Title

Transcriptome-wide mega-analyses reveal joint dysregulation of immunologic genes and transcription regulators in brain and blood in schizophrenia

Journal

Schizophrenia Research, 176(2-3)

ISSN

0920-9964

Authors

Hess, Jonathan L
Tylee, Daniel S
Barve, Rahul
et al.

Publication Date

2016-10-01

DOI

10.1016/j.schres.2016.07.006

Peer reviewed

HHS Public Access

Author manuscript

Schizophr Res. Author manuscript; available in PMC 2017 October 01.

Published in final edited form as:

Schizophr Res. 2016 October ; 176(2-3): 114–124. doi:10.1016/j.schres.2016.07.006.

Transcriptome-Wide Mega-Analyses Reveal Joint Dysregulation of Immunologic Genes and Transcription Regulators in Brain and Blood in Schizophrenia

Jonathan L. Hess[1,*], **Daniel S. Tylee**[1,*], **Rahul Barve**[1], **Simone de Jong**[2], **Roel A. Ophoff**[2,3], **Nishantha Kumarasinghe**[4,5,6,7], **Paul Tooney**[6,8,9,10], **Ulrich Schall**[4,6,9,10], **Erin Gardiner**[6,8,10], **Natalie Jane Beveridge**[6,8,9], **Rodney J. Scott**[8,9], **Surangi Yasawardene**[5], **Antionette Perera**[5], **Jayan Mendis**[5], **Vaughan Carr**[6,11], **Brian Kelly**[4,9,10], **Murray Cairns**[6,8,9,10], **the Neurobehavioural Genetics Unit**, **Ming T. Tsuang**[12], and **Stephen J. Glatt**[1,†]

[1]Psychiatric Genetic Epidemiology & Neurobiology Laboratory (PsychGENe Lab); Departments of Psychiatry and Behavioral Sciences & Neuroscience and Physiology; SUNY Upstate Medical University; Syracuse, NY, U.S.A [2]Center for Neurobehavioral Genetics, Semel Institute for Neuroscience and Behavior, David Geffen School of Medicine at the University of California Los Angeles, Los Angeles, California, U.S.A [3]Department of Psychiatry, Brain Center Rudolf Magnus, University Medical Center Utrecht, Utrecht, The Netherlands [4]School of Medicine & Public Health, The University of Newcastle, Callaghan, Newcastle, Australia [5]Department of Anatomy, Faculty of Medical Sciences, University of Sri Jayawardenepura, Nugegoda, Sri Lanka [6]Schizophrenia Research Institute, Sydney, New South Wales, Australia [7]Faculty of Medicine, Sir John Kotelawala Defence University, Ratmalana, Sri Lanka [8]School of Biomedical Sciences & Pharmacy, Faculty of Health, The University of Newcastle, New South Wales, Australia [9]Hunter Medical Research Institute, Newcastle, Australia [10]Centre for Translational Neuroscience & Mental Health, University of Newcastle, Callaghan, Newcastle, Australia [11]School of Psychiatry, University of New South Wales, Kensington, New South Wales, Australia [12]Center for Behavioral Genomics, Department of Psychiatry, Institute for Genomic Medicine, University of California, San Diego, La Jolla, CA, USA; Harvard Institute of Psychiatric Epidemiology and Genetics, Boston, USA

Abstract

The application of microarray technology in schizophrenia research was heralded as paradigm-shifting, as it allowed for high-throughput assessment of cell and tissue function. This technology was widely adopted, initially in studies of *postmortem* brain tissue, and later in studies of peripheral blood. The collective body of schizophrenia microarray literature contains apparent

[†]To whom correspondence should be addressed: SUNY Upstate Medical University, 750 East Adams Street, Syracuse, NY 13210, Phone: (315) 464-7742, stephen.glatt@psychgenelab.com.
[*]These authors contributed equally to this work.

Conflicts of Interest

The authors have no conflicts of interest to declare.

inconsistencies between studies, with failures to replicate top hits, in part due to small sample sizes, cohort-specific effects, differences in array types, and other confounders. In an attempt to summarize existing studies of schizophrenia cases and non-related comparison subjects, we performed two mega-analyses of a combined set of microarray data from *postmortem* prefrontal cortices (n = 315) and from *ex-vivo* blood tissues (n = 578). We adjusted regression models per gene to remove non-significant covariates, providing best-estimates of transcripts dysregulated in schizophrenia. We also examined dysregulation of functionally related gene sets and gene co-expression modules, and assessed enrichment of cell types and genetic risk factors. The identities of the most significantly dysregulated genes were largely distinct for each tissue, but the findings indicated common emergent biological functions (*e.g.* immunity) and regulatory factors (*e.g.*, predicted targets of transcription factors and miRNA species across tissues). Our network-based analyses converged upon similar patterns of heightened innate immune gene expression in both brain and blood in schizophrenia. We also constructed generalizable machine-learning classifiers using the blood-based microarray data. Our study provides an informative atlas for future pathophysiologic and biomarker studies of schizophrenia.

Keywords

schizophrenia; gene expression; transcriptome; brain; blood; innate immunity; support vector machine; random forests

INTRODUCTION

The molecular bases of schizophrenia (**SZ**) remain unresolved despite decades of intensifying research. This situation impedes progress toward biologically based risk assessment and diagnostic testing, early detection, and the development of rationally selected therapeutics to alter disease progression and clinical trajectory. As such, characterizing molecular correlates of SZ is of great potential interest and value. In the past 15 years, the transcriptome has received growing attention in SZ research, particularly in the effort to identify biomarkers—objective biological indicators of normal functioning or illness. Whole-transcriptome quantification (*e.g.*, by microarray) offers several attractive features: (1) it provides a relatively efficient and unbiased means of screening many analytes of a single molecule type (*i.e.*, messenger RNAs, **mRNAs**); (2) RNAs can be mapped reliably onto genes and proteins to assess a wide range of biological processes; (3) the measurement of large numbers of biological features allows for the assessment of network function; and (4) differences in mRNA expression reflect the combination of genetic and environmental factors, making it a more dynamic and responsive readout of biological function than static genetic variants. Indeed, transcriptome-wide studies of *postmortem* brain tissues have revealed altered molecular pathways and helped generate new hypotheses about the biological underpinnings of SZ. Similarly, studies of blood samples from individuals with SZ shed light on disturbances in circulating immune cells and could provide a basis for easily assessable SZ biomarkers.

Despite the vast potential and initial enthusiasm surrounding the use of microarrays in SZ research, the cross-study replication of genes and pathways found to be disrupted in SZ is

mixed (Mirnics et al., 2006). A variety of practical and technical limitations may contribute to this, including: the evolution of new array technologies over time, the likelihood of etiologic heterogeneity of SZ (Arnedo et al., 2015; Tsuang and Faraone, 1995), the use of small sample sizes, and the inability to adequately protect against type-I errors, all of which exacerbate the "winner's-curse" phenomenon that undermines replication. In light of these issues, several studies have sought to consolidate the knowledge of transcriptomic abnormalities in SZ *via* meta-analysis (Bergon et al., 2015; Mistry and Pavlidis, 2010; Mistry et al., 2013a; Pérez-Santiago et al., 2012). These studies bolstered confidence by employing consistent preprocessing methods and demonstrating some similar dysregulated genes and network features across different studies; the implicated biological functions included oxidative phosphorylation, protein and nucleotide metabolism, synaptic transmission, myelination and glial function, and immune function, each of which have been implicated in previous work (Åberg et al., 2006; Dean, 2011; Devor and Waziri, 1993; Fineberg and Ellman, 2013; Middleton et al., 2002; Potvin et al., 2008). However, the approaches employed in previous meta-analyses studies had some limitations: (1) for the detection of differentially expressed genes, meta-analysis of summary statistics is relatively underpowered compared with combined-samples re-analysis of individual level data; (2) summary statistic meta-analysis does not allow flexible and transcript-specific modeling of clinical covariates across the entire sample; and (3) meta-analysis is not amenable to co-expression network analyses.

We use the term mega-analysis to refer to a strategy of the pooling of individual-level clinical and biological data from multiple studies for statistical modeling with appropriate correction for between-study variations (Mistry et al., 2013b; Seifuddin et al., 2013). This strategy allows for explicit modeling of factors that are consistently reported across studies (*i.e.*, gender, age), as well as factors that are inconsistently reported across studies (*e.g.*, subject medication status). In this study, we conducted two separate mega-analyses to summarize existing microarray-based transcriptomic studies of SZ in *postmortem* brain and in blood tissue using mixed-effect linear modeling. We extend upon previous approaches by employing network and annotation-based analyses to assess emergent biological functions. Furthermore, we perform systematic cross-tissue comparison of dysregulated functional gene sets and co-expression networks in SZ. We also characterized gene co-expression networks that were preserved across brain and blood samples in unaffected comparison subjects. Finally, we examined the cross-study generalizability of blood-based transcriptomic classifiers that differentiated SZ cases from unaffected comparison subjects.

METHODS

Methods are described briefly here due to space constraints. A comprehensive description of all methods can be found in the Supplementary Materials.

Literature Search and Study Selection

We searched public databases (*i.e.*, NCBI dbGaP, PubMed, SCOPUS, and EMBL-EBI ArrayExpress) for microarray-based studies of gene expression in subjects with SZ, schizoaffective disorder, or psychosis. We conducted this literature search for eligible data

up to January 1, 2015; otherwise-eligible studies published after this date were not included in our analyses, but are shown in Supplementary Table 1 and compared to our findings qualitatively in the Discussion section.. Twenty-five studies of blood-based gene expression and 19 studies of brain-based gene expression were identified. The following criteria for study inclusion and sample inclusion were used: (1) we only included studies that compared cases with unaffected non-related controls, (2) we only included cases classified as SZ or schizoaffective disorder, depressive subtype, based on the original investigators' determinations, (3) we only included studies for which raw probe-level data and gene annotations were available, (4) we only included studies that utilized non-custom microarray platforms developed by Affymetrix or Illumina to minimize technical sources of heterogeneity, and (5) we only included *postmortem* brain studies with samples consisting of tissue homogenates. Ultimately, we included nine blood studies and nine brain studies (*postmortem* prefrontal cortex, **PFC**, only) were ultimately retained for analysis (Table 1). The rationale for excluding each of the 26 other studies is provided in Supplementary Table 1.

