Citation for published version (APA):
Immune signatures and disorder-specific patterns in a cross-disorder gene expression analysis


Background
Recent studies point to overlap between neuropsychiatric disorders in symptomatology and genetic aetiology.

Aims
To systematically investigate genomics overlap between childhood and adult attention-deficit hyperactivity disorder (ADHD), autism spectrum disorder (ASD) and major depressive disorder (MDD).

Method
Analysis of whole-genome blood gene expression and genetic risk scores of 318 individuals. Participants included individuals affected with adult ADHD (n = 93), childhood ADHD (n = 17), MDD (n = 63), ASD (n = 51), childhood dual diagnosis of ADHD–ASD (n = 16) and healthy controls (n = 78).

Results
Weighted gene co-expression analysis results reveal disorder-specific signatures for childhood ADHD and MDD, with moderate overlap between MDD and schizophrenia, bipolar disorder and ADHD and low, but significant, overlap between schizophrenia and ASD. These results suggest that some shared molecular genetic factors underlie a significant proportion of the risk for development of several psychiatric disorders. To date, however, it is not known whether these genomic overlaps are reflected in overlapping patterns of gene expression in tissues. In this study we therefore explored whole blood gene expression across adult and childhood ADHD, ASD, MDD and healthy controls, using weighted gene co-expression analysis to search for patterns of correlated gene expression between disorders as well as disorder-specific gene expression signatures. In addition, we generate polygenic risk scores to assess overlap between disorders on a genome-wide genetic-risk level. We next investigate whether differences in polygenic risk scores are reflected in disorder-related gene expression profiles and whether they can be used to tease apart genetic and environmental influences on gene expression.

Conclusions
Our results reveal disorder overlap and specificity at the genetic and gene expression level. They suggest new pathways contributing to distinct pathophysiology in psychiatric disorders and shed light on potential shared genomic risk factors.

Declaration of interest
G.B. acts a consultant in preclinical genomics for Eli Lilly.

Copyright and usage
© The Royal College of Psychiatrists 2016. This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY) licence.

Psychiatric disorders are leading causes of disability globally, surpassing that attributed to more physical disorders. Worldwide prevalence rates for the five major psychiatric disorders range from 0.2% for attention-deficit hyperactivity disorder (ADHD) to 5.7% for major depressive disorder (MDD), with schizophrenia, autism spectrum disorder (ASD) and bipolar disorder showing intermediate prevalence.1 Although these disorders are classified into distinct disorder categories, they show overlap in symptomatology and shared genetic risk in family studies, as well as comorbidity in individuals and families.2 For example, 41% of patients with ASD were found to have a comorbid disorder, including ADHD.3 In addition, family studies show that relatives of probands with bipolar disorder show an increased risk for both bipolar disorder and schizophrenia.4 This could indicate common aetiological pathways for some psychiatric disorders and studying them jointly could lead to new insights.

Twin study heritability estimates for psychiatric disorders are moderate for MDD (40%) and high (80–90%) for ADHD, ASD and schizophrenia, indicating a significant genetic component in their aetiology.5–7 Large genome-wide studies have now started to reveal risk variants for individual disorders but an outstanding question is whether genetic risk factors or polygenic risk scores are specific to particular disorders, or shared. For instance, certain copy number variations (CNVs) are overrepresented in both patients with schizophrenia and those with ASD.8,9 Additionally, calcium-channel activity genes have been shown to have shared involvement in ASD, schizophrenia and bipolar disorder, whereas polygenic risk analysis showed strong overlap of genome-wide risk between bipolar disorder, MDD and schizophrenia, and moderate overlap between ASD and schizophrenia.10 In another sample, high sharing was seen between bipolar disorder and schizophrenia, and also highlight two immune-related gene co-expression modules correlating inversely with MDD and adult ADHD disease status. We find no significant relationship between polygenic risk scores and gene expression signatures.

Method

Participants
Participants were from four different projects and included individuals affected with adult ADHD, childhood ADHD, MDD, ASD, childhood dual diagnosis ADHD–ASD and healthy controls (Table 1). Because of the age difference and its possible confounding effect on gene expression within the ADHD diagnostic group we decided to split the ADHD samples into a childhood (mean age 10 years) and adult (mean age 32 years) ADHD groups. Participant characteristics and inclusion/exclusion

<table>
<thead>
<tr>
<th>Aims</th>
<th>Method</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>To systematically investigate genomics overlap between childhood and adult attention-deficit hyperactivity disorder (ADHD), autism spectrum disorder (ASD) and major depressive disorder (MDD).</td>
<td>Analysis of whole-genome blood gene expression and genetic risk scores of 318 individuals. Participants included individuals affected with adult ADHD (n = 93), childhood ADHD (n = 17), MDD (n = 63), ASD (n = 51), childhood dual diagnosis of ADHD–ASD (n = 16) and healthy controls (n = 78).</td>
<td>Weighted gene co-expression analysis results reveal disorder-specific signatures for childhood ADHD and MDD, with moderate overlap between MDD and schizophrenia, bipolar disorder and ADHD and low, but significant, overlap between schizophrenia and ASD. These results suggest that some shared molecular genetic factors underlie a significant proportion of the risk for development of several psychiatric disorders. To date, however, it is not known whether these genomic overlaps are reflected in overlapping patterns of gene expression in tissues. In this study we therefore explored whole blood gene expression across adult and childhood ADHD, ASD, MDD and healthy controls, using weighted gene co-expression analysis to search for patterns of correlated gene expression between disorders as well as disorder-specific gene expression signatures. In addition, we generate polygenic risk scores to assess overlap between disorders on a genome-wide genetic-risk level. We next investigate whether differences in polygenic risk scores are reflected in disorder-related gene expression profiles and whether they can be used to tease apart genetic and environmental influences on gene expression.</td>
</tr>
</tbody>
</table>
Gene expression data preprocessing

