

Published in final edited form as:

*Am J Psychiatry*. 2010 November ; 167(11): 1364–1372. doi:10.1176/appi.ajp.2010.09121789.

## Association Between a High-Risk Autism Locus on 5p14 and Social Communication Spectrum Phenotypes in the General Population

Beate St. Pourcain, Ph.D., Kai Wang, Ph.D., Joseph T. Glessner, M.S., Jean Golding, Ph.D., D.Sc., Colin Steer, M.Sc., Susan M. Ring, Ph.D., David H. Skuse, M.D., Struan F.A. Grant, Ph.D., Hakon Hakonarson, M.D., Ph.D., and George Davey Smith, M.D., D.Sc.

Medical Research Council Centre for Causal Analyses in Translational Epidemiology, University of Bristol; Department of Social Medicine, University of Bristol; Children's Hospital of Philadelphia, Philadelphia; Centre for Child and Adolescent Health, Department of Community Based Medicine, University of Bristol; and the University College London Institute of Child Health, London.

### Abstract

**Objective**—Recent genome-wide analysis identified a genetic variant on 5p14.1 (rs4307059), which is associated with risk for autism spectrum disorder. This study investigated whether rs4307059 also operates as a quantitative trait locus underlying a broader autism phenotype in the general population, focusing specifically on the social communication aspect of the spectrum.

**Method**—Study participants were 7,313 children from the Avon Longitudinal Study of Parents and Children. Single-trait and joint-trait genotype associations were investigated for 29 measures related to language and communication, verbal intelligence, social interaction, and behavioral adjustment, assessed between ages 3 and 12 years. Analyses were performed in one-sided or directed mode and adjusted for multiple testing, trait interrelatedness, and random genotype dropout.

**Results**—Single phenotype analyses showed that an increased load of rs4307059 risk allele is associated with stereotyped conversation and lower pragmatic communication skills, as measured by the Children's Communication Checklist (at a mean age of 9.7 years). In addition a trend toward a higher frequency of identification of special educational needs (at a mean age of 11.8 years) was observed. Variation at rs4307059 was also associated with the phenotypic profile of studied traits. This joint signal was fully explained neither by single-trait associations nor by overall behavioral adjustment problems but suggested a combined effect, which manifested through multiple subthreshold social, communicative, and cognitive impairments.

**Conclusions**—Our results suggest that common variation at 5p14.1 is associated with social communication spectrum phenotypes in the general population and support the role of rs4307059 as a quantitative trait locus for autism spectrum disorder.

Autism spectrum disorders are childhood neurodevelopmental conditions characterized by an impairment of social interaction and communication and by repetitive interests, behaviors, and activities (1). Based on a broad definition of autism spectrum disorder that includes autism, pervasive developmental disorder not otherwise specified, and Asperger's

Copyright © 2010 American Psychiatric Association. All rights reserved.

Address correspondence and reprint requests to Dr. St. Pourcain (Glaser), MRC Centre for Causal Analyses in Translational Epidemiology, University of Bristol, Oakfield House, Oakfield Grove, Bristol BS8 2BN, UK; beate.stpourcain@bristol.ac.uk.

All authors report no financial relationships with commercial interests.

syndrome, the condition affects more than 1% of children in the United Kingdom (2) and is four times more common in boys than in girls (3). Twin and family studies have demonstrated the high genetic liability of autism, reporting concordance rates in monozygotic twins of up to 92% (4) and a 22-fold increase in risk for first-degree relatives (5).

It has been shown that the genetic architecture of autism spectrum disorder is highly heterogeneous, often involving rare Mendelian and de novo mutations (6). Recent genome-wide association studies (7-9), however, have convincingly demonstrated that there is also common genetic variation within the genetic diversity of the condition, which is approachable through well-powered association designs. The strongest genome-wide significant and replicated evidence for association with autism spectrum disorder has so far been observed for an intergenic single-nucleotide polymorphism (SNP) (rs4307059) on 5p14.1 (8) that resides between two genes encoding neuronal cell-adhesion molecules (approximately 132 kb upstream of *CDH10* and 910 kb downstream of *CDH9*; human genome assembly hg19).

The discovery of common risk loci naturally raises the question of whether the underlying variants also contribute to variation in phenotypes that are related to autism spectrum disorder but are milder and nonpsychopathological. Prevailing psychological theories construe autism as a dimensional disorder, with autism spectrum disorder embodying only the extreme end of a continuum reflecting developmental difficulties (10, 11). Evidence for this hypothesis has been provided by the detection of subthreshold autistic traits in family members of autistic patients (12, 13) that are heritable (14, 15); further support has been lent by studies demonstrating that autistic traits are continuously distributed within the general population (16, 17), without natural boundaries between normal and abnormal behavior (11). Thus, it is possible that common variants, such as rs4307059 or variants in linkage disequilibrium with rs4307059, may act as a quantitative trait locus (QTL) underlying a broader autism phenotype.

In this study, we examined the association between rs4307059 and a series of broader autism phenotypes in members of the Avon Longitudinal Study of Parents and Children (ALSPAC), a large U.K. birth cohort representative of the general population for which the contribution of autism spectrum disorder-related traits has been demonstrated (18). Our phenotype selection focused on the social communication spectrum of autism spectrum disorder (11), which is prognostic of behavioral adjustment (18), highly heritable (11, 19), and probably etiologically distinct from the repetitive interests, behaviors, and activities spectrum (11). Specifically, we investigated single and joint genetic association effects involving standardized measures of language, communication, verbal intelligence, social interaction, and behavioral adjustment to study the impact of rs4307059 as a QTL for autism spectrum disorder.

