Pathway Analysis in Attention Deficit Hyperactivity Disorder: An Ensemble Approach

Michael A. Mooney, PhD[1,2]; Shannon K. McWeeney, PhD[1,2,3]; Stephen Faraone, PhD[4,5];
Anke Hinney, PhD[6]; Johannes Hebebrand, MD, PhD[6]; IMAGE2 Consortium[4,5]; German
ADHD GWAS Group[6]; Joel T. Nigg, PhD[7,8]; Beth Wilmot, PhD[1,2,3]

[1]Division of Bioinformatics & Computational Biology, Department of Medical Informatics
& Clinical Epidemiology, Oregon Health & Science University, Portland, OR 97239, USA

[2]OHSU Knight Cancer Institute, Portland, OR 97239, USA

[3]Oregon Clinical and Translational Research Institute, Portland, OR 97239, USA

[4]Departments of Psychiatry and Neuroscience & Physiology, State University of New
York, Syracuse, NY 13120 USA

[5]K.G. Jebsen Centre for Psychiatric Disorders, Department of Biomedicine, University of
Bergen, Bergen, Norway

[6]Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy,
Universitätsklinikum Essen, University of Duisburg-Essen, Essen, Germany

[7]Division of Psychology, Department of Psychiatry, Oregon Health & Science University,
Portland, OR 97239, USA

[8]Department of Behavioral Neuroscience, Oregon Health & Science University, Portland,
OR 97239, USA

Corresponding author:

Beth Wilmot, PhD

3181 SW Sam Jackson Park Rd.

Mail code: CR145

Portland, OR, 97239

USA

phone: 503-494-7510

fax: 503-418-0125

wilmotb@ohsu.edu

IMAGE2 Consortium:

Stephen V. Faraone (Departments of Psychiatry and Neuroscience & Physiology, State University of New York, Syracuse, New York, USA, and K.G. Jebsen Centre for Psychiatric Disorders, Department of Biomedicine, University of Bergen, Bergen, Norway); **Marcel Romanos** and **Andreas Warnke** (Department Of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University of Wuerzburg, Wuerzburg, Germany); **Andreas Reif** (Department of Psychiatry, University of Frankfurt, Frankfurt, Germany); **Susanne Walitza** (Department of Child and Adolescent Psychiatry, University of Zurich, Zurich, Switzerland, and Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University of Wuerzburg, Wuerzburg, Germany); **Herbert Roeyers** (Ghent University, Belgium); **Barbara Franke** (Radboud University Medical Centre, Department of Genetics and Department of Psychiatry, Nijmegen, The

Netherlands); **Jan K. Buitelaar** (Radboud University Medical Centre, Nijmegen, The

Netherlands); **Klaus-Peter Lesch** (Division of Molecular Psychiatry, ADHD Clinical

Research Unit, Department of Psychiatry, Psychosomatics and Psychotherapy, University

of Wuerzburg, Germany, and Department of Neuroscience, School for Mental Health

and Neuroscience (MHENS), Maastricht University, The Netherlands); **Lindsey Kent**

(School of Medicine, University of St Andrews, St Andrews, Fife, Scotland, UK);

Alejandro Arias Vasquez (Radboud University Medical Centre, Nijmegen, Donders

Institute for Brain Cognition and Behavior, Departments of Psychiatry, Human Genetics

and Cognitive Neuroscience, The Netherlands); **Anita Thapar** (MRC Centre for

Neuropsychiatric Genetics and Genomics and Institute of Psychological Medicine and

Clinical Neurosciences, Cardiff, Wales, UK); **Christine Freitag** (Department of Child and

Adolescent Psychiatry, Psychosomatics and Psychotherapy, JW Goethe University,

Frankfurt am Main, Germany); **Kate Langley** (MRC Centre for Neuropsychiatric Genetics

& Genomics, School of Medicine, Cardiff University, Cardiff, Wales, UK); **Michael C.

O'Donovan** (MRC Centre for Neuropsychiatric Genetics and Genomics and Institute of

Psychological Medicine and Clinical Neurosciences, Cardiff, Wales, UK); **Michael J. Owen**

(MRC Centre for Neuropsychiatric Genetics and Genomics and Institute of Psychological

Medicine and Clinical Neurosciences, Cardiff, Wales, UK); **Nanda Lambregts-Rommelse**

(Karakter Child and Adolescent Psychiatry University Center and Department of

Psychiatry, Donders Institute for Brain, Cognition and Behavior, Radboud University

Nijmegen Medical Centre, Nijmegen, The Netherlands); **Richard J.L. Anney** (Department

of Psychiatry, Trinity College Dublin, Trinity Centre for Health Sciences, St. James's

Hospital, Dublin, Ireland); **Aisling Mulligan** (School of Medicine and Medical Science, University College Dublin, Dublin, Ireland); **Aribert Rothenberger** (University Medical Centre Göttingen, Germany); **Hans-Christoph Steinhausen** (University of Basel, Basel, Switzerland); **Michael Gill** (Department of Psychiatry, School of Medicine, Trinity College Dublin, Dublin, Ireland); **Philip Asherson** (Social, Genetic & Developmental Psychiatry Centre, Institute of Psychiatry, King's College London, London, UK)

German ADHD GWAS Group:

Anke Hinney, Özgür Albayrak, Anna-Lena Volckmar, Johannes Hebebrand (Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, Universitätsklinikum Essen, University of Duisburg-Essen, Essen, Germany); **Astrid Dempfle** (Institute of Medical Informatics and Statistics, Christian-Albrechts University Kiel, Kiel, Germany); **Sven Cichon, Per Hoffmann** (Institute of Neuroscience and Medicine (INM-1), Structural and Functional Organization of the Brain, Genomic Imaging, Research Center Juelich, Juelich, Germany and Institute of Human Genetics, University of Bonn, Bonn, Germany and Human Genomics Research Group, Division of Medical Genetics, Universityhospital Basel, Department of Biomedicine); **Markus M. Nöthen** (Institute of Human Genetics, University of Bonn, Bonn, Germany and Department of Genomics, Life & Brain Center, University of Bonn, Bonn, Germany); **Stefan Schreiber** (Institute of Clinical Molecular Biology, University Hospital Schleswig-Holstein, Kiel, Germany); **Susanne Möbus** (Institute for Medical Informatics, Biometry and Epidemiology (IMIBE), University of Duisburg-Essen, Essen, Germany); **H.-Erich**

Wichmann (Institute of Epidemiology I, Helmholtz Zentrum München, German Research Center for Environmental Health, Ingolstädter Landstr. 1, 85764 Neuherberg, Germany; Institute of Medical Informatics, Biometry and Epidemiology, Ludwig-Maximilians-Universität, Munich, Germany; and Institute of Medical Statistics and Epidemiology, Technical University Munich, Germany); **Beate Herpertz-Dahlmann** (Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, RWTH Aachen, University Clinics, Aachen, Germany)**; Judith Sinzig** (Department of Child and Adolescent Psychiatry, University of Cologne, Cologne, Germany and Department of Child and Adolescent Psychiatry and Psychotherapy, LVR – Clinic Bonn, Bonn, Germany); **Gerd Lehmkuhl** (Department of Child and Adolescent Psychiatry, University of Cologne, Cologne, Germany); **Tobias J. Renner** (Department of Child and Adolescent Psychiatry, Universitätsklinikum Tübingen, Tübingen, Germany); **Marcel Romanos** (Department of Child and Adolescent Psychiatry, University of Würzburg, Würzburg, Germany); **Benno G. Schimmelmann** (University Hospital of Child- and Adolescent Psychiatry, University of Bern, Bern, Switzerland)

