

Genetic background of grey matter brain networks related to attention-deficit/hyperactivity disorder

Gido H. Schoenmacker^{1,2*}, Kuaikuai Duan^{3,4}, Kelly Rootes-Murdy⁵, Wenhao Jiang⁵, Pieter J. Hoekstra⁶, Catharina A. Hartman⁷, Jaap Oosterlaan⁸, Martine Hoogman⁹, Barbara Franke^{1,9}, Jessica A. Turner^{4,5}, Jingyu Liu^{4,10}, Tom Claassen², Tom Heskes², Jan K. Buitelaar^{1,11,12}, Alejandro Arias Vásquez^{1,9}

1. Department of Psychiatry, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen, The Netherlands
2. Faculty of Science, Radboud University, Nijmegen, The Netherlands
3. Department of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, USA
4. Tri-Institutional Center for Translational Research in Neuroimaging and Data Science (TReNDS), Atlanta, USA
5. Department of Psychology, Georgia State University, Atlanta, USA
6. Department of Child and Adolescent Psychiatry, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands
7. Department of Psychiatry, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands
8. Department of Clinical Neuropsychology, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands
- 9.
10. Department of Computer Science, Georgia State University, Atlanta, USA
11. Department of Cognitive Neuroscience, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen, The Netherlands
12. Karakter Child and Adolescent Psychiatry University Centre, Radboud University Medical Center, Nijmegen, The Netherlands

* Corresponding author: gido@schoenmacker.nl

Keywords: ADHD, imaging, genetics, stability selection, candidate genes

Acknowledgements

This project has received funding from the European Union’s Seventh Framework Programme for research, technological development and demonstration under grant agreement no 602805 - AGGRESSOTYPE. Additional support for this work was received from the European Community’s Horizon 2020 Programme (H2020/2014 – 2020) under supported by the US. National Institutes of Health through the grant agreement no 728018 (Eat2beNICE). 1R01MH106655. The founder has no active role in this study. This work is part of the research programme Computing Time National Computing Facilities Processing Round pilots 2018 with project number 17666, which is (partly) financed by the Dutch Research Council (NWO). This work was carried out on the Dutch national e-infrastructure with the support of SURF Cooperative. This paper reflects only the authors’ views, and the European Union is not liable for any use that may be made of the information contained therein.

Abstract

Attention-deficit/hyperactivity disorder (ADHD) is a common neurodevelopmental disorder and is associated with structural grey matter differences in the brain. We investigated the genetic background of some of these brain differences in a sample of 899 adults and adolescents consisting of individuals with ADHD and healthy controls. Previous work in an overlapping sample identified three ADHD-related grey matter brain networks located in areas of the superior, middle, and inferior frontal gyrus as well as the cerebellar tonsil and culmen. We associated these brain networks with protein coding genes using a statistical stability selection approach. We identified ten genes, the most promising of which were NR3C2, TRHDE, SCFD1, GNAO1, and UNC5D. These genes are expressed in brain and linked to neuropsychiatric disorders including ADHD. With our results we aid in the growing understanding of the aetiology of ADHD from genes to brain to behaviour.

Introduction

Attention-deficit/hyperactivity disorder (ADHD) is a common neurodevelopmental disorder, affecting an estimated 3.4-7.2% of children (Polanczyk *et al.*, 2007, 2014; Thomas *et al.*, 2015) and 2-3% of adults (Simon *et al.*, 2009; Fayyad *et al.*, 2017) worldwide. ADHD is characterised by symptoms in the domains of inattention and/or hyperactivity-impulsivity and resulting impairment of functioning (American Psychiatric Association, 2013) and is heterogeneous in aetiology, developmental trajectory, and clinical profile (Luo *et al.*, 2019). In 60-80% of cases, ADHD presents with comorbid psychiatric disorders such as oppositional defiant and conduct disorders, autism spectrum disorder, depression, bipolar disorder, borderline personality disorder, anxiety disorders, as well as substance use disorders, sleep disorders, and learning disabilities (Jensen *et al.*, 1997; Ohnishi *et al.*, 2019; Kooij *et al.*, 2012; Katzman *et al.*, 2017; Biederman *et al.*, 1993).

Twin studies have shown that childhood ADHD is highly heritable with genetic factors explaining over 70% of variance (Faraone *et al.*, 2015). For ADHD in adults, heritability is lower than or comparable to childhood ADHD, with self-ratings showing lower heritability (<50%) than ratings by others (Faraone and Larsson, 2019). Despite this high heritability, the search for genetic factors underlying ADHD has proven difficult (Faraone *et al.*, 2005). Only the largest genome-wide association (GWA) meta-analysis to date reported genome-wide significant loci, concluding that common genetic variants account for 22% of phenotypical variance (Demontis *et al.*, 2019). The polygenic nature of ADHD means that individual genetic effects are small, with odds ratios between 1.077-1.198. There is also evidence for a shared genetic background of childhood and adult ADHD (Rovira *et al.*, 2020). In addition, there is considerable overlap between the genetics of ADHD and neurodevelopmental traits in the general population (Martin *et al.*, 2014) as well as other psychiatric disorders (Anttila *et al.*, 2018) such as autism spectrum disorder (Ghirardi *et al.*, 2019; Stergiakouli *et al.*, 2017).

Since ADHD is at least in part a disorder of the brain, it may be expected that the influence of genetic variants on ADHD is mediated by variations in brain structure and function. Indeed, differences in brain development, volume, activity, and connectivity are observed between groups of individuals with ADHD and healthy controls (Klein *et al.*, 2017; Hoogman *et al.*, 2017, 2019). Using magnetic resonance imaging (MRI), specific structural differences in grey matter (GM) tissue have been identified. On average, individuals with ADHD have 3-5% overall reduction in GM volume (Greven *et al.*, 2015; Hoogman *et al.*, 2019). More specifically, ADHD status is associated with reduced cortical surface and thickness (Hoogman *et al.*, 2019) as well as reduction in subcortical structures including

nucleus accumbens, amygdala, caudate nucleus, hippocampus, and putamen (Hoogman *et al.*, 2017). These case-control differences were most pronounced in children. In addition, ADHD in the general population has been associated with reductions in total brain volume (Hoogman *et al.*, 2012; Shaw *et al.*, 2011) as well as GM volume reductions in frontal, temporal, caudate nucleus, and cerebellar regions (Xavier Castellanos *et al.*, 2002).

Studies combining neuroimaging and genetics approaches in ADHD face all the challenges of a genetically and biologically heterogeneous disorder with small effect sizes, complicated by overall small sizes of samples due to the need for both neuroimaging and genetic information to be available. ADHD imaging genetics studies mostly focussing on single genetic variants have so far produced few promising results (Klein *et al.*, 2017). Recently, a multivariate imaging genetic analysis approach, named sparse parallel independent component analysis, has been developed to optimize the independence and sparsity of genetic sources as well as the imaging-genetic associations, and it has been applied to identify the genetic factors underlying GM volume alterations in superior frontal gyrus associated with working memory underperformance in both ADHD adults and adolescents (Duan *et al.*, 2020). Meanwhile, the first significant genome-wide genetic results on the genetic covariation between ADHD and intracranial volume were reported by combining large ($n=11,221-55,374$) GWA meta analyses, thereby circumventing the need for both imaging and genetic data to be available for the same subjects (Klein *et al.*, 2019). In contrast, in the present study we worked with raw imaging and genetics data of 899 individuals and addressed the statistical power issues differently.

