# 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).

We carried out each multiplex PCR using 5 ng genomic DNA in a final reaction volume of 2.5 μL containing .1 U Taq polymerase (Hot Star Taq; Qiagen, Valencia, California), 1X PCR buffer (Qiagen), 1 mmol/L magnesium chloride ($MgCl_2$), 250 μm deoxyribonucleotide triphosphates (dNTPs [Qiagen]), and 100 nmol/L forward and reverse primers for all SNPs (excluding rs6911855 and rs3777664, which were at 200 nmol/L). Thermocycling using a dual 384-Well GeneAmp PCR System 9700 Cycler (Applied Biosystems) was performed according to the hME protocol. Polymerase chain reaction products were subjected to 1 μL of a shrimp alkaline phosphatase (SAP; Sequenom) solution containing .15 U SAP and incubated at 37°C for 20 min followed by 85°C for 5 min.

The 2.0 μL hME reaction contained allele-specific extension primers at 600 nmol/L to 900 nmol/L each, 1X hME ACT termination mix, and .08 U/L of ThermoSequenase (Sequenom). Homogeneous MassEXTEND cycling conditions were modified to include 99 cycles of 5 sec at 94°C, 5 sec at 52°C, and 5 sec at 72°C. The pooled hME products of the two multiplexes (11 μL final volume) were purified by the addition of SpectroCLEAN resin (Sequenom) and dispensed onto SpectroCHIPs (Sequenom) using a nanolitre pipetting system (SpectroCHIP, SpectroJet; Sequenom). Reaction products were analyzed by MALDI-TOF MassARRAY (Sequenom).

Genotype error checking, including Mendelian inconsistencies, and tests of Hardy-Weinberg equilibrium were performed in MERLIN (http://www.sph.umich.edu/csg/abecasis/Merlin) (36) and Sib-Pair (http://www2.qimr.edu.au/davidD/sib-pair.html) (37).

### Association Analysis

Single nucleotide polymorphism and haplotype association analysis was performed in PBAT (http://biostat.harvard.edu/~clange/default.htm) (38), a program with the facility of analyzing quantitative traits within a multivariate framework using a family-based association test (39). As PBAT does not consider the relatedness of monozygotic twins, MZ twins' phenotypic scores were averaged where both twins participated and accordingly treated as a single observation; because MZ correlations were fairly strong, this averaging of phenotypes was not considered to cause major variance differences between MZs and siblings. The Deffenbacher *et al.* (2) significant association between *KIAA0319* and overall reading ability was found under a dominance model, so we specified dominance models for all SNP and haplotype analyses against the null hypothesis of no linkage or association. A bivariate analysis of the Contextualized Cambridge Reading Test and Schonell Graded Word Reading Test reading measures was performed and included the covariates of age, sex, and performance IQ, all of which were significant at $p < .001$. Univariate analysis of these two phenotypes was performed only if the bivariate tests were significant, thereby reducing the overall number of tests performed. The CORE principal component, which had fewer data points, was analyzed separately but adjusted for the same covariates, which were significant at $p < .001$. While 10 SNPs were analyzed, we do not correct the $p$ value for these tests, as the relatedness of SNPs would make this correction overly conservative. For a SNP explaining 1% of variance in our traits, under a dominance model and against a background sibling correlation of .30, we have roughly 76% power ($\alpha = .05$) to detect overall association with a SNP with minor allele frequency above .05 (40).

## Results

### Descriptive

Mendelian inconsistencies were identified using Sib-Pair and made up .21% of the data; a further 1.1% of the data was probable genotyping errors, as detected in MERLIN and removed from the analysis. Genotype screening showed that the population was in Hardy-Weinberg equilibrium at all SNPs. Allele frequencies were consistent with previous reports, and minor allele frequencies less than .05 were observed for rs6911855 and rs11757448 (Table 1). The physical locations of the 10 SNPs across the 6p22.2 chromosomal region previously associated with reading phenotypes are schematically presented in Figure 1, along with their intermarker linkage disequilibrium.

