Molecular Psychiatry (2016) 21, 339–347

www.nature.com/mp

ORIGINAL ARTICLE

Gene expression in major depressive disorder

R Jansen[1,14], BWJH Penninx[1,14], V Madar[2], K Xia[3], Y Milaneschi[1], JJ Hottenga[4], AR Hammerschlag[5], A Beekman[1], N van der Wee[6], JH Smit[1], AI Brooks[7], J Tischfield[7], D Posthuma[5,8], R Schoevers[9], G van Grootheest[1], G Willemsen[4], EJ de Geus[4], DI Boomsma[4], FA Wright[2,10,11], F Zou[2], W Sun[12] and PF Sullivan[12,13]

The search for genetic variants underlying major depressive disorder (MDD) has not yet provided firm leads to its underlying molecular biology. A complementary approach is to study gene expression in relation to MDD. We measured gene expression in peripheral blood from 1848 subjects from The Netherlands Study of Depression and Anxiety. Subjects were divided into current MDD ($N = 882$), remitted MDD ($N = 635$) and control ($N = 331$) groups. MDD status and gene expression were measured again 2 years later in 414 subjects. The strongest gene expression differences were between the current MDD and control groups (129 genes at false-discovery rate, FDR < 0.1). Gene expression differences across MDD status were largely unrelated to antidepressant use, inflammatory status and blood cell counts. Genes associated with MDD were enriched for interleukin-6 (IL-6)-signaling and natural killer (NK) cell pathways. We identified 13 gene expression clusters with specific clusters enriched for genes involved in NK cell activation (downregulated in current MDD, FDR $= 5.8 \times 10^{-5}$) and IL-6 pathways (upregulated in current MDD, FDR $= 3.2 \times 10^{-3}$). Longitudinal analyses largely confirmed results observed in the cross-sectional data. Comparisons of gene expression results to the Psychiatric Genomics Consortium (PGC) MDD genome-wide association study results revealed overlap with *DVL3*. In conclusion, multiple gene expression associations with MDD were identified and suggest a measurable impact of current MDD state on gene expression. Identified genes and gene clusters are enriched with immune pathways previously associated with the etiology of MDD, in line with the immune suppression and immune activation hypothesis of MDD.

Molecular Psychiatry (2016) **21,** 339–347; doi:10.1038/mp.2015.57; published online 26 May 2015

INTRODUCTION

Genome-wide association studies (GWAS) have identified associations for many psychiatric disorders, such as schizophrenia,[1] autism[2] and bipolar disorder.[3] Major depressive disorder (MDD) has modest heritability (31–42%),[4] but large GWAS for MDD have yet to identify any genome-wide significant associations.[5] This lack of success using GWAS could be overcome by increasing sample sizes, studying the heterogeneity of MDD, evaluating gene–environment interactions or epistatic effects.[6]

As a complementary approach, biological pathways involved in MDD may also be identified by studying gene expression. The transcriptome is partly regulated by the genome, as indicated by twin and expression quantitative trait locus (eQTL) studies[7,8] but also by other factors such as sex,[9] smoking,[10] psychosocial factors[11] and age.[12] Therefore, gene expression–MDD associations may reflect genetic and non-genetic effects.

All transcriptomic studies of human diseases face the issue of choice of tissue. This is particularly important for brain diseases (for example, in contrast to many cancers) given the relative inaccessibility of the critical tissue. Psychiatric researchers thus face a set of hard choices with multiple practical advantages and disadvantages. It can be argued that brain samples are optimal for transcriptomic studies of MDD as it is widely believed that the locus of MDD is in the central nervous system. However, these advantages are offset by numerous practical issues due to the fact that this is not a controlled experiment: there are few high-quality postmortem samples from people with MDD in the world (resulting in low statistical power), longitudinal studies of the same person are impossible (that is, when depressed and when well), the choice of the brain region(s) to study remain unclear and sampling adults postmortem could well miss a developmental process, and postmortem studies are bedeviled by technical issues (the manner of death has profound impact on gene expression in brain, prolonged postmortem intervals are typical, dissections are not necessarily standardized across sites and across time, and few postmortem studies have made serious attempts to account for cellular heterogeneity) as well as a host of conceptual issues (a major issue is the appropriateness of controls give that MDD case characteristics can well give rise to brain expression changes—MDD is typified by the use of psychotropic medications, obesity, poorer health status, increased licit and illicit substance use and so on). In our view, limited statistical power is a major issue: the effects are likely to be subtle and undetectable unless the sample sizes are in the hundreds or thousands.

[1]Department of Psychiatry, VU University Medical Center, Neuroscience Campus Amsterdam, Amsterdam, The Netherlands; [2]Department of Biostatistics, University of North Carolina, Chapel Hill, NC, USA; [3]Department of Psychiatry, University of North Carolina, Chapel Hill, NC, USA; [4]Department of Biological Psychology, VU University Amsterdam, Neuroscience Campus Amsterdam, Amsterdam, The Netherlands; [5]Department of Complex Trait Genetics, VU University Amsterdam, Center for Neurogenomics and Cognitive Research, Neuroscience Campus Amsterdam, Amsterdam, The Netherlands; [6]Department of Psychiatry, Leiden University Medical Center, Leiden, The Netherlands; [7]Department of Genetics and the Human Genetics Institute, RUCDR Infinite Biologics, Rutgers University, New Brunswick, NJ, USA; [8]Department of Clinical Genetics, Neuroscience Campus Amsterdam, VU University Medical Center, Amsterdam, The Netherlands; [9]Department of Psychiatry, University Medical Center Groningen, Groningen, The Netherlands; [10]Department of Statistics, North Carolina State University, Raleigh, NC, USA; [11]Department of Biological Sciences, North Carolina State University, Raleigh, NC, USA; [12]Department of Genetics, University of North Carolina, Chapel Hill, NC, USA and [13]Department of Psychiatry, University of North Carolina, Chapel Hill, NC, USA. Correspondence: Dr R Jansen, Department of Psychiatry, Neuroscience Campus Amsterdam, Amsterdam, VU University Medical Center, A.J. Ernststraat 1187, Amsterdam 1081 HL, The Netherlands.
E-mail: Ri.Jansen@ggzingeest.nl
[14]These authors contributed equally to this work.
Received 9 October 2014; revised 26 March 2015; accepted 7 April 2015; published online 26 May 2015

The use of peripheral blood samples has multiple advantages: large sample sizes are achievable, RNA can be immediately stabilized, cellular heterogeneity is readily quantified, white blood cells can reach most parts of the body (that is, are a 'sentinel' tissue reflecting the overall state of an organism), longitudinal sampling is feasible so that a person can be their own control (that is, studying a person when depressed and when well) and gene expression in blood is neither perfectly correlated nor uncorrelated with that in many brain regions.[13–15] Moreover, the etiology of MDD has been associated with pathophysiological pathways not restricted to brain areas[16,17]—for example, inflammatory,[18] immune[19] and metabolic processes[20]—which can be captured by gene expression in blood.

