Article

Association of the *KIAA0319* Dyslexia Susceptibility Gene With Reading Skills in the General Population

Silvia Paracchini, D.Phil.

Colin D. Steer, M.Sc.

Lyn-Louise Buckingham, B.Sc.

Andrew P. Morris, Ph.D.

Susan Ring, Ph.D.

Thomas Scerri, D.Phil.

John Stein, F.R.C.P.

Marcus E. Pembrey, M.D.

Jiannis Ragoussis, Ph.D.

Jean Golding, Ph.D.

Anthony P. Monaco, Ph.D.

Objective: The authors previously identified a haplotype on chromosome 6p22 defined by three single-nucleotide polymorphisms (SNPs) that was associated with dyslexia (reading disability) in two independent samples of families that included at least one sibling with severe reading impairment. The authors also showed that this haplotype is associated with a reduction in expression of the *KIAA0319* gene. In addition, a completely independent study detected an association between *KIAA0319* markers and reading disability. In the current study, the authors tested whether the *KIAA0319* gene influences reading skills in the general population, rather than having an effect restricted to reading disability.

Method: The authors genotyped four SNPs that previously showed association with reading disability in the population of 7–9-year-old children in the Avon Longitudinal Study of Parents and Children (ALSPAC), a large longitudinal cohort for which reading-related phenotypes were available for more than 6,000 individuals. The authors conducted quantitative analysis for both single markers and haplotypes.

Results: The rs2143340 SNP, which effectively tags the three-SNP risk haplotype, was significantly associated with a test for reading ability. The risk haplotype itself also showed association with poor reading performance, and as in previous research, the association was stronger when the analysis was controlled for IQ.

Conclusions: These results both support a role of the *KIAA0319* gene in the development of dyslexia and suggest that this gene influences reading ability in the general population. Moreover, the data implicate the three-SNP haplotype and its tagging SNP rs2143340 as genetic risk factors for poor reading performance.

(Am J Psychiatry 2008; 165:1576–1584)

Dyslexia (reading disability) is a specific impairment in learning to read with a prevalence of 5%–10% in school-age children (1, 2). Reading disability is a complex trait determined in large part by genetic factors (3, 4). Association studies and translocation breakpoint analyses have proposed several genes as susceptibility candidates at some of the quantitative trait loci linked to dyslexia (5–7): *DYX1C1* on chromosome 15 (8), *KIAA0319* (9, 10) and *DCDC2* (11, 12) on chromosome 6, *ROBO1* on chromosome 3 (13), and *MRPL19* and *C2ORF3* on chromosome 2 (14).

Among these, the *KIAA0319* gene is a strong candidate, being supported by evidence in both association and functional studies. Significant associations have now been reported in at least three independent samples (Table 1). Francks et al. (9) identified a 77-kb region that showed association with different quantitative reading-related phenotypes in both a large sample of families from the United Kingdom and a sample of twin-based families from the Colorado Learning Disabilities Research Center. The region covered the first four *KIAA0319* exons, the entire *TTRAP* gene, and the first exon of *THEM2*. The single-nucleotide polymorphisms (SNPs) rs4504469, rs2038137, and rs2143340 defined the major haplotypes present in the region, and the 1-1-2 haplotype (where 1 is the major allele and 2 is the minor allele), which is effectively tagged by the rs2143340 SNP, was significantly associated with reading disability in both samples. Cope et al. (10) detected associations with markers rs4504469 and rs6935076, both located within *KIAA0319*, in a completely independent sample of U.K. families by using a categorical definition of dyslexia. In particular, the 2-1 haplotype defined by these two markers was significantly associated with a protective effect, having a significantly higher frequency in good/normal readers. An association at the same locus was found in two other studies that, however, cannot be considered independent because they both derived their samples from the Colorado Learning Disabilities Research Center cohort (15, 16). Kaplan et al. (15) reported association in a set of 104 families between a reading-related measure and the JA04 microsatellite, which is located in the first *KIAA0319* exon. Deffenbacher et al. (16) performed association analysis in a set of 114 families selected for severity and detected evidence for association both at the *KIAA0319* locus and within the *DCDC2* gene, which lies about 200 kb distal to *KIAA0319*. A third study, which employed 153 families from the Colo-

This article is featured in this month's AJP **Audio** and is discussed in an editorial by Dr. Freedman (p. 1505).

TABLE 1. Most Significant Findings Previously Reported for Associations Between Reading Skills and Genetic Markers Used in the Present Study

Marker or Haplotype	Associated Allele or Haplotype (1=major allele, 2=minor allele)[a]	Study	Sample[b]	Statistical Association (p)	Phenotype[c]
Markers					
rs4504469	1	Francks et al. (9)	Oxford sample	0.002	Irregular word reading
				0.009	Phonological decoding
				0.0004	Single word reading
				0.01	Phonological awareness
		Cope et al. (10)	Cardiff individual sample	0.002	Developmental dyslexia
			Cardiff trio sample	0.04	Developmental dyslexia
rs6935076	2	Cope et al. (10)	Cardiff individual sample	0.006	Developmental dyslexia
			Cardiff trio sample	0.002	Developmental dyslexia
rs2038137	1	Francks et al. (9)	Oxford sample	0.002	Irregular word reading
				0.003	Phonological decoding
				0.0002	Single word reading
				0.007	Single word spelling
rs2143340	2	Francks et al. (9)	Oxford sample	0.01	Irregular word reading
				0.0003	Orthographic coding choice
				0.02	Single word reading
			U.S. sample	0.005	Single word reading
				0.03	Single word spelling
				0.02	Phonological awareness
Haplotypes					
rs4504469-rs2038137-rs2143340	1-1-2	Francks et al. (9)	Oxford sample	0.005	Irregular word reading
				0.00007	Orthographic coding choice
				0.003	Single word reading
				0.03	Single word spelling
			U.S. sample	0.02	Single word reading
				0.04	Single word spelling
				0.02	Phonological awareness
rs4504469-rs6935076	2-1	Cope et al. (10)	Cardiff individual sample	0.00003	Developmental dyslexia
			Cardiff trio sample	0.006	Developmental dyslexia

