Behav Genet. Author manuscript; available in PMC 2017 June 13.

Published in final edited form as:

Behav Genet. 2011 January ; 41(1): 77–89. doi:10.1007/s10519-010-9434-1.

Genetic Variation in the *KIAA0319* 5′ Region as a Possible Contributor to Dyslexia

Adrienne Elbert,
Genetics and Development Division, Toronto Western Research Institute, Toronto Western Hospital, University Health Network, MP14-302, 399 Bathurst Street, Toronto, ON M5T 2S8, Canada

Maureen W. Lovett,
Program in Neurosciences and Mental Health, Hospital for Sick Children, Toronto M5G 1X8, Canada

Tasha Cate-Carter,
Program in Neurosciences and Mental Health, Hospital for Sick Children, Toronto M5G 1X8, Canada

Ashley Pitch,
Program in Neurosciences and Mental Health, Hospital for Sick Children, Toronto M5G 1X8, Canada

Elizabeth N. Kerr, and
Program in Neurosciences and Mental Health, Hospital for Sick Children, Toronto M5G 1X8, Canada

Cathy L. Barr
Genetics and Development Division, Toronto Western Research Institute, Toronto Western Hospital, University Health Network, MP14-302, 399 Bathurst Street, Toronto, ON M5T 2S8, Canada. Program in Neurosciences and Mental Health, Hospital for Sick Children, Toronto M5G 1X8, Canada

Abstract

Reading disabilities (RD) have been linked and associated with markers on chromosome 6p with results from multiple independent samples pointing to *KIAA0319* as a risk gene and specifically, the 5′ region of this gene. Here we focus genetic studies on a 2.3 kb region spanning the predicted promoter, the first untranslated exon, and part of the first intron, a region we identified as a region of open chromatin. Using DNA from probands with RD, we screened for genetic variants and tested select variants for association. We identified 17 DNA variants in this sample of probands, 16 of which were previously reported in public databases and one previously identified in a screen of this region. Based on the allele frequencies in the probands compared to public databases, and on possible functional consequences of the variation, we selected seven variants to test for association

Correspondence to: Cathy L. Barr.

Electronic supplementary material The online version of this article (doi:10.1007/s10519-010-9434-1) contains supplementary material, which is available to authorized users.

in a sample of families with RD, in addition to four variants which had been tested previously. We also tested two markers 5′ of this region that were previously reported as associated. The strongest evidence for association was observed with alleles of the microsatellite marker located in the first untranslated exon and haplotypes of that marker. These results support previous studies indicating the 5′ region of the *KIAA0319* gene as the location of risk alleles contributing to RD.

Keywords

Dyslexia; Genetics; Reading Disabilities; KIAA0319; Transcription

Introduction

Reading Disabilities (RD), or Developmental Dyslexia, is a complex disorder that results from genetic and environmental factors. RD is characterized by difficulties with accurate and/or fluent word recognition and by poor spelling and decoding abilities and is estimated to impair 5–12% of school-aged children in North America (Shaywitz et al. 1990). The National Institute of Child Health and Human Development (NICHD), as reviewed in McGrath et al. (2006), describes RD as resulting from deficits in the cognitive constituents: phonological awareness, phonological coding, orthographic coding and rapid serial naming.

The heritability of reading difficulties ranges from 30 to 72% (DeFries and Gillis 1993; Wadsworth et al. 2000; Harlaar et al. 2005; Hawke et al. 2006; Wadsworth et al. 2007). Genes influencing reading are distributed throughout the genome, with evidence for linkage of RD, reading skills or related phenotypes reported to at least 10 chromosome regions (Paracchini et al. 2007; Scerri and Schulte-Korne 2010). Of these, the chromosome 6p region has been the most studied with linkage to RD or reading measures reported in multiple samples (Smith and Kimberling 1991; Cardon et al. 1994; Grigorenko et al. 1997; Fisher et al. 1999; Gayan et al. 1999; Grigorenko et al. 2000). Early association studies in the 6p RD linkage region have refined the candidate locus to a region containing five brain-expressed genes: neurensin-1 (*NRSN1*), doublecortin domain containing 2 (*DCDC2*), *KIAA0319*, TRAF and TNF associated protein (*TTRAP*) and thioesterase superfamily 2 (*THEM2*) (Kaplan et al. 2002; Turic et al. 2003; Deffenbacher et al. 2004). Two of the candidate genes, *KIAA0319* and *DCDC2*, have since shown the strongest evidence for association to RD and reading measures (Francks et al. 2004; Cope et al. 2005; Meng et al. 2005; Harold et al. 2006; Schumacher et al. 2006; Luciano et al. 2007; Paracchini et al. 2008; Couto et al. 2010; Dennis et al. 2009; Wilcke et al. 2009). Both *KIAA0319* and *DCDC2* have roles in neuronal migration (*DCDC2*: (Meng et al. 2005; Burbridge et al. 2008); *KIAA0319*: (Paracchini et al. 2006; Peschansky et al. 2010)) and thus are attractive candidates based on previous evidence from post-mortem studies in individuals with RD that revealed the presence of ectopias and dysplasias indicative of neuronal migration abnormalities (Galaburda et al. 1985; Humphreys et al. 1990). Therefore, DNA variation in developmental genes governing the process of neuronal migration has been suggested as a mechanism contributing to RD (Galaburda et al. 1985).

Previous studies supporting *KIAA0319* as a RD risk gene have found the most significant association to markers in the 5′ end of *KIAA0319*, suggesting that the RD causal DNA variant(s) may have a role in regulating gene expression (Francks et al. 2004; Cope et al. 2005; Harold et al. 2006). This is consistent with Francks et al. (2004), who found no evidence for a protein-coding change in *KIAA0319* that accounted for the association to RD and the finding of a decrease in *KIAA0319* expression in cell lines carrying a risk haplotype (Paracchini et al. 2006). This haplotype spans a 77 kb region across *KIAA0319* and adjacent gene *TTRAP*, represented by the alleles *G-C-C* of markers rs4504469, rs2038137 and rs2143340. Recently, the minor alleles (A) of rs9461045 and (G) of rs3212236 were shown to alter reporter gene expression from a luciferase construct in neuronal and non-neuronal cell lines (Dennis et al. 2009). Both the A allele of rs9461045 and the G allele of rs3212236 were associated with RD and multiple measures of reading ability and reside on the previously identified RD risk haplotype (Francks et al. 2004). The G allele of rs9461045 was shown to create a binding site for the transcriptional repressor OCT-1 (Dennis et al. 2009). Thus this allele was suggested as the risk allele contributing to RD.

However, there are inconsistencies in the *KIAA0319* haplotype association findings across studies. Cope et al. (2005) could not replicate the association to the *G-C-C* haplotype, and instead identified a RD risk haplotype *G-A*, consisting of different *KIAA0319* markers, rs4504469 with rs6935076. Luciano et al. (2007) also tested haplotypes consisting of the markers rs4504469, rs2038137, and rs2143340, and found evidence for association of the haplotypes *G-C-T* and *A-C-T* to poor reading, but not *G-C-C*. In Couto et al. 2010, the RD risk haplotype described by Francks et al. (2004) showed no association with RD, and instead the haplotype *A-A-T* showed association to RD. This is inconsistent with both Francks et al. (2004) and Cope et al. (2005), in which the *A-A-T* haplotype was associated with better reading performance. These inconsistencies suggest that the risk variant resides on multiple haplotypes or there may be multiple RD causal variants residing on different haplotypes.

To identify RD risk variants in *KIAA0319*, we investigated changes occurring in the gene's regulatory elements. Regions of open chromatin mark regulatory elements and DNA variation within these regions may alter gene expression regulation. One way to predict the location of putative DNA regulatory elements is by mapping the locations of DNA-bound regulatory proteins or modified histones by chromatin immunoprecipitation (ChIP) (Ren et al. 2000). The N-terminal residues of histones 3 and 4 are exposed from the nucleosome and subject to modifications such as methylation, acetylation, ubiquitination and phosphorylation, resulting in a 'histone code' that facilitates or impedes the access of transcription factors to the DNA (Strahl and Allis 2000; Hon et al. 2009).