## Method

### Sample

ALSPAC is a population-based prospective birth cohort for which there has been extensive data collection on the health and development of children and their parents (20). All pregnant women in the Bristol, England, area with an expected delivery between April 1991 and December 1992 were approached for participation in the study. A total of 14,541 women enrolled in the study, and 13,988 children were alive at 1 year. Ethical approval for the study was obtained from the ALSPAC Law and Ethics Committee and the Local Research Ethics Committees.

Of 9,650 offspring for whom DNA was available in 2009, 9,100 were successfully genotyped, which included 7,862 singletons. Of those, information was available for 7,738 on ethnicity and maternal education (assessed during the antenatal period: below O-level [9.23%], O-level [52.39%], and above O-level [38.38%]; O-levels are U.K. school-leaving qualifications taken at age 16). A total of 7,313 genotyped ethnically homogeneous white European singletons (52.07% of them males) were selected for this study.

### Autism Spectrum Disorder

The ALSPAC cohort contains a small proportion of children with autism spectrum disorder who were identified from either National Health Service community pediatric records or Education Service databases for the region (21). Eighty-six children with autism spectrum disorder were identified by age 11, giving a prevalence of 62 per 10,000 children. rs4307059 genotype information was available for 41 unrelated white European children with autism spectrum disorder.

### Genotyping

DNA was extracted as previously described (22). The genotyping of the DNA was performed by KBioscience ([www.kbioscience.co.uk](http://www.kbioscience.co.uk)) using a competitive allele-specific polymerase chain reaction system (KASPar). A genotyping call rate of 94.2% was achieved, and all genotypes were distributed in adherence to Hardy-Weinberg equilibrium (exact  $p=0.56$ ;  $N_{CC}=1,033$  [14.13%],  $N_{CT}=3,463$  [47.35%],  $N_{TT}=2,817$  [38.52%]). rs4307059 alleles were coded with respect to the major T allele (frequency=62.42%) conferring risk for autism spectrum disorder (8).

### Selection of Social Communication Spectrum Phenotypes

Broader autism phenotypes that may be related to a diagnosis of autism spectrum disorder used by Wang and colleagues (8) (see part 1 of the data supplement that accompanies the online edition of this article) were chosen with respect to early language, (social) communication, verbal intelligence, social interaction, behavioral difficulties, and special educational needs. Instruments were selected to cover most comprehensively continuous traits between ages 3 and 12 for the majority of children in ALSPAC. By age 3, individuals with autism exhibit some abnormality in at least one of the three key areas (1, 23). For each selected instrument we studied all available repeated measurements within the specified time frame. None of these measures captured specifically the repetitive interests, behaviors, and activities spectrum (see part 2 of the online data supplement).

Early language patterns were screened because a delay in the acquisition of early language is the most common initial symptom of autism spectrum disorder recognized by parents (24). Mother-reported early childhood language and communication patterns were assessed with the MacArthur Toddler Communication Questionnaire (25) summary language score (at a mean age of 3.25 years).

Social communication skills were assessed with the mother-reported Social and Communication Disorders Checklist (SCDC) (19) (at mean ages of 7.7 and 10.8 years) and communicative impairments with the mother-reported Children's Communication Checklist (CCC) (26) (at a mean age of 9.7 years). The SCDC is a brief screening instrument of social reciprocity and verbal/nonverbal communication (19), and the CCC captures important aspects of communication impairment of the broader autism phenotype in children (15).

Given the strong evidence for association between autism and cognitive functioning (3), verbal IQ was measured with the WISC-III (27) (at a mean age of 8.6 years). A short version

of the test consisting of alternate items only (except the coding task) was applied by trained psychologists during the year 8 focus clinic examination.

Mother-reported social interaction was assessed using the sociability subscale of the Emotionality-Activity-Sociability Temperament Survey (EAS) (28) (at mean ages of 3.3, 4.8, and 5.8 years) and the peer problem subscale of the Strengths and Difficulties Questionnaire (SDQ) (29) (at mean ages of 4.0, 6.8, 8.2, 9.7, and 11.8 years). There is some evidence that EAS sociability scores predict the quality of interactions in preschool children and their ability to handle conflict (30), and most of the variation in social functioning, as assessed by more specialized instruments, such as the Autism Spectrum Screening Questionnaire (31), is likely to be reflected by the SDQ peer problem subscale. Child-reported friendship scores (at mean ages of 8.6 and 10.7 years) were derived from the Cambridge Hormones and Moods Project Friendship Questionnaire (HMP-FQ) (32) to rate children's happiness with their friendships.

Behavioral difficulties were examined because social communicative deficits might be of prognostic significance with respect to behavioral adjustment at school (18). These were captured using mother-reported total behavioral difficulties as measured by the Revised Rutter Parent Scale for Preschool Children (R-RPS-PC) (33) during early life (at a mean age of 3.5 years), and later on by the SDQ (at mean ages of 4.0, 6.8, 8.2, 9.7, and 11.8 years). Children with special educational needs were identified using the Pupil Level Annual School Census (PLASC) for the 2003–2004 academic year (at a mean age of 11.8 years) supplied by the Department of Education for England, which has previously been used to identify autism cases in the ALSPAC sample (21). For each child attending a state school, the PLASC data provide information on the level of extra help being provided at school (“no special provision,” “school action,” “school action plus,” or “statement of special educational needs”; see part 3 of the online data supplement).

In total, 29 mostly continuous measures were subjected to genetic association analysis. For each phenotype, we determined whether it coded for high functionality (where higher scores indicate fewer problems or more functionality) or low functionality (where higher scores indicate more problems, less functionality, or fewer friends). Characteristics of the sample are summarized in Table 1, and trait intercorrelations are presented in Table S1 in the online data supplement.

### Missing Data

Like other large longitudinal cohorts, ALSPAC has attrition over time. Although this has been linked to sociodemographic factors, the distribution of genotypes is generally unrelated to socioeconomic status (34). In line with this notion, there was no evidence that rs4307059 genotypes, examined for their association with social communication spectrum traits, were predictors of trait missingness (based on a logistic regression approach using all available 9,100 genotypes to predict missingness for the 29 selected phenotypes; minimum p value, 0.26; data not shown).

