ZOOLOGICAL RESEARCH

Transcriptomic analyses of humans and mice provide insights into depression

Hui-Juan Li[1,2,#], Xi Su[3,4,#], Lu-Wen Zhang[3,4,#], Chu-Yi Zhang[1,2], Lu Wang[1], Wen-Qiang Li[3,4], Yong-Feng Yang[3,4], Lu-Xian Lv[3,4,5], Ming Li[1,2,6,*], Xiao Xiao[1,*]

[1] *Key Laboratory of Animal Models and Human Disease Mechanisms of the Chinese Academy of Sciences and Yunnan Province*, *Kunming Institute of Zoology*, *Chinese Academy of Sciences*, *Kunming*, *Yunnan* 650223, *China*

[2] *Kunming College of Life Science*, *University of Chinese Academy of Sciences*, *Kunming*, *Yunnan* 650204, *China*

[3] *Henan Mental Hospital*, *Second Affiliated Hospital of Xinxiang Medical University*, *Xinxiang*, *Henan* 453002, *China*

[4] *Henan Key Lab of Biological Psychiatry*, *International Joint Research Laboratory for Psychiatry and Neuroscience of Henan*, *Xinxiang Medical University*, *Xinxiang*, *Henan* 453002, *China*

[5] *Henan Province People's Hospital*, *Zhengzhou*, *Henan* 450003, *China*

[6] *KIZ-CUHK Joint Laboratory of Bioresources and Molecular Research in Common Diseases*, *Kunming Institute of Zoology*, *Chinese Academy of Sciences*, *Kunming*, *Yunnan* 650223, *China*

ABSTRACT

Accumulating studies have been conducted to identify risk genes and relevant biological mechanisms underlying major depressive disorder (MDD). In particular, transcriptomic analyses in brain regions engaged in cognitive and emotional processes, e.g., the dorsolateral prefrontal cortex (DLPFC), have provided essential insights. Based on three independent DLPFC RNA-seq datasets of 79 MDD patients and 75 healthy controls, we performed differential expression analyses using two alternative approaches for cross-validation. We also conducted transcriptomic analyses in mice undergoing chronic variable stress (CVS) and chronic social defeat stress (CSDS). We identified 12 differentially expressed genes (DEGs) through both analytical methods in MDD patients, the majority of which were also dysregulated in stressed mice. Notably, the mRNA level of the immediate early gene *FOS* (*Fos* proto-oncogene) was significantly decreased in both MDD patients and CVS-exposed mice, and CSDS-susceptible mice exhibited a greater reduction in *Fos* expression compared to resilient mice. These findings suggest the potential key roles of this gene in the pathogenesis of MDD related to stress exposure. Altered transcriptomes in the DLPFC of MDD patients might be, at least partially, the result of stress exposure, supporting that stress is a primary risk factor for MDD.

Keywords: Major depressive disorder; Stressed mice; Dorsolateral prefrontal cortex; Transcriptomic analysis; *FOS*

INTRODUCTION

Major depressive disorder (MDD) is a devastating mental

Received: 01 July 2020; Accepted: 07 September 2020; Online: 08 September 2020

Foundation items: This study was supported by the Chinese Academy of Sciences (CAS) Western Light Program and CAS Youth Innovation Promotion Association to X.X. and the CAS Pioneer Hundred Talents Program and 1000 Young Talents Program to M.L.

[#]Authors contributed equally to this work

*Corresponding authors, E-mail: limingkiz@mail.kiz.ac.cn; xiaoxiao2@mail.kiz.ac.cn

DOI: 10.24272/j.issn.2095-8137.2020.174

illness with a lifetime prevalence of around 6.8%–13.0% worldwide (Hasin et al., 2005; Malhi & Mann, 2018; Otte et al., 2016). During the past few decades, numerous studies have been carried out to define its biological basis, pathological mechanisms, clinical biomarkers, and effective therapies (Kang et al., 2012; Le-Niculescu et al., 2009; Li et al., 2020a; Liu et al., 2020; Lu et al., 2020; Ota et al., 2014; Xiao et al., 2018). So far, several hypotheses about its pathogenesis have been proposed and tested (Duman & Aghajanian, 2012; Duman et al., 2016; Penzes et al., 2011), and recent genome-wide association studies (GWAS) have also identified genetic loci linked with MDD (Howard et al., 2019; Wray et al., 2018). While these studies have promoted our understanding of this illness, many questions remain unanswered. According to epidemiological analyses, the heritability of MDD is only moderate (Sullivan et al., 2000), and environmental factors, such as negative life events and stressful experiences, are primary risk factors that cannot be easily tested through GWAS. For example, previous studies have reported that mice exposed to stress showed depressive-like symptoms (Bagot et al., 2016; Cheng et al., 2018; Hodes et al., 2015), and early attachment-figure separation can increase risk for later depression (Hennessy et al., 2010). Therefore, the complexity of MDD pathogenesis requires investigations using data beyond those containing only genotype information.

Transcriptomic analyses of human tissues allow alternative studies to search for MDD biological mechanisms, as mRNA expression of genes is mediated by the combinatorial effects of genetic variations (Westra et al., 2013; Wright et al., 2014), stress exposures (Ota et al., 2014), and other factors (Charlesworth et al., 2010; Jansen et al., 2014). In recent years, transcriptomic analyses in brain and peripheral blood tissues have reported particularly intriguing genes and pathways relevant to MDD, which might be further applied as clinical biomarkers or potential drug targets (Jansen et al., 2016; Kim et al., 2016; Labonté et al., 2017; Leday et al., 2018; Li et al., 2013; Mostafavi et al., 2014; Pantazatos et al., 2017; Ramaker et al., 2017; Scarpa et al., 2018; Seney et al., 2018; Sequeira et al., 2007; Wang et al., 2008; Wittenberg et al., 2020). However, MDD is a polygenic illness originating from pathological alterations in the brain, and the sample sizes of current transcriptomic analyses in human brains are still too small to define sufficient dysregulated genes and pathways.

Here, we integrated independent RNA-seq datasets of dorsolateral prefrontal cortex (DLPFC) tissues from MDD patients, as well as GWAS statistics, expression quantitative trait loci (eQTL) in the DLPFC of humans, and RNA-seq data in the ventromedial prefrontal cortex (vmPFC) of mice exposed to chronic stress. We conducted differential expression analysis, weighted gene co-expression network analysis (WGCNA), gene set enrichment analysis (GSEA), summary data-based Mendelian randomization (SMR) analysis, and rank-rank hypergeometric overlap (RRHO). These analyses yielded dysregulated genes, co-expression modules, and transcriptional networks associated with human MDD, many of which were also reproduced in stressed mice, thus providing insights into the pathogenesis of depression.

MATERIALS AND METHODS

All protocols and methods were approved by the Institutional Review Board of the Kunming Institute of Zoology, Chinese Academy of Sciences, China.

Differential expression analysis using DLPFC RNA-seq data in MDD patients and controls

RNA-seq datasets: Three independent raw RNA-seq datasets (SRA files) were downloaded from the Gene Expression Omnibus (GEO) database. The first GEO dataset (GSE102556) included human postmortem brain DLPFC (BA8/9) tissues from 26 MDD cases and 22 matched controls from the Douglas Bell Canada Brain Bank (DBCBQ; Douglas Mental Health Institute, Canada) (Labonté et al., 2017). All subjects in the GSE102556 dataset were of European ancestry and French-Canadian descent, and their average age of death was 47. As described by the authors, “diagnoses were obtained using DSM-IV criteria by means of SCID-I interviews adapted for psychological autopsies”, and the samples were “barcoded for multiplexing and sequenced at 50 bp paired-end on Illumina Hi-Seq 2500” (Labonté et al., 2017). The second GEO dataset (GSE101521) comprised brain tissues from 30 DSM-IV-defined MDD cases and 29 controls collected at the Division of Molecular Imaging and Neuropathology, New York State Psychiatric Institute and Columbia University (Pantazatos et al., 2017). In this dataset, RNA was extracted from DLPFC (BA9) tissues, and paired-end sequencing for total RNA was performed on an Illumina Hi-Seq 2500 with a 100 bp read length (Pantazatos et al., 2017).

The third GEO dataset (GSE80655) consisted of the middle part of the superior frontal gyrus (DLPFC, BA9) from 23 MDD patients and 24 healthy controls (Ramaker et al., 2017). These brain samples were collected under the Brain Donor Program at the University of California, Irvine, Department of Psychiatry and Human Behavior. Total RNA was extracted and sequenced on an Illumina Hi-Seq 2000 with a paired-end 50 bp read length.

