Exome chip analyses in adult attention deficit hyperactivity disorder

T Zayats1, KK Jacobsen1, R Kleppe1, CP Jacob2, S Kittel-Schneider3, M Ribases4,5,6, JA Ramos-Quiroga4,5,6,7 V Richarte4,5,6,7 M Casas4,5,6,7, NR Mota6, EH Grevet6, M Klein10,19 J Comorinas10,11, J Bralten10, T Galesloot13, AA Vasquez18,19,20 M Cichon16,17,21, MB Gabrielsen22,23, OL Holmen22, CHD Bau8, NR Mota8, EH Grevet9, M Klein10, J Corominas10,11, J Bralten10, T Galesloot12, AA Vasquez10,13,14, S Herms15,16,17, AJ Forstner15,16, J Buitelaar13, L Kiemeney12, SV Faraone1,24, B Cormand25,26,27,28, B Franke10,14, A Reif3, J Haavik1,29 and S Johansson30,31

Attention-deficit/hyperactivity disorder (ADHD) is a highly heritable childhood-onset neuropsychiatric condition, often persisting into adulthood. The genetic architecture of ADHD, particularly in adults, is largely unknown. We performed an exome-wide scan of adult ADHD using the Illumina Human Exome Bead Chip, which interrogates over 250,000 common and rare variants. Participants were recruited by the International Multicenter persistent ADHD CollaboraTion (IMpACT). Statistical analyses were divided into 3 steps: (1) gene-level analysis of rare variants (minor allele frequency (MAF) < 1%); (2) single marker association tests of common variants (MAF ≥ 1%), with replication of the top signals; and (3) pathway analyses. In total, 9365 individuals (1846 cases and 7519 controls) were examined. Replication of the most associated common variants was attempted in 9847 individuals (2077 cases and 7770 controls) using fixed-effects inverse variance meta-analysis. With a Bonferroni-corrected significance level of 1.82E-06, our analyses of rare coding variants revealed four study-wide significant loci: 6q22.1 locus (\(P = 4.46E-08\)), where NT5DC1 and COL10A1 reside; the SEC23IP locus (\(P = 6.47E-07\)); the PSD locus (\(P = 7.58E-08\)) and ZCCHC4 locus (\(P = 1.79E-06\)). No genome-wide significant association was observed among the common variants. The strongest signal was noted at rs9325032 in PPP2R28 (odds ratio = 0.81, \(P = 1.61E-05\)). Taken together, our data add to the growing evidence of general signal transduction molecules (NT5DC1, PSD, SEC23IP and ZCCHC4) having an important role in the etiology of ADHD. Although the biological implications of these findings need to be further explored, they highlight the possible role of cellular communication as a potential core component in the development of both adult and childhood forms of ADHD.

OPEN

INTRODUCTION

Attention-deficit/hyperactivity disorder (ADHD) is the most commonly diagnosed childhood-onset neuropsychiatric condition,1 persisting into adulthood in up to 60% of the cases.2 The adult form of ADHD presents a high risk for developing co-morbid psychiatric conditions, including the burden of disease and, consequently, global impairment, resistance to treatment and costs of illness.3,4

ADHD in children and adolescents is highly heritable with an estimated mean heritability of ~76%.5 Although genome-wide association studies of ADHD show that ~40% of ADHD's heritability can be accounted for by common variants,6 no