Previously, we identified a putative regulatory element in *KIAA0319* which may harbour regulatory variants (Couto et al. 2010). The 2.7 kb acetylated histone peak that spans the promoter and first untranslated exon, and part of intron 1 of *KIAA0319* was identified by chromatin immunoprecipitation (ChIP) combined with detection by microarrays (ChIP-chip) using an antibody recognizing H3Ac at K9/K14 in a human retinoblastoma cell-line Y79. H3K9ac and H3K14ac are well-studied markers of accessible chromatin, which often contain regulatory elements (Liang et al. 2004; Bernstein et al. 2005; Lim et al. 2005; Kratz

et al. 2010). This acetylated histone peak in *KIAA0319* was also observed in six mouse tissues (Couto et al. 2010), and as a 2.3 kb peak in two additional human cell lines, a neuroblastoma SH-SY5Y and a glio-blastoma, A-172 by ChIP-sequencing (unpublished).

Interestingly, this region is at the centre of *KIAA0319* association findings and overlaps four markers previously associated with RD: JA04 (Kaplan et al. 2002), rs2235676 (Francks et al. 2004), rs9477247 (Francks et al. 2004) and rs2038137 (Francks et al. 2004; Harold et al. 2006). In addition, the marker rs6935076, which is associated to RD in two samples (Cope et al. 2005; Couto et al. 2010) is less than 600 bp away from the acetylated histone region, and therefore may be in linkage disequilibrium with a regulatory variant therein.

The purpose of the current study was to screen this acetylated histone peak region identified in *KIAA0319* for putative functional variants. Our aim was to identify putative RD causal variant(s) that could better explain haplotype associations from previous RD studies. We also tested the previously reported functional variants rs9461045 and rs3212236 for association in our RD sample.

Methods

Participants and assessment

The recruitment and assessment of families participating in this study have been previously described (Luca et al. 2007; Couto et al. 2008, 2010). Briefly, subjects with reading difficulties (aged 6–16) and their families (including siblings aged 6–16) were recruited from local schools in southern Ontario with the majority from the greater Toronto area. Children were excluded if there was evidence of neurological or chronic illness, psychosis, bipolar affective disorder, Tourette syndrome or chronic tics. Children were also excluded if scores on the Performance and Verbal Scales of Wechsler Intelligence Scale for Children (WISC) III were below 80. To qualify, probands had to have English as their first language, or to have been in an English-education system for at least 5 years. The reading measures used in this study have been described previously (Wigg et al. 2004; Luca et al. 2007; Couto et al. 2008, 2010) and the distribution of reading scores for the first 232 families of this study can be found in Luca et al. (2007). Information about behaviour and possible symptoms of neurological, medical, and psychiatric disorders were obtained using a structured interview with parents (Children's Interview for Psychiatric Syndromes) (Weller et al. 2000) and a semi-structured interview with teachers (TTI) (Tannock et al. 2002), supplemented with standardized questionnaires (Conners Parent and Teacher Rating Scales Revised (Conners 1997) and Ontario Child Health Survey Scales-Revised (Boyle et al. 1993)). Single-word reading, phonological decoding skills, and spelling were assessed using the third edition of the Wide Range Achievement Test (WRAT-3), the Woodcock Reading Mastery Tests-Revised (WRMT-R), and the Test of Word Reading Efficiency (TOWRE). The WRAT-3 subtests provide assessments of single-word reading and spelling (Wilkinson 1993). Two subtests of the WRMT-R were used, one to evaluate phonological decoding (Word Attack), and the other to assess single-word reading (Word ID) (Woodcock 1987).

Standardized scores (by age norm) of three tests were used to classify probands and siblings for categorical analysis of RD: (i) the third edition of the Wide Range Achievement Test

(WRAT III)—Reading subtest, (ii) the Woodcock Reading Mastery Tests-Revised (WRMT-R)—Word Attack subtest, and (iii) WRMT-R—Word Identification subtest. Subjects were considered to be affected with categorical RD if they scored (i) 1.5 standard deviations below average on two of the three tests, or (ii) 1.0 standard deviation below average on all three tests. Of the 352 families recruited to the study, 156 probands and 25 siblings met these criteria and were used in the analyses described in this study. These criteria identified a subset of individuals in our sample whose scores on the three core reading measures, on average, fall within the lower 5% tail of normally distributed reading ability in the general population.

The majority of the families recruited (68.1%) claim European ancestry. Another 26% of the parents describe their ancestry as Canadian Caucasian (no ethnicity specified), 1.8% as Latin American, 2.9% as non-European, and 1.2% as mixed non-European. The study was approved by the research ethics board of the Hospital for Sick Children and all participants provided written consent to participate.

Screening of putative regulatory element by PCR and DNA sequencing

DNA from subjects, siblings and parents was extracted from peripheral blood lymphocytes using a high salt extraction method as previously described (Miller et al. 1988). Forty-eight Caucasian probands that met the categorical criteria for RD were used to screen the 5′ region of open chromatin identified by Couto et al. 2010. This region spans chr6: 24752922–24755657 (Mar. 2006 build NCBI36/hg18) and was screened using four sets of overlapping primers (Supplementary Table S1).

PCR reactions were performed in a total volume of 20 μl, with 60 ng of genomic DNA, 50 ng of each primer (forward and reverse), 2 μl of PCR buffer, 0.2 mM dNTP, 1.5 mM MgCl2 and 5 U of Taq polymerase. PCR thermal cycling stages were as follows: (i) 94°C for 6 min (initial denaturing of DNA), (ii) 35 cycles of 94°C for 20 s, X°C for 20 s, and 72°C for 45 s (where X°C is the annealing temperature of the specific primers), (iii) 72°C for 5 min (extension).

Prior to DNA sequencing, 5 μl of the PCR reaction was treated with exonuclease I (0.2 μl, 20 U/μl) (New England Biolabs, NEB, Ipswich, MA) and Antarctic phosphatase (0.55 μl, 5 U/μl) (NEB) in 0.55 μl H_2O, for 15 min at 37 C, then for 15 min at 80 C.

Polymorphism genotyping

SNP genotyping was conducted using TaqMan assays from Applied Biosystems Inc. (ABI, Foster City, CA) (Supplementary Tables S2 and S3). The PCR reaction mix for the TaqMan assays contained: 30 ng of DNA, 2.5 μl of TaqMan Genotyping PCR Master Mix, and 0.05 μl of allelic discrimination mix (which contained 36 μM of each primer and 8 μM of each probe) in a total reaction volume of 5 μl. The PCR thermal cycling stages were as follows: (i) 50°C for 2 min, (ii) 95°C for 10 min, then (iii) 60 cycles of: 95°C for 15 s, and 59°C for 1 min. The generated genotype data was interpreted by the ABI 7900-HT Sequence Detection System® (SDS), using the allelic discrimination end-point analysis mode of the SDS software (version 2.0; Applied Biosystems).

The microsatellite marker, JA04, was amplified using a fluorescently labeled forward primer (Supplementary Table S4), then genotyped using the Applied Biosystems AB3730 DNA Analyzer. Genotyping results were analyzed with the GeneMapper software (Applied Biosystems, Foster City, CA).

Statistical tests

The TDT statistic in Haploview (v.4.0) (Barrett et al. 2005) was used to test for association of biallelic markers, while the Extended-TDT (ETDT) was used for the analysis of the multi-allelic microsatellite (Sham and Curtis 1995). The Haploview program (v.4.0) was also used to analyze and visualize linkage disequilibrium (LD) blocks using the Gabriel et al. algorithm (Gabriel et al. 2002; Barrett et al. 2005). Haplotype transmission was analyzed using the TRANSMIT program (Clayton and Jones 1999). Merlin was used to check the genotype data for unlikely crossover events indicating erroneous genotypes (Abecasis et al. 2002). This data set was free of any detectable Mendelian errors and there were no crossovers between markers. None of the markers used in the analyses deviated from Hardy–Weinberg equilibrium (HWE).

We estimated the power of the sample using the genetic power calculator available at http://pngu.mgh.harvard.edu/~purcell/cgi-bin/dtdt.cgi. Since the assumption of the screening studies in *KIAA0319* was that we are testing putative causative polymorphisms, the test marker was assumed to be the causative marker (identical allele frequencies, in complete LD). Assuming an additive or dominant inheritance model, our family-based RD sample has acceptable power (0.8) to detect a genotype relative risk of 2.0 for a causative marker with an allele frequency ranging between 0.13 and 0.27.

