

ARTICLE

Variants in *TTC25* affect autistic trait in patients with autism spectrum disorder and general population

Dina Vojinovic¹, Nathalie Brison², Shahzad Ahmad¹, Ilse Noens^{3,4}, Irene Pappa^{5,6}, Lennart C Karssen^{1,7}, Henning Tiemeier^{1,8}, Cornelia M van Duijn^{1,9}, Hilde Peeters^{2,3,10} and Najaf Amin^{*,1,10}

Autism spectrum disorder (ASD) is a highly heritable neurodevelopmental disorder with a complex genetic architecture. To identify genetic variants underlying ASD, we performed single-variant and gene-based genome-wide association studies using a dense genotyping array containing over 2.3 million single-nucleotide variants in a discovery sample of 160 families with at least one child affected with non-syndromic ASD using a binary (ASD yes/no) phenotype and a quantitative autistic trait. Replication of the top findings was performed in Psychiatric Genomics Consortium and Erasmus Rucphen Family (ERF) cohort study. Significant association of quantitative autistic trait was observed with the *TTC25* gene at 17q21.2 (effect size = 10.2, P -value = 3.4×10^{-7}) in the gene-based analysis. The gene also showed nominally significant association in the cohort-based ERF study (effect = 1.75, P -value = 0.05). Meta-analysis of discovery and replication improved the association signal (P -value_{meta} = 1.5×10^{-8}). No genome-wide significant signal was observed in the single-variant analysis of either the binary ASD phenotype or the quantitative autistic trait. Our study has identified a novel gene *TTC25* to be associated with quantitative autistic trait in patients with ASD. The replication of association in a cohort-based study and the effect estimate suggest that variants in *TTC25* may also be relevant for broader ASD phenotype in the general population. *TTC25* is overexpressed in frontal cortex and testis and is known to be involved in cilium movement and thus an interesting candidate gene for autistic trait. *European Journal of Human Genetics* (2017) 25, 982–987; doi:10.1038/ejhg.2017.82; published online 17 May 2017

INTRODUCTION

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by deficits in social communication and social interaction and restricted and repetitive patterns of activities and behaviour with an onset in early development.¹ However ASD is a psychiatric diagnosis based on clinical criteria, and the severity of these characteristics can be measured as quantitative traits that represent a continuum that extends into the general population, with ASD at the extreme end of the distribution.² Other associated, but not core features are intellectual disability, attention-deficit disorder and medical comorbidities.³ The prevalence of ASD is estimated to be 62/10 000⁴ with boys-to-girls ratio of ~4:1.⁵ The importance of a genetic aetiology is established with heritability estimates ranging from 37 to 90%.^{6–9} Despite genetic heterogeneity, considerable progress in understanding the genetic architecture of ASD has been made by identifying monogenetic causes through genetic syndromes,¹⁰ rare chromosomal abnormalities,^{11,12} rare copy-number variants^{13–16} and rare penetrant gene mutations.³ Several genomic regions, including 2q, 3q25–27, 3p25, 6q14–21, 7q31–36 and 17q11–21¹⁷ have been linked to ASD. The role of rare genetic variants in the aetiology of ASD has been established by high-throughput technologies.^{18,19} More recently, the theory of excess of *de novo* loss-of-function variants in ASD patients has gained popularity after some initial successes.^{18–20} Around

1000 genes have been identified to be enriched with *de novo* loss-of-function mutations in ASD patients.²¹ However, *de novo* genetic variants do not contribute to the estimated heritability as these are not inherited. On the other hand, most genetic variance in ASD is attributed to common genetic variants.^{9,22} Their role has been demonstrated by several genome-wide association studies (GWAS)^{23–29} (Supplementary Table S1). Even though not many common susceptibility variants have been identified, significant association has been reported at 5p14.1,²³ at 5p15.31 between *SEMA5A* and *TAS2R1* genes,²⁸ within *MACROD2* at 20p12.1²⁶ and at 1p13.2.²⁹ However, there is a significant overlap of the discovery samples used and little replication of specific loci between studies.³⁰

Although the individual effect of common variants is modest, their joint effect may be substantial.²⁵ In this study besides assessing the effect of single variants on ASD, we evaluated the joint effect of multiple single variants in a gene in a genome-wide gene-based association analysis in patients with ASD from a Belgian Flemish cohort who were genotyped on a dense genotyping array.

MATERIALS AND METHODS

General overview of the study design and work flow are illustrated in Supplementary Figure S1.

¹Department of Epidemiology, Erasmus University Medical Center, Rotterdam, The Netherlands; ²Center for Human Genetics, University Hospitals Leuven, KU Leuven, Leuven, Belgium; ³Leuven Autism Research (LAuRes), Leuven, Belgium; ⁴Parenting and Special Education Research Unit, KU Leuven, Leuven, Belgium; ⁵School of Pedagogical and Educational Sciences, Erasmus University Rotterdam, Rotterdam, The Netherlands; ⁶Generation R Study Group, Erasmus University Medical Center, Rotterdam, The Netherlands; ⁷PolyOmica, s-Hertogenbosch, The Netherlands; ⁸Department of Child and Adolescent Psychiatry/Psychology, Erasmus University Medical Center-Sophia Children's Hospital, The Netherlands; ⁹Leiden Academic Centre for Drug Research (LACDR), Leiden University, The Netherlands

*Correspondence: Dr N Amin, Department of Epidemiology, Erasmus University Medical Center, PO Box 2040, Rotterdam 3000 CA, The Netherlands. Tel +31 10 7038125; Fax: +31 10 704 46 57; E-mail: n.amin@erasmusmc.nl

¹⁰These authors contributed equally to this work.