Abstract:

Attention deficit hyperactivity disorder (ADHD) is a common neurodevelopmental disorder affecting approximately 5% of children. Although a wealth of evidence shows a significant genetic component to the disease, definitive genetic mechanisms have not been identified. Pathway analyses, a subset of gene-set analyses, are methods to extend the knowledge gained from genome-wide association studies (GWAS) by providing functional context for genetic associations. However, a key issue is that there are numerous methods for association testing of gene sets and no real consensus regarding the best approach. The present study applied six pathway analysis methods to identify biological pathways associated with ADHD in two GWAS datasets from the Psychiatric Genomics Consortium. Each of these methods uses a different technique for aggregating individual SNP-level effects to produce a pathway-level association measure. Methods that utilize genotypes to model pathway-level effects were found to identify more replicable pathway associations than methods using summary statistics. In addition, pathways implicated by more than one analysis method were more likely to replicate. A consensus of results across methods was determined by using a simple voting scheme, and by calculating the median p-value. Pathways containing potassium channel genes and others involved in RhoA signaling, glycosaminoglycan biosynthesis, and fibroblast growth factor receptor activity were nominally significant by multiple methods in two independent datasets. These results support previous hypotheses about the role of regulation of neurotransmitter release, neurite outgrowth and axon guidance in

contributing to the ADHD phenotype and suggest the value of cross-method

convergence in evaluating pathway analysis results.

Introduction

Attention deficit hyperactivity disorder (ADHD) is a common and heritable neurodevelopmental disorder that affects approximately 5% of children worldwide. The disorder is characterized by symptoms of inattention, hyperactivity, and impulsivity, and frequently persists in impairing form into adulthood [1].

While the heritability of ADHD has been estimated to be 60-80% [2], definitive genetic mechanisms have not yet been identified. Meta-analyses of candidate gene studies have identified genes consistently associated with ADHD (DAT1, DRD4, DRD5, 5-HTT, HTR1B, SNAP25), although collectively these account for less than 5% of genetic variance in ADHD and none are diagnostic. Unsurprisingly, such studies have also highlighted the genetic heterogeneity among ADHD patients [2-6].

Genome-wide association studies (GWAS) [7-19] have revealed additional candidate genes (e.g. CDH13, SPOCK3, KCNC1, KCNIP1, KCNIP4), although these variants have not achieved genome-wide significance [13,18,20,21]. The most consistent finding is the CDH13 gene, which has been implicated in two family-based GWAS [7,8] and two case-control GWAS [10,12]. Results from studies of other neuropsychiatric disorders [22] suggest that studies with tens of thousands of subjects will likely be needed to reveal more definitive single variant associations.

Gene set methods, which test for association between groups of genes and a trait, offer a means of extending and contextualizing the knowledge gained from GWAS for several reasons. First, ADHD, like other complex diseases, is polygenic in nature, so testing for association with sets of related variants (e.g. those influencing a biochemical

pathway) can provide a functional context for multiple genetic risk factors and potentially yield new mechanisms and treatment targets.

Second, because the number of gene sets is far fewer than the number of SNPs in a GWAS, examining gene sets improves power to detect genetic correlates by reducing the multiple testing correction. A third advantage is that effects due to genetic heterogeneity can be detected. This is related to the issue of small effect sizes, since the result of genetic heterogeneity in a study population will be a mixture of small-effect variants. If multiple small effects are present within a pathway it may be possible to detect their cumulative effect using pathway analysis methods.

ADHD is an ideal candidate for pathway analysis given the evidence supporting a polygenic model of disease susceptibility [16,23-25]. A few pathway analyses, using a variety of pathway definitions and statistical methods, have been conducted on ADHD datasets. Poelmans et al. identified the top 85 genes reported in five ADHD GWAS and performed a literature search for gene functions. They reported that 45 of the 85 GWAS hits could be assigned to a neurodevelopment network involved in directed neurite outgrowth [26]. Similarly, Cristino et al. found that ADHD-associated genes are significantly more interconnected in a protein-protein interaction network than expected by chance [27].

Stergiakouli et al. performed a pathway analysis on an ADHD GWAS dataset consisting of 727 children with ADHD and 5081 controls. Using the ALIGATOR method they found that 13 significant pathways also contained an excess of CNV-affected genes.

Pathways related to cholesterol metabolism, cation channel activity, and CNS development were implicated [14].

Yang et al. applied three analysis methods, INRICH [28], DAPPLE [29], and GREAT [30], to a GWAS dataset consisting of 1400 cases and 963 controls of Chinese descent. Although results from the three methods differed somewhat, common processes, such as cell adhesion, glutamate synaptic development, and axon development, were implicated [16]

Bralten et al. performed a candidate pathway analysis using data from the International Multi-site ADHD Genetics (IMAGE) study [7], consisting of 909 trios. Three candidate gene sets (dopamine/norepinephrine pathway, serotonin pathway, and neuritic outgrowth pathway) were defined using the Ingenuity software (www.ingenuity.com) and a literature review. The three pathways combined were associated with hyperactive/impulsive symptomatology but not inattention symptomatology [31].

Hammerschlag et al. tested 17 expert-curated gene sets of pre- and post-synaptic genes in the IMAGE2 case-control dataset, which consists of 896 cases and 2455 controls [12]. However, none were more strongly associated with ADHD than random gene sets of equal size [32].

The results from these previous gene set analyses performed on ADHD datasets provide further evidence of the polygenic nature of the disorder. However, they also underscore the challenge of interpreting pathway analyses due to the variation among methods. This challenge is substantial because of the large number of ways to define a

gene set and to test for association between a gene set and a phenotype [33]. This issue is highlighted in the recent study by O'Dushlaine et al. that examined gene set (a mixture of Gene Ontology and pathway models) enrichment across five different methods to rank pathways associated with schizophrenia, major depression and bipolar disorder [34].

The present study is unlike prior pathway analyses in ADHD, in that it focuses on methods that use genotypes (rather than summary statistics) to model gene- or pathway-level association measures. Our main hypothesis was that methods utilizing genotypes would better represent the underlying genetic architecture and therefore would identify more replicable pathway associations. We applied four such methods, and compared them with two commonly used methods that rely on summary statistics.

Because of the different results expected from different pathway analysis algorithms, we aimed to discover robust pathway-level effects by identifying a consensus of pathway significance across the methods and multiple independent data sets. Our second hypothesis was that this ensemble approach for identifying robust pathway effects would confirm prior findings that neuro-developmental processes are important genetic mechanisms in ADHD.