Previous studies have identified associations between peripheral blood gene expression and MDD status (reviewed in Hepgul *et al.*[21]), but most studies have focused on candidate genes. There were some apparent consistent findings (alterations in inflammation, glucocorticoid receptor and neuroplasticity pathways) but only small sample sizes were used ($N < 181$ cases) and no meta-analysis was performed. A recent RNA-seq study in whole blood[22] ($N = 463$ cases) did not identify differential gene expression between subjects with lifetime MDD and controls, although enrichment of smaller *P*-values in the interferon-α pathway was noted. The lifetime MDD group that was studied was a mixture of individuals with current and remitted MDD, and this may have obscured MDD-related changes in gene expression, given that MDD state effects may be prominent.[21]

The goal of the present study was to identify associations between gene expression and MDD status using peripheral blood gene expression in 1848 subjects from The Netherlands Study of Depression and Anxiety (NESDA),[23] while accounting for multiple demographic and technical covariates. To our knowledge, this is the largest MDD gene expression study yet conducted. The sample included 882 subjects with current MDD, 635 subjects with remitted MDD and 331 controls, which enabled to distinguish trait and state effects. Pathway analysis of the identified genes and gene expression networks associated with MDD allowed to further interpret our findings. As a subset of 414 subjects also had gene expression and MDD status data after 2 years, longitudinal analyses were conducted to validate our findings. Finally, our results were meta-analyzed with the recent RNA-seq study in MDD cases and compared with results from the largest published Psychiatric Genomics Consortium MDD GWAS.

MATERIALS AND METHODS

Subjects and biological sampling

The research protocol was approved by the Ethical Committees of participating universities and all subjects provided written informed consent. Subjects were ascertained and sampled from NESDA,[23] a longitudinal cohort study evaluating the consequences of depressive and anxiety disorders.[24,25] Subjects were recruited from population-based samples and primary and secondary care settings. The baseline sample consisted of 2981 subjects (aged 18–65 years) and included healthy controls and individuals with a current or prior history of MDD and/or anxiety disorder. Exclusion criteria included a primary diagnosis of bipolar, psychotic, obsessive compulsive or severe addictive disorder and not being fluent in Dutch. After 2 years, a face-to-face follow-up assessment was conducted with a response of 87.1% ($N = 2596$). Peripheral blood samples were obtained at baseline and after 2 years. Blood was drawn in the morning (0830–0930 h) after an overnight fast. Heparinized whole blood was transferred into *PAX* gene blood RNA tubes (Qiagen, Valencia, CA, USA) within 60 min, incubated, and stored at − 20 °C. RNA was isolated using Qiagen kits.

MDD assessment and other demographic and clinical variables

The presence of lifetime DSM-IV MDD[26] was established using the Composite Interview Diagnostic Instrument (version 2.1).[27] Subjects had no lifetime history of MDD or anxiety disorders (controls, $N = 331$), remitted MDD (present in lifetime but not at the current interview, $N = 635$) or current MDD (within the 6 months prior to interview, $N = 882$). At the 2-year follow-up, the Composite Interview Diagnostic Instrument was repeated to evaluate the presence of MDD during and at the end of the follow-up period. Covariates used in gene expression models are primarily based on our earlier study[8] and include sex, age, body mass index, smoking status and red blood cell count (obtained at the baseline and 2-year follow-up interviews). To evaluate immune dysregulation, high-sensitivity C-reactive protein (CRP) was measured in duplicate by an ELISA based on purified protein and polyclonal anti-CRP antibodies (Dako, Glostrup, Denmark). Antidepressant use was based on drug container inspection of all medications used in the past month, classified according to the World Health Organization Anatomical Therapeutic Chemical classification, and included selective serotonin reuptake inhibitors (ATC code N06AB), serotonin–norepinephrine reuptake inhibitors (N06AX16, N06AX21), tricyclic antidepressants (N06AA) and tetracyclic antidepressants (N06AX03, N06AX05 and N06AX11).

Gene expression measurements

Gene expression in peripheral venous blood was assayed at RUCDR Infinite Biologics (http://www.rucdr.org/). Our prior paper contains full details.[8] Briefly, total RNA was extracted using automated methods, and RNA quality and quantity were assessed. Samples were randomized to plates (baseline and 2-year follow-up samples were randomized independently). Gene expression profiles were determined using Affymetrix U219 arrays (96-well format) and the GeneTitan System as per the manufacturer's protocol. After quality control (Supplementary Methods), 45 574 probesets targeting 18 417 genes were available for analysis, from 1848 baseline and 479 2-year follow-up samples (414 subjects were measured at both baseline and 2-year follow-up).

Gene expression as mediator of genetic associations with MDD status: eQTL analysis and GWAS comparisons

eQTL analysis was performed by Matrix eQTL[28] using gene expression and genotype data from NESDA subjects as reported previously[8] ($N = 1895$ from which 1848 overlap with the current sample used for the analyses of associations with MDD). Briefly, genotypes were coded as 0, 1 or 2 and for each single-nucleotide polymorphism (SNP)-probeset pair a linear regression model was fitted with probeset expression as dependent variable, and genotype as independent variable while controlling for multiple covariates. Matrix eQTL treats local ('cis') and distant ('trans') eQTLs separately for multiple testing correction to control the false-discovery rate (FDR). The cis SNPs are within a gene transcript ± 1 Mb. Gene-based analysis of the Psychiatric Genomics Consortium MDD GWAS results based on 9240 MDD cases and 9519 controls[5] (http://www.med.unc.edu/pgc/files/resultfiles/pgc.cross.mdd.zip) was conducted using VEGAS[29] (http://gump.qimr.edu.au/VEGAS) and JAG[30] (http://ctglab.nl/software), see Supplementary Methods.

Involvement of MDD-related genes in biological processes

To assess whether Gene Ontology Biological Process categories or canonical pathways were enriched among the identified MDD-related genes, we evaluate enrichment using Fisher's exact test (only using genes most strongly asociated with MDD) and Wilcoxon's tests (using all genes). Gene Ontology Biological Process categories were retrieved using the R (Vienna, Austria) package org.Hs.eg.db (2.10.1). Canonical pathways—KEGG (Kyoto Encyclopedia of Genes and Genomes), REACTOME, BioCarta and PID (Pathway Interaction Database)—were downloaded from GSEA (http://www.broadinstitute.org/gsea/msigdb/genesets.jsp?collection=CP). In total, 9902 Gene Ontology Biological Process and 1273 canonical pathways contained $\geqslant 1$ gene measured by Affymetrix U219 microarrays and were included in our analyses. The reference set used to perform Fisher's exact tests consisted of the genes measured by the U219 microarrays assigned to one or more Gene Ontology Biological Process categories or canonical pathways. Multiple comparisons were controlled using FDR < 0.1. The Wilcoxon test compares the distribution of the *P*-values corresponding to the genes in a category with the distribution of *P*-values from all other genes. Results were verified using permutation implemented in SAFE.[31]

Examination of gene expression clusters associated with MDD

The correlation structure of gene expression was examined using unsigned co-expression networks constructed using the Weighted Gene Co-expression Network Analysis package in R.[32] First, gene expression was

residualized using all covariates and MDD status. A single probe with highest mean expression per gene was selected to be included in the network analysis (18 417 genes). The correlation matrix was raised to the fifth power to meet the scale-free topology criterion and used for the average linkage hierarchical clustering algorithm. The resultant clustering tree is used to define the clusters. The minimum cluster size was set to 30 and the cut-off for merging clusters was set to 0.25. Each cluster is then characterized by its 'eigengene', the first principal component of the cluster expression data. The association between the clusters and MDD was computed with the same linear model as used for the probeset level expression data, replacing probeset gene expression by one of the eigengenes.

Longitudinal analyses

Associations between the change in MDD status and the change in gene expression were evaluated between subjects with clear MDD status change and controls: 137 subjects without MDD at both baseline and 2-year follow-up (consistent controls) and 146 subjects with current MDD at the first measurement and no current MDD at 2-year follow-up (recovered MDD). Gene expression data from these 283 subjects at baseline and 2-year follow-up were residualized with respect to the covariates also used in the cross-sectional analyses, including time (baseline vs 2-year follow-up) in the model. The difference in residualized gene expression between baseline and the 2-year follow-up was computed for every subject and for every probeset. For each probeset, this difference was used as a dependent variable in a linear model, with change in MDD status (consistent control vs recovered MDD) as an independent variable.