Received 29 August 2016; revised 3 April 2017; accepted 13 April 2017; published online 17 May 2017

Discovery sample

The discovery sample consisted of 160 nuclear Belgian Flemish families (657 individuals; Supplementary Table S2). The families were recruited to participate in the prospective study through the Expert Center for Autism (ECA) Leuven. All probands had been seen multiple times as part of their clinical care program in the ECA before recruitment. The families were asked to participate if there was at least one child with the diagnosis of non-syndromic ASD of unknown origin after a clinical genetics workup. Out of the 160 families, 55 were multiplex (two or more siblings with ASD) and 105 simplex. In six families the father had also been diagnosed with ASD such that there were 77.7% affected males and 22.3% affected females with male-to-female ratio of 3.5:1. Among them, 88.4% had normal and high intelligence, whereas 11.6% had mild, moderate or severe intellectual disability.

Diagnoses of ASD were made by a multidisciplinary team in the ECA Leuven according to DSM-IV-TR (American Psychiatric Association, 2000) criteria. Additionally, participants were assessed for quantitative autistic trait using the Dutch version of the Social Responsiveness Scale (SRS) and the Social Responsiveness Scale for Adults (SRS-A) designed to measure social impairment associated with ASD across a wide range of severity.^{31,32} Completed questionnaires were obtained for 490 probands, parents and siblings. Among the affected patients that had the SRS score available, the majority had normal and high intelligence (86%). For all participants, we received written informed consent. This study was approved by the Medical Ethical Committee of the University Hospitals Leuven.

Genotyping

Genotyping of 657 individuals from the discovery cohort was performed at the Center for Human Genetics at the KU Leuven, Belgium using the HumanOmni2.5-8 BeadChip, which contains more than 2.3 million common and less-frequent single-nucleotide polymorphisms (SNPs) from the 1000 Genome Project (minor allele frequency >2.5%). SNP calling was performed in Genome Studio 2011.1 using the genotyping module v1.9. Markers with call rate <95%, or which were monomorphic or which failed an exact test of Hardy-Weinberg equilibrium (HWE) (P -value < 1×10^{-7}) were removed from the analysis. Samples with low call rate <95% or high identity-by-state ($\geq 95\%$) were also removed from the analysis. Ethnic outliers were determined using multidimensional scaling analysis with 1000 Genomes data set (Supplementary Figure S2). All samples clustered tightly with the Europeans and no ethnic outlier was identified. In total 1 646 898 markers and 654 genotyped individuals were retained for further statistical analysis. Lastly, Mega2 tool v4.4^(ref. 33) was used to identify Mendelian inconsistencies, which were later set to missing.

Statistical analysis

Baseline descriptive analysis was performed with SPSS v21 (IBM Corporation, Armonk, NY, USA) and PEDSTATS v0.6.12.³⁴ Genome-wide association analyses of the binary ASD phenotype were performed through joint modelling of linkage and association, using the LAMP software v0.0.9 (School of Public Health, Ann Arbor, MI, USA). LAMP uses a maximum likelihood model to extract information on genetic linkage and association from samples of unrelated individuals, sib pairs, trios and larger pedigrees in settings where population stratification is not a concern (Supplementary Figure S2).³⁵ Odds ratios and 95% confidence intervals were estimated using PLINK v1.07.³⁶ The association tests for markers on sex chromosomes were performed by transmission disequilibrium test for chromosome X and by logistic regression for chromosome Y. Single-variant and gene-based genome-wide association analyses of the quantitative autistic trait adjusted for age, gender and familial relationships were performed using the RVtests software tool version 20150630 (<http://zhanxw.github.io/rvtests/>). The gene-based analysis included Combined Multivariate and Collapsing method which is robust and powerful in the presence of wide spectrum of variant allele frequencies.³⁷ The genes were defined according to human reference genome hg19. All association analyses were performed for entire discovery sample, and simplex and multiplex families separately. The standard genome-wide significance threshold of 5×10^{-8} was used to declare significance in the single-variant analyses, while the genome-wide significance threshold for the gene-based analysis was set at 2.5×10^{-6}

based on 19 650 genes tested. PLINK/SEQ v0.10 (<https://atgu.mgh.harvard.edu/plinkseq/>) was used to convert PLINK files into variant call format files. All genome maps were updated to human genome build 19 (hg19). Gene pathway enrichment analysis of all nominally significant genes (P -value < 0.01) in the gene-based analysis was performed using the web-based gene network pathway enrichment tool (<http://129.125.135.180:8080/GeneNetwork/pathway.html>).

The data were deposited in the GWAS Central database (<http://www.gwascentral.org/study/HGVST1847>).

Bioinformatic analysis

To annotate SNPs with regulatory information, we used RegulomeDB v1.1 database (<http://www.regulomedb.org/index>) that combines information from ENCODE and other sources, as well as computational predictions and manual annotations into a tool that classifies SNPs into six categories, where Category 1 variants are 'likely to affect binding and linked to expression of a gene target', whereas category 6 variants have 'minimal binding evidence'.³⁸ Furthermore, regulatory information on SNPs in haplotype blocks was explored using a HaploReg v4.1 tool.³⁹ For these analysis r^2 was set to 1 and the population of European descent was chosen.