Data & Methods

Participants and Genotype Data

Two independent, ADHD case-control, GWAS datasets from the Psychiatric Genomics Consortium, which will be referred to as the (a) *IMAGE2* (N=3351; mean age =

10.5, SD = 2.9) and (b) *German ADHD GWAS* (N=1793; mean age = 11, SD = 2.7) datasets, were used for our analysis [12,13]. Details about these datasets and the genotype QA/QC procedures are available in the Supplementary Methods.

Gene Sets

The pathways tested were obtained from the Pathway Commons database (www.pathwaycommons.org; version 4) [35], which included a total of 3074 human pathways from the following sources: Reactome (www.reactome.org; v46) [36], NCI Pathway Interaction Database (pid.nci.nih.gov; 16-AUG-2012) [37], HumanCyc (humancyc.org; 17.1) [38], and PANTHER (www.pantherdb.org/pathway/; 3.2.1) [39]. This initial collection of pathways was filtered by removing those with only a single gene, those with more than 300 genes, and duplicates (same name and same genes). If two pathways shared the same name, but contained different members, the gene members were merged to create a single pathway. Uniprot IDs were converted to Ensemble gene IDs using the mapping contained in the Ensembl database (version 74). The final set of 2233 pathways ranged in size from 2 to 284 genes (mean=31, SD=39). Because of the different requirements of each analysis method, very small pathways were not tested by all methods. Of the final set of 2233 pathways, 1980 and 2057 were tested by all methods in the IMAGE2 and German ADHD GWAS datasets, respectively. Figure 1 provides an overview of our pathway analysis workflow.

Mapping SNPs to Genes

SNPs were mapped to pathway genes if located within 1Kb of the gene boundaries. Gene and SNP locations were obtained from the Ensembl database (www.ensembl.org; v74). For the IMAGE2 dataset, 52921 SNPs were mapped to 5093 pathway genes. For the German ADHD GWAS dataset, 103128 SNPs were mapped to 6136 pathway genes.

Pathway Analysis Methods

Six pathway analysis methods were applied to both datasets. Four were previously published methods that use the original genotype data rather than SNP p-values: GRASS, PCgamma, PoDA, and NBF [40-43]. Two were previously published methods that utilize SNP p-values: GSEA [44] and Fisher's method for combining p-values [43,45]. See the supplementary methods for more details on these algorithms. To examine the individual SNP effects contributing to pathway associations, SNP-level p-values were calculated using the logistic regression procedure in Plink v1.07 [46].

Adjustment for Pathway Size

Although often overlooked, an obvious confound in interpreting pathway analysis results is that pathways with more SNPs ("larger" pathways) are more likely to be associated with the phenotype [33,44,47,48]. The degree of correlation between pathway size and pathway significance was therefore examined for all methods. When a significant correlation was seen, pathway p-values were adjusted as follows.

For each pathway, a collection of random pathways was constructed in order to calculate a null distribution of p-values. These random pathways were created to

approximately match the number of genes and SNPs in the target pathway. This was accomplished by binning all genes according to the number of SNPs assigned to each gene. Because genes with a large number of SNPs are rare, bins were merged so that each contained approximately 25 genes. Random pathways were then created by sampling the appropriate number of genes from each bin. The adjusted p-value is simply the proportion of random pathways with a p-value smaller than the p-value of the target pathway.

Results

Accounting for Pathway Size

We first considered the effect of pathway size in the IMAGE2 data set. Both the PoDA and GSEA methods have built-in permutation procedures that successfully corrected for size bias (correlation p-values > 0.2). The four other methods all had significant correlations between pathway size and significance of association to ADHD. These effects were small for PCgamma and GRASS (Pearson's correlation coefficients, r, of 0.169 and 0.068, respectively; p-values < 0.002). However, the results from Fisher's method were highly correlated with pathway size ($r = 0.95$, p-value $< 2 \times 10^{-16}$). In addition, there was a significant negative correlation between pathway size and pathway significance (the inverse of the Bayes Factor) reported by the NBF method ($r = -0.40$, p-value $< 2 \times 10^{-16}$).

Therefore, p-values from the PCgamma, GRASS, and Fisher's methods were adjusted for pathway size as described in Methods. This procedure successfully corrected the size

bias for PCgamma and Fisher's methods (correlation p-values > 0.13), but

"overcorrected" and resulted in a slight negative correlation between size and

significance for GRASS (r = -0.045, p-value = 0.036) (Supplementary Figure 1). However,

the adjusted p-value was retained. The results from the NBF method could not be

corrected because the hierarchical model used in that method does not allow for the

application of permutation-based correction.

In the German ADHD GWAS dataset, we repeated these checks. Similar results

regarding the relationship between pathway size and significance were seen (data not

shown), and therefore corrections were applied in the same way. All pathway p-values

reported below are adjusted for pathway size either inherently or by our permutation

procedure. All pathway-level association statistics (both adjusted and unadjusted) and

the number of genes and SNPs in each pathway are reported in Supplementary Tables

1-4.

Comparing Pathway Analysis Algorithms

A total of 1980 pathways were tested by all methods in the IMAGE2 dataset; the

number of pathways reported as nominally significant ranged from 88 for GSEA to 61 for

the NBF method. Pathways reported as nominally significant by Fisher's, PCgamma , and

GRASS were most likely to also be significant by at least one other method (74.6%,

74.1%, and 62.9%, respectively), while those reported as nominally significant by NBF

were least likely to be confirmed by a second method (22.9%) (Table 1).

This initial finding replicated well in the German ADHD GWAS dataset, with the PCgamma, GRASS, and Fisher's methods overlapping most with other methods (74%, 72.5%, and 72.2%, respectively) and the NBF method overlapping the least (25%).

With regard to cross-sample replication of particular pathways associated with ADHD, PCgamma had the highest proportion of nominally significant pathways that were also reported as nominally significant in the German ADHD GWAS dataset (16.8%), followed by GRASS and PoDA (~12%). GSEA, Fisher's Method, and the NBF method all had replication rates below 9% (Table 1). This finding is consistent with our hypothesis that methods utilizing genotypes would identify more replicable associations, the NBF method being an exception. The small sample size of each data set is a limitation of our study and is likely responsible, in part, for the discordance between results in the two datasets.

Next, for each pathway, p-values from both cohorts were combined, using Fisher's method [45], to create a single pooled p-value for each analysis method except NBF (which reports a Bayes factor, not a p-value). The number of methods reporting a pooled p-value ≤ 0.05 was counted and the median pooled p-value for each pathway was calculated. Table 2 shows the top 25 most significant pathways ranked by median pooled p-value. The most significant pathway by any method was the Potassium Channels pathway, with a pooled size-adjusted p-value of 4.11×10^{-5} for the GRASS algorithm.

Given the limited amount of overlap seen among the different methods, discordant pathways were examined in order to gain a better understanding of the differences

16

between methods. We use the term "discordant pathway" to mean one that is reported as significant by only a single method (9.3% of pathways tested by all methods).

We hypothesize that differences in the distribution of SNP-level p-values among pathways may explain some of the discordance across methods. For instance, some methods may be more sensitive to pathways containing a few strong to moderate SNP effects, while others are more sensitive to pathways with many small SNP effects.

