ORIGINAL ARTICLE

Peripheral blood gene expression profiles linked to monoamine metabolite levels in cerebrospinal fluid

JJ Luykx[1,2,3,11], LM Olde Loohuis[4,11], M Neeleman[1,11], E Strengman[5], SC Bakker[1], E Lentjes[6], P Borgdorff[7], EPA van Dongen[8], P Bruins[8], RS Kahn[1], S Horvath[9,10], S de Jong[4,12] and RA Ophoff[1,4,9]

The blood–brain barrier separates circulating blood from the central nervous system (CNS). The scope of this barrier is not fully understood which limits our ability to relate biological measurements from peripheral to central phenotypes. For example, it is unknown to what extent gene expression levels in peripheral blood are reflective of CNS metabolism. In this study, we examine links between central monoamine metabolite levels and whole-blood gene expression to better understand the connection between peripheral systems and the CNS. To that end, we correlated the prime monoamine metabolites in cerebrospinal fluid (CSF) with whole-genome gene expression microarray data from blood ($N = 240$ human subjects). We additionally applied gene-enrichment analysis and weighted gene co-expression network analyses (WGCNA) to identify modules of co-expressed genes in blood that may be involved with monoamine metabolite levels in CSF. Transcript levels of two genes were significantly associated with CSF serotonin metabolite levels after Bonferroni correction for multiple testing: *THAP7* ($P = 2.8 \times 10^{-8}$, $\beta = 0.08$) and *DDX6* ($P = 2.9 \times 10^{-7}$, $\beta = 0.07$). Differentially expressed genes were significantly enriched for genes expressed in the brain tissue ($P = 6.0 \times 10^{-52}$). WGCNA revealed significant correlations between serotonin metabolism and hub genes with known functions in serotonin metabolism, for example, *HTR2A* and *COMT*. We conclude that gene expression levels in whole blood are associated with monoamine metabolite levels in the human CSF. Our results, including the strong enrichment of brain-expressed genes, illustrate that gene expression profiles in peripheral blood can be relevant for quantitative metabolic phenotypes in the CNS.

Translational Psychiatry (2016) **6**, e983; doi:10.1038/tp.2016.245; published online 13 December 2016

INTRODUCTION

Gene expression profiling analysis constitutes a powerful tool to increase our understanding of neurobiological mechanisms influencing health and disease, including psychiatric disorders such as schizophrenia,[1] major depressive disorder,[2,3] bipolar disorder[4] and 22q11 deletion syndrome-associated psychosis.[5] Integrating transcriptomics with metabolomics approaches may signal neurobiological mechanisms that regulate the metabolome, as demonstrated for the mouse liver.[6] The selective permeability and cross-talk between central nervous system (CNS) and blood is controlled by the blood–brain barrier. It is currently unknown to what degree gene expression levels in blood may be informative for metabolic processes relevant to the CNS and whether peripheral gene expression levels are linked to CNS metabolism. Although studies comparing transcripts in blood with postmortem brain tissue were recently reviewed,[7] no investigations relating cerebrospinal fluid (CSF) to peripheral gene expression have been published. Moreover, limited availability of high-quality postmortem brain tissue hampers the applicability of such studies. CSF, on the other hand, allows direct probing of correlations between metabolites in the CNS and peripheral gene expression.

The primary metabolites of the monoamines—5-hydroxyindoleacetic acid (5-HIAA), homovanillic acid (HVA) and 5-medroxy-4-hydroxyphenylglycol (MHPG)—are implicated in physiological CNS processes (ranging from cognitive functioning to reproduction) and neurobehavioral traits, most notably mood disorders[8,9] and Brunner syndrome.[10] Monoamine metabolites (MMs) can be measured in CSF.[11] Evidence from genome-wide linkage and association studies[11,12] attests to the genetic underpinnings of monoamine turnover in CSF.

Here, we examined whether whole-blood gene expression profiles may be informative for the study of CNS metabolomics. Such knowledge may deepen the understanding of the applicability of peripheral gene expression studies for CNS (endo) phenotypes in psychiatry. Moreover, our design has the potential to point to biological pathways implicated in monoamine metabolism, which, in turn, may lead to candidate pathophysiological processes in mental illness. We therefore measured concentrations of MMs in CSF and collected genome-wide gene expression data in 240 healthy human subjects. We not only