Replication samples

Psychiatric Genomics Consortium (PGC). A lookup of top findings from the single-variant analyses of the binary ASD phenotype and quantitative autistic trait was performed in the latest PGC GWAS. This data set consists of a total of 6495 parent-child trios who met diagnostic criteria for ASD and had genome-wide SNP data available (<https://www.med.unc.edu/pgc>).

Erasmus Rucphen Family (ERF) study. Replication of the gene-based analysis of quantitative autistic trait was performed in the ERF study as 1250 participants from this cohort have been assessed for quantitative autistic trait using Baron-Cohen's Autism-Spectrum Quotient (AQ) test⁴⁰ and exomes of half of these participants ($n=615$) have been sequenced, thus providing a greater resolution at the gene level. Individuals whose exome were sequenced were selected based on having good quality phenotype information on a wide range of topics, and therefore random with regards to AQ scores (Supplementary Table S2). ERF is a family-based cohort originating from 22 couples and spread over 23 generations.^{41,42} The ERF study was approved by the Medical Ethics Committee of the Erasmus MC which is constituted according to the WMO (Wet Medisch-wetenschappelijk Onderzoek met mensen). A written informed consent was obtained from all study participants.

Sequencing was done at a mean depth of $74 \times$ using the Nimblegen SeqCap EZ V2 capture kit on an Illumina HiSeq2000 sequencer (Illumina, San Diego, CA, USA) using the TruSeq Version 3 protocol at the Human Genotyping facility of the Internal Medicine department, at the Erasmus MC, The Netherlands.^{43,44} The sequence reads were aligned to the human genome build 19 (hg19), using Burrows-Wheeler Aligner and the NARWHAL pipeline.^{45,46} After processing, genetic variants were called, using the Unified Genotyper tool from the GATK.⁴⁷ Variants with a low quality (QUAL < 150), which were out of HWE (P -value < 10^{-6}) or with low call rate (< 90%), as well as samples with a low call rate (< 90%), and duplicates, were removed.⁴⁴ Functional annotations were also performed using the SeattleSeq annotation 138 database (<http://snp.gs.washington.edu/SeattleSeqAnnotation138/>).

Association analyses of the quantitative autistic trait adjusted for age, gender and familial relationship were performed using the RVtests software. Meta-analysis of gene-based results of discovery sample and ERF study was performed using Fisher's combined probability test.

RESULTS

Results of genome-wide association analysis for the binary ASD phenotype and quantitative autistic trait are illustrated in Supplementary Figures S3 and S4. No single-variant surpassed the genome-wide significance threshold. Top findings from the association analyses are shown in Tables 1 and 2 for the discovery sample, and in Supplementary Tables S3 and S4 for simplex and multiplex families

Table 1 Top findings (P -value $< 1 \times 10^{-5}$) from the binary ASD phenotype association analysis

SNP	Chr	Position ^a	Function	CADD	Gene	Reference allele	Variant allele	N	Call rate	Discovery cohort			PGC		
										MAF	HWE	LOD	P	OR	P
rs6452310	5	25996104	Intergenic	0.58	MSNP1AS	G	A	654	1	0.32	0.9	6.27	7.80×10^{-8}	0.996	0.9
rs72996425	18	59451206	Intergenic	5.32	RFN125/CHD20	A	C	654	1	0.03	1	4.37	7.20×10^{-6}	0.96	0.7
rs7147213	14	52932731	Intron	2.13	TXNDC16	A	G	653	0.998	0.12	1	4.32	8.20×10^{-6}	1.03	0.34
rs7700465	5	26003284	Intergenic	3.03	MSNP1AS	G	A	654	1	0.28	0.89	4.29	8.70×10^{-6}	0.99	0.69
rs7142973	14	52926750	Intron	1.02	TXNDC16	A	G	654	1	0.12	1	4.26	9.50×10^{-6}	1.04	0.36
rs4901281	14	52980190	Intron	7.66	TXNDC16	C	A	654	1	0.12	1	4.26	9.50×10^{-6}	0.97	0.39
rs12435289	14	52981240	Intron	3.54	TXNDC16	C	A	654	1	0.12	1	4.26	9.50×10^{-6}	0.97	0.37

Abbreviations: ASD, autism spectrum disorder; Chr, chromosome; CADD, phred-like Combined Annotation Dependent Depletion scores from Kircher et al,⁷⁰ University of Washington; HWE, Hardy-Weinberg equilibrium in founders; MAF, minor allele frequency; N, number of individuals; OR, odds ratio; PGC, Psychiatric Genomics Consortium.
^aGenomic position is according the hg19 assembly.

separately. Suggestive association of binary ASD phenotype was observed with two common variants (rs6452310; P -value = 7.80×10^{-8} and rs7700465; P -value = 8.70×10^{-6} ; Table 1) at chromosome 5p14.1—a region previously known to be associated with ASD (Figure 1). The two variants were in strong linkage disequilibrium (LD) ($r^2 = 0.85$) with each other but not in LD with any of the previously identified variants in this region associated with ASD^{23,24} (r^2 ranged from 0.002 to 0.009, D' ranged from 0.02 to 0.33). None of the top variants from this analysis showed evidence of association in the replication sample (Table 1).