Quality control analysis: Quality control, alignment, and gene-expression quantification were performed through the same procedures under consistent criteria for all three RNA-seq datasets. In brief, Trimmomatic v0.36 was used to examine the sequencing quality and trim reads (Bolger et al., 2014). Clean paired-end reads were aligned to the genome GRCh38 using Hisat2 (v2.1.0) (Kim et al., 2015). The gene-level read counts were quantified based on the gene annotation file (GRCh38.91) using featureCounts and transcript per million (TPM) was calculated to quantify gene expression levels (Liao et al., 2014). Genes with an average TPM of <1.0 or non-protein coding were excluded from subsequent analyses. Finally, quality control, alignment, and gene-expression quantification yielded 12 928 protein-coding genes with TPM≥1.0 for subsequent analyses. To merge the three datasets, maximize statistical power, and identify

differentially expressed genes (DEGs), we simultaneously applied two different approaches (i.e., merged analysis and meta-analysis in the following discussion).

Identification of dysregulated genes using merged analysis: The TPM expression matrices from three different GEO datasets were merged and transformed as $\log_2$(TPM+1), and then surrogate variables (SVs) were estimated using the *sva* function (v3.28.0) (Leek et al., 2020). The *removeBatchEffect* function in *limma* (v3.36.5) was used to remove batch effects and adjust for known covariates (e.g., sex, age, postmortem interval (PMI), RNA integrity number (RIN)) and to estimate SVs and the top three expression principal components (PCs) (Ritchie et al., 2015). Normalized TPM matrices were proceeded using linear fit in the *limma* package to conduct differential expression analysis. The $\log_2$(fold-change) ($\log_2$FC) values and *P*-values of each gene were computed using the *eBays* function in the *limma* package. Genes with a false discovery rate (FDR)≤0.1 and fold-change≥1.2 were defined as DEGs following previous studies (Li et al., 2020b; Xu et al., 2018).

Identification of dysregulated genes using meta-analysis: Differential expression analysis based on read counts was also performed. Briefly, the likelihood ratio test (LRT) in the DESeq2 package (v1.20.0) (Love et al., 2014) was carried out for each dataset. For the LRT, known covariates (e.g., sex, age, PMI, RIN) were adjusted for each study, and the LRT statistics in the three datasets were used for Liptak-Stouffer's meta-analysis (Laoutidis & Luckhaus, 2015). Genes with a meta-analysis FDR≤0.1 and fold-change≥ 1.2 were considered DEGs.

Correlation analysis of two methods for calculations of DEGs: To increase the reliability of DEGs, Pearson's correlation and sign tests were conducted using $\log_2$(fold-change) derived from both the merged and meta-analyses. DEGs identified through both approaches were used for further analyses.

Weighted gene co-expression network analysis (WGCNA)

Based on the normalized-TPM matrices of 12 928 protein-coding genes in 79 MDD patients and 75 healthy controls from all three datasets, unsigned co-expression gene modules were constructed using WGCNA (v1.66). This is a common method for identifying clusters of highly correlated genes (Langfelder & Horvath, 2008). Under the assumption that a gene co-expression network is scale-free, this method uses a soft threshold to estimate the weights of edges between two random genes. It is considered more robust than an unweighted network (i.e., dichotomizing connection between genes, 1 represents co-expression, 0 represents non-coexpression) in identifying co-expression patterns (Zhang & Horvath, 2005). Briefly, we calculated the network adjacency and topological overlap matrix (TOM) using an appropriate soft-threshold power after analysis of scale-free topology for multiple soft-threshold powers. According to the default parameters, the appropriate soft-threshold power was 5, based on which the R^2 of scale topology fit equaled 0.88. Gene clusters were plotted based on TOM dissimilarity. Modules whose dissimilarity of eigengenes was less than 0.1 were merged using the *mergeCloseModules* function after dynamic modules were generated. The module eigengenes (first PC) as well as the associations between modules and diagnostic status were examined using the *moduleEigengenes* function. Modules with an adjusted *P*≤0.05 were considered significantly associated with the disorder. In addition, based on the weight of co-expression between any two genes in a specific module, we calculated the co-expression scores of each gene with the remaining genes in the same module using $\frac{\sum_{i=1}^{n} w_i}{n}$ to define the hub genes, where *w* is the weight between a specific gene and other genes in the same module and *n* is the number of genes in the tested module minus 1.

Protein-protein interactions (PPIs), biological processes, and pathway analyses

Physical PPIs among DEGs were constructed using the STRING database (v11.0) (Szklarczyk et al., 2017). To examine whether genes were enriched in any biological processes (BP) potentially relevant to MDD pathogenesis, functional prediction was performed using Gene Ontology (GO) annotation with clusterProfiler (v3.8.1) (Yu et al., 2012), and GO BP terms with an FDR-corrected *P*≤0.05 were considered statistically significant. Semantic similarity analyses were then conducted with GOSemSim (v2.6.2) (Yu et al., 2010) to narrow down the enriched GO terms based on their similarity with each other (similarity≥0.5 was considered highly similar).

SMR analysis using depression GWAS and DLPFC eQTL datasets

To identify genes whose mRNA expression levels were affected by genetic risk of MDD, we integrated the MDD GWAS statistics (Wray et al., 2018) and human DLPFC RNA-seq eQTL data (Gandal et al., 2018b), and performed SMR analysis (v1.03) (Zhu et al., 2016). This method was developed based on the Mendelian randomization (MR) framework (Smith & Ebrahim, 2003; Smith & Hemani, 2014), which considers genetic variations, mRNA expression levels of genes, and traits as the instrumental variable, exposure variable, and outcome, respectively. Genome-wide statistics of MDD GWAS (including 59 851 cases and 113 154 controls) were retrieved from recent study (Wray et al., 2018), and summary statistics of RNA-seq eQTL (calculated by adjusting for 100 hidden covariate factors) of 1 387 individuals were obtained from the PsychENCODE dataset at http://resource.psychencode.org/ (Gandal et al., 2018b). During SMR analysis, the eQTL *P*-value threshold was set at 0.01, and genes with multi-single nucleotide polymorphism (SNP)-based SMR *P*-values lower than 0.05 were empirically considered to be potential MDD risk genes (Wu et al., 2018).

Differential expression analysis using vmPFC RNA-seq data in stressed mice

We downloaded the raw vmPFC RNA-seq data of stressed mice from GSE102556 (Labonté et al., 2017). This study

included 19 control mice and 19 mice exposed to chronic variable stress (CVS) for 21 days. All CVS-stressed mice exhibited a range of depression- and anxiety-related behavioral abnormalities. The brain vmPFC tissues of these mice were dissected, and their transcriptomes were examined through 50 bp paired-end RNA-sequencing on an Illumina Hi-Seq 2500.

We also obtained the raw vmPFC RNA-seq data of stressed mice from GSE81672 (Bagot et al., 2017). In this study, mice underwent chronic social defeat stress (CSDS), and were further characterized as either susceptible (n=3) or resilient (n=4) according to their behavioral indices. Specifically, the susceptible mice showed depression- and anxiety-related behaviors following CSDS exposure. A control group comprising mice without CSDS exposure were also included (n=5). RNA-seq of the vmPFC tissues of these mice was performed on an Illumina Hi-Seq 2500 with 50 bp paired-end reads.

For both murine brain RNA-seq datasets, we conducted quality control, alignment, and gene-expression quantification using a similar pipeline as that applied for the human data, with the reference genome replaced by GRCm38. As these data were obtained from inbred mice under strict experimental control, there were no covariates to be adjusted for in comparison with the human data. We therefore used the *Wald* test in the DESeq2 package to perform differential expression analysis between stressed and control mice.

Rank-rank hypergeometric overlap (RRHO)

The extent of overlap of dysregulated genes highlighted in MDD patients and stressed mice were examined by performing unbiased RRHO (v1.26.0), as described previously (Bagot et al., 2016; Cahill et al., 2018; Plaisier et al., 2010; Seney et al., 2018; Stein et al., 2014). This is a threshold-free method that estimates overlap between two ranked lists of genes. To perform RRHO analysis, the two gene sets were firstly ranked at the genome-wide scale according to their $-\log_{10}$(P-value) and direction of change revealed by differential expression analyses. Secondly, a series of hypergeometric P-values were calculated for each gene by sliding the rank threshold to estimate the significance of overlapping genes above the expected threshold at each ranking site. Lastly, the hypergeometric P-values of genes were plotted on a heatmap.

RESULTS

Differential expression analysis at single gene level in depressed patients

Three independent RNA-seq datasets of human DLPFC tissues, including a total of 79 MDD patients and 75 healthy controls, were utilized. After quality control analysis, 12 928 protein-coding genes showing TPM≥1.0 were used for subsequent analyses. We firstly merged the TPM matrices of all individuals from all three datasets, and then normalized the data by adjusting for batch effects (by distinct studies), known covariates (i.e., sex, pH and PMI), SVs (representing unknown factors), and the top three expression PCs. The PC plot of normalized expression matrices is presented in Figure 1A, which showed generally consistent gene expression patterns across individuals from different studies. We then conducted differential expression analysis based on the normalized TPM matrices in the merged datasets using linear fit in the *limma* package. In total, 2 277 genes showed nominal significance (P≤0.05; Supplementary Table S1), and 19 DEGs were identified (FDR≤0.1, fold-change≥1.2; Figure 1B).