1K.G. Jebsen Centre for Neuropsychiatric Disorders, Department of Biomedicine, University of Bergen, Bergen, Norway; 2Section of Molecular Psychiatry, Clinical Research Unit on Disorders of Neurodevelopment and Cognition Center of Mental Health, University of Wuerzburg, Wuerzburg, Germany; 3Department of Psychiatry, Psychosomatic Medicine and Psychotherapy, University Hospital Frankfurt, Frankfurt am Main, Germany; 4Psychiatric Genetics Unit, Group of Psychiatry, Mental Health and Addictions, Vall d’Hebron Research Institute, Universitat Autònoma de Barcelona, Barcelona, Spain; 5Department of Psychiatry, Hospital Universitari Vall d’Hebron, Barcelona, Barcelona, Spain; 6Biomedical Network Research Centre on Mental Health (CIBERSAM), Madrid, Spain; 7Departments of Psychiatry and Legal Medicine, Universitat Autònoma de Barcelona, Barcelona, Spain; 8Department of Genetics, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil; 9Department of Psychiatry, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil; 10Department of Human Genetics, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen, The Netherlands; 11Department of Ophthalmology, Radboud Institute for Molecular Life Sciences, Radboud University Medical Center, Nijmegen, The Netherlands; 12Department for Health Evidence, Radboud University Institute for Health Sciences, Radboud University Medical Center, Nijmegen, The Netherlands; 13Department of Cognitive Neurosciences, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen, The Netherlands; 14Department of Psychiatry, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen, The Netherlands; 15Department of Human Genetics, Donostia Institute for Brain, Cognition and Behaviour, Donostia University Medical Center, Donostia, Spain; 16Department of Biomedicine, University of Basel, Basel, Switzerland; 17Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden; 18NIHR BRC for Mental Health, Institute of Psychiatry, Psychology and Neuroscience and SLaM NHS Trust, King’s College London, London, UK; 19MRC Social Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology and Neuroscience, King’s College London, London, UK; 20Institute of Neuroscience and Medicine, Structural and Functional Organization of the Brain (NMN-1), Research Center Juelich, Juelich, Germany; 21K.G. Jebsen Center for Genetic Epidemiology, Department of Public Health, NTNU, Norwegian University of Science and Technology, Trondheim, Norway; 22Department of Laboratory Medicine, Children’s and Women’s Health, Norwegian University of Science and Technology, Trondheim, Norway; 23Departments of Psychiatry and of Neurosurgery and Physiology, SUNY Upstate Medical University, Syracuse, NY, USA; 24Department of Genética, Microbiología i Estadística, Facultat de Biologia, Universitat de Barcelona, Barcelona, Spain; 25Centre de Investigación BiomédicaAnchors en Red de Enfermedades Raras, Barcelona, Spain; 26Institut de Biomedicina de la Universitat de Barcelona, Barcelona, Spain; 27Institut de Recerca Pediàtrica HosAnchorpital Sant Joan de Déu, Barcelona, Spain; 28Division of Psychiatry, Haukeland University Hospital, Bergen, Norway; 29K.G. Jebsen Centre for Neuropsychiatric Disorders, Department of Clinical Science, University of Bergen, Bergen, Norway and 30Center for Medical Genetics and Molecular Medicine, Haukeland University Hospital, Bergen, Norway. Correspondence: Dr T Zayats or Professor S Johansson, K.G. Jebsen Centre for Neuropsychiatric Disorders, Department of Biomedicine, University of Bergen, Jonas Lies vei 91, Bergen 5009, Norway.
E-mail: tetyana.zayats@uib.no or stefan.johansson@uib.no
Received 22 May 2016; revised 5 August 2016; accepted 17 August 2016

Citation: Transl Psychiatry (2016) 6, e923; doi:10.1038/tp.2016.196
Attention deficit hyperactivity disorder
T Zayats et al

MATERIALS AND METHODS

Subjects

All adult ADHD cases examined in this study were volunteers enrolled through the International Multicenter persistent ADHD CollaboraTion (IMpACT). All patients were diagnosed with ADHD according to DSM-IV criteria, meaning the presence of ADHD symptoms before 7 years of age. Controls were recruited either at an IMpACT site (Brazil, Germany, The Netherlands, UiB Norway and Spain) or through population studies (Germany, The Netherlands, HUNT Norway and the United Kingdom). All subjects were of European descent who were ensured by means of self-report together with multi-dimensional scaling (Supplementary Appendix 1). In Brazil, only Native Brazilians of European descent were recruited by including only those who reported having grandparents of European descent and by performing a morphological classification based on skin color and morphological traits rather than relying on self-classification only.1 The population in Southern Brazil where this study was performed is mainly of European descent and population stratification was not found in the European-derived population of Rio Grande do Sul.12,13 Individuals from Southern Brazil show predominantly European ancestry (94%) according to estimates of interethnic admixture.12,14

