Genetic Heterogeneity in ADHD: DAT1 Gene Only Affects Probands Without CD

Kaixin Zhou,1 Wai Chen,1 Jan Buitelaar,2 Tobias Banaschewski,3,4 Robert D. Oades,5 Barbara Franke,2,6 Edmund Sonuga-Barke,7 Richard Ebstein,8 Jacques Eisenberg,9 Michael Gill,10 Iris Manor,8 Ana Miranda,11 Fernando Mulas,11 Herbert Roeyers,12 Aribert Rothenberger,9 Joseph Sergeant,13 Hans-Christoph Steinhausen,14 Jessica Lasky-Su,15,16,17 Eric Taylor,1 Keesley J. Brooks,8 Xiaohui Xu,1 Benjamin M. Neale,1,18,19 Fruhling Rijndijk,1 Margaret Thompson,7 Philip Asherson,1 and Stephen V. Faraone15,16,17*

1MRC Social Genetic Developmental and Psychiatry Centre, Institute of Psychiatry, London, United Kingdom
2Department of Psychiatry, Radboud University Nijmegen Medical Center, Nijmegen, The Netherlands
3Child and Adolescent Psychiatry, University of Göttingen, Germany
4Central Institute of Mental Health, Department of Child and Adolescent Psychiatry and Psychotherapy, Mannheim, Germany
5University Clinic for Child and Adolescent Psychiatry, Essen, Germany
6Department of Human Genetics, Radboud University Nijmegen Medical Center, Nijmegen, The Netherlands
7School of Psychology, University of Southampton, Highfield, Southampton, United Kingdom
8ADHD Clinic, Gehr Mental Health Center, Petah Tikvah, Israel
9S. Herzog Memorial Hospital, Jerusalem, Israel
10Department of Psychiatry, Trinity Centre for Health Sciences, St. James’s Hospital, Dublin, Ireland
11Department of Developmental and Educational Psychology, University of Valencia, Spain
12Ghent University, Ghent, Belgium
13Department of Clinical Neuropsychology, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands
14Department of Child and Adolescent Psychiatry, University of Zurich, Switzerland
15Department of Psychiatry, SUNY Upstate Medical University, Syracuse, New York
16Department of Neuroscience, SUNY Upstate Medical University, Syracuse, New York
17Department of Physiology, SUNY Upstate Medical University, Syracuse, New York
18Broad Institute of Harvard and MIT, Cambridge, Massachusetts
19Center for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts

Previous studies have found heterogeneous association between DAT1-3′-UTR-VNTR and attention deficit hyperactivity disorder (ADHD). Various proportions of conduct disorder (CD) comorbidity in their ADHD samples may partially explain the observational discrepancies. Evidence for this comes from family and twin studies which found ADHD probands with CD (ADHD + CD) are genetically different from those without CD (ADHD − CD). Genotypes of 20 DAT1 markers were analyzed in 576 trios, consisting of 141 ADHD + CD and 435 ADHD − CD. In addition to the classical TDT test, a specific genetic heterogeneity test was performed to identify variants that have different transmission patterns in the two phenotypic subgroups. After multiple-test correction, rs40184 and rs2652511 were significant in TDT tests. Further heterogeneity test found the two SNPs had a significant transmission pattern difference between ADHD + CD and ADHD − CD children, indicating that DAT1 has a significantly greater genetic influence on ADHD without CD. Although the result needs further replications, it does highlight the importance of selecting genetically homogeneous samples for molecular genetic analyses of ADHD.

© 2007 Wiley-Liss, Inc.