RESULTS

The NESDA sample had analyzable gene expression data from 1848 subjects (Table 1) including 882 subjects with current MDD, 635 with remitted MDD (that is, MDD in the past but not at the baseline interview) and 331 controls with no lifetime history of MDD at baseline. Many subject characteristics were consistent with prior MDD findings (for example, more females and current smokers among cases). Other personal characteristics and technical features were similar across MDD groups (for example, body mass index, CRP and red blood cell count.

Cross-sectional associations between MDD status and gene expression

To assess differential gene expression between the control, remitted and current MDD groups we fitted linear models for each of the 45 574 gene expression probesets, that included 16 empirically-selected covariates (Supplementary Table S1, Supplementary Methods). First, when remitted MDD and current MDD groups were pooled as a lifetime MDD group, 10 genes were differentially expressed at a liberal threshold (FDR < 0.1). Second, five genes were associated with MDD as a three-level (control, remitted MDD and current MDD) categorical variable (FDR < 0.1). In the three pairwise MDD group comparisons, we found 129 genes (142 probesets) differentially expressed between the control and current MDD groups (FDR < 0.1, Supplementary Table S2) but no probesets were differentially expressed between the control and remitted MDD groups or between remitted MDD and current MDD groups (FDR > 0.1). Thus, the strongest gene expression differences existed between current MDD and controls, suggesting MDD state-dependency of gene expression differences. Among the 119 genes differentially expressed between the control and current MDD groups the mean expression of the remitted MDD group is intermediate between controls and current MDD for 127 of 129 genes, and for 116 genes closer to the mean expression of the current MDD group than to the control group (Figure 1 and Supplementary Figure S1). Moreover, 18 genes were significant at FDR < 0.05 (Table 2) and the strongest association was found for *KLRD1* (FDR $= 1.6 \times 10^{-3}$, Supplementary Figure S2). The histogram of *P*-values from the comparison between controls and current MDD is strongly inflated towards zero (Supplementary Figure S3) suggesting the presence of many genes with expression changes associated with current MDD that were not detectable with the current sample size. For each of the 45 574 gene expression probesets we evaluated sex $\times$ MDD and age $\times$ MDD interaction terms, but no significant associations were found (FDR > 0.1).

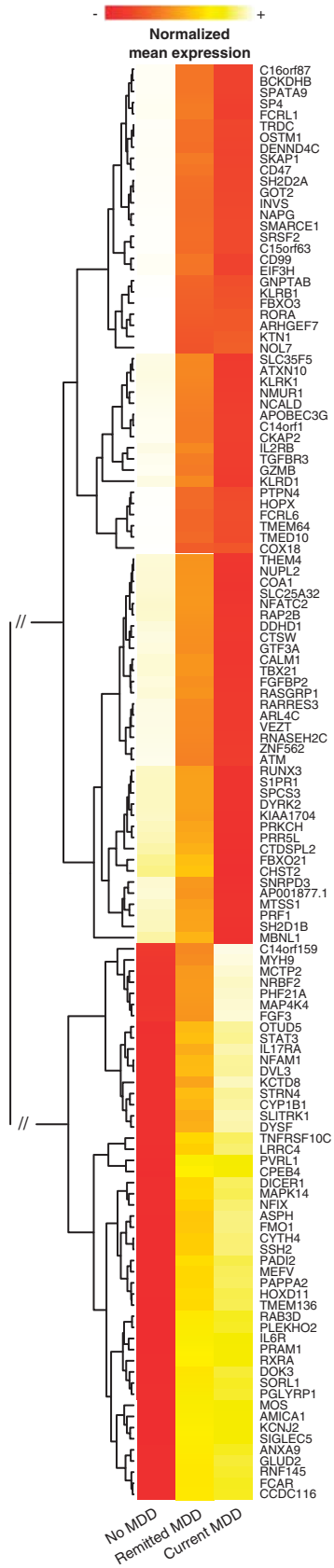
Evaluating potential confounders

Several modifications of the linear model described above (with MDD as three-level categorical variable) were made to assess the influence of confounders in the MDD-expression associations. CRP (a general indicator of systemic inflammatory status) has been shown to be slightly elevated in MDD.[33] When including CRP in the model, 127 of 129 genes were still differentially expressed between the control and current MDD groups (FDR < 0.1). The regression betas for current MDD vs controls from the models with and without CRP were highly correlated (Spearman's $\rho = 0.99$ and Spearman's $\rho = 0.99$ when restricting to the 142 probesets associated with current MDD). Delineating the effect of antidepressants on gene expression in a model for MDD is not straightforward in an observational study as antidepressant use is 'confounded-by-indication' and reflects more severe MDD. Addition of variables reflecting selective serotonin reuptake inhibitors, tricyclic antidepressants and serotonin–norepinephrine reuptake inhibitors use to the model showed that no genes were associated

Table 1. Baseline patient characteristics of the NESDA sample

Characteristic	*Current MDD*	*Remitted MDD*	*No MDD*	P-*value*
Subjects (%)	882 (47.7%)	635 (34.4%)	331 (17.9%)	
Female sex (%)	594 (67.3%)	448 (70.6%)	199 (60.1%)	0.005
Age in years, mean (s.d.)	40.6 (12.2)	43.0 (12.67)	42.6 (14.59)	0.0005
SSRI use (%)	268/882 (30.4%)	110/635 (17.3%)	NA	< 0.0001
TCA use (%)	38/882 (4.3%)	20/635 (3.1%)	NA	< 0.0001
SNRI use (%)	103/882 (11.7%)	25/635 (3.4%)	NA	< 0.0001
Current smoking (%)	407 (46.1%)	260 (40.9%)	90 (27.2%)	< 0.0001
Body mass index, mean (s.d.)	25.9 (5.43)	25.7 (4.87)	25.5 (4.67)	0.43
C-reactive protein (mg l^{-1})	3.08 (4.98)	2.89 (5.81)	2.51 (5.64)	0.26
Red blood count (million cells/microliter)	4.62 (0.41)	4.62 (0.39)	4.65 (0.44)	0.46
Hour of day at blood draw (hours)	8.2 (0.44)	8.2 (0.50)	8.2 (0.43)	0.29

Abbreviations: NA, not applicable; NESDA, The Netherlands Study of Depression and Anxiety; MDD, major depressive disorder; SNRI, serotonin–norepinephrine reuptake inhibitor; SSRI, selective serotonin reuptake inhibitor; TCA, tricyclic antidepressants.



with antidepressant use (FDR > 0.1). The models estimating current MDD effects were rerun excluding subjects using selective serotonin reuptake inhibitors ($N = 379$), tricyclic antidepressants ($N = 59$), or serotonin–norepinephrine reuptake inhibitors ($N = 128$). The regression betas for current MDD vs controls were highly correlated between the full model and the model excluding antidepressant users (Spearman's $\rho = 0.89$ and Spearman's $\rho = 0.84$ when restricting to the 142 probesets associated with current MDD). For 404 subjects with baseline data, gene expression, MDD status and blood cell counts (lymphocytes, neutrophils, monocytes, eosinophils and basophils) were measured two years later (10 subjects measured at 2-year follow-up did not have blood cell count data). The baseline analysis was repeated using the 2-year follow-up samples, with and without correcting for cell counts. Betas for current MDD vs controls from the models with cell counts were highly correlated with betas from the models without cell counts (Spearman's $\rho = 0.92$, using only the 142 probesets identified in the baseline analysis) suggesting no large impact of blood composition on our results. Overall, these additional analyses indicate that the observed current MDD–gene expression associations appear largely unrelated to immune status, antidepressant use or blood composition. However, as antidepressant use strongly correlates with MDD status, confounding effects of antidepressants cannot be excluded.