Specifically, we engaged ADHD imaging genetics power issues on three fronts. Firstly, we used previously published brain networks associated with ADHD (Liu *et al.*, 2020; Duan *et al.*, 2018). By looking at the most discriminative combinations of brain volumes instead of fixed structural regions, we increased our power to detect ADHD-related differences by limiting the number of comparisons (Xu *et al.*, 2009). Secondly, we reduced information at the genetic level by running a gene-wide association study focussing only on protein-coding genes. By looking for associations at the gene instead of variant level, we reduced the number of variables from millions to thousands. While this approach does not consider variants far outside of known gene areas (which are often implicated in genetic association studies (Schierding *et al.*, 2014; Zhang and Lupski, 2015)), protein coding genes and their immediate environment, i.e. promoter and terminator regions, present a promising initial target due to being interpretable, relatively well-characterised biologically plausible factors of disease. Lastly, instead of traditional univariate association tests that work based on effect size (the estimation of which is highly dependent on sample size (Visscher *et al.*, 2014)), we used a statistical method

which allowed us to prioritise genes based on the stability of association (Meinshausen and Bühlmann, 2010). This increased our power to detect small but consistent effects and thus reduced false positive findings.

Using this approach, we investigated the genetic background of three ADHD-related brain networks (Duan *et al.*, 2018; Liu *et al.*, 2020). We identified a set of 10 genes based on the stability of association with these three networks, examined their expression in the brain, and tested their joint association with ADHD symptoms. With this we aimed to contribute to a growing understanding of the biological pathways from genes via brain to behaviour in ADHD.

Methods

Participants

Our sample consisted of n=899 Dutch participants from the NeuroIMAGE (von Rhein *et al.*, 2015) and the International Multicentre persistent ADHD genetics Collaboration (IMpACT) (Franke *et al.*, 2010) projects. It includes both adults (n=429) and adolescents (n=470) with ADHD and healthy controls. Ethical approval was obtained from the National Institute of Health registered ethical review boards and written, informed consent was obtained from every participant. Individuals with IQ<70, epilepsy, schizophrenia, autism spectrum disorder, and neurological/genetic disorders that might mimic ADHD were excluded. A majority of the sample consisted of related siblings: there were 265 singleton participants (29%), 226 sibling duos (50%), 46 trios (15%), and 11 quartets (5%). From each participant genome-wide genetic data, brain imaging, and extensive phenotyping was collected, described below and summarised in Table 1.

Table 1: Sample description. N: sample size. Age: mean (standard deviation) in years. ADHD symptoms: mean (standard deviation) of ADHD symptom count. ADHD \leq 2: Number of participants with no more than two ADHD symptoms per domain. ADHD \geq 7: Number of participants with seven or more ADHD symptoms per domain.

	N	%Female	Age	Mean ADHD symptoms	N with ADHD\leq2 symptoms	N with ADHD\geq7 symptoms
Adults	429	53%	25.4 (9.4)	6.7 (6.1)	168	202
Adolescents	470	44%	14.6 (2.2)	6.4 (6.4)	210	220

ADHD-related measures

ADHD symptoms were assessed in accordance with DSM-IV criteria. NeuroIMAGE assessed ADHD symptoms using a diagnostic algorithm that included the Conner's rating scale (Conners, 1997); IMPaCT used the ADHD Rating Scale (DuPaul *et al.*, 1998). Both studies assessed attention/working memory using the Wechsler Adult Intelligence Scale Digit Span test (Wechsler *et al.*, 2000), a task in which subjects have to repeat a series of numbers forwards or backwards. A full description of assessment procedures and data collection can be found elsewhere (von Rhein *et al.*, 2015; Franke *et al.*, 2010).

Neuroimaging

Previous work in the adult portion of our sample identified five GM brain networks that were associated with attention tasks or ADHD symptoms (Duan *et al.*, 2018). Three of these networks were later replicated in an adolescent sample (article in preprint (Liu *et al.*, 2020)). The complete neuroimaging procedures are described elsewhere (Duan *et al.*, 2018; Liu *et al.*, 2020). In short, MRI neuroimaging was performed at two scan sites using three 1.5T scanners. Both manual and statistical quality checks were performed to exclude low quality scans. Scans were normalised, smoothed, and corrected for age, sex, and scanner. GM brain networks were estimated in adults only using independent component analysis. A brain network can be interpreted as a collection of brain regions in which voxels express similar GM patterns across participants. Network loadings, i.e. individual contribution of a network to total GM variation, were calculated for each participant. Network loadings in adolescents were calculated using projected adult networks (Liu *et al.*, 2020).

As mentioned, three GM brain networks were associated with ADHD-related measures in both adolescents and adults. Two networks were associated with digit span performance and one with inattention symptom count. The three networks considered in this study are

- #2. the inferior frontal gyrus, which was associated with both maximum forward and backward digit span performance;
- #3. the superior and middle frontal gyri, which were positively associated with maximum backward digit span performance; and
- #4. the cerebellar tonsil and culmen, which were negatively associated with inattention symptom count.

The brain networks are indicated as network #2, #3, and #4, following the numbering in the earlier papers (Liu *et al.*, 2020; Duan *et al.*, 2018). Brain networks #1 and #5 were not considered in this paper because of inconsistent results between adults and adolescents. After correction for family relations and 20 genetic principal components, there were no significant differences in network loadings between adults and adolescents (Welch's t-test; $p=.60$, $p=.92$, $p=.76$ respectively). These data are shown in Figure 1.

The network loadings are not independent: networks #2 and #3 are moderately correlated with $r=.58$ (.54-0.62 95% CI, $p<1e-10$), whereas loadings of both networks #2 and #4 as well as loadings of networks #3 and #4 are weakly correlated with $r=.15$ (.08-.21, $p<1e-10$), and $r=.18$ (.11-.24, $p<1e-10$)

respectively. Each network is also negatively correlated with ADHD symptom count with correlations of $r=-.06$ ($-.13-.00$, $p=.05$), $r=-.08$ ($-.15-.01$, $p=.01$), and $r=-.11$ ($-.18-.05$, $p<.001$) respectively.

