransmission related

Molecules	NEU	NEU.KCL	NPC	Target	Target Class	Mode of action; Notes
chlorpromazine				DRD2/DRD4/HTR2A	Family A GPCR	antagonist; Antipsychotic
YM-09151-2				DRD2/DRD4/HTR2A/ADRA1A	Family A GPCR	antagonist; Atypical antypsychotic; *Same as Nemonapride*
cycloheximide				FKBP1A	Isomerase	Blocks translation
clonidine				ADRA2A/ADRA2C/ADRA2B	Family A GPCR	Agonist; Used for blood pressure and in ADHD
valproxan				CA2/CA1 (GABA/Glu/Na ch)	Lyase	Used in epilepsy and BPD
nemonapride				DRD2/DRD3/DRD4/HTR2A/ADRA1A	Family A GPCR	Antagonist; Atypical antipsychotic
EHPG-piperazine				ADRA2A/ADRA2B/ADRA2C	Family A GPCR	Affects adrenergic signalling
zonisamide				CA2 (Ca ch./GABA/Glu)	Lyase	Sulfonamide; Antagonist (Used in epilepsy, BPD, PD)
BRD-A09053961				FYN/EGFR/LCK	Kinase	_
allantoxanamide				ADORA1/ADORA2A	Family A GPCR	Uricase inhibitor
BRD-K99554241				MTNR1A/MTNR1B	Family A GPCR	Melatonin Rec antagonist
BRD-K97591839				SLC6A3 (DAT)	E-chem transporter	DAT inhibitor
enevalproate				FNTA/FNTB/SRD5A2	Enzyme/Enzyme/Oxidoreductase	Valproate metabolite; Used in epilepsy and BPD
BRD-K72015216				SLC6A3 (DAT)/SLC6A4 (SERT)	E-chem transporter	Blocks DAT and SERT
BRD-K27237442				DRD1	Family A GPCR	5-HT2A/2C, D1 antagonists and 5-HT1A agonist
lithocholic-acid				VDR/NR1H4/GPBAR1/AKR1B10	Nuclear receptor/Family A GPCR/Enzyme	Bile acid; Fat solubilization for absorption
troglitazone				PPARG/HTR2B/TBXAS1/MAOB	Nuc Rec/Family A GPCR/Cyt P450/Oxidoreductase	PPARG agonist; Use in Type 2 Diabetes
tetrabenazine				SLC18A2 (VMAT2)/ADRA1A	E-chem transporter/Family A GPCR	Monoamine vesicle storage inhibitor
clofibric-acid				PPARA	Nuclear receptor	PPARalpha agonist
FK-866				NAMPT	Enzyme	inhibitor of nicotinamide phosphoribosyltransferase
dopamine				KDM4E/DRD5/MTNR1A/SLC6A2	Eraser/Family A GPCR/E-chem transporter	Multiple actions; DA/NE agonist
BRD-K84252391				RCOR1/KDM1A/SIGMAR1/DRD2	Eraser/Membrane receptor/Family A GPCR	Multiple actions; Binds Sigma, DA and 5TH receptors
ellagic-acid				ERBB2/AKR1B1/CCND1	Kinase/Enzyme/Kinase	Free-radical inhibiting antioxidant
pyridine-2,4-dicarboxylic-acid				KDM4E/EGLN2/FTO/EGLN1	Eraser/Enzyme/Oxidoreductase	hypoxia-inducible transcription factors (HIFs) inhibitor
glutamyldopamine				SELL	Adhesion	dopamine precursor
BRD-K23628492				BCHE	Hydrolase	non-specific choline esterase; Similar to AchE
beta-alanine				KDM4E	Eraser	Epigenetic regulator, increase GABA levels
mevalonic-acid				PGD/MVD	Enzyme	related to statin effects
spermidine				CA2	Lyase	regulate membrane potential, NOS inhibition
BRD-K98336812				TRPV1	Voltage-gated ion channel	Chlorpromazine metabolite
BRD-A27924917				GABBR2/GABBR1	Family C GPCR	GABA-B Receptor Antagonists
rosiglitazone				TBXAS1/MAOB/CA2/AGTR1	Cyt P450/Oxidoreductase/Lyase/Family A GPCR	Antihyperglycemic and anti-inflammatory activities
6-nitrodopamine				SNCA	Unclassified	Nytrosylated DA
tacedinaline				HDAC3/ HDAC2/ HDAC1	Eraser	HDAC inhibitor; Epigenetic regulator
BG-1010				HTR2B/ HTR2A	Family A GPCR	_