### Statistical Analysis

The association between rs4307059 and each of the selected phenotypes was investigated with single regression models. Normally distributed traits were analyzed with ordinary least squares regression assuming an additive genetic model. Skewed continuous traits were modeled as count data, and the error distribution was specified as quasi-Poisson to accommodate for over- or underdispersion of the data. Categorical outcomes were analyzed using multinomial regression. The genetic effect for the two latter models was assumed to be log-additive, which is consistent with previous research (8).

In all analyses, genetic effects were adjusted for maternal education, sex, and age at phenotypic assessment. Although unlikely to be a confounder, maternal education as a covariate may considerably reduce the amount of unexplained variance, especially with respect to language-related phenotypes (e.g., maternal education explained 6.5% variance [adjusted  $R^2$ ] in verbal IQ; data not shown). The statistical significance of genetic associations was assessed using F tests or likelihood ratio tests.

As our prior hypothesis assumed an underlying autism spectrum disorder QTL, implicating an association between decreased intellectual, social, and communicative functionality (and likewise more behavioral and social problems) and a higher load of the rs4307059 risk allele, the analysis was performed in one-sided or directed mode. The expected direction of the genetic effect is outlined for each phenotype in Table 2.

To account for trait interrelatedness, multiple testing, and random genotype dropout, all single-trait associations were adjusted using permutations. A Bonferroni correction would have been less suitable because it cannot account for phenotypic correlation between traits. Genotypes from all genotyped individuals were included in the permutation vector, irrespective of whether they had been phenotyped or not, in order to avoid bias due to genotype dropout because of phenotype missingness.

For the permutation analysis, we created 100,000 permuted data sets through random assignment of the genotype vector, thus preserving the correlation structure among traits. Empirical p values for each trait were obtained as the proportion of permutations for which the minimum two-sided p value for all traits was less than or equal to the trait p value observed in the original data set. A one-sided mode was applied by restricting the vector of permuted two-sided p values to those associations where the effect was consistent with the direction of the original effect.

Evidence for profile associations between social communication spectrum phenotypes and rs4307059 was sought using an adapted permutation approach. A factor analytic/principal components-based approach seemed less suitable for the following reasons. First, socioeconomic influences, such as maternal education, explain a high proportion of the phenotypic variance (especially for language-related traits; see above) and are likely to load high on the first principal components/orthogonal factors. Thus, these factors are less likely to represent genetic variation as genotypes are generally unrelated to socioeconomic influences (34). Because single SNP variation explains only a small proportion overall of the phenotypic variance, often 1% even for genome-wide significant signals (35), genetic variation may largely be represented by components or factors that are dominated by stochastic variation. Second, multiple genetic associations may result because of shared phenotypic variation, but also because of pleiotropic effects, that is, multiple phenotype effects per single gene. The choice of a permutation approach, however, allows combining the entire evidence from all traits into a joint signal, irrespective of missingness, phenotype interrelatedness, or distribution and without the need to cluster phenotypic variance. The empirical approach we used thus assessed the combined evidence for genetic association, which is consistent with an underlying QTL for autism spectrum disorder (directed mode). A combined statistic, as suggested by Fisher in 1932 (36), was generated as

$$2 \sum_{i=1}^k -\ln P_i$$

where  $k$  is the number of phenotypes and  $P$  is the two-sided p value. Empirical p values were obtained as the proportion of permutations in which the summary statistic was greater than or equal to the statistic in the original data set *and* each single effect was consistent with

an underlying autism spectrum disorder QTL. Inconsistent p values were converted to 1. Thus, any inconsistent association in the original data set would not contribute to the original summary statistic but to the random variability during the permutation process.

All analyses were carried out using the R software package ([www.r-project.org](http://www.r-project.org)).

## Results

### Single-Trait Association Analysis

Using quasi-Poisson, ordinary least squares, and multinomial regression models, we studied single associations between each of the 29 social communication spectrum phenotypes and rs4307059 in a sample of 7,313 ALSPAC children (see Table 2). A higher load of the autism spectrum disorder risk allele was related to more stereotyped conversation and decreased ability to understand pragmatic aspects of communication as measured by two highly related CCC scales (Spearman's rank-order correlation  $r_s=0.74$ ,  $p<0.001$ ; see Table S1 in the online data supplement). It should be noted that the stereotyped conversation scale is a subscale of the pragmatic communication composite scale (26). The expected log count of more functional communication scores decreased by 0.0069 per increase in T allele (95% CI= -0.010 to -0.003) for the stereotyped conversation subscale (CCC; at a mean age of 9.7 years). This was accompanied by a decrease in 0.60 (95% CI=-0.91 to -0.29) pragmatic communication scores per increase in T allele (CCC; at a mean age of 9.7 years), with the genetic variation explaining ~0.26% of the total phenotypic variance (adjusted  $R^2$ ). An association between special educational needs (at a mean age of 11.8 years) and rs4307059 might be present as a trend.

Nominal associations were also observed between rs4307059 and other CCC scales (at a mean age of 9.7 years), social communication skills (SCDC, at a mean age of 10.8 years), behavioral difficulties (R-RPS-PC, at a mean age of 3.5 years; SDQ, at mean ages of 9.7 and 11.8 years), social interactions (HMP-FQ, at a mean age of 8.6 years), and cognitive functioning (WISC-III, at a mean age of 8.6 years), such that lower functionality, more problems, and fewer friends were associated with a greater frequency of the autism spectrum disorder risk allele (see Table 2).