To examine differences in genetic effects for discordant pathways, SNP-level p-values were calculated using the logistic regression procedure in Plink v1.07 [46]. Next, each gene was assigned the minimum p-value among all SNPs in that gene. The distribution of the minimum gene-level p-value and the median gene-level p-value for each method's discordant pathways are plotted in Supplementary Figures 2 and 3. These plots show that gene-level effects within pathways implicated by one method are, in some cases, significantly different from the gene-level effects within pathways implicated by another method. For example, pathways reported as significant by only PCgamma tend to have a smaller minimum gene-level p-value compared to pathways reported as significant by only GSEA (t-test p-value < 0.0005 for both IMAGE2 and German ADHD GWAS datasets). This suggests that PCgamma is sensitive to pathways with only a few moderate SNP effects, while GSEA is sensitive to pathways with many small effects.

These observations support previous assertions [34,49] that it may be beneficial to apply multiple analysis methods to a dataset, since the results from different methods

can be complementary. Furthermore, it is likely that pathways reported as significant by multiple methods are more stable and replicable (not due to spurious genetic effects).

For example, 46 pathways were reported as nominally significant by three or more methods in the IMAGE2 data set, while 211 pathways were nominally significant by only a single method. A significantly higher proportion of the pathways identified by three or more methods replicated in the German ADHD GWAS dataset (16 of 46; 34%), compared to the pathways identified by only a single method (35 of 211; 17%) (Fisher's exact test p-value = 0.0078).

Seven pathways were reported nominally significant by more than one method in both cohorts (pathways **bold** in Table 2). Q-Q plots of SNP-level p-values for all SNPs in each of these pathways show an excess of weak effects (Figure 2). These observations are consistent with a polygenic model of disease risk for ADHD, as has been demonstrated previously [16,23-25,50].

Supplemental analyses were done to evaluate the use of imputed genotypes for pathway analysis (Supplementary Tables 5 and 6).

Specific Pathway Findings for ADHD

Pathways reported as nominally significant by at least two methods in both data sets are: Ca activated K+ channels, FGFR1b ligand binding and activation, FGFR2b ligand binding and activation, Potassium Channels, Validated targets of C-MYC transcriptional repression, RhoA signaling pathway, and Chondroitin sulfate biosynthesis. All of these

are expressed in the brain and are relevant to neuro-development. Here we present biological context for these pathways and supporting evidence for their role in ADHD.

Potassium channel genes have been implicated in a number of previous GWAS and pathway analyses of ADHD [7,8,10,16,20]. These findings from genetic studies are supported by research on the role of potassium channels in the regulation of dopaminergic neurons [51]. For instance, Fulton et al. found that a Kv1 channel blocker significantly increased dopamine release in mouse midbrain dopamine neurons, and provided evidence that the D2 dopamine autoreceptor attenuates dopamine release through regulation of Kv1 voltage-gated potassium channels [52].

Pharmacological studies provide additional support for the role of potassium channels in ADHD. Kobayashi et al. found that atomoxetine, a norepinephrine reuptake inhibitor approved for the treatment of ADHD, significantly reduced inward currents through G-protein-activated inwardly rectifying K+ (GIRK) channels expressed in Xenopus oocytes [53]. And Sasaki et al. conducted a preliminary study on the efficacy of tipepidine, reported to inhibit GIRK channel currents [54], to treat childhood ADHD. They found that ADHD Rating Scale IV scores improved significantly for 10 ADHD patients after taking 30mg of tipepidine daily for 4 weeks [55].

Figure 3 shows gene-level association measures (minimum SNP p-value) for all potassium channel genes, along with interactions from the STRING protein-protein interaction database (low-confidence interactions excluded) [56]. Also plotted are the distributions of distance scores, S, (as calculated by the PoDA algorithm) showing a significant difference between cases and controls (odds ratios of 1.41 and 1.81 for the

IMAGE2 and German ADHD GWAS study groups, respectively; FDR adjusted p-values < 3 $\times$ 10^{-13}).

One hypothesis regarding the etiology of ADHD involves a dysregulation of developmental processes, particularly axon guidance and neurite outgrowth [26,57]. A number of the pathways implicated in this study contribute to these neuro-developmental processes, namely the RhoA signaling pathway, pathways involved in proteoglycan metabolism, and pathways involved in fibroblast growth factor receptor activation. Although the role of c-Myc in neurodevelopment has not been studied extensively [58], c-Myc knockout models show significant effects on brain growth [59], and the interaction between c-Myc and RhoA in cancer is well known [60].

A recent review by Stankiewicz and others summarizes the abundance of literature describing the role of Rho family GTPases in neurodevelopment [61]. RhoA in particular has been shown to regulate neuronal survival and migration during development [62-64]. Note that 14 of 45 genes (31%) in the RhoA signaling pathway are also members of the much larger axon guidance pathway (280 genes).

Chondroitin sulfate proteoglycans (CSPGs) are thought to act as inhibitory signals to guide neuronal growth [65,66]. It has been proposed that the inhibitory effect of the Rho/ROCK pathway on neurite growth is mediated by CSPGs [67,68]. Monnier et al. demonstrated that both an inhibitor of Rho and an inhibitor of the ROCK kinase were able to block CSPG inhibition of axon growth [67]. Siebert et al. confirmed this finding and further showed that chondroitinase ABC, which removes the glycosaminoglycan chains from CSPGs, counteracts the inhibition of axon growth [66].

Interestingly, the SPOCK3 gene, which encodes a calcium-binding proteoglycan expressed in the brain, has previously been implicated in GWAS of ADHD and personality disorders [7,17].

Like CSPGs, heparan sulfate proteoglycans (HSPGs) have been shown to play a role in axon guidance and neuronal growth [69,70]. HSPGs may exert their effect through the activation of fibroblast growth factor receptor (FGFR) signaling pathways [71], which are important in neurite outgrowth [72,73] and other neuronal development processes [74]. It has also been suggested that FGRFs may interact with the ADHD-susceptibility gene CDH13 [57].

Discussion

An abundance of data on the genetics of ADHD has been produced in recent years. Although results have been inconsistent, patterns are beginning to emerge. First, multiple studies have demonstrated the polygenic nature of the disorder [16,23-25]. The observation that ADHD is likely due to the cumulative effect of many genes, each contributing only a small effect on their own, explains much of the discordance among previous genetic association studies, which have largely been underpowered to detect small effects.

The predictive value of polygenic risk scores provides hope that larger studies will be able to produce more definitive genetic associations [50]. Furthermore, when taking a higher-level view of the reported genetic associations, a number of cellular processes

have consistently been implicated. For instance, genes involved in cell-cell signaling, adhesion, and neural development have been top hits in multiple studies.

Gaining insights by taking this process-level view is precisely the goal of pathway analyses. Given the variety of algorithms for aggregating SNP-level effects, we aimed to combine the results from multiple analysis methods to identify pathways most likely associated with ADHD. We identified seven pathways reported as nominally significant by multiple analysis methods in two independent data sets (Table 2). Each of these pathways was found to contain an excess of small SNP effects consistent with a polygenic model of disease risk. Furthermore, these pathways provide additional support for previous hypotheses about the etiology of ADHD, particularly related to the regulation of neurotransmitter release, and neuro-developmental processes.