[1]Department of Psychiatry, Brain Center Rudolf Magnus, University Medical Center Utrecht, Utrecht, The Netherlands; [2]Department of Translational Neuroscience Human Neurogenetics Unit, Brain Center Rudolf Magnus, University Medical Center Utrecht, Utrecht, The Netherlands; [3]Department of Psychiatry, ZNA Hospitals, Antwerp, Belgium; [4]Center for Neurobehavioral Genetics, Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles, Los Angeles, CA, USA; [5]Department of Pathology, University Medical Center Utrecht, Utrecht, The Netherlands; [6]Department of Clinical Chemistry and Hematology, University Medical Center Utrecht, Utrecht, The Netherlands; [7]Department of Anesthesiology, Intensive Care and Pain Management, Diakonessenhuis Hospital, Utrecht, The Netherlands; [8]Department of Anesthesiology, Intensive Care and Pain Management, University Medical Center Utrecht, Utrecht, The Netherlands; [9]Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, USA and [10]Department of Biostatistics, Fielding School of Public Health, University of California, Los Angeles, Los Angeles, CA, USA. Correspondence: Professor RA Ophoff, Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA 90095 USA.
E-mail: ophoff@ucla.edu
[11]These authors contributed equally to this work.
[12]Current address: MRC Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK.
Received 7 July 2016; accepted 15 October 2016

highlight several genes and gene modules associated with MM levels, but also demonstrate that such genes are enriched for genes expressed in the brain tissue.

MATERIALS AND METHODS

Study population
Subject recruitment was described previously.[11] In brief, volunteers were recruited at outpatient pre-operative screening services in four hospitals in and around Utrecht, The Netherlands, from August 2008 until March 2010. We included patients (i) undergoing spinal anesthesia for minor elective surgical procedures, (ii) ranging between 18 and 60 years of age and (iii) with four grandparents born in The Netherlands or other Northwestern European countries (Belgium, Germany, UK, France and Denmark). Written informed consent was obtained from the participants. Each candidate participant received a personal telephone interview to exclude subjects with self-reported psychotic or major neurological disorders (stroke, brain tumors, neurodegenerative diseases) and to record any use of psycho-tropic medication.

CSF collection and monoamine metabolite measurements
From each participant, 6 ml of CSF and 5 ml of blood were collected simultaneously between 0800 h and 1600 h, as described previously.[11,13] CSF was kept at 4 °C and transported within 5 h to the laboratory at the University Medical Center Utrecht. The samples were visually inspected for blood contamination and the contaminated samples were excluded from further analyses. Each sample was immediately aliquotted and stored at − 80 °C. Concentrations of 5-HIAA, HVA and MHPG were measured using high-performance liquid chromatography with electrochemical detection according to an established method.[13] Detection failure precluded reliable concentration assessments for 21 MHPG measurements; these were excluded from further analyses. Metabolite concentrations exceeding the mean ± 3 s.d. were considered outliers and removed ($N=1$ for 5-HIAA and HVA; $N=2$ for MHPG). The MM concentrations were not normally distributed (defined by a Shapiro–Wilk test P-value cut-off of 0.05) and therefore log-transformed before analysis. Given the potential increase in power of quantitative genetic analysis of metabolite ratios,[14–16] we then also computed the three ratios between metabolite levels.

RNA quantification procedures, genome-wide expression profiling and quality control
For the isolation and purification of messenger RNA (mRNA) from whole blood, the PAXgene extraction kit (Qiagen, Valencia, CA, USA) was used. The PAXgene tubes were stored at − 20 °C and RNA was isolated within 6 months after sample collection according to the manufacturer's protocol, including an optional DNase digestion step. The total mRNA was quantified using a ribogreen assay (Invitrogen Quant-it Ribogreen Thermo Fisher Scientific, Walham, MA, USA). The quality of total RNA was determined using the Agilent 2100 Bioanalyzer (Agilent, Santa Clara, CA, USA). A threshold of RNA integrity number of 7 was applied to ensure good RNA quality. Genome-wide RNA expression profiling was obtained with the Illumina HumanRef-12 (version 3) arrays using Illumina's standard protocol at the UCLA Neuroscience Genomics Core. Raw data extraction and background correction were performed using GenomeStudio (Illumina, San Diego, CA, USA). Quality control was performed by assessment of hierarchical clustering, box plots, density distribution plots and pair-wise correlations (Supplementary Methods). The Lumi package in R was used for robust spline normalization and variance stabilizing transformation of the gene expression data.[17,18] A total of 48 803 probes were included on the Illumina microarray. To select those probes that were expressed in at least one sample, a gene filter was applied with the detection P-value generated set to 0.01 by GenomeStudio Software (Illumina).[19] Our gene expression data set consisted of three batches based on Illumina serial numbers. Normalization between batches was performed with the ComBat package for R software.[20]

Differential expression analysis
The associations of individual genes with mRNA were tested by means of linear regression using the Limma package[21,22] [20,21] in R (www.r-project. org). We previously ruled out a significant impact of time of day of sampling, storage duration and the use of psychotropic medication on MM levels,[13] but age, gender, height and season of the year at the time of fluid collection constitute relevant covariates for CSF MM studies.[13,23] For single metabolites and their ratios, a linear regression model was applied including these covariates. The season of the year was encoded as a dichotomous variable (sampling in March, April or May versus the rest of the year).[13] The gene expression levels were taken as dependent and metabolite levels as independent variables, whereas outcome measures were rate of change (β, expression increase or decrease per metabolite level unit). We applied a stringent Bonferroni correction to adjust for the fact that we evaluated 25 361 gene expression probes (that remained after filtering on the basis of quality control criteria) and six phenotypes (all three MMs and their ratios), resulting in a false positive rate of $\alpha=0.05/(25\ 361\times 6)=3.29\times 10^{-7}$.