BML-210				HDAC3/ HDAC2/ HDAC1	Eraser	HDAC inhibitor; Epigenetic regulator
BRD-A70541707				SLC6A3	E-chem transporter	_
BG-1024				ABL1/ HDAC/ Mapk14	Kinase/ Nuclear receptor/ kinase	_
oxyphenbutazone				PTGS1	Oxidoreductase	blocks prostaglandin synthesis
BRD-K99636700				_	_	_
NSC-3852				METAP2	Protease	HDAC inhibitor; Epigenetic regulator
BRD-K31344914				OPRM1	Family A GPCR	_
BG-1032				ATR	Kinase	_
geldanamycin				TRAP1/ HSP90AA1/ HSP90AB1	Other cytosolic protein	inhibits cytosolic chaperone (HSP90)
NO-ASA				PDE3A	Phosphodiesterase	Platelet aggregation Inhibitors
CCT036477				VDR	Nuclear receptor	Targets Wnt pathway
BRD-K39096267				MMP2	Protease	_
BG-1001				KCNMA1	Voltage-gated ion channel	_
BRD-K75971499				KCNH2	Voltage-gated ion channel	_
HDAC3-selective				_	_	_
amiodarone				CHRM4/ERBB2/HTR2B/FYN/THRA	GPCR/Nuclear receptor/kinase/E-chem transporter	HERG blocker
wortmannin				PRKDC/ PIK3CA/ PIK3R1/ MYLK	Kinase/ Enzyme	immunosuppressants, inhibits DNA synthesis
fenretinide				RBP4	Secreted protein	activates retinoic acid receptors inducing apoptosis
trichostatin-a				SF3B3/ HDAC3/ NCOR2	Eraser	HDAC inhibitor; Epigenetic regulator
KCl				_	_	maintains intracellular tonicity
LY-294002				PDE5A/PIM1	Phosphodiesterase/Kinase	prevent the normal catalytic reaction
MG-132				CAPN2/CTSK/PSMB2/NFKBIA	Protease	inhibits cysteine endopeptidases
BRD-A11009626				DRD2/DRD1	Family A GPCR	_
BRD-K88575585				MTNR1A	Family A GPCR	appetite suppressants
BL003-049-HDAC3				HDAC1	Eraser	HDAC inhibitor; Epigenetic regulator
BRD-K59915259				HTR7	Family A GPCR	_
etoposide				MMP2	Protease	inhibits DNA synthesis
plumbagin				EP300	Writer	augments immune response
vorinostat				HDAC3	Eraser	HDAC inhibitor; Epigenetic regulator
BG-1009				HDAC1	Eraser	HDAC inhibitor; Epigenetic regulator
p-azido-PE-TFMPP				HTR1A	Family A GPCR	_
pyrvinium				AR	Nuclear receptor	anthelmintic agent which acts to kill pinworms
SKF-83959				DRD1/DRD2	Family A GPCR	acts as an agonist at the D1-D2 dopamine receptor
ionomycin				CES2	Enzyme	increases calcium ion permeability of cell membrane
niguldipine				ADRA2C	Family A GPCR	inhibits calcium influx through cellular membranes
BG-1016				HDAC1/HDAC2/HDAC3	Eraser	HDAC inhibitor; Epigenetic regulator
BRD-A78236793				DRD2	Family A GPCR	_
spiperone				DRD2	Family A GPCR	blocks dopamine action
mevastatin				HMGCR	Oxidoreductase	inhibit the growth of bacteria
JS-K				FNTA/FNTB	Enzyme	a nitric oxide donor, induces autophagy
Ac-Leu-Leu-Nle-CHO				CAPN1/CAPNS1/CTSB	Protease	inhibits cysteine endopeptidases
BRD-K61894884				_	_	_
BRD-K47535255				CNR1/CNR2	Family A GPCR	
terreic-acid-(-)				CYP19A1	Cyt P450	inactivates MurA by covalently attaching to Cys115
BRD-K32307229				HDAC2/HDAC6/HDAC8	Eraser	HDAC inhibitor; Epigenetic regulator
BRD-K06209536				IDO1	Enzyme	Aminophenol urease inhibitors
rotenone				MT-ND4/HTR6/CYP2C19	Oxidoreductase/Family A GPCR/Cyt P450	mitochondrial NADH:ubiquinone reductase inhibitor
suramin				HDAC3	Eraser	Epigenetic regulator