[a] All the reported alleles were associated with reading problems except for the 2-1 haplotype, which showed a protective effect.
[b] Oxford sample: 126 U.K. families selected for severity. Cardiff individual sample: 223 U.K. individuals with reading disability and 273 comparison subjects. Cardiff trio sample: 143 U.K. trios of probands with reading disability and their parents. U.S. sample: 124 U.S. families selected for severity.
[c] The specific reading skills are described by Francks et al. (9). In the study by Cope et al. (10), developmental dyslexia was defined as a reading age 2.5 or more years behind that expected from chronological age.

rado Learning Disabilities Research Center, showed an association with *DCDC2* but not with *KIAA0319* (11). The different outcomes of these studies are most likely due to the different criteria employed for family selection and the small samples (5). Analysis conducted in a sample of German trios selected for spelling impairment showed association with the *DCDC2* gene (12).

No functional mutations have been identified yet for either of the two genes through resequencing of the exons and promoter regions (9, 10, 12). Harold et al. (17) conducted a comprehensive study to compare the *DCDC2* and *KIAA0319* genes by genotyping the two U.K. samples previously used by Francks et al. (Oxford sample) (9) and Cope et al. (Cardiff sample) (10) for a common set of *DCDC2* and *KIAA0319* markers. None of the *DCDC2* markers showed any association, while several markers showed significant associations and identical allelic trends in both the Oxford and Cardiff samples. The most significant association was clustered around the first exon and putative regulatory sequences of *KIAA0319* in both samples, suggesting that the functional genetic variant might act at the level of gene expression. In agreement with this finding, we have shown through in vivo studies that the 1-1-2 haplotype is associated with a relative reduction of *KIAA0319* gene expression (18).

Reading disability could be considered to represent the lower tail of a normal distribution of reading ability observed in the population (19). It is possible that the quantitative trait loci underlying dyslexia susceptibility also influence variation in reading performance in the normal range. In the study described here, we aimed to test this hypothesis, asking whether the *KIAA0319* gene influences reading skills in the general population, rather than being specifically implicated in determining reading problems in restricted groups of individuals with dyslexia.

Method

Sample Description

We used the Avon Longitudinal Study of Parents and Children (ALSPAC), a general population cohort of about 14,000 children with expected dates of delivery between

TABLE 2. Reading-Related Measures Selected From the Avon Longitudinal Study of Parents and Children (ALSPAC) Study

Phenotype[a]	Total Sample Size	Children With Problems		Score	
		N	%	Mean	SD
Dichotomous measures					
Parent-reported reading difficulties (age 7)	6,085	373	6.13		
Additional adjustment for IQ	4,860	271	5.58		
Subjects with IQ>90	3,825	131	3.42		
Not at target level on Statutory Assessment Test for reading (age 7)	7,207	961	13.33		
Additional adjustment for IQ	4,790	393	8.20		
Subjects with IQ>90	3,597	129	3.59		
Continuous measures					
Reading (age 7)	6,338			27.95	9.29
Additional adjustment for IQ	5,288			28.42	9.04
Subjects with IQ>90	4,104			30.34	8.30
Spelling (age 7)	6,244			25.59	12.57
Additional adjustment for IQ	5,243			25.91	12.47
Subjects with IQ>90	4,086			27.90	11.89
Phoneme awareness (age 7)	6,327			19.95	9.50
Additional adjustment for IQ	5,283			20.28	9.36
Subjects with IQ>90	4,103			21.93	8.95
Accuracy (age 9)	5,339			104.04	13.50
Additional adjustment for IQ	4,685			104.58	13.34
Subjects with IQ>90	3,642			107.31	12.38
Nonword reading (age 9)	5,816			5.23	2.48
Additional adjustment for IQ	5,138			5.28	2.47
Subjects with IQ>90	4,016			5.64	2.36

[a] See text for descriptions of reading measures. All analyses included adjustment for age and gender.

April 1, 1991, and Dec. 31, 1992, in the southwest of England (20). From age 7, all children were invited annually for assessments on a wide range of physical, behavioral, and neuropsychological traits, including reading-related measures. DNA was extracted from blood and mouthwash samples and processed as described previously (21). Informed written consent was obtained from the parents after complete description of the study at the time of enrolment into the ALSPAC project, with the option for them or their children to withdraw at any time. Ethical approval for the present study was obtained from the ALSPAC Law and Ethics Committee and the local research ethics committees.

