

A 77-Kilobase Region of Chromosome 6p22.2 Is Associated with Dyslexia in Families From the United Kingdom and From the United States

Clyde Francks,^{1,*} Silvia Paracchini,^{1,*} Shelley D. Smith,³ Alex J. Richardson,² Tom S. Scerri,¹ Lon R. Cardon,¹ Angela J. Marlow,¹ I. Laurence MacPhie,¹ Janet Walter,² Bruce F. Pennington,⁴ Simon E. Fisher,¹ Richard K. Olson,⁵ John C. DeFries,⁵ John F. Stein,² and Anthony P. Monaco¹

¹Wellcome Trust Centre for Human Genetics and ²Department of Physiology, University of Oxford, Oxford, United Kingdom; ³Department of Pediatrics, University of Nebraska Medical Center, Omaha; ⁴Department of Psychology, University of Denver, Denver; and ⁵Institute for Behavioral Genetics, University of Colorado, Boulder

Several quantitative trait loci (QTLs) that influence developmental dyslexia (reading disability [RD]) have been mapped to chromosome regions by linkage analysis. The most consistently replicated area of linkage is on chromosome 6p23-21.3. We used association analysis in 223 siblings from the United Kingdom to identify an underlying QTL on 6p22.2. Our association study implicates a 77-kb region spanning the gene *TTRAP* and the first four exons of the neighboring uncharacterized gene *KIAA0319*. The region of association is also directly upstream of a third gene, *THEM2*. We found evidence of these associations in a second sample of siblings from the United Kingdom, as well as in an independent sample of twin-based sibships from Colorado. One main RD risk haplotype that has a frequency of ~12% was found in both the U.K. and U.S. samples. The haplotype is not distinguished by any protein-coding polymorphisms, and, therefore, the functional variation may relate to gene expression. The QTL influences a broad range of reading-related cognitive abilities but has no significant impact on general cognitive performance in these samples. In addition, the QTL effect may be largely limited to the severe range of reading disability.

Introduction

Developmental dyslexia (reading disability [RD] [MIM 600202]) is a specific difficulty learning to read, in the absence of obvious causes, such as low general intelligence or lack of educational opportunity (Smith et al. 1996). The prevalence is estimated at ~5% in school-aged children (Smith et al. 1996). Dyslexia is highly familial and heritable, but the genetic background is complex and heterogeneous (DeFries et al. 1987; Fisher and DeFries 2002; Francks et al. 2002). Linkage studies have identified several genomic regions that may contain QTLs for dyslexia susceptibility, including regions on chromosomes 2, 3, 6, 15, and 18 (Fisher and DeFries 2002; Fisher et al. 2002). A candidate gene for dyslexia susceptibility (*DYX1C1*), located on 15q21, was recently reported to be disrupted by a balanced translocation in subjects with RD from one family (Taipale et al. 2003). Suggestive evidence was also

reported of associations, at a more general population level, between RD and putative functional SNPs within this gene (Taipale et al. 2003). However, our analysis of *DYX1C1* in 264 families with RD from the United Kingdom (Scerri et al., in press) and in 148 families from Canada (Wigg et al., in press) casts doubt on the functional significance of these SNPs, since, in these samples, RD appears to be associated with SNP alleles opposite to those implicated by Taipale et al. (2003). Therefore, the *DYX1C1* associations await convincing replication.

The most consistently replicated linkage to RD is on 6p23-21.3; this linkage has been found in five separate samples from the United Kingdom and the United States (Cardon et al. 1994, 1995; Grigorenko et al. 1997; Fisher et al. 1999, 2002; Gayan et al. 1999; Grigorenko et al. 2000; Fisher and DeFries 2002). The 6p23-21.3 region showed one of the two strongest linkage signals in our previous genomewide screen of one of these samples (223 siblings in 89 unrelated nuclear families from the United Kingdom, each with at least one proband with RD) (Fisher et al. 2002; Marlow et al. 2003). However, linkage and association studies have failed, so far, to refine the candidate interval on 6p to smaller than ~16 Mb (Fisher and DeFries 2002; Kaplan et al. 2002; Turic et al. 2003; Deffenbacher et al. 2004).

Developmental dyslexia is often characterized by im-

Received July 15, 2004; accepted for publication September 29, 2004; electronically published October 22, 2004.

Address for correspondence and reprints: Dr. Anthony P. Monaco, Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Headington, Oxford OX3 7BN. E-mail: anthony.monaco@well.ox.ac.uk

* These authors contributed equally to this work.

© 2004 by The American Society of Human Genetics. All rights reserved.
0002-9297/2004/7506-0010\$15.00

pairments in several conceptually distinct cognitive skills that are related to reading or language processing (Castles and Coltheart 1993; Gayan and Olson 2001). In our studies, we have assessed probands and their siblings for a range of reading and language-related measures, such as single-word reading ability, spelling ability, and measures of phonological and orthographic coding (sounding out unfamiliar words and recognizing the spelling of familiar words, respectively [see the “Material and Methods” section]). These measures are strongly intercorrelated in clinical and epidemiological samples (Castles and Coltheart 1993; Gayan and Olson 2003) (e.g., correlations 0.41–0.76 in 223 U.K. siblings [Marlow et al. 2001]). The measures also show significant but weaker correlations with verbal/nonverbal intelligence (correlations 0.14–0.40 in 223 U.K. siblings [Marlow et al. 2001]). Twin studies have shown that the different reading-related traits may be influenced by both overlapping and trait-specific genetic factors (Gayan and Olson 2003). Likewise, twin studies also indicate that variabilities in reading-related measures and general intelligence have both overlapping and independent genetic components (Gayan and Olson 2003).

We previously used multivariate linkage analysis to show that the QTL on 6p23-21.3 appeared to influence trait variability that was shared between all reading-related cognitive measures but that was not shared with measures of general intelligence (IQ) (Marlow et al. 2003). We therefore reasoned that shared variance between reading-related measures and IQ could be considered as noise in the specific context of further linkage and association analysis of the 6p23-21.3 QTL. In the present study, we refined our linkage mapping of the QTL to a 5.8-Mb interval by removing the variance in reading-related measures that is shared with IQ. We then selected brain-expressed genes within the new candidate interval for association analysis with SNPs. We detected significant association with IQ-adjusted RD measures in our sample of 223 U.K. siblings, within a 77-kb region of strong intermarker linkage disequilibrium (LD) that spans the gene for TTRAP (tumor necrosis factor [TNF], tumor receptor-associated factor [TRAF], and TNF receptor-associated protein), and the first four exons of the neighboring uncharacterized gene *KIAA0319*. We replicated this QTL effect in two large, independent samples of sibships collected in the United Kingdom and the United States; these samples were selected to include probands who were severely affected with RD. We found that one main risk haplotype, with a frequency of ~12%, was associated with RD in both the U.K. and U.S. samples. This is the first identification of a QTL that has been shown, by repeated replication, to be of relevance to many individuals with RD.