The exclusion of 41 ALSPAC individuals with autism spectrum disorder did not attenuate the observed signals (data not shown). An association analysis restricted to these 41 individuals was not performed because of the small sample size. Multilevel sensitivity analysis (see part 4 and Table S2 in the online data supplement) on repeated measurements provided no evidence for variation in SNP effects over time. There was also no multilevel support for time-invariant SNP effects.

### Joint Association Analysis

Closer examination of the observed cross-sectional single-trait associations revealed that for the majority of traits, the detected genetic effects were consistent with the presence of an underlying QTL for autism spectrum disorder (see Figure 1 and Table 2). An exception from this pattern was observed only for sociability (EAS) at a mean age of 4.8 years. Hence, we also investigated the combined phenotype association effect for rs4307059. Using an extended permutation approach, this provided empirical evidence for a joint association between rs4307059 and the phenotypic spectrum of traits (joint  $p=0.00066$ ). Overall, the evidence for a joint signal was stronger than for any single-trait association alone. This combined signal was driven not only by highly associated single-trait associations but also by a phenotypic profile, as even after the exclusion of the CCC scales assessing stereotyped conversation and pragmatic aspects of communication we observed a joint association between rs4307059 and the remaining measures (joint  $p=0.0024$ ). Communicative,

cognitive, and social interactive components of the spectrum appear to play a major role in this association as further exclusion of information on total behavioral difficulties (R-RPS-PC, at a mean age of 3.5 years; SDQ, at mean ages of 4.0, 6.8, 8.2, 9.7, and 11.8 years) and special educational needs (at a mean age of 11.8 years) only marginally attenuated the effect (joint  $p=0.0048$ ). None of the implicated phenotypes, including verbal IQ (WISC-III) and social communication skills (SCDC), showed evidence for association when adjusted for multiple testing as a single trait (see Table 2).

## Discussion

We report a genetic association between social communication spectrum phenotypes and a high-risk autism spectrum disorder locus (8) in members of a large U.K. birth cohort representative of the general population. We found robust evidence for single-trait associations involving children's communication and joint association effects that were related to the phenotypic profile of studied traits. These findings were consistent with the autism spectrum disorder risk profile of the SNP identified in previous association studies (8) and are unlikely to be affected by bias, multiple testing, phenotype interrelatedness, or random genotype dropout.

The strongest single-trait associations were observed for stereotyped conversation and pragmatic communication skills, as measured by two related CCC scales (see Table 2), where a higher load of the risk allele was related to more communication deficits. Severe problems with pragmatic aspects of communication appear to be universal in autism spectrum disorder and are usually ascribed to an impaired theory of mind (37)—that is, an impaired understanding of other people's minds and mental states. Assuming the dimensionality of communication problems, which is consistent with the previously observed continuum of traits contributing to the autistic spectrum in ALSPAC (18), our findings provide evidence that variation at a high-risk locus for autism spectrum disorder is also related to a broader autism phenotype in the general population that is likely to manifest as inappropriate use of language in social contexts.

It is also possible that variation at rs4307059 is linked to some aspects of behavioral adjustment as reflected by special educational needs, although the identification of special educational needs is based on a large collection of criteria, including cognitive skills and the presence or absence of attention deficit hyperactivity disorder, speech problems, and physical difficulties (21).

In addition, this study produced support for joint association effects that were not solely explained by single-trait associations, as the combined effect was stronger than for any single association alone and was still present when highly associated measures were excluded. Joint effects were related to a highly consistent pattern of single allelic associations involving language, communication, cognition, social interaction, and behavioral adjustment such that less functionality and more problems were associated with a higher frequency of the common rs4307059 risk allele, although the confidence intervals for single effects were often wide. Deviation from this pattern was detected for only one of 29 selected phenotypes. As our findings indicate, the likelihood that the same risk-associated allele would predict variation in these phenotypes within members of the general population by chance is very remote.

The exact risk structure of the joint association profile, however, is still unknown and likely to be complex. Given that even for highly associated single traits, the SNP explains only a small proportion of the phenotypic variance ( $<0.3\%$ ), as expected for genetic variation in general, the phenotypic relatedness among the measures is difficult to disentangle, especially

as the joint effect captures both pleiotropic effects and shared sources of phenotypic variation. However, exclusion of measures of overall behavioral difficulties and special educational needs (in addition to highly associated single-trait associations) attenuated the combined effect only marginally. This lends support to the hypothesis that total behavioral difficulties might be mediated through social ability since adjustment problems, particularly in school, might be the consequence of inadequate social communicative skills (18). More importantly, our findings provide evidence that the combined signal involves multiple social, communicative, and cognitive impairments, including lower verbal IQ (WISC-III) and impaired social communication skills (SCDC), none of which showed association with rs4307059 when adjusted for multiple testing. Common variation at rs4307059 may therefore be related to small impairments in overall social and communication skills, which might not be detectable as single signals (subthreshold impairments) but might increase the susceptibility for autism spectrum disorder. As such, variation at rs4307059 may provide the genetic background against which a more severe disruption of an autism spectrum disorder risk locus is more likely to result in a diagnosis of autism (38). Although the molecular mechanisms are still unidentified, it has been hypothesized that rs4307059 tags causal variants in *CDH9/CDH10*, a theory that is supported by pathway-based association analyses implicating neuronal cell-adhesion molecules in the etiology of autism spectrum disorder (8).

Taken together, our results support the hypothesis that variability at rs4307059 is related to an underlying QTL for autism spectrum disorder in the general population. Moreover, they suggest that multiple facets of a broader autism phenotype share underlying genetic risk factors, which is consistent with previous reports on shared genetic influences for social, communication, and language skills (39).