Gene network analysis
The 5000 most variably expressed probes[24] were analyzed using weighted gene co-expression network analysis (WGCNA).[25–27] We define co-expression modules as branches of a hierarchical clustering tree; each module was color-coded. To determine a representative module expression profile for each module, we summarized the (standardized) gene expression profiles of the module by their first principal component. This statistic is referred to as the Module Eigengene. The Module Eigengene value of a sample can be thought of as an average standardized gene expression value for all probes in a module per sample. The Module Eigengene is the mathematically optimal summary of a module as it captures the maximum amount of the underlying variation.[28] We correlated the Module Eigengene of each module with the MMs. WGCNA allows for reconstruction of networks of co-expressed genes and identification of hub genes. Genes with high gene expression significance (that is, highly correlated with the trait of interest) and that are strongly correlated with the Module Eigengene are central hubs and are therefore of interest to monoamine metabolism in our study. Further details are provided in the Supplementary Methods. Again, Bonferroni correction was applied for the 16 modules and six MMs and ratios, resulting in Bonferroni-corrected significance threshold of $\alpha=0.05/(16\times 6)=5.21\times 10^{-4}$.

Functional classification and genetic overlap
A threshold of 10^{-3} was used to select differentially expressed genes resulting from the six association analyses (all three MMs and their ratios). Thus, a total of 1499 Illumina probes were included, corresponding to 1394 unique genes. To determine possible enrichment of this set in genes expressed in brain tissue, we used DAVID (http://david.abcc.ncifcrf.gov/, version 6.7) and the corresponding R package RDAVIDWebService for functional annotation of gene sets included in our analysis.[29–31] Databases included to annotate gene list for tissue expression were CGAP EST QUARTILE; CGAP SAGE QUARTILE; GNF U133A QUARTILE; PIR TISSUE SPECIFICITY and UP TISSUE. To address potential biases that can confound this analysis, we also performed a permutation analysis by randomly selecting an equal number of genes and performing DAVID enrichment analysis. Owing to constraints on the use of the DAVID webserver using the package RDAVIDWebService (50 queries per day), we performed 100 permutations. In addition, as RDAVIDWebService does not allow for official gene symbols as input, we used the originally identified Illumina probes as input to the DAVID analysis ($n=1499$).

To assess the possible genetic overlap between the highlighted genes and psychiatric risk genes, we selected genes containing single-nucleotide polymorphisms from genome-wide association studies (with boundaries of each gene being expanded by 20 kb on each side) using the same P-value threshold (10^{-3}). The threshold of 10^{-3} was picked to be able to select a sufficient number of genes for an overlap analysis, while maintaining a relatively low type I error rate. For this analysis, clumped summary statistics were used for the largest psychiatric genomics consortium (PGC, http://www.med.unc.edu/pgc) genome-wide association studies for major depressive disorder,[32] bipolar disorder[33] and schizophrenia[34] in April 2015. For each trait, these genes were then overlapped with the set of differentially expressed genes, as well as the genes in the modules identified using WGCNA. The significance of overlap with genes in genome-wide association study regions was determined empirically on the basis of 10 000 permutations: for each repetition, a set of genes of the same size was randomly selected from the genes on the expression array and overlapped with psychiatric risk genes. Again, Bonferroni correction was applied for the three psychiatric traits and multiple overlap tests, resulting in Bonferroni-corrected significance threshold of $\alpha=0.05/(12\times 3)=0.001$. As a negative control for this analysis, we also included

the latest genome-wide association study results for human height[35] and performed the same analysis as for the above-mentioned psychiatric phenotypes.

RESULTS

Baseline characteristics

Quality control of gene expression data resulted in 233 subjects and 25 361 probes available for analyses. The mean (s.d.) gene expression level of these probes was 8.54 (1.10). Seventy-five percent of the study population was male; the mean (s.d.) age was 39 (11) years. The mean (s.d.) monoamine metabolite concentrations were 147 (62.0) for 5-HIAA, 24.6 (5.07) for MHPG and 211 (74.3) nmol l^{-1} for HVA.

