A Haplotype Spanning *KIAA0319* and *TTRAP* Is Associated with Normal Variation in Reading and Spelling Ability

Michelle Luciano, Penelope A. Lind, David L. Duffy, Anne Castles, Margaret J. Wright, Grant W. Montgomery, Nicholas G. Martin, and Timothy C. Bates

Background: *KIAA0319* (6p22.2) has recently been implicated as a susceptibility gene for dyslexia. We aimed to find further support for this gene by examining its association with reading and spelling ability in adolescent twins and their siblings unselected for dyslexia.

Methods: Ten single nucleotide polymorphisms (SNPs) in or near the *KIAA0319* gene were typed in 440 families with up to five offspring who had been tested on reading and spelling tasks. Family-based association analyses were performed, including a univariate analysis of the principal component reading and spelling score derived from the Components of Reading Examination (CORE) test battery and a bivariate analysis of whole-word reading tests measured in a slightly larger sample.

Results: Significant association with rs2143340 (*TTRAP*) and rs6935076 (*KIAA0319*) and with a three-SNP haplotype spanning *KIAA0319* and *TTRAP* was observed. The association with rs2143340 was found in both analyses, although the effect was in the opposite direction to that previously reported. The effect of rs6935076 on the principal component was in the same direction as past findings. Two of the three significant individual haplotypes showed effects in the opposite direction to the two prior reports.

Conclusions: These results suggest that a multilocus effect in or near *KIAA0319* may influence variation in reading ability.

Key Words: Family-based association, haplotype analysis, *KIAA0319*, reading ability, spelling ability, *TTRAP*

There are now three reports of an association of the *KIAA0319* gene (located at 6p22.3-p22.2) with dyslexia phenotypes (1–3). Dyslexias are common neurobehavioral disorders of reading with a prevalence of up to 17.5% in children (4,5). Genetic research to date has predominantly focused on case-control studies of reading disability rather than ability. Nevertheless, there is evidence to suggest that reading disability represents the low tail of a continuous distribution of reading ability in the population (e.g., 6,7). Our aim here, then, was to investigate the association of reading ability with single nucleotide polymorphisms (SNPs) in the *KIAA0319* region in an unselected population sample, thus complementing gene-finding approaches in dyslexic samples.

The region of chromosome 6p bounded broadly by markers D6S109 (8) and D6S291 (9) is the most widely replicated chromosomal region linked to reading phenotypes, with reports of significant linkage in five independent affected sibling pair samples (8–14). However, some groups have failed to replicate the linkage (15–17).

In the only linkage study of reading and spelling phenotypes in an unselected sample to date, we did not detect linkage to this region (18). Our study used a test battery (Components of Reading Examination [CORE]) (19) that captured aspects of the dual route cascaded computational model of reading (20). In a sample which partially overlapped (67%) with that measured on the CORE, linkage to the Schonell Graded Word Reading Test of regular and irregular word reading was found on 6p but in an area slightly distal (~5–10 cM) of 6p21.3 (21). The age range of this sample (16–22 years) was more restricted than that tested on the CORE (age range of 12–25) and the sample was younger on average (16 ± .73 versus 18 ± 2.8). We note, however, that this study only had the power to detect very large quantitative trait loci (QTL) effects by linkage analysis and that the location of linkage peaks has a high degree of uncertainty (22). As these linkage studies were both underpowered, partly an artefact of using an unselected sample, the mostly negative linkage findings on 6p do not preclude QTLs of small effect in this region from influencing normal variation in reading.

The first indication that the *KIAA0319* gene might explain part of the variation in reading disability linked to 6p21.3 came from the Kaplan *et al.* (14) study, showing an association of microsatellite marker JA04 situated within the *KIAA0319* gene. In a sample including 30% of families from the Kaplan *et al.* (14) study and using a high-density SNP map of the 6p21.3 region, Deffenbacher *et al.* (2) reported association ($p < .05$) between a SNP within the *KIAA0319* gene and overall reading ability using a family-based association test under a dominance model. Francks *et al.* (3) identified a 77 kilobase (kb) region (including the first four exons of the *KIAA0319* gene, as well as SNPs in *TTRAP*) on 6p22.2 associated with reading phenotypes. There was no association between SNPs in this region and intelligence quotient (IQ), and IQ acted as an important covariate, suggesting that the gene at 6p is specific for reading and not for the correlated trait of general ability. Most recently, three SNPs in the *KIAA0319* gene (including two implicated in the Francks *et al.* [3] study) have been shown to associate with reading phenotypes in both family-based and case-control analyses (1). Two of these SNPs (*rs*4504469, *rs*6935076) also formed a haplotype that was significantly associated with developmental dyslexia in both case-control and proband-parent trio tests. However, a three-SNP

From Genetic Epidemiology (ML, PAL, DLD, MJW, GWM, NGM), Queensland Institute of Medical Research, Brisbane, Australia; Macquarie Centre for Cognitive Science (AC), Macquarie University, Sydney, Australia; and Department of Psychology (TCB), University of Edinburgh, Edinburgh, United Kingdom.

Address reprint requests to Michelle Luciano, Ph.D., Queensland Institute of Medical Research, Herston, Brisbane, QLD 4029, Australia; E-mail: michelle.luciano@qimr.edu.au.

Received March 15, 2006; revised September 12, 2006; accepted October 30, 2006.

0006-3223/07/$32.00
doi:10.1016/j.biopsych.2007.03.007

haplotype (spanning *KIAA0319* and *TTRAP*) previously implicated by Francks *et al.* (3) did not show association in this sample. The function of *KIAA0319* is presently unknown, but it is a protein that is highly expressed in the brain and therefore of potential functional relevance to reading ability (23).