## Limitations

These findings must be interpreted within the context of potential limitations. First, although measures were derived from educational databases and mother and child reports, mother-reported data were overrepresented in the study and may have contributed to greater variance sharing. Second, the analysis included only white European children, and the reported genetic association therefore needs to be replicated in populations of different ethnic backgrounds. Third, instruments were chosen that focus on social communication spectrum phenotypes assessed between ages 3 and 12 for the majority of children in ALSPAC. However, they nonetheless represent a cohort-specific selection of measures, which may not be exhaustive. This is unlikely to affect the reported single-trait associations, but it may impinge on the total of the investigated phenotypic variance, which is reflected by the joint signal.

## Conclusions

To our knowledge, this is the first study to provide evidence that a common high-risk autism spectrum disorder variant is associated with a broader autism spectrum disorder phenotype in the general population. Affected phenotypes manifested as impaired communication skills and multiple subthreshold social, communicative, cognitive, and behavioral impairments. Our findings contribute to the understanding of genetic factors underlying the complex etiology of autism and support the role of rs4307059 as a QTL for autism spectrum disorder. Our study thus opens a novel perspective on identifying loci for dimensional disorders linking the genetics of disease with the genetics of population variation.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.



## Acknowledgments

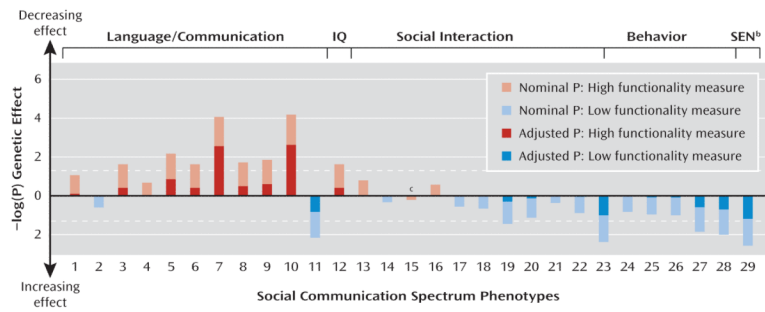
Supported by grant 74882 from the UK Medical Research Council and grant 076467/Z05/z from the Wellcome Trust. The University of Bristol provided core support for the Avon Longitudinal Study of Parents and Children and this analysis.

## References

1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, 4th ed, Text Revision (DSM -IV-TR). American Psychiatric Association; Washington, DC: 2000.
2. Baird G, Simonoff E, Pickles A, Chandler S, Loucas T, Meldrum D, Charman T. Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the Special Needs and Autism Project (SNAP). *Lancet*. 2006; 368:210–215. [PubMed: 16844490]
3. Fombonne E. Epidemiology of autistic disorder and other pervasive developmental disorders. *J Clin Psychiatry*. 2005; 66:3–8. [PubMed: 16401144]
4. Bailey A, Le Couteur A, Gottesman I, Bolton P, Simonoff E, Yuzda E, Rutter M. Autism as a strongly genetic disorder: evidence from a British twin study. *Psychol M ed*. 1995; 25:63–77.
5. Lauritsen MB, Pedersen CB, Mortensen PB. The incidence and prevalence of pervasive developmental disorders: a Danish population-based study. *Psychol Med*. 2004; 34:1339–1346. [PubMed: 15697060]
6. Sebat J, Lakshmi B, Malhotra D, Troge J, Lese-Martin C, Walsh T, Yamrom B, Yoon S, Krasnitz A, Kendall J, Leotta A, Pai D, Zhang R, Lee YH, Hicks J, Spence SJ, Lee AT, Puura K, Lehtimäki T, Ledbetter D, Gregersen PK, Bregman J, Sutcliffe JS, Jobanputra V, Chung W, Warburton D, King MC, Skuse D, Geschwind DH, Gilliam TC, Ye K, Wigler M. Strong association of de novo copy number mutations with autism. *Science*. 2007; 316:445–449. [PubMed: 17363630]
7. Ma D, Salyakina D, Jaworski JM, Konidari I, Whitehead PL, Andersen AN, Hoffman JD, Slifer SH, Hedges DJ, Cukier HN, Griswold AJ, McCauley JL, Beecham GW, Wright HH, Abramson RK, Martin ER, Hussman JP, Gilbert JR, Cuccaro ML, Haines JL, Pericak-Vance MA. A genome-wide association study of autism reveals a common novel risk locus at 5p14.1. *Ann Hum Genet*. 2009; 73:263–273. [PubMed: 19456320]
8. Wang K, Zhang H, Ma D, Bucan M, Glessner JT, Abrahams BS, Salyakina D, Imielinski M, Bradfield JP, Sleiman PM, Kim CE, Hou C, Frackelton E, Chiavacci R, Takahashi N, Sakurai T, Rappaport E, Lajonchere CM, Munson J, Estes A, Korvatska O, Piven J, Sonnenblick LI, Alvarez Retuerto AI, Herman EI, Dong H, Hutman T, Sigman M, Ozonoff S, Klin A, Owley T, Sweeney JA, Brune CW, Cantor RM, Bernier R, Gilbert JR, Cuccaro ML, McMahon WM, Miller J, State MW, Wassink TH, Coon H, Levy SE, Schultz RT, Nurnberger JI, Haines JL, Sutcliffe JS, Cook EH, Minshew NJ, Buxbaum JD, Dawson G, Grant SF, Geschwind DH, Pericak-Vance Ma, Schellenberg GD, Hakonarson H. Common genetic variants on 5p14.1 associate with autism spectrum disorders. *Nature*. 2009; 459:528–533. [PubMed: 19404256]
9. Weiss LA. Autism genetics: emerging data from genome-wide copy-number and single nucleotide polymorphism scans. *Expert Rev Mol Diagn*. 2009; 9:795–803. [PubMed: 19895225]
10. Wing, L. The continuum of autistic characteristics, in *Diagnosis and Assessment in Autism*. Schopler, E.; Mesibov, G., editors. Plenum; New York: 1988. p. 91-110.
11. Mandy W, Skuse D. Research review: what is the association between the social-communication element of autism and repetitive interests, behaviours and activities? *J Child Psychol Psychiatry*. 2008; 49:795–808. [PubMed: 18564070]
12. Bolton P, Macdonald H, Pickles A, Rios P, Goode S, Crowson M, Bailey A, Rutter M. A case-control family history study of autism. *J Child Psychol Psychiatry*. 1994; 35:877–900. [PubMed: 7962246]
13. Piven J, Palmer P, Jacobi D, Childress D, Arndt S. Broader autism phenotype: evidence from a family history study of multiple-incidence autism families. *Am J Psychiatry*. 1997; 154:185–190. [PubMed: 9016266]
14. Constantino JN, Todd RD. Intergenerational transmission of subthreshold autistic traits in the general population. *Biol Psychiatry*. 2005; 57:655–660. [PubMed: 15780853]