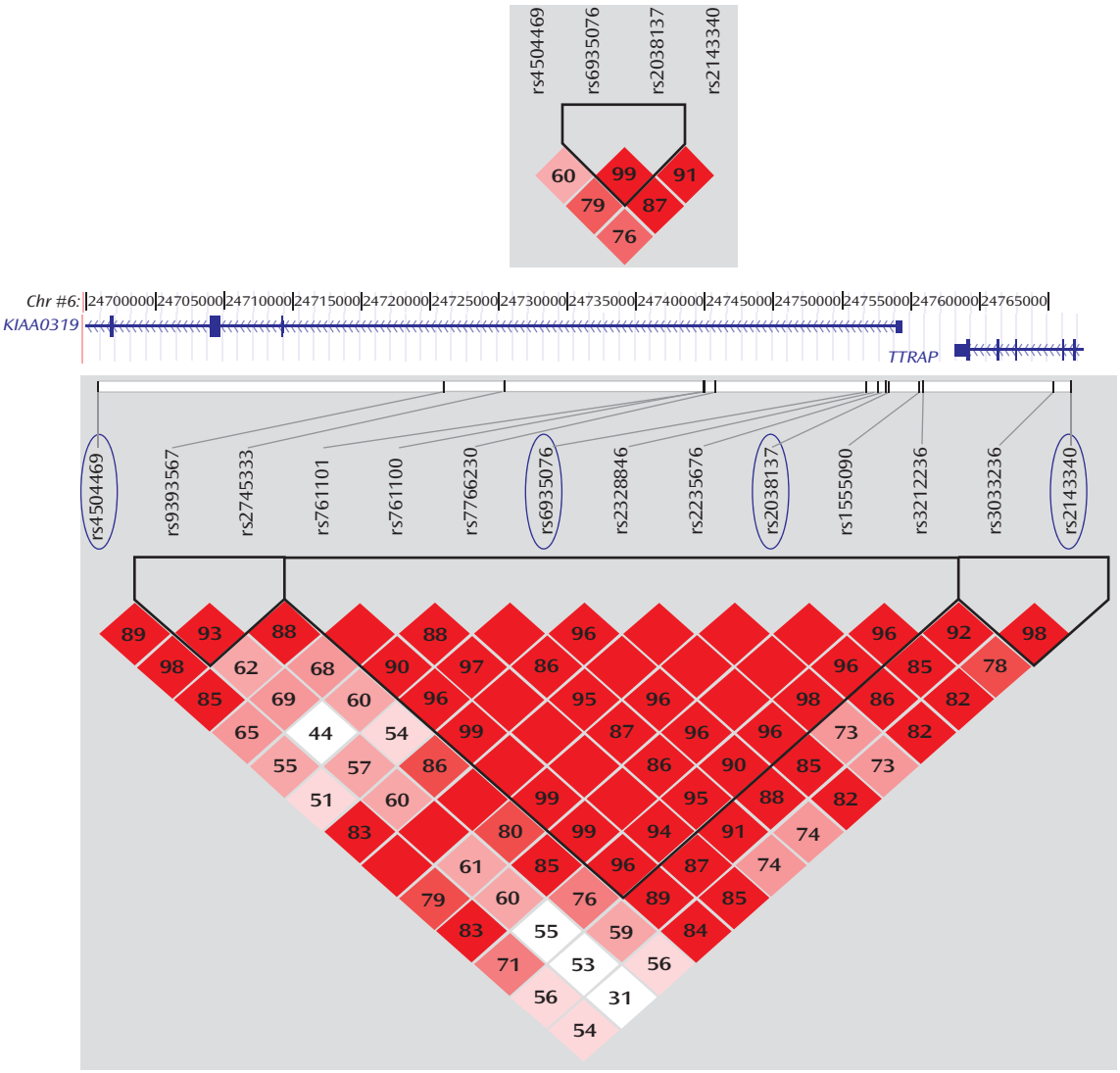
Genotyping

We genotyped all the available DNA samples from the children (N=10,621) for the four SNPs (rs4504469, rs6935076, rs2038137, and rs2143340) that showed significant associations in previous studies (9, 10). Genotyping was performed in a 4-plex reaction using the homogeneous mass-extend assay (Sequenom, San Diego) according to the manufacturer's instructions. Polymerase chain reaction was performed on 10 ng of DNA by using the following primers: rs4504469_F ACGTTGGATGAGAGCACAGCATCCCAACAC, rs4504469_R ACGTTGGATGGTGGTAGGAGATATGGGTAG, rs6935076_F ACGTTGGATGAACCGAAGCCCAGAGAAAAC, rs6935076_R ACGTTGGATGAAAAAAATTCCTGGCCAGGG , rs2038137_F ACGTTGGATGGCCCTCTTTCCTATTTCTCG, rs2038137_R ACGTTGGATGAGCTTCAGTGTCGCCAGCAG, rs2143340_F ACGTTGGATGATTTGTAGCCCTCATTTTAC, and rs2143340_R ACGTTGGATGTCTGTCTAGAACCTGGCATG. The primers used for primer extension reactions were as follows: rs4504469_E AACACCTCCCACTAGC, rs6935076_E CAGACATGAGGAGAATGA, rs2038137_E TTTCTCGGCCAGGCGC, and rs2143340_E ACAAATTTTAAAAGAGCCCTA. The success call rate was approximately 95% (95.33% for rs4504469, 93.51% for rs6935076, 96.01% for rs2038137, and 95.75% for rs2143340). To control for genotyping quality, about 3% of the samples were present as duplicates and distributed across each of the 32×384-well plates used. The duplicates gave an estimate of the genotyping error rate of about 1.5%. Each plate included four samples randomly derived from nine non-ALSPAC control DNA samples. The genotypes of these control samples were matched in all the plates except for two genotype calls, giving a genotype error estimate of 0.4%. No genotype calls were returned for any of the blank wells present on each plate. All the markers were in Hardy-Weinberg equilibrium, and their allele frequencies were consistent with those in previous reports.