KEY WORDS: comorbidity; stratification; association


INTRODUCTION

Family, twin, and adoption studies have consistently shown that attention deficit hyperactivity disorder (ADHD) is highly heritable (>75%) with population prevalence estimated at 4–8% [Asherson, 2004; Heiser et al., 2004; Faraone et al., 2005;
Thapar et al., 2005; Waldman and Gizer, 2006). However, there remains considerable uncertainty about the more specific etiology of ADHD. The rationale for early efforts to identify ADHD susceptibility genes using candidate gene approaches were largely guided by the fact that ADHD children responded favorably to stimulants such as methylphenidate and dexamphetamine [Spencer et al., 1996]. Based on the fact that these medications inhibit the dopamine transporter, the dopamine transporter gene (DAT1 or SLC6A3) became a target of investigation [Cook et al., 1995]. During the last decade, mixed positive and negative findings emerged from a large number of independent studies exploring the association between DAT1 and ADHD. The most intensively tested marker, a 40-bp variable number of tandem repeat (VNTR) located in the 3'-untranslated region (3'-UTR) was shown to have a small (OR = 1.13, 95% CI (1.03–1.31)) but significant effect in a meta-analysis that pooled family based studies together [Faraone et al., 2005]. More recently, however, a further meta-analysis of DAT1 that included a more comprehensive analysis of available studies, found no overall evidence for association but significant evidence of heterogeneity between datasets [Li et al., 2006]. The reason for the observed heterogeneity is currently unknown.

One possible explanation for such heterogeneity is that the 40-bp VNTR is not the functional variant causing susceptibility and may tag one or more causal variants to different degrees from sample to sample. Evidence for this comes from the analysis of an additional 30-bp VNTR in intron 8 that showed association of ADHD to a specific haplotype of the two VNTR markers in UK and Taiwanese samples, as well as the IMAGE project sample (pooled OR = 1.4, P = 6 x 10^-8) [Asherson et al., 2007]. The low resolution obtained by using only one marker may partially explain the inconsistent findings. Another possible source of inter-study heterogeneity is of a phenotypic nature. Like other psychiatric disorders, ADHD shows substantial phenotypic heterogeneity associated with different DSM-IV subtypes [Faraone et al., 2000a], the co-existence of various comorbid traits [Thapar et al., 2006], as well as variability in the persistence of symptoms across the lifespan [Faraone and Tsuang, 2001; Faraone et al., 2006].

Some of this heterogeneity is likely to have its origin in genetic differences between individuals [Faraone et al., 2000c]. In the current article, we set out to test whether the pattern of inter-study variability in the ADHD–DAT1 association might be explained, at least in part, by heterogeneity within ADHD samples in relation to comorbid conduct disorder (CD).

The comorbidity between ADHD and CD has been observed consistently for many years. Taylor noted that the co-morbidity between ADHD and CD raises a key nosological question: are these two disorders best seen as occurring yet separate entities, or does their consistent co-occurrence signal the existence of some genetic effects that were unique to ADHD, as well as shared environmental effects on the risk for conduct problems but not ADHD. In recent psychophysiological studies of event-related potentials and task performance, the ADHD + CD group was found to be psychophysio logically less deviant or impaired than either pure ADHD group or pure CD group indicating that ADHD + CD might represent a separate disorder distinct from ADHD – CD [Banaschewski et al., 2003; Albrecht et al., 2005].

Driven by these compelling data, we hypothesize that comorbidity with CD partially accounts for the inter-study variability in the ADHD–DAT1 association. In the current study, we set out to test this hypothesis by reanalyzing data from a previous candidate gene analysis utilizing data from the IMAGE project which provided nominal association evidence at the DAT1 locus [Brookes et al., 2006].

**MATERIALS AND METHODS**

**Subjects and Genotyping**

In this study, we used data from our previous analysis of 51 candidate genes in which the association with DAT1 reached gene-wide significance and suggested the existence of two independent associations at the 3'- and 5'-ends of the gene [Brookes et al., 2006]. The original nuclear family sample was stratified into independent ADHD + CD and ADHD – CD trios. This allowed us to formally carry out a standard chi-square test of genetic heterogeneity based on the transmission patterns in both groups.

European Caucasian subjects were recruited from 12 specialist clinics in eight countries: Belgium, Germany, Holland, Ireland, Israel, Spain, Switzerland, and United Kingdom. Ethical approval for the study was obtained from National Institute of Health registered ethical review boards for each center. All ADHD probands and their siblings were aged 5 to 17 at the time of entry into the study and access was required to one or both biological parents for DNA collection. Entry criteria for probands were a clinical diagnosis of DSM-IV combined subtype ADHD and having one or more full siblings available for ascertainment of clinical information and DNA collection. Exclusion criteria applying to both probands and siblings included autism, epilepsy, IQ < 70, brain disorders, and any genetic or medical disorder associated with externalizing behaviors that might mimic ADHD. For the analyses completed here, we selected the subset of 976 proband–parent trios for whom we had complete genotype data for markers across the DAT1 gene.