Involvement of MDD-related genes in biological processes

Enrichment analysis was performed using the *P*-values and corresponding gene ranking computed for the differences in

Table 2. *P*-values, FDR and betas of 18 genes differentially expressed between control and current MDD groups (FDR < 0.05)

Gene	*P*	*FDR*	*Beta*
KLRD1	3.42E–08	0.0016	−0.344
IL2RB	1.06E–06	0.0097	−0.291
GZMB	1.45E–06	0.0101	−0.303
LRRC4	1.61E–06	0.0101	0.294
CPEB4	1.77E–06	0.0101	0.293
CALM1	4.69E–06	0.0237	−0.242
TGFBR3	8.62E–06	0.0363	−0.283
PVRL1	1.08E–05	0.0363	0.271
SNRPD3	1.17E–05	0.0363	−0.265
APOBEC3G	1.21E–05	0.0363	−0.270
PTPN4	1.22E–05	0.0363	−0.268
KLRK1	1.31E–05	0.0363	−0.265
DYSF	1.40E–05	0.0363	0.254
TNFRSF10C	1.43E–05	0.0363	0.271
NCALD	1.43E–05	0.0363	−0.260
C14orf1	1.79E–05	0.0429	−0.268
COX18	1.92E–05	0.0438	−0.276
SSH2	2.19E–05	0.0475	0.253

Negative beta means downregulated expression in the current MDD group.

Figure 1. Mean expression across control, remitted MDD and current MDD groups. For the 129 genes differentially expressed between the current MDD and control groups, mean expression was color coded for the control, remitted MDD and current MDD groups (scaled per gene). Genes were ordered according to the dendrogram of a hierarchical cluster analysis using the mean expression across the three groups. Clearly the mean expression of the remitted MDD group is often between the mean expression of control and current MDD groups. MDD, major depressive disorder.

Table 3. Pathway analysis of genes associated with current MDD

Pathway	*Genes in gene set*	*Genes in pathway*	*Genes overlap*	P-*value*	*FDR*	*Overlapping genes*
Enrichment of genes upregulated in MDD						
REACTOME IL-6 signaling	28	10	2	0.00057	0.058	*IL6R, STAT3*
PID IL-6-mediated signaling events	28	47	3	0.00063	0.058	*IL6R, STAT3* and *MAPK14*
BIOCARTA map kinase inactivation of SMRT	28	11	2	0.00069	0.058	*RXRA, MAPK14*
BIOCARTA IL-6 signaling	28	22	2	0.0028	0.14	*IL6R, STAT3*
BIOCARTA ERBB2 in signal transduction	28	22	2	0.0028	0.14	*IL6R, STAT3*
Enrichment of genes downregulated in MDD						
KEGG natural killer cell mediated cytotoxicity	42	122	6	0.000046	0.014	*GZMB, KLRK1, PRF1, SH2D1B, KLRD1* and *NFATC2*
PID downstream signaling in naïve CD8+ T cells	42	63	4	0.00036	0.054	*IL2RB, GZMB, PRF1* and *NFATC2*
KEGG graft-versus-host disease	42	30	3	0.00055	0.056	*GZMB, PRF1* and *KLRD1*
BIOCARTA granzyme A mediated apoptosis pathway	42	11	2	0.0015	0.093	*GZMB, PRF1*
REACTOME ionotropic activity of kainate receptors	42	11	2	0.0015	0.093	*CALM1, NCALD*
BIOCARTA D4-GDI signaling pathway	42	12	2	0.0019	0.093	*GZMB, PRF1*
BIOCARTA CTL mediated immune response	42	13	2	0.0022	0.094	*GZMB, PRF1*

Abbreviations: FDR, false-discovery rate; IL-6, interleukin-6; MDD, major depressive disorder. Fisher's exact test *P*-values were computed for KEGG, REACTOME, BioCarta and PID Canonical pathways using the genes upregulated in MDD and the genes downregulated in MDD as gene sets.

gene expression between the current MDD and control groups. No enrichment was found for Gene Ontology Biological Process pathways. The 50 upregulated genes were enriched with five canonical pathways (FDR < 0.1, Table 3), including the REACTOME interleukin (IL-6)-signaling pathway (FDR $= 0.06$). The 79 genes downregulated in current MDD were enriched with seven canonical pathways (FDR < 0.1, Table 3), including the KEGG pathway 'Natural killer cell mediated cytotoxicity' (FDR $= 0.014$). Global statistics for enrichment were computed using all *P*-values (Wilcoxon's test for a difference in distribution between *P*-values from a gene category and all other *P*-values), but no categories were significant. Significance of gene category enrichment was confirmed by applying a permutation paradigm implemented in SAFE[31] software, which showed that the identified categories were also significant when using permutation based *P*-values.

Comparison with previous transcriptome studies in MDD cases

Mostafavi *et al.*[22] studied gene expression differences between 463 MDD cases and 459 controls using RNA-seq measurements. The *P*-values and directions of effects from this study were compared and meta-analyzed with our findings. From the 129 genes identified in our study, none was significant in the RNA-seq study (FDR > 0.1, corrected for 129 tests). From the 29 top genes associated with MDD in the RNA-seq study (FDR < 0.25), *SRSF5* ($P = 0.005$ in our study, $P = 0.000164$ in Mostafavi *et al.* study) and *PIPOX* ($P = 0.01$ in our study, $P = 0.0002$ in Mostafavi *et al.* study) were the most notable. When meta-analyzing the *P*-values from the two studies, (weighted *Z*-score method), 12 genes were differentially expressed at FDR < 0.1 (*CALM1, FCRL6, APOBEC3G, SRSF5, RAP2B, PIPOX, PRR5L, HAGHL, ARL4C, NMUR1, KLRD1* and *KCNJ2*, Supplementary Table S3). For six of these genes $P < 0.05$ in both studies and thus the meta-analysis results are not driven by one study. Several other studies have identified genome-wide gene expression differences between MDD cases and controls. Glahn *et al.*[34] (peripheral blood, $N = 215$ cases) reported 10 genes, Shelton *et al.*[35] (prefrontal cortex, $N = 14$ cases) found 30 genes and Garbett *et al.*[36] (fibroblast, $N = 16$ cases) identified 139 genes. None of these genes were significant in our data when taking directions of effects into account (FDR > 0.1, Supplementary Table S4).

Gene expression as a mediator of DNA associations with MDD status

For the 129 MDD associated genes in our study, gene-based GWAS *P*-values were computed using VEGAS and JAG based on the Psychiatric Genomics Consortium MDD GWAS results.[5] Both methods yielded a significant association between *DVL3* and MDD status (Vegas: FDR $= 6 \times 10^{-4}$; JAG: FDR $= 8 \times 10^{-4}$, corrected for 129 tests). SNPs in *DVL3* showed the strongest association with MDD in the Psychiatric Genomics Consortium results (rs1969253, $P = 4.8 \times 10^{-6}$).[5] We then identified genetic variants influencing the expression of the 129 current MDD associated genes. There were 76 genes with at least one local eQTL, and no distant eQTLs were detected (FDR < 0.05, Supplementary Table S5). For each gene we selected the SNP with the strongest association, and looked up corresponding Psychiatric Genomics Consortium MDD GWAS *P*-values, but none was notable (FDR > 0.1).