Phenotype Measures

We selected reading-related measures that would match the assessments we used in our previous analysis (9), including two dichotomous measures—reading difficulties and reading Statutory Assessment Tests—and five continuous measures—reading, spelling, phoneme awareness, accuracy, and nonword reading. These measures were available for a range of about 5,300 to 7,200 children (Table 2). The measure of reading difficulties is based on the answers to self-completion questionnaires that the ALSPAC team asked the parents to fill in when the children were 7 years old. The relevant section of the questionnaire asked whether the school had recognized reading difficulties in their child and, if so, had made special arrangements.

FIGURE 1. Linkage Disequilibrium and Genomic Location of Genetic Markers Assessed for Association With Reading Measures in the Avon Longitudinal Study of Parents and Children (ALSPAC) Cohort[a]



[a] The top graph shows D′, a measure of linkage disequilibrium, for the four single-nucleotide polymorphisms (SNPs) employed in the present study. The bottom graph shows the four SNPs (highlighted by blue ovals) in the context of their genomic locations (adapted from the University of California, Santa Cruz, Genome Browser; http://genome.ucsc.edu/) and linkage disequilibrium structure at the *KIAA0319* locus in combination with other markers across the dyslexia-associated region. D′ was calculated on genotype data collected previously (17) in our set of U.K. families (Oxford sample). Linkage disequilibrium was calculated by using Haploview 3.32 (http://www.broad.mit.edu/mpg/haploview/). Linkage disequilibrium blocks, indicated by solid black lines, are identified according to the definition of Gabriel et al. (28).

About 6% of the children required special arrangements. Results of the school-administered reading Statutory Assessment Tests were obtained from the relevant local educational authorities at about age 7 years. The test included reading and comprehension tasks. About 13% of the children did not achieve the target levels in those reading tests. The continuous reading measure was obtained at age 7 years by using the Wechsler Objective Reading Dimensions test (22). Spelling was assessed at age 7 years and was based on the children's performance in spelling 15 words of increasing difficulty chosen specifically for this age group (P. Bryant and T. Nunes, personal communication). Phoneme awareness was assessed at age 7 years and was based on the children's performance on the phoneme deletion task, in which children are asked to repeat a word after removing one or more phonemes (23). Accuracy and nonword reading were assessed at age 9 years by means of additional tests aimed to measure reading abilities (24, 25). Measures of total and verbal IQ were based on the WISC-III administered at age 8 (26).

Statistical Analysis

We included in the analysis only individuals with self-reported white European ancestry. We performed single-marker and haplotype-based analyses of the raw continuous phenotypic measures in a linear regression framework,

TABLE 3. Association of Genetic Marker rs2143340 With Reading-Related Measures in the Avon Longitudinal Study of Parents and Children (ALSPAC) Cohort

		Adjusted Mean for Each Genotype (1=major allele, 2=minor allele)[b]				Allelic Trend	
Phenotype[a]	N	11	12	22	p[c]	Odds Ratio or β[d]	p[c]
Dichotomous measures							
Parent-reported reading difficulties (age 7)	5,921	5.65	5.53	4.62		0.95	
Additional adjustment for IQ	4,728	3.81	3.70	2.64		0.93	
Subjects with IQ>90	3,716	2.88	3.11	—[e]		0.92	
Not at target level on Statutory Assessment Test for reading (age 7)	7,090	12.32	11.81	14.51		1.00	
Additional adjustment for IQ	4,703	3.86	3.12	5.95		0.94	
Subjects with IQ>90	3,528	2.79	2.23	4.99		0.95	
Continuous measures							
Reading (age 7)	6,221	0.023	–0.042	–0.169	0.009	–0.074	0.003
Additional adjustment for IQ	5,196	0.070	0.016	–0.113	0.009	–0.065	0.003
Subjects with IQ>90	4,030	0.285	0.201	0.050	0.003	–0.094	0.001
Spelling (age 7)	6,130	0.030	–0.011	–0.139		–0.055	0.03
Additional adjustment for IQ	5,154	0.058	0.006	–0.110	0.04	–0.062	0.02
Subjects with IQ>90	4,014	0.220	0.153	0.020	0.03	–0.077	0.008
Phoneme awareness (age 7)	6,211	0.019	–0.016	–0.070		–0.038	
Additional adjustment for IQ	5,192	0.049	0.024	–0.029		–0.029	
Subjects with IQ>90	4,030	0.230	0.182	0.134		–0.048	
Accuracy (age 9)	5,254	0.021	0.010	–0.113		–0.028	
Additional adjustment for IQ	4,610	0.057	0.057	–0.036		–0.014	
Subjects with IQ>90	3,584	0.260	0.247	0.086		–0.035	
Nonword reading (age 9)	5,718	0.017	–0.020	–0.118		–0.046	
Additional adjustment for IQ	5,049	0.035	–0.003	–0.066		–0.041	
Subjects with IQ>90	3,944	0.187	0.128	0.025		–0.065	0.03

[a] See text for descriptions of reading measures. All analyses included adjustment for age and gender.
[b] Adjusted means reflect prevalence of adverse outcome (dichotomous measures) or mean standardized score (continuous measures). Online data supplement Table 1 displays 95% confidence intervals.
[c] Only nominally significant p values ($p \leq 0.05$) are shown.
[d] Odds ratios are used for dichotomous measures, and β values are used for quantitative measures. These reflect the change in odds or standardized score for each additional copy of the minor allele. Odds ratios greater than 1 and β values less than 0 indicate a detrimental effect.
[e] Not tested.

adjusting for age and gender. Analysis of the dichotomous measures was performed by using logistic regression.