Distribution of brain network loadings

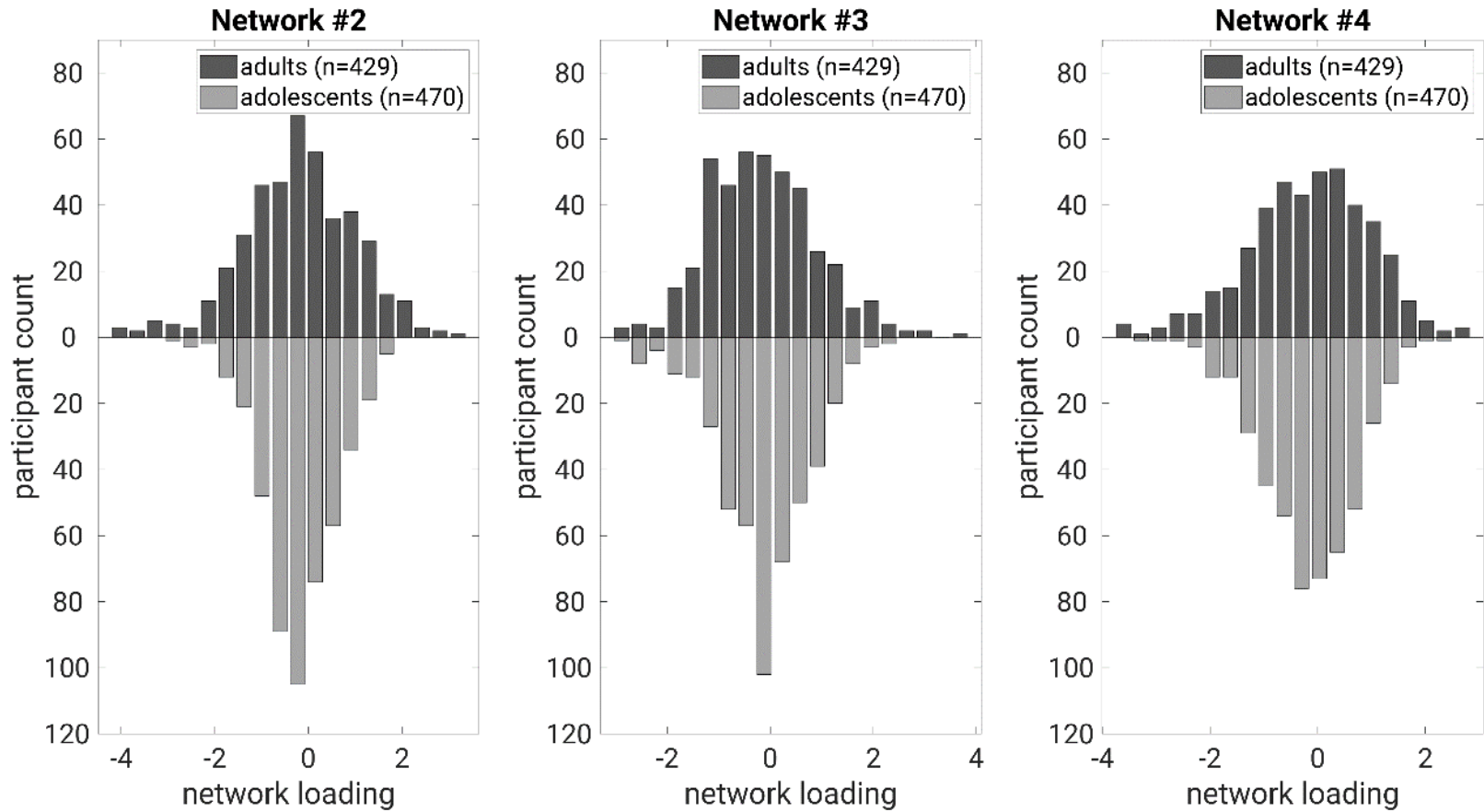


Figure 1: Raw (non-normalised) network loadings for adults (top) and adolescents (bottom) brain networks. Data have been corrected for age, sex, scan site, family relation, and 20 genetic principal components. No significant differences appear between adults and adolescents.

Genetics

For NeuroIMAGE, genome-wide genotyping was performed using the Perlegen and HumanCytoSNP-12 genotyping platforms. For IMPACT, it was performed using the PsychChip platform. Variant-level quality control included >95% call rate, >1% minor allele frequency (MAF), and a Hardy-Weinberg Equilibrium p-value >1e-6. Participants were excluded if the individual call rate was below 95% (K. Brookes *et al.*, 2006; Neale *et al.*, 2010; von Rhein *et al.*, 2015). Imputation was performed using the ENIGMA Consortium imputation protocol (Thompson *et al.*, 2014), which in short consists of imputation using the 1000 Genomes reference population (1000 Genomes Project Consortium *et al.*, 2015) and Minimac imputation software (Fuchsberger *et al.*, 2015; Howie *et al.*, 2012).

Genetic variants were mapped to gene level using the quantitative genetic scoring (QGS) method ((Schoenmacker *et al.*); under review). This method calculates phenotype-agnostic scores for selected genetic regions representing a relative difference from a reference population. QGS values for 12,407 protein coding genes were calculated using variants (MAF >1%) with the European section of 1000 Genomes as a reference population (Consortium *et al.*, 2010; 1000 Genomes Project Consortium *et al.*, 2015). Only protein coding genes were considered to limit the number of variables in our analyses. Gencode release 29 (Frankish *et al.*, 2019) was used for genes and gene locations and a flanking region of 10kb before and 35kb after every gene was included. QGS values were normalised before analysis.

Statistical analyses

Three independent stability selection analyses were performed, one for each brain network. Stability selection was performed using the randomised lasso procedure (Meinshausen and Bühlmann, 2010). One random lasso permutation step consisted of the following procedure. First, a random subsample of size $n/2$ was selected (without repeats). The next steps were performed on this subsample. Second, every predictor (i.e. protein coding gene QGS value) was multiplied by a random weight between .5-1. This random weight is a hyperparameter and its purpose is to better detect the stability of correlated variables. Third, corrected network loadings were regressed against all genes in a lasso-penalised multiple regression with a series of m decreasing penalisation strengths. This means that progressively more predictors are selected by the lasso model. The result is a 12,407 by m binary matrix containing whether a particular gene was selected at a particular penalisation strength.

This procedure was performed for 5000 subsamples for each network. Afterwards, the 5000 binary matrices were summed together and divided by the number of permutations to produce a stability path per gene, consisting of a stability score (i.e. the selection probability) for each penalisation strength.

The penalisation strengths were determined at the start of an analysis and not varied. They consisted of a geometric series beginning with a value just large enough so that no genes are selected and decrease until (on average) about 500 genes are selected. This value of 500 genes corresponds to an upper bound of the per-family error rate $E(V)$ of 100. This upper limit on the error rate was based on the expected number of non-zero variables for a given penalisation strength (Meinshausen and Bühlmann, 2010).

We focused on genes reaching a stability score of .5, or in other words, genes that were selected in more than half of the random subsamples. Any gene with a stability $>.5$ we call consistently associated or “consistent” for short. These consistent genes represent a prioritised selection of the associations with the highest stability.

For all consistent genes, univariate association to the relevant brain network was calculated using linear regression. These associations are not independent from the stability selection and are expected to be strong. Similarly, for all consistent genes univariate associations with ADHD symptoms were calculated. These associations are mostly independent from the stability selection (but not completely due to the -.11 to -.06 correlation between brain networks and ADHD symptom count) and are expected to be weak.

Association of the set of consistent genes with ADHD symptoms was tested using a permutation approach where a set of s genes (where s is the number of consistent genes) were randomly selected and tested for association with ADHD symptoms. This was repeated 10,000 times. In addition, a polygenic risk score (PRS; (Purcell *et al.*, 2009)) approach was used to test the proportion of variance in ADHD symptoms explained using 5, 10, 50, 100, 250, and 500 genes with the highest stability. To this end, the sum of these genes multiplied by their respective brain network association effect size was calculated. Then this sum score was associated to ADHD symptom count. Significance, again, was tested using a randomised gene permutation approach and corrected for multiple testing.

Brain expression of the consistent genes was tested using Human Protein Atlas data (available at proteinatlas.org) (Uhlén *et al.*, 2015). The Human Protein Atlas classifies protein-coding genes into several categories: elevated expression in the brain (compared to non-brain tissue); non-specific expression including brain; elevated non-brain expression; and no expression in brain at all. A binomial test was used to determine whether a significant number of consistent genes fell into the first two categories (elevated brain expression or non-specific brain expression).

Results

Stability

In total ten consistent genes were identified, meaning that these had a selection probability ≥ 0.5 . The results are shown in Table 2. The overall most consistent gene was RELL1 for brain network #4, being selected nearly 60% of the time. As a result, it was also significantly associated with this network ($\beta = -.14$, 95%CI (-.20--.07), $p = 3.5e-5$). Note that overall, eight out of ten regression coefficients for the consistent genes were negative, including RELL1.