Results of gene-based genome-wide association analysis are illustrated in Supplementary Figure S5 for the discovery sample, and in Supplementary Table S5 for simplex and multiplex families separately. The gene-based association analysis revealed significant association of quantitative autistic trait and *TTC25* gene (P -value = 3.4×10^{-7}) on chromosome 17 (Table 3). This association was not driven by any single variant but by nine variants, four of which showed nominally significant association with quantitative autistic trait (P -value < 0.05) (Supplementary Table S6). The combined effect of these variants on the SRS score was large (effect size = 10.2). The functional annotation of nine variants revealed that five variants are likely to have regulatory functions (Category 1 RegulomeDB score; Supplementary Table S6). Furthermore, the surrounding variants in strong LD ($r^2 = 1$) with nine variants lie in enhancer histone marks, and protein-binding regions and change regulatory motifs based on the variant allele changes (Supplementary Table S7). The gene was also nominally associated with autistic trait in the ERF cohort (P -value = 0.045). The combined effect of 15 variants in the *TTC25* gene (Supplementary Table S8) in the replication sample was much smaller (effect size = 1.75). Meta-analysis of discovery and replication samples resulted in an improved association signal (P -value = 1.5×10^{-8}).

Pathway analysis

Pathway enrichment analysis based on nominally significant genes in the gene-based association analysis (Supplementary Table S9) showed significantly enriched SMAD protein signal transduction (P -value = 2×10^{-5}) pathway and series of digestive system development enrichment categories (Supplementary Table S10).

DISCUSSION

In this study, we have identified a novel gene *TTC25* associated with quantitative autistic trait. The association of the gene *TTC25* in a cohort-based study suggests that the gene may be relevant for broader ASD phenotype in the general population. Further, we identified

SMAD protein signal transduction pathway and series of digestive system development categories as being significantly enriched with genes nominally associated with quantitative autistic trait. Moreover, our study provides additional evidence for the previously identified association of the intergenic loci at 5p14.1 with the binary ASD phenotype.

TTC25 gene is located on chromosome 17q21.2 and encodes Tetratricopeptide Repeat Domain 25. 17q21.2 locus has previously been linked to ASD in a genome-wide linkage scan,^{48,49} although a gene was never implicated. *TTC25* is overexpressed in testis, frontal cortex and rectum (<http://www.genecards.org/cgi-bin/carddisp.pl?gene=TTC25>). *TTC25* is involved in cilium movement, organization and morphogenesis.⁵⁰ Cilia are specialized organelles protruding from the cell surface of almost all mammalian cells.⁵¹ Mutations in ciliary proteins cause ciliopathies which can affect many organs at different levels of severity and are characterized by a wide spectrum of phenotypes.⁵¹ In the vertebrate nervous system, the primary cilium is increasingly viewed as hub for certain neural developmental signalling pathways, and growing data suggest this is also true for several types of adult neuronal signalling.⁵² The capacity of the brain to interpret the sensory input is often affected in ciliopathies, resulting in neurological disorders; cognitive impairment, anosmia, intellectual disability, ASD and obesity are apparent in various degrees in many of the ciliopathies.^{51,53} Further Joubert syndrome (JS) is a well-known ciliopathy of the central nervous system.^{52,54} Features of ASD, such as problems in social behaviour, communication problems and repetitive behaviours, have been described in up to 40% of JS patients^{55–57} and about 25% of JS patients meet criteria for the DSM-IV diagnosis of ASD.^{55,58} Multiple variants mapped to this gene in our sample appear to have a regulatory function.

Our identification of SMAD protein signal transduction pathway as being significantly enriched with genes nominally associated with quantitative autistic trait reinforces the role of the transforming growth factor- β (TGF β) in ASD. The Smad pathways are the major mediators of transcriptional responses induced by the TGF β family, which control cell-fate determination, cell cycle arrest, apoptosis and actin rearrangements.⁵⁹ While decreased levels of TGF β have been reported in blood samples from individuals with ASD⁶⁰ and associated with more severe behavioural scores in ASD children,⁶¹ higher levels of TGF β have been reported in postmortem brain and cerebrospinal fluid samples of ASD patients.⁶² In addition, series of digestive system development categories were enriched with genes nominally associated with ASD. Gastrointestinal (GI) disturbances are 4-fold more common in ASD⁶³ and available scientific evidence supports

Table 2 Top findings (P -value $< 1 \times 10^{-5}$) from the quantitative autistic trait association analysis

SNP	Chr	Position ^a	Function	CADD	Gene	Reference allele	Variant Allele	N	AF	Discovery cohort			PGC				
										Call Rate	HWE	U_Stat	Sqrt_V_Stat	Effect	P	OR	P
rs77782810	14	93169444	Downstream	0.81	None	T	G	486	0.09	0.99	0.35	2.11	0.42	12.13	4.28×10^{-7}	1.03	0.54
rs75822891	12	551394	3-prime-UJR	7.37	CCDC77	A	C	486	0.02	1	1	1.11	0.23	21.81	8.78×10^{-7}	0.89	0.12
rs61867303	10	106594496	Intron	3.92	SORCS3	G	A	486	0.04	1	0.03	1.44	0.30	16.34	1.19×10^{-6}	1.03	0.70
rs202183827	14	93162861	Intergenic	1.14	None	G	A	486	0.10	0.99	0.12	2.11	0.45	10.25	3.30×10^{-6}	—	—
rs75820392	12	526569	Intron	1.64	CCDC77	C	T	486	0.02	0.99	1	0.98	0.21	21.86	3.61×10^{-6}	1.08	0.44
rs76459666	14	93161327	Intergenic	4.74	None	C	T	486	0.1	0.99	0.13	2.11	0.46	10.11	3.88×10^{-6}	1.01	0.85
rs12433198	14	78042307	Intron	2.51	SPTLC2	G	A	486	0.08	1	0.52	1.81	0.40	11.22	6.49×10^{-6}	1.13	0.01
rs115601680	6	46378057	Intron	0.12	RCAN2	C	T	486	0.01	1	1	0.60	0.13	32.90	8.99×10^{-6}	—	—
rs77988177	20	21972232	Intergenic	2.47	None	G	A	486	0.03	1	0.27	1.06	0.24	18.52	9.18×10^{-6}	0.94	0.52

Abbreviations: AF, allele frequency; Chr, chromosome; CADD, phred-like Combined Annotation Dependent Depletion scores from Kircher et al;⁷⁰ University of Washington; Effect, effect of variant allele; N, number of individuals; HWE, Hardy-Weinberg equilibrium in founders; OR, odds ratio; PGC, Psychiatric Genomics Consortium; SNP, single-nucleotide polymorphism.