For the single-marker analyses, genotypes at each SNP were coded as 0, 1, or 2, according to the number of copies of the minor allele. Two tests were performed at each SNP: 1) a genotype-based test, separately comparing the risks of the heterozygote and the rare homozygote to the common homozygote genotype, and 2) an allelic trend test, comparing the same risks under a multiplicative disease model. The allelic trend test is generally more powerful than the genotype test, unless there are strong dominance effects or a heterozygote advantage; these have not been previously reported for reading disability with SNPs in *KIAA0319*.

Haplotypes were estimated by maximum likelihood through implementation of the E-M algorithm, and the most likely phase configuration was recorded for each individual. To reduce bias due to phase uncertainty, we excluded individuals with incomplete data on the SNPs defining the haplotypes. Since the purpose of this study was to replicate previous results, we tested only the haplotypes that had shown significant associations, rather than testing all the individual haplotypes defined by the same marker combinations, in order to reduce the number of tests. Therefore, we tested the 1-1-2 haplotype defined by markers rs4504469, rs2038137, and rs2143340 that was previously identified as a risk factor (9) and haplotype 2-1 defined by markers rs4504469 and rs6935076, which was previously observed to be protective (10). For the same reason, we did not test all the haplotypes generated by the four markers. We also generated a "global" p value by testing simultaneously all haplotypes derived from either the three-marker or two-marker combinations.

In our previous analysis we found that the association was stronger in subsets of samples selected for severity and when the phenotypic traits were adjusted with measures of general intelligence (IQ) (9), a finding that agrees with results of our multivariate analysis, which showed that the quantitative trait locus on chromosome 6p influences reading ability specifically but not IQ (27). Accordingly, we also performed the analysis in subgroups of the sample at the extremes (lower and higher 25th, 10th, and 5th percentiles) of the distribution in a logistic regression framework, controlling for IQ through two different approaches. For both the entire sample and the subgroups, we performed the analysis both by adjusting the reading measures for IQ and by selecting only individuals who had a total raw IQ higher than 90. This latter approach aimed to reproduce the ascertainment criteria used in our previous study to recruit individuals with reading disability and a high IQ (9). The reported association p values were not corrected for multiple testing because the tests were not

independent, owing to the high correlation across markers (Figure 1). Moreover, this is a replication study, and rather than conduct an extensive analysis of this locus, we tested only specific SNPs and haplotypes in an attempt to reduce the number of comparisons.

The statistical analysis was performed with the Stata program, version 8.2 (Stata Corp., College Station, Tex.). Intermarker linkage disequilibrium was evaluated by using Haploview 3.32 (http://www.broad.mit.edu/mpg/haploview/).

Results

In the single-marker analysis, the SNP rs2143340 showed the most significant association. In particular, the minor allele was significantly associated with poor performance on the continuous measures of reading and spelling (Table 3). The association with the reading test was nominally significant in the entire sample and became stronger after selection of individuals with high IQ. The association with spelling was of a lower magnitude but, similarly, increased in significance after control for IQ. The allelic trend model produced consistently stronger associations. Analysis with the other continuous measures showed a consistent trend of association between the minor allele of rs2143340 and poor performance ($\beta<0$, Table 3). The other three markers showed only a few marginally significant associations (data supplement Table 1). There was no particular trend of association for marker rs2038137, while rs4504469 showed mainly a trend of association between the major allele and poor performance in the reading-related measures, as previously reported (9, 10). The trend for rs6935076 was an association between the major allele and poor performance on the continuous measures, in disagreement with what was previously observed in the Cardiff sample, where a significant association with dyslexia was reported for the minor allele (Table 1) (10, 17). In our previous analysis of rs6935076 conducted in the Oxford sample, we did not detect any significant or consistent trend of association (17).

In the haplotype analysis, the 1-1-2 haplotype showed a nominally significant association with poor performance on the continuous measures of reading and spelling (Table 4). A consistent trend of association between the 1-1-2 haplotype and poor performance was observed for all the continuous measures ($\beta<0$, Table 4). The 2-1 haplotype did not show any significant associations. The test for "global" associations showed nominally significant associations for both the rs4504469-rs2038137-rs2143340 and the rs4504469-rs6935076 haplotypes (Table 4). The three-marker haplotypes were associated with a wider range of measures, including reading, spelling, phoneme awareness, and nonword reading, while the two-marker haplotypes showed associations solely with the reading and phoneme tests.