The program Evolutionary-based Haplotype Analysis Package (eHAP, Seltman et al. 2003) was used to determine the phylogenetic relationship between haplotypes based on the genotype data. This tool was used to build a cladogram (Fig. 1).

Transcription factor binding site predictions

The effects of putative functional alleles on transcription factor binding motifs were predicted in silico with Transcription Element Search System (TESS), which uses transcription factor binding sequences, consensus strings and positional weightmatrices from the TRANSFAC (version 6), JASPAR IMD (version 1.1) (Vlieghe et al. 2006), and CBIL-Gibbs Mat (version 1.1) databases (Schug 2008). The allele from the transcribed strand (in this case, the reverse strand) was submitted for analysis, along with 10–15 bp flanking sequence. The basic search parameters were used. “Perfect-match” (secondary Lg-likelihood deficit <0.1), and pretty-good-match (secondary Lg-likelihood deficit <1.0), Transcription Factor Binding Site (TFBS) results were included, which were then filtered for human transcription factors with roles in brain development or function.

Results

Screening of the 5′ region and selection of variants for genotyping

The purpose of this study was to identify putative functional polymorphisms within the 5′ region of *KIAA0319* that could explain the evidence for association of this gene to dyslexia. To identify polymorphisms, a region of *KIAA0319* previously identified as a region of open chromatin spanning the promoter, first untranslated exon and part of intron 1 (chr6: 24753560–24755536, Mar. 2006 build NCBI36/hg18) was sequenced using overlapping PCR products amplified from 48 Caucasian RD probands. Seventeen polymorphisms were observed (Table 1), consisting of 14 annotated SNPs reported in public databases, one SNP previously reported in a screen of this region (Dennis et al. 2009), a microsatellite, and a 26 bp deletion (rs71815143). This 26 bp deletion consisted of a repetitive sequence with a periodicity of 13 bp. The SNP rs6456625 (C/T), and the unannotated SNP (C/T; chr6: 24754493 (Mar 2006 NCBI36/hg18)) occurred within this deletion polymorphism.

Allele frequencies for the polymorphisms were calculated from the sequencing data in the 48 RD probands. Allele frequencies were available for seven of these polymorphisms in dbSNP (Table 1). Four of the SNPs had previously been tested for association in our RD sample rs9467247, rs2038137, rs2235677 and rs2235676 (Couto et al. 2010), and showed no significant association with RD (Table 1). Of the remaining DNA variants, markers were selected for genotyping in the RD family sample, if there was (i) a minor allele frequency greater than 0.05 (rs1883593, rs62400773, and rs13220596 excluded), (ii) a difference in allele frequencies between probands and those reported in dbSNP (http://www.ncbi.nlm.nih.gov/projects/SNP/) for CEU, (Caucasian individuals with Northern and Western European ancestry from the CEPH collection) (rs2038135 and rs2038136 not selected for genotyping). The marker rs2038135 was also in strong LD with rs2038137, which was not significantly associated with RD.

The SNPs rs35882364/rs35853270 were also not selected for genotyping. The SNPs rs35882364 and rs35853270 are adjacent single thymidine deletions and there were no allele frequencies available in the database for these SNPs. It is unclear whether rs35882364/rs35853270 represent two SNPs (as reported in the database), or one SNP. In the screened probands, the sequence at this locus was observed as “TT” or “T0”, where ‘0’ represents a deleted T. Since the “00” genotype was not observed, it is possible that one of the two SNPs is not polymorphic, or that both SNPs have low MAF. Therefore, no genotyping assay could be designed and these SNPs were not tested for association in the RD family sample. Moreover, there were no allele-specific TFBS predicted changes for these polymorphisms.

Since the functional risk allele may affect transcription factor binding affinity, the TFBS changes for all of the polymorphisms were predicted in silico using TESS. It is important to note that TESS utilizes (1) databases containing experimentally validated TFBS sequences to predict the TFBS, and (2) tools to evaluate the significance of the predicted sites, but that neither of these allow for total accuracy in TFBS predictions. Some TFBS will be missed, due to the incompleteness of experimentally validated TFBS sequences, while other TFBS will be predicted, but not be biologically relevant. To assess which TFBS are most likely to be biologically functional, the transcription factors were filtered for function/expression in

the brain (Table 1). Brain-relevant TFBS changes were predicted to occur for most polymorphisms including: Sp1, HOXA5, FOXL1, FOXC1, C/EBP, POU5F1, Stat6, s-Ets-2, RAR, AP-1 and GATA2 (Table 1). Of the polymorphisms selected for genotyping, several were of particular interest because of the predicted change in TFBS. These include the HOXA5 binding site predicted for the minor allele (C) of rs1883594, but not the major allele (G). HOXA5 is a transcriptional regulator during development and is expressed early in embryonic neurons (Joksimovic et al. 2005). The insertion allele of rs71815143 is predicted to contain three Sp1 binding sites and a CCAAT/enhancer binding protein (C/EBP) binding site, which are not predicted for the deletion allele. Sp1 is a ubiquitous transcriptional activator, except in the CNS, where it is expressed exclusively in early forebrain neurons (Mao et al. 2009). C/EBP recruits transcriptional activators to enhancer and promoter regions (Calella et al. 2007).

Association analysis

Seven polymorphisms within the screened region of *KIAA0319* (rs1883592, rs1883594, rs3756821, rs718 15143, rs6456625 and 6: 24754493 in Table 1, and JA04 in Table 2), along with two SNPs outside the region (rs9461045 and rs3212236 in Table 3) were tested for association to categorical RD as single markers in the entire sample of families meeting the criteria ($n = 156$ nuclear families including 25 affected siblings). The markers rs9461045 and rs3212236 were identified as functional RD risk variants in a previous study (Dennis et al. 2009).

The most significant evidence for association was with the marker JA04, in particular we observed biased transmission of allele 14 (25GT) ($\chi^{2} = 7.367$, $p = 0.0067$). The A allele of rs1883592, the G allele of rs3756821 and the insertion allele of rs71815143 showed trends for biased transmission ($p = 0.081$, $p = 0.080$ and $p = 0.061$ respectively) to RD probands, whereas the marker rs1883594 and 6: 24754493 showed no evidence for association (Table 1). The markers rs9461045 and rs3212236, previously reported as associated by Dennis et al. (2009), showed no evidence for association to RD in our sample (Table 3). The *p*-values presented have not been corrected for multiple testing.

Haplotype analysis

We examined haplotypes of the polymorphisms tested in this study with additional markers to determine if these combinations could explain the previous reported haplotype findings, and further to determine if the risk is related to combinations of alleles as opposed to single alleles. The JA04 microsatellite was selected for haplotype analysis because it showed the strongest statistical evidence for association to RD and was predicted to function in gene transcription regulation. Although rs71815143 did not show strong statistical evidence for association to RD, this variant was included in haplotype analysis because of the predicted effects of this deletion on TFBS. No haplotype combinations were analyzed other than the ones presented here. Transmit TDT analysis was used to test whether the alleles of JA04 and rs71815143 contributed to the previously identified RD risk haplotypes generated by the markers rs4504469, rs2038137 and rs2143340 (Francks haplotype) and rs4504469 and rs6935076 (Cope haplotype), which were genotyped in our RD sample in a previous study (Couto et al. 2010). The results of the haplotype analyses can be found in Tables 4 and 5.

Results showed that the associated alleles of JA04 and rs71815143 contributed only to one Francks haplotype (Table 4). This haplotype consisted of the alleles A-A-T, which had a trend for association in a previous study ($p = 0.091$, Couto et al. unpublished). Including the two markers JA04 and rs71815143 in this haplotype (A-A-25GT-ins-T) increased the evidence for association of this haplotype to RD ($p = 0.0093$).

Couto et al. 2010 had previously identified a risk haplotype (A-G; $p = 0.018$) consisting of the markers reported by Cope et al. (2005). Including the markers JA04 and rs71815143 (A-G-25GT-ins) increased the evidence for association of this haplotype ($p = 0.0010$) (Table 5). This was the only haplotype carrying the alleles biased in transmission in the single marker analyses of JA04 and rs71815143. The haplotype consisting of alternative alleles (G-A-11GT-del) showed a trend for under-transmission to RD ($p = 0.08$).