All participants provided signed informed consent in accordance with the Declaration of Helsinki. The study was approved by ethics committees in each collaborating country at the corresponding recruitment center: in Germany, by the Ethics Committee at the University of Würzburg (Würzburg, Germany); in Norway, by the regional Ethics Committee for medical and health research ethics, western Norway at the University of Bergen (Bergen, Norway) and the Norwegian Data Inspectorate, the Regional and National Committees for Medical and Health Research Ethics and the Norwegian Directorate of Health at the University of Trondheim (Trondheim, Norway); in Spain, by the Ethics Committee of the Hospital Universitari Vall d’Hebron (HUVH, Barcelona, Spain) at the Department of Psychiatry from the Hospital Universitari Vall d’Hebron (HUVH, Barcelona, Spain); in the Netherlands, by the regional Ethics Committee (Centrale Commissie Mensgebonden Onderzoek: CMO Regio Arnhem—Nijmegen; Protocol number llBC.04.0303), the Institutional Review Board of the Radboud University Medical Center (Nijmegen, the Netherlands) and by the regional Ethics Committee (Centrale Commissie Mensgebonden Onderzoek: CMO Regio Arnhem Nijmegen; 2008/163; ABR: NL23894.091.08) at the outpatient clinic of GGZ Delfland in Delft, to Parnassia, psycho-medical centre in The Hague or at the department of Psychiatry at the Radboud University Nijmegen Medical Centre in Nijmegen, the Netherlands.; in Brazil, by the Ethics Committee of the Hospital de Clínicas de Porto Alegre (Porto Alegre, Brazil) at the ADHD Outpatient Program at the Hospital de Clínicas de Porto Alegre (HCPA); in UK, by the National Research Ethics Committee at the National Adult ADHD Outpatient Clinic at the South London and Maudsley NHS Trust, (London, UK) and the Wellcome Trust Case Control Consortium (WTCCC).

For the discovery phase of the study, samples were available from four European IMpACT sites: Germany, The Netherlands, Norway and Spain. A replication sample, to examine the most significant SNVs from the combined variant analysis identified in the discovery stage, was recruited from the following four IMpACT sites: Brazil, Germany, The Netherlands and the UK. Additional samples were recruited through the WTCCC. A detailed description of all samples is provided in the Supplementary Appendix 1.

Genotyping, genotype calling and quality control

All subjects included in the discovery stage were genotyped on the InfiniumHumanExome array (Illumina, San Diego, CA, USA). The German cases, the full Dutch sample and the full Spanish sample were genotyped on HumanExome-12v1.1_A and the UiB Norwegian sample was genotyped on HumanExome-12v1,8 version of the chip. All genotypes were processed using the Illumina GenomeStudio V2011.1 software, with additional genotype assignments implemented in the zCall.15 Best practice guidelines were used to perform quality control (QC) of genotype calls in GenomeStudio.16 Further QC was carried out on all data sets using PLINK17 with the following steps: (1) genotyping rate threshold was set to 98% both for individuals and for SNVs; (2) Hardy–Weinberg test threshold was set to P < 1.00E–05, (3) overall heterozygosity of individuals was screened based on common (minor allele frequency (MAF) ≥ 1%) and rare (MAF < 1%) SNVs separately, with outliers defined as those outside the range of mean ± 3s.d. of the total heterozygosity observed in the population; (4) relatedness (PI_HAT) threshold was set to 10% and (5) ethnic homogeneity was ensured by means of multi-dimensional scaling with HapMap3 populations. Genotype calling and all QC steps were performed for each data set individually. Those samples that were collected in the same country (namely German cases and German controls, Dutch cases and Dutch controls, Norwegian UiB and HUNT samples) were merged in PLINK and additional QC steps were implemented. Specifically, the screening for heterozygosity, cryptic relatedness and population outliers was performed once more as described above. As the end result, four data sets were produced: a combined German sample, a combined Dutch sample, a combined Norwegian sample and a Spanish sample.

For the replication stage, the most significant three common (MAF > 1%) SNVs were genotyped by KASP assay (LGC Genomics, UK) in IMpACT samples (Brazil, Germany, The Netherlands and the UK). The Dutch replication sample was also genotyped on Perlegen (Perlegen Sciences) and HumanCytoSNP-12 array (Illumina).