### Association Analyses

Results of the family-based association tests for the individual SNPs are shown in Table 2. Significant association was found for rs6935076 in the univariate (CORE principal component) analysis ($p = .008$), with the minor allele related to poorer reading performance (.003 SD). In both bivariate (Contextualized Cambridge Reading Test, Schonell Graded Word Reading Test) and univariate analyses, rs2143340 was significant, with respective $p$ values of .009 and .02. Univariate analyses of the Contextualized Cambridge Reading Test and Schonell Graded Word Reading Test were also significant at a two-trait Bonferroni adjusted significance level of .025 (Contextualized Cambridge Reading Test, $p = .002$; Schonell Graded Word Reading Test, $p = .007$). This SNP explained around 1% of variance in each measure, with

**Table 1.** SNP Marker Descriptive Information, Including Genetic Map Position, Gene Location, and Minor Allele Frequency

| SNP ID | Position Mb[a] | Gene | Gene Location | Allele | | Minor Allele Frequency[b] |
|---|---|---|---|---|---|---|
| | | | | Major (1) | Minor (2) | |
| rs2793422 | 24.526 | *MRS2L* | Exon 8 | G | A | .37 |
| rs4504469 | 24.697 | *KIAA0319* | Exon 4 | C | T | .41 |
| rs6911855 | 24.715 | *KIAA0319* | Intron1 | C | T | .02 |
| rs11757448 | 24.721 | *KIAA0319* | Intron1 | G | A | .04 |
| rs2179515 | 24.736 | *KIAA0319* | Intron1 | G | A | .37 |
| rs6935076 | 24.752 | *KIAA0319* | Intron1 | C | T | .36 |
| rs2038137 | 24.754 | *KIAA0319* | Intron1 | C | A | .39 |
| rs2143340 | 24.767 | *TTRAP* | Intron2 | T | C | .15 |
| rs3777664 | 24.802 | *THEM2* | Intron1 | T | C | .31 |
| rs1053598 | 24.813 | *C6orf62* | 3′UTR | C | T | .27 |

NCBI, National Center for Biotechnology Information; SNP, single nucleotide polymorphism.
[a]Mb = Mega-base position on NCBI Build 35.
[b]Maximum number of genotypes = 2443.

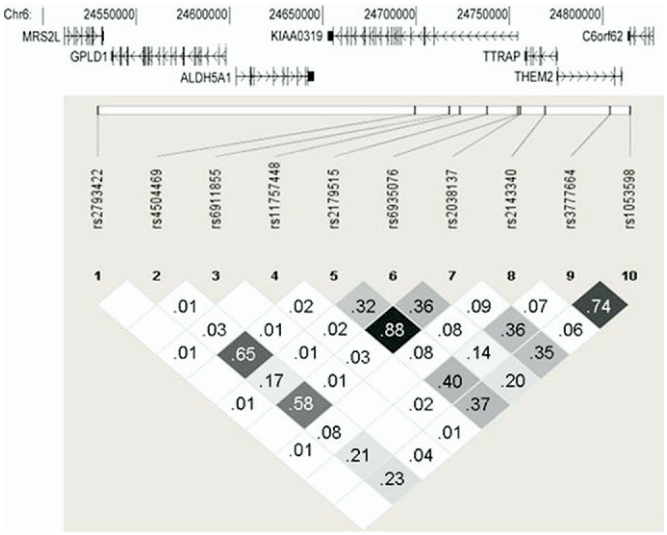


**Figure 1.** The location of the 10 SNPs across the 302 kb region containing the genes *MRSL2*, *GPLD1*, *ALDH5A1*, *KIAA0319*, *TTRP*, *THEM2*, and *C6orf62* (adapted from UCSC Genome Browser, NCBI Build 35, http://genome.ucsc.edu/). Transcriptional direction of and exon location within each gene are indicated by arrow direction and vertical lines, respectively. Inter-SNP linkage disequilibrium (shown below the gene structure) was generated using Haploview 3.2 (http://www.broad.mit.edu/personal/jcbarret/haploview) and depicts the $r^2$ between SNPs, with darker shading representing stronger linkage disequilibrium (only nonzero $r^2$ are shown in the cells). kb, kilobase; NCBI, National Center for Biotechnology Information; SNP, single nucleotide polymorphism; UCSC, University of California, Santa Cruz.

the direction of the effect such that the minor allele conferred a reading advantage (mean effect of .06 SD). A similar proportion of variance in the CORE principal component was accounted for by this marker, and consistent with the bivariate results, the minor allele conferred a reading advantage of .04 SD. The probability that a person with the dominant genotype for the rs2143340 minor allele was below average, average, or above average in reading ability was compared with the major allele homozygote. Below and above average reading groups were defined by respective scores <1 SD and >1 SD from the mean (covariate corrected) Schonell Graded Word Reading Test score, since this test showed the largest mean effect size. The probabilities of belonging to the below average or average reading groups were lower for those with the dominant genotype and higher for those in the above average reading group, with ratios of .97 (below average and average) and 1.2 (above average) observed between dominant versus homozygous groups.

