Data supplement to de Jong et al. Cross-disorder gene expression analysis: immune signatures and disorder-specific patterns. Br J Psychiatry doi: 10.1192/bjp.bp.115.175471

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).

<u>BioNed: Biomarkers for Childhood Neuropsychiatric Disorders project</u>

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 a 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

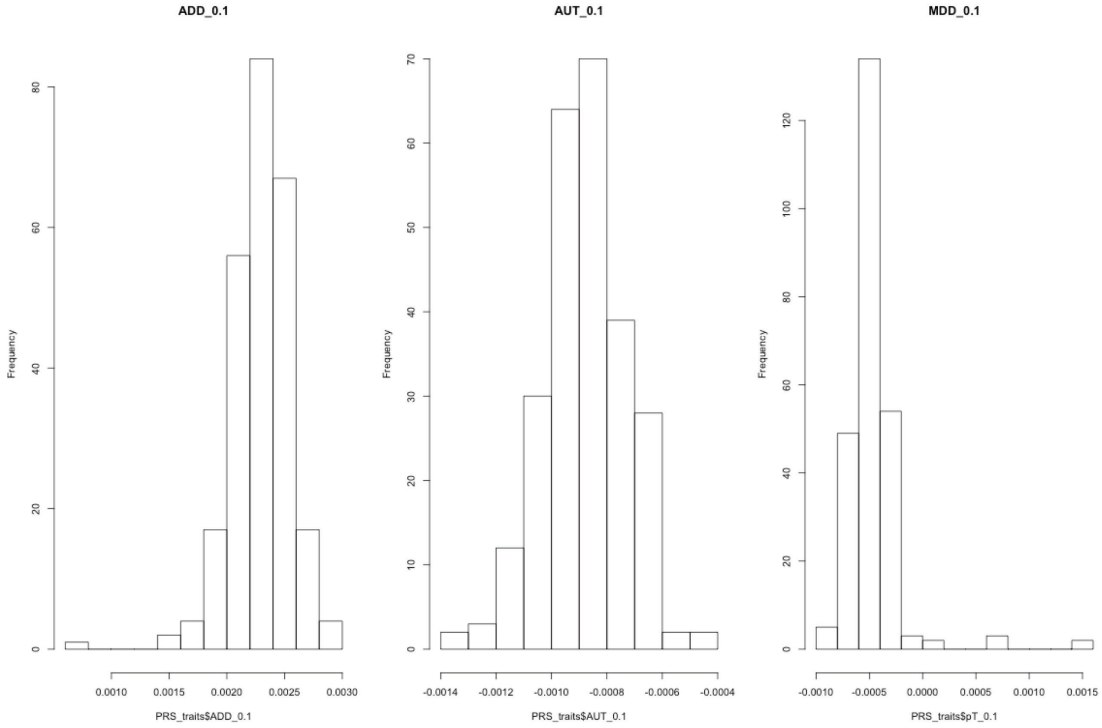


Fig. DS2

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.