Statistical analyses

Statistical analyses were divided into three main stages: (1) examination of rare coding SNVs (MAF < 1%), (2) examination of common SNVs (MAF ≥ 1%) with replication in an independent sample and (3) exploration of biological pathways possibly involved in the development of adult ADHD based on the results from the previous two steps. Prior to the analyses, each subject’s genetic substructure characteristics were estimated with principal components analysis implemented in EIGENSTRAT software for each data set individually.18 Calculation of principal components was performed based only on individuals and SNVs revealing high genotyping rate (≥99%) and common variants (MAF ≥ 1%) after removal of strand ambiguous SNVs and those in high LD (r2 < 0.2). Long stretches of LD were also removed prior to calculation of principal components.19

Examination of rare coding variants

Rare variants were defined as those with MAF < 1%. The variants were combined per gene and tested for association with adult ADHD in RAREMETAL.20 The software first analyzes each data set individually and then combines the generated summary statistics across all the data (meta-analysis). The individual analyses generated single-variant score statistics.
as well as covariance matrices to reflect linkage disequilibrium (LD) among the variants. We performed two meta-analyses at the gene level: the MAF-weighted burden test and the variable threshold test (VT) across all data sets. Only those genes were tested in which at least two variants were observed across all data sets. All individual tests were adjusted for the number of significant principal components as calculated in EIGENSTRAT. Variants of those genes that reached a P-value below 1.00E−05 in the individual association tests were manually evaluated in GenomeStudio in the German, Dutch and Norwegian data sets. GenomeStudio data was not available for the Spanish data set. Single-nucleotide polymorphisms displaying less than optimal clustering were excluded. Bonferroni correction was applied to adjust for multiple testing in the final step of the analyses. Significant association signals, revealed by these analyses of rare variants, were also examined in the context of association among common variants in meta-analysis of this study as well as in that of Psychiatric Genomics Consortium (PGC).

Examination of common variants
Common SNVs were defined as those with MAF ≥ 1%. The variants were tested in PLINK software, assuming an additive model. Each data set was first analyzed separately, including significant EIGENSTRAT principal components as covariates. The results were then combined with the use of fixed-effects inverse variance meta-analysis in METAL, applying genomic control. Only those SNVs present in at least two data sets were meta-analyzed. QQ-plots were constructed to ensure proper distribution of observed statistics.

Replication was attempted for the top three SNVs detected in the discovery phase. In the Brazilian, Dutch and German samples, the association of each top variant was assessed in PLINK with logistic regression assuming an additive model. The association in the British sample was assessed by the chi-square test after constructing 2 × 2 contingency tables of allele counts, as we did not have access to raw genotype data of the controls. Each data set was analyzed individually. The results were combined in the fixed-effects inverse variance meta-analysis implemented in PLINK. The three top common variants were also looked up in the largest published ADHD genome-wide association meta-analysis of PGC.

In silico functional analyses of the significant loci
We performed a comprehensive assessment of all rare SNVs contributing to the association signals that survived the study-wide correction for multiple testing to explore their potential functional impact. Such effect was evaluated in two ways: (1) influence of the genetic variation on the encoded protein and (2) regulatory DNA effects (sequences which regulate gene transcription, gene expression and/or DNA replication). The severity of missense SNVs on the function of the encoded proteins was assessed with sequence alignment, using protein blast and protein homology modeling. Applying homology models of human NT5DC1, the structural details surrounding the sites of missense mutations were investigated in Discovery studio Visualizer (v 4.1; Illumina). In addition, SIFT and PolyPhen predictions were also carried out in Ensembl Variant Effect Predictor (VEP).

The possible regulatory effects were assessed in RegulomeDB, utilizing genomic annotation databases that employ the data curated by the Encyclopedia of DNA Elements (ENCODE) and provide information on functionally important elements. RegulomeDB identifies DNA features and regulatory elements of the human genome.