The research diagnosis of DSM-IV ADHD and CD was reached following completion of a parent interview with the parent account childhood symptoms (PACS) [Taylor et al., 1986]. This is a semi-structured, standardized, investigator-based interview developed as an instrument to provide an objective measure of child behavior. For the diagnosis of ADHD, a standardized algorithm was applied to PACS data to derive each of the 18 DSM-IV ADHD items, providing operational definitions for each behavioral symptom. These were
then combined with items that scored 2 or 3 from the teacher rated Conners’ ADHD subscale, to generate the total number of items from the DSM-IV symptom checklist. The definition of situational pervasiveness required that some symptoms occurred within two or more different situations from the PACS interview, or the presence of one or more symptoms scoring 2 or more from the ADHD subscale of the teacher rated Conners. The diagnosis of DSM-IV CD was based on the PACS data alone. All the CD diagnosis was made without the existence of oppositional defiant disorder (ODD). Of the 576 probands included in this analysis, 435 were diagnosed without CD (ADHD – CD) and 141 with CD (ADHD + CD).

As shown in Table I, there is no significant age or sex difference between the two phenotypic subgroups. Table II describes the demographic distribution of the 576 trios. The ADHD + CD proband incidence rates are between 9 and 45% across our study sites. Using Pearson’s chi-squared test of equal incidence across centers, we found the incidence rates significantly different between IMAGE centers through 10,000 simulations (P = 0.0002), suggesting a high degree of inter-site phenotypic heterogeneity in terms of CD.

We selected 18 common SNPs (minor allele frequency > 0.05) that had been successfully genotyped at the DAT1 locus in the original study. Genotypes from two VNTR markers including the most intensively investigated 3′-UTR 40-bp VNTR and the intron 8 30-bp VNTR were also included in this analysis [Asherson et al., 2007].

Statistical Tests

The Transmission Disequilibrium Test (TDT) was used to test for single marker associations [Ewens and Spielman, 2005]. For haplotype-specific TDT, the haplotype counts were obtained by summing the EM-estimated fractional likelihoods of each individual as implemented in Haploview [Barrett et al., 2005]. In addition, we used ETDT to test the multi-allele haplotype transmission disequilibrium [Sham and Curtis, 1995]. All the association tests were performed in each phenotypic subgroup separately and then in the full sample.

To detect genetic heterogeneity between the two subgroups, we used a standard chi-square test. Under the null hypothesis of no genetic heterogeneity, a variant is expected to have the same effect in both clinical subgroups, so that an allele of the marker will be expected to have the same transmission to non-transmission ratio (T/NT) from heterozygotic parents in each subgroup separately and then in the full sample.

To detect genetic heterogeneity between the two subgroups, we used a standard chi-square test. Under the null hypothesis of no genetic heterogeneity, a variant is expected to have the same effect in both clinical subgroups, so that an allele of the marker will be expected to have the same transmission to non-transmission ratio (T/NT) from heterozygotic parents in each subgroup separately and then in the full sample.

Table III summarizes the single marker TDT test results in each phenotypic subgroup and the full sample. The two VNTRs only showed marginal associations (P < 0.1) in both the ADHD – CD group and the full sample. Five SNPs (rs40184, rs2652511, rs11564750, rs10070282, rs2550946) showed nominally significant (P < 0.05) associations in both the ADHD – CD group and the full sample. In particular, rs40184 (P = 0.0004) from the 3′-end LD block in DAT1 and rs2652511 (P = 0.002) from the 5′-end LD block were still significant in the ADHD – CD group after the stringent Bonferroni multiple test correction (P < 0.0025). More importantly, because these two SNPs locate in separate LD blocks with D’ < 0.15 between them (95% CI: 0.09–0.2), we can view them as independent association signals at the DAT1 locus. Strikingly, none of the markers showed any association signal within the ADHD + CD group.