Whole-genome gene expression data of a total of 424 individuals were generated using the Illumina HT-12.v4 BeadChips at the SGDP/BRC BioResource Illumina core lab according to the manufacturer’s protocol. We rigorously quality controlled and preprocessed the data using a standard pipeline (https://github.com/snewhouse/BRC_MH_Bioinformatics), excluding sample and probe outliers, after which robust spline normalisation and log2 transformation were applied. After excluding samples based on low-quality expression profiles we included only participants with full phenotype data (age, gender, diagnosis, date of collection, ethnicity, psychoactive medication use per individual, and RNA integrity number (RIN) and RNA concentration per collection, ethnicity, psychoactive medication use per individual, with full phenotype data (age, gender, diagnosis, date of collection, ethnicity, psychoactive medication use per individual, and RNA integrity number (RIN) and RNA concentration per collection). This left 318 participants and 5638 probes for analysis (Table 1). To minimise project collection and sample handling batch effects, the data were corrected for three unknown variables (SV) using the SVA R package13 using a linear model. Surrogate batch effects, the data were corrected for three unknown variables (SV)1 was not correlated to known covariates, but with all projects except DeCC, comorbidity of other psychiatric disorders (another psychiatric diagnosis was an exclusion criterion in the DeCC study). Each project has ethical approval and full informed consent for each participant (details in online supplement DS1). Whole blood samples were collected using PAXGene tubes for RNA and EDTA for DNA. All RNA samples were processed within one batch to generate whole-genome gene expression data.

Weighted gene co-expression network reconstruction

Corrected gene expression data were analysed using weighted gene co-expression analysis (WGCNA).15,16 We constructed a signed weighted co-expression network based on the matrix of pairwise Pearson correlation coefficients, which were raised to a fixed power ($\beta = 12$) by the criteria described by Zhang & Horvath.16 Soft-thresholding results in a 5638 × 5638 dimensional weighted adjacency matrix containing pairwise connection strengths. Subsequently, a topological overlap measure is calculated based on the number of shared neighbours. Modules were then defined as branches of a hierarchical clustering tree using a dissimilarity measure (1 – topological overlap). Each module is subsequently assigned a colour. To define a representative module expression profile for each module, we summarised the (standardised) gene expression profiles of the module by their first principal component. This statistic is referred to as the module eigengene: it can be thought of as an average gene expression value for all probes in a module per sample. In order to identify hub genes, we calculated a connectivity measure (‘module membership’) per probe by correlating expression values with the relevant module eigengene. Probes with high module membership are defined as ‘hub genes’ of their module.

Module eigengenes and phenotypes of interest

The module eigengenes were correlated to phenotype of interest using a linear model.17 Participants were assigned a main diagnosis of a particular disorder (controls, ASD, MDD, adult ADHD, childhood ADHD, ADHD–ASD) + all other participants. We investigated the effect of possible covariates on the association of module eigengenes with our phenotypes of interest. This led us to correct for gender, RIN and RNA concentration. Medication use defined as dichotomous measures of use of antidepressants ($n = 73$), stimulants ($n = 49$), antipsychotics ($n = 7$), benzodiazepines ($n = 5$) and mood stabilisers ($n = 5$) were investigated as a possible covariate. Considering the nature of the sample collection of the childhood ADHD–ASD and childhood ADHD samples, analyses were repeated also including age as covariate, allowing us to tease out diagnosis and age-specific effects. We used a Bonferroni threshold for significance (the number of tests was calculated as the number of modules × number of diagnoses × 2).

Characterisation of modules of interest

Modules of interest were tested for enrichment of blood cell type lists using the userListEnrichment function in WGCNA package with five as minimum number of genes in a pathway. This function compares the number of overlapping genes to the maximal possible overlap and applies Bonferroni correction. Enrichment analyses were performed in WebGestalt using Wikipathways database.18 In addition, we performed connectivity mapping on selected modules to investigate overlap with drug-induced gene expression changes through the LINCS/CMap database (http://apps.lincscloud.org).19 We entered the gene content of modules of interest as upregulated, resulting in a list of compounds of which the application to cell lines results in a similar gene expression pattern.