Analysis of subgroups of individuals selected at the extremes of the distribution showed strengthening of the association of both rs2143340 and the 1-1-2 haplotype with the continuously scored reading phenotype in the 25th percentile selection of the entire sample (data supplement Table 2 and data supplement Table 3). Specifically, the analysis of the rs2143340 marker in this subgroup showed the strongest association observed in this study (allelic trend: $p=0.0002$). The same trend was not observed in more stringent percentile selections for the continuous reading measure. Other nominally significant associations were observed for other phenotypic subgroups, but overall, we did not observe a regular trend consistent with gradually increasing stringency in other phenotypic subgroups. An explanation for this lack of consistency could be the loss of power to detect the *KIAA0319* effect in small samples. Similarly, we did not detect significant associations with the dichotomous measures when we tested small proportions of individuals with reading problems in a logistic regression framework. Most likely, when specific ascertainment criteria are not applied, quantitative analysis in a large sample reflecting all the phenotypic variation is more powerful than looking at smaller numbers of individuals at the extremes of the distribution.

Discussion

The results reported here are in agreement with our previous findings, which showed associations of the 1-1-2 haplotype and the minor allele of rs2143340 with poor reading performance (9). Moreover, all the associations we detected in the present study involving both the rs2143340 marker and the risk haplotype were strengthened after we controlled for IQ, and this was particularly true for the analysis of individuals with IQs above 90, supporting our hypothesis that IQ can be considered a confounder for the quantitative trait locus on chromosome 6p (9). Conversely, we did not observe any associations between the 2-1 haplotype and good reading performance, as previously suggested (Table 1) (10). Nevertheless, the global test for the rs4504469-rs6935076 haplotypes showed some significant association, as would be expected owing to the correlation and physical proximity of the four SNPs (Figure 1).

The allelic combination of markers rs4504469, rs2038137, and rs2143340 captured most of the genetic variation in the dyslexia-associated region by describing the most common haplotypes in our U.K. sample (9). We hypothesize that in the white European population, the 1-1-2 haplotype carries a functional mutation relevant to dyslexia. Association data in our previous (9) and present study show that the haplotypic analysis does not detect any association stronger than what is observed in single-marker analysis. We conclude that the analysis of rs2143340 alone is sufficient and more powerful to detect the risk haplotype effect. This can be explained by the fact that rs2143340 efficiently tags the risk haplotype in the white European population, and missing genotypes in any of the three markers can contribute to loss of power in reconstructing haplotypes.

TABLE 4. Association of Two Specific Haplotypes With Reading-Related Measures in the Avon Longitudinal Study of Parents and Children (ALSPAC) Cohort

	rs4504469-rs6935076 2-1 Haplotype				
Phenotype[a]	N	Odds Ratio or β[b]	95% CI	p[c]	Global p[c,d]
Dichotomous measures					
Parent-reported reading difficulties (age 7)	5,709	0.88	0.75 to 1.04		
Additional adjustment for IQ	4,546	0.89	0.73 to 1.08		
Subjects with IQ>90	3,569	0.86	0.65 to 1.13		
Not at target level on Statutory Assessment Test for reading (age 7)	6,844	0.95	0.85 to 1.05		
Additional adjustment for IQ	4,518	0.93	0.78 to 1.11		
Subjects with IQ>90	3,384	0.95	0.73 to 1.25		
Continuous measures					
Reading (age 7)	5,982	–0.003	–0.041 to 0.034		0.008
Additional adjustment for IQ	4,990	–0.009	–0.042 to 0.024		0.04
Subjects with IQ>90	3,868	–0.017	–0.058 to 0.024		0.04
Spelling (age 7)	5,896	–0.007	–0.044 to 0.031		
Additional adjustment for IQ	4,951	–0.002	–0.038 to 0.035		
Subjects with IQ>90	3,853	–0.013	–0.057 to 0.030		
Phoneme awareness (age 7)	5,972	0.007	–0.031 to 0.045		0.03
Additional adjustment for IQ	4,986	0.014	–0.022 to 0.050		
Subjects with IQ>90	3,868	0.02	–0.024 to 0.064		
Accuracy (age 9)	5,065	–0.007	–0.047 to 0.033		
Additional adjustment for IQ	4,440	–0.012	–0.048 to 0.024		
Subjects with IQ>90	3,452	–0.009	–0.054 to 0.036		
Nonword reading (age 9)	5,511	–0.009	–0.048 to 0.030		
Additional adjustment for IQ	4,862	–0.012	–0.050 to 0.026		
Subjects with IQ>90	3,798	–0.012	–0.057 to 0.033		

[a] See text for descriptions of reading measures. All analyses included adjustment for age and gender.
[b] Odds ratios are used for dichotomous measures, and β values are used for quantitative measures. These reflect the change in odds or standardized score for each additional copy of the haplotype. Odds ratios greater than 1 and β values less than 0 indicate a detrimental effect.
[c] Only nominally significant p values ($p \leq 0.05$) are shown.
[d] From simultaneous test of all haplotypes.

We have shown before that the 1-1-2 haplotype is associated with a reduction of *KIAA0319* gene expression in cell line models and that *KIAA0319* is required for neuronal migration during the development of the neocortex (18). We have hypothesized that a suboptimal level of the *KIAA0319* protein could be the cause of subtle cortical abnormalities that might lead to the development of dyslexia. A major question is, How is it possible that alteration in a gene involved in such a wide-ranging mechanism as neuronal migration can affect specific cognitive functions and not lead to a global impairment? Probably, the *KIAA0319* risk haplotype alone is not enough to have a significant impact on cognition, but when in the presence of particular combinations of multiple factors, it might specifically affect precise functions, such as reading abilities. Further studies are now required to identify such factors, of both genetic and environmental nature, and to test whether the *KIAA0319* gene might influence other phenotypes.