We also calculated the DEGs based on read counts. As described in the methods section, differential expression analysis was performed in each dataset using the LTR method in DESeq2 (adjusting for covariates), and a meta-analysis of the three datasets was then carried out using the Liptak-Stouffer's method. This approach identified 904 genes with nominal significance (meta-analysis P≤0.05; Supplementary Table S2) and 20 DEGs (meta-analysis FDR≤0.1, fold-change≥1.2; Figure 1C).

Pearson's correlation analysis using genome-wide statistics also showed that the correlation coefficients of $\log_2$(fold-change) obtained via both methods were highly consistent (Pearson's r=0.81; Figure 1D). Among the 12 928 genes included in the analysis, 10 273 showed skewed expression levels compared with the controls in the same direction using both methods (sign test $P<2.2\times10^{-16}$). Notably, mRNA levels in 710 genes were nominally altered (P≤0.05) compared with controls using both analytical methods, with their statistics being highly consistent ($\log_2$FC, Pearson's r=0.95; Supplementary Table S3). Twelve genes were further identified as significant DEGs in both analyses (i.e., *FOS*, *EGR1*, *FOSB*, *EGR2*, *NR4A1*, *NR4A3*, *RLBP1*, *CCL3L3*, *PLA2G5*, *TRIB1*, *CYR61*, and *CH25H*; Table 1; Figure 1E). Notably, many of these DEGs (e.g., *FOS*, *FOSB*, *EGR1* and *EGR2*) belong to the family of immediate early genes (IEGs) that respond rapidly and transiently to cellular stimuli and influence neuronal physiology (Curran & Franza, 1988; Gallo et al., 2018). Here, their protein products formed a dense PPI network (Figure 1F).

Construction of gene co-expression modules in brain DLPFC and identification of MDD-associated modules

Examining transcriptional networks associated with MDD should provide insights into genes facilitating disease-related alterations (Gandal et al., 2018a; Gerring et al., 2019). We constructed co-expression modules in 154 DLPFC samples (79 MDD patients and 75 healthy controls) through WGCNA using normalized expression (after removing batch effects, known covariates, SVs, and top three PCs) in the 12 928 protein-coding genes. We performed hierarchical clustering of genes using the dynamic tree cut approach and then merged modules with dissimilarities<0.1 according to their eigengenes. We identified 47 co-expression modules, except for the grey module, which consisted of unclassified genes (Figure 2A; Supplementary Table S4). To estimate whether a group of genes within a module exhibited case-control

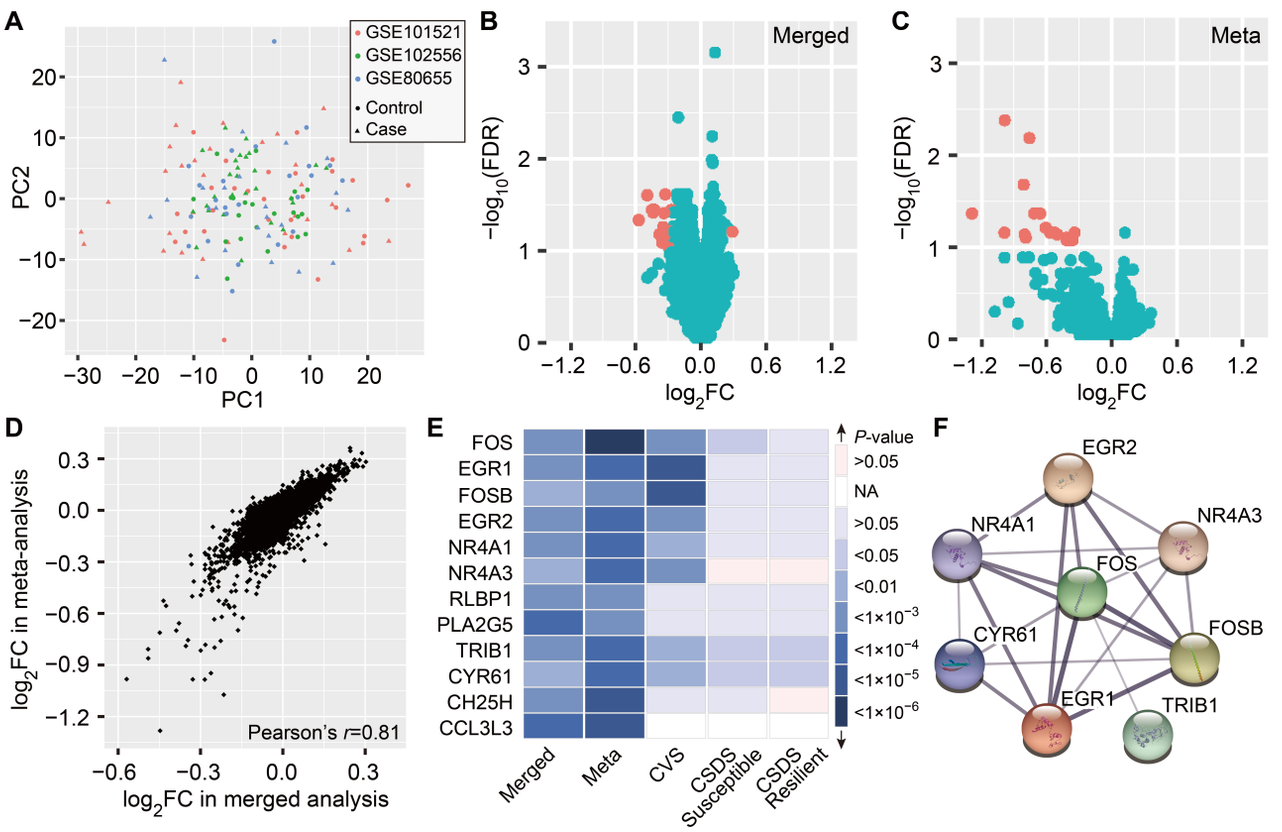

Figure 1 Differential expression analysis in depressed patients

A: PC plot of normalized expression matrices in three RNA-seq datasets. B: Volcano plot of genome-wide results using merged analysis. Red dots indicate DEGs (FDR≤0.1 and fold-change≥1.2). C: Volcano plot of genome-wide results using meta-analysis. Red dots indicate DEGs (FDR≤0.1 and fold-change≥1.2). D: Correlation analysis of $\log_2$FC derived from both analytical methods. E: Differential expression analysis of 12 DEGs identified in MDD patients by both methods. First two columns show results in MDD patients, last three columns show results in stressed mice. Blue indicates down-regulated genes, red indicates up-regulated genes in corresponding analysis. F: Eight out of 12 DEGs showed direct protein-protein interactions (PPIs). FC: Fold-change; FDR: False discovery rate; CVS: Chronic variable stress; CSDS: Chronic social defeat stress.

differences, we calculated module eigengenes (first PC), and found that three modules were significantly associated with diagnosis of MDD after multiple corrections (P≤0.05/47; Figure 2B; Supplementary Table S5), i.e., skyblue3 (P=1.36×10^{-4}, module size=97), ivory (P=5.03×10^{-4}, module size=85), and mediumpurple3 (P=9.09×10^{-4}, module size=91). For these three MDD-associated modules, Supplementary Table S6 shows the co-expression scores of each gene with the remaining genes in the same module, and Figure 2C highlights the top 20 hub genes in each module.

We then investigated whether genes in each MDD-associated module were enriched in any biological processes. GSEA showed that genes in the skyblue3 module were significantly enriched in the GO terms "regulation of synapse structure or activity", "regulation of synapse organization", "dendritic spine organization", and "circadian rhythm"; genes in the ivory module were significantly enriched in "learning", "rhythmic process", and "response to steroid hormone" (FDR≤0.05; Figure 2D); and genes in mediumpurple3 module were not enriched in any biological processes.

We also analyzed the enrichment of DEGs highlighted through single gene level analyses (FDR≤0.1, fold-change≥1.2) in these modules, and found that many DEGs (*FOS*, *EGR1*, *EGR2*, *NR4A3*, *NR4A1*, *CYR61*, and *CH25H*) were included in the ivory module. This module also contained many genes that showed nominal significance in differential expression analysis between cases and controls (P≤0.05; Supplementary Table S7). Furthermore, GO analysis found that these dysregulated genes were significantly enriched in "regulation of synaptic plasticity", "learning or memory", "regulation of neuron death", "response to calcium ion", "positive regulation of inflammatory response", "rhythmic process", and "response to steroid hormone" (FDR≤0.05; Figure 2D).