Fig. 5 Molecules antagonizing or mimicking the MDD-related expression profile. Connectivity map (Cmap)-based identification of molecules antagonizing or mimicking the MDD-related DAG module 10 gene expression profile: Dissimilarity (representing antagonizing effect, top panel), and similarity (representing disease mimicking effect, bottom panel) between drug-induced transcriptomic profiles derived from neuronal cell lines and module 10 gene-sets, estimated by tau score (see methods), are shown. Significant Tau values (<−90 for dissimilarity; >90 for similarity) are shown in increasing shades of blue and orange respectively. Molecule names in red are associated with either monoaminergic or catecholaminergic systems while those in green with other known neurotransmitters. The target and target class for each molecule was predicted using SwissTargetPrediction (see methods). Mode of action and notes taken from PubChem.

pathology of MDD is state-dependent and mostly resolved during remission of MDD, and that remission is further associated with restructuring of intracellular elements and extracellular cell-to-cell contacts.

CRH-, *SST*- and *VIP*-expressing GABAergic cellular changes across MDD episodes and remission

The cell-type deconvolution approach showed that *CRH*-, *SST*- and *VIP*- expressing interneurons are dynamically affected in synchrony with the episodes and remission phases of MDD. Reduced *CRH* expression, consistent with dysregulated hypothalamus pituitary adrenal stress axis in MDD, was previously reported in subcortical regions [27], as well as in cortical layers [27]. Reduced *SST* [20, 59] and *VIP* [60–62] expression were reported in several independent cohorts. Importantly, *VIP*, *CRH* and *SST* are markers of GABAergic interneurons with overlapping patterns of expression, and that directly or indirectly inhibit pyramidal-cell dendrites [63]. These results are therefore consistent with a state-dependent MDD-related pathology that affects the regulation of excitatory input onto pyramidal cells, suggesting altered information processing in MDD [64]. Reduced dendritic inhibition may also contribute to the elevated sgACC activity during MDD-episodes.

The cellular deconvolution suggests variable densities of cell types between MDD episodes and remissions (Fig. 3D, E), which is questionable for non-dividing neurons. Deconvolution analysis assumes that cell-type gene expression contributes linearly to its fraction in bulk tissue [65]. Gene expression however, depends on other parameters, such as cell size and functional status [66]. For instance, we showed lower *SST* expression per cell in the sgACC of MDD subjects, rather than fewer SST-positive cells [67]. The association of change in expression per cell for *CRH*-, *SST*- and *VIP*-expressing interneurons is also supported by analyses showing that normalizing for putative cell-type differences enhanced the MDD-related significance of cell-to-cell signaling and neurotransmission, and that multiple pathways implicated in cell structure (hence volume and function) were affected in phase with MDD-episodes and remission.