Material and Methods

The U.K. Family Sample (Samples 1 and 2)

We identified 264 unrelated nuclear families at the dyslexia clinic at the Royal Berkshire Hospital in Reading, United Kingdom. The majority (>90%) of the first 192 families were recruited on the basis of having (1) at least one proband whose single-word reading ability was >2 SDs below that predicted by tests of verbal or nonverbal reasoning (Marlow et al. 2001; Fisher et al. 2002) (see below for specific tests used) and (2) evidence of RD in one or more siblings of the proband with dyslexia (e.g., on the basis of parental reports or school history) (Marlow et al. 2001; Fisher et al. 2002). Many of the probands in these families had above-average verbal/nonverbal reasoning and roughly normal reading-related abilities (Marlow et al. 2003). To increase the representation of probands with poor reading in the sample, the majority (~80%) of the remaining 72 families were recruited through probands that were required to have single-word reading ability ≥ 1 SD below that predicted for their age, with a minimum IQ of 90 (again, see below for the specific tests). These families were recruited through a minimum of just one proband with RD, with no requirement for RD in siblings of the probands. All of the remaining families in both recruitment phases were collected via referral to the clinic of at least one reading-disabled proband by a qualified clinician. In practice, all of the sibships that were collected under the second recruitment scheme included a proband that met criteria for the first scheme (and 51% of first-scheme sibships included a proband that met second-scheme criteria). A battery of reading-related and cognitive tests was administered to all U.K. probands and siblings, regardless of recruitment, for the purpose of quantitative genetic analysis (see below). Subjects were given a full description of the experimental procedures, plus the option to ask questions or to withdraw at any time. Written informed consent was then obtained.

Blood samples or buccal swabs were donated by all available children and parents for the purpose of genomic DNA extraction. The 264 families comprised a total of 630 siblings for whom psychometric test scores and genomic DNA were available. Sibship sizes used in our analyses ranged from 2 to 5. Sibling ages ranged from 5.6 years to 30.6 years (median 11.7 years). The sample consisted uniformly of families of European origin.

We refer to the first 89 U.K. families (all from the first recruitment scheme) as “sample 1”; these were used in our previous genomewide linkage screen (Fisher et al. 2002) and also in the linkage analysis and initial association analysis that we report in the present study. Of these families, 82 were originally analyzed by Fisher et al. (1999) in a targeted linkage study of chromosome

6p. The remaining 175 U.K. families (“sample 2,” a mixture of both recruitment schemes) were used in the present study to test for replication of the associations found in sample 1.

The Colorado Family Sample (Sample 3)

The Colorado sample (sample 3) comprised 159 families drawn from the Colorado twin study of RD (DeFries et al. 1987; Gayan and Olson 2003). We identified twin pairs from the records of 27 Colorado school districts and selected those in which at least one member had a positive school history of reading difficulty (Gayan and Olson 2003). Each proband and all siblings were then administered a battery of psychometric tests (see below) for the purpose of quantitative genetic analysis. The sample investigated in the present study also included nontwin siblings in a significant proportion of families, which yielded 369 total siblings. Sibling ages ranged from 8.0 years to 18.9 years (median 10.8 years). We did not use pairs of MZ twins in this study. Genomic DNA samples were extracted from all siblings and available parents, and some samples were preamplified using the GenomiPhi kit (Amersham). Of the 159 families analyzed, 109 were used for our previous genome screen (Fisher et al. 2002). All families were of European descent. We used sample 3 for a further test of replication of our association findings.

Phenotype Measures: U.K. Samples (Samples 1 and 2)

We administered a battery of psychometric tests to all probands and cosiblings in each family, and we age-adjusted and standardized their scores against a normative control data set, as described elsewhere (Marlow et al. 2001; Fisher et al. 2002). These included measures of single-word reading ability (READ), spelling ability (SPELL), phonological decoding ability (PD) (use of rule-like letter-sound relations to derive the pronunciation of pseudowords), phonemic awareness (PA) (awareness of the phonemic structure of language), orthographic coding (OC-irreg) (use of word-specific spelling patterns to recognize real irregular words), orthographic coding assessed by a forced word choice test (OC-choice), and tests of verbal (SIM) and nonverbal (MAT) reasoning. Each measure was based on the percentage accuracy for an individual psychometric test (Marlow et al. 2001). All of these measures are significantly intercorrelated and familial (Marlow et al. 2001; Fisher et al. 2002). We used the sum of SIM and MAT as a proxy measure of IQ (Thompson 1982), and we linearly regressed reading-related scores on this IQ measure, for the total combined U.K. sample (samples 1 and 2), to obtain IQ-adjusted Z scores (mean 0 [1 SD]).

Phenotype Measures: U.S. Sample (Sample 3)

A battery of psychometric tests was administered to all twins and cosiblings in the U.S. sample. These tests were aimed at tapping the same range of cognitive skills as in the U.K. sample, although the actual tests differed in all but one case (OC-choice was common to both samples). We therefore had conceptually equivalent measures for sample 3 as in the U.K. samples, for the measures READ, SPELL, PD, and PA, together with the same measure of OC-choice as in the U.K. samples. The tests have been described elsewhere (Fisher et al. 2002; Gayan and Olson 2003). Unlike the U.K. measures, some of the Colorado measures were derived as composites from more than one test and/or from accuracy combined with latency data (Gayan and Olson 2003). Briefly, READ was derived as a composite of accuracy on two word recognition tests (Gayan and Olson 2003). PD was a composite of accuracy and latency on one- and two-syllable nonword reading. PA was a composite of accuracy on a phoneme-deletion task and weighted percentage correct on a phoneme-segmentation and -transposition task. SPELL and OC-choice were based on accuracy for individual tests of spelling and orthographic choice tasks, respectively. The reading-related measures were age-adjusted and standardized against an appropriate normative twin-based population (Gayan and Olson 2003). All of these measures are significantly intercorrelated and familial (Fisher et al. 2002; Gayan and Olson 2003). A standardized measure of full-scale IQ (Gayan and Olson 2003) was used for linear IQ adjustment of reading-related measures, in the same way as for the U.K. samples.