Exploration of biological pathways: GO terms
To evaluate whether any biological processes may be implicated by our results, we performed gene-set analyses using gene ontology (GO) terms. The analyses were performed in MAGMA. First, a degree of association was calculated for each gene based on both common and rare variants by the use of principal component regression. Rare variants were defined as those with MAF < 1% and a burden score was created for each gene, computed as weighted sum of all rare variants in that gene. All data sets were analyzed individually, adjusted for significant EIGENSTRAT components. Fixed effects meta-analysis was then performed, using the square root of the sample sizes as weights. To evaluate each gene’s contribution to the examined gene sets (GO terms), the P-value of each gene from the meta-analysis was converted to a Z-value and used as an outcome variable for a regression model with gene-set membership as a predictor. Gene size, gene-sets’ gene density as well as global LD were added as covariates to adjust for possible confounding effects and prevent spurious association. We restricted our pathway analysis only to those sets that contained SNVs in at least 10 genes in our total aADHD data (all four data sets). Bonferroni correction was applied to adjust for multiple testing.

Gene expression in brain tissue
Gene expression in the human brain across the lifetime was assessed for the loci surviving the Bonferroni correction. Data was accessed from the Human Brain Transcriptome project (http://hbatlas.org), the Expression Atlas (https://www.ebi.ac.uk/gxa/home) and Allen Brain Atlas (http://www.brain-map.org).

RESULTS
Subjects and genotyping
After QC, 9365 individuals (1846 adult ADHD cases and 7519 controls) were available for the analyses. Supplementary Appendix 1 and Table 1 present the details of the four discovery data sets.

Examination of rare coding variants
As it is still unknown what is the most powerful statistical approach for testing whether a mixture of neutral, risk and protective variants is present within a single gene, we leaned on the view of rare functional variants being under selective pressure and utilized two methods to explore such scenario: (1) the MAF-weighted burden test that defines rare variants with a fixed threshold and places weights inverse to the MAFs of the variants within a gene and (2) the VT that empirically determines the MAF threshold and makes no prior assumption about the relationship between MAFs and effect sizes of the variants within a gene. Out of 14 431 genes with observed variants, 13 715 genes contained at least two variants. We adjusted our significance level for the use of two methods and the 13 715 genes containing more than two rare SNVs (27 430 tests in total), which yielded a Bonferroni-corrected study-wide significance threshold of 1.82E−06. Four loci survived this correction for multiple testing. The most significant association was observed using the VT method, present at 6q22.1 (P = 4.46E−08, MAF threshold = 0.00027), where the 5′-nucleotide domain containing 1 (NT5DC1) gene resides along with the collagen type X alpha 1 (COL10A1) gene. This locus was also the most significant finding using the MAF-weighted burden test (ST1). Supplementary Table 2 presents details of the variants contributing to association signal in 6q22.1 locus.

The three additional study-wide significant associations were observed for the SEC23 interacting protein gene (SEC23IP; P = 6.47E−07, VT method), the pleckstrin and Sec7 domain containing gene (PSD; P = 7.58E−07, MAF-weighted burden test)
and zinc finger, CCHC domain containing 4 gene (ZCCHC4, \( P = 1.79 \times 10^{-6} \), VT method). ST1 summarizes the most significant genes \( (P \leq 1.00 \times 10^{-5}) \) observed in the performed gene-based analyses of rare coding variants.

These four study-wide significant loci revealed no compelling association signals among single point tests of common variants in this study nor in PGC (ST2 and ST3).

Examination of common variants

The most significant signal in our meta-analysis of common variants was observed at rs9325032 in the PPP2R2B gene (odds ratio = 0.81, \( P = 1.61 \times 10^{-5} \)). No variant reached genome-wide significance \( (P < 5.00 \times 10^{-8}) \). ST3 details the top SNVs with association \( P \)-values < 1.00 × 10\(^{-3}\) in the meta-analysis. SF1 depicts the QQ-plots.

Replication was attempted for the top three common SNVs (ST4). No variant reached genome-wide significance (SF2). However, rs9325032 (PPP2R2B gene) showed a modest trend of significance \( (P = 0.033, \text{ST4A}) \), although significant heterogeneity (for example, differences in estimated MAFs) and discordant direction of effect was observed among samples (ST4A). Comparisons using the data for childhood ADHD collected by the PGC revealed no significant findings for the top SNVs in the current study (ST4B).