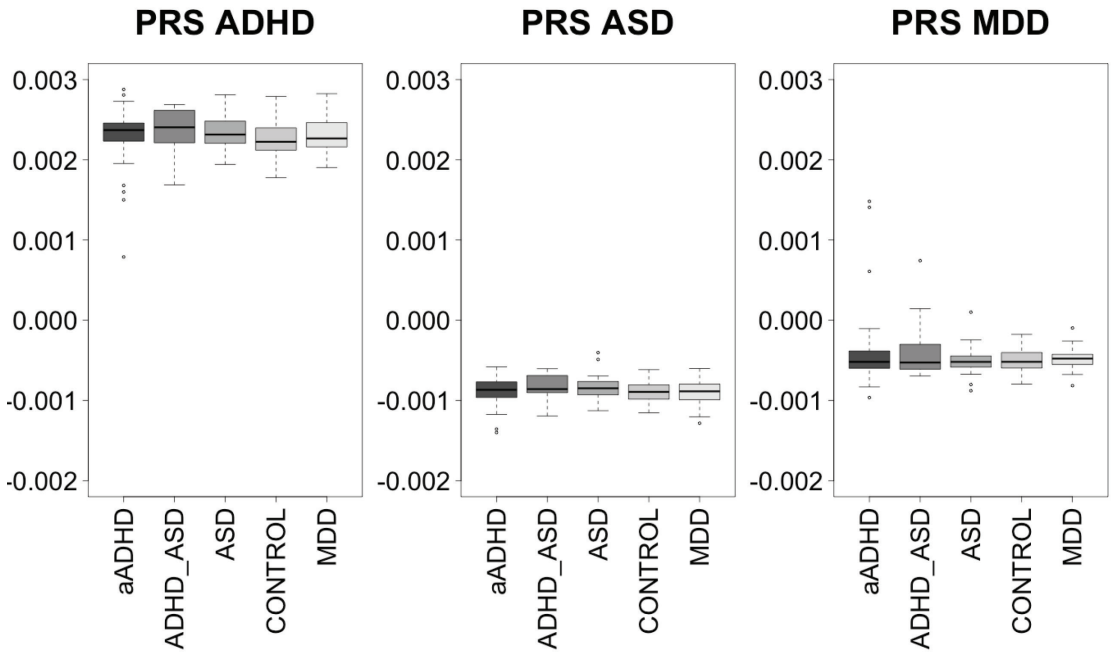


Fig. DS3

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.