Table 2: Consistently associated genes. Below are the ten genes that had a stability ≥ 0.5 . The “Beta” column represents the univariate regression coefficient for the gene with the relevant brain network. The 95% confidence interval for the beta is given, as well as the uncorrected p-value. The “ADHD p-value” column represents the uncorrected p-value for the univariate association to ADHD symptom count.

Gene	Chromosome	Start position	End position	Brain network	Stability	Beta	CI low	CI high	P-value	ADHD p-value
UNC5D	8	35092975	35654068	2	0.57	-0.14	-0.20	-0.07	3.15E-05	0.58
YEATS4	12	69753483	69784576	2	0.55	-0.13	-0.19	-0.06	1.58E-04	0.61
GNAO1	16	56225259	56391369	3	0.55	-0.13	-0.20	-0.07	5.08E-05	0.42
SCFD1	14	31091318	31205018	3	0.51	-0.12	-0.19	-0.06	1.77E-04	0.04
DDO	6	110712974	110736765	3	0.50	-0.11	-0.17	-0.04	1.20E-03	0.68
TENT4B	16	50186829	50269221	3	0.50	-0.12	-0.19	-0.06	1.80E-04	0.25
RELL1	4	37592422	37687998	4	0.60	-0.14	-0.20	-0.07	3.49E-05	0.01
NR3C2	4	148999913	149365850	4	0.58	0.12	0.06	0.19	2.17E-04	0.51
TRHDE	12	72481046	73064537	4	0.58	0.13	0.06	0.19	1.51E-04	0.07
KRT73	12	53001354	53012343	4	0.53	-0.12	-0.19	-0.06	1.87E-04	0.12

For network #2, the most consistent gene was UNC5D, being selected almost 57% of the time. For network #3 this was GNAO1 with almost 55%. The full stability paths are shown in Figure 2. The stability for all 12,406 genes for all three networks can be found in supplementary Table 1. This table also contains the Ensembl stable gene IDs.

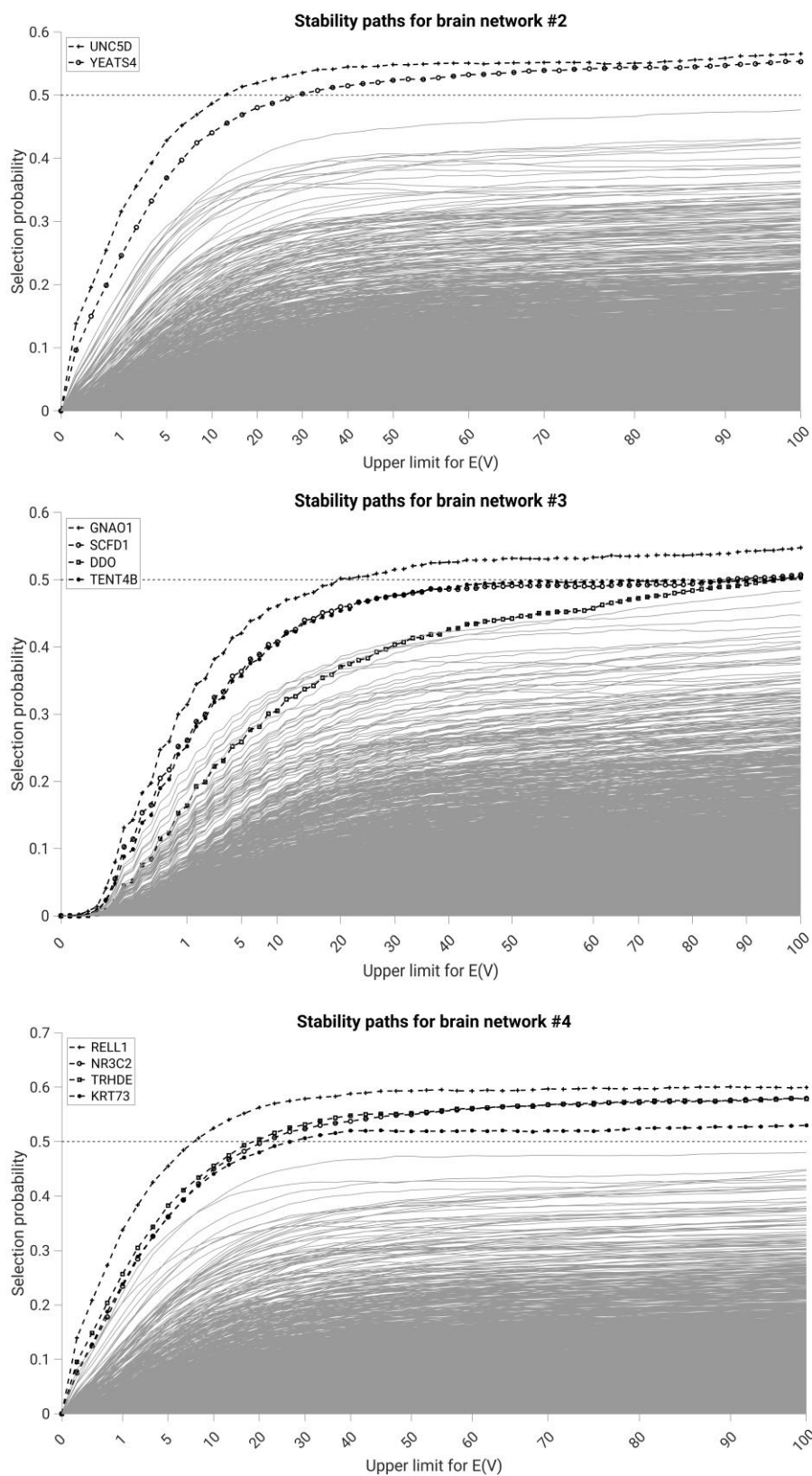


Figure 2: (a) Stability paths for network #2; (b) Stability paths for network #3. (c) Stability paths for network #3. The x-axis shows the upper limit for the expected number of false positives $E(V)$.

Expression in brain

Out of the ten consistent genes, three genes (GNAO1, TRHDE, and UNC5D) had elevated brain expression. An additional six were non-specifically expressed in the brain. This means that brain-expressed genes were significantly over-represented in our ten consistent genes ($p=.03$). The only consistent gene not expressed in the brain was KRT73.

Association with ADHD symptoms

In addition to the highest stability score, RELL1 also had the lowest univariate association p-value with ADHD symptoms in our sample with $p=0.01$ (see Table 2). The ten consistent genes together were significantly associated with ADHD symptoms when compared to ten randomly selected genes ($p=.03$). In addition, the PRS results showed nominally significant associations with ADHD symptoms for all networks. The genes of network #4 were significantly associated with ADHD symptoms with a maximum $R^2=.014$ ($p=.001$) using 50 genes and the genes of network #3 were significantly associated with ADHD symptoms with a maximum $R^2=.010$ ($p=.009$) using 500 genes. The full results are shown in Figure 3.