**Table 2.** Association of SNPs with the Cambridge and Schonell Whole-Word Reading Tests and the CORE Principal Component

| | Bivariate: Cambridge, Schonell | CORE Principal Component |
|---|---|---|
| | FBAT *p* Value | FBAT *p* Value |
| rs2793422 | .06 | .40 |
| rs4504469 | .85 | .57 |
| rs6911855 | .44 | .65 |
| rs11757448 | .46 | .92 |
| rs2179515 | .93 | .98 |
| rs6935076 | .22 | **.01** |
| rs2038137 | .69 | .84 |
| rs2143340 | **.01** | **.02** |
| rs3777664 | .52 | .61 |
| rs1053598 | .54 | .86 |

Significant values are highlighted in bold.

Cambridge, Contextualized Cambridge Reading Test; CORE, Components of Reading Examination; FBAT, family-based association test; Schonell, Schonell Graded Word Reading Test; SNP, single nucleotide polymorphism.

**Table 3.** Haplotype Analysis in PBAT for the Bivariate Analysis of Cambridge and Schonell Reading Tests and the Univariate Test of the CORE Principal Component

| | | Bivariate | PC |
|---|---|---|---|
| | Haplotype Frequency | FBAT *p* Value | FBAT *p* Value |
| rs4504469, rs6935076 | | | |
| 11 | .281 | .77 | .22 |
| 12 | .291 | .93 | .17 |
| 21 | .361 | .84 | .83 |
| 22 | .067 | .09 | .06 |
| rs4504469, rs2038137, rs2143340 | | | |
| 111 | .389 | .35 | **.02** |
| 221 | .337 | .54 | .67 |
| 112 | .131 | **.04** | .13 |
| 211 | .072 | **.02** | **.02** |
| 121 | .047 | .45 | .65 |
| 212 | .013 | .45 | .19 |
| 222 | .005 | .03 | .06 |
| 122 | .004 | .10 | .03 |

Significant values ($p < .05$) are highlighted in bold for haplotypes with a frequency > 5%.

Global 2-SNP haplotype association: Bivariate, $p = .24$; PC, $p = .36$.

Global 3-SNP haplotype association: Bivariate, $p = .005$; PC, $p = .01$.

Note: 1 represents the major allele.

Cambridge, Contextualized Cambridge Reading Test; CORE, Components of Reading Examination; FBAT, family-based association test; PC, principal component; Schonell, Schonell Graded Word Reading Test; SNP, single nucleotide polymorphism.

The two-SNP haplotype (rs4504469, rs6935076) found by Cope *et al.* (1) to be related to developmental dyslexia and the three-SNP haplotype (rs4504469, rs2038137, rs2143340) originally identified by Francks *et al.* (3) were tested for association. The global tests of association were significant for the three-SNP haplotype in the bivariate (Contextualized Cambridge Reading Test, Schonell Graded Word Reading Test) and univariate (CORE principal component) analyses. Individual haplotype test results are shown in Table 3. The 2-1-1 and 1-1-2 haplotypes (respective frequencies of 7% and 13%) were significant in the bivariate analysis. Univariate analysis showed the Contextualized Cambridge Reading Test reading score to be significantly ($p = .011$) associated with the 2-1-1 haplotype, while the Contextualized Cambridge Reading Test and Schonell Graded Word Reading Test reading measures showed respective *p* values of .01 and .04 for their association with the 1-1-2 haplotype. The 2-1-1 and 1-1-2 haplotypes had been shown by Francks *et al.* (3) to be related to a number of reading and spelling measures. While the 2-2-2 haplotype was also significant in the bivariate analysis, its low frequency precluded it from further investigation. In the CORE analysis, two of the three SNP haplotypes with a frequency greater than 5% (1-1-1, 2-1-1) showed association at the .05 level of significance. Haplotype 1-1-1 was significant in the studies by Cope *et al.* (1) and Francks *et al.* (3), while haplotype 2-1-1 was only implicated by Francks *et al.* (3)—although it had a minor allele frequency of 5% in their sample. Haplotype 1-1-1 conferred

a reading disadvantage of .02 SD, whereas 1-1-2 conferred a reading advantage of .06 SD with respect to all other haplotypes. The 2-1-1 haplotype, which was significant in both the univariate and bivariate analyses, was associated with lower mean reading scores relative to all other haplotypes (mean difference of .03 SD across all measures).