Integrative analysis of DEGs with GWAS risk genes in depression

We investigated whether DEGs in the DLPFC of MDD patients were associated with genetic risk of this illness by performing

Table 1 Identification of 12 differentially expressed genes (DEGs) in human MDD and their associations in stressed mice

Human gene		Merged analysis			Meta-analysis			Mouse	CVS mice		CSDS susceptible		CSDS resilient		WGCNA module
Ensembl ID	Name	log_2FC	*P*-value	FDR	log_2FC	*P*-value	FDR	Name	log_2FC	*P*-value	log_2FC	*P*-value	log_2FC	*P*-value	
ENSG00000170345	*FOS*	−0.569	3.35E−04	4.67E−02	−0.982	3.25E−07	4.21E−03	*Fos*	−0.917	2.13E−04	−1.238	2.83E−02	−0.735	0.156	Ivory
ENSG00000138135	*CH25H*	−0.346	5.76E−04	5.56E−02	−0.755	1.02E−06	6.57E−03	*Ch25h*	−0.542	0.070	−0.263	0.730	0.581	0.235	Ivory
ENSG00000276085	*CCL3L3*	−0.491	3.54E−05	2.50E−02	−0.808	4.88E−06	2.10E−02	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Ivory
ENSG00000122877	*EGR2*	−0.448	1.32E−04	3.61E−02	−0.712	1.75E−05	4.31E−02	*Egr2*	−0.995	2.80E−04	−0.950	0.155	−0.465	0.434	Ivory
ENSG00000142871	*CYR61*	−0.378	1.08E−03	6.65E−02	−0.657	2.00E−05	4.31E−02	*Cyr61*	−0.872	1.59E−03	−0.568	2.05E−02	−0.518	1.43E−02	Ivory
ENSG00000173334	*TRIB1*	−0.266	1.65E−04	3.61E−02	−0.340	5.45E−05	6.98E−02	*Trib1*	−0.311	4.67E−03	−0.330	2.77E−02	−0.309	1.65E−02	Skyblue3
ENSG00000120738	*EGR1*	−0.435	1.92E−04	3.86E−02	−0.526	7.01E−05	6.98E−02	*Egr1*	−0.564	2.02E−06	−0.358	0.271	−0.180	0.558	Ivory
ENSG00000123358	*NR4A1*	−0.425	1.16E−04	3.60E−02	−0.556	7.62E−05	7.03E−02	*Nr4a1*	−0.465	7.07E−03	−0.549	0.056	−0.329	0.189	Ivory
ENSG00000119508	*NR4A3*	−0.330	1.54E−03	7.65E−02	−0.500	9.03E−05	7.30E−02	*Nr4a3*	−0.382	3.39E−04	0.002	0.996	0.215	0.593	Ivory
ENSG00000140522	*RLBP1*	−0.284	7.47E−04	6.15E−02	−0.414	1.08E−04	7.90E−02	*Rlbp1*	−0.086	0.195	−0.108	0.322	−0.019	0.824	Brown
ENSG00000125740	*FOSB*	−0.353	1.95E−03	8.18E−02	−0.790	1.10E−04	7.90E−02	*Fosb*	−0.830	3.44E−06	−0.275	0.111	−0.045	0.771	Lightyellow
ENSG00000127472	*PLA2G5*	−0.325	2.89E−05	2.45E−02	−0.414	1.36E−04	8.36E−02	*Pla2g5*	−0.166	0.257	−0.137	0.528	−0.341	0.102	Brown

FC: Fold-change; FDR: False discovery rate; CVS: Chronic variable stress; CSDS: Chronic social defeat stress. N/A: Not available.

overlap analysis between DEGs and genetic risk genes. In brief, based on SMR analysis and integration of MDD GWAS statistics and human DLPFC eQTL results (Gandal et al., 2018b; Wray et al., 2018), we found that the mRNA expression levels of 490 genes in the DLPFC were nominally associated with GWAS genetic risk of MDD (SMR $P \leq 0.05$, Supplementary Table S8). However, none of the 12 DEGs showed nominal associations in SMR analysis. We further examined whether any genes were included in both the 490 genetic risk genes and 710 nominal DEGs ($P \leq 0.05$); however, only nine genes showed consistent expression change direction based on genetic risk and diagnostic effects (Supplementary Table S9).

Differential expression analysis of mice under chronic variable stress

Stressful and negative life events are major environmental risk factors of MDD. Thus, it is of great interest to examine whether stress exposure can lead to MDD-relevant transcriptomic alterations in brain tissues. Due to the lack of human brain transcriptomic information with stress exposure as the only environmental stimuli, we analyzed the transcriptomes of vmPFC tissues from mice under CVS (Labonté et al., 2017), which included 19 CVS-stressed mice and 19 unstressed mice. We then performed differential expression analysis on 11 695 protein-coding genes expressed in both humans and mice to gain insights into stress-correlated brain transcriptomic characteristics. The RRHO test revealed substantial overlap of down-regulated genes in CVS-stressed mice and MDD patients compared with their respective controls (maximum Fisher's exact test (FET) $P<1\times10^{-66}$; Figure 3A, B), especially the significantly down-regulated genes in humans identified through meta-analysis. These results suggest that analyses of murine brain transcriptomes undergoing CVS should provide valuable information for the brain transcriptomes of human MDD.

Supplementary Table S10 lists the genes nominally dysregulated in both MDD patients and CVS-stressed mice ($P \leq 0.05$). Among the 12 DEGs in MDD patients defined by both merged analysis and meta-analysis (11 genes were covered by RNA-seq analysis in mice), eight genes also showed consistent and nominal significance in CVS mice (i.e., *FOS*, *EGR1*, *FOSB*, *EGR2*, *NR4A1*, *NR4A3*, *TRIB1*, and *CYR61*; Figure 1E; Table 1). These data together suggest a significant overlap of down-regulated genes in MDD patients and CVS-stressed mice.

Differential expression analysis of mice under chronic social defeat stress

We also analyzed the transcriptomes of vmPFC tissues from mice under CSDS (Bagot et al., 2017). As mentioned earlier, the stressed mice were further characterized as either susceptible or resilient, with susceptible mice showing depression- and anxiety-related behaviors following CSDS exposure. Supplementary Table S11 shows whether the 710 nominal DEGs in humans were also differentially expressed in CSDS-susceptible and -resilient mice. Our analysis revealed

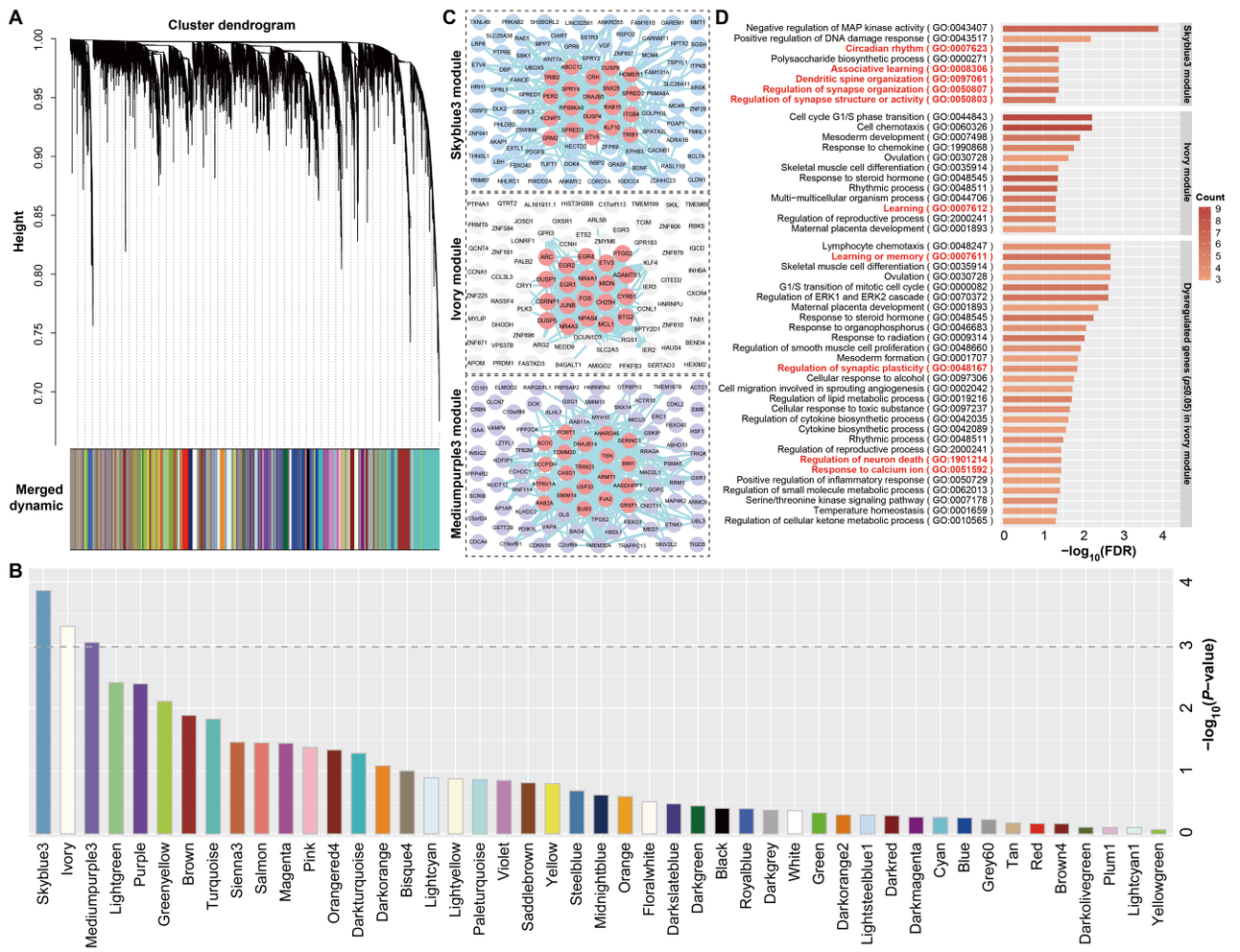

Figure 2 Gene co-expression modules in DLPFC of brain, MDD-associated modules, and biological pathway enrichment