A putative causal role for immune function and inflammation in upper cortical layer cells in MDD

Bayesian-network analysis uses changes in association probabilities between gene modules across multiple network iterations to deduct potential causal flow [34, 38, 68]. Here, this analysis shows that activation of immune function, inflammation, and oxidative stress originating from or affecting inhibitory-neurons, glial, endothelial and immune cells in upper cortical layers may have causal roles across episode and remission phases. This suggests the presence of sustained immune activation and other stressors (oxidative-stress, inflammation), which affect *CRH*-, *VIP*- and *SST*-expression neurons in upper cortical layers in synchrony with episodes and remission, consistent with phasic changes in neuronal structure and bioenergetics, and with the shift from innate and adaptive immune activation during episodes to innate immunity only during remission. In this way, the trait pathology associated with immune/inflammation may increase the biological vulnerability of the sgACC, setting it up for relapse into episode state-like pathology, demonstrating a novel type of plasticity associated with MDD.

Inferring causality needs interventional data [69]. In lack of such data in postmortem studies, we used Bayesian network to infer causality and validated the results using drug-induced transcriptomic perturbation data. It is intriguing that the gene expression profile of the Bayesian network module most associated with MDD could be antagonized or mimicked in cell-lines by drugs targeting the dopamine or monoamine systems, consistent with their roles in MDD and therapeutics. Surprisingly, this was obtained from acute drug exposure in cell-based systems lacking the complexity of the brain or diseases timeframe. This suggests that complex biological events across time and cell-systems in MDD may recruit a reduced set of cellular processes, as identified by conserved gene coexpression modules across states. This is consistent with the suggested therapeutic or pro-disease targets (Fig. 5), which belong to biological regulators (epigenetic, nuclear-receptor, and protein-modification), in addition to single target molecules.

Together the integrated approach used in this study suggests a temporal sequence assembling known hypotheses of MDD, namely immune- and cellular stress-related phenomena as causal upstream factors in series of state-dependent plasticity events affecting excitatory and inhibitory neurotransmission, as well as serotonergic, glutamatergic, and GABAergic dysfunctions.

Limitations

First, the heterogeneity of MDD was not investigated at symptoms or variation in disease trajectory level but rather focused on different contrasts at the group level. Second, besides the fact that sex plays an important role in MDD [26], the data were not stratified based on sex as this would have greatly reduced the power of the analysis by reducing the sample sizes of all five experimental groups. Third, as remission is predominantly associated with medication, the effect of medication was not removed from the analysis. Additionally, as most subjects were on multiple drug class (anticonvulsant, antidepressants, antipsychotics, and unknowns), it was not possible to untangle their effect from the data. Fourth, due to lack of additional samples, independent validation of the state-dependent and cell-specific findings were not performed. Instead, we used available genomic datasets to provide supporting evidence. Fifth, to avoid overfitting, the differential expression analysis considered only those variables which accounted for >5% of total variability in the data. Hence, amidst these limitations, additional relevant factors and patterns of changes may still be uncovered. Sixth, as some findings were previously reported (i.e., immune function, DTI neurons), one can question the novelty of the present study. However, reproducing results in independent cohorts, as in the present study, is an important internal validation mandatory for postmortem studies. Finally, the functional predictions and suggested causal links obtained through bioinformatics analyses should be interpreted as hypothesis-generating.

DATA AVAILABILITY

All datasets analyzed during the current study are available as supplementary tables. Raw data (count matrix, fastq.gz, or.bam) are available from the corresponding author on reasonable request.

REFERENCES

1. Malhi GS, Mann JJ. Depression. Lancet (Lond, Engl). 2018;392:2299–312.
2. Otte C, Gold SM, Penninx BW, Pariante CM, Etkin A, Fava M, et al. Major depressive disorder. Nat Rev Dis Prim. 2016;2:16065.
3. Sibille E, French B. Biological substrates underpinning diagnosis of major depression. Int J Neuropsychopharmacol. 2013. 2013. https://doi.org/10.1017/S1461145713000436.
4. Scharnowski F, Nicholson AA, Pichon S, Rosa MJ, Rey G, Eickhoff SB, et al. The role of the subgenual anterior cingulate cortex in dorsomedial prefrontal–amygdala neural circuitry during positive-social emotion regulation. Hum Brain Mapp. 2020. 2020. https://doi.org/10.1002/hbm.25001.
5. Haas BW, Omura K, Constable RT, Canli T. Emotional conflict and neuroticism: personality-dependent activation in the amygdala and subgenual anterior cingulate. Behav Neurosci. 2007. 2007. https://doi.org/10.1037/0735-7044.121.2.249.
6. Masten CL, Eisenberger NI, Borofsky LA, Mcnealy K, Pfeifer JH, Dapretto M. Subgenual anterior cingulate responses to peer rejection: a marker of adolescents' risk for depression. Dev Psychopathol. 2011. 2011. https://doi.org/10.1017/S0954579410000799.
7. Mayberg HS, Liotti M, Brannan SK, McGinnis S, Mahurin RK, Jerabek PA, et al. Reciprocal limbic-cortical function and negative mood: converging PET findings in depression and normal sadness. Am J Psychiatry. 1999. 1999.