Gene expression networks associated with MDD status

To identify clusters of genes with correlated expression, we applied Weighted Gene Co-expression Network Analysis and identified 13 clusters. Four clusters were significantly different between controls and current MDD (FDR $< 3.2 \times 10^{-3}$, Supplementary Table S6). The most significant cluster (FDR $= 5.8 \times 10^{-5}$) contained 64 genes, all downregulated in current MDD, and 21 were among the 129 genes associated with current MDD. Supplementary Figure S4 shows that genes in the cluster strongly related to MDD have stronger and more connections to the other genes in the network, as compared to genes weakly related to MDD. These 64 genes were enriched for several canonical pathways including the PID pathway 'IL12-mediated signaling events' (FDR $= 1.4 \times 10^{-7}$) and the KEGG pathway 'Natural killer cell mediated cytotoxicity' (FDR $= 8.3 \times 10^{-7}$). Another cluster significantly associated with current MDD (FDR $= 3.2 \times 10^{-3}$) contains 447 genes, of which 411 were upregulated in current MDD (including 22 of the 129 genes associated with current MDD). This cluster was enriched with several canonical pathways including the REACTOME pathway 'Signaling by interleukins', (FDR $= 4.3 \times 10^{-7}$) and the PID pathway 'IL-6-mediated signaling events' (FDR $= 4.8 \times 10^{-5}$).

Table 4. The 129 genes associated with current MDD in the baseline cross-sectional analysis were validated in a longitudinal design.

Gene	P *longitudinal*	*Beta longitudinal*	P *cross sectional*	*Beta cross sectional*
RUNX3	4.70E-04	−0.166	1.82E-04	−0.224
IL6R	8.70E-04	0.126	1.96E-04	0.205
NFATC2	1.40E-03	−0.178	2.88E-04	−0.225
RASGRP1	2.00E-03	−0.201	1.04E-04	−0.230
ARHGEF7	2.30E-03	−0.188	1.51E-04	−0.223
RNF145	3.60E-03	0.205	1.65E-04	0.247
TNFRSF10C	3.80E-03	0.191	1.43E-05	0.271
SORL1	4.40E-03	0.094	2.58E-04	0.220
FCAR	6.10E-03	0.152	2.18E-04	0.241
SH2D2A	6.50E-03	−0.173	1.18E-04	−0.243
IL2RB	6.60E-03	−0.175	1.06E-06	−0.291
AMICA1	7.80E-03	0.182	4.34E-05	0.248
TMEM136	8.60E-03	0.152	2.08E-04	0.242
OSTM1	1.10E-02	−0.371	2.47E-04	−0.231
C14orf1	1.10E-02	-0.185	1.79E-05	−0.268
PADI2	1.10E-02	0.110	1.95E-04	0.248
ATXN10	1.10E-02	−0.162	3.88E-05	−0.257
VEZT	1.30E-02	−0.191	1.53E-04	−0.239
MAP4K4	1.40E-02	0.104	1.76E-04	0.225

This table shows the 19 genes significantly associated with MDD change (FDR < 0.1, corrected for 142 (= #probesets) tests). For these genes, direction of effects were concordant with the cross-sectional analysis (β longitudinal > 0: the mean expression decreased more in the recovered MDD group than in the consistent control group, β cross sectional > 0: genes up-regulated in the current MDD group).

Cross-sectional analyses in 2-year follow-up measurements

The cross-sectional analyses applied to the baseline measurements was repeated using 479 samples measured at 2-year follow-up (147 current MDD, 54 remitted MDD, 278 controls, 414 subjects measured at both baseline and 2-year follow-up). The comparison between gene expression of the current MDD and control groups did not result in any significant associations after correction (FDR > 0.1), which may be due to a lack of power. When only selecting the *P*-values of the 142 probesets identified in the baseline analyses, 3 genes were significant (FDR < 0.1, *PRF1, SH2D2A* and *TRDC,* corrected for 142 tests) and 18 genes at FDR < 0.15 (*KLRD1, GZMB, NCALD, NMUR1, PRF1, PRAM1, KCNJ2, ARL4C, SH2D2A, FGFBP2, TBX21, PRR5L, RUNX3, TRDC, DYSF, PADI2, DENND4C* and *RAB3D*). Also, the betas for current MDD effects of the 142 probesets were significantly correlated between the baseline and follow-up analyses (Spearman's $\rho = 0.6$, 111 probesets have the same direction of effect in baseline and follow-up analyses). Thus, multiple genes were associated with current MDD at both baseline and 2-year follow-up measurements, and for most genes the same trend was present at both time points. The smaller sample size at follow-up may explain why most genes were not significantly replicated at that measurement.

Differences in change of gene expression over 2 years

Genes associated with current MDD in the baseline cross-sectional analyses were validated in a longitudinal design. The difference in 2-year change in gene expression was compared between 137 subjects without MDD at both baseline and 2-year follow-up (consistent control), and 146 subjects with a clear MDD change as they had current MDD at the first measurement and no MDD at 2-year follow-up (recovered MDD). From the 129 genes identified in the cross-sectional study, 19 were significantly associated with MDD change (*IL2RB, TNFRSF10C, C14orf1, ATXN10, AMICA1, RASGRP1, SH2D2A, ARHGEF7, VEZT, RNF145, MAP4K4, RUNX3, PADI2, IL6R, TMEM136, FCAR, OSTM1, SORL1* and *NFATC2,* FDR < 0.1, corrected for 142 tests, Table 4), with the direction of effects concordant with the cross-sectional analyses: for genes upregulated in the current MDD group the mean expression decreased more in the recovered MDD group than in the consistent control group; for the genes downregulated in the current MDD group the mean expression increased more in the recovered MDD group than in the consistent control group. Also, for the cross-sectionally identified differential 142 probesets, the betas from the cross-sectional analyses (current MDD vs controls) were strongly correlated (Spearman's $\rho = 0.79$) with the betas from the longitudinal model (consistent control vs recovered MDD group), and 92% of the betas were in the same direction when comparing cross-sectional and longitudinal results.

DISCUSSION

We report the largest systematic transcriptome study of MDD yet conducted. Gene expression differences between MDD cases and controls were found to be numerous but tend to be small. We found evidence that gene expression associated with current MDD may implicate many genes or gene networks but several with effect sizes too small to be detected by our sample size. This is supported by the enrichment of the MDD associated genes with natural killer (NK) cell and IL-6 pathways, and by the network analysis that identified two clusters of genes that were significantly associated with current MDD. Interestingly, we found *DVL3* to be upregulated in current MDD, and note that *DVL3* had the strongest (but not genome-wide significant) associations in a mega GWAS for MDD.[5] The strongest effects were found in the comparison between current MDD and controls, and not in the comparison between the control and remitted MDD groups, which indicates that MDD state effects in gene expression are stronger than trait effects.