A subsidiary analysis including individuals who reported at least 75% Anglo-Celtic ancestry (based on grandparents' ethnicity) was performed to check whether association effects were consistent in the more homogenous sample. The single SNP association results showed the same direction of effect in this subsample, although for rs2143340 association was just short of significance with $p$ values of .06 and .07 for the respective bivariate and univariate analyses. Haplotype analysis supported association of the 1-1-1 haplotype in the univariate analysis ($p$ = .02) and haplotype 2-1-1 in the bivariate analysis ($p$ = .04). Haplotype 1-1-1 decreased reading scores by a mean effect size of .06 SD, while haplotype 2-1-1 decreased reading scores by an average effect size of .04 SD.

## Discussion

This study examined the association of SNPs and haplotypes within and near the *KIAA0319* gene with reading measures in a twin sample representative of the general population for reading and spelling ability. Two sets of reading measures were collected on different occasions: one composed of two measurements of whole-word reading (premorbid IQ indices) and the other comprised tests of regular, irregular, and nonword reading and spelling. To overcome statistical problems related to missing data, these two sets of measures were analyzed separately. A bivariate analysis of the Schonell Graded Word Reading Test and Contextualized Cambridge Reading Test reading measures (collected in a larger sample) and a univariate analysis of the principal component derived from the six reading and spelling phenotypes were performed. Family-based association tests of individual SNPs showed association with rs2143340 (*TTRAP* gene) and rs6935076 (*KIAA0319* gene) and with the three-SNP haplotype—spanning *KIAA0319* and *TTRAP*—identified by Francks *et al.* (3).

The most consistent result of our study was the individual association with rs2143340, which was significant in all analyses, albeit in the opposite direction to that reported by Francks *et al.* (3). While this result appears puzzling, it has been shown that multiplicative or additive effects of multiple risk loci can produce associations with an observed variant in opposite directions if the correlation between the observed variant and the unobserved variant differs between populations due to sampling variation (41). Alternatively, this result may represent type I error: Cope *et al.* (1) failed to replicate the individual association with rs2143340. This SNP was not in the *KIAA0319* gene but in *TTRAP*, a gene that had also shown association to a number of reading measures in a study by Deffenbacher *et al.* (2). *TTRAP* plays a role in nuclear factor-kappa B (NF-κB) transcription by encoding a tumor necrosis factor receptor-associated protein (42), with downstream effects on long-term potentiation and synaptic plasticity, known to influence learning and memory. The association to *TTRAP* was not significant or of reduced significance if IQ was removed as a covariate in the analysis (data not shown), indicating that the gene in this region has effects specific to reading ability and unrelated to variance in general cognition. This is consistent with the Francks *et al.* (3) finding that significance levels weakened in their SNP analyses, which did not adjust reading disability phenotypes for the effects of IQ.

Association with rs6935076 (in *KIAA0319*) was supported in the univariate analysis and in the subsample of Anglo-Celtics. The effect was in the same direction as that reported by Cope *et al.* (1), with the minor allele negatively affecting reading ability. It can be argued that the CORE test battery is more sensitive to reading disability than the Contextualized Cambridge Reading Test and Schonell Graded Word Reading Test, and this may account for their lack of association in the bivariate analysis. The two-SNP haplotype (rs4504469, rs6935076) to which this SNP contributes did not show association with any of the measures. This haplotype had been identified by Cope *et al.* (1) through a stepwise logistic regression of the 10 SNPs. Our study, the first attempt to replicate this finding, found no support for this haplotype in an unselected sample.

The three-SNP haplotype (rs4504469, rs2038137, rs2143340) originally identified by Francks *et al.* (3) was associated with reading ability in our sample. In terms of individual haplotypes, the 1-1-1 haplotype supported by both Francks *et al.* (3) and Cope *et al.* (1) was significant in our study for the reading principal component but not the whole-word reading measures. Interestingly, Francks *et al.* (3) found this haplotype to be significant only for their phonological decoding measure, and this may explain why our whole-word reading measures, which predominantly tap irregular word reading (or lexical processes), were not related to this haplotype. Consistent with the findings of Francks *et al.* (3) and Cope *et al.* (1), the 1-1-1 haplotype in our sample was associated with decreased reading scores. While Francks *et al.* (3) considered the 1-1-2 haplotype to be the main risk haplotype for reading disability, in our study this haplotype was only significant for the bivariate analysis and the effect on the measures was in the opposite direction. Similarly, while we replicated the 2-1-1 haplotype in both analyses, the direction of the effect was also the reverse of what they found. With so many statistical tests, it is possible that some of our significant results are type I error. The 1-1-1 haplotype is perhaps the most convincing replication, as it has now been found in three samples, each showing the same direction of effect. This haplotype was also significant in our more homogenous subsample of Anglo-Celtics, even though the individual effect of rs2143340 was just short of significance in this analysis.