<table>
<thead>
<tr>
<th>Table 1 Sample characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diagnosis</td>
</tr>
<tr>
<td>Controls</td>
</tr>
<tr>
<td>Autism spectrum disorder</td>
</tr>
<tr>
<td>Major depressive disorder</td>
</tr>
<tr>
<td>Adult attention-deficit hyperactivity disorder</td>
</tr>
<tr>
<td>Childhood attention-deficit hyperactivity disorder</td>
</tr>
<tr>
<td>Autism spectrum disorder–attention-deficit hyperactivity disorder</td>
</tr>
</tbody>
</table>

RIN, RNA integrity number.
Genotype information and previous genome-wide association study (GWAS) findings

For most individuals genotype data were available and after quality control and imputation polygenic risk score analysis and genetic pathway analysis was carried out to investigate the relationship of genome-wide risk to disorder-specific gene expression findings in this study (see online supplement, Supplements DS2–4 and Figs DS1 and DS2).

Results

The WGCNA on 318 participants and 5638 probes resulted in seven gene co-expression modules, ranging from 2077 probes in the turquoise module to 80 probes in the red module. The grey module contains 675 probes not belonging to any other module, representing background noise. The network dendrogram is given in online Fig. DS3 and all probes and corresponding module assignments in online Table DS1. The module eigengenes representing a summary of all genes in a given module were related to our traits of interest: participants with a main diagnosis of a particular disorder (controls, ASD, MDD, adult ADHD, childhood ADHD, ADHD–ASD) vs. all other participants, covarying for gender, RIN, concentration in one model and also including age in a second model. This results in 7 phenotypes × 6 module eigengenes = 42 tests per model, therefore 84 tests in total when considering both models. We applied a Bonferroni threshold for significance of \( P = 0.05/84 = 6 \times 10^{-4} \). Results for the first model are listed in Table 2. We did not find significant gene expression effects for the ASD and ADHD–ASD groups. In addition, medication use (antidepressants, stimulants, antipsychotics, mood stabilisers or benzodiazepines) was not significantly associated to any gene co-expression module (online Fig. DS4).

The green and red immune modules are inversely correlated to psychiatric disorders

The green (186 probes) and red module (80 probes) eigenvalues represent per individual correlated negatively with MDD and positively with adult ADHD status (Table 2, Fig. 1). Even though the MDD sample is female-dominated and the adult ADHD sample male-dominated, gender was included as a covariate and therefore does not fully account for this effect. Results also remain significant after additional correction for age. Of the 93 individuals with ADHD, 7 had comorbid MDD. However, these individuals still resembled the pattern of the adult ADHD group more than the MDD group (online Fig. DS5). Wikipathways enrichment analyses through Webgestalt reveal significant enrichment for regulation of toll-like receptor signalling pathway (eight genes, adjusted \( P = 0.02 \)) in the green module. The red module, however, is enriched for immune-related pathways: type II interferon signalling (nine genes, adjusted \( P = 1.8 \times 10^{-6} \)), type III interferon signalling (three genes, adjusted \( P = 4 \times 10^{-3} \)).

Connectivity mapping (http://apps.lincscloud.org) \( ^{19} \) revealed 200 compounds corresponding to upregulation of genes in the green module, and 3273 to upregulation in the red module. This implies that application of these compounds to cell lines results in a similar gene expression profile as we find for adult ADHD, and opposite to that for MDD. Among the compounds, there were several tricyclic (amitriptyline, desipramine, nortriptyline, protriptyline, trimipramine) and other ( nefazodone, trazodone) antidepressants. In addition, there are steroid-related drugs (among others: betamethasone, cucurbitacin-I, sarmentogenin, wortmannin, alcomethasone, alfaldolone, altrenogest, androstanol, beclamethasone, cortisol, corticosterone) and other inflammatory drugs (celecoxib, diclofenac, propraglandin). Full connectivity mapping results are given in online Table DS2.

Figure 1 depicts gene significance (correlation between phenotype and module eigengene) for MDD and adult ADHD of the red and green module content. Size and colour indicate module membership, our definition of connectivity. Gene significance is highly correlated with gene connectivity in the green (MDD: \( r = -0.45, P = 1.7 \times 10^{-3} \), adult ADHD: \( r = 0.36, P = 6.2 \times 10^{-7} \)) and red (MDD: \( r = -0.43, P = 5.9 \times 10^{-5} \), adult ADHD: \( r = 0.45, P = 2.9 \times 10^{-5} \)) modules, indicating that highly connected hub genes generally show the strongest associations with the phenotypes. The six main green hub genes are YTHD1, AK1, PAK2, C20orf94 (now SIX4P), PAPD5 and SPATA13. The expression of transcription factor YTH1 indeed shows one of the highest negative correlations with MDD status, together with TMEM189–UBE2V1, PSMD1, WDR62 and LOC730052 (now UBE2V1P2). MEF2R, TPM3, TBP1, 2NF364 (now RNF115) and RAD21 show highest positive correlation with adult ADHD. For the red module, the hub genes are the interferon genes IFIT3, IFIT5 and XAF1, PARP14 and two probes for UBE2L6. Of these, PARP14 is highly inversely correlated to MDD and adult ADHD, as are MX1 and EIF2UK2. In addition, two probes for STAT1 show high gene significance for MDD and OA2 for adult ADHD.

