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.
George Davey Smith, M.D., D.Sc.

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.

(Am J Psychiatry 2010; 167:1364–1372)

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 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

This article is featured in this month’s AJP Audio, is the subject of a CME course (p. 1415), and is discussed in an editorial by Dr. Kendler (p. 1291).
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.kbiosciences.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; Nc = 1,033 [14.13%, Nc = 3,463 [47.35%, Nc = 1,033 [14.13%, Nc = 887 [12.39%]; 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. 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.

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).
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 avail-

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.
able 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 R²] 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-locus 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 interrelat-
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

### TABLE 2. Association Between rs4307059 and Social Communication Spectrum Traits in the Avon Longitudinal Study of Parents and Children Birth Cohort (N=7,313)

<table>
<thead>
<tr>
<th>Measure</th>
<th>Instrument(^b)</th>
<th>Mean Age(^c) (Years)</th>
<th>(\beta)-OLS(^d)</th>
<th>(\beta)-QP(^e) Effect(^f)</th>
<th>SE</th>
<th>Nominal p(^g)</th>
<th>Adjusted p(^h)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Continuous spectrum phenotype</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Language/Communication</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Language</td>
<td>CDI</td>
<td>3.3</td>
<td>-0.0030</td>
<td>&lt;0</td>
<td>0.0022</td>
<td>0.087</td>
<td>0.78</td>
</tr>
<tr>
<td>2. Social communication</td>
<td>SDCC</td>
<td>7.7</td>
<td>0.018</td>
<td>&gt;0</td>
<td>0.026</td>
<td>0.25</td>
<td>0.98</td>
</tr>
<tr>
<td>3. Intelligibility and fluency</td>
<td>CCC</td>
<td>9.7</td>
<td>-0.0022</td>
<td>&lt;0</td>
<td>0.0011</td>
<td>0.024</td>
<td>0.39</td>
</tr>
<tr>
<td>4. Syntax</td>
<td>CCC</td>
<td>9.7</td>
<td>-0.00027</td>
<td>&lt;0</td>
<td>0.00032</td>
<td>0.21</td>
<td>0.97</td>
</tr>
<tr>
<td>5. Inappropriate initiation</td>
<td>CCC</td>
<td>9.7</td>
<td>-0.0045</td>
<td>&lt;0</td>
<td>0.0018</td>
<td>0.0068</td>
<td>0.14</td>
</tr>
<tr>
<td>6. Coherence</td>
<td>CCC</td>
<td>9.7</td>
<td>-0.0023</td>
<td>&lt;0</td>
<td>0.0012</td>
<td>0.024</td>
<td>0.39</td>
</tr>
<tr>
<td>7. Stereotyped conversation</td>
<td>CCC</td>
<td>9.7</td>
<td>-0.0069</td>
<td>&lt;0</td>
<td>0.0018</td>
<td>0.000086</td>
<td>0.0023</td>
</tr>
<tr>
<td>8. Use of conversational context</td>
<td>CCC</td>
<td>9.7</td>
<td>-0.0029</td>
<td>&lt;0</td>
<td>0.0014</td>
<td>0.019</td>
<td>0.32</td>
</tr>
<tr>
<td>9. Conversational rapport</td>
<td>CCC</td>
<td>9.7</td>
<td>-0.0027</td>
<td>&lt;0</td>
<td>0.0012</td>
<td>0.014</td>
<td>0.25</td>
</tr>
<tr>
<td>10. Pragmatic aspects of communication</td>
<td>CCC</td>
<td>9.7</td>
<td>-0.60</td>
<td>&lt;0</td>
<td>0.16</td>
<td>0.000066</td>
<td>0.0017</td>
</tr>
<tr>
<td>11. Social communication</td>
<td>SCDC</td>
<td>10.8</td>
<td>0.076</td>
<td>&gt;0</td>
<td>0.031</td>
<td>0.070</td>
<td>0.15</td>
</tr>
<tr>
<td><strong>IQ</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Verbal IQ</td>
<td>WISC-III</td>
<td>8.6</td>
<td>-0.64</td>
<td>&lt;0</td>
<td>0.33</td>
<td>0.024</td>
<td>0.39</td>
</tr>
<tr>
<td><strong>Social interaction</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. Sociability</td>
<td>EAS</td>
<td>3.3</td>
<td>-0.058</td>
<td>&lt;0</td>
<td>0.058</td>
<td>0.16</td>
<td>0.93</td>
</tr>
<tr>
<td>14. Peer problems</td>
<td>SDQ</td>
<td>4.0</td>
<td>0.0015</td>
<td>&gt;0</td>
<td>0.019</td>
<td>0.47</td>
<td>1.00</td>