The LD across the markers in the 5′ region was determined in our RD family-based sample using Haploview (version 4.0). The markers rs9461045 and rs3212236 which are 5′ of *KIAA0319*, and the marker rs6935076, which is located 600 bp distal to the acetylated H3 peak in the first intron, were included in the analysis. The marker rs6935076 was associated in this RD sample in a previous study (Couto et al. 2010). Table 6 shows the strong LD across the region, thus the region can be analyzed as one LD block.

To determine whether a combination of alleles across the putative regulatory element showed association with RD, haplotype analysis was performed for markers in this LD block. Haploview identified tag-markers: rs2038137, rs3756821, and rs9461045. In addition, we included the markers with evidence for association: rs6935076, JA04 and rs71815143. The haplotype that is biased in transmission ($p = 0.0013$) carries the risk alleles of the individually associated markers (Table 7).

Strong LD as measured by D' across the associated region makes it difficult to determine which markers distinguish the risk haplotype from unassociated haplotypes. Therefore, eHap was used to determine the phylogenetic relationships between the polymorphisms in the affected RD probands. The analyzed haplotypes consisted of the markers: rs6935076, rs2038137, rs71815143, rs3756821 and rs9461045 (Fig. 1).

The results show that the alleles distinguishing the associated haplotype from an ancestral haplotype are the insertion allele of rs71815143 and alleles 13–19 of JA04. However, the ancestral haplotype was only observed at a frequency of 0.02, thus the frequency is too low for association testing. Therefore, it is unclear whether the ancestral haplotype is also associated with RD.

Discussion

The purpose of this study was to identify putative functional polymorphism(s) that explain the previous association to RD. We hypothesized that causal risk variants in *KIAA0319* function by altering transcription, and are located in the region of open chromatin identified by Couto et al. 2010. In the current study, this region was screened, and putative functional variants were selected and tested for association to RD.

Two putative functional variants in the 5′ region of *KIAA0319* were identified: a microsatellite (JA04) and a 13 bp repeat (rs71815143). An unannotated polymorphism was identified at 6: 24754493 within the deletion allele of rs71815143. Although the insertion allele of rs71815143 showed a trend for association ($p = 0.061$), the polymorphism at 6: 24754493 in the deletion allele of rs71815143 showed no association with RD. This SNP was also reported in the Dennis et al. (2009) study, and no evidence for association was reported.

Although rs71815143 was not statistically associated with RD ($p = 0.061$), TFBS predictions suggest it may have allele-specific affects on gene expression. The risk allele of rs71815143 may function to augment risk for RD by influencing *KIAA0319* expression. The risk allele was predicted to bind Sp1 and C/EBP. C/EBP activates gene expression by binding promoters and enhancer elements, and Sp1 is a transcriptional activator known to activate RNA synthesis from selected promoters by RNA polymerase II (Briggs et al. 1986; Landschulz et al. 1988). In addition, Sp1 can bind and activate gene expression from TATA-less promoters (Badran et al. 2005), such as the promoter of *KIAA0319*.

Allele 14 (25 GT repeats) of the microsatellite showed the most evidence for association to RD ($p = 0.0067$). This may not replicate the previous finding of Kaplan et al. (2002) who reported association with the most common allele of JA04, which would be 11GT repeats in our sample. However, we cannot be absolutely sure that the most common allele in that sample is the same as the most common in ours because the allele size was not specified in that paper.

Microsatellites have been implicated in the regulation of transcription (Hamada et al. 1984; Naylor and Clark 1990; Gebhardt et al. 1999; Rothenburg et al. 2001; Vinces et al. 2009). The occurrence of these repeats in the 5′ region of genes more often than expected by chance has led some to speculate that they are involved in rapid gene evolution and responsiveness to changing environmental conditions (King et al. 1997). There are numerous ways that the microsatellite may increase or decrease gene expression: (i) relief of DNA supercoiling; (ii) altering nucleosome placement; or (iii) forming a joint to bridge the promoter with more distal regulatory elements. Any combinations of these scenarios could be occurring in the *KIAA0319* gene. Alternating purine/pyrimidine repeats, such as GT repeats, are conducive to a transition from the normal, B form of DNA, to the Z form. Unlike the B form, Z-DNA double helix is left-handed, contains 12 bp per turn (instead of 10.4 bp) and exists only transiently. Different lengths of GT repeats form different amounts of Z-DNA, at different rates, suggesting that polymorphic dinucleotide repeats may be one of the genetic factors contributing to variation in gene expression. The transition to one helical turn of Z-DNA can relieve 1.8 helical turns of right-handed negative superhelical turns (Aboul-ela et al. 1992). In eukaryotes, this transition to Z-DNA can result in increased or decreased gene expression.

Z-DNA can also alter nucleosome placement. It has been shown that while the B-form was stably packaged into nucleosomes, the Z-form prevented nucleosome assembly in the in vitro conditions (Nickol et al. 1982). The clearance of nucleosomes increases the promoter's accessibility to transcription machinery. Liu et al. (2001, 2006) showed in vivo that when the

CSF1 gene is silenced, the promoter of *CSF1* and a nearby GT-repeat (B-DNA) are assembled within a nucleosome. Upon recruitment of BRG1, a subunit of SWI2/SNF2-like chromatin-remodeling complex BAF the GT repeat converted to Z-DNA, disrupting the position of the nucleosomes, and gene activation. BRG1 binding induced Z-DNA transition by inducing superhelical torsion (Liu et al. 2001, 2006).

In this study, the 25GT repeat allele of the microsatellite allele showed evidence for association to RD. However, it is unclear how the 25GT length, and not other lengths, could increase RD risk. Studies on the B-to-Z DNA transition suggest that longer alleles may form longer stretches of Z-DNA and elicit a stronger effect on gene expression. Another possibility is that the 25GT length represents a threshold at which the repeat is long enough to transition to Z-DNA in an energetically favourable manner. However, the TDT results for the microsatellite did not support that longer alleles of the microsatellite are associated with RD.

In spite of their predicted effects on transcription, evidence from the current study suggests that rs71815143 and JA04 are not functional variants that explain the association in other RD samples. This is because neither of these markers could resolve the inconsistency in haplotype associations in previous studies. Instead, it is likely these variants are in strong LD with the causative allele(s).

We expect that the causative allele will segregate onto multiple risk haplotypes. This would resolve the inconsistency in the haplotype findings in the literature, as one explanation is that the functional risk allele segregates on multiple haplotypes at different frequencies, in different RD samples. The associated alleles of the microsatellite (25GT) and rs71815143 (insertion allele) only segregated with the A-A-T (Francks) haplotype and the A-G (Cope) haplotype. These haplotypes have the alternative alleles of those previously identified by the other groups, suggesting that neither of these markers can explain the association finding in the Francks et al. (2004) or Cope et al. (2005) studies.

Luciano et al. (2007) had suggested that the haplotype of the alternative alleles, G-C-T, of the markers used in the Francks haplotype was the major risk haplotype, as some evidence for over-transmission was observed in Luciano et al. (2007), Cope et al. (2005) and Francks et al. (2004). This haplotype carries the non-risk alleles of rs71815143 (deletion) and the microsatellite (11GT and 12GT), and showed no bias in transmission in our RD sample. Therefore, these variants are not likely to be the common functional risk variants in the other RD samples studied.

It is possible that the risk alleles reside outside the region screened here. The Dennis et al. (2009) study identified an RD-associated SNP (rs9461045) that creates a binding site for the transcriptional repressor OCT-1. New elements such as this created by the risk alleles would only be identified in chromatin from tissues or cell lines carrying the risk alleles and with markers of repressed chromatin.

In summary, several putative functional variants were identified including a microsatellite (*JA04*) and rs71815143, of which *JA04* showed evidence for association to RD. The microsatellite may create risk for RD by transitioning from B to Z-DNA, which can increase

or decrease gene expression, depending on the gene context. The risk alleles of these two markers both contributed to a risk haplotype in our sample; however this risk haplotype is inconsistent with the haplotypes associated in other RD samples. Therefore, these two putative functional variants are not likely to be the main functional risk variants. The functional variants identified by Dennis et al. (2009) were not associated in our sample, and therefore we find no support in our families for these alleles as the functional variants contributing to RD.

The *p*-values for the single marker association tests would not withstand a correction for multiple testing. Given the strong LD observed across the 5′ region of *KIAA0319*, haplotype analysis is a more appropriate test to determine whether there is evidence for association to RD in this region. The haplotype analysis results suggest that there is evidence for significant haplotype associations to RD, even in light of multiple testing.