Differential gene expression analyses

After Bonferroni correction for multiple testing, two genes were significantly associated with monoamine metabolite levels: THAP7 and DDX6 at $P = 2.8 \times 10^{-8}$ ($\beta = 0.08$) and 2.9×10^{-7} ($\beta = 0.07$), respectively (Table 1); both transcripts were associated with serotonin (5-HT) metabolite levels. For the other metabolite levels, no genome-wide significant results were detected.

Gene co-expression network analysis

WGCNA allows one to detect co-expression modules, their representatives (known as Module Eigengenes) and the identification of intramodular hub genes. The most significant associations between modules and MM levels were observed for 5-HT metabolism: the Module Eigengenes of the brown module (757 probes) and the turquoise modules (1430 probes) were significantly correlated with levels of 5-HIAA and the ratios 5-HIAA: MHPG and 5-HIAA:HVA (Table 2). The black module exhibited only nominally significant evidence of association with monoamine levels. None of the other modules had a significant association with metabolites or related ratios.

Genes whose gene expression levels are highly correlated with the Module Eigengenes can be interpreted as centrally located

hubs in the module.[28] Interestingly, four intramodular hub genes in the MM-related modules were previously implicated in monoamine metabolism or psychiatric disorders. Within the brown module, the gene encoding the serotonin 2A receptor (HTR2A) and GNAI3 were among the top four most highly ranking genes that correlated with the 5-HIAA:HVA ratio ($r = -0.30$, $P = 1.5 \times 10^{-6}$; and $r = -0.31$, $P = 1.4 \times 10^{-6}$, respectively; Figures 1a and b and Supplementary 1A). Within the turquoise module, GSK3B and TCF4 were among the genes most strongly correlated with the 5-HIAA/HVA and 5-HIAA:MHPG ratios ($r = 0.23$, $P = 4.9 \times 10^{-4}$; and $r = 0.21$, $P = 1.6 \times 10^{-3}$, respectively; Figures 1c and d and Supplementary 1B). The expression levels of these genes explain between 4.4 and 9.6% of the variation in CSF monoamine metabolite concentrations. Within the black module, a module with nominally significant evidence of association with monoamine levels, the gene encoding catechol-O-methyltransferase (COMT) showed a strong correlation with the 5-HIAA:HVA ratio ($r = 0.29$, $P = 5.5 \times 10^{-6}$; Figure 1e and Supplementary 1C).

Brain tissue enrichment

Using DAVID, we found brain tissue enrichment for the 1499 differentially expressed probes ($P_{\text{Bonferroni corrected}} = 5.98 \times 10^{-52}$; Supplementary 2). This highly significant result is unlikely to be due to a bias in representation of genes on the array as the permutation results using randomly selected sets of genes show no such enrichment of genes expressed in the brain (Supplementary 2). Moreover, performing the same analysis using the 1394 unique genes corresponding to the identified probes yields nearly identical results (Supplementary 2).

We did not find significant evidence for genetic overlap between differentially expressed genes and psychiatric risk genes or genes in loci associated with height (negative control; Supplementary 3). Nonetheless, we detected a nominally significant overlap (permutation test, $P = 0.04$) between genes within the black co-expression module and genes located at schizophrenia susceptibility loci.

DISCUSSION

Our results demonstrate that gene expression profiles in whole blood are significantly correlated with monoamine metabolite levels in CSF, most notably the serotonin metabolite 5-HIAA. Moreover, the genes highlighted in our study are enriched for genes expressed in human brain tissue.

To our knowledge, this is the first study to link peripheral transcriptome levels to monoamine levels in the CNS. Single gene correlations in our study were particularly strong for serotonin metabolism: THAP7 and DDX6 were both found to be associated with the 5-HIAA:MHPG ratio. THAP7 on 22q11 is a ubiquitously expressed gene encoding a chromatin-associated, histone tail-binding protein, while DDX6 is a ubiquitously expressed RNA helicase gene. These genes have not been linked to monoamine metabolism previously. As reported before,[16,36] we observed

Table 1. Significant results from the differential expression analysis highlights two genes

Differential expression associations with the 5-HIAA:MHPG ratio (N = 202 subjects)			
Gene	Average gene expression	β^a	Unadjusted P-value
THAP7	8.6	0.081	2.8×10^{-8}
DDX6	8.1	0.069	2.9×10^{-7}

^{a}Log fold changes of expression (expression increase or decrease per metabolite level unit).