Data Import, Normalization, Quality Control and Probe Matching

Data from each study were processed and normalized independently. Affymetrix arrays underwent robust multi-array average (**RMA**) normalization (Irizarry et al., 2003), with additional GC-correction whenever possible (*e.g.*, not compatible with Affymetrix Human Exon 1.0ST array). Both Affymetrix and Illumina array data were quantile-normalized and *log*-2 transformed. We mapped probes to HGNC gene symbols and collapsed expression values of multiple probes to individual genes through median summarization. Finally, for each gene within each individual study, expression values were *z*-transformed in order to normalize the range and variance of expression *across* datasets generated on different array platforms; the effects of normalization and transformation are depicted in Figure S1. In order to identify potential differences in the proportions of leukocyte subtypes between SZ cases and unaffected comparison subjects, we performed deconvolution analysis using previously described methods (Abbas et al., 2009) followed by an independent samples *t*-test with family-wise Benjamini-Hochberg (**BH**) correction for multiple testing.

Mixed-Effect Linear Modeling and Gene Set Analysis

Expression and covariate data from individual studies were combined, creating separate brain ($n = 315$ and blood datasets ($n = 578$). Independent mega-analyses were performed on these datasets using mixed-effect linear modeling. The brain analysis included covariates for age (continuous), ancestry (Caucasian, Asian, African-American), gender (male, female), *postmortem* interval (continuous) and tissue pH (continuous). The blood analysis included covariates for age, sample-type (whole blood, leukocytes, peripheral blood mononuclear cells), ancestry (Caucasian, Asian), gender, and anti-psychotic status (yes, no; as defined by original study authors). A total of 20 767 genes were analyzed from brain studies and 19 737 genes contained sufficient data for analysis in blood studies. With the exception of 217 genes in our blood data set, all genes that were included in our mega-analyses had been measured in more than one study. For multiple-test correction, we examined Bonferroni-corrected *p*-values in order to conservatively define differentially expressed genes; for downstream analyses, we used a more permissive False Discovery Rate (**FDR**) *q*-value < 0.10 to control

the family-wise error rate at 10% while allowing more transcripts to move forward for cross-tissue comparison and enrichment analysis (Storey, 2003). Permutation-based gene set analyses (Väremo et al., 2013) were performed separately for each tissue using the summary statistics (*p*- and *t*-values) derived from the single-gene analyses.

Expression Quantitative Trait Loci and GWAS Enrichment Analysis

For various gene lists of interest (*i.e.*, differentially expressed or participants in a network module), we sought to assess whether those genes disproportionately represented: (1) expression quantitative trait loci (eQTLs) previously identified in brain or blood cells (National Center for Biotechnology Information [NCBI] eQTL Browser); and (2) SNPs associated with SZ based on prior association studies (Ripke et al., 2014).

Constructing Networks of Co-Expressed Genes in Brain and Blood

We performed unsupervised weighted unsigned gene co-expression network analysis (**WGCNA**; (Langfelder et al., 2011) using the *blockwiseModules* function separately in brain and blood datasets. We used linear mixed effect models (as described above) to predict module eigengene values, in order to identify SZ-associated modules. For SZ- associated modules, we identified highly connected hub genes, assessed functional enrichment (described below), and performed cross-tissue comparison of module genes. Additionally, we assessed network module preservation within each tissue (across diagnostic groups) and between tissues (only in unaffected comparison subjects).

Enrichment Analysis of Biological Annotations and Cell-Type Signatures

Hypergeometric testing was used to assess network modules for: (1) functional enrichment based on the contents of the DAVID Knowledgebase (v.6.7; Huang et al., 2009) (2) enrichment with brain cell-specific signatures (Dougherty et al., 2010); and (3) immune cell-specific signatures (Abbas et al., 2009; Watkins et al., 2009). Family-wise BH correction was applied per database to control for multiple testing.

Gene Set and Network Module Heterogeneity Analyses

For SZ-associated gene sets and network modules, we sought to assess whether the same SZ samples were driving the between-groups difference observed for each feature using previously described clustering methods (Lottaz et al., 2007).

Machine-Learning Classification using Blood Transcriptomic Data

We used Random Forest and Ensemble Support Vector Machine approaches to construct and validate classifiers using independent data matrices carved from the blood mega-analysis dataset. Training (n = 413 samples run on Illumina arrays) and validation sets (n = 165 run on Affymetrix arrays) were generated; this manufacturer-based separation was chosen to pose a maximal challenge to classifier generalizability. Sets of the top significantly dysregulated genes (k = 20, 60, 150) were identified by linear mixed-modeling in the training set; these sets reflect the minimum, maximum, and average size of optimal classifiers of neuropsychiatric disorders from blood transcriptome data based on our past

experience (Glatt et al., 2013, 2012; Tsuang et al., 2005; Tylee et al., 2015). Classifiers were fit to the training data and subsequently tested in the independently withheld validation set.

RESULTS

Dysregulated Genes and Gene Sets in Brain

For our brain mega-analysis, 92 genes were dysregulated in SZ at an FDR $q < 0.10$ and two genes (*RHOBTB3* and *ABCA1*) reached a Bonferroni-corrected level of significance (Supplementary Table 3A). Among the 92 dysregulated genes, 73 genes were up-regulated and 19 genes were down-regulated in SZ (two-tailed sign test $p < 3.9\times10^{-9}$). Gene set analysis identified 745 sets (among 9254 examined) with at least one significant test hypothesis (Bonferroni $p < 0.05$); the vast majority of gene sets (720) were up-regulated among SZ cases, whereas one gene set showed a non-directional effect, and 19 gene sets showed a down-regulated effect (Supplementary Table 3B). The list of up-regulated sets included innate immune and inflammatory signaling pathways (TNF-α, NF-kB, p38 MAPK, IL-6 *via* STAT3, IL-2 *via* STAT5, IFN-γ, several TLR signaling cascades, protozoal infection, implicated in lupus); cellular stress responses (hypoxia, UV exposure, unfolded protein response, apoptosis/p53 cascades); response to androgens; metabolism of cholesterol and fatty acids; RNA metabolism and binding; several pathways related to cell survival, growth, and oncogenesis (*EGFR, ERBB2*, Insulin, *KRAS, MAPK/MEK, MTOR, MYC, PDGFR, PIGF, VEGF*); many gene sets targeted by miRNAs and transcription factors; genes involved in development and cellular differentiation; and several chromosomal loci, among others. Down-regulated gene sets included olfactory signaling pathways, genes with promotor CpG-site methylation in neural precursor cells, and several chromosomal loci (most notably 22q11). Five gene sets showed significant evidence for heterogeneity, such that approximately 20% of SZ cases appeared to drive the group main effect, and different individuals drove the effects for different functional sets (Figure 1, **Panel A**).

Dysregulated Genes and Gene Sets in Blood

We estimated the abundance of 17 leukocyte subtypes in SZ cases and unaffected comparison subjects and found no significant difference between groups, though a trend toward increased activated cytotoxic T cells was observed among SZ cases (uncorrected $p = 0.054$, Supplementary Table 2)

Within our blood mega-analysis, 2 238 genes were dysregulated in SZ at an FDR $q < 0.10$ and 220 reached a Bonferroni-corrected level of significance (Supplementary Table 4A). Among the 2 238 genes, 1 110 were up-regulated and 1 128 were down-regulated in SZ (two-tailed sign test $p = 0.66$); the absence of a systematic directional effect in the single-gene analysis did not preclude a directional effect at the level of gene sets, and gene set analysis identified 526 gene sets (among 9256 examined) with at least one significant test hypothesis (Bonferroni $p < 0.05$); the majority of gene sets (390) were up-regulated among SZ cases, whereas 21 gene sets showed a non-directional effect, and 115 gene sets showed a down-regulated effect (Supplementary Table 4B). The list of up-regulated sets included innate immune and inflammatory signaling pathways (TNF-α, NF-kB, IL-6, several TLR signaling cascades, protozoal infection, implicated in lupus); cellular stress responses

(hypoxia, UV exposure); response to androgens; glycolytic metabolism; several pathways related to cell survival, growth, and oncogenesis (*EGFR, ERBB2, PDGFR, PIGF, PTEN, VEGF*); many gene sets targeted by miRNAs and transcription factors; genes involved in development and cellular differentiation; and several chromosomal loci, among others. Down-regulated gene sets included those involved in DNA repair; metabolism and nonsense-mediated decay of mRNA; influenza RNA replication; citric acid cycle, mitochondrial function, and oxidative phosphorylation; ribosomal function and the regulation of protein translation; genes whose expression is typically driven by *MYC* and *EIF4E*; and several chromosomal loci, among others. One gene set (*MYC* targets) showed significant evidence for heterogeneity, such that approximately 50% of SZ cases appear to drive the group main effect (Figure 1, **Panel B**).

Cross-Tissue Comparison of Dysregulated Genes and Gene Sets

Cross-tissue comparison of dysregulated gene lists is depicted in Figure 2; 10 genes were dysregulated in both the brain and blood mega-analyses, reflecting a non-significant overlap (hypergeometric test $p = 0.68$). Seven genes were coordinately up-regulated in SZ across tissues ($p = 0.15$) and 1 gene was coordinately down-regulated. Two genes showed directionally discordant effects across tissues. The cross-tissue overlap of significantly dysregulated gene sets is depicted in Figure 1, **Panel A**. Two hundred and sixty-three gene sets were common to both tissues (Bonferroni $p = 1.7\times10^{-158}$); 255 of these sets were commonly up-regulated ($p = 2.8\times10^{-198}$) and 4 sets were commonly down-regulated ($p = 4.6\times10^{-4}$). One gene set showed directionally discordant effects across tissues. Gene sets that were dysregulated in both tissues are shown in **bold font** in Supplementary Tables 3B and 4B.