We found association for RD with haplotypes of polymorphisms in the 5′ region. Although the main functional risk variant across different samples has not been identified, the results of this study, and those of others (Paracchini et al. 2006; Dennis et al. 2009) suggest that the functional RD risk variants reside in the 5′ region and alter gene expression of *KIAA0319*. It is becoming more evident with complex traits/disorders that functional risk variants modify gene expression, as opposed to changing protein coding sequences (Jais 2005; Nica and Dermitzakis 2008). Studies of *KIAA0319* in RD can serve as a model for the study of other complex traits and disorders. Future studies in *KIAA0319* will determine the functional effects of these polymorphisms and haplotypes of these alleles on gene expression.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin—rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet. 2002; 30(1):97–101. [PubMed: 11731797]

Aboul-ela F, Bowater RP, Lilley DM. Competing B-Z and helix-coil conformational transitions in supercoiled plasmid DNA. J Biol Chem. 1992; 267(3):1776–1785. [PubMed: 1730717]

Badran BM, Kunstman K, Stanton J, Moschitta M, Zerghe A, Akl H, Burny A, Wolinsky SM, Willard-Gallo KE. Transcriptional regulation of the human CD3 gamma gene: the TATA-less CD3 gamma promoter functions via an initiator and contiguous Sp-binding elements. J Immunol. 2005; 174(10): 6238–6249. [PubMed: 15879122]

Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics. 2005; 21(2):263–265. [PubMed: 15297300]

Bernstein BE, Kamal M, Lindblad-Toh K, Bekiranov S, Bailey DK, Huebert DJ, McMahon S, Karlsson EK, Kulbokas EJ III, Gingeras TR, Schreiber SL, Lander ES. Genomic maps and comparative analysis of histone modifications in human and mouse. Cell. 2005; 120(2):169–181. [PubMed: 15680324]

Boyle MH, Offord DR, Racine Y, Fleming JE, Szatmari P, Sanford M. Evaluation of the revised Ontario Child Health Study scales. J Child Psychol Psychiatry. 1993; 34(2):189–213. [PubMed: 8444992]

Briggs MR, Kadonaga JT, Bell SP, Tjian R. Purification and biochemical characterization of the promoter-specific transcription factor, Sp1. Science. 1986; 234(4772):47–52. [PubMed: 3529394]

Burbridge TJ, Wang Y, Volz AJ, Peschansky VJ, Lisann L, Galaburda AM, Lo Turco JJ, Rosen GD. Postnatal analysis of the effect of embryonic knockdown and overexpression of candidate dyslexia susceptibility gene homolog Dcdc2 in the rat. Neuroscience. 2008; 152(3):723–733. [PubMed: 18313856]

Calella AM, Nerlov C, Lopez RG, Sciarretta C, von Bohlen und Halbach O, Bereshchenko O, Minichiello L. Neurotrophin/Trk receptor signaling mediates C/EBPalpha, -beta and NeuroD recruitment to immediate-early gene promoters in neuronal cells and requires C/EBPs to induce immediate-early gene transcription. Neural Dev. 2007; 2:4. [PubMed: 17254333]

Cardon LR, Smith SD, Fulker DW, Kimberling WJ, Pennington BF, DeFries JC. Quantitative trait locus for reading disability on chromosome 6. Science. 1994; 266(5183):276–279. [PubMed: 7939663]

Clayton D, Jones H. Transmission/disequilibrium tests for extended marker haplotypes. Am J Hum Genet. 1999; 65(4):1161–1169. [PubMed: 10486335]

Conners, CK. Conners' Rating Scales—Revised. Multi-Health Systems Inc; Toronto: 1997.

Cope N, Harold D, Hill G, Moskvina V, Stevenson J, Holmans P, Owen MJ, O'Donovan MC, Williams J. Strong evidence that KIAA0319 on chromosome 6p is a susceptibility gene for developmental dyslexia. Am J Hum Genet. 2005; 76(4):581–591. [PubMed: 15717286]

Couto JM, Gomez L, Wigg K, Cate-Carter T, Archibald J, Anderson B, Tannock R, Kerr EN, Lovett MW, Humphries T, Barr CL. The KIAA0319-like (KIAA0319L) gene on chromosome 1p34 as a candidate for reading disabilities. J Neurogenet. 2008; 22(4):295–313. [PubMed: 19085271]

Couto JM, Livne-Bar I, Huang K, Xu Z, Cate-Carter T, Feng Y, Wigg K, Humphries T, Tannock R, Kerr EN, Lovett MW, Bremner R, Barr CL. Association of reading disabilities with regions marked by acetylated H3 histones in KIAA0319. Am J Med Genet B Neuropsychiatr Genet. 2010 Mar 5; 153B(2):447–62. [PubMed: 19588467]

Deffenbacher KE, Kenyon JB, Hoover DM, Olson RK, Pennington BF, DeFries JC, Smith SD. Refinement of the 6p21.3 quantitative trait locus influencing dyslexia: linkage and association analyses. Hum Genet. 2004; 115(2):128–138. [PubMed: 15138886]

DeFries, JC., Gillis, JJ. Genetics of reading disability. In: Plomin, R., McClearn, GE., editors. Nature, nurture, and psychology. American Psychological Association; Washington, DC: 1993. p. 121-145.

Dennis MY, Paracchini S, Scerri TS, Prokunina-Olsson L, Knight JC, Wade-Martins R, Coggill P, Beck S, Green ED, Monaco AP. A common variant associated with dyslexia reduces expression of the KIAA0319 gene. PLoS Genet. 2009; 5(3):e1000436. [PubMed: 19325871]

Fisher SE, Marlow AJ, Lamb J, Maestrini E, Williams DF, Richardson AJ, Weeks DE, Stein JF, Monaco AP. A quantitative-trait locus on chromosome 6p influences different aspects of developmental dyslexia. Am J Hum Genet. 1999; 64(1):146–156. [PubMed: 9915953]

Francks C, Paracchini S, Smith SD, Richardson AJ, Scerri TS, Cardon LR, Marlow AJ, MacPhie IL, Walter J, Pennington BF, Fisher SE, Olson RK, DeFries JC, Stein JF, Monaco AP. 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. 2004; 75(6):1046–1058. [PubMed: 15514892]

Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ, Altshuler D. The structure of haplotype blocks in the human genome. Science. 2002; 296(5576):2225–2229. [PubMed: 12029063]

Galaburda AM, Sherman GF, Rosen GD, Aboitiz F, Geschwind N. Developmental dyslexia: four consecutive patients with cortical anomalies. Ann Neurol. 1985; 18(2):222–233. [PubMed: 4037763]

Gayan J, Smith SD, Cherny SS, Cardon LR, Fulker DW, Brower AM, Olson RK, Pennington BF, DeFries JC. Quantitative-trait locus for specific language and reading deficits on chromosome 6p. Am J Hum Genet. 1999; 64(1):157–164. [PubMed: 9915954]

Gebhardt F, Zanker KS, Brandt B. Modulation of epidermal growth factor receptor gene transcription by a polymorphic dinucleotide repeat in intron 1. J Biol Chem. 1999; 274(19):13176–13180. [PubMed: 10224073]

Grigorenko EL, Wood FB, Meyer MS, Hart LA, Speed WC, Shuster A, Pauls DL. Susceptibility loci for distinct components of developmental dyslexia on chromosomes 6 and 15. Am J Hum Genet. 1997; 60(1):27–39. [PubMed: 8981944]

Grigorenko EL, Wood FB, Meyer MS, Pauls DL. Chromosome 6p influences on different dyslexia-related cognitive processes: further confirmation. Am J Hum Genet. 2000; 66:715–723. [PubMed: 10677331]

Hamada H, Seidman M, Howard BH, Gorman CM. Enhanced gene expression by the poly(dT-dG).poly(dC-dA) sequence. Mol Cell Biol. 1984; 4(12):2622–2630. [PubMed: 6098815]

Harlaar N, Spinath FM, Dale PS, Plomin R. Genetic influences on early word recognition abilities and disabilities: a study of 7-year-old twins. J Child Psychol Psychiatry. 2005; 46(4):373–384. [PubMed: 15819646]