The haplotypic effects reflected the individual effect of rs2143340, a SNP in *TTRAP* that showed very little linkage disequilibrium with the other two SNPs located in *KIAA0319*. In fact, the linkage disequilibrium between this SNP and any of the HapMap identified SNPs in *KIAA0319* was low, with a maximum $r^2$ of .72. However, Paracchini *et al.* (43) recently showed the 1-1-2 risk haplotype (other haplotypes were not investigated) to be associated with reduced expression of *KIAA0319* but not *TTRAP* or *THEM2*, so while we implicate the SNP in *TTRAP*, it is not clear whether the putative gene is *TTRAP* or a gene (or intergenic controlling element) nearby.

Our analysis of the shared variance (derived from the principal components analysis) between different components of reading and spelling—namely nonlexical and lexical processes—indicates that the causative gene in this region has general effects on reading processes. This is in line with recent findings that both shared and separate genetic factors influence the nonlexical and lexical routes of reading (44) and also agrees with emerging linkage and association evidence that genes on 6p influence orthographic and phonological decoding measures (14,35). In the present study, we have implicated a general reading gene.

However, our previous finding (44) of specific genetic influences on nonlexical and lexical routes suggests testing candidate genes with individual reading component measures where the overall reading composite does not show association.

The positive findings from our study have important implications for the study of reading disabilities. While Francks *et al.* (3) suggested that the association effect on 6p was more marked in probands with more severe reading disability, our results show that this effect is detectable (albeit with a smaller effect size) in a sample with relatively few reading impaired participants (15.5% of our sample scored below 1 SD from the age, sex, and IQ-corrected reading component mean). Furthermore, this effect has been detected using a shorter test battery than is normally used in a clinical setting and via telephone (for CORE battery) rather than face-to-face administration. Our findings certainly add weight to the Shaywitz *et al.* (7) argument that dyslexia simply represents the low tail of a normal distribution of reading ability in the population. The implication is that the same genes influence poor and exceptional reading ability, affording new genetic study designs that include using the extremes of the distribution or, as we have done, the entire range of scores in reading ability.

*Phenotype collection was funded by Australian Research Council Grants (A79600334, A79906588, A79801419, DP0212016, DP0343921). ML is supported by an Australian Research Council Postdoctoral Fellowship (DP0449598).*

*We thank the twins and their parents for their cooperation; Anjali Henders and Megan Campbell for managing sample processing and DNA extraction; Alison Mackenzie for coordinating the reading test battery mailout; Marlene Grace, Ann Eldridge, and the many research interviewers for data collection; Dr. Manuel Ferreira, Jonathan Hansen, and Dr. Andrew Birley for helpful statistical advice; and Professor Max Coltheart, Professor Peter Visscher, and Dr. Dale Nyholt for their useful comments on various aspects of this work.*