Thus, while *KIAA0319* may be a susceptibility gene for developmental dyslexia, the causal alleles are unknown. In the present study, we test the association between 10 SNPs in or near the *KIAA0319* gene—including those forming the two-SNP and three-SNP haplotypes tested by Francks *et al.* (3) and Cope *et al.* (1)—with reading and spelling phenotypes in a sample of adolescent twins and their families unselected for reading impairment. Taking the view that reading disability represents the low tail of the reading ability distribution in the population, we expected the association of the *KIAA0319* gene with dyslexia to replicate in our sample.

Methods and Materials

Sample

Twins were initially recruited from primary schools in the greater Brisbane area by media appeals and word of mouth, as part of ongoing studies of melanoma risk factors and cognition (24,25). Data were also gathered from nontwin siblings of twins, with families comprising up to five siblings (including twins). The representativeness of a range of traits, including mole count (26) and intellectual ability (27), indicated that the sample was typical of the Queensland population. Based on the reported ancestry by the twins' parents, the majority of the sample was Caucasian (~98%) and predominantly of Anglo-Celtic (~82%) descent. Blood was obtained from twins, siblings, and from 85.8% of parents for blood grouping and DNA extraction. Zygosity of same-sex twins was diagnosed using nine polymorphic DNA microsatellite markers (AmpF1STR Profiler Plus Amplification Kit; Applied Biosystems, Foster City, California) and three blood groups (ABO, MNS, and Rh), giving a probability of correct assignment greater than 99.99%. Ethical approval for this study was received from the Human Research Ethics Committee, Queensland Institute of Medical Research. Written informed consent was obtained from each participant and their parent/guardian (if younger than 18 years) prior to phenotype and blood collection.

Measures and Procedure

Data collection was performed in two stages within the context of an ongoing study of cognition in adolescent twins and their siblings (25). Intelligence quotient data (Multidimensional Aptitude Battery [MAB]) (28) and measures of whole-word recognition (Schonell Graded Word Reading Test, Contextualized Cambridge Reading Test) were collected in 855 individuals (twins and their nontwin siblings) as close as possible to their 16th birthday, although siblings were a year older on average (age range of 16–22 years). The sample included 237 dizygotic (DZ) twin pairs and a further 62 DZ families with at least one additional nontwin sibling (73 siblings). Of the 126 monozygotic (MZ) twin families, 33 comprised an additional sibling and 6 had two siblings. A further 14 unpaired DZ twins or siblings were included in analyses. Male participants made up 48.5% of the sample. Participants were excluded if there was a parental report of either twin having a history of significant head injury, neurological or psychiatric illness, substance abuse or dependence, or current use of medication with known effects on the central nervous system (not including short-term treatment). Participants had normal or corrected-to-normal vision (> 6/12 Snellen equivalent). The shortened version of the MAB included three verbal subtests (information, arithmetic, vocabulary) and two performance subtests (spatial, object assembly). The Contextualized Cambridge Reading Test (29,30) assesses the pronunciation of irregular words (which must be within the reader's lexicon to be pronounced correctly) embedded within sentences, whereas the Schonell Graded Word Reading Test (31) consists of a list of both irregular words and regular words (these can be pronounced using grapheme-phoneme conversion rules or through lexical access). Both these reading tests are used as neuropsychological tests of premorbid IQ, reflecting a strong correlation with IQ (correlations up to .65 for the MAB IQ subtests); for more detail on these word recognition tests and correlations with IQ, see Wainwright *et al.* (32). The Schonell Graded Word Reading Test reading data were negatively skewed and thus transformed by a logarithmic function of the reverse distribution.

In the second stage of the study (33), more sensitive measures of reading ability were obtained on 82.9% of the cognition sample. As these tests were administered at a later date to existing participants and to new participants before they were tested in the cognition study (i.e., <16 years), the age range for this sample widened to between 14 and 23 years, with a mean age of 18.2 ± 2.4 years. Regular word, irregular word, and nonword reading were assessed using the CORE (19), a 120-word extended version of the Castles and Coltheart (34) test with additional items included to increase the difficulty level for an older sample. This test was administered over the telephone by a trained researcher.

Regular- and irregular-word spelling were measured by 18 regular words and 18 irregular words from the CORE, which were orally presented in a mixed order and without time constraint, with the dependent variable being number of words spelled correctly to oral challenge. The nonlexical spelling assessment required subjects to produce a regularized spelling for the 18 words given in the irregular spelling test. Each word was presented verbally, and the letter string used for spelling was recorded and scored for phonological correctness from a list of acceptable regularizations. Words were repeated on request. Test scores on each of the three reading subtests and three spelling tests were calculated as a simple sum of correct items and were log-odds transformed to normalize their distribution. A previous multivariate linkage study of six reading and spelling phenotypes showed that each measure contributed to the linkage at 6p (35), so to reduce the multiple testing problem, we derived a principal component factor score from our measures which explained almost 70% of variance.

Genotyping

The 10 SNPs reported by Cope *et al.* (1) in their case-control analysis were selected for genotyping, although one of these SNPs (rs6939068) failed genotyping procedures (e.g., assay design) and was replaced with rs11757448 (located 143 base pair [bp] upstream). Genotyping forward and reverse polymerase chain reaction (PCR) primers and a primer extension probe were designed using SpectroDESIGNER software (Sequenom, San Diego, California) and purchased from Bioneer Corporation (Daejeon, Korea). The 10 assays were assembled into two multiplex sets of five SNPs each. Single nucleotide polymorphisms were typed on the Sequenom MassARRAY platform using a modified protocol for high multiplex homogeneous MassEXTEND (hME) reactions (Sequenom, application notes).