Turquoise module represents overlapping signatures between psychiatric disorders

The turquoise module (2077 probes) is positively correlated with both MDD and childhood ADHD status (Table 2). This effect does not remain significant for childhood ADHD after correction for age. Enrichment analyses reveal significant categories to be cytoplasmic ribosomal proteins (69 genes, adjusted \( P = 2.2 \times 10^{-6} \)), electron transport chain (45 genes, adjusted \( P = 10 \times 1.6^{-6} \)) and oxidative phosphorylation (25 genes, adjusted \( P = 1 \times 10^{-2} \)). The correlation between gene significance and

<table>
<thead>
<tr>
<th>Module</th>
<th>Probes, n</th>
<th>Controls</th>
<th>Adult attention-deficit hyperactivity disorder</th>
<th>Childhood attention-deficit hyperactivity disorder</th>
<th>Autism spectrum disorder–attention-deficit hyperactivity disorder</th>
<th>Autism spectrum disorder</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turquoise</td>
<td>2077</td>
<td>ns</td>
<td>( 4.4 \times 10^{-3} )</td>
<td>ns</td>
<td>( 4.0 \times 10^{-5} )</td>
<td>ns</td>
</tr>
<tr>
<td>Red</td>
<td>80</td>
<td>ns</td>
<td>( -4.3 \times 10^{-7} )</td>
<td>( 4.3 \times 10^{-7} )</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>Green</td>
<td>186</td>
<td>ns</td>
<td>( -5.8 \times 10^{-7} )</td>
<td>( 4.8 \times 10^{-7} )</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>Blue</td>
<td>1672</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
<td>( -4.0 \times 10^{-7} ) c</td>
<td>ns</td>
</tr>
</tbody>
</table>

ns, not significant.

a. Module eigengenes were tested for association with phenotypes of interest including gender, RNA integrity Number RIN and RNA concentration as covariates.
b. Positive associations.
c. Negative associations.
connectivity is $r = -0.48$, $P < 2.2 \times 10^{-16}$ for childhood ADHD and $r = -0.48$, $P < 2.2 \times 10^{-16}$ for MDD. Major hub genes in this module are ZRANB2, TMEM126B, RPL15, PCNP and SRP9. The genes with highest MDD gene significance are LOC1001291, ANXA1, RNF13, GMFG and H3F3B. For childhood ADHD these are SUMO3, RPL22, UTP3, POLE3 and CCDC50.

### Disorder-specific signature of the blue module

The blue module (1672) is negatively correlated to childhood ADHD status (Table 2). This effect does not remain significant after correction for age. Enrichment analyses reveal the following significant Wikipathways: integrin-mediated cell adhesion (26 genes, adjusted $P = 8 \times 10^{-4}$), focal adhesion (35 genes, adjusted $P = 1 \times 10^{-3}$), prostate cancer (21 genes, adjusted $P = 6 \times 10^{-3}$) and IL17 signalling pathway (13 genes, adjusted $P = 0.04$). The correlation between childhood ADHD gene significance and connectivity is $r = -0.5$, $P < 2.2 \times 10^{-16}$. Blue hub genes are WAS, MOBK1L2A, C15orf39, G6PD and GNAI2. Genes with highest negative correlation to childhood ADHD disease status are TMUB2, STAT5B, DENND3, CA4, FCGR3B.

### Relationship gene expression signatures and previous GWAS findings

We did not find enrichment of GWAS signal or association of the module eigengenes with polygenic risk scores for the modules of interest (online supplements DS2–4 and Figs DS1 and DS2).

### Discussion

#### Main findings

This study of whole-blood gene expression of several psychiatric disorders aimed to reveal overlapping gene expression patterns and disorder-specific signatures. The WGCNA on 318 participants revealed seven gene co-expression modules. Of these, two small modules are inversely related to MDD and adult ADHD, the large turquoise module is associated with both MDD and childhood ADHD and finally, the blue module shows a disorder-specific signature for childhood ADHD. No significant results were found for the ASD, ADHD–ASD and control groups. Even though there is some evidence for increased immune-related comorbidities in ASD20,21 we do not find an immune gene expression signature for these groups, as we do for MDD and adult ADHD.

#### Interpretation and comparison with findings from other studies

Two small modules, red and green, are negatively correlated to MDD status, but positively correlated to adult ADHD status. GWAS on both disorders have not yielded genome-wide significant hits to date22,23. Previous literature about genetic overlap between MDD and ADHD has yielded conflicting results, reporting a genetic correlation between MDD and ADHD,11 but no significant overlap in polygenic risk scores.10 A reason for these
discrepancies could be the lack of distinction between adult ADHD and childhood ADHD in previous studies, which could add to phenotypic heterogeneity. The group with childhood ADHD may include individuals whose condition will become less severe with age, whereas the adult group with ADHD contains individuals whose condition is chronic. Another reason why we chose to analyse participants with childhood and adult ADHD in separate groups is because of anticipated biological heterogeneity between children and adults in the context of peripheral gene expression. An explanation for our findings would be gender differences in prevalence of MDD and adult ADHD, however, analyses were corrected for gender and stratified analyses yielded the same results (data not shown). We also performed a correlation with the module eigengenes and an indicator of current state of depression (Beck Depression Inventory) within participants, which did not show a significant association (data not shown).