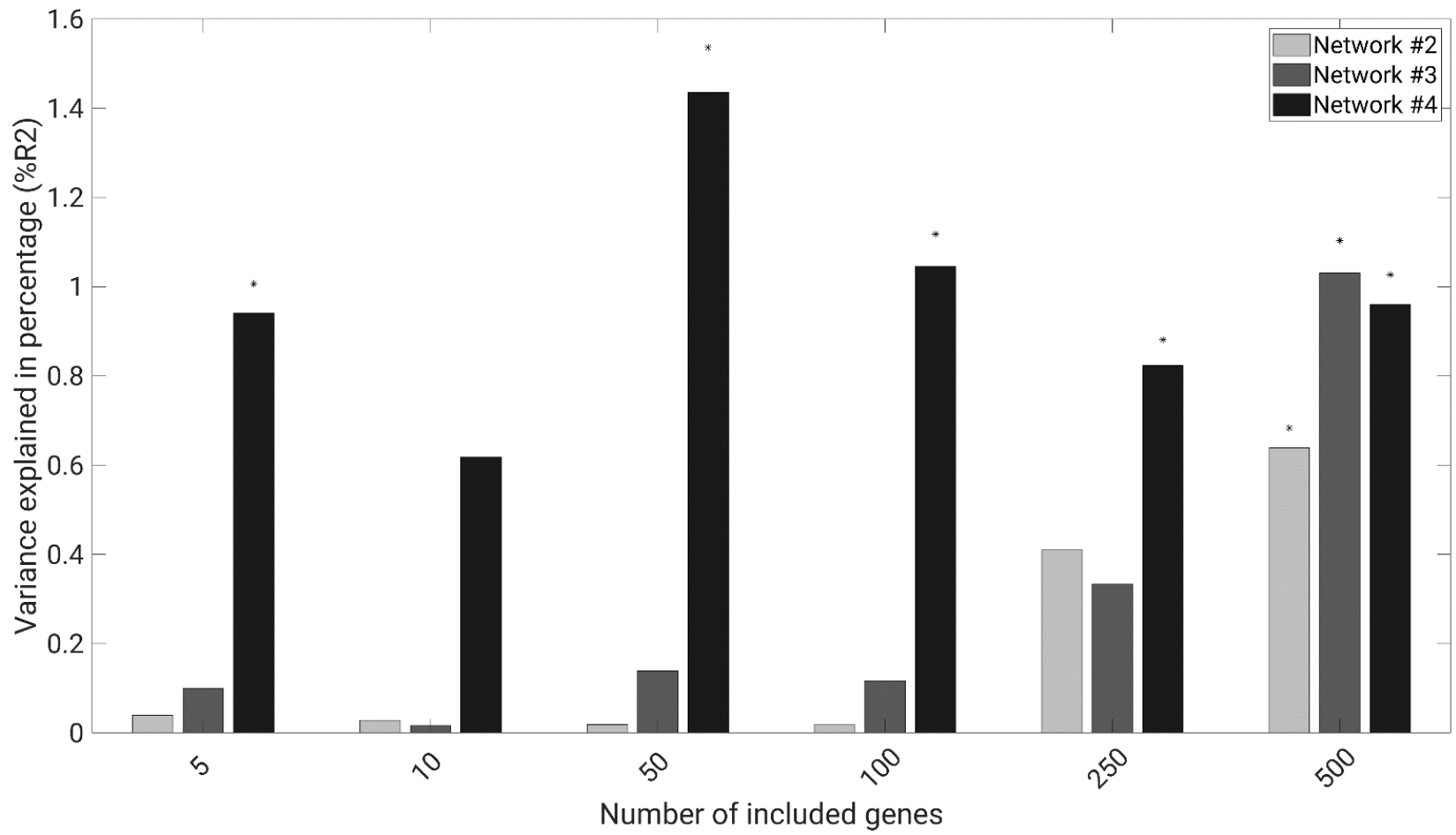


Figure 3: Results of ADHD symptom PRS using the top genes from stability selection. Bars with an asterisk are nominally significant at $p \leq 0.05$.

Discussion

We examined the genetic background of three ADHD-related grey matter brain networks using a resampling-based stability selection approach. We found ten consistently related genes to be associated with those networks, nine of which were expressed in the brain. Additionally, this set of ten genes was significantly associated with ADHD symptoms both as a gene set and as a variable in PRS.

The most consistent results were found for network #4, the cerebellar tonsil and culmen. The RELT-Like Protein 1 (RELL1; stability score 0.60) gene showed both the most consistent association with a brain network and the most significant association with ADHD symptoms. RELL1 encodes a receptor expressed in lymphoid tissues which is involved in apoptosis (Moua *et al.*, 2017). RELL1 is unspecifically expressed in the brain and a common variant in this gene has been tentatively implicated in general cognitive function (rs111283315; $p=1.86e-6$) (Locke *et al.*, 2015). The second-most consistent gene for this network was Nuclear Receptor Subfamily 3 Group C Member 2 (NR3C2; stability score 0.58), which encodes the mineralocorticoid receptor, a binding point for corticosteroid hormones (ter Heegde *et al.*, 2015). Reduced mineralocorticoid receptor expression in the brain has been associated with schizophrenia, bipolar disorder, depression, and suicide (Klok *et al.*, 2011; Medina *et al.*, 2013; Qi *et al.*, 2013; Xing *et al.*, 2004; Young *et al.*, 1998) and various classes of antidepressants increase expression in rats (Yau *et al.*, 1995; Reul *et al.*, 1994; Seckl and Fink, 1992; Bjartmar *et al.*, 2000). The gene itself has been associated with ADHD and the HPA-axis (Kortmann *et al.*, 2013). Significantly, the gene was also predictive of ADHD in conjunction with psychosocial stress in a partially overlapping sample (Van Der Meer *et al.*, 2017). The third-most consistent gene was Thyrotropin Releasing Hormone Degrading Enzyme (TRHDE; stability score 0.58), which (as the name suggests) inactivates thyrotropin releasing hormone. The TRHDE gene overlaps with Tryptophan Hydroxylase 2 (TPH2), sharing 159 variants in our data. TPH2 variants have been associated with ADHD (Waider *et al.*, 2011; K Brookes *et al.*, 2006; Baehne *et al.*, 2009; Sheehan *et al.*, 2005) and – although conflicting findings also exist (Geissler *et al.*, 2017; Sheehan *et al.*, 2007; Johansson *et al.*, 2010) – a meta-analysis concluded that TPH2 variants may affect executive functioning and inattention as well as other neuropsychiatric disorders (Ottenhof *et al.*, 2018). In a recent study, TRHDE has been associated with a slower increase in cerebellum white matter (Brouwer *et al.*, 2020).

Two consistent genes were associated with network #2 consisting of the inferior frontal gyrus. The overall fourth-most consistent gene Unc-5 Netrin Receptor D (UNC5D; stability score 0.57) encodes a receptor for netrin NTN4, which plays a role in neurite outgrowth in the cortex of rodents (Hayano *et*

et al., 2014). In humans, UNC5D has been associated with neuronal function (Srikanth *et al.*, 2018; Sasaki *et al.*, 2008) and implicated in ASD (Walker and Scherer, 2013; Hussman *et al.*, 2011; Gamsiz *et al.*, 2013). The YEATS Domain Containing 4 (YEATS4; stability score 0.55) gene encodes glioma amplified sequence 41, a highly preserved transcription factor with the highest expression in the brain (Munnia *et al.*, 2001), which is essential for cell viability and growth (Park and Roeder, 2006). YEATS4 has been associated with depression and its treatment response (Yamagata *et al.*, 2017; Amare *et al.*, 2018).

