QTL Association Analysis of the DRD4 Exon 3 VNTR Polymorphism in a Population Sample of Children Screened With a Parent Rating Scale for ADHD Symptoms

Sarah Curran,1,2* J. Mill,1 P. Sham,1 F. Rijsdijk,1 K. Marusic,1 E. Taylor,1,2 and P. Asherson1
1Social, Genetic, and Developmental Psychiatry Research Centre, Institute of Psychiatry, London, United Kingdom
2Department of Child and Adolescent Psychiatry, GKT Medical School, London, United Kingdom

Current developments in molecular genetics have led to a rapid increase in research aimed at the identification of genetic variation that influences complex human phenotypes. One phenotype that has aroused a great deal of interest is the behavioral trait hyperactivity and the related clinical disorder attention-deficit hyperactivity disorder (ADHD). The driving force behind the molecular genetic research in this area is the overwhelming evidence from quantitative genetic studies that show high heritability (h^2 \approx 0.7–0.9) for the behaviors characterizing the diagnosis of ADHD, whether the disorder is viewed as a categorical entity or a continuous trait. To date, molecular studies have aimed at identifying susceptibility genes for ADHD, defined using operational diagnostic criteria, and have focused on variation within genes that regulate dopamine neurotransmission. Several studies report ADHD to be associated with the 7-repeat allele of a 48 bp repeat polymorphism (DRD4–7) in exon 3 of the dopamine D4 receptor gene (DRD4). In this study, we take a dimensional perspective of ADHD and examine the relationship of this DRD4 polymorphism in a sample of children selected from the general population on the basis of high and low scores on the five ADHD items of the Strengths and Difficulties Questionnaire (SDQ) as rated by their parents. We found a significant relationship between DRD4–7 and high-scoring individuals [chi-square = 8.63; P = 0.003; OR = 2.09 (95% CI 1.24 < OR < 3.54), F-statistic = 7.245; P = 0.008]. © 2001 Wiley-Liss, Inc.

KEY WORDS: QTL; quantitative trait; ADHD; DRD4 receptor gene, association study, childhood behavior

INTRODUCTION

Attention-deficit hyperactivity disorder (ADHD) is a disorder prevalent in 2%–5% of school-age children and clinically defined as the persistent behavior patterns of overactivity, inattentiveness, and impulsivity that are pervasive across social situations and accompanied by substantial social impairments. The long-term outcome of the categorically defined clinical disorder is poor [Taylor et al., 1996; Mannuzza and Klein, 2000] with an increased risk of social isolation and persistent psychopathology in adolescence and adulthood affecting up to 60% of cases [Barkley, 1996; Cantwell, 1996; Hill and Schoener, 1996]. However, there is some evidence that dimensional methods of representing these behavioral variations may prove to have even greater predictive validity [Fergusson et al., 1995]. Sex differences are probably overestimated in clinical samples, but more subtle differences do exist. In epidemiological samples, the male:female ratio is approximately 3:1, whereas in clinical samples the male:female ratio varies from 5–9:1, probably reflecting referral bias and comorbidity with other externalizing disorders such as conduct disorders and aggression [Gaub and Carlson, 1997].