Down-regulation of NK cell, upregulation of IL-6 pathways

A bidirectional reinforcing communication between the brain and the immune system has been proposed as part of the etiology of MDD.[16,37,38] The most robust markers for immune suppression and immune activation in MDD are decreased NK cell numbers and cytotoxicity (NKCC)[39,40] and elevated IL-6, tumor necrosis factor-α (TNF-α) and CRP protein levels.[19,41,42] Meta-analyses of longitudinal studies suggest a causal pathway from inflammation to MDD.[33,43] Small gene expression studies ($N < 100$) have shown an

increased expression of *IL-6* and *TNF-α* in MDD.[44,45] We did not observe differential expression of *IL-6*, *CRP* or *TNF-α* in MDD, which may be due to low expression of these genes in blood. In addition, CRP is synthesized in the liver, and whole-blood *CRP* gene expression does not probably reflect blood CRP protein levels. However, the genes upregulated in MDD include one of the receptors of *TNF*-α (*TNFRSF10C*), *MAPK14*, the *IL-6* receptor and *STAT3,* and are therefore significantly enriched with genes in the PID IL-6 signaling pathway. Network analysis shows that the latter two genes are part of a large network (447 genes) significantly upregulated in MDD, which is enriched with 19 genes from the REACTOME pathway 'Signaling by interleukins', and 11 genes from the PID IL-6 signaling pathway. Thus, in line with the known increased levels of some inflammatory blood protein markers in MDD, we identified the upregulation of a large gene expression network enriched with inflammatory interleukin genes. This suggests that IL-6, CRP and TNF-α may be part of a larger protein network upregulated in MDD, which may not have been identified due to the limited capacity of measuring large amounts of proteins in MDD cases.

NK cells are a subset of lymphocytes that are important in innate immunity, defending against viruses and tumors. Serotonin increases NKCC[46] and long-term treatment with drugs affecting serotonergic mechanisms (selective serotonin reuptake inhibitors) augments NKCC[47] and may protect NK cells from oxidative damage.[48,49] NKCC increases with the resolution of depressive symptoms[50] that suggests a state rather than trait effect. To our knowledge, this is the first report of differential NKCC-related gene expression in MDD cases versus controls. These findings may reflect differences in NK cell counts, NKCC or other NK cell function. Six genes involved in NKCC were downregulated in MDD *(GZMB, KLRK1, PRF1, SH2D1B, KLRD1, NFATC2*) and are part of a larger downregulated cluster (64 genes) that contains nine genes associated with NKCC. Thus, our study confirms the immune suppression and immune activation hypothesis of MDD. This hypothesis was mainly based on observations at the protein and cell level—here we show that this hypothesis extends to the transcriptome level, most likely involving large clusters of genes. Dysregulation of these genes may be the cause or the consequence of MDD status, or be part of a bidirectional relationship between immune system and MDD.[41] The immune system genes we identified may serve as targets to further investigate this immune-MDD relationship.

DVL3 associated with MDD in genomic and transcriptomic studies

Gene-based comparisons of gene expression results to the largest published GWAS for MDD[5] to date revealed a suggestive overlap with *DVL3* (rs1969253 GWAS $P = 4.8 \times 10^{-6}$). *DVL3* is a part of the Wnt (Wingless-related integration site) signaling pathway, which is crucial in the regulation of hippocampal neurogenesis.[51] The role of the Wnt signaling pathway in mood disorders was recently reviewed.[52,53] *DVL3* transcripts have found to be decreased in individuals with MDD in the nucleus accumbens[54] and frontal regions[55] and are upregulated in leukocytes of individuals reporting social isolation.[56] Although the GWAS finding for *DVL3* SNPs did not reach genome-wide significance, the parallel findings in expression studies do suggest a tentative role for involvement of *DVL3* in the etiology of MDD. Although rs1969253 is not associated with *DVL3* expression in whole blood,[8] it is associated with *DVL3* average brain expression[57] ($P = 0.0028$, $N = 134$) and hippocampal expression[57] ($P = 0.009$, $N = 122$). We may thus hypothesize that rs1969253 is influencing MDD status via intermediate *DVL3* brain expression and the Wnt pathway. From this point of view our findings in blood may be caused by correlated blood and brain gene expression. The parallel discovery of *DVL3* in GWAS and our gene expression study shows how, in addition to GWAS, the exploration of additional molecular layers can contribute to understanding the genetics of MDD, similar to what was reported earlier in methylation[58] and RNA-seq[22] studies.

Meta-analysis of gene expression associations with MDD

Comparison of our results with the recent RNA-seq study in MDD[22] revealed 12 genes with consistent effects. We note in particular the findings for *SRSF5*, *CALM1* and *NMUR1*. Serine/arginine-rich splicing factor 5 (*SRSF5*) was highlighted in a recent RNA-seq study in bipolar[59] and is involved in circadian regulation. Calmodulin 1 (*CALM1*) has a role in neurotransmission and calmodulin-related gene expression is altered in lateral habenula[60] and frontal cortex[61,62] of MDD cases. *NMUR1* is a receptor of Neuromedin-U, a neuropeptide associated with anxiety and depression-like behavior in mice.[63,64] Follow-up work will be needed to further confirm the role of these genes in MDD.

Longitudinal analyses indicates reversibility of MDD effects in gene expression

For the 129 genes identified in cross-sectional analyses, a substantial proportion (15%) showed significant reversibility in expression after 2 years in the MDD group who remitted from their earlier episode. However, for several other genes we could not illustrate such a change over time. This could be caused by the fact that the power of our longitudinal analyses was limited due to the rather small sample size. In addition, gene expression patterns may not change in a very large extent as remitted patients still experience more (sub threshold) symptoms as compared with healthy controls. Also, MDD trait effects in gene expression will not reverse at all. The fact that not all baseline gene expression differences may revert when a MDD patient remits, is in line with the cross-sectional analyses in which the remitted MDD group had gene expression intermediate between the control and current MDD groups for most genes associated with current MDD. These findings, together with the lack of DNA variants underlying the gene expression–MDD associations indicate that MDD state effects in gene expression are more prominent than trait effects, as suggested earlier.[21] Our results do indicate that gene expression studies seem most informative if conducted in currently symptomatic patients. The very important implication of the found reversibility of gene expression profiles is that it points to potential modification by interventions. Consequently, it seems very worthwhile for future research studies to examine gene expression patterns as outcome parameters and examine reversibility of gene expression patterns after remittance.

Strengths and limitations

We report the largest study of gene expression in MDD yet conducted, controlling for various potential confounders and demographic covariates. The distinction between remitted and current MDD groups allowed to distinguish state and trait effects. Longitudinal analyses enabled to confirm some of the observed state effects, by indicating that for several identified genes, expression patterns indeed reversed over 2 years of time in the group of MDD remitted patients. Despite the large sample size, our findings were not highly significant, especially in longitudinal analyses. Also, due to the strong correlation between MDD status and antidepressant use, confounding effects of antidepressants may be present, although our analysis suggests that antidepressant use has no additional impact on gene expression and does not perturb associations between gene expression and MDD. Although meta-analysis with a recent RNA-seq study identified several significant genes, most genes were only significant in one study, and findings from other studies in MDD cases were not replicated. Future larger sample sizes are needed to confirm our findings. Nevertheless, we identified multiple associations between gene expression (clusters) and MDD status. Furthermore,

as the results of pathway analysis of the MDD-related genes and gene clusters are in line with the immune suppression and immune activation hypothesis of MDD, the identified immune genes and immune gene clusters should be targets for future research.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

Gene expression data were funded by the US National Institute of Mental Health (RC2 MH089951) as a part of the American Recovery and Reinvestment Act of 2009. The Netherlands Study of Depression and Anxiety (NESDA) and the Netherlands Twin Register (NTR) were funded by The Netherlands Organization for Scientific Research (MagW/ZonMW grants 904-61-090, 985-10-002,904-61-193,480-04-004, 400-05-717, 912-100-20; Spinozapremie 56-464-14192; Geestkracht program grant 10-000-1002); the Center for Medical Systems Biology (NWO Genomics), Biobanking and Biomolecular Resources Research Infrastructure (BBMRI-NL), VU University's Institutes for Health and Care Research and Neuroscience Campus Amsterdam, NBIC/BioAssist/RK (2008.024); the European Science Foundation (EU/QLRT-2001-01254); the European Community's Seventh Framework Program (FP7/2007-2013); ENGAGE (HEALTH-F4-2007-201413); and the European Science Council (ERC, 230374). Transport, extraction, cDNA preparation and generation of microarray data for the NTR samples were carried out under a supplement to the NIMH Center for Collaborative Genomics Research on Mental Disorders (U24 MH068457, Principal Investigator JT). RJ was supported by the Biobank-based Integrative Omics Study (BIOS) consortium, which is funded by the Biobanking and Biomolecular Research Infrastructure (BBMRI-NL, NWO project 184.021.007). Gene expression data used for this study are available at dbGaP, accession number phs000486.v1.p1 (http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id = phs000486.v1.p1).