Enrichment of eQTL and GWAS Association Signals among Dysregulated Genes

eQTLs (NCBI eQTL Browser) were significantly enriched among dysregulated genes ($q < 0.1$) identified in the blood mega-analysis, but not those from the brain analysis (Supplementary Table 5). Specifically, 50 genes identified as dysregulated in the blood analysis previously showed evidence of eQTL regulation in lymphoblastoid cell lines ($p < 0.007$), while 607 showed evidence of eQTL regulation in the frontal cortex ($p < 0.001$); 20 genes were common to both lists, including *TMEM30A*, *SCAMP3*, *HSD17B12*, *TRPV2*, *BCR*, *SLC2A8*, *SLK*, *BCAT2*, *NUP93*, *DDX55*, *ANP32A*, *RTN4*, *ETS2*, *MDH2*, *DHRS1*, *UROS*, *MRPL43*, *HERPUD2*, *CYB561*, and *RAE1*. Within our lists of dysregulated genes in each tissue, we did not find significant quantitative enrichment with genes harboring (or located proximal to) SZ-associated SNPs (Ripke et al., 2014) as compared with randomly permuted gene lists of the same size (brain $p = 0.07$, blood $p = 0.99$). Among the 108 independent loci (located proximal to 311 genes) showing genome-wide significant association in the largest SZ GWAS from the Psychiatric Genomics Consortium (**PGC**; Ripke et al., 2014), we observed overlap with 1 differentially expressed brain gene (Clusterin, *CLU*, up-regulated in SZ cases) and 36 differentially expressed blood genes (hypergeometric overlap $p < 0.99$; Supplementary Table 6).

Network Co-expression Analysis identifies SZ-Associated Modules in Brain and Blood

We detected 21 modules of co-expressed genes in the brain, and while none were associated with SZ at a BH-corrected threshold of significance, three (arbitrarily labeled “green”, “salmon”, and “yellow”) were nominally associated and were examined in downstream analyses (Figures 3 and S2). The green module was diminished in SZ cases and enriched with synapse-and neuronal projection- and development-related genes (Figure 3D). This module also over-represented signatures of neuronal cell types known to express D1 and D2 dopaminergic receptors, and cortical neurons and immune cells (Figure 3C). The salmon module was enriched with immunologic terms and also over-represented signatures of cortical and cerebellar astrocytes, cerebellar oligodendrocytes, Bergman glia, and brain stem cholinergic motor neurons (Figure 3G and 3H). The yellow module was enriched with metabolic and electron transport function and over-represented cortical and cerebellar astrocytes, cerebellar oligodendrocytes, and Bergman glia, and cortical oligodendrocyte progenitors (Figure S2C and S2D).

We detected 33 modules in the blood, nine of which were associated with SZ at a corrected level of significance. Among these, “darkolivegreen”, “greenyellow”, “grey60”, “pink”, “turquoise”, and “yellow” were enhanced in SZ, while the “blue”, “steelblue”, and “cyan” modules were diminished among SZ cases (Figures S3 – S5). Notably, two enhanced modules (darkolivegreen and grey60) were enriched with innate immune function and over-represented granulocytes (Supplementary Table 7). The yellow module also over-represented the granulocyte and B lymphocyte signature. The steelblue module was associated with antigen presentation and self-recognition and over-represented natural killer cells. The pink module enriched with mitochondrial functions. The turquoise and greenyellow modules enriched with post-translational modifications and protein trafficking, with the former module showing over-representation of B lymphocytes, and also enriched in splicing complex function. All immune cell enrichments are shown in Supplementary Table 7. Hub genes and functional enrichments of SZ-associated blood modules are shown in Figure S4 and S5, respectively.

Enrichment of GWAS Association Signal in SZ-Associated Modules

The green and yellow brain modules, as well as the blue, grey60, turquoise, and yellow blood modules contained genes with known SNPs more strongly associated with SZ compared with randomly permuted genes (family-wise BH $p < 0.05$; Figure S6). Differentially expressed genes containing known SNPs that are significantly associated with SZ are indicated in Supplementary Table 6.

Cross-Tissue Overlap of Genes and Functional Annotations in SZ-Associated Modules

Cross-tissue comparison revealed overlapping genes between the green brain module and the blue and cyan blood modules (hypergeometric $p < 6.1\times10^{-14}$ and 1.94×10^{-7}); there was also significant sharing of genes between the salmon brain module and yellow blood module ($p < 1.3\times10^{-3}$; Supplementary Table 8A). A list of genes for modules that were significantly overlapped between brain and blood is provided in Supplementary Table 8B. The salmon brain module shared three functional annotations with two blood modules: genes mapping to

major histocompatibility complex region at cytoband 6p21.3 (cyan), and response to biotic stimulus and defense response (darkolivegreen).

Network Co-Expression Analysis Identifies Cross-Group and Cross-Tissue Preservation

Among the 26 modules that were identified in brain samples from unaffected comparison subjects, 24 were highly preserved and two were moderately preserved among SZ samples (Figure S7, **Panel A**); the least preserved module ("darkturquoise", *Z summary* = 5.9) was enriched with extracellular space, immune functions, cytokine (*CXCL1*, *CSF3*, *BMP2*, *CXCL2*, *IL12A*, *TNFSF9*, *CXCL10*) and growth factor activity (*CXCL1*, *CSF3*, *BMP2*, *ENDOU*, *IL12A*, *HBEGF*, *HGF*) and signal transducer activity. For the thirty-two modules identified in unaffected comparison subject blood samples, all were highly preserved among SZ samples (Figure S7, **Panel B**). The "skyblue" module was the least preserved (*Z summary* = 14) and was enriched with genes functioning in RNA processing and metabolism. In a contrast of brain and blood, 13 of 30 network modules identified in unaffected comparison subjects' brain samples showed strong evidence for preservation (*Z summary* > 10), 15 showed moderate preservation (10 < *Z summary* ≥ 2), and two showed no evidence for preservation (*Z summary* < 2; Figure S7, **Panel C**). The "pink" module identified in brain showed the strongest evidence of preservation in blood (*Z summary* = 25) and was enriched with genes associated with zinc finger transcription factors and nuclear components. The full list of significantly enriched biological annotations associated with abovementioned modules can be found in Supplementary Tables 9 – 11, respectively.

Machine-Learning Classification using Blood Transcriptome Data

The results of blood-based transcriptomic classification analyses are shown in Supplementary Table 12. Random Forest classifiers performed with high receiver operating characteristic area under the curve (AUC; 0.92 to 0.96) in the training matrix and retained moderate AUCs (0.72 to 0.77) in the independent validation matrix. Ensemble Support Vector Machine classifiers also performed with high AUCs in the training matrix (0.90 to 0.99) and moderate AUCs in the validation matrix (0.72 to 0.75). Assuming that a random binomial classifier (*e.g.*, coin-flip) would obtain no-better-than chance performance if employed within an identical bootstrapping and aggregation framework (AUC = 0.50), the validation sample predictions reflect better than chance performance (binomial test *p*-values ranging from 5×10^{-6} to 2×10^{-9}).

DISCUSSION

Our study detected many dysregulated genes that surpassed rigorous corrections for multiple testing, particularly in the larger blood dataset. In combination with gene set and network-based analyses, identified emergent biological functions robustly altered in SZ brain and peripheral blood transcriptome. Our study is preceded by a recent meta-analysis that sought to identify a cross-tissue signature of SZ using *postmortem* brain and *ex vivo* peripheral blood microarray data (Bergon et al., 2015). Comparing the differentially expressed genes identified in the present study with those reported by Bergon *et al.*, we observe only eight brain and 40 blood genes implicated by both studies; differences in the results could be attributed to the following differences between studies: (1) inclusion of different *postmortem*

brain regions (multiple regions *vs*. PFC-only in our analysis) and the number of blood studies included (more included in present study); (2) the total number of genes tested (more in the present study); (3) the approaches taken to reduce between-study variation; (4) the approach to statistical modeling and the specific covariates modeled; and (5) the thresholds for declaring significance. Despite these differences, we note broad similarities with respect to the biological pathways and functions implicated in each study (*e.g.*, genes that mediate immunologic functions, mitochondrial processes, and protein metabolism). However, it is important to acknowledge that functional annotation-based approaches will necessarily be limited to the biological terms that are well-annotated within their respective databases, and that many databases over-represent domains of cancer biology and immunology, which could contribute to bias in studies such as ours.

In addition, we note similarities between our mega-analysis of *postmortem* PFC homogenates and a prior microarray study of SZ based on laser-microdissections of dorsolateral PFC ($n = 24$ schizophrenia, $n = 12$ schizoaffective, $n = 24$ unaffected comparison; (Arion et al., 2015). From this study, we examined the top 35 differentially expressed genes detected in SZ cases in layers 3/5; among these, our PFC analyses also showed down-regulation of *DEF8*. . However, the effect we observed in PFC homogenates for *DEF8* was not strong enough to survive multiple testing (uncorrected Ps $< 6\times10^{-4}$, q-values < 0.12). We also cross-referenced their findings with the results from our blood analysis and found four genes with consistent down-regulation in SZ at a $p< 0.05$ (*TINF2*, *RPS10*, *NDUFA8*, and *EMG1*). These overlaps suggest that a subset of genes show generalizable differences across tissues and cell types.

We foresee transcript-level mega-analyses making important contributions to our knowledge of differential transcript expression and splicing patterns associated with SZ. However, there is an issue with combining data from multiple platforms that needs to be overcome for such an analysis to be feasible, namely the alignment of probe data across platforms. Local alignment methods such as BLAST can identify probe clusters with high sequence homology, however, alignment might not smooth out unwanted variation in probe design and libraries between platforms.

In the present study, genes dysregulated in SZ brain tissue were associated with diverse biological functions, but featured prominently among these were up-regulated inflammatory and cellular stress responses, cell growth and oncogenesis pathways, and metabolic pathways. These findings were recapitulated and further resolved by network analysis, which implicated three modules reflecting neurodevelopment (diminished in SZ), inflammation (enhanced), and lipid metabolism (enhanced), with the latter two modules enriched with markers of glial cells and the former enriched with neuronal cell types. Our mega-analysis helps clarify conflicting reports of NF-kB dysregulation (Rao et al., 2013; Roussos et al., 2013) by demonstrating that transcriptional targets of this signaling pathway are up-regulated in a large sample. Furthermore, our observations are consistent with the idea that excessive expression and signaling via damage/pathogen-associated molecular pattern receptors may contribute to brain inflammation in SZ (Fillman et al., 2013; Venkatasubramanian and Debnath, 2013).