A: In total, 47 co-expression modules were identified. B: Three co-expression modules were significantly associated with MDD after multiple corrections (i.e., $P \leq 0.05/47$), i.e., skyblue3, ivory, and mediumpurple3, respectively. C: Number of genes in three modules was 97, 85, and 91, respectively. Top 5% of gene-gene links in respective modules are displayed according to weights calculated by WGCNA. In addition, based on co-expression scores of each gene in three MDD-associated modules, the top 20 genes are in red. D: Gene Ontology (GO) analysis results for genes in skyblue3 and ivory modules as well as nominally dysregulated genes in ivory module.

the opposing patterns of gene expression alterations in humans (i.e., dysregulated genes identified using meta-analysis) and CSDS-resilient mice (enriched dots at top-left corner in heatmap plot) (Figure 3B), and also highlighted a significant overlap in genes down-regulated in both MDD patients (i.e., dysregulated genes identified using meta-analysis) and CSDS-susceptible mice (enriched dots at top-right corner in heatmap plot) (maximum FET $P<1\times10^{-95}$; Figure 3B). These results suggest that susceptibility to CSDS-induced behavioral impairments was likely affected by MDD-relevant genes, confirming the putative roles of stress exposure in the pathogenesis of MDD.

We then examined whether the 12 DEGs in depressed patients (11 genes covered by RNA-seq analysis in mice) were also correlated with depression- and anxiety-like symptoms in CSDS mice. However, the *P*-values of most genes did not reach nominal significance, probably due to limited statistical power. Furthermore, only three genes (i.e., *FOS*, *TRIB1*, and *CYR61*; Figure 1E; Table 1) showed nominal expression differences in CSDS-susceptible mice compared with control mice ($P \leq 0.05$). The expression levels of *FOS* did not significantly differ between CSDS-resilient mice and control mice ($P=0.156$; Table 1), whereas the expression levels of *TRIB1* and *CYR61* were nominally significant in CSDS-resilient mice, although their *P*-values were smaller than those in CSDS-susceptible mice (Figure 1E; Table 1). Overall, despite some caveats likely caused by the limited sample size and unique characteristics of CSDS-resilient mice, analysis of CSDS-susceptible mice obtained similar results as those of CVS mice, further supporting the significant

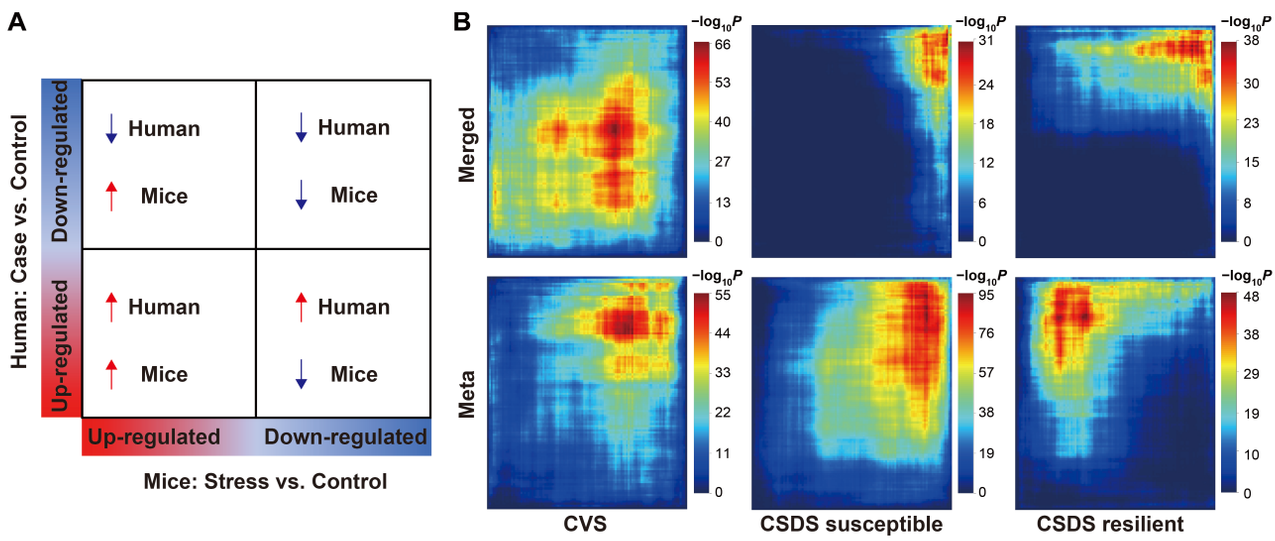

Figure 3 Comparison analysis of dysregulated genes between MDD patients and stressed mice

A: Schematic indicating interpretation of Rank-Rank Hypergeometric Overlap (RRHO) test plots. Signals in bottom left quadrant and top right quadrant represent an overlap of up-regulated and down-regulated genes found in both MDD patients and stressed mice, respectively. Signals in top left quadrant represent genes up-regulated in MDD patients but down-regulated in stressed mice; signals in bottom right quadrant represent genes down-regulated in MDD patients but up-regulated in stressed mice. Degree of significance is depicted in color bar of RRHO maps. B: RRHO maps comparing transcriptional profiles of MDD patients and stressed mice. CVS: Chronic variable stress; CSDS: Chronic social defeat stress.

overlap of down-regulated genes in depressed patients and stressed mice.

DISCUSSION

Extensive transcriptomic analyses of MDD patients have been conducted based on currently available microarray or RNA-seq datasets in brain and peripheral tissues. While previous studies have identified several dysregulated genes, satisfactory consistency between studies is lacking, calling for further investigations using larger samples. In the present study, we conducted differential expression analysis using RNA-seq data from three independent cohorts of MDD patients and controls. Although our sample size is smaller than that used for transcriptomic analyses of other psychiatric disorders (Gandal et al., 2018b), the total sample size is larger than most previous MDD brain transcriptomic studies (Kim et al., 2016; Labonté et al., 2017; Li et al., 2013; Pantazatos et al., 2017; Ramaker et al., 2017; Scarpa et al., 2018; Seney et al., 2018; Sequeira et al., 2007; Wang et al., 2008). Using two alternative analytical approaches, we identified 12 DEGs, which showed different mRNA levels among diagnostic groups. Notably, several identified DEGs have also been implicated in earlier studies (Pantazatos et al., 2017; Scarpa et al., 2018), although the magnitudes of significance differed among studies due to varied sample sizes.

We did not observe significant overlap between DEGs (or nominally dysregulated genes) and MDD genetic risk genes, suggesting that genetic risk might have only subtle effects on transcriptomic change, at least in the human DLPFC tissues. Nevertheless, the human DEGs were successfully replicated in stressed mice, highlighting the pivotal roles of stress exposure in altered brain transcriptomes. This result confirms the contention that stress is an important risk factor for depression. Consistently, recent research also reported shared transcriptomic signatures between MDD patients and mice under chronic stress (Scarpa et al., 2020). Therefore, cross-species analyses may provide crucial insights into the neurobiological basis of depression.

Based on WGCNA, we found three co-expression modules associated with MDD. Genes in these modules were significantly enriched in GO terms, including spine-, synapse-, and learning-associated processes, thereby confirming the impairment of cognitive function as a typical MDD characteristic, as applied in diagnostic and therapeutic practices (Gotlib & Joormann, 2010; Solé et al., 2015). In addition, our results were consistent with previous studies showing that synaptic dysfunctions play important roles in the pathogenesis of depression (Duman & Aghajanian, 2012; Russo & Nestler, 2013). Intriguingly, some genes in the MDD-associated modules also participate in the regulation of circadian rhythm. It is acknowledged that sleep disturbance is a core symptom of depression (Nutt et al., 2008; Riemann et al., 2001), and serotoninergic neurotransmission dysfunction appears to play a pivotal role in insomnia in MDD patients (Adrien, 2002), thus explaining the neurobiological basis underlying the relationship between depression and circadian rhythm. Therefore, our results further support these pathological mechanisms in MDD.