The present study shows that the effect of *KIAA0319* alleles conferring susceptibility to reading disability is not restricted to specific groups of dyslexic individuals. Sample selection and ascertainment criteria play a critical role in the outcome of association studies, which are normally based on samples of a few hundred probands. Individuals are usually recruited if their reading performance falls below a predetermined threshold defining dyslexia. There are not universally accepted criteria to fix such a threshold, and selection of different subgroups of individuals with reading disability might be at the basis of disparities in the outcomes of association studies (5). Negative association reports in previous studies (11, 12) might be due to small samples that did not have enough power to detect the *KIAA0319* effect or in which the *KIAA0319* gene was not the major causative genetic factor. The same observation regarding sample size can be made for a recently published study that evaluated the influence of the *KIAA0319* gene on the reading skills of an unselected sample of 440 families in Australia (29). Associations were detected with both the rs2143340 marker and the 1-1-2 haplotype, but they cannot be regarded as identical allelic replications since they showed opposite trends; both the rs2143340 minor allele and the 1-1-2 haplotype were associated with better reading performance. The associations were only marginally significant and became further attenuated when the analysis was repeated only in individuals with at least 75% Anglo-Celtic ancestry (about 82% of the sample), suggesting that population admixture could also be a factor at the basis of this incongruous association. In support of this explanation, we observed that when we conducted the analysis on our entire sample, including individuals of nonwhite European ancestry, the associations were notably decreased (data not shown). For example, in the allelic trend tests, the strength of the association between rs2143340 and the continuous reading measure decreased from p=0.003 in the sample based only on individuals of white European ancestry to p=0.02 in the entire

rs4504469-rs2038137-rs2143340 1-1-2 Haplotype				
N	Odds Ratio or β[b]	95% CI	p[c]	Global p[c,d]
5,765	0.99	0.80 to 1.24		
4,592	0.98	0.75 to 1.28		
3,609	0.93	0.63 to 1.36		
6,942	1.02	0.88 to 1.17		
4,589	0.98	0.77 to 1.25		
3,426	0.96	0.65 to 1.42		
6,064	–0.067	–0.118 to –0.016	0.01	0.009
5,061	–0.057	–0.103 to –0.012	0.02	0.005
3,928	–0.081	–0.138 to –0.024	0.005	0.03
5,978	–0.046	–0.097 to 0.006		0.03
5,022	–0.051	–0.101 to –0.001	0.05	0.02
3,914	–0.068	–0.128 to –0.009	0.03	0.05
6,054	–0.043	–0.094 to 0.008		0.03
5,057	–0.03	–0.080 to 0.019		
3,928	–0.047	–0.107 to 0.013		
5,132	–0.021	–0.076 to 0.034		
4,500	–0.009	–0.060 to 0.041		
3,501	–0.026	–0.088 to 0.036		
5,582	–0.034	–0.088 to 0.019		0.05
4,924	–0.029	–0.082 to 0.023		0.05
3,848	–0.052	–0.113 to 0.010		

sample (gene-ethnicity interaction: p=0.04). This shows that a small number of individuals with different ancestry (about 4.6% of the sample analyzed) can significantly modify association outcomes, in agreement with previous reports (30, 31). The rs2143340 itself lies within the neighboring *TTRAP* gene and probably does not have any functional role. More likely, it is in strong linkage disequilibrium in the white European population with a causal variant or variants expected to influence *KIAA0319* gene expression. As shown in Figure 1, rs2143340 is in linkage disequilibrium with markers across *KIAA0319*, including SNPs located at the promoter site, where gene expression regulatory regions are expected to reside. Conversely, the linkage disequilibrium between rs2143340 and *KIAA0319* promoter markers is lost in populations of Asian and African origin (www.hapmap.org; data supplement Figure 1). If a functional mutation is indeed located at the *KIAA0319* promoter, it will then co-occur by chance more frequently with the major allele of rs2143340 in these other, non-European populations. This observation can explain the opposite trend of association detected in admixed samples when compared to samples of U.K. origin.

In conclusion, this study shows that the effect of a quantitative trait locus can be detected in an unselected sample, providing that the sample is large enough and reproduces the entire spectrum of a phenotypic distribution. These data not only provide further support for a role of the *KIAA0319* gene in the development of dyslexia, but they also show that this gene influences reading ability in a wider context. Specifically, both the 1-1-2 haplotype and, in particular, the rs2143340 tagging SNP have been confirmed to be genetic risk factors for reading problems at the chromosome 6p quantitative trait locus in the U.K. population. These data suggest the need for functional studies both to identify the causal genetic variant responsible for the risk haplotype effect and to clarify the function of the *KIAA0319* gene, in order to make it possible to dissect the biological pathway at the basis of the reading process and, more broadly, cognitive functions.

Received Dec. 10, 2007; revisions received April 1 and June 17, 2008; accepted June 19, 2008 (doi: 10.1176/appi.ajp.2008.07121872). From the Wellcome Trust Centre for Human Genetics and the Department of Physiology, University of Oxford; the Department of Community Based Medicine and Department of Social Medicine, University of Bristol, Bristol, U.K.; and the Institute of Child Health, University College London. Address correspondence and reprint requests to Dr. Paracchini, Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, OX3 7BN Oxford, U.K.; silviap@well.ox.ac.uk (e-mail).