Methods that test for the cumulative effect of multiple genes increase the strength of secondary analyses, and allow researchers to extract additional information from currently available datasets. Our results and others [26] have shown the ability to place individual genetic associations within a meaningful biological context that will help focus future research and guide the development of hypotheses about the mechanisms of ADHD susceptibility.

Supplementary information is available at *Molecular Psychiatry's* website.

Conflict of Interest

Barbara Franke received a speaker fee from Merz. All other authors declare no conflict of interest.

References

1. Faraone SV, Biederman J, Mick E. The age-dependent decline of attention deficit hyperactivity disorder: a meta-analysis of follow-up studies. Psychol Med. 2006 Feb;36(2):159-65. PMID: 16420712

2. Faraone SV, Perlis RH, Doyle AE, Smoller JW, Goralnick JJ, Holmgren MA *et al.* Molecular genetics of attention-deficit/hyperactivity disorder. Biol Psychiatry. 2005 Jun 1;57(11):1313-23. Epub 2005 Jan 21. Review. PMID: 15950004

3. Faraone SV, Khan SA. Candidate gene studies of attention-deficit/hyperactivity disorder. J Clin Psychiatry. 2006;67 Suppl 8:13-20. Review. PMID: 16961425

4. Kebir O, Tabbane K, Sengupta S, Joober R. Candidate genes and neuropsychological phenotypes in children with ADHD: review of association studies. J Psychiatry Neurosci. 2009 Mar;34(2):88-101. Review. PMID: 19270759

5. Gizer IR, Ficks C, Waldman ID. Candidate gene studies of ADHD: a meta-analytic review. Hum Genet. 2009 Jul;126(1):51-90. doi: 10.1007/s00439-009-0694-x. Epub 2009 Jun 9. Review. PMID: 19506906

6. Stergiakouli E, Thapar A. Fitting the pieces together: current research on the genetic basis of attention-deficit/hyperactivity disorder (ADHD). Neuropsychiatr Dis Treat. 2010 Sep 7;6:551-60. doi: 10.2147/NDT.S11322. PMID: 20856918

7. Neale BM, Lasky-Su J, Anney R, Franke B, Zhou K, Maller JB *et al*. Genome-wide association scan of attention deficit hyperactivity disorder. Am J Med Genet B Neuropsychiatr Genet. 2008 Dec 5;147B(8):1337-44. doi: 10.1002/ajmg.b.30866. PMID: 18980221

8. Lasky-Su J, Neale BM, Franke B, Anney RJ, Zhou K, Maller JB *et al.* Genome-wide association scan of quantitative traits for attention deficit hyperactivity disorder identifies novel associations and confirms candidate gene associations. Am J Med Genet B Neuropsychiatr Genet. 2008 Dec 5;147B(8):1345-54. doi: 10.1002/ajmg.b.30867. PMID: 18821565

9. Lasky-Su J, Anney RJ, Neale BM, Franke B, Zhou K, Maller JB *et al.* Genome-wide association scan of the time to onset of attention deficit hyperactivity disorder. Am J Med Genet B Neuropsychiatr Genet. 2008 Dec 5;147B(8):1355-8. doi: 10.1002/ajmg.b.30869. PMID: 18937294

10. Lesch KP, Timmesfeld N, Renner TJ, Halperin R, Röser C, Nguyen TT *et al.* Molecular genetics of adult ADHD: converging evidence from genome-wide

association and extended pedigree linkage studies. J Neural Transm. 2008 Nov;115(11):1573-85. doi: 10.1007/s00702-008-0119-3. Epub 2008 Oct 7. PMID: 18839057

11. Mick E, Todorov A, Smalley S, Hu X, Loo S, Todd RD *et al.* Family-based genome-wide association scan of attention-deficit/hyperactivity disorder. J Am Acad Child Adolesc Psychiatry. 2010 Sep;49(9):898-905.e3. doi: 10.1016/j.jaac.2010.02.014. Epub 2010 May 14. PMID: 20732626

12. Neale BM, Medland S, Ripke S, Anney RJ, Asherson P, Buitelaar J *et al.* Case-control genome-wide association study of attention-deficit/hyperactivity disorder. J Am Acad Child Adolesc Psychiatry. 2010 Sep;49(9):906-20. PMID: 20732627

13. Hinney A, Scherag A, Jarick I, Albayrak Ö, Pütter C, Pechlivanis S *et al.* Genome-wide association study in German patients with attention deficit/hyperactivity disorder. Am J Med Genet B Neuropsychiatr Genet. 2011 Dec;156B(8):888-97. doi: 10.1002/ajmg.b.31246. PMID: 22012869

14. Stergiakouli E, Hamshere M, Holmans P, Langley K, Zaharieva I; deCODE Genetics *et al*. Investigating the contribution of common genetic variants to the risk and

pathogenesis of ADHD. Am J Psychiatry. 2012 Feb;169(2):186-94. PMID: 22420046

15. Ebejer JL, Duffy DL, van der Werf J, Wright MJ, Montgomery G, Gillespie NA *et al.* Genome-wide association study of inattention and hyperactivity-impulsivity measured as quantitative traits. Twin Res Hum Genet. 2013 Apr;16(2):560-74. doi: 10.1017/thg.2013.12. PMID: 23527680

16. Yang L, Neale BM, Liu L, Lee SH, Wray NR, Ji N *et al.* Polygenic transmission and complex neuro developmental network for attention deficit hyperactivity disorder: genome-wide association study of both common and rare variants. Am J Med Genet B Neuropsychiatr Genet. 2013 Jul;162B(5):419-30. doi: 10.1002/ajmg.b.32169. Epub 2013 May 31. PMID: 23728934

17. Weber H, Scholz CJ, Jacob CP, Heupel J, Kittel-Schneider S, Erhardt A *et al.* SPOCK3, a risk gene for adult ADHD and personality disorders. Eur Arch Psychiatry Clin Neurosci. 2014 Aug;264(5):409-21. doi: 10.1007/s00406-013-0476-2. Epub 2013 Nov 29. PMID: 24292267

18. Sánchez-Mora C, Ramos-Quiroga JA, Bosch R, Corrales M, Garcia-Martínez I, Nogueira M *et al.* Case-Control Genome-Wide Association Study of Persistent Attention-Deficit Hyperactivity Disorder Identifies FBXO33 as a Novel

Susceptibility Gene for the Disorder. Neuropsychopharmacology. 2015 Mar;40(4):915-26. doi: 10.1038/npp.2014.267. PMID: 25284319

19. Zayats T, Athanasiu L, Sonderby I, Djurovic S, Westlye LT, Tamnes CK, *et al.* Genome-wide analysis of attention deficit hyperactivity disorder in norway. PLoS One. 2015 Apr 13;10(4):e0122501. doi: 10.1371/journal.pone.0122501. PMID: 25875332

20. Franke B, Neale BM, Faraone SV. Genome-wide association studies in ADHD. Hum Genet. 2009 Jul;126(1):13-50. doi: 10.1007/s00439-009-0663-4. Epub 2009 Apr 22. Review. PMID: 19384554

21. Neale BM, Medland SE, Ripke S, Asherson P, Franke B, Lesch KP *et al.* Meta-analysis of genome-wide association studies of attention-deficit/hyperactivity disorder. J Am Acad Child Adolesc Psychiatry. 2010 Sep;49(9):884-97. doi: 10.1016/j.jaac.2010.06.008. Epub 2010 Aug 1. PMID: 20732625

22. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. Nature. 2014 Jul 24;511(7510):421-7. doi: 10.1038/nature13595. Epub 2014 Jul 22. PMID: 25056061

23. Hamshere ML, Langley K, Martin J, Agha SS, Stergiakouli E, Anney RJ *et al.* High loading of polygenic risk for ADHD in children with comorbid aggression. Am J Psychiatry. 2013 Aug;170(8):909-16. doi: 10.1176/appi.ajp.2013.12081129. PMID: 23599091

24. Groen-Blokhuis MM, Middeldorp CM, Kan KJ, Abdellaoui A, van Beijsterveldt CE, Ehli EA *et al.* Attention-deficit/hyperactivity disorder polygenic risk scores predict attention problems in a population-based sample of children. J Am Acad Child Adolesc Psychiatry. 2014 Oct;53(10):1123-9.e6. doi: 10.1016/j.jaac.2014.06.014. Epub 2014 Aug 19. PMID: 25245356

25. Martin J, Hamshere ML, Stergiakouli E, O'Donovan MC, Thapar A. Genetic risk for attention-deficit/hyperactivity disorder contributes to neurodevelopmental traits in the general population. Biol Psychiatry. 2014 Oct 15;76(8):664-71. doi: 10.1016/j.biopsych.2014.02.013. Epub 2014 Feb 25. PMID: 24673882

26. Poelmans G, Pauls DL, Buitelaar JK, Franke B. Integrated genome-wide association study findings: identification of a neurodevelopmental network for attention deficit hyperactivity disorder. Am J Psychiatry. 2011 Apr;168(4):365-77. doi: 10.1176/appi.ajp.2010.10070948. Epub 2011 Feb 15. Review. PMID: 21324949

27. Cristino AS, Williams SM, Hawi Z, An JY, Bellgrove MA, Schwartz CE, *et al.* Neurodevelopmental and neuropsychiatric disorders represent an interconnected molecular system. Mol Psychiatry. 2014 Mar;19(3):294-301. doi: 10.1038/mp.2013.16. Epub 2013 Feb 26. PMID: 23439483

28. Lee PH, O'Dushlaine C, Thomas B, Purcell SM. INRICH: interval-based enrichment analysis for genome-wide association studies. Bioinformatics. 2012 Jul 1;28(13):1797-9. doi: 10.1093/bioinformatics/bts191. Epub 2012 Apr 17. PMID: 22513993

29. Rossin EJ, Lage K, Raychaudhuri S, Xavier RJ, Tatar D, Benita Y *et al.* Proteins encoded in genomic regions associated with immune-mediated disease physically interact and suggest underlying biology. PLoS Genet. 2011 Jan 13;7(1):e1001273. doi: 10.1371/journal.pgen.1001273. PMID: 21249183

30. McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB *et al.* GREAT improves functional interpretation of cis-regulatory regions. Nat Biotechnol. 2010 May;28(5):495-501. doi: 10.1038/nbt.1630. Epub 2010 May 2. PMID: 20436461

31. Bralten J, Franke B, Waldman I, Rommelse N, Hartman C, Asherson P *et al.* Candidate genetic pathways for attention-deficit/hyperactivity disorder (ADHD)

show association to hyperactive/impulsive symptoms in children with ADHD. J

Am Acad Child Adolesc Psychiatry. 2013 Nov;52(11):1204-1212.e1. doi:

10.1016/j.jaac.2013.08.020. Epub 2013 Sep 5. PMID: 24157394

32. Hammerschlag AR, Polderman TJ, de Leeuw C, Tiemeier H, White T, Smit AB *et al.*

Functional gene-set analysis does not support a major role for synaptic function

in attention deficit/hyperactivity disorder (ADHD). Genes (Basel). 2014 Jul

22;5(3):604-14. doi: 10.3390/genes5030604. PMID: 25055203

33. Mooney MA, Nigg JT, McWeeney SK, Wilmot B. Functional and genomic context

in pathway analysis of GWAS data. Trends Genet. 2014 Sep;30(9):390-400. doi:

10.1016/j.tig.2014.07.004. Epub 2014 Aug 22. PMID: 25154796

34. Network and Pathway Analysis Subgroup of Psychiatric Genomics Consortium.

Psychiatric genome-wide association study analyses implicate neuronal, immune

and histone pathways. Nat Neurosci. 2015 Feb;18(2):199-209. doi:

10.1038/nn.3922. PMID: 25599223

35. Cerami EG, Gross BE, Demir E, Rodchenkov I, Babur O, Anwar N *et al.* Pathway

Commons, a web resource for biological pathway data. Nucleic Acids Res. 2011

Jan;39(Database issue):D685-90. PMID: 21071392

36. Croft D, Mundo AF, Haw R, Milacic M, Weiser J, Wu G *et al.* The Reactome pathway knowledgebase. Nucleic Acids Res. 2014 Jan;42(Database issue):D472-7. doi: 10.1093/nar/gkt1102. Epub 2013 Nov 15. PMID: 24243840

37. Schaefer CF, Anthony K, Krupa S, Buchoff J, Day M, Hannay T *et al.* PID: the Pathway Interaction Database. Nucleic Acids Res. 2009 Jan;37(Database issue):D674-9. PMID: 18832364

38. Romero P, Wagg J, Green ML, Kaiser D, Krummenacker M, Karp PD. Computational prediction of human metabolic pathways from the complete human genome. Genome Biol. 2005;6(1):R2. PMID: 15642094

39. Mi H, Muruganujan A, Thomas PD . PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. Nucleic Acids Res. 2013 Jan;41(Database issue):D377-86. PMID: 23193289

40. Chen LS, Hutter CM, Potter JD, Liu Y, Prentice RL, Peters U *et al.* Insights into colon cancer etiology via a regularized approach to gene set analysis of GWAS data. Am J Hum Genet. 2010 Jun 11;86(6):860-71. doi: 10.1016/j.ajhg.2010.04.014. PMID: 20560206

41. Biernacka JM, Jenkins GD, Wang L, Moyer AM, Fridley BL. Use of the gamma method for self-contained gene-set analysis of SNP data. Eur J Hum Genet. 2012 May;20(5):565-71. doi: 10.1038/ejhg.2011.236. Epub 2011 Dec 14. PMID: 22166939

42. Braun R, Buetow K. Pathways of distinction analysis: a new technique for multi-SNP analysis of GWAS data. PLoS Genet. 2011 Jun;7(6):e1002101. doi: 10.1371/journal.pgen.1002101. Epub 2011 Jun 9. PMID: 21695280

43. Evangelou M, Dudbridge F, Wernisch L. Two novel pathway analysis methods based on a hierarchical model. Bioinformatics. 2014 Mar 1;30(5):690-7. doi: 10.1093/bioinformatics/btt583. Epub 2013 Oct 11. PMID: 24123673

44. Wang K, Li M, Bucan M. Pathway-based approaches for analysis of genomewide association studies. Am J Hum Genet. 2007 Dec;81(6):1278-83. PMID: 17966091