The green module shows enrichment for a cell signalling category and harbours some interesting highly connected hub genes that also show high gene significance, most notably, YY1, WDR82 and AAK1. YY1 is a transcription factor involved in many processes including transforming growth factor beta (TGFB) signalling, but has also shown to be active in histone modification. In addition WDR82 (WD repeat domain 82) is part of SET1A/SET1B histone H3K4 methyltransferase complexes, which genetic pathway analyses has shown to be strongly enriched for association in MDD, bipolar disorder and schizophrenia. This implication of epigenetic processes provides a mechanism by which environmental influences can exert their influence on gene expression and thereby contribute to the pathology of neuropsychiatric disorders. The green hub gene adaptor-associated kinase 1 (AAK1) is of interest because it is a positive regulator of the Notch pathway. This pathway is traditionally implicated in cell-fate determination during development, but has important function in tissue homeostasis and neuronal plasticity later in life. In addition, it has been implicated to play a role in immune functioning.

Even though the immune system has not been implicated in adult ADHD before, there are many studies on its association with MDD although the relationship is still somewhat controversial. Expression differences in cytokines have been shown to differentiate patients with MDD from ones with bipolar disorder and controls and studies have shown regulation of the serotonin receptor through cytokines and neurotrophins. Importantly, cytokines and polymorphisms in interleukin genes have been shown to predict antidepressant treatment response. The majority of patients with MDD (70%) and a few (13%) of the participants with adult ADHD in the current study were on antidepressant medication although medication use did not correlate significantly with any gene expression module. However, connectivity mapping revealed that the upregulation of genes in red and green modules as seen in our adult ADHD group coincides with those seen in response to application of a number of tricyclic antidepressants, indicated for the treatment of depression and ADHD with comorbid depression. Results also contained some anti-inflammatory drugs, converging with the module enrichments for immune system genes. This could support the hypothesis that anti-inflammatory drugs such as non-steroidal anti-inflammatory drugs (NSAIDs) might have a role in the treatment of MDD and our data suggests, also ADHD. Literature on the effects of NSAIDs in patients with MDD is, however, mixed. Our results suggest potential drug repositioning opportunities for NSAIDs for both MDD and ADHD.

Whereas MDD and adult ADHD showed opposite patterns of gene expression changes (the green and red modules), we found an overlapping gene expression signature between MDD and childhood ADHD, in the form of the turquoise module. One possible interpretation of this finding is that childhood ADHD increases risk for developing MDD in later life. In order to fully untangle this, we would need childhood diagnoses of ADHD in the MDD cohort, but this was not available.

Finally, we find a disorder-specific signature of childhood ADHD in the blue module, which is highly enriched for cell signalling genes. There are two hub genes located on the X-chromosome; G6PD and WAS. Even though we corrected for gender in our analyses, the childhood ADHD group consisted solely of males, which might explain downregulation of X-chromosome genes, although the blue module is not enriched for genes on sex chromosomes. The significance of the association of childhood ADHD with the turquoise and blue modules disappears when adding age as a covariate in the analyses, but in the absence of a healthy child control group, age and childhood ADHD cohort membership are highly correlated. However, the fact that the childhood dual diagnosis ADHD–ASD group does not show a correlation with these modules suggests that age cannot fully explain these results.

We did not find enrichment of GWAS signal or association of the module eigengenes with polygenic risk scores for the modules of interest (online supplements DS2–4 and Figs DS1 and DS2). This could indicate that the differences in gene expression are driven by environmental rather than genetic factors or, perhaps, that polygenic scores are not yet strong enough in disorders such as MDD and ADHD. One possible known environmental influence on gene expression is smoking. Even though we did not have access to smoking behaviour for all participants, a Fisher’s exact test of enrichment of smoking-related genes did not reveal a significant enrichment for the modules of interest (data not shown). Likewise, a sample-handling or collection effect is unlikely because of initial batch correction and the fact that the healthy controls were for the most part from the same project as the participants with MDD and yet do not show an effect for the relevant modules. In addition, the lack of genetic association could also be the result of the initial GWAS results being underpowered to detect variants associated with MDD, ADHD and ASD.

Directions for further study

Future research could extend this study to include a broader range of psychiatric disorders, such as psychotic disorders (schizophrenia) and anxiety disorders (obsessive–compulsive disorder, generalised anxiety disorder), in order to better understand the genomic correlates of different syndromes. However, our results have several limitations including power to detect effects in relatively small samples and the reliance upon cross-sectional study designs. Our findings in adult ADHD and MDD will require replication and assessment in different study designs to assess potential therapeutic applications. Also, in this study we examined gene expression in blood but it will be important to determine whether the pattern of results holds true for brain tissue. In ASD, for instance, changes in the expression of a number of genes has been reported to be altered in post-mortem studies.

In conclusion, in a study of gene expression in peripheral blood of patients with psychiatric disorders and healthy controls, we identified both cross-disorder and disorder-specific signatures for adult ADHD and MDD. With the caveats discussed above, they suggest new pathways contributing to distinct pathophysiology in psychiatric disorders and shed light on potential shared genomic risk factors.