Single Marker Genetic Heterogeneity Test

The genetic heterogeneity tests between ADHD + CD and ADHD – CD are also summarized in Table III. The two significant SNPs (rs40184, rs2652511) from the association test in the ADHD – CD group also showed significant genetic heterogeneity between groups (P = 0.016 and P = 0.023). In addition, rs27072 showed significant heterogeneity (P = 0.034), but has only a marginal association with the ADHD – CD group (P = 0.06) and no association with the ADHD + CD group (P = 0.168). Owing to the multiple LD block structures at the DAT1 locus and the heterogeneity signals seen in the two blocks, we opted to perform simulations to estimate the statistical significance of observing two such independent signals. In each simulation, only the proband diagnosis was permuted to preserve the LD structure. The heterogeneity test was performed on each permuted data set and the highest statistic from each LD block was collected. From 10,000 simulations, 181 were observed with multiple significant heterogeneity signals (P < 0.05) from different LD blocks, simultaneously. Thus the empirical significance of our finding is P = 0.018.

RESULTS

Single Marker Association Analysis

The 18 SNP and two VNTR markers analyzed here covered the whole DAT1 genomic region with an average density of 1 marker/2.85 kb. Figure 1 shows the pairwise linkage disequilibrium (LD) measure D’ in the founders as estimated by GOLD [Abecasis and Cookson, 2000]. The three-block structure pattern is consistent with that identified in a recent comprehensive study of the DAT1 genomic region [Greenwood et al., 2006]. Both VNTR markers fall into the 3′-end block with the 40-bp VNTR in tight LD with other SNPs and the 30-bp VNTR at the boundary of this block. This pattern suggests that the previous association studies using these two VNTRs can only pick up association signals from the 3′-end of the DAT1 locus.

Table III summarizes the single marker TDT test results in each phenotypic subgroup and the full sample. The two VNTRs only showed marginal associations (P < 0.1) in both the ADHD – CD group and the full sample. Five SNPs (rs40184, rs2652511, rs11564750, rs10070282, rs2550946) showed nominally significant (P < 0.05) associations in both the ADHD – CD group and the full sample. In particular, rs40184 (P = 0.0004) from the 3′-end LD block in DAT1 and rs2652511 (P = 0.002) from the 5′-end LD block were still significant in the ADHD – CD group after the stringent Bonferroni multiple test correction (P < 0.0025). More importantly, because these two SNPs locate in separate LD blocks with D’ < 0.15 between them (95% CI: 0.09–0.2), we can view them as independent association signals at the DAT1 locus. Strikingly, none of the markers showed any association signal within the ADHD + CD group.

Single Marker Genetic Heterogeneity Test

The genetic heterogeneity tests between ADHD + CD and ADHD – CD are also summarized in Table III. The two significant SNPs (rs40184, rs2652511) from the association test in the ADHD – CD group also showed significant genetic heterogeneity between groups (P = 0.016 and P = 0.023). In addition, rs27072 showed significant heterogeneity (P = 0.034), but has only a marginal association with the ADHD – CD group (P = 0.06) and no association with the ADHD + CD group (P = 0.168). Owing to the multiple LD block structures at the DAT1 locus and the heterogeneity signals seen in the two blocks, we opted to perform simulations to estimate the statistical significance of observing two such independent signals. In each simulation, only the proband diagnosis was permuted to preserve the LD structure. The heterogeneity test was performed on each permuted data set and the highest statistic from each LD block was collected. From 10,000 simulations, 181 were observed with multiple significant heterogeneity signals (P < 0.05) from different LD blocks, simultaneously. Thus the empirical significance of our finding is P = 0.018.
ADHD groups. This was due to significant over-transmission in the square test of genetic heterogeneity between the two sub-
sission in the ADHD/C0 (35.3%) is significantly over-transmitted in the ADHD group (haplotype specific \( P = 0.0009 \)). On the contrary, the most common GG haplotype (35.3%) is significantly under-transmitted in the ADHD – CD group (haplotype specific \( P = 1.6 \times 10^{-6} \)). The AA haplotype showed a moderate protective effect. In a 3-df likelihood test of transmission disequilibrium, the two locus haplotype trans-
mission is significantly distorted in the ADHD – CD group (\( P = 8.5 \times 10^{-5} \)). Finally, the transmission patterns of the GG haplotype were significantly different (\( P = 0.0009 \)) in the chi-
square test of genetic heterogeneity between the two sub-
groups. This was due to significant over-transmission in the ADHD – CD subgroup and non-significant under-transmission in the ADHD + CD subgroup.