8. Mayberg HS, Lozano AM, Voon V, McNeely HE, Seminowicz D, Hamani C, et al. Deep brain stimulation for treatment-resistant depression. Neuron 2005;45:651–60.
9. Merkl A, Neumann WJ, Huebl J, Aust S, Horn A, Krauss JK, et al. Modulation of beta-band activity in the subgenual anterior cingulate cortex during emotional empathy in treatment-resistant depression. Cereb Cortex. 2016. 2016. https://doi.org/10.1093/cercor/bhv100.
10. Hasler G, van der Veen JW, Tumonis T, Meyers N, Shen J, Drevets WC. Reduced Prefrontal Glutamate/Glutamine and γ-Aminobutyric Acid Levels in Major Depression Determined Using Proton Magnetic Resonance Spectroscopy. Arch Gen Psychiatry. 2007;64:193.
11. Sanacora G, Mason GF, Rothman DL, Behar KL, Hyder F, Petroff OAC, et al. Reduced cortical γ-aminobutyric acid levels in depressed patients determined by proton magnetic resonance spectroscopy. Arch Gen Psychiatry. 1999. 1999. https://doi.org/10.1001/archpsyc.56.11.1043.
12. Sanacora G, Mason GF, Rothman DL, Krystal JH. Increased occipital cortex GABA concentrations in depressed patients after therapy with selective serotonin reuptake inhibitors. Am J Psychiatry. 2002. 2002. https://doi.org/10.1176/appi.ajp.159.4.663.
13. Sanacora G, Rothman DL, Mason G, Krystal JH. Clinical Studies Implementing Glutamate Neurotransmission in Mood Disorders. Ann N Y Acad Sci. 2003;1003:292–308.
14. Sequeira A, Mamdani F, Ernst C, Vawter MP, Bunney WE, Lebel V, et al. Global brain gene expression analysis links Glutamatergic and GABAergic alterations to suicide and major depression. PLoS One. 2009. 2009. https://doi.org/10.1371/journal.pone.0006585.
15. Duric V, Banasr M, Stockmeier CA, Simen AA, Newton SS, Overholser JC, et al. Altered expression of synapse and glutamate related genes in post-mortem hippocampus of depressed subjects. Int J Neuropsychopharmacol. 2013. 2013. https://doi.org/10.1017/S1461145712000016.
16. Ongur D, Drevets WC, Price JL. Glial reduction in the subgenual prefrontal cortex in mood disorders. Proc Natl Acad Sci. 2002. 2002. https://doi.org/10.1073/pnas.95.22.13290.
17. Medina A, Watson SJ, Bunney W, Myers RM, Schatzberg A, Barchas J, et al. Evidence for alterations of the glial syncytial function in major depressive disorder. J Psychiatr Res. 2016. 2016. https://doi.org/10.1016/j.jpsychires.2015.10.010.
18. Sibille E, Morris HM, Kota RS, Lewis DA. GABA-related transcripts in the dorsolateral prefrontal cortex in mood disorders. Int J Neuropsychopharmacol. 2011;14:721–34.
19. Tripp A, Kota RS, Lewis DA, Sibille E. Reduced somatostatin in subgenual anterior cingulate cortex in major depression. Neurobiol Dis. 2011;42:116–24.
20. Tripp A, Oh H, Guilloux J-PP, Martinowich K, Lewis DA, Sibille E. Brain-derived neurotrophic factor signaling and subgenual anterior cingulate cortex dysfunction in major depressive disorder. Am J Psychiatry. 2012;169:1194–202.
21. Sibille E, Wang Y, Joeyen-Waldorf J, Gaiteri C, Surget A, Oh S, et al. A molecular signature of depression in the amygdala. Am J Psychiatry. 2009;166:1011–24.
22. Scifo E, Pabba M, Kapadia F, Ma T, Lewis DA, Tseng GC, et al. Sustained molecular pathology across episodes and remission in major depressive disorder. Biol Psychiatry. 2018. 2018. https://doi.org/10.1016/j.biopsych.2017.08.008.
23. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA. 2005;102:15545–50.
24. Pantazatos SP, Huang Y-Y, Rosoklija GB, Dwork AJ, Arango V, Mann JJ. Whole-transcriptome brain expression and exon-usage profiling in major depression and suicide: evidence for altered glial, endothelial and ATPase activity. Mol Psychiatry. 2017;22:760–73.
25. Ramaker RC, Bowling KM, Lasseigne BN, Hagenauer MH, Hardigan AA, Davis NS, et al. Post-mortem molecular profiling of three psychiatric disorders. Genome Med. 2017. 2017. https://doi.org/10.1186/s13073-017-0458-5.
26. Labonté B, Engmann O, Purushothaman I, Menard C, Wang J, Tan C, et al. Sex-specific transcriptional signatures in human depression. Nat Med. 2017. 2017. https://doi.org/10.1038/nm.4386.
27. Ding Y, Chang L-C, Wang X, Guilloux J-P, Parrish J, Oh H, et al. Molecular and genetic characterization of depression: overlap with other psychiatric disorders and aging. Mol Neuropsychiatry. 2015;1:1–12.
28. Bioconductor - GeneOverlap. https://bioconductor.org/packages/release/bioc/html/GeneOverlap.html. Accessed 22 April 2021.
29. Baron M, Veres A, Wolock SL, Faust AL, Gaujoux R, Vetere A, et al. A single-cell transcriptomic map of the human and mouse pancreas reveals inter- and intra-cell population structure. Cell Syst. 2016. 2016. https://doi.org/10.1016/j.cels.2016.08.011.
30. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, et al. Comprehensive integration of single-cell data. Cell. 2019. 2019. https://doi.org/10.1016/j.cell.2019.05.031.
31. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods. 2015;12:453–7.
32. Shen-Orr SS, Tibshirani R, Khatri P, Bodian DL, Staedtler F, Perry NM, et al. Cell type–specific gene expression differences in complex tissues. Nat Methods. 2010;7:287–9.
33. Zhao W, Langfelder P, Fuller T, Dong J, Li A, Hovarth S. Weighted gene coexpression network analysis: State of the art. J Biopharm Stat. 2010.
34. Agrahari R, Foroushani A, Docking TR, Chang L, Duns G, Hudoba M, et al. Applications of Bayesian network models in predicting types of hematological malignancies. Sci Rep. 2018. 2018. https://doi.org/10.1038/s41598-018-24758-5.
35. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15.
36. Zhang B, Horvath S. A general framework for weighted gene co-expression network analysis. Stat Appl Genet Mol Biol. 2005;4.
37. Langfelder P, Horvath S. Eigengene networks for studying the relationships between co-expression modules. BMC Syst Biol. 2007;1:54.
38. Zhang B, Gaiteri C, Bodea LG, Wang Z, McElwee J, Podtelezhnikov AA, et al. Integrated systems approach identifies genetic nodes and networks in late-onset Alzheimer's disease. Cell. 2013. 2013. https://doi.org/10.1016/j.cell.2013.03.030.
39. Subramanian A, Narayan R, Corsello SM, Peck DD, Natoli TE, Lu X, et al. A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. Cell. 2017. 2017. https://doi.org/10.1016/j.cell.2017.10.049.
40. Daina A, Michielin O, Zoete V. SwissTargetPrediction: updated data and new features for efficient prediction of protein targets of small molecules. Nucleic Acids Res. 2019. 20 May 2019. https://doi.org/10.1093/nar/gkz382.
41. Shen-Orr SS, Tibshirani R, Khatri P, Bodian DL, Staedtler F, Perry NM, et al. Cell type-specific gene expression differences in complex tissues. Nat Methods. 2010. 2010. https://doi.org/10.1038/nmeth.1439.