Supplement DS1

Description of different projects

DECC: Depression Case Control project

From this project we included gene expression data of 63 MDD cases and 57 controls. The Depression Case-Control (DeCC) study is a case-control study that recruited unrelated patients from three clinical UK sites: London, Cardiff and Birmingham [1]. Subjects were identified from psychiatric clinics, hospitals and general medical practices and from volunteers responding to media advertisements. All participants experienced two or more episodes of major depression of at least moderate severity. The diagnosis of MDD was ascertained using the Schedules for Clinical Assessment in Neuropsychiatry (SCAN) [2] interview. Subjects were excluded if they or a first-degree relative ever fulfilled criteria for mania, hypomania or schizophrenia. The controls were screened for lifetime absence of any psychiatric disorder using a modified version of the Past History Schedule [3]. Participants were excluded if they, or a first-degree relative, ever fulfilled the criteria for depression or any other psychiatric disorder. Approval was obtained from the local research ethics committees/institutional research boards of all of the participating sites in U.K.: London, Cardiff and Birmingham and all individuals gave written informed consent [4].

ADHD: Attention-Deficit Hyperactivity Disorder project

From this project we included gene expression data of 93 aADHD cases.

From this project we included gene expression data of 93 ADHD cases comprising adults attending a National Adult ADHD Outpatient Clinic. Self-report and informant based versions of the Barkley Adult ADHD Rating Scale (BAARS-IV) [5] were used to measure ADHD symptoms. Consisting of 18 DSM-IV items related to inattention and hyperactivity–impulsivity, respondents indicated how frequently they experienced behaviours on a scale of 0 to 3 (never or rarely, sometimes, often, very often) during the past 6 months. Total scores were calculated for each symptom dimension.
Informant ratings were provided by a family member or close friend. Diagnosis was based on psychiatric interview and the Conners Adult ADHD Diagnostic Interview for DSM-IV (CAADID) completed by a psychiatrist and a community psychiatric nurse. The CAADID is a structured interview divided into Part I (Patient History Questionnaire) and Part II (Diagnostic Criteria Interview), which are administered separately. Each of the 18 items is scored “yes”, if the behavioural symptom is present often within the past 6 months and outcomes are total current ADHD symptom score, and separate totals for inattentive and hyperactive-impulsive symptom domains. Participants attending assessment appointments were consented by a member of the BRC Bioresource team following the receipt of an information sheet and a detailed explanation of the BRC Bioresource initiative. Ethical approval was granted by the National Research Ethics Committee, London (12/LO/07990).

BioNed: Biomarkers for Childhood Neuropsychiatric Disorders project

*From this project we included gene expression data of 15 ASD_ADHD cases, 7 ASD cases, 17 cADHD cases and 7 controls.*

All participants were male with a clinical diagnosis of autism, atypical autism, Asperger’s syndrome and/or a diagnosis of combined type ADHD (or hyperkinetic disorder) made according to ICD-10 or DSM-IV (American Psychiatric Association, 2000; World Health Organization, 2005). Exclusion criteria included any known comorbid medical condition such as Major Mood Disorder (including bipolar I and II); severe OCD, severe Generalised Anxiety Disorder, Conduct Disorder and genetic disorders (e.g. Fragile X Syndrome). Participants were also excluded if they had a diagnosis of epilepsy or had received a head injury/neurological insult that may affect cognitive functioning. Individuals who were taking psychotropic medication (except for stimulants in ADHD) and who had a Full-Scale IQ (FSIQ) less than 70 measured using the Wechsler Abbreviated Scale of Intelligence (WASI, Wechsler, 1999) were not included. Participants taking stimulant medication were included, but a 24-48 hour washout period was mandatory prior to testing. Upon recruitment, participants were assigned to one of three research groups (ASD, ADHD, ASD+ADHD) using a multi-source, multi-measure approach, taking into consideration clinical status of the patient as well as additional standardised psychological measures. The
Social Communication (SCQ; Rutter et al. 2003), the Autism Diagnostic Interview-Revised (ADI-R; Lord et al. 1994), and the Autism Diagnostic Observation Schedule-Generic (ADOS-G; Lord et al. 2000) were used for the assessment of ASD cases. The Conners 3rd edition parent short form (Conners 3-PS; Conners 2008), the Parent Account of Childhood Symptoms (PACS; Taylor et al. 1991; Chen and Taylor 2006) were used to confirm ADHD cases. Comorbid ASD+ADHD cases were required to meet full research diagnostic criteria for both ASD and ADHD. The study protocol was approved by a medical ethics committee (NHS REC Ref: 08/H0803/161). Parental written consent was given before the experiment began.