Twin studies examining the relative importance of genetic and environmental influences have consistently shown ADHD to be among the most highly heritable behaviors in childhood, whether assessed dichotomously as a categorical disorder or dimensionally as a continuum, with heritability in the order of 70%–90% [Cantwell, 1972; Goodman and Stevenson 1989a, 1989b; Biederman et al., 1990, 1992; Gillis et al., 1992; Stevenson, 1992; Thapar, 1995; Gjone et al., 1996;
Levy et al., 1996, 1997; Silberg et al., 1996; Eaves et al., 1997; Sherman et al., 1997]. Although ADHD is diagnosed using operational criteria to define diagnostic categories, measures of activity and attention are continuously distributed in the general population and many studies have found a high level of correspondence between quantitative measures of hyperactivity and the categorical diagnosis [Edelbrock, 1986; Bird et al., 1987; Biederman et al., 1993, 1996; Chen et al., 1994; Boyle et al., 1997; Hudziak, 1997]. In these studies, there does not appear to be any obvious bimodality that separates ADHD children from non-ADHD children, suggesting that ADHD may be at one extreme of a quantitative dimension. In fact, twin studies have generally applied dimensional rating scales of hyperactivity, with clinical cutoffs applied when diagnostic categories were required. These studies all show high heritabilities regardless of where these cutoffs have been made and regardless of whether diagnostic or continuous criteria have been applied. Furthermore, mathematical modeling applied to such studies supports the hypothesis that ADHD is the extreme of a behavior that varies genetically throughout the entire population [Gillis et al., 1992; Stevenson, 1992; Levy et al., 1997]. Taken together, these findings suggest strongly that a dimensional perspective on hyperactivity is a valid approach to the identification of quantitative measures and the categorical diagnosis [Edelbrock, 1986; Bird et al., 1987; Biederman et al., 1993, 1996; Chen et al., 1994; Boyle et al., 1997; Hudziak, 1997]. In these studies, there does not appear to be any obvious bimodality that separates ADHD children from non-ADHD children, suggesting that ADHD may be at one extreme of a quantitative dimension. In fact, twin studies have generally applied dimensional rating scales of hyperactivity, with clinical cutoffs applied when diagnostic categories were required. These studies all show high heritabilities regardless of where these cutoffs have been made and regardless of whether diagnostic or continuous criteria have been applied. Furthermore, mathematical modeling applied to such studies supports the hypothesis that ADHD is the extreme of a behavior that varies genetically throughout the entire population [Gillis et al., 1992; Stevenson, 1992; Levy et al., 1997]. Taken together, these findings suggest strongly that a dimensional perspective on hyperactivity is a valid approach to the identification of quantitative trait loci (QTLs) and should provide a powerful complementary strategy to the study of clinically defined subtypes.

Though the evidence for a large genetic component in ADHD has come from twin studies based on questionnaire-derived symptom scores, molecular genetic studies have to date focussed on ADHD diagnosed using categorical criteria. As far as we are aware, there have been no published studies taking a QTL perspective for the ascertainment and selection of cases. QTL methods are statistically powerful approaches, whereby continuous traits may be genetically mapped to multiple discrete chromosomal locations (known as QTLs), by examining the relationship between genetic markers and the relative amount of a given trait. While originally developed to study quantitative phenotypes in plants and animals, the approach has been successfully applied to continuous traits in humans. Recent examples in rodents include the localization of loci that contribute to hypertension, diabetes, obesity, and atherosclerosis [Aitman et al., 1999; Mu et al., 1999], as well as behavioral phenotypes such as alcohol consumption [Crabbe et al., 1999] and learning [Caldarone et al., 1997]. Moisan et al. [1996] mapped a major QTL to rat chromosome 8 for hyperactivity scores. In humans, chromosomal synteny between the mouse model and humans was used to map a gene for the complex trait of obesity, and QTL linkage scans have shown promising results for type 2 diabetes mellitus and hypertension [Lembertas et al., 1997; Hsueh et al., 2000; Watanabe et al., 2000]. Applications of QTL techniques have been used to study reading ability [Fisher SE et al., 1998, 1999; Gayan et al., 1999] and intelligence [Fisher PJ et al., 1999; Hill et al., 1999] and are likely to be useful in conditions such as ADHD, in which quantitative measures of the phenotype can be applied.