### DISCUSSION

In this study, TDT tests in ADHD + CD and ADHD – CD subgroups gave very different results. Two SNPs (rs40184 and rs2652511) from the DAT1 gene were significantly associated with ADHD – CD after the application of the stringent Bonferroni correction for multiple tests. In the haplotype-specific TDT test at these two loci, a highly significant (\( P = 1.6 \times 10^{-6} \)) association was observed between the GG haplotype and the ADHD – CD phenotype. In contrast, none of the 20 markers were associated with ADHD + CD even without correction. When the two phenotypic subgroups were combined, there were only nominally significant signals, but none of them remained after multiple test correction. One potential explanation for this observation is that age or gender difference between the subgroups. However, as shown in Table I, neither age nor sex is significantly different between the two subgroups. Another possible explanation might be the genetic mechanisms associated with varying levels of comorbid CD across samples. We applied a formal statistical test to address whether observing the different association results in ADHD + CD and ADHD – CD was due to genetic heterogeneity or simply occurred by chance. The test statistic was calculated with heterozygous parents’ allele T/NT to ensure it was robust for population stratification. Our single marker genetic heterogeneity tests identified significant SNPs in two LD blocks at the DAT1 locus. Further permutation results suggested the chance of observing such a pattern of genetic heterogeneity under the LD background as being very low (\( P = 0.018 \)). Thus, we are confident that the significance of the subgroup difference in the transmission pattern really indexes the impact of phenotypic heterogeneity. Considering the fact that a significant association was only found within the ADHD – CD subgroup, our data may further indicate that variation in DAT1 is only implicated in the genetic etiology of ADHD children who do not have CD.

### Haplotype Analysis

Table IV presents haplotype-specific TDT tests of the two significant SNPs (rs40184, rs2652511) in both subgroups, and the heterogeneity test. As expected, all four haplotype-specific TDT tests were non-significant in the ADHD + CD subgroup (\( P > 0.1 \)). On the contrary, the most common GG haplotype (35.3%) is significantly over-transmitted in the ADHD – CD group (haplotype specific \( P = 1.6 \times 10^{-6} \)). The AA haplotype showed a moderate protective effect. In a 3-df likelihood test of transmission disequilibrium, the two locus haplotype trans-
mission is significantly distorted in the ADHD – CD group (\( P = 8.5 \times 10^{-5} \)). Finally, the transmission patterns of the GG haplotype were significantly different (\( P = 0.0009 \)) in the chi-
square test of genetic heterogeneity between the two sub-
groups. This was due to significant over-transmission in the ADHD – CD subgroup and non-significant under-transmission in the ADHD + CD subgroup.

<table>
<thead>
<tr>
<th>Study site</th>
<th>ADHD – CD</th>
<th>ADHD + CD</th>
<th>Full sample</th>
<th>ADHD + CD incidence rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Site 1</td>
<td>45</td>
<td>20</td>
<td>65</td>
<td>31</td>
</tr>
<tr>
<td>Site 2</td>
<td>20</td>
<td>5</td>
<td>25</td>
<td>17</td>
</tr>
<tr>
<td>Site 3</td>
<td>63</td>
<td>13</td>
<td>76</td>
<td>17</td>
</tr>
<tr>
<td>Site 4</td>
<td>13</td>
<td>8</td>
<td>21</td>
<td>17</td>
</tr>
<tr>
<td>Site 5</td>
<td>16</td>
<td>5</td>
<td>21</td>
<td>17</td>
</tr>
<tr>
<td>Site 6a</td>
<td>41</td>
<td>33</td>
<td>74</td>
<td>45</td>
</tr>
<tr>
<td>Site 7</td>
<td>20</td>
<td>2</td>
<td>22</td>
<td>9</td>
</tr>
<tr>
<td>Site 8</td>
<td>51</td>
<td>23</td>
<td>74</td>
<td>31</td>
</tr>
<tr>
<td>Site 9</td>
<td>35</td>
<td>13</td>
<td>48</td>
<td>27</td>
</tr>
<tr>
<td>Site 10</td>
<td>87</td>
<td>14</td>
<td>101</td>
<td>14</td>
</tr>
<tr>
<td>Site 11</td>
<td>44</td>
<td>5</td>
<td>49</td>
<td>10</td>
</tr>
<tr>
<td>Total</td>
<td>435</td>
<td>141</td>
<td>576</td>
<td>24</td>
</tr>
</tbody>
</table>