</tr>
<tr>
<td>15. Sociability</td>
<td>EAS</td>
<td>4.8</td>
<td>0.015</td>
<td>&lt;0</td>
<td>0.049</td>
<td>0.62</td>
<td>1.00</td>
</tr>
<tr>
<td>16. Sociability</td>
<td>EAS</td>
<td>5.8</td>
<td>-0.032</td>
<td>&lt;0</td>
<td>0.051</td>
<td>0.27</td>
<td>0.99</td>
</tr>
<tr>
<td>17. Peer problems</td>
<td>SDQ</td>
<td>6.8</td>
<td>0.016</td>
<td>&gt;0</td>
<td>0.027</td>
<td>0.28</td>
<td>0.99</td>
</tr>
<tr>
<td>18. Peer problems</td>
<td>SDQ</td>
<td>8.2</td>
<td>0.019</td>
<td>&gt;0</td>
<td>0.024</td>
<td>0.22</td>
<td>0.97</td>
</tr>
<tr>
<td>19. Friendships score</td>
<td>HMP-FQ</td>
<td>8.6</td>
<td>0.093</td>
<td>&gt;0</td>
<td>0.052</td>
<td>0.036</td>
<td>0.50</td>
</tr>
<tr>
<td>20. Peer problems</td>
<td>SDQ</td>
<td>9.7</td>
<td>0.040</td>
<td>&gt;0</td>
<td>0.028</td>
<td>0.075</td>
<td>0.74</td>
</tr>
<tr>
<td>21. Friendships score</td>
<td>HMP-FQ</td>
<td>10.7</td>
<td>0.008</td>
<td>&gt;0</td>
<td>0.045</td>
<td>0.43</td>
<td>1.00</td>
</tr>
<tr>
<td>22. Peer problems</td>
<td>SDQ</td>
<td>11.8</td>
<td>0.037</td>
<td>&gt;0</td>
<td>0.032</td>
<td>0.13</td>
<td>0.88</td>
</tr>
<tr>
<td><strong>Behavior</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23. Total behavioral difficulties</td>
<td>R-RPS-PC</td>
<td>3.5</td>
<td>0.28</td>
<td>&gt;0</td>
<td>0.11</td>
<td>0.0042</td>
<td>0.094</td>
</tr>
<tr>
<td>24. Total behavioral difficulties</td>
<td>SDQ</td>
<td>4.0</td>
<td>0.091</td>
<td>&gt;0</td>
<td>0.087</td>
<td>0.15</td>
<td>0.92</td>
</tr>
<tr>
<td>25. Total behavioral difficulties</td>
<td>SDQ</td>
<td>6.8</td>
<td>0.12</td>
<td>&gt;0</td>
<td>0.092</td>
<td>0.11</td>
<td>0.84</td>
</tr>
<tr>
<td>26. Total behavioral difficulties</td>
<td>SDQ</td>
<td>8.2</td>
<td>0.13</td>
<td>&gt;0</td>
<td>0.10</td>
<td>0.10</td>
<td>0.83</td>
</tr>
<tr>
<td>27. Total behavioral difficulties</td>
<td>SDQ</td>
<td>9.7</td>
<td>0.21</td>
<td>&gt;0</td>
<td>0.098</td>
<td>0.014</td>
<td>0.26</td>
</tr>
<tr>
<td>28. Total behavioral difficulties</td>
<td>SDQ</td>
<td>11.8</td>
<td>0.24</td>
<td>&gt;0</td>
<td>0.10</td>
<td>0.010</td>
<td>0.20</td>
</tr>
<tr>
<td><strong>Categorical spectrum phenotype</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Special educational needs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29. Special educational needs</td>
<td>11.8</td>
<td>0.0027</td>
<td></td>
<td></td>
<td></td>
<td>0.063</td>
<td></td>
</tr>
<tr>
<td>School action</td>
<td>1.030</td>
<td>&gt;1</td>
<td>0.070</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School action plus</td>
<td>1.48</td>
<td>&gt;1</td>
<td>0.17</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statement of special educational needs</td>
<td>1.11</td>
<td>&gt;1</td>
<td>0.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) This subsample comprised genotyped ethnically homogeneous white European singletons.

\(^b\) CDI = MacArthur Toddler Communication Questionnaire; SDCC = 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.

\(^c\) Mean age at phenotypic assessment.

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

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

\(^f\) Expected direction of the effect assuming an underlying autism spectrum disorder QTL.

\(^g\) 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).

\(^h\) One-sided empirical p value.

\(^i\) Inconsistent effect.

\(^j\) Odds ratios are based on a multinomial regression (reference category, “no special provision”).
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.

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 of
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.
References


10. Wing L: The continuum of autistic characteristics, in Diagnosis and Assessment in Autism. Edited by Schopler E, Mesibov G. New York, Plenum, 1988, pp 91–110


Wide association analysis identifies 20 loci that influence adult height. Nat Genet 2008; 40:575–583