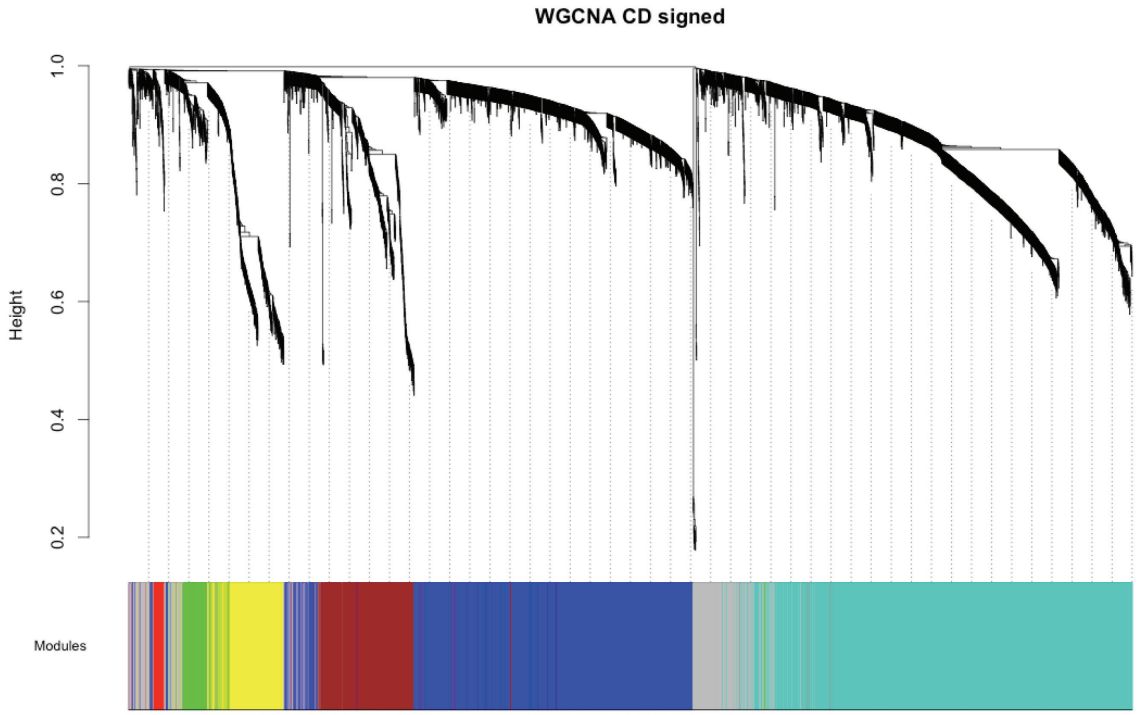


Fig. DS4 Module Trait Correlations

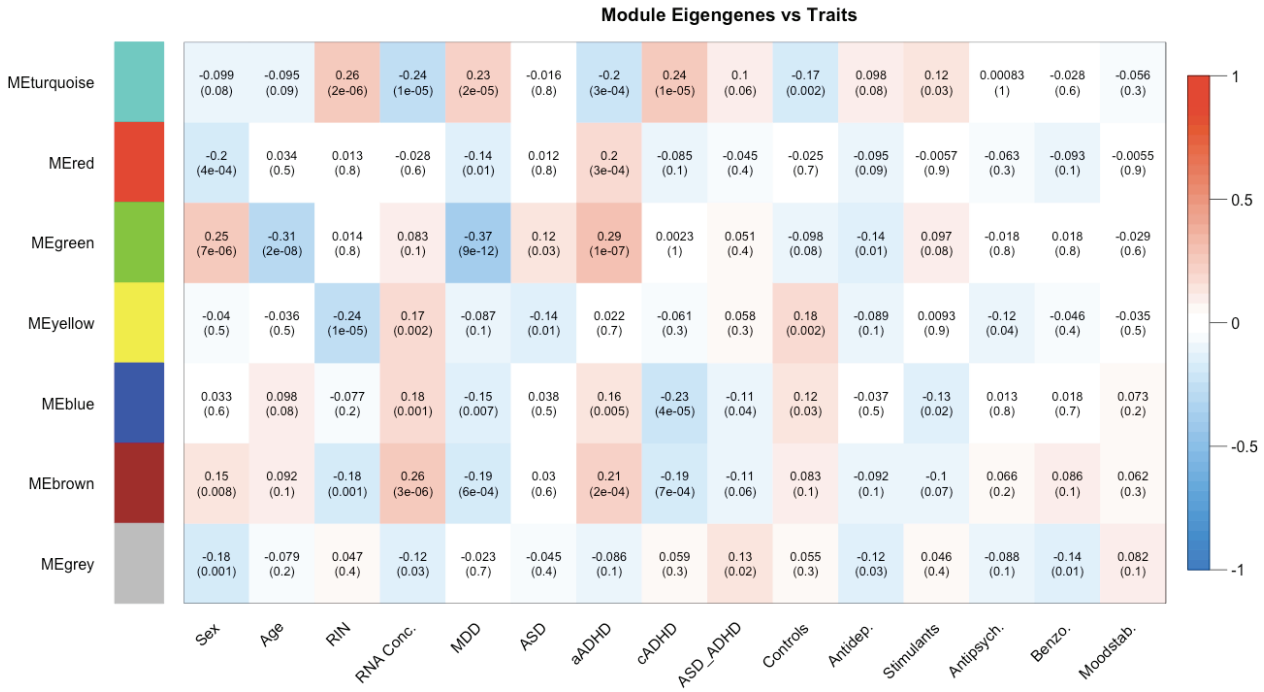
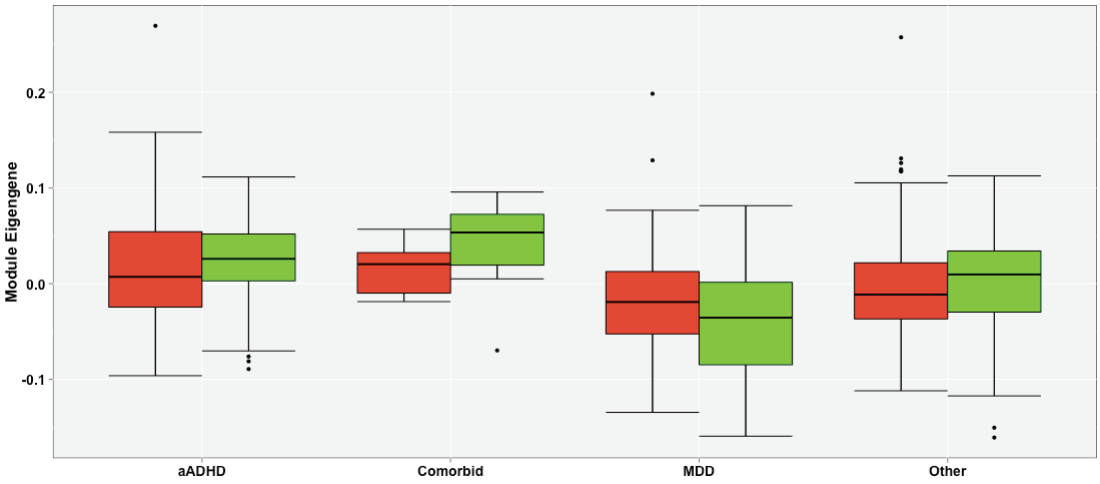