15. Bishop D, Maybery M, Wong D, Maley A, Hallmayer J. Characteristics of the broader phenotype in autism: a study of siblings using the Children's Communication Checklist-2. *Am J Med Genet B Neuropsychiatr Genet.* 2006; 141B:117–122. [PubMed: 16389586]
16. Constantino JN, Todd RD. Autistic traits in the general population: a twin study. *Arch Gen Psychiatry.* 2003; 60:524–530. [PubMed: 12742874]
17. Posserud M, Lundervold AJ, Gillberg C. Autistic features in a total population of 7–9-year-old children assessed by the ASSQ (Autism Spectrum Screening Questionnaire). *J Child Psychol Psychiatry.* 2006; 47:167–175. [PubMed: 16423148]
18. Skuse D, Mandy W, Steer C, Miller L, Goodman R, Lawrence K, Emond A, Golding J. Social communication competence and functional adaptation in a general population of children: preliminary evidence for sex-by-verbal IQ differential risk. *J Am Acad Child Adolesc Psychiatry.* 2008; 48:128–137. [PubMed: 19106766]
19. Skuse D, Mandy W, Scourfield J. Measuring autistic traits: heritability, reliability, and validity of the Social and Communication Disorders Checklist. *Br J Psychiatry.* 2005; 187:568–572. [PubMed: 16319410]
20. Golding J, Pembrey M, Jones R, ALSPAC Study Team. ALSPAC: the Avon Longitudinal Study of Parents and Children, I: study methodology. *Paediatr Perinat Epidemiol.* 2001; 15:74–87. [PubMed: 11237119]
21. Williams E, Thomas K, Sidebotham H, Emond A. Prevalence and characteristics of autistic spectrum disorders in the ALSPAC cohort. *Dev Med Child Neurol.* 2008; 50:672–677. [PubMed: 18754916]
22. Jones RW, Ring S, Tyfield L, Hamvas R, Simmons H, Pembrey M, Golding J. A new human genetic resource: a DNA bank established as part of the Avon Longitudinal Study of Pregnancy and Childhood (ALSPAC). *Eur J Hum Genet.* 2000; 8:653–660. [PubMed: 10980570]
23. World Health Organization. *Mental Disorders: A Glossary and Guide to Their Classification in Accordance With the 10th Revision of the International Classification of Diseases.* World Health Organization; Geneva: 1993.
24. De Giacomo A, Fombonne E. Parental recognition of developmental abnormalities in autism. *Eur Child Adolesc Psychiatry.* 1998; 7:131–136. [PubMed: 9826299]
25. Fenson, L.; Dale, P.; Reznick, J. *MacArthur Communicative Development Inventories.* Singular Publishing Group; San Diego: 1993.
26. Bishop D. Development of the Children's Communication Checklist (CCC): a method for assessing qualitative aspects of communicative impairment in children. *J Child Psychol Psychiatry.* 1998; 39:879–891. [PubMed: 9758196]
27. Wechsler, D.; Golombok, J.; Rust, J. *WISC-III UK Wechsler Intelligence Scale for Children: UK Manual.* 3rd ed.. Psychological Corp; Sidcup, UK: 1992.
28. Buss, A.; Plomin, R. *Temperament: Early Developing Personality Traits.* Lawrence Erlbaum Associates; Hillsdale, NJ: 1984.
29. Goodman R. The Strengths and Difficulties Questionnaire: a research note. *J Child Psychol Psychiatry.* 1997; 38:581–586. [PubMed: 9255702]
30. Graziano WG, Jensen-Campbell LA, Sullivan-Logan GM. Temperament, activity, and expectations for later personality development. *J Pers Soc Psychol.* 1998; 74:1266–1277. [PubMed: 9599442]
31. Posserud M, Lundervold AJ, Steijnen MC, Verhoeven S, Stormark KM, Gillberg C. Factor analysis of the Autism Spectrum Screening Questionnaire. *Autism.* 2008; 12:99–112. [PubMed: 18178599]
32. Goodyer IM, Wright C, Altham PM. Recent friendships in anxious and depressed school age children. *Psychol Med.* 1989; 19:165–174. [PubMed: 2727204]
33. Elander J, Rutter M. Use and development of the Rutter Parents' and Teachers' Scales. *Int J Methods Psychiatr Res.* 1996; 6:63–78.
34. Davey-Smith G, Lawlor DA, Harbord R, Timpson N, Day I, Ebrahim S. Clustered environments and randomized genes: a fundamental distinction between conventional and genetic epidemiology. *PLoS Med.* 2007; 4:e352. [PubMed: 18076282]