We identified several genes potentially involved in MDD and stress response, among which *FOS* was significantly

correlated with MDD diagnosis and stress exposure in both humans and mice. *FOS* is an IEG known to respond rapidly and transiently to cellular stimuli (Curran & Franza, 1988; Gallo et al., 2018). So far, accumulating studies have revealed the vital roles of IEGs in a variety of physiological processes, e.g., IEGs in certain types of cells and brain regions are tightly linked with learning and memory processes, as well as synaptic plasticity (Fleischmann et al., 2003; Gallo et al., 2018; Guzowski et al., 1999; Minatohara et al., 2015; Yasoshima et al., 2006), and have been widely applied as neuronal activity markers (Zhu et al., 2017). Studies using murine models of depression have also shown that IEGs are the most representative responders to stress exposure. For example, previous study found that Δ*FosB* overexpression in mice produces a resilience effect against CSDS (Donahue et al., 2014). Our results also support the potential role of IEGs as a connecting node between stress and MDD-relevant pathological changes (Covington et al., 2010; Hodges et al., 2014). Additionally, we showed that IEGs were significantly down-regulated in MDD patients, thereby confirming the abnormal cognitive function of such patients and stressed model animals (Millan et al., 2012), which likely resulted from synaptic dysfunctions in brain regions engaged in cognitive and memory processes (Forrest et al., 2018; Kang et al., 2012). Intriguingly, we also showed that *Fos* was significantly down-regulated following CSDS exposure, but only in mice susceptible to CSDS rather than in the resilient mice. Therefore, *Fos* may be a crucial susceptibility factor of stress-induced depression and may facilitate stress response-related MDD pathogenesis. However, further research is required.

While this study has presented intriguing findings, limitations should also be acknowledged. As described earlier, MDD is a polygenic disorder that likely involves hundreds or thousands of genes in its pathogenesis. Due to the limited sample size, our differential expression analyses only identified 12 genes surpassing the threshold of genome-wide significance. It is reasonable to argue that there are still important DEGs to be discovered and further investigations using larger samples are warranted. In addition, some included MDD patients had taken medications that may have affected their brain transcriptomes, which is a potential study limitation. Furthermore, the non-significant overlap between DEGs (or nominally dysregulated genes) and MDD genetic risk genes remains to be explained. It may have resulted due to technical reasons (e.g., small sample size, selection of DLPFC) or may be a reflection of sophisticated transcriptome regulation in MDD. Given that risk genetic variations of psychiatric disorders often affect mRNA expression of nearby or distal genes in the brain (Edwards et al., 2013; French & Edwards, 2020; Li et al., 2019; Liu et al., 2019; Yang et al., 2020a; Yang et al., 2020b), it is unlikely that this non-significant overlap provides a final answer. Nevertheless, recent single-cell transcriptomic analyses have found that gene expression profiles of different types of cells in the brain differ significantly (Kanton et al., 2019; Zhong et al., 2018); therefore, it is possible that genetic risk may only exert impact in certain types of cells, which may have been overshadowed in the current analyses of bulk tissues. In addition, the current study utilized data obtained from postnatal brain tissues, and most donors were adults. Whether MDD genetic risk affects the brain transcriptome during a particular stage of neurodevelopment is also an important question to answer. Last, but not least, growing evidence suggests that genetic risk may selectively affect mRNA abundance of particular isoforms of a gene (Cai et al., 2020; Li et al., 2016; Ma et al., 2020), yet many such isoforms were not annotated in the current reference genome. Therefore, further investigations on the impact of genetic risk on isoforms or transcripts differentially expressed in MDD patients compared with controls are necessary.

In summary, we observed transcriptomic dysregulations in depression across humans and mice, many of which responded to stress exposure. Despite several questions remaining to be addressed, our data provide valuable insights into the pathogenesis of depression.

DATA AVAILABILITY

The brain eQTL dataset used in this study was generated as part of the PsychENCODE Consortium, supported by: U01MH103392, U01MH103365, U01MH103346, U01MH 103340, U01MH103339, R21MH109956, R21MH105881, R21MH105853, R21MH103877, R21MH102791, R01MH 111721, R01MH110928, R01MH110927, R01MH110926, R01MH110921, R01MH110920, R01MH110905, R01MH 109715, R01MH109677, R01MH105898, R01MH105898, R01MH094714, P50MH106934, U01MH116488, U01MH 116487, U01MH116492, U01MH116489, U01MH116438, U01MH116441, U01MH116442, R01MH114911, R01MH 114899, R01MH114901, R01MH117293, R01MH117291, and R01MH117292 awarded to: Schahram Akbarian (Icahn School of Medicine at Mount Sinai), Gregory Crawford (Duke University), Stella Dracheva (Icahn School of Medicine at Mount Sinai), Peggy Farnham (University of Southern California), Mark Gerstein (Yale University), Daniel Geschwind (University of California, Los Angeles), Fernando Goes (Johns Hopkins University), Thomas M. Hyde (Lieber Institute for Brain Development), Andrew Jaffe (Lieber Institute for Brain Development), James A. Knowles (University of Southern California), Chunyu Liu (SUNY Upstate Medical University), Dalila Pinto (Icahn School of Medicine at Mount Sinai), Panos Roussos (Icahn School of Medicine at Mount Sinai), Stephan Sanders (University of California, San Francisco), Nenad Sestan (Yale University), Pamela Sklar (Icahn School of Medicine at Mount Sinai), Matthew State (University of California, San Francisco), Patrick Sullivan (University of North Carolina), Flora Vaccarino (Yale University), Daniel Weinberger (Lieber Institute for Brain Development), Sherman Weissman (Yale University), Kevin White (University of Chicago), Jeremy Willsey (University of California, San Francisco), and Peter Zandi (Johns Hopkins University).

SUPPLEMENTARY DATA

Supplementary data to this article can be found online.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

X.X. and M.L. received the funding, conceived the study, and interpreted the results. H.J.L., X.S., and L.Z. collected the data and conducted the primary analyses. C.Y.Z., L.W., W.L., Y.Y., and L.L. helped with data collection and analyses. X.X. and M.L. drafted the manuscript. All authors read and approved the final version of the manuscript.

REFERENCES

Adrien J. 2002. Neurobiological bases for the relation between sleep and depression. *Sleep Medicine Reviews*, **6**(5): 341−351.

Bagot RC, Cates HM, Purushothaman I, Lorsch ZS, Walker DM, Wang JS, et al. 2016. Circuit-wide transcriptional profiling reveals brain region-specific gene networks regulating depression susceptibility. *Neuron*, **90**(5): 969−983.

Bagot RC, Cates HM, Purushothaman I, Vialou V, Heller EA, Yieh L, et al. 2017. Ketamine and Imipramine reverse transcriptional signatures of susceptibility and induce resilience-specific gene expression profiles. *Biological Psychiatry*, **81**(4): 285−295.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**(15): 2114−2120.

Cahill KM, Huo ZG, Tseng GC, Logan RW, Seney ML. 2018. Improved identification of concordant and discordant gene expression signatures using an updated rank-rank hypergeometric overlap approach. *Scientific Reports*, **8**(1): 9588.

Cai X, Yang ZH, Li HJ, Xiao X, Li M, Chang H. 2020. A human-specific schizophrenia risk tandem repeat affects alternative splicing of a human-unique isoform *AS3MT*^{d2d3} and mushroom dendritic spine density. *Schizophrenia Bulletin*, doi: 10.1093/schbul/sbaa098.

Charlesworth JC, Curran JE, Johnson MP, Göring HH, Dyer TD, Diego VP, et al. 2010. Transcriptomic epidemiology of smoking: the effect of smoking on gene expression in lymphocytes. *BMC Medical Genomics*, **3**: 29.

Cheng Y, Sun M, Chen L, Li YJ, Lin L, Yao B, et al. 2018. Ten-eleven translocation proteins modulate the response to environmental stress in mice. *Cell Reports*, **25**(11): 3194−3203.

Covington III HE, Lobo MK, Maze I, Vialou V, Hyman JM, Zaman S, et al. 2010. Antidepressant effect of optogenetic stimulation of the medial prefrontal cortex. *Journal of Neuroscience*, **30**(48): 16082−16090.

Curran T, Franza BR Jr. 1988. Fos and Jun: the AP-1 connection. *Cell*, **55**(3): 395−397.

Donahue RJ, Muschamp JW, Russo SJ, Nestler EJ, Carlezon WA Jr. 2014. Effects of striatal ΔFosB overexpression and ketamine on social defeat stress-induced anhedonia in mice. *Biological Psychiatry*, **76**(7): 550−558.

Duman RS, Aghajanian GK. 2012. Synaptic dysfunction in depression: potential therapeutic targets. *Science*, **338**(6103): 68−72.