The blood mega-analysis yielded more significantly dysregulated genes compared with the brain mega-analysis; this was a function of both more samples and a larger magnitude of effect sizes among the top 1% of genes ranked by *p*-value (|covariate adjusted mean difference|brain = 0.41 ± 0.05; |covariate adjusted mean difference|blood = 0.46 ± 0.05; *t*-test p-value < 2×10^{-22}); the observation that the transcriptomic signature of SZ is more prominent in blood tissue is a curious one that remains open to interpretation; one explanation may be that blood tissue simply allows for a wider range in the intensity of gene expression as compared with the brain. Alternatively the blood may more prominently reflect the effects of inadequately controlled covariates (*e.g.*, smoking, antipsychotics). Blood co-expression modules were generally preserved between SZ cases and unaffected comparison subjects and SZ samples, yet several modules identified in the full sample were associated with the SZ diagnosis and support the assertion that disturbances in innate immunity, antigen presentation, granulocytic, natural killer cells and lymphocytic functions are altered in SZ.

Relatively little cross-tissue overlap was observed at the level of dysregulated gene lists, yet our gene set and network-based approaches identified cross-tissue transcriptomic convergence, particularly with respect to innate immune functions, antigen presentation, cellular growth pathways, and common regulatory mechanisms (particularly the up-regulation of many miRNA targets, suggesting that miRNA-based regulation of gene expression may be deficient in SZ). Taken together, these findings suggest that different genes are dysregulated in each tissue, but that cross-tissue convergence may be observable at the level of emergent function.

Notably, one gene from the brain analysis and 36 genes from the blood analysis harbored SNPs that reached genome-wide significance in the largest available association meta-analysis of SZ from the PGC (Ripke et al., 2014). Identifiers of SZ-associated genes from GWAS can be found in Supplementary Table 6. We did not, however, observe significant enrichment of SZ GWAS signals within the list of dysregulated genes for either tissue, yet a network-based approach revealed that SZ-associated modules in both brain and blood were enriched with SZ GWAS association signal. Additionally, we observed significant over-representation of genes with lymphoblastoid cell and *postmortem* PFC eQTLs (NCBI eQTL Browser) within our lists of dysregulated genes in the SZ blood samples, allowing the possibility that some peripheral transcriptomic differences may be governed by genetic variants with known regulatory activity in both neural and immunologic cell types (Sanders et al., 2013). However, the majority of dysregulated genes in both tissues were not associated with known eQTLs or SZ-associated loci from GWAS studies, suggesting that genetic regulatory elements may play a relatively small, indirect, or developmentally dependent role in shaping SZ-associated transcriptomic differences.

Another essential outcome of our work is identification of numerous gene co-expression modules that are preserved across brain and blood tissues in non-psychotic individuals. These results align with the findings of our previous review on the topic (Tylee et al., 2013) and also supports the pursuit of blood-based transcriptomic classification tools for CNS disorders like SZ. Our machine-learning classifier work makes several important contributions to the psychiatric biomarker literature. To our knowledge, this was the first

attempt at classifier construction within a sample composed of multiple independent studies, reflecting different distributions of sex, age, ancestry, and medication usage. Our classifiers performed with moderate (approximately 70%) accuracy in an independently withheld sample composed of distinct studies (rather than a withheld subset of the same study sample); thus these classifiers appear robust to differences in experimental factors that vary across study sites (*e.g.*, microarray platform). Future studies should employ higher-resolution transcriptomic data (*i.e.*, RNA sequencing) and more sophisticated feature-selection algorithms to explore the upper limits of blood-based classification accuracy and to assess classifier specificity when discriminating different psychiatric conditions (*e.g.*, SZ *vs.* bipolar disorder); prospective transcriptomic studies could also be useful for predicting treatment response (Mamdani et al., 2011), thus paving the way for treatment selection biomarkers.

We identified hundreds of significant molecular signatures of SZ. According to evidence from previous genomic and transcriptomic studies (Cross Disorder Group of the Psychiatric Genomics Consortium, 2013; de Jong et al., 2016), it is possible that a fraction of molecular signatures are also represented in other psychiatric disorders. We did not empirically evaluate a separate disease group in our analysis to determine if these signatures are specific to SZ, however, we have compared our SZ-associated gene modules from brain and blood to a module associated with major depressive disorder in the literature (Jansen et al., 2015). We found a high degree of overlap between the 74 gene steelblue module and a 64 gene module associated with MDD. both from peripheral blood samples (overlap = 33 genes, hypergeometric p-value < 1.08E-64, Supplementary Table 13). Based on GO terminology, these 33 overlapping genes are involved in cellular defense response, chemotaxis, and immunity. The blue and turquoise modules from blood exhibited weaker overlap with the MDD-associated module (Ps < 0.68 and < 0.51). This finding lends support to the possibility that SZ-associated markers are not all constrained to diagnostic boundaries, which is a valuable finding for understanding pathophysiology of psychiatric disorders as a larger unit.

Accumulating evidence from various lines of research links immune dysregulation and SZ, including: (1) the most strongly implicated locus in the largest SZ GWAS study lies within the major histocompatibility complex (**MHC**) region (Ripke et al., 2014), which encodes genes involved in cellular antigen presentation and reflects a critical bridge between innate and adaptive immune functions; (2) increased prevalence of autoimmune disorders is found among individuals with SZ and their relatives in epidemiologic studies (Eaton et al., 2006); (3) increased levels of cytokines in peripheral blood (Miller et al., 2011) and cerebrospinal fluid of SZ patients are correlated with elevated levels of an endogenous NMDA receptor antagonist (Schwieler et al., 2015) and with cytoarchitectural and structural changes in the brain (Ellman et al., 2010; Fung et al., 2014); and (4) pharmacological evidence that antipsychotics dampen inflammation and may act as a restoration loop into dopaminergic circuitry (Kumarasinghe et al., 2013; Müller and Schwarz, 2010; Sugino et al., 2009). The present study demonstrated that many of the same inflammatory signaling cascades are up-regulated in both tissues, and also supports previous accounts of acute SZ-related changes in adaptive immune cells (Maino et al., 2007; Ripke et al., 2014; Steiner et al., 2010) and alterations in white matter and glial cell populations (Cotter et al., 2002; Duncan et al.,

2014). In light of this growing body of literature, it seems likely that global dysregulation of immune and inflammatory function is present in a subset of individuals with SZ and it is plausible that some genetic risk factors for SZ may exert their effects through immunologic cell types and their emergent functions. A new landmark finding in SZ genetics demonstrated that risk for SZ increases linearly with expression of the MHC region gene *C4*, which in turn leads to excessive synaptic pruning, thus providing clear mechanistic evidence that immune genes can mediate and perturb brain development (Sekar et al., 2016). Acute dysregulation of inflammatory signaling systems could also contribute to the SZ phenotype within the fully developed brain, through changes in neuroplasticity and neurotransmission. The implication of dysregulated inflammatory functions in the *postmortem* SZ brain underscores the need for interdisciplinary basic science research characterizing the cross-talk between these signaling cascades and those controlling typical neurodevelopmental processes and normal functioning in the fully developed brain.

In summary, our study makes several important contributions, including statistical improvements to obtain the best-estimate of SZ-associated differential expression, a thorough cross-tissue assessment of transcriptomic dysregulation in SZ, and generalizable classification of SZ cases and comparison subjects measured on different microarray chip technologies. However, the present study inherits many of the same limitations shared by all *postmortem* brain studies of SZ. These data are cross-sectional, representing transcriptomic profiles from a single time point, occurring after disease onset, death, and (probably in the vast majority of cases) years of anti-psychotic treatment. As such, it is not possible to know whether the observed differences are causal contributors or downstream consequences of SZ pathophysiology. While we attempted to control for the effects of medication, we must acknowledge the likelihood that some observations currently attributed to diagnostic status were influenced by group differences in medication use or other uncontrolled covariates (*e.g.*, tobacco use). Studies in animal models, or antipsychotic-naïve, smoking-matched subjects will be essential for resolving these possibilities. Despite these limitations, this study provides an atlas of dysregulated genes, biological processes, and co-expression networks associated with SZ in brain and blood tissues.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Role of Funding Sources

The sponsors did not influence the design, interpreting, writing, or decision to submit for publication.

Jonathan L. Hess, Daniel S. Tylee, Rahul Barve, and Stephen J. Glatt were supported by grants from the U.S. National Institutes of Health, the Gerber Foundation, the Sidney R. Baer, Jr. Foundation, and NARSAD: The Brain & Behavior Research Foundation. Microarray for *postmortem* brain samples were supplied by the following: Stanley Medical Research Institute (Sanghyeon Kim), Tadafumi Kato, Vahram Haroutunian, Maxwell Bustamante, Kristen Brennand, and Fred Gage. Blood-based gene expression data were provided by the Neurobehavioral Genetics Unit (Chief Investigators: Vaughan Carr, Rodney Scott, Brian Kelly, Paul Tooney, Murray Cairns; Associate Investigators: Christopher Oldmeadow, Jing Qin Wu), which is supported by a grant from the New South Wales Ministry of Health administered by the Hunter Medical Research Institute. Nishantha Kumarasinghe was supported by a doctoral scholarship from the World Health Organization, as well as financial and in-kind support

from the Schizophrenia Research Institute, Australia utilizing funding from the NSW Ministry of Health, the University of New Castle, Australia and University of Sri Jayewardenpura, Sri Lanka. Blood-based gene expression data were provided by Simone de Jong of King's College London and Roel Ophoff of the University of California, Los Angeles.