35. Weedon MN, Lango H, Lindgren CM, Wallace C, Evans DM, Mangino M, Freathy RM, Perry JRB, Stevens S, Hall AS, Samani NJ, Shields B, Prokopenko I, Farrall M, Dominiczak A, Johnson T, Bergmann S, Beckmann JS, Vollenweider P, Waterworth DM, Mooser V, Palmer CNA, Morris AD, Ouwehand WH, Caulfield M, Munroe PB, Hattersley AT, McCarthy MI, Frayling TM. Genome-wide association analysis identifies 20 loci that influence adult height. *Nat Genet.* 2008; 40:575–583. [PubMed: 18391952]
36. Fisher, RA. *Statistical Methods for Research Workers.* Oliver and Boyd; Edinburgh: 1932.
37. Baron-Cohen S. Social and pragmatic deficits in autism: cognitive or affective? *J Autism Dev Disord.* 1988; 18:379–402. [PubMed: 3049519]
38. Skuse DH. Rethinking the nature of genetic vulnerability to autistic spectrum disorders. *Trends Genet.* 2007; 23:387–395. [PubMed: 17630015]
39. Dworzynski K, Ronald A, Hayiou-Thomas M, Rijdsdijk F, Happé F, Bolton P, Plomin R. Aetiological relationship between language performance and autistic-like traits in childhood: a twin study. *Int J Lang Commun Disord.* 2007; 42:273–292. [PubMed: 17514542]



**FIGURE 1. Association Profile Between rs4307059 and Social Communication Spectrum Phenotypes<sup>a</sup>**

<sup>a</sup> All p values are one-sided; dashed lines correspond to a p value threshold of 0.05. All spectrum phenotypes and their p values are numbered as outlined in Table 2.

<sup>b</sup> SEN=special educational needs.

<sup>c</sup> Inconsistent with an underlying autism spectrum disorder QTL.

**TABLE 1**  
 Sample Characteristics of Social Communication Spectrum Traits in a Subsample of the Avon Longitudinal Study of Parents and Children Birth Cohort (N=7,313)<sup>d</sup>

Measure	Instrument <sup>b</sup>	Mean Age <sup>c</sup> (Years)	Median	SD	N	%	Code <sup>d</sup>	Phenotype Distribution <sup>e</sup>
Continuous spectrum phenotype								
<b>Language/Communication</b>								
1. Language	CDI	3.3	310	35.50	5,948	81.33	H	S
2. Social communication	SCDC	7.7	2	3.71	5,319	72.73	L	S
3. Intelligibility and fluency	CCC	9.7	36	1.89	5,181	70.85	H	S
4. Syntax	CCC	9.7	32	0.51	5,161	70.57	H	S
5. Inappropriate initiation	CCC	9.7	27	2.38	5,164	70.61	H	S
6. Coherence	CCC	9.7	36	2.01	5,169	70.68	H	S
7. Stereotyped conversation	CCC	9.7	27	2.44	5,149	70.41	H	S
8. Use of conversational context	CCC	9.7	30	2.09	5,093	69.64	H	S
9. Conversational rapport	CCC	9.7	33	1.95	5,094	69.66	H	S
10. Pragmatic aspects of communication	CCC	9.7	153	7.75	5,071	69.34	H	NM
11. Social communication	SCDC	10.8	1	3.54	4,920	67.28	L	S
<b>IQ</b>								
12. Verbal IQ	WISC-III	8.6	107	16.51	4,925	67.35	H	NM
<b>Social interaction</b>								
13. Sociability	EAS	3.3	18	3.08	6,099	83.40	H	NM
14. Peer problems	SDQ	4.0	1	1.48	5,754	78.68	L	S
15. Sociability	EAS	4.8	18	2.59	5,932	81.12	H	NM
16. Sociability	EAS	5.8	18	2.62	5,562	76.06	H	NM
17. Peer problems	SDQ	6.8	1	1.40	5,442	74.42	L	S
18. Peer problems	SDQ	8.2	1	1.53	5,028	68.75	L	S
19. Friendships score	HMP-FQ	8.6	3	2.43	4,770	65.23	L	NM
20. Peer problems	SDQ	9.7	1	1.49	5,138	70.26	L	S
21. Friendships score	HMP-FQ	10.7	3	2.14	4,821	65.92	L	NM
22. Peer problems	SDQ	11.8	0	1.55	4,726	64.62	L	S
<b>Behavior</b>								

Measure	Instrument <sup>b</sup>	Mean Age <sup>c</sup> (Years)	Count	%	N	Total	Code	Phenotype Distribution <sup>e</sup>
23. Total behavioral difficulties	R-RPS-PC	3.5	12	5.59	6,007	82.14	L	NM
24. Total behavioral difficulties	SDQ	4.0	8	4.53	5,642	77.15	L	NM
25. Total behavioral difficulties	SDQ	6.8	6	4.67	5,432	74.28	L	NM
26. Total behavioral difficulties	SDQ	8.2	7	5.14	5,023	68.69	L	NM
27. Total behavioral difficulties	SDQ	9.7	6	4.79	5,124	70.07	L	NM
28. Total behavioral difficulties	SDQ	11.8	5	4.92	4,725	64.61	L	NM
<b>Categorical spectrum phenotype</b>								
<b>Special educational needs</b>								
29. Special educational needs		11.8	5,197	85.80	6,057	82.83	L	C
No special provision			5,24	8.65				
School action			1,90	3.14				
School action plus			146	2.41				
Statement of special educational needs								

<sup>a</sup>This subsample comprised genotyped ethnically homogeneous white European singletons.

<sup>b</sup>CDI=MacArthur Toddler Communication Questionnaire; SCDC=Social and Communication Disorders Checklist; CCC=Children's Communication Checklist; WISC-III=Wechsler Intelligence Scale for Children, 3rd ed.; EAS=Emotionality-Activity-Sociability Temperament Survey; SDQ=Strengths and Difficulties Questionnaire; HMP-FQ=Cambridge Hormones and Moods Project Friendship Questionnaire; R-RPS-PC=Revised Rutter Parent Scale for Preschool Children.

<sup>c</sup>Mean age at phenotypic assessment.