U and V statistics are score statistics and their covariance matrix.

^aGenomic position is according the hg19 assembly.

combination of changes in the areas of immune function, gut microbiome and gut and brain signalling pathways.⁶⁴ Recent studies in animal models suggests that GI difficulties may originate from the same genetic changes that lead to the behavioural characteristics of ASD.⁶⁵

In addition, we found the association of ASD with the known region on 5p14.1. This is one of the few replicated GWA regions that implicates long noncoding RNA gene, *MSNPIAS* (moesin pseudogene 1, antisense) in ASD risk.⁶⁶ This region has also been associated with social communication spectrum phenotypes in the general population supporting the role of 5p14.1 as a quantitative trait locus for ASD.⁴⁹ *MSNPIAS* shows a very high sequence homology to the chromosome X transcript of *MSN* that is involved in brain development.⁶⁶ *MSNPIAS* is highly overexpressed (12.7-fold) in the postmortem cerebral cortex of individuals with ASD.⁶⁶ Interestingly, our top hit did not replicate in the PGC sample and vice versa. The multiple different variants discovered in the 5p14.1 region^{23,24} may suggest that multiple alleles in the same region are implicated in ASD.

Although our study sample was small, the strength of our study is that we used a data set with high-quality phenotypes, in which participants were assessed for both binary ASD phenotype and a quantitative autistic trait. The use of quantitative endophenotype provides additional power to find genetic signals by focusing on less complex aspects of complex phenotypes such as ASD.⁶⁷ The identification of new loci associated with quantitative autistic trait in our study validates this approach. Another strength is phenotypic homogeneity in the sense that the majority of patients in the current study had a normal intelligence, unlike most ASD cohorts with typical rates of intellectual disability ranging from 30 to 50%.⁶⁸ Further, the genomes of our study participants were genotyped on a very dense SNP array that contains not only common but also less-frequent SNPs. This gave us the opportunity to make a comprehensive overview of how common and less-frequent variants, both individually and taken together, affect ASD. Our study shows the advantage of a gene-based test as the more powerful approach compared to single-variant analysis and demonstrated the use of gene-based pathway and enrichment analysis in understanding the molecular mechanisms of the disorder. One of the possible limitations of our study is that the assessment of quantitative autistic trait in the discovery and replication cohort used two different questionnaires, the SRS and the AQ. As the questionnaires have both been designed to measure the severity of social responsiveness problems across clinical cases and the general population and as their ratings are significantly correlated,⁶⁹ we were able to compare results from the two cohorts.

To conclude, our study has identified a novel gene *TTC25* to be associated with autistic trait in the ASD population where majority of patients have a normal intelligence. The replication of *TTC25* association in a cohort-based study suggests that this gene may also be relevant for broader ASD phenotype in the general population. However, whether these findings hold true also for ASD patients with intellectual disability remains to be evaluated. *TTC25* is overexpressed in frontal cortex and testis and is known to be involved in cilium movement and thus an interesting candidate gene for autistic trait. Furthermore, we discovered significantly enriched SMAD protein signal transduction pathway and series of digestive system development categories in the pathway analysis of quantitative autistic trait. Our finding provides new insights into the genetic background of quantitative autistic trait.