Duman RS, Aghajanian GK, Sanacora G, Krystal JH. 2016. Synaptic plasticity and depression: new insights from stress and rapid-acting antidepressants. *Nature Medicine*, **22**(3): 238−249.

Edwards SL, Beesley J, French JD, Dunning AM. 2013. Beyond GWASs: illuminating the dark road from association to function. *The American Journal of Human Genetics*, **93**(5): 779−797.

Fleischmann A, Hvalby O, Jensen V, Strekalova T, Zacher C, Layer LE, et al. 2003. Impaired long-term memory and NR2A-type NMDA receptor-dependent synaptic plasticity in mice lacking c-Fos in the CNS. *Journal of Neuroscience*, **23**(27): 9116−9122.

Forrest MP, Parnell E, Penzes P. 2018. Dendritic structural plasticity and neuropsychiatric disease. *Nature Reviews Neuroscience*, **19**(4): 215−234.

French JD, Edwards SL. 2020. The role of noncoding variants in heritable disease. *Trends in Genetics*, doi: 10.1016/j.tig.2020.07.004.

Gallo FT, Katche C, Morici JF, Medina JH, Weisstaub NV. 2018. Immediate early genes, memory and psychiatric disorders: focus on c-Fos, Egr1 and Arc. *Frontiers in Behavioral Neuroscience*, **12**: 79.

Gandal MJ, Haney JR, Parikshak NN, Leppa V, Ramaswami G, Hartl C, et al. 2018a. Shared molecular neuropathology across major psychiatric disorders parallels polygenic overlap. *Science*, **359**(6376): 693−697.

Gandal MJ, Zhang P, Hadjimichael E, Walker RL, Chen C, Liu S, et al. 2018b. Transcriptome-wide isoform-level dysregulation in ASD, schizophrenia, and bipolar disorder. *Science*, **362**(6420): eaat8127.

Gerring ZF, Gamazon ER, Derks EM. 2019. A gene co-expression network-based analysis of multiple brain tissues reveals novel genes and molecular pathways underlying major depression. *PLoS Genetics*, **15**(7): e1008245.

Gotlib IH, Joormann J. 2010. Cognition and depression: current status and future directions. *Annual Review of Clinical Psychology*, **6**: 285−312.

Guzowski JF, McNaughton BL, Barnes CA, Worley PF. 1999. Environment-specific expression of the immediate-early gene *Arc* in hippocampal neuronal ensembles. *Nature Neuroscience*, **2**(12): 1120−1124.

Hasin DS, Goodwin RD, Stinson FS, Grant BF. 2005. Epidemiology of major depressive disorder: results from the National Epidemiologic Survey on Alcoholism and Related Conditions. *Archices of General Psychiatry*, **62**(10): 1097−1106.

Hennessy MB, Deak T, Schiml-Webb PA. 2010. Early attachment-figure separation and increased risk for later depression: potential mediation by proinflammatory processes. *Neuroscience & Biobehavioral Reviews*, **34**(6): 782−790.

Hodes GE, Pfau ML, Purushothaman I, Ahn HF, Golden SA, Christoffel DJ, et al. 2015. Sex differences in nucleus accumbens transcriptome profiles associated with susceptibility versus resilience to subchronic variable stress. *Journal of Neuroscience*, **35**(50): 16362−16376.

Hodges TE, Green MR, Simone JJ, McCormick CM. 2014. Effects of social context on endocrine function and Zif268 expression in response to an acute stressor in adolescent and adult rats. *International Journal of Developmental Neuroscience*, **35**(1): 25−34.

Howard DM, Adams MJ, Clarke TK, Hafferty JD, Gibson J, Shirali M, et al. 2019. Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nature Neuroscience*, **22**(3): 343−352.

Jansen R, Batista S, Brooks AI, Tischfield JA, Willemsen G, van Grootheest G, et al. 2014. Sex differences in the human peripheral blood transcriptome. *BMC Genomics*, **15**: 33.

Jansen R, Penninx BWJH, Madar V, Xia K, Milaneschi Y, Hottenga JJ, et al. 2016. Gene expression in major depressive disorder. *Molecular Psychiatry*, **21**(3): 339−347.

Kang HJ, Voleti B, Hajszan T, Rajkowska G, Stockmeier CA, Licznerski P, et al. 2012. Decreased expression of synapse-related genes and loss of

synapses in major depressive disorder. *Nature Medicine*, **18**(9): 1413−1417.
Kanton S, Boyle MJ, He ZS, Santel M, Weigert A, Sanchís-Calleja F, et al. 2019. Organoid single-cell genomic atlas uncovers human-specific features of brain development. *Nature*, **574**(7778): 418−422.
Kim D, Langmead B, Salzberg SL. 2015. HISAT: a fast spliced aligner with low memory requirements. *Nature Methods*, **12**(4): 357−360.
Kim S, Hwang Y, Webster MJ, Lee D. 2016. Differential activation of immune/inflammatory response-related co-expression modules in the hippocampus across the major psychiatric disorders. *Molecular Psychiatry*, **21**(3): 376−385.
Labonté B, Engmann O, Purushothaman I, Menard C, Wang JS, Tan CF, et al. 2017. Sex-specific transcriptional signatures in human depression. *Nature Medicine*, **23**(9): 1102−1111.
Langfelder P, Horvath S. 2008. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*, **9**: 559.
Laoutidis ZG, Luckhaus C. 2015. The Liptak-Stouffer test for meta-analyses. *Biological Psychiatry*, **77**(1): e1−e2.
Leday GGR, Vértes PE, Richardson S, Greene JR, Regan T, Khan S, et al. 2018. Replicable and coupled changes in innate and adaptive immune gene expression in two case-control studies of blood microarrays in major depressive disorder. *Biological Psychiatry*, **83**(1): 70−80.
Leek JT, Johnson WE, Parker HS, Fertig EJ, Jaffe AE, Storey JD, et al. 2020. sva: Surrogate Variable Analysis. R package version 3.34.0.
Le-Niculescu H, Kurian SM, Yehyawi N, Dike C, Patel SD, Edenberg HJ, et al. 2009. Identifying blood biomarkers for mood disorders using convergent functional genomics. *Molecular Psychiatry*, **14**(2): 156−174.
Li HJ, Chang H, Song XQ, Liu WP, Li LY, Wang L, et al. 2019. Integrative analyses of major histocompatibility complex loci in the genome-wide association studies of major depressive disorder. *Neuropsychopharmacology*, **44**(9): 1552−1561.
Li HJ, Qu N, Hui L, Cai X, Zhang CY, Zhong BL, et al. 2020a. Further confirmation of *netrin 1 receptor* (*DCC*) as a depression risk gene *via* integrations of multi-omics data. *Translational Psychiatry*, **10**(1): 98.
Li JZ, Bunney BG, Meng F, Hagenauer MH, Walsh DM, Vawter MP, et al. 2013. Circadian patterns of gene expression in the human brain and disruption in major depressive disorder. *Proceedings of the National Academy of Sciences of the United States of America*, **110**(24): 9950−9955.
Li M, Jaffe AE, Straub RE, Tao R, Shin JH, Wang YH, et al. 2016. A human-specific *AS3MT* isoform and *BORCS7* are molecular risk factors in the 10q24.32 schizophrenia-associated locus. *Nature Medicine*, **22**(6): 649−656.
Li WQ, Cai X, Li HJ, Song M, Zhang CY, Yang YF, et al. 2020b. Independent replications and integrative analyses confirm *TRANK1* as a susceptibility gene for bipolar disorder. *Neuropsychopharmacology*, doi: 10.1038/s41386-020-00788-4.
Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, **30**(7): 923−930.
Liu WP, Li WQ, Cai X, Yang ZH, Li HJ, Su X, et al. 2020. Identification of a functional human-unique 351-bp *Alu* insertion polymorphism associated with major depressive disorder in the 1p31.1 GWAS risk loci. *Neuropsychopharmacology*, **45**(7): 1196−1206.
Liu WP, Yan H, Zhou DY, Cai X, Zhang YYN, Li SY, et al. 2019. The depression GWAS risk allele predicts smaller cerebellar gray matter volume and reduced *SIRT1* mRNA expression in Chinese population. *Translational Psychiatry*, **9**(1): 333.
Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, **15**(12): 550.
Lu QB, Sun JF, Yang QY, Cai WW, Xia MQ, Wu FF, et al. 2020. Magnetic brain stimulation using iron oxide nanoparticle-mediated selective treatment of the left prelimbic cortex as a novel strategy to rapidly improve depressive-like symptoms in mice. *Zoological Research*, **41**(4): 381−394.
Ma L, Semick SA, Chen Q, Li C, Tao R, Price AJ, et al. 2020. Schizophrenia risk variants influence multiple classes of transcripts of sorting nexin 19 (*SNX19*). *Molecular Psychiatry*, **25**(4): 831−843.
Malhi GS, Mann JJ. 2018. Depression. *The Lancet*, **392**(10161): 2299−2312.
Millan MJ, Agid Y, Brüne M, Bullmore ET, Carter CS, Clayton NS, et al. 2012. Cognitive dysfunction in psychiatric disorders: characteristics, causes and the quest for improved therapy. *Nature Reviews Drug Discovery*, **11**(2): 141−168.
Minatohara K, Akiyoshi M, Okuno H. 2015. Role of immediate-early genes in synaptic plasticity and neuronal ensembles underlying the memory trace. *Frontiers in Molecular Neuroscience*, **8**: 78.
Mostafavi S, Battle A, Zhu X, Potash JB, Weissman MM, Shi J, et al. 2014. Type I interferon signaling genes in recurrent major depression: increased expression detected by whole-blood RNA sequencing. *Molecular Psychiatry*, **19**(12): 1267−1274.
Nutt D, Wilson S, Paterson L. 2008. Sleep disorders as core symptoms of depression. *Dialogues in Clinical Neuroscience*, **10**(3): 329−336.
Ota KT, Liu RJ, Voleti B, Maldonado-Aviles JG, Duric V, Iwata M, et al. 2014. REDD1 is essential for stress-induced synaptic loss and depressive behavior. *Nature Medicine*, **20**(5): 531−535.
Otte C, Gold SM, Penninx BW, Pariante CM, Etkin A, Fava M, et al. 2016. Major depressive disorder. *Nature Reviews Disease Primers*, **2**: 16065.
Pantazatos SP, Huang YY, Rosoklija GB, Dwork AJ, Arango V, Mann JJ. 2017. Whole-transcriptome brain expression and exon-usage profiling in major depression and suicide: evidence for altered glial, endothelial and ATPase activity. *Molecular Psychiatry*, **22**(5): 760−773.
Penzes P, Cahill ME, Jones KA, VanLeeuwen JE, Woolfrey KM. 2011. Dendritic spine pathology in neuropsychiatric disorders. *Nature Neuroscience*, **14**(3): 285−293.
Plaisier SB, Taschereau R, Wong JA, Graeber TG. 2010. Rank-rank hypergeometric overlap: identification of statistically significant overlap between gene-expression signatures. *Nucleic Acids Research*, **38**(17): e169.
Ramaker RC, Bowling KM, Lasseigne BN, Hagenauer MH, Hardigan AA, Davis NS, et al. 2017. Post-mortem molecular profiling of three psychiatric disorders. *Genome Medicine*, **9**(1): 72.
Riemann D, Berger M, Voderholzer U. 2001. Sleep and depression-results from psychobiological studies: an overview. *Biological Psychology*, **57**(1-3): 67−103.
Ritchie ME, Phipson B, Wu D, Hu YF, Law CW, Shi W, et al. 2015. *Limma* powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*, **43**(7): e47.
Russo SJ, Nestler EJ. 2013. The brain reward circuitry in mood disorders.