Genotyping

We had microsatellite marker genotype data for the families in sample 1 and sample 3 that were included in our previous genomewide screens (Fisher et al. 2002). Nineteen additional polymorphic microsatellite markers were genotyped as part of the present study in sample 1, in accordance with standard protocols (Fisher et al. 2002). Between 6 and 16 microsatellite markers were also genotyped for families in sample 2. Microsatellite genotype data were used only for linkage analysis or to increase the accuracy of identity-by-descent-sharing estimates for the modeling of residual linkage during association analysis (see below). Because of their high number of alleles and higher mutability compared with SNPs, microsatellites were not tested for trait-marker association.

All SNP genotyping was performed using the Sequenom system, in accordance with the manufacturer's instructions. PCR primers and primer-extension probes were designed with the SpectroDESIGNER software, with the exception of the probes for three insertion/deletion polymorphisms, which were designed manually. All primer and probe sequences are available on request.

In total, 57 SNPs were genotyped in sample 1, of which 48 spanned a 225-kb interval that surrounded our initial positive-association findings (see the “Results” section), with an average intermarker distance of 5 kb. Twenty of these SNPs—those that showed association in sample 1 or that were included in the same assay multiplexes—were genotyped in samples 2 and 3. We identified and eliminated Mendelian errors in our genotype data and used MERLIN (Abecasis et al. 2002) to identify and eliminate genotypes that indicated unlikely recombination events. We also used MERLIN to infer missing genotypes from relatives in unambiguous situations. No SNPs deviated significantly from Hardy-Weinberg equilibrium (data not shown).

Linkage Analysis

Quantitative linkage analysis of IQ-adjusted reading-related measures was performed using the 223 siblings from the first 89 U.K. families (sample 1), using polymorphic microsatellite genotype data from our previous genomewide linkage screen (Fisher et al. 2002), supplemented with data from 19 new microsatellite markers, and using the deCODE genetic map (Kong et al. 2002). Linkage analysis was performed within the variance-components (VC) framework (see below), as implemented in the package MERLIN (Abecasis et al. 2002), with no dominance variance and with the estimation of a single-trait mean. VC linkage analysis within this sample has been validated elsewhere by simulation (Fisher et al. 2002).

Intermarker LD Analysis and Haplotypes

Pairwise intermarker LD statistics for SNPs were calculated using HAPLOXT from the GOLD package (Abecasis and Cookson 2000), by use of founder-haplotype estimates derived from MERLIN (Abecasis et al. 2002). The GoldSurfer package was used to evaluate simultaneously pairs of LD features with a three-dimensional graphical interface (Pettersson et al., in press). To gauge the haplotype diversity within this region, SNP-based haplotypes that spanned LD region B (see the “Results” section) were estimated using MERLIN. Three SNPs (*rs4504469*, *rs2038137*, and *rs2143340*) within LD region B were identified, which, together, distinguished all of the observed haplotypes spanning region B with frequencies >3% (data not shown). We next used genotype data for these three SNPs as input for the haplotype-estimating method of Rohde and Fürst (2001), which uses a combination of an expectation-maximization algorithm and nuclear-family data to assign the most likely haplotypes to each individual, regardless of sporadic missing data. The best-estimate haplotypes from this method were used in marker-trait association analysis. Marker *rs2143340*, which distinguishes the risk haplotype from

all other common haplotypes, was also genotyped in a panel of 190 random human control individuals from the European Collection of Cell Cultures (ECACC).

Marker-Trait Association Analysis

Marker-trait association was assessed using the “total-association” option of the package QTDT (Abecasis et al. 2000). In this analysis, maximum-likelihood modeling is first used to fit a null VC model to quantitative-trait data in sibships in which variance is partitioned into unshared environmental, shared environmental/polygenic, and QTL-specific components, with a single-trait mean. Subsequently, a full-association model is fitted that includes the same VC as the null model but includes the mean effects of the SNP alleles to model association (Abecasis et al. 2000). Residual familial and linkage effects are therefore explicitly modeled within this framework and do not bias the association test. The likelihood-ratio test statistic is distributed asymptotically as a χ^2 with 1 df. This “total-association” test can be biased by population stratification, so we verified our results by using the QTDT package to perform a less powerful within-sibship orthogonal association test that is robust to stratification (the “orthogonal” test) (Abecasis et al. 2000). The orthogonal test produced significant associations within the same genomic interval as the total-association test, despite having lower power than the total-association test (data not shown). We performed a genotype-permutation procedure within QTDT to verify the results from the orthogonal test and to correct for any biasing effects of trait nonnormality; the significance levels derived from these permutations were similar to the asymptotic predictions for this test (data not shown). This suggested that trait nonnormality and/or phenotypic selection had not substantially biased the orthogonal test. There is no comparable way to obtain empirical significance levels for the total-association test, but we reasoned that the relatively unbiased behavior of the orthogonal test indicated that the total test should also be robust in our samples, since both tests are similar in design (Abecasis et al. 2000). We present the total-association results without allowing for dominance effects; when we allowed for dominance, the results remained effectively unchanged (data not shown).

Selection for RD Severity

For both the U.K. and Colorado data sets, we aimed to increase our power to detect the QTL, in association analysis, by selecting only sibships that included individuals scoring below a threshold that was derived from those measures that showed the best evidence of linkage across 6p23-21.3 in each sample. Our multivariate linkage analysis had previously suggested that the 6p21.3-23 QTL influenced variability that was shared between

all RD traits but that was not shared with IQ (Marlow et al. 2003). We therefore hypothesized that differences in the strengths of univariate linkages with the different RD traits were largely stochastic in nature and were influenced by trait- and sample-specific variations in sample sizes and patterns of missing genotype/phenotype data. Nonetheless, we reasoned that, within each sample, the relative strengths of univariate linkages could be used to indicate which specific phenotypes should be employed for severity selection, to maximize the power to detect association effects within any given sample. In the U.K. samples (samples 1 and 2), we used a threshold of -0.5 on a score calculated as the mean of individuals' IQ-adjusted PD and OC-irreg (see linkage data for sample 1 in fig. 1; linkage in sample 2 was not significant [data not shown]). In the U.S. sample, we used a threshold of 0 on the mean of individuals' IQ-adjusted READ and PD (these measures showed the strongest evidence of linkage in this sample [data not shown]). The threshold of 0 does not correspond to the unselected population mean, since the IQ adjustment was performed within the sample (see the "Material and Methods" section). The choice of a threshold value is necessarily arbitrary. Furthermore, the U.K. and Colorado samples could not be compared directly, since they were ascertained, tested, and standardized differently. In our choice of thresholds, we aimed to balance the increased potential power from phenotype selection against the decreased power from reduced sample size, as well as to have comparable sample sizes in the U.K. and Colorado selected samples (table 3). The phenotype selection yielded 126 families that included 313 siblings from the combined U.K. samples (samples 1 and 2) and 124 families that included 290 siblings from the U.S. sample (sample 3). The selected and total samples did not differ

significantly for their IQs in either the U.K. or U.S. data sets. The P values that we present are not corrected for multiple hypothesis testing; a Bonferroni correction would be too conservative for multiple correlated traits and nonindependent genetic markers. Appropriate corrections for correlated phenotypes and genotypes are presently unknown.