**In silico** functional analyses of the significant loci

In total, 32 rare coding SNVs contributed to the gene-based association signals passing study-wide Bonferroni correction in the four loci (6q22.1, SEC23IP, PSD and ZCCHC4). No single common variant reached genome-wide significance and, thus, common variants were not assessed further for functionality. First, we evaluated the effect of rare missense variants on the encoded protein in the most significant locus of NT5DC1. The NT5DC1 protein contains a known 5′-nucleotidase domain (aa11–337) and a haloacid dehalogenase-like domain (aa198–324). Table 2 summarizes the results of sequence alignment using protein blast and protein homology modeling (template NP_689942.2) to assess the severity of the observed NT5DC1 missense variants in this locus. Structural inspection of the SNVs in the NT5DC1 locus revealed their potential to form a possible phosphorylation site (rs150257749) and to induce conformational change (rs150293032) (Table 2 and SF3B), both of which may influence the function of the protein. A homology model of human NT5DC1 and structural inspection of the three missense mutation sites in human NT5DC1 are presented in SF3. The NT5DC1 sequence alignment is shown in SF4.

While examining the impact on their respective encoded proteins across all the 32 variants, six were predicted to be deleterious by both PolyPhen and SIFT: one in NT5DC1 (rs150293032), two in PSD (rs142273937 and rs200819772), two in SEC23IP (rs142663854 and rs142266445) and one in ZCCHC4 (rs151252286). Detailed features of these missense variants are presented in ST5.

Apart from surveying the possible direct effects of the variants on their encoded proteins, we also explored their regulatory potential using RegulomeDB. These analyses revealed that the SNVs with the most likely regulatory effect reside within the PSD gene \( (rs1477203944, rs148732359 \text{ and } rs142273937) \) and probably affect the binding properties of polymerase (RNA) II (DNA directed) polypeptide A (POLR2A) as well as may alter CTCF regulatory element binding \( (rs140739855 \text{ and } rs200819772) \) (ST6). In addition, the binding of POLR2A and CTCF may also be affected by the variants in the other two study-wide significant loci: rs73357833 in SEC23IP, rs201763036 in ZCCHC4 (ST6). No significant eQTL effects were observed. Detailed RegulomeBD features of all variants are presented in ST6.
Exploration of biological pathways: GO terms
To evaluate whether certain biological processes may be implicated by our results, we performed gene-set analyses using GO terms. Among the GO terms, 1,844 terms contained data for 10 or more genes in our data set. This brings the Bonferroni-corrected significance threshold to a $P$-value of 2.17E-05. The strongest association was observed for ‘mRNA 3’-end processing (GO:00031124 term, $P = 1.07E−04$). ST7 summarizes the top GO terms ($P \leq 1.00E−03$) observed in the meta-analysis of both common and rare variants.

Gene expression in brain tissue
The expression of the four study-wide significant loci (NT5DC1, SEC23IP, PSD and ZCCHC4) was evaluated throughout life in the Human Brain Transcriptome atlas. The transcriptional trajectories of all four loci revealed variability between pre- and post-natal stages, in line with the general patterns of gene expression in the brain across the lifetime. Supplementary Figure 6 depicts the expression levels of these loci throughout life.

We have also examined the expression patterns of the four aforementioned loci with regards to the brain regions. NT5DC1 revealed the highest expression in substantia nigra as well as other midbrain and hindbrain structures (FANTOM5 and GTEx projects, adult expression levels). The other three loci showed high expression in the cortex and basal ganglia input nuclei (ZCCHC4, FANTOM5 data), and throughout brain (SEC23IP, PSD, mouse data in Allen Brain Atlas).

DISCUSSION
This study aimed to shed light to the genetic architecture of adult ADHD, the form of this disorder that has received relatively little attention so far. Our main findings are the novel, study-wide significant candidate loci for adult ADHD at 6q22.1, where NT5DC1 and COL10A1 reside, as well as the SEC23IP, PSD and ZCCHC4 loci.

NT5DC1 encodes a member of the haloacid dehalogenase superfamily of enzymes, the exact physiological role of which is still largely unknown. On the basis of its sequence similarity to other haloacid dehalogenases, NT5DC1 may be involved in the dephosphorylation of intracellular signaling molecules and our in silico analyses of the SNVs contributing to its association signal suggest that these variants may alter its function (Table 2). The NT5DC1 locus, however, also contains a shorter COL10A1 gene that is embedded in reverse orientation in NT5DC1 intron 6 (S5F). With our sample size it is not possible to statistically distinguish the effects of these two genes from each other, but the strongest association is seen when testing the larger NT5DC1 gene ($P = 4.48E−08$).