Nature Reviews Neuroscience, **14**(9): 609−625.

Scarpa JR, Fatma M, Loh YHE, Traore SR, Stefan T, Chen TH, et al. 2020. Shared transcriptional signatures in major depressive disorder and mouse chronic stress models. *Biological Psychiatry*, **88**(2): 159−168.

Scarpa JR, Jiang P, Gao VD, Fitzpatrick K, Millstein J, Olker C, et al. 2018. Cross-species systems analysis identifies gene networks differentially altered by sleep loss and depression. *Science Advances*, **4**(7): eaat1294.

Seney ML, Huo ZG, Cahill K, French L, Puralewski R, Zhang J, et al. 2018. Opposite molecular signatures of depression in men and women. *Biological Psychiatry*, **84**(1): 18−27.

Sequeira A, Klempan T, Canetti L, ffrench-Mullen J, Benkelfat C, Rouleau GA, et al. 2007. Patterns of gene expression in the limbic system of suicides with and without major depression. *Molecular Psychiatry*, **12**(7): 640−655.

Smith GD, Ebrahim S. 2003. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease?. *International Journal of Epidemiology*, **32**(1): 1−22.

Smith GD, Hemani G. 2014. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Human Molecular Genetics*, **23**(R1): R89−R98.

Solé B, Jiménez E, Martinez-Aran A, Vieta E. 2015. Cognition as a target in major depression: new developments. *European Neuropsychopharmacology*, **25**(2): 231−247.

Stein JL, de la Torre-Ubieta L, Tian Y, Parikshak NN, Hernández IA, Marchetto MC, et al. 2014. A quantitative framework to evaluate modeling of cortical development by neural stem cells. *Neuron*, **83**(1): 69−86.

Sullivan PF, Neale MC, Kendler KS. 2000. Genetic epidemiology of major depression: review and meta-analysis. *American Journal of Psychiatry*, **157**(10): 1552−1562.

Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, et al. 2017. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Research*, **45**(D1): D362−D368.

Wang SS, Kamphuis W, Huitinga I, Zhou JN, Swaab DF. 2008. Gene expression analysis in the human hypothalamus in depression by laser microdissection and real-time PCR: the presence of multiple receptor imbalances. *Molecular Psychiatry*, **13**(8): 786−799.

Westra HJ, Peters MJ, Esko T, Yaghootkar H, Schurmann C, Kettunen J, et al. 2013. Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nature Genetics*, **45**(10): 1238−1243.

Wittenberg GM, Greene J, Vértes PE, Drevets WC, Bullmore ET. 2020. Major depressive disorder is associated with differential expression of innate immune and neutrophil-related gene networks in peripheral blood: a quantitative review of whole-genome transcriptional data from case-control studies. *Biological Psychiatry*, doi: 10.1016/j.biopsych.2020.05.006.

Wray NR, Ripke S, Mattheisen M, Trzaskowski M, Byrne EM, Abdellaoui A, et al. 2018. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nature Genetics*, **50**(5): 668−681.

Wright FA, Sullivan PF, Brooks AI, Zou F, Sun W, Xia K, et al. 2014. Heritability and genomics of gene expression in peripheral blood. *Nature Genetics*, **46**(5): 430−437.

Wu Y, Zeng J, Zhang FT, Zhu ZH, Qi T, Zheng ZL, et al. 2018. Integrative analysis of omics summary data reveals putative mechanisms underlying complex traits. *Nature Communications*, **9**(1): 918.

Xiao X, Zheng FF, Chang H, Ma YN, Yao YG, Luo XJ, et al. 2018. The gene encoding protocadherin 9 (*PCDH9*), a novel risk factor for major depressive disorder. *Neuropsychopharmacology*, **43**(5): 1128−1137.

Xu L, Chen Y, Mayakonda A, Koh L, Chong YK, Buckley DL, et al. 2018. Targetable BET proteins- and E2F1-dependent transcriptional program maintains the malignancy of glioblastoma. *Proceedings of the National Academy of Sciences of the United States of America*, **115**(22): E5086−E5095.

Yang ZH, Cai X, Qu N, Zhao LJ, Zhong BL, Zhang SF, et al. 2020a. Identification of a functional 339 bp Alu insertion polymorphism in the schizophrenia-associated locus at 10q24.32. *Zoological Research*, **41**(1): 84−89.

Yang ZH, Zhou DY, Li HJ, Cai X, Liu WP, Wang L, et al. 2020b. The genome-wide risk alleles for psychiatric disorders at 3p21.1 show convergent effects on mRNA expression, cognitive function, and mushroom dendritic spine. *Molecular Psychiatry*, **25**(1): 48−66.

Yasoshima Y, Sako N, Senba E, Yamamoto T. 2006. Acute suppression, but not chronic genetic deficiency, of c-*fos* gene expression impairs long-term memory in aversive taste learning. *Proceedings of the National Academy of Sciences of the United States of America*, **103**(18): 7106−7111.

Yu GC, Li F, Qin YD, Bo XC, Wu YB, Wang SQ. 2010. GOSemSim: an R package for measuring semantic similarity among GO terms and gene products. *Bioinformatics*, **26**(7): 976−978.

Yu GC, Wang LG, Han YY, He QY. 2012. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*, **16**(5): 284−287.

Zhang B, Horvath S. 2005. A general framework for weighted gene co-expression network analysis. *Statistical Applications in Genetics and Molecular Biology*, **4**: 17.

Zhong SJ, Zhang S, Fan XY, Wu Q, Yan LY, Dong J, et al. 2018. A single-cell RNA-seq survey of the developmental landscape of the human prefrontal cortex. *Nature*, **555**(7697): 524−528.

Zhu SS, Cordner ZA, Xiong JL, Chiu CT, Artola A, Zuo YN, et al. 2017. Genetic disruption of ankyrin-G in adult mouse forebrain causes cortical synapse alteration and behavior reminiscent of bipolar disorder. *Proceedings of the National Academy of Sciences of the United States of America*, **114**(39): 10479−10484.

Zhu ZH, Zhang FT, Hu H, Bakshi A, Robinson MR, Powell JE, et al. 2016. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nature Genetics*, **48**(5): 481−487.