Harold D, Paracchini S, Scerri T, Dennis M, Cope N, Hill G, Moskvina V, Walter J, Richardson AJ, Owen MJ, Stein JF, Green ED, O'Donovan MC, Williams J, Monaco AP. Further evidence that the KIAA0319 gene confers susceptibility to developmental dyslexia. Mol Psychiatry. 2006; 11(12): 1085–1091. 1061. [PubMed: 17033633]

Hawke JL, Wadsworth SJ, DeFries JC. Genetic influences on reading difficulties in boys and girls: the Colorado twin study. Dyslexia. 2006; 12(1):21–29. [PubMed: 16512171]

Hon G, Wang W, Ren B. Discovery and annotation of functional chromatin signatures in the human genome. PLoS Comput Biol. 2009; 5(11):e1000566. [PubMed: 19918365]

Humphreys P, Kaufmann WE, Galaburda AM. Developmental dyslexia in women: neuropathological findings in three patients. Ann Neurol. 1990; 28(6):727–738. [PubMed: 2285260]

Jais PH. How frequent is altered gene expression among susceptibility genes to human complex disorders? Genet Med. 2005; 7(2):83–96. [PubMed: 15714075]

Joksimovic M, Jeannotte L, Tuggle CK. Dynamic expression of murine HOXA5 protein in the central nervous system. Gene Expr Patterns. 2005; 5(6):792–800. [PubMed: 15922675]

Kaplan DE, Gayan J, Ahn J, Won TW, Pauls D, Olson RK, DeFries JC, Wood F, Pennington BF, Page GP, Smith SD, Gruen JR. Evidence for linkage and association with reading disability on 6p21.3-22. Am J Hum Genet. 2002; 70(5):1287–1298. [PubMed: 11951179]

King DG, Soller M, Kashi Y. Evolutionary tuning knobs. Endeavour. 1997; 21(1):36–40.

Kratz A, Arner E, Saito R, Kubosaki A, Kawai J, Suzuki H, Carninci P, Arakawa T, Tomita M, Hayashizaki Y, Daub CO. Core promoter structure and genomic context reflect histone 3 lysine 9 acetylation patterns. BMC Genomics. 2010; 11:257. [PubMed: 20409305]

Landschulz WH, Johnson PF, Adashi EY, Graves BJ, McKnight SL. Isolation of a recombinant copy of the gene encoding C/EBP. Genes Dev. 1988; 2(7):786–800. [PubMed: 2850264]

Liang G, Lin JC, Wei V, Yoo C, Cheng JC, Nguyen CT, Weisenberger DJ, Egger G, Takai D, Gonzales FA, Jones PA. Distinct localization of histone H3 acetylation and H3-K4 methylation to the transcription start sites in the human genome. Proc Natl Acad Sci USA. 2004; 101(19):7357–7362. [PubMed: 15123803]

Lim JH, West KL, Rubinstein Y, Bergel M, Postnikov YV, Bustin M. Chromosomal protein HMGN1 enhances the acetylation of lysine 14 in histone H3. EMBO J. 2005; 24(17):3038–3048. [PubMed: 16096646]

Liu R, Liu H, Chen X, Kirby M, Brown PO, Zhao K. Regulation of CSF1 promoter by the SWI/SNF-like BAF complex. Cell. 2001; 106(3):309–318. [PubMed: 11509180]

Liu H, Mulholland N, Fu H, Zhao K. Cooperative activity of BRG1 and Z-DNA formation in chromatin remodeling. Mol Cell Biol. 2006; 26(7):2550–2559. [PubMed: 16537901]

Luca P, Laurin N, Misener VL, Wigg KG, Anderson B, Cate-Carter T, Tannock R, Humphries T, Lovett MW, Barr CL. Association of the dopamine receptor D1 gene, DRD1, with inattention symptoms in families selected for reading problems. Mol Psychiatry. 2007; 12:776–785. [PubMed: 17310237]

Luciano M, Lind PA, Duffy DL, Castles A, Wright MJ, Montgomery GW, Martin NG, Bates TC. A Haplotype Spanning KIAA0319 and TTRAP Is Associated with Normal Variation in Reading and Spelling Ability. Biol Psychiatry. 2007; 26:26.

Mao XR, Moerman-Herzog AM, Chen Y, Barger SW. Unique aspects of transcriptional regulation in neurons–nuances in NFkappaB and Sp1-related factors. J Neuroinflammation. 2009; 6:16. [PubMed: 19450264]

McGrath LM, Smith SD, Pennington BF. Breakthroughs in the search for dyslexia candidate genes. Trends Mol Med. 2006; 12(7):333–341. [PubMed: 16781891]

Meng H, Smith SD, Hager K, Held M, Liu J, Olson RK, Pennington BF, DeFries JC, Gelernter J, O'Reilly-Pol T, Somlo S, Skudlarski P, Shaywitz SE, Shaywitz BA, Marchione K, Wang Y, Paramasivam M, LoTurco JJ, Page GP, Gruen JR. DCDC2 is associated with reading disability and modulates neuronal development in the brain. Proc Natl Acad Sci USA. 2005; 102(47):17053–17058. [PubMed: 16278297]

Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res. 1988; 16(3):1215. [PubMed: 3344216]

Naylor LH, Clark EM. d(TG)n.d(CA)n sequences upstream of the rat prolactin gene form Z-DNA and inhibit gene transcription. Nucleic Acids Res. 1990; 18(6):1595–1601. [PubMed: 2158081]

Nica AC, Dermitzakis ET. Using gene expression to investigate the genetic basis of complex disorders. Hum Mol Genet. 2008; 17(R2):R129–R134. [PubMed: 18852201]

Nickol J, Behe M, Felsenfeld G. Effect of the B–Z transition in poly(dG-m5dC). poly(dG-m5dC) on nucleosome formation. Proc Natl Acad Sci USA. 1982; 79(6):1771–1775. [PubMed: 6952228]

Paracchini S, Thomas A, Castro S, Lai C, Paramasivam M, Wang Y, Keating BJ, Taylor JM, Hacking DF, Scerri T, Francks C, Richardson AJ, Wade-Martins R, Stein JF, Knight JC, Copp AJ, Loturco J, Monaco AP. The chromosome 6p22 haplotype associated with dyslexia reduces the expression of KIAA0319, a novel gene involved in neuronal migration. Hum Mol Genet. 2006; 15(10):1659–1666. [PubMed: 16600991]

Paracchini S, Scerri T, Monaco AP. The genetic lexicon of dyslexia. Annu Rev Genomics Hum Genet. 2007; 19:19.

Paracchini S, Steer CD, Buckingham LL, Morris AP, Ring S, Scerri T, Stein J, Pembrey ME, Ragoussis J, Golding J, Monaco AP. Association of the KIAA0319 dyslexia susceptibility gene with reading skills in the general population. Am J Psychiatry. 2008

Peschansky VJ, Burbridge TJ, Volz AJ, Fiondella C, Wissner-Gross Z, Galaburda AM, Turco JJ, Rosen GD. The effect of variation in expression of the candidate dyslexia susceptibility gene homolog Kiaa0319 on neuronal migration and dendritic morphology in the rat. Cereb Cortex. 2010; 20(4):884–897. [PubMed: 19679544]

Ren B, Robert F, Wyrick JJ, Aparicio O, Jennings EG, Simon I, Zeitlinger J, Schreiber J, Hannett N, Kanin E, Volkert TL, Wilson CJ, Bell SP, Young RA. Genome-wide location and function of DNA binding proteins. Science. 2000; 290(5500):2306–2309. [PubMed: 11125145]

Rothenburg S, Koch-Nolte F, Rich A, Haag F. A polymorphic dinucleotide repeat in the rat nucleolin gene forms Z-DNA and inhibits promoter activity. Proc Natl Acad Sci USA. 2001; 98(16):8985–8990. [PubMed: 11447254]

Scerri TS, Schulte-Korne G. Genetics of developmental dyslexia. Eur Child Adolesc Psychiatry. 2010; 19(3):179–197. [PubMed: 20091194]

Schug J. Using TESS to predict transcription factor binding sites in DNA sequence. Current protocols in bioinformatics. 2008; Chapter 2(Unit 2.6)

Schumacher J, Anthoni H, Dahdouh F, Konig IR, Hillmer AM, Kluck N, Manthey M, Plume E, Warnke A, Remschmidt H, Hulsmann J, Cichon S, Lindgren CM, Propping P, Zucchelli M, Ziegler A, Peyrard-Janvid M, Schulte-Korne G, Nothen MM, Kere J. Strong genetic evidence of DCDC2 as a susceptibility gene for dyslexia. Am J Hum Genet. 2006; 78(1):52–62. [PubMed: 16385449]

Seltman H, Roeder K, Devlin B. Evolutionary-based association analysis using haplotype data. Genet Epidemiol. 2003; 25(1):48–58. [PubMed: 12813726]

Sham PC, Curtis D. An extended transmission/disequilibrium test (TDT) for multi-allele marker loci. Ann Hum Genet. 1995; 59(Pt 3):323–336. [PubMed: 7486838]

Shaywitz SE, Shaywitz BA, Fletcher JM, Escobar MD. Prevalence of reading disability in boys and girls. Results of the Connecticut Longitudinal Study. JAMA. 1990; 264(8):998–1002. [PubMed: 2376893]

Smith, SD., Kimberling, WJ. Reading disabilities. In: Pennington, B., editor. Genetic and neurological influences. Kluwer Academic; Boston: 1991.