References

Abbas AR, Wolslegel K, Seshasayee D, Modrusan Z, Clark HF. Deconvolution of blood microarray data identifies cellular activation patterns in systemic lupus erythematosus. PLoS One. 2009; 4(7):e6098. [PubMed: 19568420]

Åberg K, Saetre P, Lindholm E, Ekholm B, Pettersson U, Adolfsson R, Jazin E. Human QKI, a new candidate gene for schizophrenia involved in myelination. Am. J. Med. Genet. - Neuropsychiatr. Genet. 2006; 141 B(1):84–90.

Arion D, Corradi JP, Tang S, Datta D, Boothe F, He a, Cacace a M, Zaczek R, Albright CF, Tseng G, Lewis Da. Distinctive transcriptome alterations of prefrontal pyramidal neurons in schizophrenia and schizoaffective disorder. Mol. Psychiatry. 2015; 20(11):1397–1405. [PubMed: 25560755]

Arnedo J, Svrakic DM, Del Val C, Romero-Zaliz R, Hernendez-Cuervo H, Fanous AH, Pato MT, Pato CN, De Erausquin GA, Robert Cloninger C, Zwir I. Uncovering the hidden risk architecture of the schizophrenias: Confirmation in three independent genome-wide association studies. Am. J. Psychiatry. 2015; 172(2):139–153. [PubMed: 25219520]

Bergon A, Belzeaux R, Comte M, Pelletier F, Herve M, Gardiner EJ, Beveridge NJ, Liu B, Carr V, Scott RJ, Kelly B, Cairns MJ, Kumarasinghe N, Schall U, Blin O, Boucraut J, Tooney PA, Fakra E, Ibrahim EC. CX3CR1 is dysregulated in blood and brain from schizophrenia patients. Schizophr. Res. 2015; 168(1–2):434–443. [PubMed: 26285829]

Cotter D, Mackay D, Chana G, Beasley C, Landau S, Everall IP. Reduced neuronal size and glial cell density in area 9 of the dorsolateral prefrontal cortex in subjects with major depressive disorder. Cereb. Cortex. 2002; 12(4):386–394. [PubMed: 11884354]

Dean B. Understanding the role of inflammatory-related pathways in the pathophysiology and treatment of psychiatric disorders: evidence from human peripheral studies and CNS studies. Int. J. Neuropsychopharmacol. 2011; 14(7):997–1012. [PubMed: 21156092]

Devor EJ, Waziri R. A familial/genetic study of plasma serine and glycine concentrations. Biol. Psychiatry. 1993; 34(4):221–225. [PubMed: 8399818]

Dougherty JD, Schmidt EF, Nakajima M, Heintz N. Analytical approaches to RNA profiling data for the identification of genes enriched in specific cells. Nucleic Acids Res. 2010; 38(13):4218–4230. [PubMed: 20308160]

Duncan LE, Holmans PA, Lee PH, O'Dushlaine CT, Kirby AW, Smoller JW, Öngür D, Cohen BM. Pathway analyses implicate glial cells in schizophrenia. PLoS One. 2014; 9(2):e89441. [PubMed: 24586781]

Eaton WW, Byrne M, Ewald H, Mors O, Chen C-Y, Agerbo E, Mortensen PB. Association of schizophrenia and autoimmune diseases: linkage of Danish national registers. Am. J. Psychiatry. 2006; 163(3):521–528. [PubMed: 16513876]

Ellman LM, Deicken RF, Vinogradov S, Kremen WS, Poole JH, Kern DM, Tsai WY, Schaefer CA, Brown AS. Structural brain alterations in schizophrenia following fetal exposure to the inflammatory cytokine interleukin-8. Schizophr. Res. 2010; 121(1–3):46–54. [PubMed: 20553865]

Fillman SG, Cloonan N, Catts VS, Miller LC, Wong J, McCrossin T, Cairns M, Weickert CS. Increased inflammatory markers identified in the dorsolateral prefrontal cortex of individuals with schizophrenia. Mol. Psychiatry. 2013; 18(2):206–214. [PubMed: 22869038]

Fineberg AM, Ellman LM. Inflammatory cytokines and neurological and neurocognitive alterations in the course of schizophrenia. Biol. Psychiatry. 2013; 73(10):951–966. [PubMed: 23414821]

Fung SJ, Joshi D, Fillman SG, Weickert CS. High white matter neuron density with elevated cortical cytokine expression in schizophrenia. Biol. Psychiatry. 2014; 75(4):e5–e7. [PubMed: 23830667]

Glatt SJ, Tsuang MT, Winn M, Chandler SD, Collins M, Lopez L, Weinfeld M, Carter C, Schork N, Pierce K, Courchesne E. Blood-based gene expression signatures of infants and toddlers with autism. J. Am. Acad. Child Adolesc. Psychiatry. 2012; 51(9):934–944. [PubMed: 22917206]

Glatt SJ, Tylee DS, Chandler SD, Pazol J, Nievergelt CM, Woelk CH, Baker DG, Lohr JB, Kremen WS, Litz BT, Tsuang MT. Blood-based gene-expression predictors of PTSD risk and resilience among deployed marines: a pilot study. Am. J. Med. Genet. B. Neuropsychiatr. Genet. 2013; 162B(4):313–326. [PubMed: 23650250]

Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. 2009; 4(1):44–57. [PubMed: 19131956]

Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics. 2003; 4(2):249–264. [PubMed: 12925520]

Kumarasinghe N, Beveridge NJ, Gardiner E, Scott RJ, Yasawardene S, Perera A, Mendis J, Suriyakumara K, Schall U, Tooney PA. Gene expression profiling in treatment-naive schizophrenia patients identifies abnormalities in biological pathways involving AKT1 that are corrected by antipsychotic medication. Int. J. Neuropsychopharmacol. 2013; 16(7):1483–1503. [PubMed: 23442539]

Langfelder P, Luo R, Oldham MC, Horvath S. Is my network module preserved and reproducible? PLoS Comput. Biol. 2011; 7(1):e1001057. [PubMed: 21283776]

Lottaz C, Toedling J, Spang R. Annotation-based distance measures for patient subgroup discovery in clinical microarray studies. Bioinformatics. 2007; 23(17):2256–2264. [PubMed: 17586546]

Maino K, Gruber R, Riedel M, Seitz N, Schwarz M, Müller N. T- and B-lymphocytes in patients with schizophrenia in acute psychotic episode and the course of the treatment. Psychiatry Res. 2007; 152(2–3):173–180. [PubMed: 17448544]

Mamdani F, Berlim MT, Beaulieu M-M, Labbe A, Merette C, Turecki G. Gene expression biomarkers of response to citalopram treatment in major depressive disorder. Transl. Psychiatry. 2011; 1:e13. [PubMed: 22832429]

Middleton FA, Mirnics K, Pierri JN, Lewis DA, Levitt P. Gene expression profiling reveals alterations of specific metabolic pathways in schizophrenia. J. Neurosci. 2002; 22(7):2718–2729. [PubMed: 11923437]

Miller BJ, Buckley P, Seabolt W, Mellor A, Kirkpatrick B. Meta-analysis of cytokine alterations in schizophrenia: Clinical status and antipsychotic effects. Biol. Psychiatry. 2011; 70(7):663–671. [PubMed: 21641581]

Mirnics K, Levitt P, Lewis DA. Critical appraisal of DNA microarrays in psychiatric genomics. Biol. Psychiatry. 2006; 60(2):163–176. [PubMed: 16616896]

Mistry M, Gillis J, Pavlidis P. Meta-analysis of gene coexpression networks in the post-mortem prefrontal cortex of patients with schizophrenia and unaffected controls. BMC Neurosci. 2013a; 14:105. [PubMed: 24070017]

Mistry M, Gillis J, Pavlidis P. Genome-wide expression profiling of schizophrenia using a large combined cohort. Mol. Psychiatry. 2013b; 18(2):215–225. [PubMed: 22212594]

Mistry M, Pavlidis P. A cross-laboratory comparison of expression profiling data from normal human postmortem brain. Neuroscience. 2010; 167(2):384–395. [PubMed: 20138973]

Müller N, Schwarz MJ. Immune System and Schizophrenia. Curr. Immunol. Rev. 2010; 6(3):213–220. [PubMed: 21057585]

National Center for Biotechnology Information [NCBI] eQTL Browser. US Natl. Libr. Med. http://www.ncbi.nlm.nih.gov/projects/gap/eqtl/index.cgi

Pérez-Santiago J, Diez-Alarcia R, Callado LF, Zhang JX, Chana G, White CH, Glatt SJ, Tsuang MT, Everall IP, Meana JJ, Woelk CH. A combined analysis of microarray gene expression studies of the human prefrontal cortex identifies genes implicated in schizophrenia. J. Psychiatr. Res. 2012; 46(11):1464–1474. [PubMed: 22954356]

Potvin S, Stip E, Sepehry AA, Gendron A, Bah R, Kouassi E. Inflammatory Cytokine Alterations in Schizophrenia: A Systematic Quantitative Review. Biol. Psychiatry. 2008; 63(8):801–808. [PubMed: 18005941]

Rao JS, Kim H-W, Harry GJ, Rapoport SI, Reese EA. Increased neuroinflammatory and arachidonic acid cascade markers, and reduced synaptic proteins, in the postmortem frontal cortex from schizophrenia patients. Schizophr. Res. 2013; 147(1):24–31. [PubMed: 23566496]