Fig. DS5 Supplementary Information IV: Green and Red ME expression for aADHD, MDD and comorbid subjects



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

1. Cohen-Woods S, Gaysina D, Craddock N, et al. (2009) Depression Case Control (DeCC) Study fails to support involvement of the muscarinic acetylcholine receptor M2 (CHRM2) gene in recurrent major depressive disorder. Hum Mol Genet 18:1504–9. doi: 10.1093/hmg/ddp051

2. Wing JK (1990) SCAN. Arch Gen Psychiatry 47:589. doi: 10.1001/archpsyc.1990.01810180089012

3. McGuffin P, Katz R, Aldrich J (1986) Past and present state examination: the assessment of "lifetime ever" psychopathology. Psychol Med 16:461–5.

4. Lewis CM, Ng MY, Butler AW, et al. (2010) Genome-wide association study of major recurrent depression in the U.K. population. Am J Psychiatry 167:949–57. doi: 10.1176/appi.ajp.2010.09091380

5. Barkley Adult ADHD Rating Scale—IV (BAARS-IV). http://www.guilford.com/books/Barkley-Adult-ADHD-Rating-Scale-IV-BAARS-IV/Russell-Barkley/9781609182038. Accessed 6 May 2015

6. Oldfield RC (1971) The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9:97–113.

7. Wechsler D (1999) Wechsler Abbreviated Scale of Intelligence™. http://www.pearsonclinical.com/education/products/100000593/wechsler-abbreviated-scale-of-intelligence-wasi.html#. Accessed 16 Apr 2015

8. Lord C, Rutter M, Le Couteur A (1994) Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. J Autism Dev Disord 24:659–85.

9. Lord C, Risi S, Lambrecht L, et al. (2000) The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. J Autism Dev Disord 30:205–23.

10. Purcell S, Neale B, Todd-Brown K, et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81:559–575.

11. Patterson N, Price AL, Reich D (2006) Population structure and eigenanalysis. PLoS Genet 2:e190. doi: 10.1371/journal.pgen.0020190

12. Kuhn RM, Haussler D, Kent WJ (2013) The UCSC genome browser and associated tools. Brief Bioinform 14:144–61. doi: 10.1093/bib/bbs038

13. Pedroso I, Lourdusamy A, Rietschel M, et al. (2012) Common genetic variants and gene-expression changes associated with bipolar disorder are over-represented in brain signaling pathway genes. Biol Psychiatry 72:311–7. doi: 10.1016/j.biopsych.2011.12.031

14. Cross-Disorder Group of the Psychiatric Genomics Consortium (2013) Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. Lancet 381:1371–9. doi: 10.1016/S0140-6736(12)62129-1

15. Euesden J, Lewis CM, O'Reilly PF (2014) PRSice: Polygenic Risk Score software. Bioinformatics. doi: 10.1093/bioinformatics/btu848