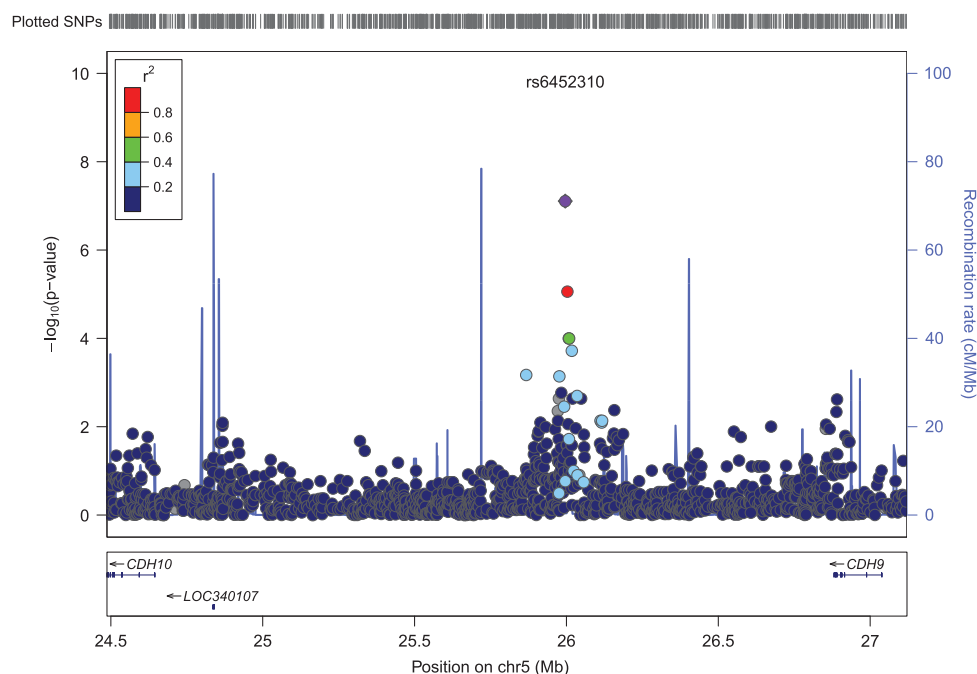


Figure 1 Regional association and recombination rate plot of the 5p14.1 region and the binary ASD phenotype in the discovery cohort. The left y axis represents $-\log_{10} P$ -values for association with binary ASD phenotype in the discovery cohort. The right y axis represents the recombination rate, and the x axis represents chromosomal position (genomic position is according to the hg19 assembly). The most significantly associated single SNP in this region (rs6452310) is denoted with a purple diamond. Surrounding SNPs are shaded according to their pairwise correlation (r^2) with rs6452310. The gene annotations are shown below the figure.

Table 3 Top results (P -value $< 1 \times 10^{-3}$) from the gene-based association analysis of quantitative autistic trait in the discovery cohort

Gene	Chr	N	NumVar	NumSite	AF ^a	Effect	P
TTC25	17	486	13	9	0.338	10.23	3.44×10^{-7}
SH2D5	1	486	9	5	0.440	-10.88	1.49×10^{-4}
PSMC5	17	486	4	3	0.458	-11.13	6.81×10^{-4}
LOC729177	6	486	22	14	0.434	-9.15	9.37×10^{-4}

Abbreviations: AF, allele frequency; Chr, chromosome; Effect, effect size; N, sample size; NumVar, number of variants; NumSite, number of sites.

^aAllele frequency is calculated as an average of individual allele frequency adjusted for relatedness (<http://zhanxw.github.io/rvtests/>).

CONFLICT OF INTEREST

IN receives royalties from Hogrefe Uitgeverij for the SRS-A (although not for the questionnaires used in this study). The remaining authors declare no conflict of interest.

ACKNOWLEDGEMENTS

The Erasmus Rucphen Family study (ERF) study as a part of EUROSPAN (European Special Populations Research Network) was supported by European Commission FP6 STRP grant number 018947 (LSHG-CT-2006-01947) and also received funding from the European Community's Seventh Framework Programme (FP7/2007-2013)/grant agreement HEALTH-F4-2007-201413 by the European Commission under the programme 'Quality of Life and Management of the Living Resources' of 5th Framework Programme (no. QL2-CT-2002-01254). High-throughput analysis of the ERF data was supported by joint grant from Netherlands Organization for Scientific Research and the Russian Foundation for Basic Research (NWO-RFBR 047.017.043). Exome sequencing analysis in ERF was supported by the ZonMw grant (project 91111025). We are grateful to all study participants and their relatives, general practitioners and neurologists for their contributions and to P Veraart for her

help in genealogy, J Vergeer for the supervision of the laboratory work and P Snijders for his help in the data collection. NA is supported by The Netherlands Brain Foundation (project number F2013(1)-28). HP is a Senior Clinical Investigator of The Research Foundation—Flanders (FWO). We acknowledge PGC consortium for providing the pre-publication results.

- 1 Association AP *Diagnostic and Statistical Manual of Mental Disorders*. 5th edn Arlington, VA: American Psychiatric Publishing, 2013.
- 2 Constantino JN, Todd RD: Autistic traits in the general population: a twin study. *Arch Gen Psychiatry* 2003; **60**: 524–530.
- 3 Devlin B, Scherer SW: Genetic architecture in autism spectrum disorder. *Curr Opin Genet Dev* 2012; **22**: 229–237.
- 4 Elsabbagh M, Divan G, Koh YJ et al: Global prevalence of autism and other pervasive developmental disorders. *Autism Res* 2012; **5**: 160–179.
- 5 Werling DM, Geschwind DH: Sex differences in autism spectrum disorders. *Curr Opin Neurol* 2013; **26**: 146–153.
- 6 Hallmayer J, Cleveland S, Torres A et al: Genetic heritability and shared environmental factors among twin pairs with autism. *Arch Gen Psychiatry* 2011; **68**: 1095–1102.
- 7 Bailey A, Le Couteur A, Gottesman I et al: Autism as a strongly genetic disorder: evidence from a British twin study. *Psychol Med* 1995; **25**: 63–77.
- 8 Sandin S, Lichtenstein P, Kuja-Halkola R, Larsson H, Hultman CM, Reichenberg A: The familial risk of autism. *JAMA* 2014; **311**: 1770–1777.
- 9 Gauglerl T, Klei L, Sanders SJ et al: Most genetic risk for autism resides with common variation. *Nat Genet* 2014; **46**: 881–885.
- 10 Betancur C: Etiological heterogeneity in autism spectrum disorders: more than 100 genetic and genomic disorders and still counting. *Brain Res* 2011; **1380**: 42–77.
- 11 Marshall CR, Noor A, Vincent JB et al: Structural variation of chromosomes in autism spectrum disorder. *Am J Hum Genet* 2008; **82**: 477–488.
- 12 Shen Y, Dies KA, Holm IA et al: Clinical genetic testing for patients with autism spectrum disorders. *Pediatrics* 2010; **125**: e727–e735.
- 13 Pinto D, Pagnamenta AT, Klei L et al: Functional impact of global rare copy number variation in autism spectrum disorders. *Nature* 2010; **466**: 368–372.
- 14 Gai X, Xie HM, Perin JC et al: Rare structural variation of synapse and neurotransmission genes in autism. *Mol Psychiatry* 2012; **17**: 402–411.
- 15 Sebat J, Lakshmi B, Malhotra D et al: Strong association of *de novo* copy number mutations with autism. *Science* 2007; **316**: 445–449.
- 16 Girirajan S, Brkanac Z, Coe BP et al: Relative burden of large CNVs on a range of neurodevelopmental phenotypes. *PLoS Genet* 2011; **7**: e1002334.