AIMS: Autism Interventions

From this project we included gene expression data of 44 ASD cases and 14 controls. All participants were right handed (measured using The Edinburgh Handedness inventory [6] and native English speakers. Exclusion criteria included; pre-existing medical conditions or complications (e.g. head trauma, epilepsy); use of medication affecting brain function; mental retardation; a history of major psychiatric disorder (e.g. psychosis); chromosomal abnormality (e.g. fragile X, Tuberous Sclerosis, VCFS); and any MRI contraindications. Intellectual ability was assessed using the WASI [7]. All participants had an IQ greater than 70 (i.e. were within the high-functioning range of the autistic spectrum). For the autistic group, inclusion was based on a clinical diagnosis of autism using the International Statistical Classification of Diseases, 10th Revision (ICD-10) research criteria and confirmed using the ADI-R [8] (all cases reached ADI-R algorithm cut-offs in the domains of impaired reciprocal social interaction, communication, and repetitive behaviors and stereotyped patterns, although failure to reach cutoff in a single domain by 1 point was permitted). Current symptoms were assessed using the ADOS [9], but not used as an inclusion criterion. The study was given ethical approval by the National Research Ethics Committee, Suffolk, UK. All volunteers gave written informed consent.
Supplement DS2

Genotype data preprocessing

The majority of subjects with gene expression (n=252, 80%) were also genotyped as part of their respective projects. The MDD project subjects (n=57 cases, n=54 controls) were genotyped on the Illumina 610k BeadChip, the ASD (n=34 cases, 14 controls) and ADHD_ASD (n=12 ADHD_ASD cases, n=7 ASD cases, n=1 cADHD case) on the Illumina HumanCoreExome BeadChip and ADHD (n=73 aADHD cases) on the Illumina OmniExpress BeadChip. All data were quality controlled separately in Plink v1.07 [10] using the same parameters (SNPs were excluded when missingness >1%, MAF <0.01 or HWE <0.00001. Individuals were excluded when missingness >1%). Sex and relatedness checks were carried out, in addition to Eigensoft analyses to confirm self-reported ethnicities [11]. When necessary SNP positions were lifted over from hg18 to hg19 build using UCSC LiftOver tool [12]. To eliminate between chip genotype coverage differences, we imputed all datasets to the 1000Genomes, Phase1.v3 (SHAPEIT, no singletons) using the Michigan Imputation Server (https://imputationserver.sph.umich.edu), utilizing SHAPEIT and Minimac software. Imputed data was merged and filtered for quality using R^2>0.3 and MAF>0.01.

Supplement DS3

FORGE genetic pathway analysis

We examined significance of gene expression modules as pathways in published GWAS. We used FORGE to combine p-values per gene and subsequently pathways [13]. To assess LD structure we used the 1000 Genomes data after liftover to hg19 and pruned with a R^2=0.9 threshold using Priority Pruner, prioritizing low p-value SNPs within the PGC cross disorder results [14]. For the Forge.pl run a maximum of 100,000 permutations was set and the algorithm was run with a fuzzy border option (5’ 35kb, 3’ 10kb). Subsequent gsa.pl runs used the Z statistic (fixed after permutations) for each gene, or the raw SNP p-value case of only one SNP per gene. FORGE genetic pathway analysis did not reveal enrichment of MDD, ADHD or ASD GWAS signal for any of the modules.
Supplement DS4

Polygenic Risk Scores and gene expression

Genome-wide (excluding MHC region) Polygenic Risk Scores for ADHD, ASD and MDD were generated with PRSice software [15] using PGC cross-disorder $p$-values as training sets [14] with the exception of the MDD summary statistics, which we have used the leave-one-out scores excluding RADIANT (of which DECC subjects are part). Polygenic risk scores (PRS) for ADHD, ASD and MDD were calculated for a subset (n=252) of samples (Supplementary X and XI below) and we applied a t-test between each phenotype and all other subjects, except cADHD because of lack of samples. The significance threshold was set at $p<0.05/25=0.002$. The PRS did not differ significantly between groups for any of the disorders. None of the PRS were not significantly associated to any of the module eigengenes, and did not change original results when taken along as covariates.
Fig. DS1

Polygenic Risk Score distributions
Polygenic Risk Scores in different disorder groups

Polygenic Risk Scores for ADHD (panel A), ASD (Panel B) and MDD (Panel C) were calculated for a subset of 252 samples using a p<0.1 cutoff.

Fig. DS2

PRS ADHD

PRS ASD

PRS MDD
**Fig. DS3**

**WGCNA dendrogram.** Network reconstruction identifies 7 distinct modules of co-expressed genes in whole blood of 318 cross disorder subjects. The dendrogram was produced by average linkage hierarchical clustering of genes using topological overlap. Modules of co-expressed genes were assigned numbers corresponding to the branches indicated by the horizontal bar beneath the dendrogram.
Fig. DS4 Module Trait Correlations