45. Fisher RA. Statistical Methods for Research Workers (4th Edition). Oliver and Boyd, London, 1932.

46. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007 Sep;81(3):559-75. Epub 2007 Jul 25. PMID: 17701901

47. Wang K, Li M, Hakonarson H. Analysing biological pathways in genome-wide association studies. Nat Rev Genet. 2010 Dec; 11(12):843–54. PubMed: 21085203

48. Ramanan VK, Shen L, Moore JH, Saykin AJ. Pathway analysis of genomic data: concepts, methods, and prospects for future development. Trends Genet. 2012 Jul; 28(7):323–32. PubMed: 22480918

49. Gui H, Li M, Sham PC, Cherny SS. Comparisons of seven algorithms for pathway analysis using the WTCCC Crohn's Disease dataset. BMC Res Notes. 2011 Oct 7;4:386. doi: 10.1186/1756-0500-4-386. PMID: 21981765

50. Wray NR, Lee SH, Mehta D, Vinkhuyzen AA, Dudbridge F, Middeldorp CM. Research review: Polygenic methods and their application to psychiatric traits. J Child Psychol Psychiatry. 2014 Oct;55(10):1068-87. doi: 10.1111/jcpp.12295. Epub 2014 Aug 1. PMID: 25132410

51. Dragicevic E, Schiemann J, Liss B. Dopamine midbrain neurons in health and Parkinson's disease: Emerging roles of voltage-gated calcium channels and ATP-sensitive potassium channels. Neuroscience. 2015 Jan 22;284C:798-814. doi:

10.1016/j.neuroscience.2014.10.037. Epub 2014 Oct 30. Review. PMID: 25450964

52. Fulton S, Thibault D, Mendez JA, Lahaie N, Tirotta E, Borrelli E *et al.* Contribution of Kv1.2 voltage-gated potassium channel to D2 autoreceptor regulation of axonal dopamine overflow. J Biol Chem. 2011 Mar 18;286(11):9360-72. doi: 10.1074/jbc.M110.153262. PMID: 21233214

53. Kobayashi T, Washiyama K, Ikeda K. Inhibition of G-protein-activated inwardly rectifying K+ channels by the selective norepinephrine reuptake inhibitors atomoxetine and reboxetine. Neuropsychopharmacology. 2010 Jun;35(7):1560-9. doi: 10.1038/npp.2010.27. PMID: 20393461

54. Hamasaki R, Shirasaki T, Soeda F, Takahama K. Tipepidine activates VTA dopamine neuron via inhibiting dopamine D_2 receptor-mediated inward rectifying K^+ current. Neuroscience. 2013 Nov 12;252:24-34. doi: 10.1016/j.neuroscience.2013.07.044. Epub 2013 Jul 26. PMID: 23896570

55. Sasaki T, Hashimoto K, Tachibana M, Kurata T, Okawada K, Ishikawa M *et al.* Tipepidine in children with attention deficit/hyperactivity disorder: a 4-week, open-label, preliminary study. Neuropsychiatr Dis Treat. 2014 Jan 24;10:147-51. doi: 10.2147/NDT.S58480. eCollection 2014. PMID: 24493927

56. Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A *et al.* STRING v9.1: protein-protein interaction networks, with increased coverage and integration. Nucleic Acids Res. 2013 Jan;41(Database issue):D808-15. PMID: 23203871

57. Rivero O, Sich S, Popp S, Schmitt A, Franke B, Lesch KP. Impact of the ADHD-susceptibility gene CDH13 on development and function of brain networks. Eur Neuropsychopharmacol. 2013 Jun;23(6):492-507. doi: 10.1016/j.euroneuro.2012.06.009. Epub 2012 Jul 12. Review. PMID: 22795700

58. Mainwaring LA, Bhatia B, Kenney AM. Myc on my mind: a transcription factor family's essential role in brain development. Oncotarget. 2010 Jun;1(2):86-8. No abstract available. PMID: 21297222

59. Wey A, Knoepfler PS. c-myc and N-myc promote active stem cell metabolism and cycling as architects of the developing brain. Oncotarget. 2010 Jun;1(2):120-30. PMID: 20651942

60. Sauzeau V, Berenjeno IM, Citterio C, Bustelo XR. A transcriptional cross-talk between RhoA and c-Myc inhibits the RhoA/Rock-dependent cytoskeleton.

Oncogene. 2010 Jul 1;29(26):3781-92. doi: 10.1038/onc.2010.134. PMID: 20453885

61. Stankiewicz TR, Linseman DA. Rho family GTPases: key players in neuronal development, neuronal survival, and neurodegeneration. Front Cell Neurosci. 2014 Oct 7;8:314. doi: 10.3389/fncel.2014.00314. eCollection 2014. PMID: 25339865

62. Katayama K, Melendez J, Baumann JM, Leslie JR, Chauhan BK, Nemkul N *et al.* Loss of RhoA in neural progenitor cells causes the disruption of adherens junctions and hyperproliferation. Proc Natl Acad Sci U S A. 2011 May 3;108(18):7607-12. doi: 10.1073/pnas.1101347108. PMID: 21502507

63. Sanno H, Shen X, Kuru N, Bormuth I, Bobsin K, Gardner HA *et al.* Control of postnatal apoptosis in the neocortex by RhoA-subfamily GTPases determines neuronal density. J Neurosci. 2010 Mar 24;30(12):4221-31. doi: 10.1523/JNEUROSCI.3318-09.2010. PMID: 20335457

64. Cappello S, Böhringer CR, Bergami M, Conzelmann KK, Ghanem A, Tomassy GS *et al.* A radial glia-specific role of RhoA in double cortex formation. Neuron. 2012 Mar 8;73(5):911-24. doi: 10.1016/j.neuron.2011.12.030. PMID: 22405202

65. Maeda N, Ishii M, Nishimura K, Kamimura K. Functions of chondroitin sulfate and heparan sulfate in the developing brain. Neurochem Res. 2011 Jul;36(7):1228-40. doi: 10.1007/s11064-010-0324-y. Epub 2010 Nov 26. Review. PMID: 21110089

66. Siebert JR, Osterhout DJ. The inhibitory effects of chondroitin sulfate proteoglycans on oligodendrocytes. J Neurochem. 2011 Oct;119(1):176-88. doi: 10.1111/j.1471-4159.2011.07370.x. Epub 2011 Aug 16. PMID: 21848846

67. Monnier PP, Sierra A, Schwab JM, Henke-Fahle S, Mueller BK. The Rho/ROCK pathway mediates neurite growth-inhibitory activity associated with the chondroitin sulfate proteoglycans of the CNS glial scar. Mol Cell Neurosci. 2003 Mar;22(3):319-30. PMID: 12691734

68. Siebert JR, Conta Steencken A, Osterhout DJ. Chondroitin sulfate proteoglycans in the nervous system: inhibitors to repair. Biomed Res Int. 2014;2014:845323. doi: 10.1155/2014/845323. Epub 2014 Sep 18. PMID: 25309928

69. de Wit J, Verhaagen J. Proteoglycans as modulators of axon guidance cue function. Adv Exp Med Biol. 2007;600:73-89. Review. PMID: 17607948

70. Nishimura K, Ishii M, Kuraoka M, Kamimura K, Maeda N. Opposing functions of chondroitin sulfate and heparan sulfate during early neuronal polarization.