The remaining four consistent genes were associated with network #3, the superior and middle frontal gyri. The most consistent gene was G Protein Subunit Alpha O1 (GNAO1; stability score 0.55), which encodes for a heterotrimeric G protein. GNAO1 function contributes significantly to synaptic neurotransmission and neurodevelopment and is associated with neurological disorders such as epilepsy, movement disorders, and developmental delay (Feng *et al.*, 2018, 2019; Larrivee *et al.*, 2020; Saitsu *et al.*, 2016; Danti *et al.*, 2017). In addition, variants in GNAO1 have been associated with depression, neuroticism, circadian rhythm, and cognitive performance (Baselmans *et al.*, 2019; Jansen *et al.*, 2019; Lee *et al.*, 2018). The second-most consistent gene Sec1 Family Domain Containing 1 (SCFD1; stability score 0.51) has been prioritized as a causal gene for ADHD in different brain regions including cerebellum and hippocampus (Fahira *et al.*, 2019) and has been associated with Alzheimer's disease (Stamati *et al.*, 2019). The D-Aspartate Oxidase (DDO; stability score 0.50) gene plays a crucial role in the central nervous and neuroendocrine systems (Katane *et al.*, 2015) and has been associated with depression (Howard *et al.*, 2018). Lastly, the Terminal Nucleotidyltransferase 4B (TENT4B; stability score 0.50) gene is involved in RNA metabolism and has been associated with schizophrenia (Goes *et al.*, 2015).

In short, three of our ten selected genes have previously been implicated in ADHD. Other associated phenotypes include depression (4/10), cognitive performance (2/10), and schizophrenia (2/10). The strong link between our consistent genes and neuropsychiatric phenotypes suggests that the stability selection prioritisation method is able to identify promising candidates for ADHD imaging genetics. This is further reinforced by the significant over-representation of brain-expressed genes in our consistent set, as well as the small (i.e. variance explained of around 1%) but significant association with ADHD symptoms in our sample.

This significant association with ADHD symptom count could in part be explained by the correlation ($-.11 < r < -.06$) between network loadings and ADHD symptoms. We did not statistically correct for this possibility, because we did not expect genes to have a fully independent effect on behaviour. In

fact, a post-hoc mediation analysis (data not presented) showed that the effect of the selected genes on ADHD is fully mediated by the brain networks.

Eight out of ten genetic associations with the brain networks were negative, meaning that (because of the negative correlation between networks and ADHD) their association with ADHD was positive. This imbalance in direction of effect was predicted our previous work (Schoenmacker *et al.*) and may be explained by the fact that genetic rarity – encoded by QGS as difference from the general population – is often associated with risk. Recent findings suggest that evolutionary pressure against ADHD may exist in modern environments (Esteller-Cucala *et al.*, 2020) and our results lend modest support to the idea that genetic variation is more often a risk for ADHD than a protective factor.

In summary, we identified consistent associations between ten genes and ADHD-related brain networks and discussed their function and association, but we did not investigate underlying causative biological mechanisms. Employing purely statistical approach, our findings represent a list of candidates and should be seen as a starting point for further study. Further steps may include functional characterisation or bioinformatics approaches such as molecular landscape analysis. Only protein coding genes were considered, therefore, a majority of genetic variance was excluded from our analysis a priori. The currently not included genetic regions might be reconsidered for inclusion if /when larger samples become available.

By employing statistical data reduction methods in both the imaging and genetic domain and combining these with an association method based on consistency of association instead of effect size, we have created a list of ten candidate genes for ADHD imaging genetics. Three of these genes – NR3C2, TRHDE/TPH2, and SCFD1 – were previously associated with ADHD, and three have elevated expression in the brain: GNAO1, TRHDE/TPH2, and UNC5D. Our results and candidate list aid the growing understanding of the aetiology of ADHD from genes to brain to behaviour.

Bibliography

- 1000 Genomes Project Consortium, T. 1000 G.P. *et al.* (2015) A global reference for human genetic variation. *Nature*, **526**, 68–74.
- Amare, A.T. *et al.* (2018) Association of the polygenic scores for personality traits and response to selective serotonin reuptake inhibitors in patients with major depressive disorder. *Front. Psychiatry*, **9**.
- American Psychiatric Association (2013) Diagnostic and statistical manual of mental disorders (DSM-5®) American Psychiatric Pub.
- Anttila, V. *et al.* (2018) Analysis of shared heritability in common disorders of the brain. *Science (80-.)*, **360**.
- Baehne, C.G. *et al.* (2009) Tph2 gene variants modulate response control processes in adult ADHD patients and healthy individuals. *Mol. Psychiatry*, **14**, 1032–1039.
- Baselmans, B.M.L. *et al.* (2019) Multivariate genome-wide analyses of the well-being spectrum. *Nat. Genet.*, **51**, 445–451.
- Biederman, J. *et al.* (1993) Patterns of psychiatric comorbidity, cognition, and psychosocial functioning in adults with attention deficit hyperactivity disorder. *Am. J. Psychiatry*, **150**, 1792–1798.
- Bjartmar, L. *et al.* (2000) Selective effects on NGFI-A, MR, GR and NGFI-B hippocampal mRNA expression after chronic treatment with different subclasses of antidepressants in the rat. *Psychopharmacology (Berl.)*, **151**, 7–12.
- Brookes, K. *et al.* (2006) The analysis of 51 genes in DSM-IV combined type attention deficit hyperactivity disorder: association signals in DRD4, DAT1 and 16 other genes. *Mol. Psychiatry*, **11**, 934–953.
- Brookes, K. *et al.* (2006) The analysis of 51 genes in DSM-IV combined type attention deficit hyperactivity disorder: Association signals in DRD4, DAT1 and 16 other genes. *Mol. Psychiatry*, **11**, 934–953.
- Brouwer, R.M. *et al.* (2020) Dynamics of Brain Structure and its Genetic Architecture over the Lifespan. *bioRxiv*, 2020.04.24.031138.
- Conners, C. (1997) Conners' Rating Scales--revised: User's Manual Multi-Health Systems, Incorporated.
- Consortium, T. 1000 G.P. *et al.* (2010) A map of human genome variation from population scale sequencing. *Nature*, **467**, 1061.
- Danti, F.R. *et al.* (2017) GNAO1 encephalopathy: Broadening the phenotype and evaluating treatment and outcome. *Neurol. Genet.*, **3**.
- Demontis, D. *et al.* (2019) Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nat. Genet.*, **51**, 63–75.
- Duan, K. *et al.* (2018) Neural correlates of cognitive function and symptoms in attention-deficit/hyperactivity disorder in adults. *NeuroImage Clin.*, **19**, 374–383.
- Duan, K. *et al.* (2020) Sparse parallel independent component analysis and its application to identify linked genomic and gray matter alterations underlying working memory impairment in attention-deficit/hyperactivity disorder. *bioRxiv*, 2020.07.11.198622.