\( ^a \)Two study sites from UK were merged because one of them has only a few families.
quantitative theory has its origins in twin model fitting results that suggest a liability threshold model in which genetic factors associated with CD are entirely shared by ADHD and ADHD + CD is a genetically more severe form of ADHD [Thapar et al., 2001]. According to this view, the same set of causal genes determines both phenotypes but has stronger effects on the ADHD + CD subgroup. On the other hand, key evidence supporting the qualitative theory comes from familial studies showing that relatives of ADHD – CD children have a similar risk for CD as relatives of normal controls [Faraone et al., 1998, 2000c]. From this perspective, it is reasonable to infer at the molecular level, that ADHD + CD genes do not have to confer risk for ADHD – CD or vice versa. In the current study, we tested gene effects in both subgroups and found DAT1 conferring a significant risk for ADHD – CD but a non-significant protective effect for ADHD + CD. This result does not support the liability threshold model, which predicts a gene having unidirectional effects on both groups, but is fully compatible with the prediction of the qualitative theory, which suggests DAT1 as a potential ADHD – CD-specific gene.

A few studies have reported post hoc association tests with dopamine genes in ADHD + CD children after having non-significant results with the full ADHD sample. Two earlier studies did find DRD4 was only associated in ADHD children with conduct problems [Holmes et al., 2002; Kirley et al., 2003]. Another recent cohort study found no association between DRD4 and ADHD + CD [Mill et al., 2006]. However, none of them reported association tests in the ADHD – CD subgroup because they all had an a priori hypothesis that genes conferring risk for ADHD – CD would also affect ADHD + CD with bigger effect size. Moreover, there was no statistical evidence in these studies indicating DRD4 to have more influence on ADHD + CD because the ADHD – CD subgroup has been ignored in their secondary analysis. Therefore, it is questionable that the significant observations in the ADHD + CD groups of the aforementioned studies were simply due to random allelic fluctuations when the original sample

![Fig. 1. Linkage disequilibrium pattern and block structure at DAT1 locus. The diagram was plotted on the D' scale estimated from the founders.](image)

**TABLE III. Single Marker TDT Test and Heterogeneity Test for 18 Common SNPs and Two VNTRs at the DAT1 Locus**

<table>
<thead>
<tr>
<th>Marker</th>
<th>Allele 1</th>
<th>Allele 2</th>
<th>Position</th>
<th>MAF</th>
<th>ADHD – CD</th>
<th>ADHD + CD</th>
<th>Full sample</th>
<th>TDT test P-values</th>
<th>Likelihood ratio test of heterogeneity</th>
<th>ADHD – CD</th>
<th>ADHD + CD</th>
<th>T</th>
<th>NT</th>
<th>T</th>
<th>NT</th>
<th>$\chi^2$</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs7732456</td>
<td>A</td>
<td>C</td>
<td>1445965</td>
<td>0.0656</td>
<td>0.2167</td>
<td>0.1137</td>
<td>0.0722</td>
<td>62</td>
<td>49</td>
<td>17</td>
<td>9</td>
<td>0.3722</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40bp-VNTR</td>
<td>10R</td>
<td>9R</td>
<td>1447327</td>
<td>—</td>
<td>0.0698</td>
<td>0.1678</td>
<td>0.3380</td>
<td>191</td>
<td>156</td>
<td>54</td>
<td>50</td>
<td>0.5756</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs1042098</td>
<td>A</td>
<td>G</td>
<td>1447815</td>
<td>0.2809</td>
<td>0.4369</td>
<td>0.0691</td>
<td>0.4369</td>
<td>193</td>
<td>162</td>
<td>57</td>
<td>49</td>
<td>0.9146</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs27072</td>
<td>A</td>
<td>G</td>
<td>1447522</td>
<td>0.1577</td>
<td>0.0604</td>
<td>0.1678</td>
<td>0.3380</td>
<td>105</td>
<td>134</td>
<td>44</td>
<td>32</td>
<td>0.0337</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs2652511</td>
<td>A</td>
<td>G</td>
<td>1499389</td>
<td>0.4100</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>152</td>
<td>146</td>
<td>44</td>
<td>32</td>
<td>0.9146</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The position of the variants is based on NCBI assembly 35. Minor Allele Frequency (MAF) is based on founder genotypes. ADHD + CD indicates the ADHD comborid with CD group, ADHD – CD indicates the ADHD only group. Both TDT and heterogeneity tests were based on allele 1. Nominal significant P-values are marked in bold. The numbers of transmitted and non-transmitted allele 1 in each subgroup are also provided to differentiate the risk alleles in each subgroup. The markers can be divided into three LD blocks as separated by solid lines.