Polymorphism Screening

All of the known coding exons and predicted promoters of *TTRAP*, *KIAA0319*, *THEM2*, and *ALDH5A1*, as well as up to 289 bp of flanking sequence around each exon or promoter, were PCR amplified in 32 individuals with dyslexia from sample 1, who had the lowest mean scores derived from the IQ-adjusted measures PD and OC-irreg. The PCR products were screened using the WAVE DNA Fragment Analysis System (Transgenomic), which performs denaturing high-performance liquid chromatography (DHPLC) with hybridized amplicon duplexes across a range of temperatures. Any samples that showed heteroduplex formation were then resequenced. Fluorescence-based dideoxy sequencing was performed using BigDye Terminator sequencing kits (Applied Biosystems), followed by ABI 3700 capillary electrophoresis (Applied Biosystems). Sequencing primer details are available on request.

Results

Refinement of Linked Genomic Interval

Quantitative sib-pair linkage analysis, controlling for IQ as a linear covariate, provided increased evidence of linkage in our 223 U.K. siblings (sample 1), compared with analyses without IQ adjustment (fig. 1) (PD LOD

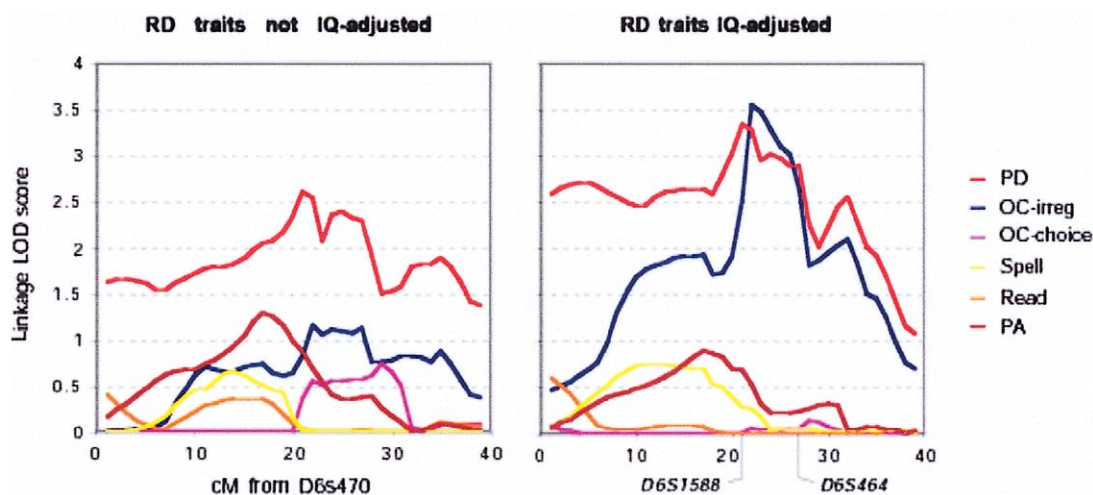


Figure 1 Linkage across 6p24.3-21.1 with reading-related measures in sample 1 (223 U.K. siblings, 195 total sibling pairs), before and after adjustment for IQ. IQ adjustment increased the strength of linkage and refined the QTL position.