- DuPaul,G.I. *et al.* (1998) The ADHD Rating Scale-IV Checklist, Norms and Clinical Interpretation Guilford Press.
- Esteller-Cucala,P. *et al.* (2020) Genomic analysis of the natural history of attention-deficit/hyperactivity disorder using Neanderthal and ancient Homo sapiens samples. *Sci. Rep.*, **10**, 1–11.
- Fahira,A. *et al.* (2019) Prediction of causal genes and gene expression analysis of attention-deficit hyperactivity disorder in the different brain region, a comprehensive integrative analysis of ADHD. *Behav. Brain Res.*, **364**, 183–192.
- Faraone,S. V. *et al.* (2015) Attention-deficit/hyperactivity disorder. *Nat. Rev. Dis. Prim.* 2015 11, **1**, 1–23.
- Faraone,S. V. *et al.* (2005) Molecular genetics of attention-deficit/hyperactivity disorder. *Biol. Psychiatry*, **57**, 1313–1323.
- Faraone,S. V. and Larsson,H. (2019) Genetics of attention deficit hyperactivity disorder. *Mol. Psychiatry*, **24**, 562–575.
- Fayyad,J. *et al.* (2017) The descriptive epidemiology of DSM-IV Adult ADHD in the World Health Organization World Mental Health Surveys. *ADHD Atten. Deficit Hyperact. Disord.*, **9**, 47–65.
- Feng,H. *et al.* (2018) A mechanistic review on GNAO1-associated movement disorder. *Neurobiol. Dis.*, **116**, 131–141.
- Feng,H. *et al.* (2019) Mouse models of GNAO1-associated movement disorder: Allele- and sex-specific differences in phenotypes. *PLoS One*, **14**.
- Franke,B. *et al.* (2010) Multicenter analysis of the SLC6A3/DAT1 VNTR haplotype in persistent ADHD suggests differential involvement of the gene in childhood and persistent ADHD. *Neuropsychopharmacology*, **35**, 656–664.
- Frankish,A. *et al.* (2019) GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res.*, **47**, D766–D773.
- Fuchsberger,C. *et al.* (2015) minimac2: faster genotype imputation. *Bioinformatics*, **31**, 782–784.
- Gamsiz,E.D. *et al.* (2013) Intellectual disability is associated with increased runs of homozygosity in simplex autism. *Am. J. Hum. Genet.*, **93**, 103–109.
- Geissler,J.M. *et al.* (2017) No genetic association between attention-deficit/hyperactivity disorder (ADHD) and Parkinson’s disease in nine ADHD candidate SNPs. *ADHD Atten. Deficit Hyperact. Disord.*, **9**, 121–127.
- Ghirardi,L. *et al.* (2019) Genetic and environmental contribution to the overlap between ADHD and ASD trait dimensions in young adults: A twin study. *Psychol. Med.*, **49**, 1713–1721.
- Goes,F.S. *et al.* (2015) Genome-wide association study of schizophrenia in Ashkenazi Jews. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.*, **168**, 649–659.
- Greven,C.U. *et al.* (2015) Developmentally stable whole-brain volume reductions and developmentally sensitive caudate and putamen volume alterations in those with attention-deficit/hyperactivity disorder and their unaffected siblings. *JAMA Psychiatry*, **72**, 490–499.
- Hayano,Y. *et al.* (2014) Netrin-4 regulates thalamocortical axon branching in an activity-dependent fashion. *Proc. Natl. Acad. Sci. U. S. A.*, **111**, 15226–15231.
- ter Heegde,F. *et al.* (2015) The brain mineralocorticoid receptor and stress resilience. *Psychoneuroendocrinology*, **52**, 92–110.

- Hoogman, M. *et al.* (2019) Brain Imaging of the Cortex in ADHD: A Coordinated Analysis of Large-Scale Clinical and Population-Based Samples. *Am. J. Psychiatry*, **176**, 531–542.
- Hoogman, M. *et al.* (2012) Current self-reported symptoms of attention Deficit/Hyperactivity Disorder are associated with total brain volume in healthy adults. *PLoS One*, **7**.
- Hoogman, M. *et al.* (2017) Subcortical brain volume differences in participants with attention deficit hyperactivity disorder in children and adults: a cross-sectional mega-analysis. *The Lancet Psychiatry*, **4**, 310–319.
- Howard, D.M. *et al.* (2018) Genome-wide association study of depression phenotypes in UK Biobank identifies variants in excitatory synaptic pathways. *Nat. Commun.*, **9**, 1–10.
- Howie, B. *et al.* (2012) Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat. Genet.*, **44**, 955–959.
- Hussman, J.P. *et al.* (2011) A noise-reduction GWAS analysis implicates altered regulation of neurite outgrowth and guidance in autism. *Mol. Autism*, **2**, 1.
- Jansen, P.R. *et al.* (2019) Genome-wide analysis of insomnia in 1,331,010 individuals identifies new risk loci and functional pathways. *Nat. Genet.*, **51**, 394–403.
- Jensen, P.S. *et al.* (1997) Comorbidity in ADHD: Implications for research, practice, and DSM-V. *J. Am. Acad. Child Adolesc. Psychiatry*, **36**, 1065–1079.
- Johansson, S. *et al.* (2010) Common variants in the TPH1 and TPH2 regions are not associated with persistent ADHD in a combined sample of 1,636 adult cases and 1,923 controls from four European populations. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.*, **9999B**, n/a-n/a.
- Katane, M. *et al.* (2015) Characterization of the enzymatic and structural properties of human D-aspartate oxidase and comparison with those of the rat and mouse enzymes. *Biol. Pharm. Bull.*, **38**, 298–305.
- Katzman, M.A. *et al.* (2017) Adult ADHD and comorbid disorders: Clinical implications of a dimensional approach. *BMC Psychiatry*, **17**.
- Klein, M. *et al.* (2017) Brain imaging genetics in ADHD and beyond – Mapping pathways from gene to disorder at different levels of complexity. *Neurosci. Biobehav. Rev.*, **80**, 115–155.
- Klein, M. *et al.* (2019) Genetic markers of ADHD-related variations in intracranial volume. *Am. J. Psychiatry*, **176**, 228–238.
- Klok, M.D. *et al.* (2011) Decreased expression of mineralocorticoid receptor mRNA and its splice variants in postmortem brain regions of patients with major depressive disorder. *J. Psychiatr. Res.*, **45**, 871–878.
- Kooij, J.J.S. *et al.* (2012) Distinguishing Comorbidity and Successful Management of Adult ADHD. *J. Atten. Disord.*, **16**.
- Kortmann, G.L. *et al.* (2013) The role of a mineralocorticoid receptor gene functional polymorphism in the symptom dimensions of persistent ADHD. *Eur. Arch. Psychiatry Clin. Neurosci.*, **263**, 181–188.
- Larrivee, C.L. *et al.* (2020) Mice with GNAO1 R209H movement disorder variant display hyperlocomotion alleviated by risperidone. *J. Pharmacol. Exp. Ther.*, **373**, 24–33.
- Lee, J.J. *et al.* (2018) Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat. Genet.*, **50**, 1112–1121.