We therefore set out in this study to establish whether the reported association between the 7-repeat allele of the variable number tandem repeat polymorphism in exon 3 of DRD4 (DRD4–7) and diagnosed cases of ADHD would replicate using a QTL approach in an epidemiological sample. A dopamine hypothesis of ADHD and hyperactivity has long been held, based first on studies of the neurochemistry of attention in animals and humans and secondly on the dramatic amelioration of ADHD symptoms by stimulant medications such as methylphenidate, whose main mode of action is via the dopamine system, though their specific site of action is unclear [Shawwitz et al., 1977; Shekim et al., 1979; Shen and Wang 1984; Zanetkin et al., 1984; McCracken, 1991; Goldman-Rakic, 1992; Castellanos, 1997]. It is therefore of considerable interest that genetic studies from several independent groups have found evidence for association and linkage with dopamine system genes, including the D4 and D5 receptor genes (DRD4 and DRD5) and the dopamine transporter gene (DAT1) in samples of individuals meeting diagnostic criteria for ADHD. The most robust of these findings so far is an association between ADHD and DRD4–7, with nine studies providing evidence for this association [LaHoste et al., 1996; Rowe et al., 1998; Smalley et al., 1998; Swanson et al., 1998; Faraone et al., 1999; Holmes et al., 2000; Muggia et al., 2000; Tahir et al., 2000; Mill et al., 2001; Sunohara et al., 2000], whereas four found no association [Castellanos et al., 1998; Eisenberg et al., 2000; Hawi et al., 2000; Kotler et al., 2000]. A recent meta-analysis of published and unpublished data suggests a respectable but modest odds ratio of 1.9 (95% CI = 1.4–2.2; $P = 0.00000008$) from 7 case control studies and 1.4 (95% CI = 1.1–1.6; $P = 0.02$) from 14 family-based studies [Faraone et al., 2001]. The functional significance of the DRD4 polymorphism is still uncertain [Asghari et al., 1995; Jovanovic et al., 1999; Watts et al., 1999; Kazmi et al., 2000]. It remains possible that the 7-repeat is in linkage disequilibrium (LD) with a functional variant, which may confound association studies and lead to discrepant association findings. The possibility that functional promoter polymorphisms explain the findings has been suggested, but is perhaps unlikely following negative findings for three such polymorphisms, in a series of trio samples that showed linkage and association to DRD4–7 [Barr et al., 2001].

**MATERIALS AND METHODS**

**Sample Ascertainment**

In this study, we screened an epidemiological sample of 5- to 15-year-old children from southern England, using the parent version of the Strengths and Difficulties Questionnaire (pSDQ) [Goodman, 1997]. The SDQ was chosen as our primary measure for several reasons. It is a brief 30-item questionnaire that can be easily administered by post to the parents of individuals in the age group selected. The SDQ is currently being used in several epidemiological surveys and has shown good...
predictive and discriminative validity against psychiatric diagnoses and good concurrent validity with the Rutter and Achenbach Screening Questionnaires. Furthermore, parent-rated SDQ scores of hyperactivity are known to be highly heritable [Martin, 1999]. Parent-rated SDQ has a sensitivity of 33.3% (5–10 years) and 45% (11–15 years) to predict any hyperkinetic disorder according to ICD-10, and a sensitivity of 29.9% (5–10 years) and 41% (11–15 years) to predict any ADHD disorder according to DSM-IV [Goodman et al., 2000]. The sensitivities for a combination of parent- and teacher-rated questionnaires to predict these clinical disorders are much higher, ranging from 75% to 85%, with predictions from teacher-rated questionnaires alone marginally better than parents alone.

Children were recruited to this study by contacting parents who were taking part in an unrelated study being carried out at the SGDP Research Centre. A large database of adults unselected for any phenotype had been established through a collaborative project with the Medical Research Council–funded GP Research framework [Sham et al., 2001]. The ethnicity of this population is 95% white Caucasian. Adults were contacted if they indicated they had children and agreed to take part in another study. Through this database, 12,000 adults were identified, asked whether they agreed to participate in the study, and, if so, to complete pSDQs on each of their children aged 5–15 years. The response rate was 30%, with 33% of responders willing to participate in this study. A total of 3,097 completed pSDQ scores were returned.

**Sample Selection**

Obtaining adequate statistical power to detect linkage to or association with genes for complex disorders can be difficult. One power-enhancing strategy that considers restraints such as genotyping and has been adopted in several QTL studies is the selective sampling of phenotypically extreme individuals [Risch and Zang 1995; Van Gestel et al., 2000]. However, for a given population size, optimal selection criteria must take into account the range of possible underlying genetic models and balance out the effects of extreme selection with the requirement for large sample size. Furthermore, simulations by Allison et al. [1998] show that the ceteris paribus assumption (i.e., the more extreme the selection, the more power one achieves) is not always true, in particular where biallelic QTLs have relatively symmetric allele frequencies and smaller mean displacement among genotypes. Using simulations, we have investigated the optimum selection criteria for QTL association studies under a range of likely genetic models and find that a 20%–27% cutoff at both ends of the distribution provides the best solution (data not shown).