1. Cope N, Harold D, Hill G, Moskvina V, Stevenson J, Holmans P, *et al.* (2005): Strong evidence that KIAA0319 on chromosome 6p is a susceptibility gene for developmental dyslexia. *Am J Hum Genet* 76:581–591.
2. Deffenbacher KE, Kenyon JB, Hoover DM, Olson RK, Pennington BF, DeFries JC, *et al.* (2004): Refinement of the 6p21.3 quantitative trait locus influencing dyslexia: Linkage and association analyses. *Hum Genet* 115: 128–138.
3. Francks C, Paracchini S, Smith SD, Richardson AJ, Scerri TS, Cardon LR, *et al.* (2004): A 77-kilobase region of chromosome 6p22.2 is associated with dyslexia in families from the United kingdom and from the United States. *Am J Hum Genet* 75:1046–1058.
4. Francks C, MacPhie IL, Monaco AP (2002): The genetic basis of dyslexia. *Lancet Neurol* 1:483–490.
5. Shaywitz SE, Shaywitz BA (2005): Dyslexia (specific reading disability). *Biol Psychiatry* 57:1301–1309.
6. Plomin R, Kovas Y (2005): Generalist genes and learning disabilities. *Psychol Bull* 131:592–617.
7. Shaywitz SE, Escobar MD, Shaywitz BA, Fletcher JM, Makuch R (1992): Evidence that dyslexia may represent the lower tail of a normal distribution of reading ability. *N Engl J Med* 326:145–150.
8. Grigorenko EL, Wood FB, Meyer MS, Hart LA, Speed WC, Shuster A, *et al.* (1997): Susceptibility loci for distinct components of developmental dyslexia on chromosomes 6 and 15. *Am J Hum Genet* 60:27–39.
9. Fisher SE, Marlow AJ, Lamb J, Maestrini E, Williams DF, Richardson AJ, *et al.* (1999): A quantitative-trait locus on chromosome 6p influences different aspects of developmental dyslexia. *Am J Hum Genet* 64:146–156.
10. Cardon LR, Smith SD, Fulker DW, Kimberling WJ, Pennington BF, DeFries JC (1994): Quantitative trait locus for reading disability on chromosome 6. *Science* 266:276–279.
11. Fisher SE, DeFries JC (2002): Developmental dyslexia: Genetic dissection of a complex cognitive trait. *Nat Rev Neurosci* 3:767–780.
12. Fisher SE, Francks C, Marlow AJ, MacPhie IL, Newbury DF, Cardon LR, *et al.* (2002): Independent genome-wide scans identify a chromosome 18 quantitative-trait locus influencing dyslexia. *Nat Genet* 30:86–91.
13. Gayan J, Smith SD, Cherny SS, Cardon LR, Fulker DW, Brower AM, *et al.* (1999): Quantitative-trait locus for specific language and reading deficits on chromosome 6p. *Am J Hum Genet* 64:157–164.
14. Kaplan DE, Gayan J, Ahn J, Won TW, Pauls D, Olson RK, *et al.* (2002): Evidence for linkage and association with reading disability on 6p21.3-22. *Am J Hum Genet* 70:1287–1298.
15. Chapman NH, Igo RP, Thomson JB, Matsushita M, Brkanac Z, Holzman T, *et al.* (2004): Linkage analyses of four regions previously implicated in dyslexia: Confirmation of a locus on chromosome 15q. *Am J Med Genet B Neuropsychiatr Genet* 131:67–75.
16. Field LL, Kaplan BJ (1998): Absence of linkage of phonological coding dyslexia to chromosome 6p23-p21.3 in a large family data set. *Am J Hum Genet* 63:1448–1456.
17. Petryshen TL, Kaplan BJ, Liu MF, Field LL (2000): Absence of significant linkage between phonological coding dyslexia and chromosome 6p23-21.3, as determined by use of quantitative-trait methods: Confirmation of qualitative analyses. *Am J Hum Genet* 66:708–714.
18. Bates TC, Luciano M, Castles A, Coltheart M, Wright MJ, Martin NG (2007): Replication of reported linkages for dyslexia and spelling and suggestive evidence for novel regions on chromosomes 4 and 17. *Eur J Hum Genet* 15(2):194–203.
19. Bates TC, Castles A, Coltheart M, Gillespie N, Wright MJ, Martin NG (2004): Behaviour genetic analyses of reading and spelling: A component processes approach. *Aust J Psychol* 56:115–126.
20. Coltheart M, Rastle K, Perry C, Langdon R, Ziegler J (2001): DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychol Rev* 108:204–256.
21. Luciano M, Wright MJ, Duffy DL, Wainwright MA, Zhu G, Evans DM, *et al.* (2005): Genome-wide scan of IQ finds significant linkage to a Quantitative Trait Locus on 2q. *Behav Genet* 36:45–55.
22. Roberts SB, MacLean CJ, Neale MC, Eaves LJ, Kendler KS (1999): Replication of linkage studies of complex traits: An examination of variation in location estimates. *Am J Hum Genet* 65:876–884.
23. Londin ER, Meng H, Gruen JR (2003): A transcription map of the 6p22.3 reading disability locus identifying candidate genes. *BMC Genomics* 4:25.
24. McGregor B, Pfitzner J, Zhu G, Grace M, Eldridge A, Pearson J, *et al.* (1999): Genetic and environmental contributions to size, color, shape, and other characteristics of melanocytic naevi in a sample of adolescent twins. *Genet Epidemiol* 16:40–53.
25. Wright M, De Geus E, Ando J, Luciano M, Posthuma D, Ono Y, *et al.* (2001): Genetics of cognition: Outline of a collaborative twin study. *Twin Res* 4:48–56.
26. Zhu G, Duffy DL, Eldridge A, Grace M, Mayne C, O'Gorman L, *et al.* (1999): A major quantitative-trait locus for mole density is linked to the familial melanoma gene CDKN2A: A maximum-likelihood combined linkage and association analysis in twins and their sibs. *Am J Hum Genet* 65:483–492.
27. Luciano M, Wright MJ, Geffen GM, Geffen LB, Smith GA, Evans DM, *et al.* (2003): A genetic two-factor model of the covariation among a subset of Multidimensional Aptitude Battery and WAIS-R subtests. *Intelligence* 31:589–605.
28. Jackson DN (1984): *Manual for the Multidimensional Aptitude Battery*. Port Huron, MI: Research Psychologists Press.
29. Beardsall L, Huppert FA (1994): Improvement in NART word reading in demented and normal older persons using the Cambridge Contextual Reading Test. *J Clin Exp Neuropsychol* 16:232–242.
30. Nelson HE (1982): *National Adult Reading Test*. Berkshire, UK: NFER - Nelson Publishing Company.
31. Schonell FJ, Schonell PE (1960): *Diagnostic and Attainment Testing*. Edinburgh: Oliver & Boyd.
32. Wainwright MA, Wright MJ, Geffen GM, Geffen LB, Luciano M, Martin NG (2004): Genetic and environmental sources of covariance between reading tests used in neuropsychological assessment and IQ subtests. *Behav Genet* 34:365–376.
33. Wright MJ, Martin NG (2004): Brisbane Adolescent Twin Study: Outline of study methods and research projects. *Aust J Psychol* 56:65–78.