- Liu, J. *et al.* (2020) Gray matter networks associated with cognitive deficit in ADHD across adolescence and adulthood. *medRxiv*, 2020.04.22.20059808.
- Locke, A.E. *et al.* (2015) Genetic studies of body mass index yield new insights for obesity biology. *Nature*, **518**, 197–206.
- Luo, Y. *et al.* (2019) A review of heterogeneity in attention deficit/hyperactivity disorder (ADHD). *Front. Hum. Neurosci.*, **13**.
- Martin, J. *et al.* (2014) Genetic risk for attention-deficit/hyperactivity disorder contributes to neurodevelopmental traits in the general population. *Biol. Psychiatry*, **76**, 664–671.
- Medina, A. *et al.* (2013) Glucocorticoid and mineralocorticoid receptor expression in the human hippocampus in major depressive disorder. *J. Psychiatr. Res.*, **47**, 307–314.
- Van Der Meer, D. *et al.* (2017) Predicting attention-deficit/hyperactivity disorder severity from psychosocial stress and stress-response genes: A random forest regression approach. *Transl. Psychiatry*, **7**, e1145.
- Meinshausen, N. and Bühlmann, P. (2010) Stability selection. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, **72**, 417–473.
- Moua, P. *et al.* (2017) RELT family members activate p38 and induce apoptosis by a mechanism distinct from TNFR1. *Biochem. Biophys. Res. Commun.*, **491**, 25–32.
- Munni, A. *et al.* (2001) Expression, cellular distribution and protein binding of the glioma amplified sequence (gas41), a highly conserved putative transcription factor. *Oncogene*, **20**, 4853–4863.
- Neale, B.M. *et al.* (2010) Meta-analysis of genome-wide association studies of attention-deficit/hyperactivity disorder. *J. Am. Acad. Child Adolesc. Psychiatry*, **49**, 884–897.
- Ohnishi, T. *et al.* (2019) Psychiatric Comorbidities in Adult Attention-deficit/Hyperactivity Disorder: Prevalence and Patterns in the Routine Clinical Setting. *Innov. Clin. Neurosci.*, **16**, 11–16.
- Ottenshof, K.W. *et al.* (2018) TPH2 polymorphisms across the spectrum of psychiatric morbidity: A systematic review and meta-analysis. *Neurosci. Biobehav. Rev.*, **92**, 29–42.
- Park, J.H. and Roeder, R.G. (2006) GAS41 Is Required for Repression of the p53 Tumor Suppressor Pathway during Normal Cellular Proliferation. *Mol. Cell. Biol.*, **26**, 4006–4016.
- Polanczyk, G. *et al.* (2007) The worldwide prevalence of ADHD: A systematic review and meta-regression analysis. *Am. J. Psychiatry*, **164**, 942–948.
- Polanczyk, G. V. *et al.* (2014) ADHD prevalence estimates across three decades: An updated systematic review and meta-regression analysis. *Int. J. Epidemiol.*, **43**, 434–442.
- Purcell, S.M. *et al.* (2009) Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature*, **460**, 748.
- Qi, X.R. *et al.* (2013) Aberrant stress hormone receptor balance in the human prefrontal cortex and hypothalamic paraventricular nucleus of depressed patients. *Psychoneuroendocrinology*, **38**, 863–870.
- Reul, J.M.H.M. *et al.* (1994) Hypothalamic-Pituitary-Adrenocortical Axis Changes in the Rat after Long-Term Treatment with the Reversible Monoamine Oxidase-A Inhibitor Moclobemide. *Neuroendocrinology*, **60**, 509–519.
- von Rhein, D. *et al.* (2015) The NeuroIMAGE study: a prospective phenotypic, cognitive, genetic and MRI study in children with attention-deficit/hyperactivity disorder. Design and descriptives. *Eur.*

Child Adolesc. Psychiatry, **24**, 265–281.

- Rovira, P. *et al.* (2020) Shared genetic background between children and adults with attention deficit/hyperactivity disorder. *Neuropsychopharmacology*, 1–10.
- Saitsu, H. *et al.* (2016) Phenotypic spectrum of GNAO1 variants: Epileptic encephalopathy to involuntary movements with severe developmental delay. *Eur. J. Hum. Genet.*, **24**, 129–134.
- Sasaki, S. *et al.* (2008) The cortical subventricular zone-specific molecule Svet1 is part of the nuclear RNA coded by the putative Netrin receptor gene Unc5d and is expressed in multipolar migrating cells. *Mol. Cell. Neurosci.*, **38**, 474–483.
- Schierding, W. *et al.* (2014) The missing story behind Genome Wide Association Studies: Single nucleotide polymorphisms in gene deserts have a story to tell. *Front. Genet.*, **5**.
- Schoenmacker, G. *et al.* Quantitative Genetic Scoring, or how to put a number on an arbitrary genetic region.
- Seckl, J.R. and Fink, G. (1992) Antidepressants increase glucocorticoid and mineralocorticoid receptor mRNA expression in rat hippocampus in vivo. *Neuroendocrinology*, **55**, 621–626.
- Shaw, P. *et al.* (2011) Cortical development in typically developing children with symptoms of hyperactivity and impulsivity: Support for a dimensional view of attention deficit hyperactivity disorder. *Am. J. Psychiatry*, **168**, 143–151.
- Sheehan, K. *et al.* (2007) No association between TPH2 gene polymorphisms and ADHD in a UK sample. *Neurosci. Lett.*, **412**, 105–107.
- Sheehan, K. *et al.* (2005) Tryptophan hydroxylase 2 (TPH2) gene variants associated with ADHD. *Mol. Psychiatry*, **10**, 944–949.
- Simon, V. *et al.* (2009) Prevalence and correlates of adult attention-deficit hyperactivity disorder: Meta-analysis. *Br. J. Psychiatry*, **194**, 204–211.
- Srikanth, P. *et al.* (2018) Convergence of independent DISC1 mutations on impaired neurite growth via decreased UNC5D expression. *Transl. Psychiatry*, **8**, 1–16.
- Stamati, P. *et al.* (2019) Does SCFD1 rs10139154 Polymorphism Decrease Alzheimer’s Disease Risk? *J. Mol. Neurosci.*, **69**, 343–350.
- Stergiakouli, E. *et al.* (2017) Shared genetic influences between dimensional ASD and ADHD symptoms during child and adolescent development. *Mol. Autism*, **8**, 18.
- Thomas, R. *et al.* (2015) Prevalence of attention-deficit/hyperactivity disorder: A systematic review and meta-analysis. *Pediatrics*, **135**, e994–e1001.
- Thompson, P.M. *et al.* (2014) The ENIGMA Consortium: Large-scale collaborative analyses of neuroimaging and genetic data. *Brain Imaging Behav.*, **8**, 153–182.
- Uhlén, M. *et al.* (2015) Tissue-based map of the human proteome. *Science (80-.)*, **347**.
- Visscher, P.M. *et al.* (2014) Statistical Power to Detect Genetic (Co)Variance of Complex Traits Using SNP Data in Unrelated Samples. *PLoS Genet.*, **10**, e1004269.
- Waider, J. *et al.* (2011) Tryptophan hydroxylase-2 (TPH2) in disorders of cognitive control and emotion regulation: A perspective. *Psychoneuroendocrinology*, **36**, 393–405.
- Walker, S. and Scherer, S.W. (2013) Identification of candidate intergenic risk loci in autism spectrum disorder. *BMC Genomics*, **14**, 499.

- Wechsler,D. *et al.* (2000) WAIS-III: Nederlandstalige bewerking Swets Test Publishers; Harcourt Test Publishers.
- Xavier Castellanos,F. *et al.* (2002) Developmental trajectories of brain volume abnormalities in children and adolescents with attention-deficit/hyperactivity disorder. *J. Am. Med. Assoc.*, **288**, 1740–1748.
- Xing,G.-Q. *et al.* (2004) Decreased expression of mineralocorticoid receptor mRNA in the prefrontal cortex in schizophrenia and bipolar disorder. *Int. J. Neuropsychopharmacol.*, **7**, 143–153.
- Xu,L. *et al.* (2009) Source-based morphometry: The use of independent component analysis to identify gray matter differences with application to schizophrenia. *Hum. Brain Mapp.*, **30**, 711–724.
- Yamagata,H. *et al.* (2017) Identification of commonly altered genes between in major depressive disorder and a mouse model of depression. *Sci. Rep.*, **7**.
- Yau,J.L.W. *et al.* (1995) Glucocorticoids, hippocampal corticosteroid receptor gene expression and antidepressant treatment: relationship with spatial learning in young and aged rats. *Neuroscience*, **66**, 571–581.
- Young,E.A. *et al.* (1998) The Role of Mineralocorticoid Receptors in Hypothalamic-Pituitary-Adrenal Axis Regulation in Humans. *J. Clin. Endocrinol. Metab.*, **83**, 3339–3345.
- Zhang,F. and Lupski,J.R. (2015) Non-coding genetic variants in human disease. *Hum. Mol. Genet.*, **24**, R102–R110.