*SNPs rs2652511 and rs40184 were used in the haplotype analysis.
was subdivided. In the current study, both ADHD + CD and ADHD – CD subgroups were used in the heterogeneity test. It is the comparison between them that provided the statistical evidence indicating DAT1 as a potential ADHD – CD-specific gene. This result suggests that the ADHD – CD subgroup is equally important to the ADHD + CD subgroup in terms of mapping genes underlying ADHD genetic heterogeneity and should not be precluded.

More generally, our findings highlight the importance of establishing genetically homogeneous samples for ADHD molecular genetic studies. In this study, we stratified our ADHD sample based on the prior knowledge that ADHD + CD may have a different genetic etiology compared to ADHD – CD. Our results do suggest that strong ADHD – DAT1 associations can only be detected with the ADHD – CD subgroup sample. Similarly, other phenotypic features such as DSM-IV subtype and symptom persistence may also be determined by genetic difference to some extent [Faraone et al., 2000c; Thapar et al., 2006]. As shown by some association studies, DAT1 and DRD5 have stronger effects on the development of inattentive subtype symptoms [Waldman et al., 1998; Lowe et al., 2004]. These results suggest that investigating specific phenotype subgroups may help to dissect the molecular genetic basis of ADHD. The main difficulty of using such stratified subgroup designs is the recruitment burden to get sufficient statistical power. For example, only about a quarter of the total ADHD sample in the current study also had CD. There will be little or no extra power to detect ADHD + CD associations with only a quarter of the total sample, unless the gene effect is specific to ADHD + CD or contributes a significantly higher risk to ADHD + CD. Nevertheless, with the rapid growth of large-scale international cooperation in ADHD molecular studies, we will be able to achieve sizable phenotypically homogeneous samples and test specific hypotheses with them.

In summary, the genetic heterogeneity test in this study found that variations in the DAT1 gene confer significantly different risks to ADHD children with and without CD. Combined with the association findings in each subgroup, we found that DAT1 was only associated with ADHD children without CD. This result may partially explain the observational inter-study heterogeneity in previous ADHD – DAT1 associations. It also highlights the importance of using genetically homogeneous samples in ADHD molecular genetic studies. Given the presence of significant CD incidence rate difference among our recruiting sites and the large amount of statistical tests performed in the current investigation, we believe further replications are needed to confirm the above conclusions [Sullivan, 2007].

ACKNOWLEDGMENTS

The IMAGE project is a multi-site, international effort supported by NIH grant R01MH62873 to S.V. Faraone. We thank all the families who kindly participated in this research.

REFERENCES


<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Allele frequency (%)</th>
<th>T</th>
<th>NT</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>22.7</td>
<td>138.8</td>
<td>189.5</td>
<td>0.005</td>
</tr>
<tr>
<td>AG</td>
<td>23.0</td>
<td>168.2</td>
<td>173.5</td>
<td>P &gt; 0.1</td>
</tr>
<tr>
<td>GA</td>
<td>19.1</td>
<td>144.6</td>
<td>170.1</td>
<td>P &gt; 0.1</td>
</tr>
<tr>
<td>GG</td>
<td>35.3</td>
<td>275.4</td>
<td>173.9</td>
<td>1.6 × 10^-6</td>
</tr>
</tbody>
</table>

The allele frequency was estimated by EM algorithm. The P-values in the table are haplotype-specific 1-df TDT test. A 3-df ETDT test resulted in P = 8.5 × 10^-5 for the ADHD – CD group and P = 0.8 for the ADHD + CD group. A 1-df chi-square test of heterogeneity on the GG haplotype resulted in P = 0.0009.