42. Foroushani A, Agrahari R, Docking R, Chang L, Duns G, Hudoba M, et al. Large-scale gene network analysis reveals the significance of extracellular matrix pathway and homeobox genes in acute myeloid leukemia: An introduction to the Pigengene package and its applications. BMC Med Genomics. 2017. 2017. https://doi.org/10.1186/s12920-017-0253-6.
43. Anisman H, Merali Z. Cytokines, stress and depressive illness: Brain-immune interactions. Ann Med. 2003;35:2–11.
44. Darmanis S, Sloan SA, Zhang Y, Enge M, Caneda C, Shuer LM, et al. A survey of human brain transcriptome diversity at the single cell level. Proc Natl Acad Sci USA. 2015;112:7285–90.
45. He Z, Han D, Efimova O, Guijarro P, Yu Q, Oleksiak A, et al. Comprehensive transcriptome analysis of neocortical layers in humans, chimpanzees and macaques. Nat Neurosci. 2017. 2017. https://doi.org/10.1038/nn.4548.
46. Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, Wrobel MJ, et al. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. Science 2006;313:1929–35.
47. Gross J, Turecki G. Suicide and the Polyamine System. CNS Neurol Disord - Drug Targets. 2013. 2013. https://doi.org/10.2174/18715273113129990095.
48. Miller AH, Raison CL. The role of inflammation in depression: from evolutionary imperative to modern treatment target. Nat Rev Immunol. 2016;16:22–34.
49. Wang JQ, Mao L. The ERK pathway: molecular mechanisms and treatment of depression. Mol Neurobiol. 2019;56:6197–205.
50. Liu CH, Zhang GZ, Li B, Li M, Woelfer M, Walter M, et al. Role of inflammation in depression relapse. J Neuroinflammation. 2019;16:90.
51. E. Leonard B. The Concept of Depression as a Dysfunction of the Immune System. Curr Immunol Rev. 2010. 2010. https://doi.org/10.2174/157339510791823835.
52. Warner-Schmidt JL, Duman RS. VEGF as a potential target for therapeutic intervention in depression. Curr Opin Pharmacol. 2008;8:14–9.
53. Allen J, Romay-Tallon R, Brymer KJ, Caruncho HJ, Kalynchuk LE. Mitochondria and mood: Mitochondrial dysfunction as a key player in the manifestation of depression. Front Neurosci. 2018;12:386.
54. Culmsee C, Michels S, Scheu S, Arolt V, Dannlowski U, Alferink J. Mitochondria, microglia, and the immune system — How are they linked in affective disorders? Front Psychiatry. 2019. 2019. https://doi.org/10.3389/fpsyt.2018.00739.
55. Czarny P, Wigner P, Galecki P, Sliwinski T. The interplay between inflammation, oxidative stress, DNA damage, DNA repair and mitochondrial dysfunction in depression. Prog Neuro-PsychopharmacologyBiol Psychiatry. 2018;80:309–21.
56. Duman RS, Sanacora G, Krystal JH. Altered connectivity in depression: GABA and glutamate neurotransmitter deficits and reversal by novel treatments. Neuron. 2019;102:75–90.
57. Banasr M, Dwyer JM, Duman RS. Cell atrophy and loss in depression: reversal by antidepressant treatment. Curr Opin Cell Biol. 2011;23:730–7.
58. Rajkowska G, Miguel-Hidalgo JJ, Wei J, Dilley G, Pittman SD, Meltzer HY, et al. Morphometric evidence for neuronal and glial prefrontal cell pathology in major depression. Biol Psychiatry. 1999. 1999. https://doi.org/10.1016/S0006-3223(99)00041-4.
59. Fee C, Banasr M, Sibille E. Somatostatin-positive gamma-aminobutyric acid interneuron deficits in depression: cortical microcircuit and therapeutic perspectives. Biol Psychiatry. 2017;82:549–59.