Table 2. Module Eigengene Pearson correlations with CSF monoamine metabolites levels

Module Eigengene^a correlations with monoamine metabolite levels			
Monoamine metabolite levels	Co-expression module	r	Unadjusted P-value
5-HIAA:HVA ratio (N = 231 subjects)	Brown module (757 transcripts)	− 0.28	2.0×10^{-5}
5-HIAA:MHPG ratio (N = 209 subjects)	Brown module (757 transcripts)	− 0.27	9.9×10^{-5}
5-HIAA:MHPG ratio (N = 209 subjects)	Turquoise module (1430 transcripts)	0.25	3.3×10^{-4}

Abbreviation: CSF, cerebrospinal fluid. ^{a}The Module Eigengene is defined as the first principal component of the module.
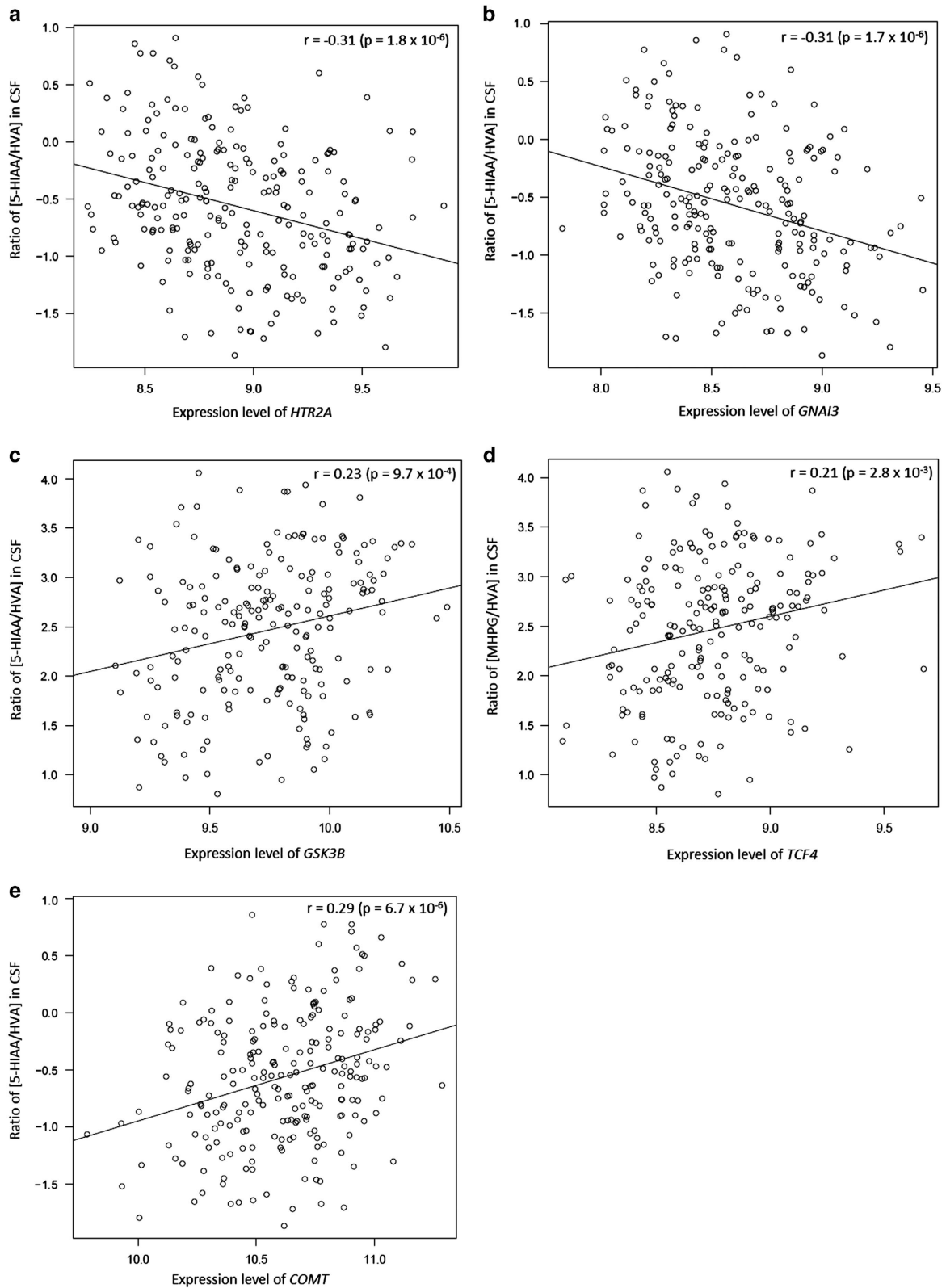
4