34. Castles A, Coltheart M (1993): Varieties of developmental dyslexia. *Cognition* 47:149–180.
35. Marlow AJ, Fisher SE, Richardson AJ, Francks C, Talcott JB, Monaco AP, *et al.* (2001): Investigation of quantitative measures related to reading disability in a large sample of sib-pairs from the UK. *Behav Genet* 31:219–230.
36. Abecasis GR, Cherney SS, Cookson WO, Cardon LR (2002): Merlin-rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 30:97–101.
37. Duffy DL (2005): *SIB-PAIR: A Program for Elementary Genetical Analyses, 0.99.9. ed.* Brisbane, Australia: Queensland Institute of Medical Research. Available at: http://www2.qimr.edu.au/davidD/sib-pair.html. Accessed July 9, 2006.
38. Lange C, DeMeo DL, Silverman EK, Weiss ST, Laird NM (2004): PBAT: Tools for family-based association studies. *Am J Hum Genet* 74:367–369.
39. Laird NM, Horvath S, Xu X (2000): Implementing a unified approach to family-based tests of association. *Genet Epidemiol* 19(suppl 1): S36–S42.
40. Purcell S, Cherny SS, Sham PC (2003): Genetic Power Calculator: Design of linkage and association genetic mapping studies of complex traits. *Bioinformatics* 19:149–150.
41. Lin PI, Vance JM, Pericak-Vance MA, Martin ER (2007): No gene is an island: The flip-flop phenomenon. *Am J Hum Genet* 80:531–8.
42. Pype S, Declercq W, Ibrahimi A, Michiels C, Van Rietschoten JG, Dewulf N, *et al.* (2000): TTRAP, a novel protein that associates with CD40, tumor necrosis factor (TNF) receptor-75 and TNF receptor-associated factors (TRAFs), and that inhibits nuclear factor-kappa B activation. *J Biol Chem* 275:18586–18593.
43. Paracchini S, Thomas A, Castro S, Lai C, Paramasivam M, Wang Y, *et al.* (2006): The chromosome 6p22 haplotype associated with dyslexia reduces the expression of KIAA0319, a novel gene involved in neuronal migration. *Hum Mol Genet* 15:1659–1666.
44. Bates TC, Luciano M, Castles A, Wright MJ, Coltheart M, Martin NG (2007): Genetics of reading and spelling: Shared genes across modalities, but different genes for lexical and nonlexical processing. *Reading and Writing* 20:147–171.