Similarly to 6q22.1 locus, the protein encoded by SEC23IP belongs to a family of enzymes as well-intracellular phospholipase A$_1$ family that degrades phospholipids and is involved in membrane trafficking. Altogether with actin cytoskeleton organization, membrane trafficking is also a mechanism through which PSD-encoded protein (a guanine nucleotide exchange factor) mediates a number of neuronal functions, including dendritic spine formation and stabilization. Furthermore, ZCCHC4-encoded zinc-finger DNA-binding protein may also be linked to cell signal transduction as suggested epigenetic silencer of RAS gene family. RAS genes reportedly inhibit apoptosis, a physiological cell death by which unsuccessful neurons are deleted from the central nervous system in the developing brain. Thus, in this study, we noted association signals in several genes involved in signal transduction, highlighting the possible role of cellular communication in the development of adult ADHD. In line with these observations, previous studies have implicated all four of our study-wide significant loci in disorders of the central nervous system, including ADHD and related neurodevelopmental disorders, ADHD co-morbid conditions as well as syndromes where ADHD symptoms are present. In addition, paralogs of the NT5DC1 gene have also been reported to have role in the development of ADHD and other neuropsychiatric disorders. These findings fit with the higher transcript level of NT5DC1 reported in substantia nigra, a major localization of dopaminergic neurons.

Despite being the largest genetic study on adult ADHD, it is still likely to be underpowered to detect variants with very small effect sizes. Thus, no genome-wide significant signal was observed for common polymorphisms (MAF $\geq$ 1%). However, gene-based statistical methods to analyze rare variants combined could potentially improve power. These aggregation tests are based on various assumptions about the underlying genetic architecture and their power depends on the true, and, in case of ADHD, unknown, disease model. In this study, we used two methods to detect gene-based association: (1) the VT method that leans on the view of functional allelic variants being a subject to purifying selection pressure and empirically calculates a threshold below which the variants are more likely to be functional, based on the observed data; and (2) MAF-weighted burden test, which makes the same assumption of functional variants being under selective pressure as the VT method. As all the four study-wide significant loci reveal association signals with both methods (ST1B), it may suggest that the results appear to be robust with regard to different gene-based tests operating under the same hypothesis and that such applications may help us in uncovering underlying genetic architecture of such complex disorders as ADHD.

Taken together, all four study-wide significant loci have been implicated in the function of neural circuitry and communication, cellular mechanisms previously linked to the development of mental disorders, including childhood ADHD. Thus, dysregulation of cellular communication could be a core component in the development of both adult and childhood forms of ADHD. Nonetheless, these findings should be subjected to further examination in larger samples before their role in ADHD can be firmly established.

CONFLICT OF INTEREST
JH has received speaker fees from Lilly, Novartis and Jansen-Cilag. HL has served as a speaker for Eli-Lilly and Shire and has received a research grant from Shire. BF has received consulting fees from Merck. In the past year, SVF received consulting income, travel expenses and/or research support from Ironshore, Arbor, Shire, Akili Interactive Labs, Alcobra, VAYA, and Neurovance and research support from the National Institutes of Health (NIH). His institution is seeking a patent for the use of sodium–hydrogen exchange inhibitors in the treatment of ADHD. In previous years, he received consulting fees or was on Advisory Boards or participated in continuing medical education programs sponsored by: Shire, Alcobra, Otsuka, McNeil, Janssen, Novartis, Pfzer and Eli-Lilly. SYF receives royalties from books published by Guilford Press: Straight Talk about Your Child’s Mental Health and Oxford University Press: Schizophrenia: The Facts. The remaining authors declare no conflict of interest.