Figure 1. Depicted are the five genes within modules that were significantly associated with monoamine metabolites and that have been implicated in psychiatric illness or monoamine metabolism. (**a**) Correlation between the 5-HIAA/HVA concentration ratio in CSF and expression levels of the gene *HTR2A*. (**b**) Correlation between the 5-HIAA/HVA concentration ratio in CSF and expression levels of the gene *GNAI3*. (**c**) Correlation between the 5-HIAA/HVA concentration ratio in CSF and expression levels of the gene *GSK3B*. (**d**) Correlation between the MHPG/HVA concentration ratio in CSF and expression levels of the gene *TCF4*. (**e**) Correlation between the 5-HIAA/HVA concentration ratio in CSF and expression levels of the gene *COMT*. Each dot represents one healthy subject. Summary statistics are shown in the figures. *COMT*, catechol-*O*-methyltransferase; CSF, cerebrospinal fluid.

stronger associations when using metabolite ratios instead of single MM levels. We postulate that incorporating metabolite ratios may improve the proportion of signal compared with noise by correcting for known and unknown covariates at the individual level.

Our systems biology analysis of the gene expression data using WGCNA managed to identify two co-expression modules that are significantly associated with metabolite levels in CSF. Several intramodular hub genes of these modules have previously been associated with psychiatric (intermediate) phenotypes (Figure 1). For example, HTR2A encodes the 5-hydroxytryptamine 2A receptor, a major postsynaptic target for serotonin in the human brain that has been functionally linked to psychosis and antipsychotic response.[37] In the module that is significantly correlated with the 5-HIAA:MHPG ratio, the hub genes GSK3B and TCF4 seem particularly relevant to neuropsychiatric disease. GSK3B encodes glycogen synthase kinase 3β, an enzyme inhibited by lithium.[38,39] Preclinical and clinical genetic data implicate the gene in lithium-sensitive behaviors in mice.[38] TCF4 encodes a basic Helix-Loop-Helix transcription factor that is highly expressed in human brain and has been implicated in schizophrenia.[34,40–42] Moreover, TCF4 is in a locus that has been shown to contribute to the shared genetic risk of five major psychiatric disorders.[43] Although not surviving stringent correction for multiple testing in our study, we find nominal evidence for a link between the prime dopamine metabolite (HVA), the black module and schizophrenia. Within this module, COMT, known to encode a paramount dopamine and norepinephrine-degrading enzyme, is one of the most strongly associated hub genes.

Although the mechanisms underlying the association between gene expression in blood and monoamine metabolism in CSF have yet to be uncovered, several pathways that connect peripheral with central nervous system tissues may have a role. Serotonin is mainly synthesized in intestinal enterochromaffin cells and transported into blood platelets by the serotonin transporter (SLC6A4).[44] The transporter in blood is structurally and functionally similar to the CNS transporter.[44] In blood, these transporters rapidly convey a substantial proportion of peripheral 5-HT into platelets, keeping plasma concentrations low.[44,45] In addition, the prime 5-HT catabolic enzyme—monoamine oxidase —degrades both central and peripheral 5-HT. Peripheral and central 5-HT metabolic pathways are thus likely to be highly intertwined or interdependent, which, in turn, may explain why gene expression in blood is informative for CNS monoamine metabolism, in particular, of 5-HIAA.

When interpreting the results, several limitations should be borne in mind. First, a replication set was lacking for the current analyses, which is due to the laborious and invasive nature of collecting CSF. Second, we did not measure the blood:CSF albumin ratio, precluding us from making inferences about the integrity of the blood–brain and blood–CSF barriers. Third, we cannot rule out a (partially) peripheral origin of monoamine metabolites measured in CSF. Examining gene expression in other cell types with a more direct link to CNS (for example, from olfactory cells[46]) may help deciphering this complexity. Future studies that take into account these considerations will provide a more refined understanding of the interplay between CNS metabolism and regulation of peripheral gene expression, which is likely controlled by the blood–brain barriers. Other advancements that will benefit future studies include a larger sample size (to increase statistical power for discovery), the application of high-throughput RNA sequencing (to obtain high-resolution transcriptome data, including in CSF itself), and collection of metabolomic data of both CSF and plasma (to acquire a more elaborate selection of metabolites). The role of the blood–CSF barrier and the interaction between peripheral systems and CNS should also be investigated in patients suffering from mood disorders, Parkinson's disease and other neurodegenerative disorders.[47–49]

On the basis of our unique data set, we find significant relationships between gene expression levels in whole blood and monoamine levels in CSF. Our finding that gene expression in blood can be linked to monoamine metabolite levels in CSF provides evidence that gene expression profiles in blood constitute a non-CNS component relevant to quantitative metabolic phenotypes of the CNS. Whether this connection across the blood–brain barrier is causally involved in CNS function and pathophysiology (for example, mental illness) remains to be discovered.