REFERENCES

1 Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 2014; **511**: 421–427.
2 Gaugler T, Klei L, Sanders SJ, Bodea CA, Goldberg AP, Lee AB *et al.* Most genetic risk for autism resides with common variation. *Nat Genet* 2014; **46**: 881–885.
3 Nurnberger JI Jr, Koller DL, Jung J, Edenberg HJ, Foroud T, Guella I *et al.* Identification of pathways for bipolar disorder: a meta-analysis. *JAMA Psychiatry* 2014; **71**: 657–664.
4 Sullivan PF, Neale MC, Kendler KS. Genetic epidemiology of major depression: review and meta-analysis. *Am J Psychiatry* 2000; **157**: 1552–1562.
5 Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium, Ripke S, Wray NR, Lewis CM, Hamilton SP, Weissman MM *et al.* A mega-analysis of genome-wide association studies for major depressive disorder. *Mol Psychiatry* 2013; **18**: 497–511.
6 Levinson DF, Mostafavi S, Milaneschi Y, Rivera M, Ripke S, Wray NR *et al.* Genetic studies of major depressive disorder: why are there no GWAS findings, and what can we do about it. in press.
7 Westra H-J, Peters MJ, Esko T, Yaghootkar H, Schurmann C, Kettunen J *et al.* Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat Genet* 2013; **45**: 1238–1243.
8 Wright FA, Sullivan PF, Brooks AI, Zou F, Sun W, Xia K *et al.* Heritability and genomics of gene expression in peripheral blood. *Nat Genet* 2014; **46**: 430–437.
9 Jansen R, Batista S, Brooks AI, Tischfield JA, Willemsen G, van Grootheest G *et al.* Sex differences in the human peripheral blood transcriptome. *BMC Genomics* 2014; **15**: 33.
10 Charlesworth JC, Curran JE, Johnson MP, Göring HH, Dyer TD, Diego VP *et al.* Transcriptomic epidemiology of smoking: the effect of smoking on gene expression in lymphocytes. *BMC Med Genomics* 2010; **3**: 29.
11 Fredrickson BL, Grewen KM, Coffey KA, Algoe SB, Firestine AM, Arevalo JMG *et al.* A functional genomic perspective on human well-being. *Proc Natl Acad Sci USA* 2013; **110**: 13684–13689.
12 Van den Akker EB, Passtoors WM, Jansen R, van Zwet EW, Goeman JJ, Hulsman M *et al.* Meta-analysis on blood transcriptomic studies identifies consistently coexpressed protein-protein interaction modules as robust markers of human aging. *Aging Cell* 2014; **13**: 216–225.
13 Sullivan PF, Fan C, Perou CM. Evaluating the comparability of gene expression in blood and brain. *Am J Med Genet Neuropsychiatr Genet* 2006; **141B**: 261–268.
14 Davies MN, Lawn S, Whatley S, Fernandes C, Williams RW, Schalkwyk LC. To what extent is blood a reasonable surrogate for brain in gene expression studies: estimation from mouse hippocampus and spleen. *Front Neurosci* 2009; **3**: 54.
15 Cai C, Langfelder P, Fuller TF, Oldham MCLuo R, van den Berg LH *et al.* Is human blood a good surrogate for brain tissue in transcriptional studies? *BMC Genomics* 2010; **11**: 589.
16 Belmaker RH, Agam G. Major depressive disorder. *N Engl J Med* 2008; **358**: 55–68.
17 Penninx BWJH, Milaneschi Y, Lamers F, Vogelzangs N. Understanding the somatic consequences of depression: biological mechanisms and the role of depression symptom profile. *BMC Med* 2013; **11**: 129.
18 Zunszain PA, Hepgul N, Pariante CM. Inflammation and depression. *Curr Top Behav Neurosci* 2013; **14**: 135–151.
19 Blume J, Douglas SD, Evans DL. Immune suppression and immune activation in depression. *Brain Behav Immun* 2011; **25**: 221–229.
20 Vogelzangs N, Beekman ATF, van Reedt Dortland AKB, Schoevers RA, Giltay EJ, de Jonge P *et al.* Inflammatory and metabolic dysregulation and the 2-year course of depressive disorders in antidepressant users. *Neuropsychopharmacology* 2014; **39**: 1624–1634.
21 Hepgul N, Cattaneo A, Zunszain PA, Pariante CM. Depression pathogenesis and treatment: what can we learn from blood mRNA expression? *BMC Med* 2013; **11**: 28.
22 Mostafavi S, Battle A, Zhu X, Potash JB, Weissman MM, Shi J *et al.* Type I interferon signaling genes in recurrent major depression: increased expression detected by whole-blood RNA sequencing. *Mol Psychiatry* 2014; **19**: 1267–1274.
23 Penninx BWJH, Beekman ATF, Smit JH, Zitman FG, Nolen WA, Spinhoven P *et al.* The Netherlands Study of Depression and Anxiety (NESDA): rationale, objectives and methods. *Int J Methods Psychiatr Res* 2008; **17**: 121–140.
24 Boomsma DI, Willemsen G, Sullivan PF, Heutink P, Meijer P, Sondervan D *et al.* Genome-wide association of major depression: description of samples for the GAIN Major Depressive Disorder Study: NTR and NESDA biobank projects. *Eur J Hum Genet EJHG* 2008; **16**: 335–342.
25 Sullivan PF, de Geus EJC, Willemsen G, James MR, Smit JH, Zandbelt T *et al.* Genome-wide association for major depressive disorder: a possible role for the presynaptic protein piccolo. *Mol Psychiatry* 2009; **14**: 359–375.
26 American Psychiatric Association *Diagnostic and Statistical Manual of Mental Disorders*. American Psychiatric Association: Washington, DC, USA, 1994.
27 Wittchen HU. Reliability and validity studies of the WHO–Composite International Diagnostic Interview (CIDI): a critical review. *J Psychiatr Res* 1994; **28**: 57–84.
28 Shabalin AA. Matrix eQTL: ultra fast eQTL analysis via large matrix operations. *Bioinforma Oxf Engl* 2012; **28**: 1353–1358.
29 Liu JZ, Mcrae AF, Nyholt DR, Medland SE, Wray NR, Brown KM *et al.* A versatile gene-based test for genome-wide association studies. *Am J Hum Genet* 2010; **87**: 139–145.
30 Lips ES, Cornelisse LN, Toonen RF, Min JL, Hultman CM, International Schizophrenia Consortium *et al.* Functional gene group analysis identifies synaptic gene groups as risk factor for schizophrenia. *Mol Psychiatry* 2012; **17**: 996–1006.
31 Barry WT, Nobel AB, Wright FA. Significance analysis of functional categories in gene expression studies: a structured permutation approach. *Bioinforma Oxf Engl* 2005; **21**: 1943–1949.
32 Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 2008; **9**: 559.
33 Wium-Andersen MK, Ørsted DD, Nielsen SF, Nordestgaard BG. Elevated C-reactive protein levels, psychological distress, and depression in 73, 131 individuals. *JAMA Psychiatry* 2013; **70**: 176–184.
34 Glahn DC, Curran JE, Winkler AM, Carless MA, Kent JW, Charlesworth JC *et al.* High dimensional endophenotype ranking in the search for major depression risk genes. *Biol Psychiatry* 2012; **71**: 6–14.
35 Shelton RC, Claiborne J, Sidoryk-Wegrzynowicz M, Reddy R, Aschner M, Lewis DA *et al.* Altered expression of genes involved in inflammation and apoptosis in frontal cortex in major depression. *Mol Psychiatry* 2011; **16**: 751–762.
36 Garbett KA, Vereczkei A, Kálmán S, Brown JA, Taylor WD, Faludi G *et al.* Coordinated messenger RNA/microRNA changes in fibroblasts of patients with major depression. *Biol Psychiatry* 2015; **77**: 256–265.
37 Eyre H, Baune BT. Neuroplastic changes in depression: a role for the immune system. *Psychoneuroendocrinology* 2012; **37**: 1397–1416.
38 Jones KA, Thomsen C. The role of the innate immune system in psychiatric disorders. *Mol Cell Neurosci* 2013; **53**: 52–62.
39 Herbert TB, Cohen S. Depression and immunity: a meta-analytic review. *Psychol Bull* 1993; **113**: 472–486.
40 Zorrilla EP, Luborsky L, McKay JR, Rosenthal R, Houldin A, Tax A *et al.* The relationship of depression and stressors to immunological assays: a meta-analytic review. *Brain Behav Immun* 2001; **15**: 199–226.
41 Howren MB, Lamkin DM, Suls J. Associations of depression with C-reactive protein, IL-1, and IL-6: a meta-analysis. *Psychosom Med* 2009; **71**: 171–186.