All authors report no competing interests.

Supported by grants to Prof. Monaco from the Wellcome Trust and by core support for ALSPAC from the U.K. Medical Research Council, Wellcome Trust, and University of Bristol.

The authors thank the families who took part in this study, the midwives who helped with recruitment, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists, and nurses; Laura L. Miller for assistance with statistical analysis; John Broxholme for information technology support; and Dianne Newbury for suggestions regarding the manuscript.

References

1. Pennington BF: The genetics of dyslexia. J Child Psychol Psychiatry 1990; 31:193–201
2. Shaywitz SE, Shaywitz BA, Fletcher JM, Escobar MD: Prevalence of reading disability in boys and girls: results of the Connecticut Longitudinal Study. JAMA 1990; 264:998–1002
3. DeFries JC, Alarcón M: Genetics of specific reading disability. Ment Retard Dev Disabil Res Rev 1996; 2:39–47
4. Fisher SE, DeFries JC: Developmental dyslexia: genetic dissection of a complex cognitive trait. Nat Rev Neurosci 2002; 3: 767–780
5. Paracchini S, Scerri T, Monaco AP: The genetic lexicon of dyslexia. Ann Rev Genomics Hum Genet 2007; 8:57–79
6. Williams J, O'Donovan MC: The genetics of developmental dyslexia. Eur J Hum Genet 2006; 14:681–689
7. Fisher SE, Francks C: Genes, cognition and dyslexia: learning to read the genome. Trends Cogn Sci 2006; 10:250–257
8. Taipale M, Kaminen N, Nopola-Hemmi J, Haltia T, Myllyluoma B, Lyytinen H, Muller K, Kaaranen M, Lindsberg PJ, Hannula-Jouppi K, Kere J: A candidate gene for developmental dyslexia encodes a nuclear tetratricopeptide repeat domain protein dynamically regulated in brain. Proc Natl Acad Sci USA 2003; 100: 11553–11558
9. 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; 75:1046–1058
10. 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; 76:581–591
11. 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; 102:17053–17058; correction, 2005; 102:18763
12. 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; 78:52–62
13. Hannula-Jouppi K, Kaminen-Ahola N, Taipale M, Eklund R, Nopola-Hemmi J, Kaariainen H, Kere J: The axon guidance receptor gene ROBO1 is a candidate gene for developmental dyslexia. PLoS Genet 2005; 1(4):e50
14. Anthoni H, Zucchelli M, Matsson H, Muller-Myhsok B, Fransson I, Schumacher J, Massinen S, Onkamo P, Warnke A, Griesemann H, Hoffmann P, Nopola-Hemmi J, Lyytinen H, Schulte-Korne G, Kere J, Nothen MM, Peyrard-Janvid M: A locus on 2p12 containing the co-regulated MRPL19 and C2ORF3 genes is associated to dyslexia. Hum Mol Genet 2007; 16:667–677
15. 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; 70:1287–1298
16. 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; 115:128–138
17. 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; 11:1085–1091
18. 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; 15:1659–1666
19. Shaywitz SE, Escobar MD, Shaywitz BA, Fletcher JM, Makuch R: Evidence that dyslexia may represent the lower tail of a normal distribution of reading ability. N Engl J Med 1992; 326:145–150
20. Golding J, Pembrey M, Jones R: ALSPAC—the Avon Longitudinal Study of Parents and Children, I: study methodology. Paediatr Perinat Epidemiol 2001; 15:74–87
21. Jones RW, Ring S, Tyfield L, Hamvas R, Simmons H, Pembrey M, Golding J: A new human genetic resource: a DNA bank established as part of the Avon Longitudinal Study of Pregnancy and Childhood (ALSPAC). Eur J Hum Genet 2000; 8:653–660
22. Rust J, Golombok S, Trickey G: WORD: Wechsler Objective Reading Dimensions Manual. Sidcup, UK, Psychological Corp, 1993
23. Rosner J, Simon DP: The Auditory Analysis Test: an initial report. J Learning Disabilities 1971; 4:40–48
24. Neale M: Neale Analysis of Reading Ability—Revised: Manual for Schools. Windsor, UK, NFER-Nelson, 1997
25. Nunes T, Bryant P, Olsson J: Learning morphological and phonological spelling rules: an intervention study. Scientific Studies of Reading 2003; 7:298–307
26. Wechsler D, Golombok S, Rust J: WISC-IIIUK: Wechsler Intelligence Scale for Children. Sidcup, UK, Psychological Corp, 1992
27. Marlow AJ, Fisher SE, Francks C, MacPhie IL, Cherny SS, Richardson AJ, Talcott JB, Stein JF, Monaco AP, Cardon LR: Use of multivariate linkage analysis for dissection of a complex cognitive trait. Am J Hum Genet 2003; 72:561–570
28. 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; 296:2225–2229
29. 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; 62:811–817
30. Marchini J, Cardon LR, Phillips MS, Donnelly P: The effects of human population structure on large genetic association studies. Nat Genet 2004; 36:512–517
31. Hirschhorn JN, Daly MJ: Genome-wide association studies for common diseases and complex traits. Nat Rev Genet 2005; 6: 95–108