- 17 Freitag CM: The genetics of autistic disorders and its clinical relevance: a review of the literature. *Mol Psychiatry* 2007; **12**: 2–22.
- 18 Sanders SJ, Murtha MT, Gupta AR et al: *De novo* mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* 2012; **485**: 237–241.
- 19 O’Roak BJ, Vives L, Girirajan S et al: Sporadic autism exomes reveal a highly interconnected protein network of *de novo* mutations. *Nature* 2012; **485**: 246–250.
- 20 O’Roak BJ, Deriziotis P, Lee C et al: Exome sequencing in sporadic autism spectrum disorders identifies severe *de novo* mutations. *Nat Genet* 2011; **43**: 585–589.
- 21 Samocha KE, Robinson EB, Sanders SJ et al: A framework for the interpretation of *de novo* mutation in human disease. *Nat Genet* 2014; **46**: 944–950.
- 22 Cross-Disorder Group of the Psychiatric Genomics C: Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet* 2013; **381**: 1371–1379.
- 23 Wang K, Zhang H, Ma D et al: Common genetic variants on 5p14.1 associate with autism spectrum disorders. *Nature* 2009; **459**: 528–533.
- 24 Ma D, Salyakina D, Jaworski JM et al: A genome-wide association study of autism reveals a common novel risk locus at 5p14.1. *Ann Hum Genet* 2009; **73**: 263–273.
- 25 Anney R, Klei L, Pinto D et al: Individual common variants exert weak effects on the risk for autism spectrum disorders. *Hum Mol Genet* 2012; **21**: 4781–4792.
- 26 Anney R, Klei L, Pinto D et al: A genome-wide scan for common alleles affecting risk for autism. *Hum Mol Genet* 2010; **19**: 4072–4082.
- 27 Connolly JJ, Glessner JT, Hakonarson H: A genome-wide association study of autism incorporating autism diagnostic interview-revised, autism diagnostic observation schedule, and social responsiveness scale. *Child Dev* 2013; **84**: 17–33.
- 28 Weiss LA, Arking DE, Daly MJ, Chakravarti A: Gene Discovery Project of Johns H and the Autism C: A genome-wide linkage and association scan reveals novel loci for autism. *Nature* 2009; **461**: 802–808.
- 29 Xia K, Guo H, Hu Z et al: Common genetic variants on 1p13.2 associate with risk of autism. *Mol Psychiatry* 2014; **19**: 1212–1219.
- 30 Holt R, Monaco AP: Links between genetics and pathophysiology in the autism spectrum disorders. *EMBO Mol Med* 2011; **3**: 438–450.
- 31 Roeyers H, Thys M, Druart C, De Schryver M, Schittekatte M: *SRS Screeningslijst Voor Autismspectrumstoornissen*. Amsterdam: Hogrefe Uitgevers B.V., 2011.
- 32 Noens I, De la Marche W, Scholte E: *SRS Screeningslijst Voor Autismspectrumstoornissen bij Volwassenen*. Amsterdam: Hogrefe Uitgevers BV, 2012.
- 33 Mukhopadhyay N, Almasy L, Schroeder M, Mulvihill WP, Weeks DE: Mega2: data-handling for facilitating genetic linkage and association analyses. *Bioinformatics* 2005; **21**: 2556–2557.
- 34 Wigginton JE, Abecasis GR: PEDSTATS: descriptive statistics, graphics and quality assessment for gene mapping data. *Bioinformatics* 2005; **21**: 3445–3447.
- 35 Li M, Boehnke M, Abecasis GR: Joint modeling of linkage and association: identifying SNPs responsible for a linkage signal. *Am J Hum Genet* 2005; **76**: 934–949.
- 36 Purcell S, Neale B, Todd-Brown K et al: PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**: 559–575.
- 37 Li B, Leal SM: Methods for detecting associations with rare variants for common diseases: application to analysis of sequence data. *Am J Hum Genet* 2008; **83**: 311–321.
- 38 Boyle AP, Hong EL, Hariharan M et al: Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res* 2012; **22**: 1790–1797.
- 39 Ward LD, Kellis M: HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res* 2012; **40**: D930–D934.
- 40 Baron-Cohen S, Wheelwright S, Skinner R, Martin J, Clubley E: The autism-spectrum quotient (AQ): evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *J Autism Dev Disord* 2001; **31**: 5–17.
- 41 Aulchenko YS, Heutink P, Mackay I et al: Linkage disequilibrium in young genetically isolated Dutch population. *Eur J Hum Genet* 2004; **12**: 527–534.
- 42 Pardo LM, MacKay I, Oostra B, van Duijn CM, Aulchenko YS: The effect of genetic drift in a young genetically isolated population. *Ann Hum Genet* 2005; **69**: 288–295.
- 43 Vojinovic D, Adams HH, van der Lee SJ et al: The dystrophin gene and cognitive function in the general population. *Eur J Hum Genet* 2015; **23**: 837–843.
- 44 Amin N, Jovanova O, Adams HH et al: Exome-sequencing in a large population-based study reveals a rare Asn396Ser variant in the LIPG gene associated with depressive symptoms. *Mol Psychiatry* 2016; **22**: 634.