<table>
<thead>
<tr>
<th>Module Eigengenes vs Traits</th>
<th>Set</th>
<th>Spill</th>
<th>RHs</th>
<th>RNA Cons</th>
<th>MOD</th>
<th>ASD</th>
<th>aADHD</th>
<th>cADHD</th>
<th>ASD, aADHD</th>
<th>CON</th>
<th>ADD, aADHD</th>
<th>ODD, ADD, aADHD</th>
<th>Carrotpte</th>
<th>AttnDeficit</th>
<th>SystDysfunct</th>
<th>AnxDepress</th>
<th>BzDep</th>
<th>MDD, anxious</th>
<th>MDD, manic</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEturquoise</td>
<td>-0.109 (0.09)</td>
<td>-0.035 (0.05)</td>
<td>0.26 (2e-05)</td>
<td>-0.24 (1e-05)</td>
<td>0.23 (2e-05)</td>
<td>-0.016 (3.1)</td>
<td>-0.2 (2e-04)</td>
<td>0.24 (1e-05)</td>
<td>0.1 (1e-06)</td>
<td>-0.17 (2e-04)</td>
<td>0.098 (1e-08)</td>
<td>0.12 (1e-03)</td>
<td>0.00083 (1)</td>
<td>-0.029 (0.5)</td>
<td>0.006 (0.3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MEred</td>
<td>-0.2 (4e-04)</td>
<td>0.034 (0.5)</td>
<td>0.013 (0.3)</td>
<td>-0.028 (8e-06)</td>
<td>-0.14 (3e-01)</td>
<td>0.012 (3.9)</td>
<td>0.2 (8e-04)</td>
<td>-0.185 (1e-01)</td>
<td>0.040 (1e-04)</td>
<td>0.025 (6e-07)</td>
<td>0.009 (9e-06)</td>
<td>0.059 (1e-01)</td>
<td>0.003 (0.3)</td>
<td>-0.083 (0.1)</td>
<td>-0.0355 (0.9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MEgreen</td>
<td>0.24 (7e-06)</td>
<td>-0.31 (2e-06)</td>
<td>0.014 (0.4)</td>
<td>0.083 (1e-01)</td>
<td>-0.37 (4e-12)</td>
<td>0.12 (3e-03)</td>
<td>0.29 (1e-07)</td>
<td>0.0023 (1)</td>
<td>0.061 (1e-04)</td>
<td>0.006 (1e-06)</td>
<td>-0.14 (4e-01)</td>
<td>0.007 (1e-06)</td>
<td>-0.016 (0.4)</td>
<td>0.014 (0.8)</td>
<td>0.029 (0.6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MEyellow</td>
<td>-0.04 (5e-05)</td>
<td>-0.035 (0.5)</td>
<td>-0.24 (1e-05)</td>
<td>0.17 (1002)</td>
<td>-0.087 (1)</td>
<td>0.14 (0.41)</td>
<td>0.022 (0.7)</td>
<td>-0.361 (0.3)</td>
<td>0.065 (0.3)</td>
<td>0.18 (0.002)</td>
<td>-0.003 (0.1)</td>
<td>0.0003 (0.1)</td>
<td>-0.012 (0.4)</td>
<td>0.046 (0.6)</td>
<td>-0.0205 (0.5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MEblue</td>
<td>0.033 (0.6)</td>
<td>0.096 (0.006)</td>
<td>-0.027 (2.2)</td>
<td>0.18 (0.001)</td>
<td>-0.15 (0.007)</td>
<td>0.038 (0.5)</td>
<td>0.16 (0.005)</td>
<td>-0.23 (4e-05)</td>
<td>-0.11 (4e-04)</td>
<td>-0.12 (6e-03)</td>
<td>-0.037 (2e-02)</td>
<td>0.013 (0.8)</td>
<td>0.014 (0.6)</td>
<td>0.073 (0.2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MEBrown</td>
<td>0.18 (0.006)</td>
<td>0.032 (0.1)</td>
<td>-0.18 (0.001)</td>
<td>0.26 (3e-06)</td>
<td>-0.19 (4e-04)</td>
<td>0.03 (4.4)</td>
<td>0.21 (2e-04)</td>
<td>-0.19 (7e-06)</td>
<td>-0.31 (8e-06)</td>
<td>0.083 (0.1)</td>
<td>-0.092 (2e-01)</td>
<td>-0.1 (1.0e-07)</td>
<td>0.095 (0.2)</td>
<td>0.066 (0.1)</td>
<td>0.062 (0.3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MEGrey</td>
<td>-0.13 (0.001)</td>
<td>-0.079 (0.2)</td>
<td>0.047 (0.4)</td>
<td>-0.043 (0.1)</td>
<td>-0.033 (0.7)</td>
<td>-0.068 (1)</td>
<td>0.079 (0.3)</td>
<td>-0.13 (0.02)</td>
<td>0.029 (0.1)</td>
<td>0.055 (0.3)</td>
<td>-0.12 (2e-03)</td>
<td>0.054 (0.1)</td>
<td>-0.088 (0.1)</td>
<td>-0.14 (0.01)</td>
<td>0.062 (0.1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Fig. DSS Supplementary Information IV: Green and Red ME expression for aADHD, MDD and comorbid subjects

![Graph showing ME expression for different conditions](image-url)
Separate tables
Table DS1: Gene list & module assignments (.xlsx).
Table DS2: Connectivity mapping results for Green and Red modules via LINCS/CMap database (http://apps.lincscloud.org) (.xlsx).

SUPPLEMENTARY REFERENCES


Immune signatures and disorder-specific patterns in a cross-disorder gene expression analysis
BJP published online May 5, 2016 Access the most recent version at DOI: 10.1192/bjp.bp.115.175471