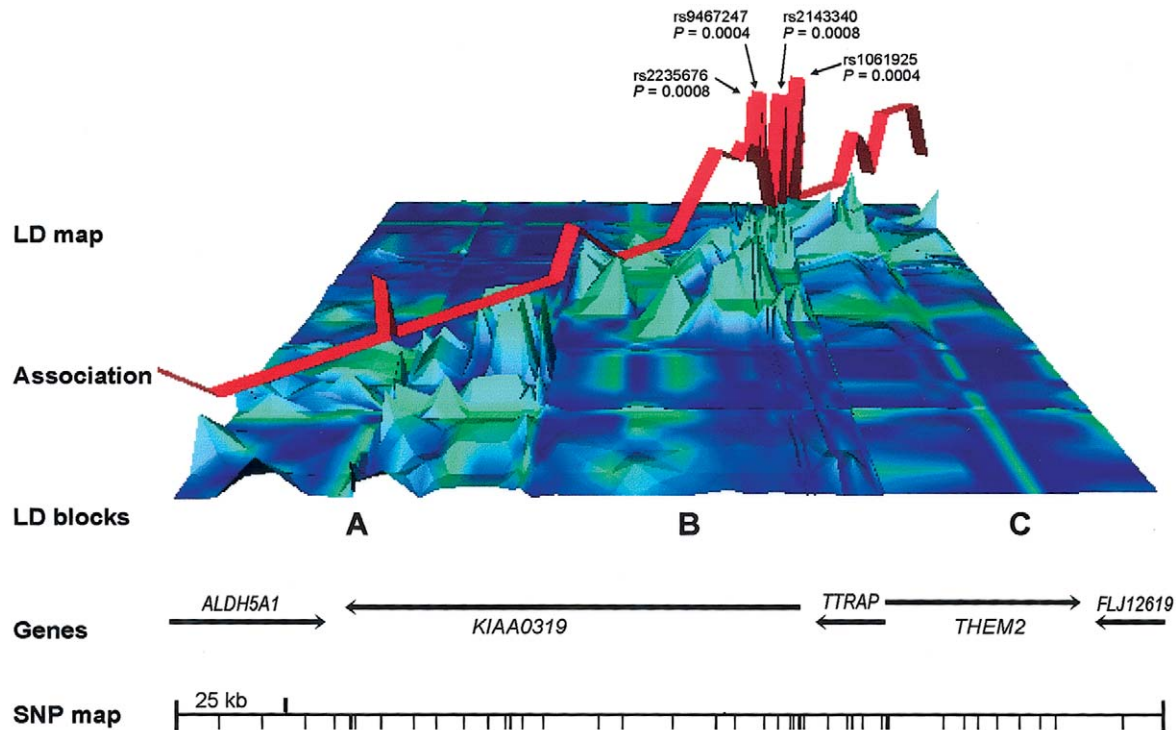


Figure 2 Representation of the 225-kb region surrounding the QTL in sample 1. The LD map shows a three-dimensional color-coded plot for pairwise intermarker LD statistics (48 biallelic markers). The colors represent D' values: green indicates high LD, and blue indicates low LD. The height of the peaks represents the $-\log_{10}$ significance of pairwise LD; high peaks indicate significant LD. Three major LD regions are distinguishable: A, B, and C. The red ribbon represents the significance of SNP associations ($-\ln[P \text{ value}]$) with the IQ-adjusted orthographic coding (OC-irreg) phenotype (see also table A1 [online only]). The SNPs that showed the strongest associations (and the corresponding P values) are indicated by arrows. The genes within the region and the locations of the SNPs are shown at the bottom.

2.62 before IQ adjustment, LOD 3.34 after IQ adjustment; OC-irreg LOD 1.16 before IQ adjustment, LOD 3.48 after IQ adjustment). The adjustment for IQ focused our association mapping on the now clearly defined 1-LOD-unit interval obtained from linkage analysis of IQ-adjusted orthographic coding ability (OC-irreg [see the “Material and Methods” section]); this interval spanned 5.8 Mb from marker *D6S1588* to marker *D6S464* (fig. 1). IQ data were not available from 38 siblings in sample 1, and part of the striking increase in LOD score with IQ-adjusted OC-irreg, compared with analysis of the unadjusted trait, arises from the loss of these individuals. However, even compared with analysis of unadjusted traits in the 186 siblings with IQ data (not shown), the IQ adjustment resulted in an increase of 1.24 LOD units for OC-irreg and 0.77 LOD units for PD.

Association and LD Analysis

There were 80 known genes within the 5.8-Mb interval between *D6S1588* and *D6S464* in June 2002 database releases. Of these known genes, 40 encode histone proteins, which are not obvious functional candidates for a specific cognitive disorder. Of the remaining 40

genes, 8 were known to be expressed in brain tissue and were selected for association studies (distal to proximal: *ALDH5A1*, *KIAA0319*, *TTRAP*, *THEM2*, *C6orf32*, *SCGN*, *BTN3A1*, and *BTN2A1* [LocusLink IDs: 7915, 9856, 51567, 55856, 9750, 10590, 11119, and 11120, respectively]). Summaries of the known functions of most of these genes are also available in the articles by Londin et al. (2003) and Deffenbacher et al. (2004). These eight genes fall within four distinct genomic clusters in the 5.8-Mb interval (University of California–Santa Cruz Genome Bioinformatics Web site); this facilitated association mapping, which relies on LD between markers that are in close proximity to one another (cluster 1: *ALDH5A1*, *KIAA0319*, *TTRAP*, and *THEM2*; cluster 2: *C6orf32*; cluster 3: *SCGN*; and cluster 4: *BTN3A1* and *BTN2A1*).

After the analysis of 15 SNPs in sample 1 within these gene clusters, we obtained significant evidence of association of SNP *rs1061925* with several IQ-adjusted measures ($P = .0004$ for OC-irreg; $P = .0114$ for READ; $P = .0269$ for OC-choice) (table A1 [online only]). We then analyzed 42 additional SNPs in sample 1 (including 4 biallelic insertion/deletions), within a 225-kb region surrounding *rs1061925* (fig. 2 and table A1 [online only]). SNP associations with several reading-related measures

were significant within a 77-kb region of strong inter-marker LD (region B [fig. 2; online-only tables A1 and A3]) that spans *TTRAP* and the first four exons of the neighboring gene *KIAA0319* (fig. 2). Our linkage analysis suggested that the QTL effect in sample 1 was most apparent in those sibships that included individuals who were severely impaired on IQ-adjusted OC-irreg and IQ-adjusted PD (fig. 1). Accordingly, we found that the significant marker-trait associations within LD region B (fig. 2) were derived predominantly from a subset of sibships that included at least one proband who scored < -0.5 SD on a mean score calculated from IQ-adjusted OC-irreg and IQ-adjusted PD (table 1).

We then analyzed 20 of the polymorphisms in a second, independent set of 407 siblings from 175 unrelated U.K. nuclear families (sample 2). Each sibship again included at least one proband with RD. The evidence of marker-trait association within sample 2 was weak, and a combined analysis with sample 1 produced attenuated evidence of association (data not shown). However, we then analyzed the subset of sample 2 (201 siblings) that included at least one sib who met the -0.5 SD criterion for the mean of IQ-adjusted PD and OC-irreg, and again the evidence of association within LD region B was significant for many SNPs and for a range of IQ-adjusted RD measures (table 2). This replicated our initial findings in sample 1, and it confirmed that the effect is strongest in the lower tail of reading-related cognitive ability. Combined analysis of the subsets from sample 1 and sample 2 that met the -0.5 SD criterion (313 siblings in 126 families) yielded highly significant evidence of associations within LD region B with most reading-related measures (table 3) (e.g., *rs9467247* and OC-irreg, $P = .0006$; *rs9467247* and READ, $P = .0003$; *rs1061925* and OC-choice, $P = .0005$; and *rs1061925* and READ, $P = .0008$).

We next analyzed 21 of the SNPs in an epidemiological sample of 369 siblings from 159 unrelated nuclear twin-based families from Colorado, each of which included at least one proband with RD (sample 3) (DeFries et al. 1987; Gayan and Olson 2003). Previous analyses of subsets of this sample, which used different markers from those described here, produced suggestive evidence of marker-trait associations within *TTRAP* and *KIAA0319*, although these studies also produced similarly suggestive evidence of association elsewhere within 6p23-21.3 (Kaplan et al. 2002; Deffenbacher et al. 2004). Our association analysis of sample 3 produced only weakly significant evidence of marker-trait association within the putative QTL (data not shown). However, we aimed once more to replicate the selection scheme used for the U.K. samples, by choosing families in which at least one sibling scored poorly on their average of IQ-adjusted READ and PD (see the "Material and Methods" section), since these measures showed the

best evidence of linkage in this sample. The evidence of association was again significant for SNPs within the same genomic interval as that implicated in the U.K. samples (table 3) (e.g., *rs9467247* and READ, $P = .0038$; *rs9467247* and PA, $P = .042$; *rs3033236* and READ, $P = .0023$; *rs3033236* and SPELL, $P = .015$). We also tried other threshold-based selection methods, in addition to those described here, and found that all methods that yielded similar selected sample sizes also produced broadly comparable effects for the association data, but the associations were consistently strongest if the prior selection had been performed using the RD traits that showed the strongest linkage within each sample (data not shown). In addition, association analysis with the IQ-unadjusted RD phenotypes gave a pattern of results that was similar to that of the association analysis with IQ-adjusted traits, although the overall level of significance was lower.