42 Dowlati Y, Herrmann N, Swardfager W, Liu H, Sham L, Reim EK *et al.* A meta-analysis of cytokines in major depression. *Biol Psychiatry* 2010; **67**: 446–457.
43 Valkanova V, Ebmeier KP, Allan CL. CRP, IL-6 and depression: a systematic review and meta-analysis of longitudinal studies. *J Affect Disord* 2013; **150**: 736–744.
44 Tsao C-W, Lin Y-S, Chen C-C, Bai C-H, Wu S-R. Cytokines and serotonin transporter in patients with major depression. *Prog Neuropsychopharmacol Biol Psychiatry* 2006; **30**: 899–905.
45 Cattaneo A, Gennarelli M, Uher R, Breen G, Farmer A, Aitchison KJ *et al.* Candidate genes expression profile associated with antidepressants response in the GENDEP study: differentiating between baseline 'predictors' and longitudinal 'targets'. *Neuropsychopharmacology* 2013; **38**: 377–385.
46 Betten A, Dahlgren C, Hermodsson S, Hellstrand K. Serotonin protects NK cells against oxidatively induced functional inhibition and apoptosis. *J Leukoc Biol* 2001; **70**: 65–72.
47 Hernandez ME, Martinez-Fong D, Perez-Tapia M, Estrada-Garcia I, Estrada-Parra S, Pavón L. Evaluation of the effect of selective serotonin-reuptake inhibitors on lymphocyte subsets in patients with a major depressive disorder. *Eur Neuropsychopharmacol J Eur Coll Neuropsychopharmacol* 2010; **20**: 88–95.
48 Evans DL, Folds JD, Petitto JM, Golden RN, Pedersen CA, Corrigan M *et al.* Circulating natural killer cell phenotypes in men and women with major depression. Relation to cytotoxic activity and severity of depression. *Arch Gen Psychiatry* 1992; **49**: 388–395.
49 Frank MG, Hendricks SE, Johnson DR, Wieseler JL, Burke WJ. Antidepressants augment natural killer cell activity: in vivo and in vitro. *Neuropsychobiology* 1999; **39**: 18–24.
50 Cruess DG, Douglas SD, Petitto JM, Have TT, Gettes D, Dubé B *et al.* Association of resolution of major depression with increased natural killer cell activity among HIV-seropositive women. *Am J Psychiatry* 2005; **162**: 2125–2130.
51 Hussaini SMQ, Choi C-I, Cho CH, Kim HJ, Jun H, Jang M-H. Wnt signaling in neuropsychiatric disorders: ties with adult hippocampal neurogenesis and behavior. *Neurosci Biobehav Rev* 2014; **47**: 369–38310.1016/j.neubiorev.2014.09.005.
52 Sani G, Napoletano F, Forte AM, Kotzalidis GD, Panaccione I, Porfiri GM *et al.* The wnt pathway in mood disorders. *Curr Neuropharmacol* 2012; **10**: 239–253.
53 Pilar-Cuéllar F, Vidal R, Díaz A, Castro E, dos Anjos S, Pascual-Brazo J *et al.* Neural plasticity and proliferation in the generation of antidepressant effects: hippocampal implication. *Neural Plast* 2013; **2013**: 537265.
54 Wilkinson MB, Dias C, Magida J, Mazei-Robison M, Lobo M, Kennedy P *et al.* A novel role of the WNT-dishevelled-GSK3β signaling cascade in the mouse nucleus accumbens in a social defeat model of depression. *J Neurosci* 2011; **31**: 9084–9092.
55 Pirooznia M, Seifuddin F, Judy J, Goes FS, Potash JB, Zandi PP. Metamoodics: meta-analysis and bioinformatics resource for mood disorders. *Mol Psychiatry* 2014; **19**: 748–749.
56 Cole SW, Hawkley LC, Arevalo JM, Sung CY, Rose RM, Cacioppo JT. Social regulation of gene expression in human leukocytes. *Genome Biol* 2007; **8**: R189.
57 Ramasamy A, Trabzuni D, Guelfi S, Varghese V, Smith C, Walker R *et al.* Genetic variability in the regulation of gene expression in ten regions of the human brain. *Nat Neurosci* 2014; **17**: 1418–1428.
58 Nieratschker V, Massart R, Gilles M, Luoni A, Suderman MJ, Krumm B *et al.* MORC1 exhibits cross-species differential methylation in association with early life stress as well as genome-wide association with MDD. *Transl Psychiatry* 2014; **4**: e429.
59 Akula N, Barb J, Jiang X, Wendland JR, Choi KH, Sen SK *et al.* RNA-sequencing of the brain transcriptome implicates dysregulation of neuroplasticity, circadian rhythms and GTPase binding in bipolar disorder. *Mol Psychiatry* 2014; **19**: 1179–1185.
60 Li K, Zhou T, Liao L, Yang Z, Wong C, Henn F *et al.* βCaMKII in lateral habenula mediates core symptoms of depression. *Science* 2013; **341**: 1016–1020.
61 Novak G, Seeman P, Tallerico T. Increased expression of calcium/calmodulin-dependent protein kinase IIbeta in frontal cortex in schizophrenia and depression. *Synap N Y N* 2006; **59**: 61–68.
62 Kang HJ, Voleti B, Hajszan T, Rajkowska G, Stockmeier CA, Licznerski P *et al.* Decreased expression of synapse-related genes and loss of synapses in major depressive disorder. *Nat Med* 2012; **18**: 1413–1417.
63 Telegdy G, Adamik A. Anxiolytic action of neuromedin-U and neurotransmitters involved in mice. *Regul Pept* 2013; **186**: 137–140.
64 Tanaka M, Telegdy G. Neurotransmissions of antidepressant-like effects of neuromedin U-23 in mice. *Behav Brain Res* 2014; **259**: 196–199.

Supplementary Information accompanies the paper on the Molecular Psychiatry website (http://www.nature.com/mp)