URLs used: http://david.abcc.ncifcrf.gov; http://www.med.unc.edu/pgc

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

1 de Jong S, van Eijk KR, Zeegers DW, Strengman E, Janson E, Veldink JH et al. Expression QTL analysis of top loci from GWAS meta-analysis highlights additional schizophrenia candidate genes. Eur J Hum Genet 2012; 20: 1004–1008.

2 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.

3 Jansen R, Penninx BW, Madar V, Xia K, Milaneschi Y, Hottenga JJ et al. Gene expression in major depressive disorder. Mol Psychiatry 2016; 21: 339–347.

4 Seifuddin F, Pirooznia M, Judy JT, Goes FS, Potash JB, Zandi PP. Systematic review of genome-wide gene expression studies of bipolar disorder. BMC Psychiatry 2013; 13: 213.

5 Jalbrzikowski M, Lazaro MT, Gao F, Huang A, Chow C, Geschwind DH et al. Transcriptome profiling of peripheral blood in 22q11.2 deletion syndrome reveals functional pathways related to psychosis and autism spectrum disorder. PLoS One 2015; 10: e0132542.

6 Kleemann R, Verschuren L, van Erk MJ, Nikolsky Y, Cnubben NH, Verheij ER et al. Atherosclerosis and liver inflammation induced by increased dietary cholesterol intake: a combined transcriptomics and metabolomics analysis. Genome Biol 2007; 8: R200.

7 Tylee DS, Kawaguchi DM, Glatt SJ. On the outside, looking in: a review and evaluation of the comparability of blood and brain "-omes". Am J Med Genet B Neuropsychiatr Genet 2013; 162B: 595–603.

8 Berger M, Gray JA, Roth BL. The expanded biology of serotonin. Ann Rev Med 2009; 60: 355–366.

9 Ryden E, Johansson C, Blennow K, Landen M, Lower CSF. HVA and 5-HIAA in bipolar disorder type 1 with a history of childhood ADHD. J Neural Transm 2009; 116: 1667–1674.

10 Brunner HG, Nelen M, Breakefield XO, Ropers HH, van Oost BA. Abnormal behavior associated with a point mutation in the structural gene for monoamine oxidase A. Science 1993; 262: 578–580.

11 Luykx JJ, Bakker SC, Lentjes E, Neeleman M, Strengman E, Mentink L et al. Genome-wide association study of monoamine metabolite levels in human cerebrospinal fluid. Mol Psychiatry 2013; 19: 228–234.

12 Freimer NB, Service SK, Ophoff RA, Jasinska AJ, McKee K, Villeneuve A et al. A quantitative trait locus for variation in dopamine metabolism mapped in a primate model using reference sequences from related species. Proc Natl Acad Sci USA 2007; 104: 15811–15816.

13 Luykx JJ, Bakker SC, Lentjes E, Boks MP, van Geloven N, Eijkemans MJ et al. Season of sampling and season of birth influence serotonin metabolite levels in human cerebrospinal fluid. PLoS One 2012; 7: e30497.

14 Nicholson G, Rantalainen M, Li JV, Maher AD, Malmodin D, Ahmadi KR et al. A genome-wide metabolic QTL analysis in Europeans implicates two loci shaped by recent positive selection. PLoS Genet 2011; 7: e1002270.

15 Suhre K, Wallaschofski H, Raffler J, Friedrich N, Haring R, Michael K et al. A genome-wide association study of metabolic traits in human urine. Nat Genet 2011; 43: 565–569.

16 Gieger C, Geistlinger L, Altmaier E, Hrabe de Angelis M, Kronenberg F, Meitinger T et al. Genetics meets metabolomics: a genome-wide association study of metabolite profiles in human serum. PLoS Genet 2008; 4: e1000282.

17 Du P, Kibbe WA, Lin SM. lumi: a pipeline for processing Illumina microarray. Bioinformatics 2008; 24: 1547–1548.

6

18 Lin SM, Du P, Huber W, Kibbe WA. Model-based variance-stabilizing transformation for Illumina microarray data. *Nucleic Acids Res* 2008; **36**: e11.

19 De Jong S, Neeleman M, Luykx JJ, ten Berg MJ, Strengman E, Den Breeijen HH *et al*. Seasonal changes in gene expression represent cell-type composition in whole blood. *Hum Mol Genet* 2014; **23**: 2721–2728.

20 Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 2007; **8**: 118–127.

21 Smyth GK. *Limma: Linear Models for Microarray Data*. Springer: New York City, NY, USA, 2005.

22 Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 2004; **3**, Article3.

23 Luykx JJ, Bakker SC, van Geloven N, Eijkemans MJ, Horvath S, Lentjes E *et al*. Seasonal variation of serotonin turnover in human cerebrospinal fluid, depressive symptoms and the role of the 5-HTTLPR. *Transl Psychiatry* 2013; **3**: e311.