ACKNOWLEDGMENTS
We are grateful to the participants of the study. Special acknowledgment is due to Benjamin Neale and Christiana de Leeuw for invaluable insight and suggestions regarding the manuscript data. We thank the UK Exome Chip consortium for access to control the data. This work was supported by the K.G. Jebsen Foundation for Medical Research, University of Bergen, the Western Norwegian Health Authorities (Helse Vest), as well as European Community’s Seventh Framework Programme (FP7/2007–2013) under grant agreement no 602805 and from the European Community’s Horizon 2020 Programme (H2020/2014–2020) under grant agreement no 643051 (MiND). The Nord-Trøndelag Health Study (The HUNT Study) is a collaboration between the HUNT Research Centre (Faculty of Medicine, Norwegian University of Science and Technology NTNU), the Nord-Trøndelag County Council, the Central Norway Health Authority and the Norwegian Institute of Public Health. This work was supported by the European College of Neuropsychopharmacology (Grant reference: Translational Psychiatry (2016), 1 – 7
REFERENCES
17. 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; 81: 559–575.
24. Teshler Galesloot. This research was financially supported by BBMRI-NL, a Research Infrastructure financed by the Dutch government (NWO 184.021.007). SVF is supported by the K.G. Jebsen Centre for Neuropsychiatric Disorders, University of Bergen, Bergen, Norway, the European Community’s Seventh Framework Programme (FP7/2007–2013) under grant agreement no. 278948 (TACTICS), no. 620808 (Aggressotype) and no. 620450 (IMAGEMEND), and from the European Community’s Horizon 2020 Programme (H2020/2014–2020) under grant agreement no. 643051 (MiND). In addition, the work was supported by the grant for the ENIGMA Consortium (grant number US4 EBO20403) from the BD2K Initiative of a cross-NIH partnership. NeuroIMAGE. This study used the sample from the NeuroIMAGE project. NeuroIMAGE was performed between 2009 and 2012, and is the follow-up study of the Dutch part of the International Multisite ADHD Genetics (IMAGE) project, a multi-site, international effort. Funding support for the IMAGE project was provided by NIH grants R01MH63673 and R101MH081803 to Dr Faraone. This work was further supported by an NWO Large Investment Grant 17501007010 and NWO Brain & Cognition an Integrative Approach grant (433-09-242) to Dr Buitelaar, and grants from the University and University Nijmegen Medical Center, University Medical Center Groningen and Accare, and VU University Amsterdam. NeuroIMAGE also receives funding from the European Community’s Seventh Framework Programme (FP7/2007–2013) under grant agreement no. 278948 (TACTICS), no. 620808 (Aggressotype) and no. 620450 (IMAGEMEND). Nijmegen Biomedical Study (NBS): The NBS is a population-based survey conducted at the Department for Health Evidence, and the Department of Laboratory Medicine of the Radboud university medical center. Principal investigators of the Nijmegen Biomedical Study are Lambertus Kieney, André Verbeek, Dorine Swinkels and Barbara Franke. The NBS data are managed by the NBS project team and are available upon request; see www.nijmegenbiomedischetude.nl for an overview of the data available in this study. Current practical coordinator of the NBS is Dr T.E. Galesloot. Readers can contact her to request the data (Tessel. Galesloot@radboudumc.nl). This research was financially supported by BBMRI-NL, a Research Infrastructure financed by the Dutch government (NWO 184.021.007). SVF is supported by the K.G. Jebsen Centre for Neuropsychiatric Disorders, University of Bergen, Bergen, Norway, the European Community’s Seventh Framework Programme (FP7/2007–2013) under grant agreement no. 620805 and NiMh grant R01MH94946. BC was supported by the Spanish Ministerio de Economía y Competitividad (SAF2012-33484, SAF2015-68341-R), AGAUR (2014SGR932), European Commission’s Seventh Framework Programme (under grant agreement no. 620805, Aggressotype) and European Community’s H2020 Programme (under grant agreements 643051, MiND, and 667302, CoCA). MR is a recipient of a Miguel de Servert contract from the Instituto de Salud Carlos III, Spain (CP09/00119). This work was funded by Instituto de Salud Carlos III (PII/00571, PII/01629, PII/011339, PII/01700), and co-financed by the European Regional Development Fund (ERDF), Agency for Research of the University of Catalonia (AGAUR) and University of Barcelona (2014SGR1357) and European Community’s H2020 Programme (under grant agreement no. 643051, MiND).
Supplementary Information accompanies the paper on the Translational Psychiatry website (http://www.nature.com/tp)