Strahl BD, Allis CD. The language of covalent histone modifications. Nature. 2000; 403(6765):41–45. [PubMed: 10638745]

Tannock, R., Hum, M., Masellis, M., Humphries, T., Schachar, R. Teacher telephone interview for children's academic performance, attention, behavior and learning: DSM-IV Version (TTIIV). The Hospital for Sick Children; Toronto: 2002. Unpublished document

The FBAT Program v1.4.1. http://www.biostat.harvard.edu/~fbat/default.html

Turic D, Robinson L, Duke M, Morris DW, Webb V, Hamshere M, Milham C, Hopkin E, Pound K, Fernando S, Grierson A, Easton M, Williams N, Van Den Bree M, Chowdhury R, Gruen J, Stevenson J, Krawczak M, Owen MJ, O'Donovan MC, Williams J. Linkage disequilibrium mapping provides further evidence of a gene for reading disability on chromosome 6p21.3-22. Mol Psychiatry. 2003; 8(2):176–185. [PubMed: 12610650]

Vinces MD, Legendre M, Caldara M, Hagihara M, Verstrepen KJ. Unstable tandem repeats in promoters confer transcriptional evolvability. Science. 2009; 324(5931):1213–1216. [PubMed: 19478187]

Vlieghe D, Sandelin A, De Bleser PJ, Vleminckx K, Wasserman WW, van Roy F, Lenhard B. A new generation of JASPAR, the open-access repository for transcription factor binding site profiles. Nucleic Acids Res. 2006; 34(Database issue):D95–D97. [PubMed: 16381983]

Wadsworth SJ, Olson RK, Pennington BF, DeFries JC. Differential genetic etiology of reading disability as a function of IQ. J Learn Disabil. 2000; 33(2):192–199. [PubMed: 15505948]

Wadsworth SJ, DeFries JC, Olson RK, Willcutt EG. Colorado longitudinal twin study of reading disability. Ann Dyslexia. 2007; 57(2):139–160. [PubMed: 18060583]

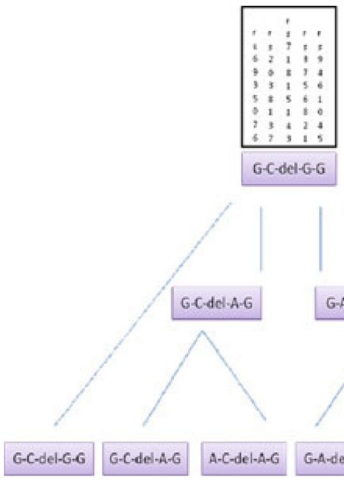
Weller EB, Weller RA, Fristad MA, Rooney MT, Schecter J. Children's Interview for Psychiatric Syndromes (ChIPS). J Am Acad Child Adolesc Psychiatry. 2000; 39(1):76–84. [PubMed: 10638070]

Wigg KG, Couto JM, Feng Y, Anderson B, Cate-Carter TD, Macciardi F, Tannock R, Lovett MW, Humphries TW, Barr CL. Support for EKN1 as the susceptibility locus for dyslexia on 15q21. Mol Psychiatry. 2004; 9(12):1111–1121. [PubMed: 15249932]

Wilcke A, Weissfuss J, Kirsten H, Wolfram G, Boltze J, Ahnert P. The role of gene DCDC2 in German dyslexics. Ann Dyslexia. 2009; 59(1):1–11. [PubMed: 19238550]

Wilkinson, GS. Wide range achievement test 3-revision 3. Jastak Associates; Wilmington: 1993.

Woodcock, RW. Woodcock Reading Mastery Tests—Revised. American Guidance Service Inc; Circle Pines: 1987.



JA04 alleles (GT length)	17-20 GT	11-19GT	11-12 GT	19 GT	13-30 GT	12-23 GT
Freq in affected	0.02	0.07	0.33	0.02	0.34	0.22
Association to RD (with JA04 allele)	-	-	A-C-11GT-del-A-G protective	-	G-A-25GT-ins-G-G risk	-
p-value	-	-	0.0402	-	0.0013	-

Fig. 1.

Phylogenetic tree of polymorphisms in the *KIAA0319* region. Phylogenetic analysis determined the evolutionary relationship between the haplotypes, consisting of the markers rs6935076 (G/A), rs2038137 (C/A), rs71815143 (ins/del), rs3756821 (G/A) and rs9461045 (G/A). The ancestral haplotype (G-C-del-G-G) was selected based on chimpanzee genomic sequence. The risk haplotype (G-A-ins-G-G) is derived from a rare haplotype G-A-del-G-G (freq = 0.02). The range of JA04 allele lengths present on the haplotypes are also depicted. Only the risk haplotype carries the longer alleles (24–30GT). The risk allele of JA04 (25GT) and of rs71815143 (ins) are only present on the risk haplotype

Table 1

Summary of identified polymorphisms, predicted TFBS and TDT results (excluding JA04)

Polymorphism (UCSC Mar 2006)	Alleles (strand)	Frequency in RD probands	Frequency in NCBI dbSNP (build30-CEU)	Predicted TFBS	Association with RD		
					Transmission #	χ^2	*P*-value
rs1883592	G	0.41	None reported	–	62	3.041	$p = 0.081$
chr6: 24755208	A	0.59	None reported	–	83		
rs9467247	G	0.78	None reported	–	40	0.108	$p = 0.724$
chr6: 24755198	T	0.22	None reported	–	43		(Couto)
rs1883593	T	0.96	0.905	POU1F1 (GHF-1)	–	–	MAF <0.05
chr6: 24755170	A	0.04	0.095	–	–		
rs1883594	C	0.29	None reported	HOXA5*	73	1.075	$p = 0.300$
chr6: 24755163	G	0.71	None reported	–	61		
rs35882364 and	TT	0.36	None reported	Nkx2-5	–	–	–
rs35853270	T	0.64	None reported	–	–		
chr6: 24755133							
rs3756821	G	0.48	0.638	LBP-1	99	3.023	$p = 0.080$
chr6: 24754800	A	0.52	0.362	–	76		
rs6456625	G	0.33	None reported	Sp1*, ETF	82	2.778	$p = 0.090$
chr6: 24754483	A	0.67	None reported	–	62		
rs71815143	Insertion	0.33	None reported	3 Sp1*, ETF, C/EBP*	87	3.503	$p = 0.061$
chr6: 24754499–24754524	Deletion	0.67	None reported	–	64		
24754493 (new SNP)	Insertion	0.33	None reported	3 Sp1*, ETF, C/EBP*	T 87: U 64	3.503	$p = 0.061$
	Deletion	0.53	None reported	–	T 81: U 90	0.474	$p = 0.491$
	C						
	Deletion	0.14	None reported	–	T 44: U 58	1.922	$p = 0.166$
	T						
rs62400773	G	1.00	None reported	–	–	–	MAF <0.05
chr6: 24754176	A	0.0	None reported	FOXC1*, GATA2*	–		
rs2038135	A	0.38	0.362	PEBP2beta, AML1	–	–	–
chr6: 24753946	C	0.62	0.638	RAR*, AP-1*	–		
rs2038136	C	0.40	0.440 s-Ets-2*, IL-6	Rebp, Stat6*	–	–	–

Polymorphism (UCSC Mar 2006)	Alleles (strand)	Frequency in RD probands	Frequency in NCBI dbSNP (build30-CEU)	Predicted TFBS	Association with RD		
					Transmission #	χ^2	*P*-value
chr6: 24753926	G	0.60	0.560	–	–		
rs2038137	A	0.38	None reported	–	87	1.050	$p = 0.306$
chr6: 24753922	C	0.62	None reported	–	74		(Couto)
rs2235676	C	0.94	0.881	–	35	1.328	$p = 0.249$
chr6: 24753676	A	0.06	0.119	POU5F1*, F2F	26		(Couto)
rs13220596	G	0.98	1.000	–	–		MAF <0.05
chr6: 24753654	A	0.02	0.000	FOXL1*, FOXC1*, E4BP4, C/EBP*, bZIP, NFIL3, Ubx	–	–	–
rs2235677	T	0.38	0.250	F2F	84	2.703	$p = 0.100$
chr6: 24753589	C	0.62	0.750	–	64		(Couto)

Brain-relevant transcription factors are marked by asterisks

'Couto' indicates polymorphisms genotyped in a previous study by Couto et al. 2010.