Ripke S, Neale BM, Corvin A, Walters JTR, Farh K-H, Holmans PA, Lee P, Bulik-Sullivan B, Collier DA, Huang H, Pers TH, Agartz I, Agerbo E, Albus M, Alexander M, Amin F, Bacanu SA, Begemann M, Belliveau RA Jr, Bene J, Bergen SE, Bevilacqua E, Bigdeli TB, Black DW, Bruggeman R, Buccola NG, Buckner RL, Byerley W, Cahn W, Cai G, Campion D, Cantor RM, Carr VJ, Carrera N, Catts SV, Chambert KD, Chan RCK, Chen RYL, Chen EYH, Cheng W, Cheung EFC, Ann Chong S, Robert Cloninger C, Cohen D, Cohen N, Cormican P, Craddock N, Crowley JJ, Curtis D, Davidson M, Davis KL, Degenhardt F, Del Favero J, Demontis D, Dikeos D, Dinan T, Djurovic S, Donohoe G, Drapeau E, Duan J, Dudbridge F, Durmishi N, Eichhammer P, Eriksson J, Escott-Price V, Essioux L, Fanous AH, Farrell MS, Frank J, Franke L, Freedman R, Freimer NB, Friedl M, Friedman JI, Fromer M, Genovese G, Georgieva L, Giegling I, Giusti-Rodríguez P, Godard S, Goldstein JI, Golimbet V, Gopal S, Gratten J, de Haan L, Hammer C, Hamshere ML, Hansen M, Hansen T, Haroutunian V, Hartmann AM, Henskens FA, Herms S, Hirschhorn JN, Hoffmann P, Hofman A, Hollegaard MV, Hougaard DM, Ikeda M, Joa I, Julià A, Kahn RS, Kalaydjieva L, Karachanak-Yankova S, Karjalainen J, Kavanagh D, Keller MC, Kennedy JL, Khrunin A, Kim Y, Klovins J, Knowles JA, Konte B, Kucinskas V, Ausrele Kucinskiene Z, Kuzelova-Ptackova H, Kähler AK, Laurent C, Lee Chee Keong J, Hong Lee S, Legge SE, Lerer B, Li M, Li T, Liang K-Y, Lieberman J, Limborska S, Loughland CM, Lubinski J, Lönnqvist J, Macek Jr M, Magnusson PKE, Maher BS, Maier W, Mallet J, Marsal S, Mattheisen M, Mattingsdal M, McCarley RW, McDonald C, McIntosh AM, Meier S, Meijer CJ, Melegh B, Melle I, Mesholam-Gately RI, Metspalu A, Michie PT, Milani L, Milanova V, Mokrab Y, Morris DW, Mors O, Murphy KC, Murray RM, Myin-Germeys I, Müller-Myhsok B, Nelis M, Nenadic I, Nertney DA, Nestadt G, Nicodemus KK, Nikitina-Zake L, Nisenbaum L, Nordin A, O'Callaghan E, O'Dushlaine C, O'Neill FA, Oh S-Y, Olincy A, Olsen L, Van Os J, Endophenotypes International Consortium P, Pantelis C, Papadimitriou GN, Papiol S, Parkhomenko E, Pato MT, Paunio T, Pejovic-Milovancevic M, Perkins DO, Pietiläinen O, Pimm J, Pocklington AJ, Powell J, Price A, Pulver AE, Purcell SM, Quested D, Rasmussen HB, Reichenberg A, Reimers MA, Richards AL, Roffman JL, Roussos P, Ruderfer DM, Salomaa V, Sanders AR, Schall U, Schubert CR, Schulze TG, Schwab SG, Scolnick EM, Scott RJ, Seidman LJ, Shi J, Sigurdsson E, Silagadze T, Silverman JM, Sim K, Slominsky P, Smoller JW, So H-C, Spencer CA, Stahl EA, Stefansson H, Steinberg S, Stogmann E, Straub RE, Strengman E, Strohmaier J, Scott Stroup T, Subramaniam M, Suvisaari J, Svrakic DM, Szatkiewicz JP, Söderman E, Thirumalai S, Toncheva D, Tosato S, Veijola J, Waddington J, Walsh D, Wang D, Wang Q, Webb BT, Weiser M, Wildenauer DB, Williams NM, Williams S, Witt SH, Wolen AR, Wong EHM, Wormley BK, Simon Xi H, Zai CC, Zheng X, Zimprich F, Wray NR, Stefansson K, Visscher PM, Trust Case-Control Consortium W, Adolfsson R, Andreassen OA, Blackwood DHR, Bramon E, Buxbaum JD, Børglum AD, Cichon S, Darvasi A, Domenici E, Ehrenreich H, Esko T, Gejman PV, Gill M, Gurling H, Hultman CM, Iwata N, Jablensky AV, Jönsson EG, Kendler KS, Kirov G, Knight J, Lencz T, Levinson DF, Li QS, Liu J, Malhotra AK, McCarroll SA, McQuillin A, Moran JL, Mortensen PB, Mowry BJ, Nöthen MM, Ophoff RA, Owen MJ, Palotie A, Pato CN, Petryshen TL, Posthuma D, Rietschel M, Riley BP, Rujescu D, Sham PC, Sklar P, St Clair D, Weinberger DR, Wendland JR, Werge T, Daly MJ, Sullivan PF, O'Donovan MC. Biological insights from 108 schizophrenia-associated genetic loci. Nature. 2014; 511(7510):421–427. [PubMed: 25056061]

Roussos P, Katsel P, Davis KL, Giakoumaki SG, Lencz T, Malhotra AK, Siever LJ, Bitsios P, Haroutunian V. Convergent findings for abnormalities of the NF-κB signaling pathway in schizophrenia. Neuropsychopharmacology. 2013; 38(3):533–539. [PubMed: 23132271]

Sanders AR, Göring HHH, Duan J, Drigalenko EI, Moy W, Freda J, He D, Shi J, Gejman PV. Transcriptome study of differential expression in schizophrenia. Hum. Mol. Genet. 2013; 22(24): 5001–5014. [PubMed: 23904455]

Schwieler L, Larsson MK, Skogh E, Kegel ME, Orhan F, Abdelmoaty S, Finn A, Bhat M, Samuelsson M, Lundberg K, Dahl M-L, Sellgren C, Schuppe-Koistinen I, Svensson C, Erhardt S, Engberg G. Increased levels of IL-6 in the cerebrospinal fluid of patients with chronic schizophrenia--significance for activation of the kynurenine pathway. J. Psychiatry Neurosci. 2015; 40(2):126–133. [PubMed: 25455350]

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. [PubMed: 23945090]

Sekar A, Bialas AR, de Rivera H, Davis A, Hammond TR, Kamitaki N, Tooley K, Presumey J, Baum M, Van Doren V, Genovese G, Rose SA, Handsaker RE, Daly MJ, Carroll MC, Stevens B, McCarroll SA. Schizophrenia risk from complex variation of complement component 4. Nature. 2016; 530(7589):177–183. [PubMed: 26814963]

Steiner J, Jacobs R, Panteli B, Brauner M, Schiltz K, Bahn S, Herberth M, Westphal S, Gos T, Walter M, Bernstein H-G, Myint AM, Bogerts B. Acute schizophrenia is accompanied by reduced T cell and increased B cell immunity. Eur. Arch. Psychiatry Clin. Neurosci. 2010; 260(7):509–518. [PubMed: 20107825]

Storey JD. The positive false discovery rate: a Bayesian interpretation and the q -value. Ann. Stat. 2003; 31(6):2013–2035.

Sugino H, Futamura T, Mitsumoto Y, Maeda K, Marunaka Y. Atypical antipsychotics suppress production of proinflammatory cytokines and up-regulate interleukin-10 in lipopolysaccharide-treated mice. Prog. Neuropsychopharmacol. Biol. Psychiatry. 2009; 33(2):303–307. [PubMed: 19138716]

Tsuang MT, Faraone SV. The case for heterogeneity in the etiology of schizophrenia. Schizophr. Res. 1995; 17(2):161–175. [PubMed: 8562491]

Tsuang MT, Nossova N, Yager T, Tsuang M-M, Guo S-C, Shyu KG, Glatt SJ, Liew CC. Assessing the validity of blood-based gene expression profiles for the classification of schizophrenia and bipolar disorder: a preliminary report. Am. J. Med. Genet. B. Neuropsychiatr. Genet. 2005; 133B(1):1–5. [PubMed: 15645418]

Tylee DS, Chandler SD, Nievergelt CM, Liu X, Pazol J, Woelk CH, Lohr JB, Kremen WS, Baker DG, Glatt SJ, Tsuang MT. Blood-based gene-expression biomarkers of post-traumatic stress disorder among deployed marines: A pilot study. Psychoneuroendocrinology. 2015; 51:472–494. [PubMed: 25311155]

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(7):595–603. [PubMed: 24132893]

Väremo L, Nielsen J, Nookaew I. Enriching the gene set analysis of genome-wide data by incorporating directionality of gene expression and combining statistical hypotheses and methods. Nucleic Acids Res. 2013; 41(8):4378–4391. [PubMed: 23444143]

Venkatasubramanian G, Debnath M. The TRIPS (Toll-like receptors in immuno-inflammatory pathogenesis) Hypothesis: a novel postulate to understand schizophrenia. Prog. Neuropsychopharmacol. Biol. Psychiatry. 2013; 44:301–311. [PubMed: 23587629]

Watkins NA, Gusnanto A, de Bono B, De S, Miranda-Saavedra D, Hardie DL, Angenent WGJ, Attwood AP, Ellis PD, Erber W, Foad NS, Garner SF, Isacke CM, Jolley J, Koch K, Macaulay IC, Morley SL, Rendon A, Rice KM, Taylor N, Thijssen-Timmer DC, Tijssen MR, van der Schoot CE, Wernisch L, Winzer T, Dudbridge F, Buckley CD, Langford CF, Teichmann S, Göttgens B, Ouwehand WH. A HaemAtlas: characterizing gene expression in differentiated human blood cells. Blood. 2009; 113(19):e1–e9. [PubMed: 19228925]