In our study, we decided a priori to select cases for DNA studies and further phenotypic measurement, using thresholds of the 10-point scale derived from the five ADHD items within the pSDQ. High scorers included all individuals scoring 7 and above out of the possible 10, and low scorers included those scoring 0 or 1. Teacher-rated SDQ questionnaires (tSDQ) were requested for selected children and mouth swabs taken for DNA extraction.

The frequency distribution for the five ADHD items from the pSDQ, for the total sample of 3,097 individuals, is shown in Figure 1. The frequencies of the various scores in this study are compatible with the results of a large epidemiological sample of approximately 10,300 U.K. children in the same age range [Goodman et al., 2000]. As can be seen, 17% of the sample scoring 7/10 and above were selected as high scorers, and 31% scoring less than or equal to 1/10 were low scorers. In total, DNA was obtained from 224 individuals (133 high scorers and 91 low scorers) for genotyping studies. Among the high scorers, the male:female ratio was 2:1, which is much lower than the male:female ratio found in clinical samples. In
contrast, the ratio was reversed for low scorers on the pSDQ scorers with male:female of 2:3.

DNA Collection and Amplification of DRD4 Exon 3 VNTR

DNA was collected by mail using cotton wool buds and two 15 ml tubes containing 2.5 ml of a storage/preservative solution [STE buffer (100 mM NaCl, 10 mM Tris-HCl, pH 8, 10 mM EDTA, pH 8), with 0.2 mg/ml proteinase K and 0.5% SDS] [Freeman et al., 1997]. The average yield per individual was 120 μg DNA.

The exon 3 VNTR was amplified with an initial 5-min denaturing step at 95°C, followed by 35 cycles of 93°C for 1 min, 55°C for 1 min, and 72°C for 1 min, and a final extension phase of 72°C for 10 min. Primers used were 5'GGTCTGCGGTGGAGTCTG-3' and 5'-GGCATACTGGGTTCTACT-3'. Reactions were performed in 22 μl volumes and included 50 ng of genomic DNA, 1.5 mM MgCl₂, 0.2 mM dNTPs (incorporating a 50/50 deaza dGTP/dGTP mix), 10% DMSO, 10 mM GeneAmp 10× PCR Gold Buffer (PE Applied Biosystems), and 1 unit of AmpliTaq Gold (PE Applied Biosystems). PCR products were run out on a 2% agarose gel stained with ethidium bromide and analyzed under UV light. Homozygous genotypes were repeated if clear and strong bands were not observed. The ability of this protocol to detect the long 7-repeat allele in heterozygotes, which shows marked differential amplification with the common 2-, 3-, and 4-repeat alleles, has been examined in our laboratory by comparison with fluorescently tagged products visualized on an ABI 310 (PE Applied Biosystems) and found to be sufficiently sensitive. Each individual was genotyped twice, and discrepancies were rechecked with two further PCRs.

Association Analysis With pSDQ Data

DRD4 allele frequencies for the high- and low-scoring groups are shown in Table I. In the comparison of the two groups, we found a significant difference between the frequencies of DRD4 alleles (chi-square = 14.37; df = 5; P = 0.013). We then tested the specific hypothesis suggested by previous studies, i.e., DRD4–7 is the associated allele, by comparing the frequency of the 7-repeat allele to the other alleles combined [chi-square = 8.63; df = 1; P = 0.003; OR = 2.09 (95% CI 1.24 < OR < 3.54)]. Genotype frequencies for the 7-repeat versus the rest were found to be in Hardy-Weinberg equilibrium (chi-square = 1.28; df = 1; P = 0.24).

Further analysis performed using simple linear regression in SPSS, with the presence of the 7-repeat allele as the independent variable and pSDQ scores as the dependent variable, was also significant (F-statistic = 7.245; P = 0.008; standardized coefficient for the 7-repeat allele as beta = 0.178; t = 2.692; P = 0.008). Controlling for sex as a confounding variable, the P value was 0.009 and showed that female sex had a slightly negative influence with a standardized coefficient of −0.262. Using all alleles in a stepwise multiple regression, the only solution found to be significant was with DRD4–7 alone, rather than in combination with any of the other alleles.