Neuroscience. 2010 Sep 15;169(4):1535-47. doi:

10.1016/j.neuroscience.2010.06.027. Epub 2010 Jun 19. PMID: 20600662

71. Jastrebova N, Vanwildemeersch M, Rapraeger AC, Giménez-Gallego G, Lindahl U,

Spillmann D. Heparan sulfate-related oligosaccharides in ternary complex

formation with fibroblast growth factors 1 and 2 and their receptors. J Biol

Chem. 2006 Sep 15;281(37):26884-92. Epub 2006 Jun 28. PMID: 16807244

72. Anderson AA, Kendal CE, Garcia-Maya M, Kenny AV, Morris-Triggs SA, Wu T *et al.*

A peptide from the first fibronectin domain of NCAM acts as an inverse agonist

and stimulates FGF receptor activation, neurite outgrowth and survival. J

Neurochem. 2005 Oct;95(2):570-83. Epub 2005 Aug 31. PMID: 16135080

73. Beesley PW, Herrera-Molina R, Smalla KH, Seidenbecher C. The Neuroplastin

adhesion molecules: key regulators of neuronal plasticity and synaptic function. J

Neurochem. 2014 Nov;131(3):268-83. doi: 10.1111/jnc.12816. Epub 2014 Aug

14. Review. PMID: 25040546

74. Woodbury ME, Ikezu T. Fibroblast growth factor-2 signaling in neurogenesis and

neurodegeneration. J Neuroimmune Pharmacol. 2014 Mar;9(2):92-101. doi:

10.1007/s11481-013-9501-5. Epub 2013 Sep 21. Review. PMID: 24057103

Tables

Table 1. Number of Nominally Significant Pathways in the IMAGE2 Dataset

Method	Proportion of nominally significant pathways (p ≤ 0.05) confirmed in at least one other method*	Proportion of nominally significant pathways (p ≤ 0.05) confirmed in the German ADHD GWAS dataset
PCgamma	63 / 85 (74.1 %)	15 / 89 (16.8 %)
GRASS	39 / 62 (62.9 %)	8 / 65 (12.3 %)
PoDA	46 / 75 (61.3 %)	9 / 75 (12 %)
GSEA	45 / 88 (51.1 %)	8 / 93 (8.6 %)
FM	59 / 79 (74.6 %)	5 / 83 (6.0 %)
NBF	14 / 61 (22.9 %)	1 / 84 (1.1 %)

* Denominators in this column are slightly smaller because they reflect only those

pathways tested by all methods. Here we refer to a pathway as *confirmed* in two

different ways: 1) when it is nominally significant (p ≤ 0.05) by a second analysis method

in the IMAGE2 dataset (center column); or 2) when it is nominally significant using the

same analysis method in an independent dataset (the German ADHD GWAS dataset;

right column).

Table 2. Top 25 Most Significant Pathways

Pathway	Pathway Size (SNP Count) IMAGE2 / German ADHD GWAS	Methods with Nominal Significance	Median Pooled P-value
Ca activated K+ channels *	262 / 487	5	0.0010
FGFR1b ligand binding and activation	56 / 126	5	0.0011
FGFR2b ligand binding and activation	64 / 145	5	0.0023
Potassium Channels	1065 / 2117	4	0.0026
Signaling mediated by p38-gamma and p38-delta	58 / 108	4	0.0043
Validated targets of C-MYC transcriptional repression *	243 / 548	5	0.0060

RhoA signaling pathway	295 / 507	4	0.0075
tnf/stress related signaling	111 / 217	5	0.0089
histidine degradation III *	41 / 79	3	0.0113
Dermatan sulfate biosynthesis	85 / 206	4	0.0116
Chondroitin sulfate biosynthesis	219 / 451	4	0.0157
Metabolism of Angiotensinogen to Angiotensins *	70 / 143	3	0.0160
Clearance of Nuclear Envelope Membranes from Chromatin *	39 / 83	4	0.0165
Histidine catabolism *	21 / 52	4	0.0184
Translesion synthesis by DNA polymerases bypassing lesion on DNA template *	11 / 13	4	0.0189
FGFR1 ligand binding and activation	69 / 166	5	0.0197
RAC1 signaling pathway	282 / 446	4	0.0197
Regulation of signaling by CBL	180 / 322	3	0.0224
Caspase-mediated cleavage of cytoskeletal proteins	128 / 208	4	0.0238
FGFR2 ligand binding and activation	81 / 201	4	0.0240
DNA Damage Bypass *	11 / 13	3	0.0254
human cytomegalovirus and map kinase pathways	69 / 143	3	0.0295
Rapoport-Luebering glycolytic shunt	5 / 12	3	0.0298
Thromboxane A2 receptor signaling *	821 / 1586	3	0.0299
LKB1 signaling events	305 / 482	3	0.0304

Pathways in **bold** were reported nominally significant by multiple methods in both the IMAGE2 and German ADHD GWAS datasets. Pathways marked with an * were also nominally significant by at least one method in the post-imputation analysis (Supplementary Tables 1 and 2). For each pathway the following information is provided: the number of SNPs assigned to the pathway in each data set, the number of analysis methods reporting the pathway nominally significant, and the median pooled p-value across all the analysis methods.

Figures

Figure 1. Pathway analysis workflow. Pathways tested were retrieved from the Pathway Commons database. Genotyped (and imputed) SNPs were mapped to genes in the pathways, and six pathway analysis algorithms were used to test for association with ADHD. A random pathway permutation procedure was used to adjust pathway significance for pathway size. Finally, pathways were ranked based on the number of methods reporting significance and the median p-value across methods.

Figure 2. Q-Q plots for seven pathways found nominally significant in both cohorts. Each pathway shows an excess of small SNP effects consistent with a polygenic model of disease risk.

Figure 3. A) The Potassium Channels pathway genes overlaid onto the STRING protein-protein interaction network (low confidence interactions, STRING score < 0.5, were removed). Node size is proportion to the IMAGE2 gene p-value, while label size is proportional to the German ADHD GWAS gene p-value. Green node border indicates a gene p-value <= 0.05 in the IMAGE2 dataset, and a green label indicates the same in the German ADHD GWAS dataset. Gray border or label indicates no SNPs present in a particular gene. B) and C) Pathway of Distinction Analysis (PoDA) *S* scores showing a difference in the distribution between cases and controls in both the IMAGE2 and German ADHD GWAS datasets, respectively.

Supplementary Figure 1. The relationship between gene set association significance and the number of SNPs assigned to the gene set. The gene set size bias affects the applied methods to varying degrees. After correction the comparability of results from different methods is greatly improved.

Supplementary Figure 2. Density plots showing the distribution of gene-level p-values across pathways reported as nominally significant by only one method in the IMAGE2 dataset. The differences among the p-value distributions for each method suggest the methods are sensitive to different types of pathway-level genetic effects (i.e. different compositions of individual SNP effects).

Supplementary Figure 3. Density plots showing the distribution of gene-level p-values across pathways reported as nominally significant by only one method in the German ADHD GWAS dataset.