Haplotype association analysis revealed one main risk haplotype that spanned the QTL (LD region B), that was common to the U.K. and U.S. populations, that had a frequency of $\sim 12\%$ in all samples, and that was effectively distinguished from all other common haplotypes by the SNP *rs2143340* (table 4). We genotyped *rs2143340* in 190 control individuals of European ancestry (ECACC) and found the minor allele frequency to be effectively the same as in the parents from the U.K. and U.S. samples (15%–16%). This frequency is higher than the 12% frequency for the main risk haplotype, because the minor allele of *rs2143340* occasionally falls on other haplotypes. We observed an increase in the minor allele frequency of *rs2143340* (up to 28% [table 5]) only in U.K. and U.S. siblings that were selected for RD phenotypic severity (table 5). These sample selections were based on the same combinations of two phenotypic measures that showed the strongest evidence of linkage in either sample (see above and table 5).

Candidate-Gene Screening

The region of association spans the genomic extent of *TTRAP* but is also directly upstream of the proximal neighboring gene *THEM2* (thioesterase superfamily member 2), and it encompasses the first four exons of the distal gene *KIAA0319* (uncharacterized brain-expressed transcript) (fig. 2). We screened the exons and predicted promoters of *TTRAP*, *KIAA0319*, *THEM2*, and another neighboring gene (*ALDH5A1* [fig. 2]) for mutations in 32 probands with severe RD. However, we did not find any variants that have obvious disruptive effects on these genes (table A2 [online only]). No coding polymorphisms were detected that tagged the main risk haplotype. The only SNP that we identified within the trait-associated region (LD region B [fig. 2]) that had an effect on protein sequence was *rs4504469* (table A1 [on-

Table 1
Marker-Trait Association P Values in Selected Siblings from Sample 1 (112 Siblings, 42 Families)

MARKER	LD REGION	MINOR ALLELE FREQUENCY ^a (%)	RISK ALLELE ^b	P VALUE FOR TRAIT					
				OC-irreg	OC-choice	PD	READ	SPELL	PA
<i>rs2817220</i>	A	.23	Major	.0326		.0067			.0004
<i>rs22252525</i>	A	.13							
<i>rs807517</i>	A	.24	Major			.0223			.0017
<i>rs1054899</i>	A	.34							
<i>rs2760184</i>	A	.04							
<i>rs807527</i>	A	.10							
<i>rs699463</i>	A	.24	Major						.0068
<i>rs807530</i>	A	.41							
<i>rs807535</i>	A	.10							
<i>rs807540</i>	A	.08	Major	.0203			.0118	.0085	
<i>rs807545</i>	A	.06	Major					.0244	
<i>rs807521</i>	A	.06	Major					.0306	
<i>rs807525</i>	A	.06	Major					.0308	
<i>rs807507</i>	A	.43							
<i>rs807508</i>	A	.06	Major					.0304	
<i>k_ex5</i>	A	.30	Major						.044
<i>rs4504469</i>	B	.47	Major	.0003			.0231		.0005
<i>rs4576240</i>	B	.05							
<i>rs4472344</i>	B	.03							
<i>rs4236032</i>	B	.05							
<i>rs2745333</i>	B	.45	Major	.0002			.0178		.0001
<i>rs2179515</i>	B	.40	Major	.0002			.0021		
<i>rs761101</i>	B	.39	Major	.0002		.0255	.0003		
<i>rs6456624</i>	B	.44	Major	.00006	.036	.0231	.0007		
<i>rs2328846</i>	B	.41	Major	.00003			.0001		.0377
<i>rs2235676</i>	B	.14	Minor	.0008	.0442		.0029		
<i>rs2038137</i>	B	.43	Major	.00006			.0002	.0285	
<i>k_pr_del</i>	B	.41	Major	.00002			.0001		.0377
<i>rs9467247</i>	B	.19	Minor	.0002	.0167		.0005		
<i>rs1555090</i>	B	.40	Major	.00003			.0001		.0377
<i>rs2294689</i>	B	.05							
<i>rs3033236</i>	B	.17	Minor	.0066	.0288		.0046		
<i>rs3212232</i>	B	.26							
<i>rs2143340</i>	B	.17	Minor	.0012	.0128		.0011		
<i>rs2056998</i>	B	.05							
<i>rs1061925</i>	B	.11	Minor	.00006	.0213		.0003		
<i>th_ex1</i>	B	.11	Minor	.0004	.0214		.0011		
<i>tt_th_del</i>	B	.06							
<i>rs3181228</i>	B	.15							
<i>rs2328847</i>	C	.01							
<i>rs1555088</i>	C	.34	Major	.0178					
<i>rs926529</i>	C	.36	Major	.0177					
<i>rs1885211</i>	C	.46	Major	.009					
<i>rs2092404</i>	C	.34	Major	.0137			.0405		
<i>th_ex3</i>	C	.04							
<i>rs3756814</i>	C	.43	Minor	.0178					
<i>rs6456632</i>	C	.15							
<i>rs2067573</i>	C	.05							

NOTE.—The selection included all the sibships with at least one sibling scoring <−0.5 SD on an average score of IQ-adjusted OC-irreg and PD. All reading-related measures are IQ-adjusted.

^a Minor allele frequency calculated in the founders only.