- 45 Brouwer RW, van den Hout MC, Grosveld FG, van Ijcken WF: NARWHAL, a primary analysis pipeline for NGS data. *Bioinformatics* 2012; **28**: 284–285.
- 46 Bucan M, Abrahams BS, Wang K et al: Genome-wide analyses of exonic copy number variants in a family-based study point to novel autism susceptibility genes. *PLoS Genet* 2009; **5**: e1000536.
- 47 McKenna A, Hanna M, Banks E et al: The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 2010; **20**: 1297–1303.
- 48 McCauley JL, Li C, Jiang L et al: Genome-wide and ordered-subset linkage analyses provide support for autism loci on 17q and 19p with evidence of phenotypic and interlocus genetic correlates. *BMC Med Genet* 2005; **6**: 1.
- 49 Lowe JK, Werling DM, Constantino JN, Cantor RM, Geschwind DH: Social responsiveness, an autism endophenotype: genome-wide significant linkage to two regions on chromosome 8. *Am J Psychiatry* 2015; **172**: 266–275.
- 50 Xu Y, Cao J, Huang S et al: Characterization of tetratricopeptide repeat-containing proteins critical for cilia formation and function. *PLoS One* 2015; **10**: e0124378.
- 51 D’Angelo A, Franco B: The dynamic cilium in human diseases. *Pathogenetics* 2009; **2**: 3.
- 52 Louvi A, Grove EA: Cilia in the CNS: the quiet organelle claims center stage. *Neuron* 2011; **69**: 1046–1060.
- 53 Basten SG, Giles RH: Functional aspects of primary cilia in signaling, cell cycle and tumorigenesis. *Cilia* 2013; **2**: 6.
- 54 Doherty D: Joubert syndrome: insights into brain development, cilium biology, and complex disease. *Semin Pediatr Neurol* 2009; **16**: 143–154.
- 55 Ozonoff S, Williams BJ, Gale S, Miller JN: Autism and autistic behavior in Joubert syndrome. *J Child Neurol* 1999; **14**: 636–641.
- 56 Artigas-Pallares J, Gabau-Vila E, Guitart-Feliubadaló M: (Syndromic autism: II. Genetic syndromes associated with autism) El autismo sindrómico: II. Síndromes de base genética asociados a autismo. *Rev Neurol* 2005; **40**: S151–S162.
- 57 Holroyd S, Reiss AL, Bryan RN: Autistic features in Joubert syndrome: a genetic disorder with agenesis of the cerebellar vermis. *Biol Psychiatry* 1991; **29**: 287–294.
- 58 Alvarez Retuerto AI, Cantor RM, Gleeson JG et al: Association of common variants in the Joubert syndrome gene (AH11) with autism. *Hum Mol Genet* 2008; **17**: 3887–3896.
- 59 Wrana JL: Signaling by the TGFbeta superfamily. *Cold Spring Harb Perspect Biol* 2013; **5**: a011197.
- 60 Okada K, Hashimoto K, Iwata Y et al: Decreased serum levels of transforming growth factor-beta1 in patients with autism. *Prog Neuropsychopharmacol Biol Psychiatry* 2007; **31**: 187–190.
- 61 Ashwood P, Enstrom A, Krakowiak P et al: Decreased transforming growth factor beta1 in autism: a potential link between immune dysregulation and impairment in clinical behavioral outcomes. *J Neuroimmunol* 2008; **204**: 149–153.
- 62 Vargas DL, Nascimbene C, Krishnan C, Zimmerman AW, Pardo CA: Neuroglial activation and neuroinflammation in the brain of patients with autism. *Ann Neurol* 2005; **57**: 67–81.
- 63 McElhanon BO, McCracken C, Karpen S, Sharp WG: Gastrointestinal symptoms in autism spectrum disorder: a meta-analysis. *Pediatrics* 2014; **133**: 872–883.
- 64 Coury DL, Ashwood P, Fasano A et al: Gastrointestinal conditions in children with autism spectrum disorder: developing a research agenda. *Pediatrics* 2012; **130**: S160–S168.
- 65 Margolis KG, Li Z, Stevanovic K et al: Serotonin transporter variant drives preventable gastrointestinal abnormalities in development and function. *J Clin Invest* 2016; **126**: 2221–2235.
- 66 Kerin T, Ramanathan A, Rivas K, Grepo N, Coetzee GA, Campbell DB: A noncoding RNA antisense to moesin at 5p14.1 in autism. *Sci Transl Med* 2012; **4**: 128ra140.
- 67 Almasy L: The role of phenotype in gene discovery in the whole genome sequencing era. *Hum Genet* 2012; **131**: 1533–1540.
- 68 Geschwind DH: Advances in autism. *Annu Rev Med* 2009; **60**: 367–380.
- 69 Armstrong K, Iarocci G: Brief report: the autism spectrum quotient has convergent validity with the social responsiveness scale in a high-functioning sample. *J Autism Dev Disord* 2013; **43**: 2228–2232.
- 70 Kircher M, Witten DM, Jain P, O’Roak BJ, Cooper GM, Shendure J: A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet* 2014; **46**: 310–315.

Supplementary Information accompanies this paper on European Journal of Human Genetics website (<http://www.nature.com/ejhg>)