60. Simon RA, Barazanji N, Jones MP, Bednarska O, Icenhour A, Engström M, et al. Vasoactive intestinal polypeptide plasma levels associated with affective symptoms and brain structure and function in healthy females. Sci Rep. 2021;11:1406.
61. Gjerris A, Rafaelsen OJ, Vendsborg P, Fahrenkrug J, Rehfeld JF. Vasoactive intestinal polypeptide decreased in cerebrospinal fluid (CSF) in atypical depression. Vasoactive intestinal polypeptide, cholecystokinin and gastrin in CSF in psychiatric disorders. J Affect Disord. 1984;7:325–37.
62. Soria V, Martínez-Amorós È, Escaramís G, Valero J, Pérez-Egea R, García C, et al. Differential association of circadian genes with mood disorders: CRY1 and NPAS2 are associated with unipolar major depression and clock and VIP with bipolar disorder. Neuropsychopharmacology 2010;35:1279–89.
63. Schulz JM, Knoflach F, Hernandez MC, Bischofberger J. Dendrite-targeting interneurons control synaptic NMDA-receptor activation via nonlinear α5-GABAA receptors. Nat Commun. 2018;9:3576.
64. Prévot T, Sibille E. Altered GABA-mediated information processing and cognitive dysfunctions in depression and other brain disorders. Mol Psychiatry. 2021;26:151–67.
65. Zaitsev K, Bambouskova M, Swain A, Artyomov MN. Complete deconvolution of cellular mixtures based on linearity of transcriptional signatures. Nat Commun. 2019. 2019. https://doi.org/10.1038/s41467-019-09990-5.
66. Wang J, Huang M, Torre E, Dueck H, Shaffer S, Murray J, et al. Gene expression distribution deconvolution in single-cell RNA sequencing. Proc Natl Acad Sci USA. 2018. 2018. https://doi.org/10.1073/pnas.1721085115.
67. Seney ML, Tripp A, McCune S, A. Lewis D, Sibille E. Laminar and cellular analyses of reduced somatostatin gene expression in the subgenual anterior cingulate cortex in major depression. Neurobiol Dis. 2015. 2015. https://doi.org/10.1016/j.nbd.2014.10.005.
68. Mostafavi S, Gaiteri C, Sullivan SE, White CC, Tasaki S, Xu J, et al. A molecular network of the aging human brain provides insights into the pathology and cognitive decline of Alzheimer's disease. Nat Neurosci. 2018. 2018. https://doi.org/10.1038/s41593-018-0154-9.
69. Meinshausen N, Hauser A, Mooij JM, Peters J, Versteeg P, Bühlmann P. Methods for causal inference from gene perturbation experiments and validation. Proc Natl Acad Sci USA. 2016;113:7361–8.

ACKNOWLEDGEMENTS

The study was supported by a project grant from the Canadian Institute of Health Research (CIHR) PJT-153175.

AUTHOR CONTRIBUTIONS

RS and ES conceptualized the study and together wrote the manuscript. RS performed all the sequencing library preparation, quality check and bioinformatics analysis. DFN participated in in-silico validations. TT and AS performed the QPCR validation. HZ participated in Bayesian network analysis. RM participated in cmap analysis and DAL provided the resources.

COMPETING INTERESTS

ES is founder and Acting Chief Scientific Officer of Damona Pharmaceuticals, a drug development company with small molecules in the pipeline for treatment of cognitive deficits across brain disorders and aging. All other authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41380-021-01347-z.

Correspondence and requests for materials should be addressed to Rammohan Shukla or Etienne Sibille.

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

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