^b Only alleles with nominally significant marker-trait association ($P < .05$) are shown.

line only)), within exon 4 of *KIAA0319*. However, the minor allele of this SNP has a frequency of 0.47 in our samples and therefore is not unique to the risk haplotype that we identified. Accordingly, the significances of phe-

notype associations for *rs4504469* were less than those of other SNPs that distinguished the risk haplotype more effectively (e.g., *rs1061925* [table 3]). The SNP *rs2143340*, which distinguishes the risk haplotype from

Table 2**Marker-Trait Association P Values in Selected Siblings from Sample 2 (201 Siblings, 84 Families)**

MARKER	LD REGION	P VALUE FOR TRAIT ^a			
		OC-choice	PD	READ	SPELL
<i>rs699463</i>	A				
<i>rs4504469</i>	B		.0256	.019	
<i>rs2179515</i>	B		.0433	.0283	.0431
<i>rs761101</i>	B			.0418	
<i>rs6456624</i>	B				
<i>rs2328846</i>	B		.0159	.0464	.0467
<i>rs2235676</i>	B	.0037			
<i>rs2038137</i>	B		.0128	.021	.0125
<i>k_pr_del</i>	B		.0269	.0377	.0207
<i>rs9467247</i>	B	0.0075			.0396
<i>rs1555090</i>	B		.0248	.0396	.0277
<i>rs3033236</i>	B				
<i>rs2143340</i>	B	.0053			
<i>rs1061925</i>	B	.0021			
<i>tt_th_del</i>	B				
<i>rs926529</i>	C		.0326		
<i>rs1885211</i>	C				
<i>th_ex3</i>	C		.0399		
<i>rs3756814</i>	C				
<i>rs6456632</i>	C				

NOTE.—The selection included all the sibships with at least one sibling scoring <-0.5 SD on an average score of IQ-adjusted OC-irreg and PD. All reading-related measures are IQ-adjusted.

^a OC-irreg and PA showed no significant association in this analysis.

all other common haplotypes, is within intron 2 of *TTRAP* and lies on an untranslated alternative transcript of this gene (Vega Human Genome Browser Web site).

Discussion

We have used quantitative-trait linkage and association analysis to identify a QTL that influences RD on 6p22.2, within the genomic region that has shown the most consistently replicated linkage to RD among studies internationally. We have refined this QTL to a 77-kb region of strong intermarker LD that spans the gene for *TTRAP* and part of the neighboring uncharacterized gene *KIAA0319*. The QTL is also directly upstream of a third uncharacterized gene, *THEM2*. We identified a single main risk haplotype that spanned the QTL, with a frequency of ~12% in our RD samples, which were drawn from populations of European descent in the United Kingdom and United States. We did not detect any coding polymorphisms within *KIAA0319*, *TTRAP*, or *THEM2* that have overt disruptive effects on these genes or that characterize the main risk haplotype. The variation at this QTL that is functionally relevant for RD may therefore influence the expression, splicing, or transcript stability of any of the *KIAA0319*, *TTRAP*, or *THEM2* genes. The strong LD between all polymorphisms within the region

of marker-trait association (fig. 2) means that the functionally relevant variation for RD may be difficult to identify by further genetic mapping, as has proven to be the case in studies of QTLs for other complex traits, including asthma and Crohn disease (Rioux et al. 2001; Zhang et al. 2003).

RT-PCR analysis and expression profiling show that *TTRAP* is expressed in most or all tissues (Londin et al. 2003) (GNF SymAtlas Web site). The *TTRAP* protein interacts with cytoplasmic TNF receptor-associated factors (TRAFs), as well as with the cytoplasmic domains of specific members of the TNF receptor superfamily (Pype et al. 2000). *TTRAP* may therefore be involved as a regulatory factor in TNF signal transduction. Overexpression of *TTRAP* inhibits activity of the TNF-responsive transcription factor NF- κ B (Pype et al. 2000). TNF signaling via NF- κ B regulates many normal physiological and pathogenic processes, including immune and autoimmune responses and tumorigenesis (Aggarwal 2003). NF- κ B activity in neurons also responds to synaptic signaling and may bring about long-term changes in neuronal function that subserve learning (Meffert et al. 2003), whereas inhibition of NF- κ B activity in forebrain neurons results in neurodegenerative-like phenotypes in mice (Fridmacher et al. 2003). Thus, *TTRAP* may potentially influence both normal and pathogenic neuronal processes via its effects on NF- κ B activity. *TTRAP* may also interact with members of the ETS family of transcription factors (Pei et al. 2003) and can also be conjugated to the small ubiquitin-like modifier SUMO-1, which modifies protein-protein interactions, subcellular localization, and stability (Lee et al. 2003).

KIAA0319 is expressed most strongly in brain tissue, with weaker expression in a restricted number of other tissues (Londin et al. 2003) (GNF SymAtlas Web site). *KIAA0319* is largely uncharacterized; the predicted protein contains a putative transmembrane domain and four polycystic kidney disease domains that have homology with extracellular domains of the polycystic kidney disease protein PKD1, in which they are involved in cell-adhesive functions (Streets et al. 2003). *THEM2* is also largely uncharacterized. The predicted protein has homology with thioesterase enzymes, and the gene is expressed in most tissues (Londin et al. 2003).

Our association data show clearly that the 6p22 QTL influences a broad range of reading-related cognitive skills in our samples (table 3). Fluctuating levels of univariate linkage of 6p23-21.3 to different reading-related phenotypes within a given study sample have been interpreted as evidence that this QTL affects some reading-related cognitive processes but not others (Grigorenko et al. 1997). The attraction of this interpretation lay in the prospect of being able to “dissect” reading-related cognition by the assignment to distinct genetic

Table 3

Marker-Trait Association P Values in Two Independent Selected Samples

MARKER	LD REGION	Risk Allele ^d	SELECTED SIBSHIPS FROM SAMPLES 1 AND 2, COMBINED (U.K.) (n = 313 SIBLINGS [126 FAMILIES])						SELECTED SIBSHIPS FROM SAMPLE 3 (U.S.) ^a (n = 290 SIBLINGS [124 FAMILIES])					
			P Value for Trait ^b						Risk Allele ^d	P Value for Trait ^c				
			OC-irreg	OC-choice	PD	READ	SPELL	PA		PD	READ	SPELL	PA	
rs699463	A	Major	.0032		.0231	.0279	.0153	.0112						
rs4504469	B	Major	.0011		.0082	.004		.01						
rs2179515	B	Major	.0012		.0131	.0004	.0232							
rs761101	B	Major	.0025		.0057	.0006	.0325							
rs6456624	B	Major	.0005		.0045	.0003	.0157							
rs2328846	B	Major	.0007		.0017	.0003	.0155							
rs2235676	B	Minor	.0023	.0009		.0041			Minor		.0131			
rs2038137	B	Major	.0013		.0026	.0002	.0061							
k_pr_del ^e	B	Major	.0011		.0032	.0002	.0086							
rs9467247	B	Minor	.0006	.0003	.0373	.0003	.0016		Minor		.0038			.0421
rs1555090	B	Major	.001		.0029	.0003	.0131							
rs3033236	B	Minor	.0134	.0104		.0073			Minor	.0421	.0023	.0145	.0204	
rs2143340	B	Minor	.01	.0003		.0115			Minor		.0050	.0273	.0124	
rs1061925	B	Minor	.0009	.0005		.0008			Minor		.0102	.0340		
tt_th_del ^e	B								Minor					.0181
rs926529	C	Major			.0132									
rs1885211	C													
th_ex_3 ^e	C													
rs3756814	C	Minor	.0332											
rs6456632	C	Major				.0415								

NOTE.—All the reading-related measures were IQ-adjusted.