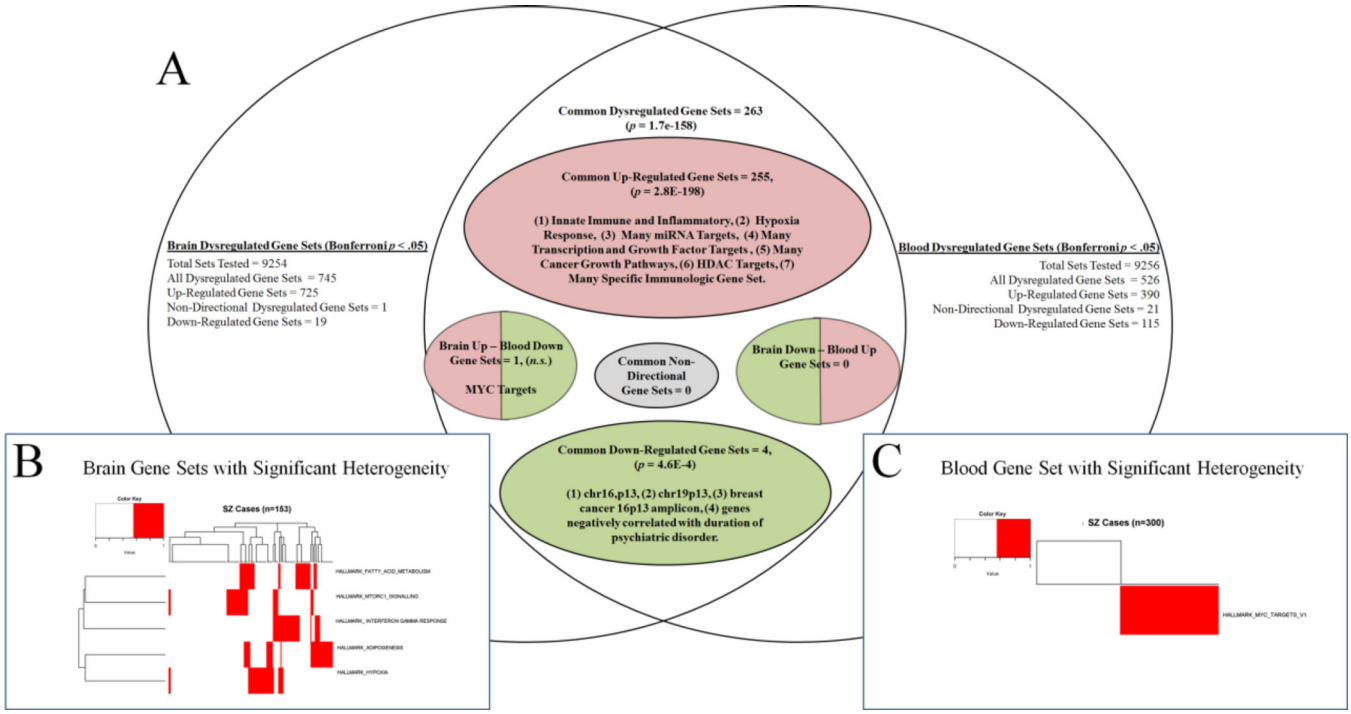

Figure 1. Cross-Tissue Comparison of Significantly Dysregulated Gene Sets (Bonferroni p <0.05)

The results of the permutation-based gene-set analyses are shown (Panel A), highlighting gene sets that were significantly dysregulated in both tissues. A total of 745 gene sets were dysregulated (out of 9254 tested) based on single-gene test statistics from the brain mega-analysis. A total of 526 gene sets were dysregulated (out of 9256 tested) based on single-gene test statistics from the blood mega-analysis. For the purpose of cross-tissue comparison, gene sets with either an absolute effect (*i.e.*, all genes in the target set) or a mixed effect (*i.e.*, a subset of the genes in the target set) were considered to be directionally dysregulated. Among the dysregulated gene sets, 263 were common to both tissues (Bonferroni-corrected hypergeometric $p < 1.7\times10^{-158}$, and among these, 255 showed evidence of up-regulation across tissues ($p < 2.8\times10^{-198}$), while only 4 showed evidence of down-regulation across tissues ($p < 4.6\times10^{-4}$). Detailed methods for gene set analysis and the reference databases can be found in the Supplementary Methods. Among dysregulated gene sets corresponding to the Molecular Signature Database's (Broad Institute) Hallmark category, we assessed whether SZ cases showed significant evidence for heterogeneity using a previously developed approach described in the Supplementary Methods. Within the brain data, we observed significant 2-group clustering of SZ cases based on the expression values corresponding to 5 gene sets showing a main-effect of upregulation in SZ (Panel B); cases are shown such that the individuals belonging to the cluster driving the up-regulation effect are depicted in red. These results suggest that different SZ cases contribute to the observed dysregulation in distinct biological pathways. Within the blood data, we observed significant 2-group clustering for a single gene set which showed a main-effect of down-regulation among SZ cases (Panel C).

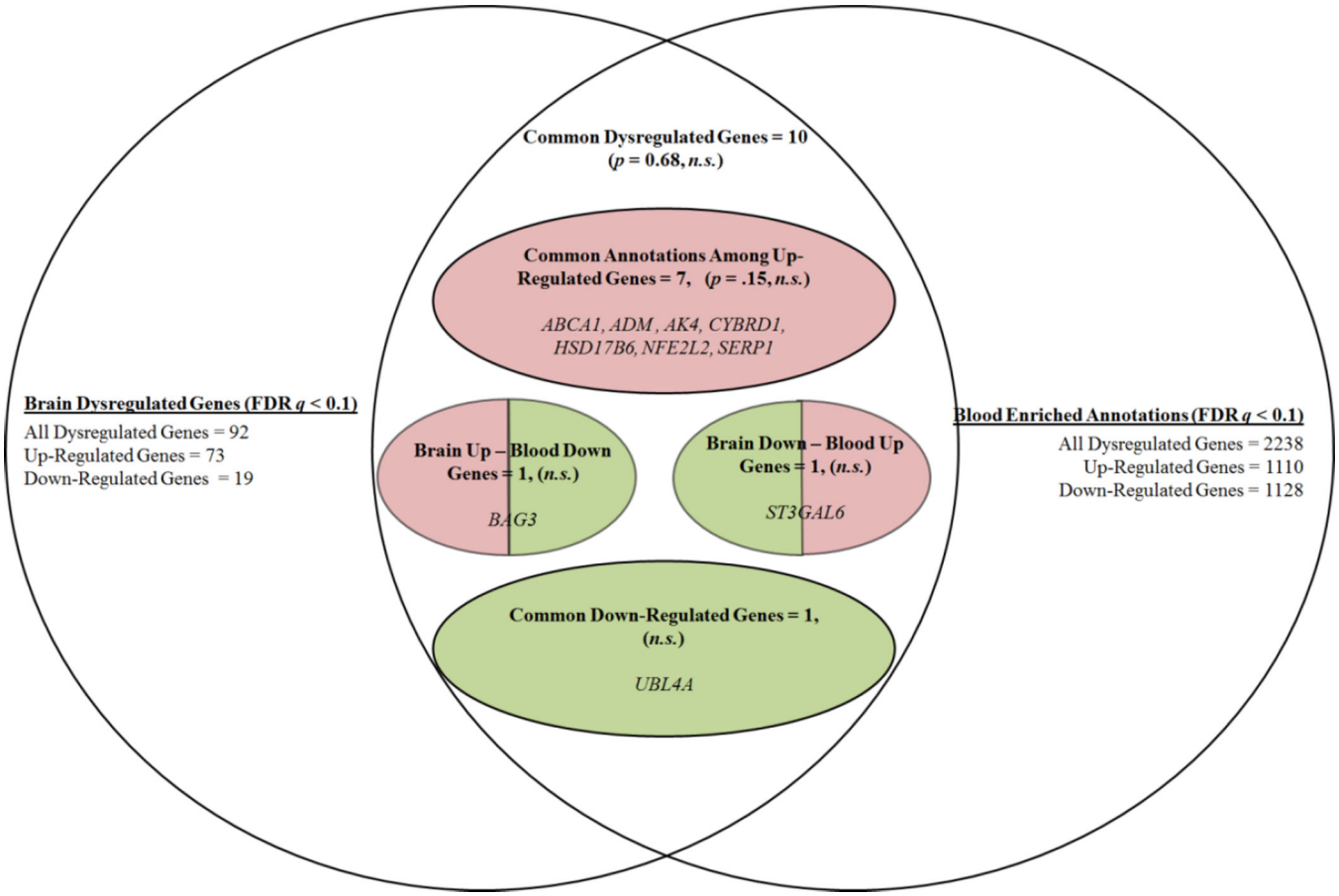

Figure 2. Cross-Tissue Comparison of Significantly Dysregulated Genes (FDR *q* <0.05)

(**A**) Based on the mega-analyses, we identified the most significantly dysregulated genes in brain (n = 92) and blood (n = 2238) at a relatively conservative threshold (FDR $q < .10$). A total of 10 genes were common to both lists; 7 genes were coordinately up-regulated and 1 gene was coordinately down-regulated across tissues. The degree of cross-tissue overlap for each of the displayed intersections was non-significant based on hypergeometric test statistics.