Table 2

TDT Analysis of microsatellite marker JA04

Alleles	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
GT(n)	11	12	13	14	15	16	17	18	19	20	22	23	24	25	26	27	28	29	30
Allele freq	0.204	0.154	0.002	0.003	0.007	0.048	0.064	0.134	0.045	0.008	0.003	0.021	0.050	0.079	0.047	0.067	0.040	0.023	0.004
T	52	51	1	1	2	16	23	40	18	4	2	9	14	34	18	21	12	7	4
U	65	54	0	2	4	13	17	46	22	3	3	4	25	15	13	27	6	7	3
χ^2	1.444	0.086				0.310	0.900	0.419	0.400			1.923	3.103	7.367	0.806	0.750	2.00	0	
p-value	0.2295	0.7697				0.5775	0.3428	0.5177	0.5271			0.1656	0.0782	0.0067	0.3692	0.3865	0.1574	1.000	

Global χ^2 for 18 degrees of freedom = 27.58 ($p = 0.0687$)

T transmitted, *U* untransmitted

Table 3

TDT analysis of previously reported functional SNPs

Polymorphism (UCSC Mar 2006)	Alleles (strand)	Frequency in dbSNP	Predicted brain-relevant TFBS	Association with RD		
				Transmission #	χ^2	*p*-value
rs9461045	G	0.742	–	54	0.086	0.77
chr6: 24757040	A	0.258	Oct-1	51		
rs3212236	A	0.761	–	53	0.835	0.36
chr6: 24756434	G	0.239	–	44		

Table 4

KIAA0319 Francks haplotype with JA04 and 13 bp repeat

Haplotype—marker alleles					Haplotype frequency	Transmission				
rs4504469	**rs2038137**	**JA04**	**rs71815143**	**rs2143340**		**Obs**	**Exp.**	**Var (O–E)**	**χ^2**	***p*-value (1 d.f.)**
2 (A)	2 (A)	14 (25 GT)	1 (ins)	1 (T)	0.060	22.903	16.952	5.2399	6.7599	0.0093
2 (A)	2 (A)	16 (27 GT)	1 (ins)	1 (T)	0.053	13	14.5	8.801	0.2556	0.6132
1 (G)	1 (C)	1 (11 GT)	2 (del)	1 (T)	0.147	39.861	46.366	21.501	1.9676	0.1607
1 (G)	1 (C)	2 (12 GT)	2 (del)	1 (T)	0.122	30.038	33.06	12.026	0.7598	0.3834
1 (G)	1 (C)	8 (18 GT)	2 (del)	2 (C)	0.095	28	30	13.575	0.2946	0.5873

Global χ^2 (51 d.f.) = 112.45; $p = 0.000002$

χ^2 on haplotypes with freq >10% (2 d.f.) = 3.1161; $p = 0.2105$

χ^2 on haplotypes with freq >5% (5 d.f.) = 9.7402; $p = 0.0829$

Table 5

KIAA0319 Cope haplotype with JA04 and 13 bp repeat

Haplotype—marker alleles				Haplotype frequency	Transmission				
rs4504469	**rs6935076**	**JA04**	**rs71815143**		**Obs**	**Exp.**	**Var (O-E)**	**χ^2**	***p*-value (1 d.f.)**
1 (G)	1 (G)	7 (17 GT)	2 (del)	0.056	21.509	18.606	7.4419	1.1324	0.2873
1 (G)	1 (G)	8 (18 GT)	2 (del)	0.135	43.886	48.329	21.597	0.9141	0.3390
1 (G)	2 (A)	1 (11 GT)	2 (del)	0.156	47.668	56.211	24.346	2.9980	0.0834
1 (G)	2 (A)	2 (12 GT)	2 (del)	0.130	45.918	47.691	17.601	0.1784	0.6727
2 (A)	1 (G)	14 (25 GT)	1 (ins)	0.064	31.559	22.613	7.4282	10.772	0.0010
2 (A)	1 (G)	16 (27 GT)	1 (ins)	0.053	16.467	17.533	10.130	0.1123	0.7376

Global χ^2 (39 d.f.) = 80.518; $p = 0.0001$

χ^2 on haplotypes with freq >10% (3 d.f.) = 5.6353; $p = 0.1308$

χ^2 on haplotypes with freq >5% (6 d.f.) = 15.447; $p = 0.0171$

Table 6

Linkage disequilibrium between markers

	rs9461045	rs3212236	rs1883592	rs9467247	rs1883594	rs3756821	rs6456625	rs71815143	rs2038137	rs2235676	rs2235677	rs6935076
rs9461045		1	1	1	0.98	0.95	1	1	0.98	1	1	0.98
rs3212236	0.99		0.98	1	0.96	0.95	1	0.98	0.96	1	0.98	0.98
rs1883592	0.12	0.11		1	0.99	1	0.99	1	1	1	1	0.99
rs9467247	1	0.99	0.12		0.97	0.96	1	1	0.98	1	1	0.98
rs1883594	0.11	0.11	0.98	0.12		0.99	0.99	0.99	0.99	0.97	0.99	0.99
rs3756821	0.18	0.19	0.35	0.18	0.34		0.99	0.99	0.96	0.98	0.98	0.96
rs6456625	0.12	0.12	0.98	0.12	0.98	0.34		1	1	1	1	0.99
rs71815143	0.12	0.12	0.99	0.13	0.99	0.35	1		1	1	1	0.99
rs2038137	0.13	0.13	0.87	0.14	0.87	0.37	0.87	0.88		1	1	0.97
rs2235676	0.68	0.66	0.08	0.69	0.08	0.13	0.08	0.08	0.09		1	0.97
rs2235677	0.12	0.12	0.99	0.13	0.98	0.34	0.98	0.99	0.89	0.08		0.99
rs6935076	0.15	0.15	0.26	0.14	0.26	0.70	0.26	0.26	0.28	0.10	0.26	

D' values are shown in the top right and r^2 at the bottom left of the table

Table 7

Haplotype analysis of *KIAA0319* 5′ region

Haplotype—marker alleles						Haplotype frequency	Transmission				
rs6935076	**rs2038137**	**JA04**	**rs71815143**	**rs3756821**	**rs9461045**		**Obs**	**Exp.**	**Var (O–E)**	**χ^2**	***p*-value (1 d.f.)**
2 (A)	1 (C)	1 (11 GT)	2 (del)	2 (A)	1 (G)	0.188	53	63.5	26.197	4.2085	0.0402
2 (A)	1 (C)	2 (12 GT)	2 (del)	2 (A)	1 (G)	0.144	49	52	20.073	0.4484	0.5031
1 (G)	2 (A)	14 (25 GT)	1 (ins)	1 (G)	1 (G)	0.078	33	23.5	8.7136	10.3570	0.0013
1 (G)	2 (A)	16 (27 GT)	1 (ins)	1 (G)	1 (G)	0.060	19	19.5	9.8129	0.0255	0.8731
1 (G)	1 (C)	8 (18 GT)	2 (del)	1 (G)	2 (A)	0.134	45	46.991	20.097	0.1972	0.6570

Global χ^2 (38 d.f.) = 55.0060; $p = 0.0366$

χ^2 on haplotypes with freq >5% (5 d.f.) = 11.62; $p = 0.0404$