We used a maximum-likelihood method implemented in Mx [Neale, 1997] to estimate the correlation between pSDQ scores in the complete data set of 3,000 individuals, with the presence of DRD4–7. This allowed for more accurate estimation of the correlation taking into account variance differences in measures due to selection on pSDQ. We found the correlation coefficient (r) = 0.13, and therefore the variance (r²) in pSDQ scores accounted for by the association with DRD4–7 to be 2%.

Association Analysis With Parent and Teacher SDQ Data

A maximum-likelihood method was again employed to estimate the correlation between pSDQ and tSDQ ADHD scores using MX (r = 0.48). Parent and teacher SDQ data were available on a total of 143 out of the 226 individuals that were genotyped. The variance of the parent scores (1.67) was not greatly different from that of the teacher scores (1.74) and therefore the same weighting was given for both scores.

Taking individuals scoring ≥ 7/10 on the pSDQ ADHD scale and ≥ 4/10 on the tSDQ ADHD scale (we were not stringent here as we had selected the sample on parent scores), we categorized individuals into pervasively high or pervasively low (pSDQ ≤ 1; tSDQ < 4) groups. There were 51 individuals in the high pervasive category and 57 individuals in the low pervasive category. The association of DRD4–7 with scores in the high pervasive versus low pervasive subgroups did not reach nominal significance levels, but showed a trend in the expected direction [chi-square = 2.64; P = 0.1; OR = 1.78 (95% CI 0.84 < OR < 3.82)] and gave rise to DRD4–7 frequencies that were proportionally similar to those in the total pSDQ sample (Table II).

**DISCUSSION**

With consistent evidence from the field of behavior genetics that ADHD has an important genetic component, there is now much international interest in trying to identify susceptibility genes for the disorder.

<table>
<thead>
<tr>
<th>TABLE I. DRD4 Allele Counts and Frequencies (in Brackets) in High and Low Scorers on pSDQ ADHD Items*</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-repeat</td>
</tr>
<tr>
<td>3-repeat</td>
</tr>
<tr>
<td>4-repeat</td>
</tr>
<tr>
<td>5-repeat</td>
</tr>
<tr>
<td>7-repeat</td>
</tr>
<tr>
<td>8-repeat</td>
</tr>
<tr>
<td>High-score group</td>
</tr>
<tr>
<td>16 (16.0%)</td>
</tr>
<tr>
<td>8 (3.0%)</td>
</tr>
<tr>
<td>170 (63.9%)</td>
</tr>
<tr>
<td>1 (0.4%)</td>
</tr>
<tr>
<td>70 (26.3%)</td>
</tr>
<tr>
<td>1 (0.4%)</td>
</tr>
<tr>
<td>Low-score group</td>
</tr>
<tr>
<td>18 (9.9%)</td>
</tr>
<tr>
<td>8 (4.4%)</td>
</tr>
<tr>
<td>124 (68.1%)</td>
</tr>
<tr>
<td>2 (1.1%)</td>
</tr>
<tr>
<td>26 (14.3%)</td>
</tr>
<tr>
<td>4 (2.2%)</td>
</tr>
</tbody>
</table>

*The high group score ≥ 7, the low group ≤ 1 on pSDQ. No 6-repeats were found in this sample.
Although molecular studies to date have focused on clinical samples of ADHD, there is substantial support for a dimensional perspective on ADHD, viewing it as the extreme of a continuously distributed trait, and therefore good reason to adopt potentially powerful quantitative trait methods. In this study, we provide further support for the QTL approach by replicating the finding of association between DRD4–7 and clinical ADHD, in a sample of children selected on the basis of high (top 17%) and low (bottom 31%) parent-rated ADHD scores, from a population sample unselected for phenotype. The selection criteria used approximations to the cutoffs suggested by statistical simulations to maximize power. The significance of the result was similar using a simple chi-square test to compare the low- and high-scoring groups as for a linear regression test using continuous trait scores (though of course not all data from the screened population were available for this analysis).