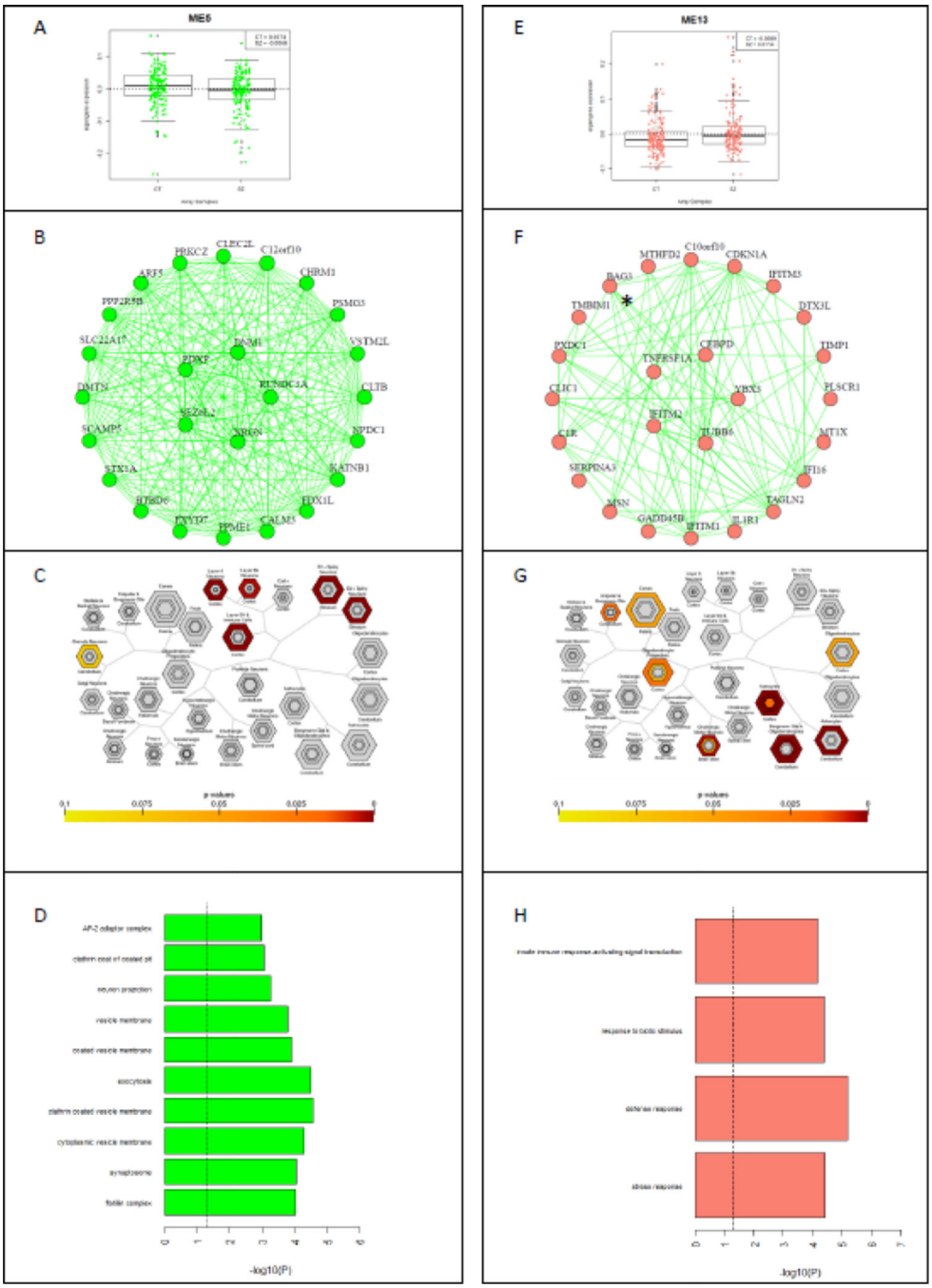

Figure 3. Co-expression Modules Nominally Associated with SZ in Brain ($p < 0.05$)

Comparison of module eigengene expression values (unadjusted for covariates) between SZ cases and unaffected comparisons within the "green" and "salmon" co-expression modules identified by the *WGCNA* R package (**A** and **E**, respectively), which were nominally associated with SZ from linear mixed model (uncorrected $p < 0.05$). We cross-referenced the set of dysregulated genes identified in the brain mega-analysis ($q < .1$) with the top 25 genes in each module ranked by intramodular connectivity (overlaps denoted by asterisk *) (**B** and **F**). In panels B and F, the top 5 "hub" genes are found in the innermost circle. Modules were

biologically characterized by testing for enrichment of brain cell-type signatures (**C** and **G**) and annotations by a pathway-based approach (**D** and **H**). In panels D and H, annotations that surpassed a BH $p < 0.05$ (cutoff depicted by vertical dotted line) from hypergeometric tests are shown (represented as −log10[P]).

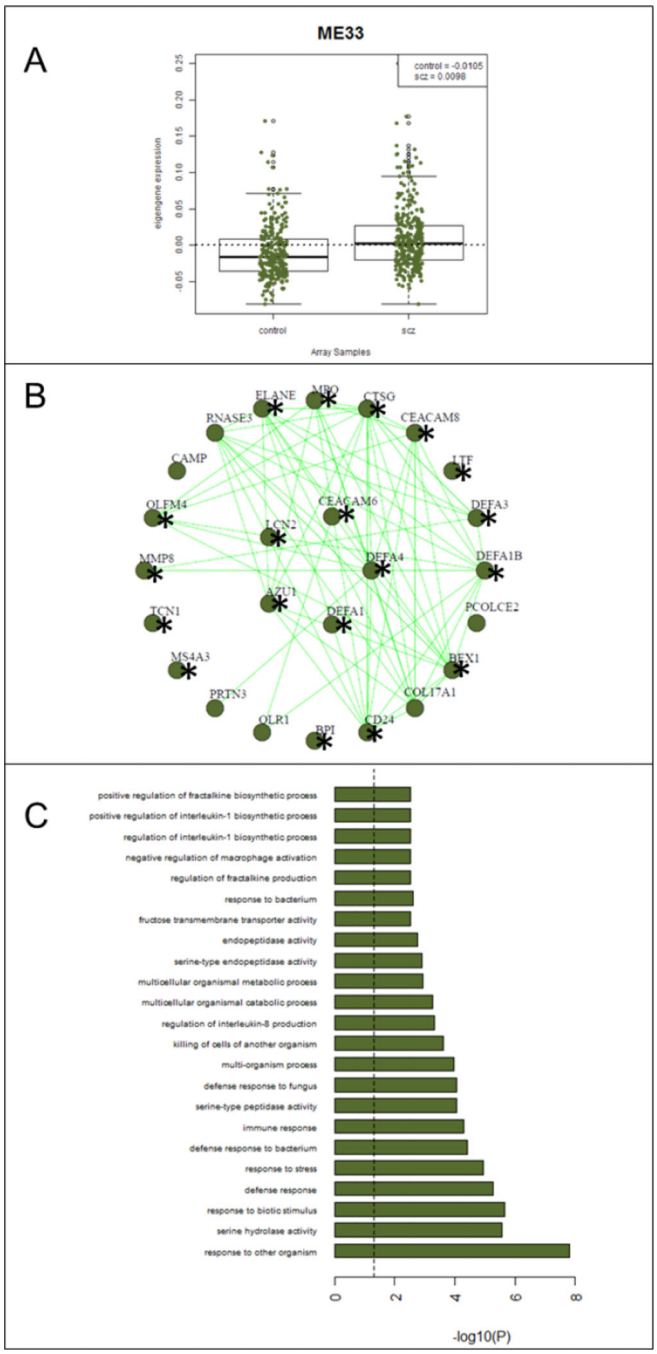

Figure 4. A Co-expression Module that was Significantly Associated with SZ in Blood (BH $p < 0.05$)

Comparison of module eigengene expression values (unadjusted for covariates) between SZ cases and unaffected comparisons within the "darkolivegreen" co-expression module identified by the *WGCNA* R package (**A**), which was significantly associated with SZ based on linear mixed model test (BH $p < 4.4\times10^{-6}$). We cross-referenced the set of dysregulated genes identified in the blood mega-analysis ($q < .1$) with the top 25 genes ranked by intramodular connectivity in this module (overlaps denoted by asterisk *) (**B**). In panel B,

the top 5 “hub” genes are found in the innermost circle. To biological characterize this module, a pathway-based approach was used to test for significant enrichment biological annotations mapping to “darkolivegreen” genes (**C**). In panel C, annotations that surpassed a BH $p < 0.05$ (cutoff depicted by vertical dotted line) from hypergeometric tests are shown (represented as −log10[P]).

Table 1

Schizophrenia Microarray Studies Included in Mega-Analysis.

Brain Studies	**Array Type**	**Cases (*n*)**	**Controls (*n*)**	**% Female**	**% Medicated**	**Predominant Ancestry**	**Genes Analyzed**
Altar -Stanley Medical Research Institute (SMRI) – Collection A	Affymetrix U133a	8	10	22%	100%	Caucasian	12 153
Altar - SMRI - Collection C	Affymetrix U133a	11	11	32%	100%	Caucasian	12 366
Dobrin - SMRI Collection	Affymetrix U133 Plus 2.0	24	25	27%	100%	Caucasian	20 286
Cohen *et al.*, 2009	Affymetrix HG 1.0 ST	4	4	0%	100%	NA	17 168
Glatt *et al.*, 2005	Affymetrix U133a	16	25	27%	100%	Caucasian	12 410
Maycox *et al.*, 2009	Affymetrix U133 Plus 2.0	28	23	40%	NA	Caucasian	20 767
Narayan *et al.*, 2008	Affymetrix U133 Plus 2.0	28	28	16%	NA	NA	20 766
Lanz (GSE53987)	Affymetrix U133 Plus 2.0	14	19	48%	NA	Caucasian	20 767
Katsel *et al.*, 2005	Affymetrix U133a	20	17	38%	NA	NA	13 832
Total: 9		153	162				20 767
Blood Studies	**Array Type**	**Cases (*n*)**	**Controls (*n*)**	**% Female**	**% Medicated**	**Predominant Ancestry**	**Genes Analyzed**
Tsuang *et al.*, 2005	Affymetrix U133 Plus 2.0	10	15	48%	36%	Asian	22 014
Tsuang *et al.*, 2005	Affymetrix U133a	20	9	60%	64%	Asian	13 977
Glatt *et al.*, 2009	Affymetrix Exon 1.0 ST	13	8	33%	57%	Caucasian	18 850
Glatt *et al.*, 2011	Affymetrix U133 Plus 2.0	8	12	50%	40%	Caucasian	22 014
Van Beveren *et al.*, 2012	Affymetrix U133 Plus 2.0	41	29	0%	NA	NA	22 014

Brain Studies	Array Type	Cases (*n*)	Controls (*n*)	% Female	% Medicated	Predominant Ancestry	Genes Analyzed
de Jong *et al.*, 2012	Illumina HumanHT-8 v3	15	21	27%	0%	Caucasian	18 516
de Jong *et al.*, 2012	Illumina HumanHT-12 v3	106	95	41%	45%	Caucasian	20 156
Kumarasinghe *et al*, 2013	Illumina HumanHT-12 v3	9	11	40%	0%	Asian	20 156
Gardiner *et al.*, 2013	Illumina HumanHT-12 v3	78	78	48%	48%	Caucasian	20 156
Total: 9		300	278				19 737*

NA = Not Available

*Among the 23 755 mixed-effect models assessed in blood studies, only 19 737 converged to produce valid statistical results; failures to converge were related to the number of missing values for a particular gene's expression level.