<sup>d</sup>Phenotype coding: H=higher scores indicate less problems or more functionality; L=higher scores indicate more problems, less functionality, or fewer friends.

<sup>e</sup>NM=assumed normal; S=skewed; C=categorical.

**TABLE 2**  
 Association Between rs4307059 and Social Communication Spectrum Traits in the Avon Longitudinal Study of Parents and Children Birth Cohort (N=7,313)<sup>d</sup>

Measure	Instrument <sup>b</sup>	Mean Age <sup>c</sup> (Years)	$\beta$ -OLS <sup>d</sup>	$\beta$ -Q <sup>e</sup>	ASD QTL Effect <sup>f</sup>	SE	Nominal $p^g$	Adjusted $p^h$
<b>Continuous spectrum phenotype</b>								
<b>Language/Communication</b>								
1. Language	CDI	3.3	-0.0030		<0	0.0022	0.087	0.78
2. Social communication	SCDC	7.7	0.018		>0	0.026	0.25	0.98
3. Intelligibility and fluency	CCC	9.7	-0.0022		<0	0.0011	0.024	0.39
4. Syntax	CCC	9.7	-0.00027		<0	0.00032	0.21	0.97
5. Inappropriate initiation	CCC	9.7	-0.0045		<0	0.0018	0.0068	0.14
6. Coherence	CCC	9.7	-0.0023		<0	0.0012	0.024	0.39
7. Stereotyped conversation	CCC	9.7	-0.0069		<0	0.0018	0.000086	0.0023
8. Use of conversational context	CCC	9.7	-0.0029		<0	0.0014	0.019	0.32
9. Conversational rapport	CCC	9.7	-0.0027		<0	0.0012	0.014	0.25
10. Pragmatic aspects of communication	CCC	9.7	-0.60		<0	0.16	0.000066	0.0017
11. Social communication	SCDC	10.8	0.076		>0	0.031	0.0070	0.15
<b>IQ</b>								
12. Verbal IQ	WISC-III	8.6	-0.64		<0	0.33	0.024	0.39
<b>Social interaction</b>								
13. Sociability	EAS	3.3	-0.058		<0	0.058	0.16	0.93
14. Peer problems	SDQ	4.0	0.0015		>0	0.019	0.47	1.00
15. Sociability	EAS	4.8	0.015 <sup>i</sup>		<0	0.049	0.62 <sup>j</sup>	1.00 <sup>k</sup>
16. Sociability	EAS	5.8	-0.032		<0	0.051	0.27	0.99
17. Peer problems	SDQ	6.8	0.016		>0	0.027	0.28	0.99
18. Peer problems	SDQ	8.2	0.019		>0	0.024	0.22	0.97
19. Friendships score	HMP-FQ	8.6	0.093		>0	0.052	0.036	0.50
20. Peer problems	SDQ	9.7	0.040		>0	0.028	0.075	0.74
21. Friendships score	HMP-FQ	10.7	0.008		>0	0.045	0.43	1.00
22. Peer problems	SDQ	11.8	0.037		>0	0.032	0.13	0.88
<b>Behavior</b>								

Measure	Instrument <sup>b</sup>	Mean Age <sup>c</sup> (Years)	$\beta$ -OLS <sup>d</sup>	$\beta$ -QP <sup>e</sup>	ASD QTL Effect <sup>f</sup>	SE	Nominal $p^g$	Adjusted $p^h$
23. Total behavioral difficulties	R-RPS-PC	3.5	0.28		>0	0.11	0.0042	0.094
24. Total behavioral difficulties	SDQ	4.0	0.091		>0	0.087	0.15	0.92
25. Total behavioral difficulties	SDQ	6.8	0.12		>0	0.092	0.11	0.84
26. Total behavioral difficulties	SDQ	8.2	0.13		>0	0.10	0.10	0.83
27. Total behavioral difficulties	SDQ	9.7	0.21		>0	0.098	0.014	0.26
28. Total behavioral difficulties	SDQ	11.8	0.24		>0	0.10	0.010	0.20
Categorical spectrum phenotype								
<b>Special educational needs</b>								
29. Special educational needs		11.8					0.0027	0.063
School action				1.030	>1	0.070		
School action plus				1.48	>1	0.17		
Statement of special educational needs				1.11	>1	0.14		

<sup>a</sup>This subsample comprised genotyped ethnically homogeneous white European singletons.

<sup>b</sup>CDE=MacArthur Toddler Communication Questionnaire; SCDC=Social and Communication Disorders Checklist; CCC=Children's Communication Checklist; WISC-III=Wechsler Intelligence Scale for Children, 3rd ed.; EAS=Emotionality-Activity-Sociability Temperament Survey; SDQ=Strengths and Difficulties Questionnaire; HMP-FQ=Cambridge Hormones and Moods Project Friendship Questionnaire; R-RPS-PC=Revised Rutter Parent Scale for Preschool Children.

<sup>c</sup>Mean age at phenotypic assessment.

<sup>d</sup> $\beta$ =genetic effect with respect to the risk allele (T) adjusted for age, sex, and maternal education;  $\beta$ -OLS=ordinary least squares estimate.

<sup>e</sup> $\beta$ =genetic effect with respect to the risk allele (T) adjusted for age, sex, and maternal education;  $\beta$ -QP=quasi-Poisson estimate (log counts).

<sup>f</sup>Expected direction of the effect assuming an underlying autism spectrum disorder QTL.

<sup>g</sup>One-sided p value based on F tests or likelihood ratio tests (autism spectrum disorder QTL consistent effect: Pone-sided = Ptwo-sided/2; autism spectrum disorder QTL inconsistent effect: Pone-sided = 1 - Ptwo-sided/2).

<sup>h</sup>One-sided empirical p value.

<sup>i</sup>Inconsistent effect.

<sup>j</sup>Odds ratios are based on a multinomial regression (reference category, "no special provision").