The main limitations to this study are, first, that it is difficult to rule out population stratification as a reason for the positive findings. Although our genotypes are in Hardy-Weinberg equilibrium, we cannot systematically eliminate the possibility of recent population admixture. Secondly, the sample is selected on the basis of scores on a parent scale that only moderately predicts clinical ADHD, even though it does show high heritability, and thus the significance of the findings to the clinical disorder remains uncertain. However, combining parent and teacher ratings gives improved prediction of the clinical disorder. It is therefore notable that although the association did not reach nominal significance when parent and teacher scores were combined \( P = 0.1 \), the size of the effect observed in the subsample used for this analysis was of the same order found for the analysis of pSDQ data alone (pervasive subsample being 50% of total sample). A power calculation showed that in the pervasive subsample, we only had 80% power to replicate a genotypic association with an odds ratio of 4, at an alpha level of 0.05. Taken together, these findings and the previous well-replicated association between DRD4–7 and clinical ADHD support the notion that our QTL approach has the ability to detect a finding that is clinically relevant.

Although the QTL approach to ADHD appears to be a promising one, there remain some problem areas that we need to concern ourselves with at the outset. Phenotypic measurement remains a controversial issue. The task here is considerably more challenging than a simple measurement of blood pressure to investigate QTLs for hypertension. Findings from behavioral genetic studies show considerable variation depending on the source of information (parent, teacher, individual), or the type and number of assessment instruments, all of which have modest correlations.

Two recent sets of twin data exemplify the important issues. Thapar et al. [2000] found heritability for ADHD considered as a dimensional trait to be considerably higher when estimated by parent ratings on the Rutter scale \( (0.84) \) than when estimated by parent ratings on the Du Paul scale \( (0.47) \). These results contrasted with another [Martin, 1999] recent twin study, in which heritability estimates for parent ratings of ADHD items from the SDQ were 0.58, while those for parent ratings on the short Conners scale were 0.78. In this study, bivariate analysis of parent SDQ and Conners ratings showed a large shared genetic component (50%) between the two rating scales, and that all of the genetic variance for the SDQ scale was included in this, but that the Conners scale detected a further 24%. A similar result was found for teacher ratings on the two questionnaires, suggesting that for ADHD, scales containing more items capture a greater proportion of genetic variance. Both studies appear to show reduced or absent contrast effects with the longer Du Paul and Conners rating scales. Thapar et al. [2000] formally tested for this and showed rater contrast effects when using the Rutter, but not the Du Paul scale (apart from inattention items alone on the Du Paul).

The use of questionnaire rating scales versus clinical interviews in behavior genetic studies may introduce further variation; questionnaires belonging to the psychometric tradition that assumes disorders are extremes on the same underlying continuum that describes variation in the normal range, whereas clinical interviews derive from the medical tradition for categorizing disorders. Genetic and environmental effects might differ, because the individuals fitting into the clinical extremes may have a different balance of environmental and genetic risk factors.

Finally, behavior genetic studies of ADHD have in general used parent ratings alone. However, estimates of heritability that combine parent and teacher ratings [Thapar et al., 2000] or use structural equation modeling to create a latent trait from multiple raters and rating scales [Simonoff et al., 1998] are as high or better (in the case of the latent trait) than parent- and teacher-rated scales alone. Given that parent-only-rated scales predict a clinical diagnosis of ADHD much less often than when combined with a teacher rating and are more subject to rater bias, the most robust method may be to use multiple measures and informants to create a heritable latent trait.

In conclusion, we have replicated the DRD4–7 association with ADHD using a QTL approach in a general population sample. This suggests that DRD4–7 or a functional variant in LD with that allele is a QTL that exerts influence over the range of hyperactivity scores in a population, rather than a risk factor for narrow diagnostic subtypes. We further suggest that these findings support the use of QTL approaches to the identification of novel QTLs for ADHD. With the problem of nonreplication of association findings already evident, we suggest that a unified approach to phenotypic measurement should be considered.
ACKNOWLEDGMENTS

This research was funded by a Wellcome Trust Training Fellowship (to S.C.).

REFERENCES