24 de Jong S, Boks MP, Fuller TF, Strengman E, Janson E, de Kovel CG *et al*. A gene co-expression network in whole blood of schizophrenia patients is independent of antipsychotic-use and enriched for brain-expressed genes. *PLoS One* 2012; **7**: e39498.

25 Langfelder P, Zhang B, Horvath S. Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut package for R. *Bioinformatics* 2008; **24**: 719–720.

26 Zhang B, Horvath S. A general framework for weighted gene co-expression network analysis. *Stat Appl Genet Mol Biol* 2005; **4**, Article17.

27 Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 2008; **9**: 559.

28 Horvath S, Dong J. Geometric interpretation of gene coexpression network analysis. *PLoS Comput Biol* 2008; **4**: e1000117.

29 Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2009; **4**: 44–57.

30 Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* 2009; **37**: 1–13.

31 Fresno C, Fernandez EA. RDAVIDWebService: a versatile R interface to DAVID. *Bioinformatics* 2013; **29**: 2810–2811.

32 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.

33 Psychiatric GCBDWG. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nat Genet* 2011; **43**: 977–983.

34 Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 2014; **511**: 421–427.

35 Wood AR, Esko T, Yang J, Vedantam S, Pers TH, Gustafsson S *et al*. Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat Genet* 2014; **46**: 1173–1186.

36 Suhre K, Shin SY, Petersen AK, Mohney RP, Meredith D, Wagele B *et al*. Human metabolic individuality in biomedical and pharmaceutical research. *Nature* 2011; **477**: 54–60.

37 Serretti A, Drago A, De Ronchi D. HTR2A gene variants and psychiatric disorders: a review of current literature and selection of SNPs for future studies. *Curr Med Chem* 2007; **14**: 2053–2069.

38 O'Brien WT, Huang J, Buccafusca R, Garskof J, Valvezan AJ, Berry GT *et al*. Glycogen synthase kinase-3 is essential for beta-arrestin-2 complex formation and lithium-sensitive behaviors in mice. *J Clin Invest* 2011; **121**: 3756–3762.

39 Wilkinson MB, Dias C, Magida J, Mazei-Robison M, Lobo M, Kennedy P *et al*. A novel role of the WNT-dishevelled-GSK3beta signaling cascade in the mouse nucleus accumbens in a social defeat model of depression. *J Neurosci* 2011; **31**: 9084–9092.

40 Navarrete K, Pedroso I, De Jong S, Stefansson H, Steinberg S, Stefansson K *et al*. TCF4 (e2-2; ITF2): a schizophrenia-associated gene with pleiotropic effects on human disease. *Am J Med Genet B Neuropsychiatr Genet* 2013; **162B**: 1–16.

41 Steinberg S, de Jong S, Andreassen OA, Werge T, Borglum AD, Mors O *et al*. Common variants at VRK2 and TCF4 conferring risk of schizophrenia. *Hum Mol Genet* 2011; **20**: 4076–4081.

42 Stefansson H, Ophoff RA, Steinberg S, Andreassen OA, Cichon S, Rujescu D *et al*. Common variants conferring risk of schizophrenia. *Nature* 2009; **460**: 744–747.

43 Cross-Disorder Group of the Psychiatric Genomics Consortium. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet* 2013; **381**: 1371–1379.

44 Yubero-Lahoz S, Robledo P, Farre M, de laTorre R. Platelet SERT as a peripheral biomarker of serotonergic neurotransmission in the central nervous system. *Curr Med Chem* 2013; **20**: 1382–1396.

45 Watts SW, Morrison SF, Davis RP, Barman SM. Serotonin and blood pressure regulation. *Pharmacol Rev* 2012; **64**: 359–388.

46 Horiuchi Y, Kano S, Ishizuka K, Cascella NG, Ishii S, Talbot CC Jr *et al*. Olfactory cells via nasal biopsy reflect the developing brain in gene expression profiles: utility and limitation of the surrogate tissues in research for brain disorders. *Neurosci Res* 2013; **77**: 247–250.

47 Chatzittofis A, Nordstrom P, Hellstrom C, Arver S, Asberg M, Jokinen J. CSF 5-HIAA, cortisol and DHEAS levels in suicide attempters. *Eur Neuropsychopharmacol* 2013; **23**: 1280–1287.

48 Lee BH, Kim YK. Potential peripheral biological predictors of suicidal behavior in major depressive disorder. *Prog Neuropsychopharmacol Biol Psychiatry* 2011; **35**: 842–847.

49 Olivola E, Pierantozzi M, Imbriani P, Liguori C, Stampanoni Bassi M, Conti M *et al*. Serotonin impairment in CSF of PD patients, without an apparent clinical counterpart. *PLoS One* 2014; **9**: e101763.

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