^a Single-tailed P values for testing the same marker allele as in the U.K. samples.

^b IQ showed no significant association in this analysis.

^c OC-choice and IQ showed no significant association in this analysis.

^d Only alleles that showed nominally significant marker-trait association (P < .05) are shown.

^e Polymorphisms identified by mutation screening analysis, which were not present in the public databases.

loci of different cognitive components of reading. However, the specificity of the 6p QTL for distinct cognitive reading-related processes was not well replicated by linkage analysis as study sample sizes were increased or as new samples were analyzed (Fisher et al. 1999; Gayan et al. 1999; Grigorenko et al. 2000; Fisher and DeFries 2002; Francks et al. 2002). In our previous multivariate linkage analysis of 6p23-21.3 (performed in sample 1), we were unable to drop any single reading-related measure from the multivariate linkage model without significantly reducing the model fit, even though, in separate univariate tests, some of the measures showed no significant linkage at this locus (Marlow et al. 2003). Now, having performed association analysis directly at the presumed underlying QTL, using powerful association analysis based on modeling allelic mean effects, the true pattern of QTL pleiotropy across the full range of reading-related measures becomes apparent (tables 1–3).

It is worth noting that our association data, like previous linkage data, still apparently contain chance fluctuations in measured effect sizes for different RD traits; for example, the trait OC-choice shows highly significant association with the main risk haplotype in the

selected U.K. sample, but no significant association in the selected U.S. sample, despite being based on the same phenotypic test (table 3). Also, there are many instances in which a single SNP may show significant association with certain RD measures in one sample and association with a different set of RD measures in a different sample (tables 1–3). Again, we interpret this as stochastic variation that is influenced by relatively small individual sample sizes and by sample-specific patterns of missing genotype/phenotype data. It is worth noting that when the selected samples 1 and 2 are combined (table 3), the associations within LD region B become highly significant across most of this LD region, as well as across most of the RD measures, as would be expected if the underlying effect is truly pleiotropic. Our prior multivariate linkage analysis, together with the nature of the LD landscape in this region, suggest that the associations that we have found within LD region B can be considered as replications in the different samples, even though the same SNP/phenotype combinations are not always significant in all of the samples.

The pleiotropic effect of this QTL on a range of reading-related measures shows that the development of these cognitive abilities are at least partly interdepend-

Table 4**Association P Values of Most Common *rs4504469*, *rs2038137*, and *rs2143340* Haplotypes with Reading-Related Measures**

HAPLOTYPE	SELECTED SUBSET FROM SAMPLES 1 AND 2, COMBINED (U.K.)						SELECTED SUBSET FROM SAMPLE 3 (U.S.)						
	Frequency (%)	P Value for Trait ^a					Mean Effect ^c	Frequency (%)	P Value for Trait ^b				Mean Effect ^c
		OC-irreg	OC-choice	PD	READ	SPELL			PD	READ	SPELL	PA	
111	.41			.0306			-.10	.46					-.01
221	.36	.0019	.0314	.0017	.0004	.0165	.24	.31					-.01
112	.12	.0045	.00007		.0024	.0246	-.34	.12	.017	.0367	.0183		-.23
211	.05						.03	.05	.0187	.0119	.0243		.30
121	.04						-.04	.04	.0245	.0413			.33
Others	.02						-.10	.02					-.07

NOTE.—All reading-related measures are IQ-adjusted. 1 = major allele; 2 = minor allele.

^a PA showed no significant association in this analysis.^b OC-choice showed no significant association in this analysis.^c Mean phenotypic effect relative to all other haplotypes, calculated across the different measures.

ent. Even if this QTL produces a primary core deficit that affects one specific aspect of reading-related cognition, such a primary deficit may cause a general inhibitory effect on the ability to learn all reading-related skills. It remains possible that other RD susceptibility loci may cause even more specific cognitive deficits than this QTL—deficits limited to just a single reading-related measure. However, any developmental interdependence between different reading-related cognitive skills may make such a simple trait-locus correspondence unlikely for any locus. Furthermore, the functional roles of individual proteins in CNS development are manifold, multiregional, multitemporal, and interaction-dependent. Thus, it may be naive to postulate that a single genetic locus can somehow encode a single cognitive phenotype (Pennington 1997; Fisher, in press). In addition, a subject's performance on an individual cognitive test is unlikely to correspond to activity in any unitary system within the brain (Goldberg and Weinberger 2004). Therefore, the kind of relatively specific cognitive disruptions that can be associated with brain lesions in adults may not provide good models for understanding the developmental genetics of reading-related cognition (Paterson et al. 1999). However, it may become possible, through association analysis of this

QTL in large epidemiological samples, to reliably detect subtle differences in the relative mean effects of the locus on different RD measures and thereby to identify aspects of reading-related cognition that may be influenced particularly strongly by this QTL.

In our association models, the main “risk” haplotype that spanned the QTL had an average effect of -0.34 SD on IQ-adjusted reading-related measures within the phenotypically selected U.K. subset; the haplotype had an average effect of -0.23 SD in the Colorado subset, when compared with all other haplotypes (table 4). However, these effect-size estimates apply only to phenotypically selected samples. To obtain a precise estimate of the contribution of this QTL to RD susceptibility, it will be important to analyze its effect in other large clinical and epidemiological samples and also in relation to other QTLs for RD, as these are identified, since multilocus interactions are a possibility in RD etiology.

We detected the 6p22 QTL most strongly after selection of sibships for phenotypic severity. Consistent with this, the minor allele for SNP *rs2143340*, which distinguishes the “risk” haplotype from all other common haplotypes, was only substantially increased in frequency (up to 28%) in the probands who were most severely affected with RD (table 5). This allele was present at a base frequency of $\sim 16\%$, even in siblings in the upper range of reading ability. We genotyped *rs2143340* in a panel of 190 control individuals and again found a 16% frequency for the minor allele. There are a number of possible explanations for these findings. Under the simplest QTL model, which involves only additive effects that are assumed to be constant across the full range of ability, we calculate that an allele with 12% frequency (the risk haplotype frequency) and a phenotypic effect of -0.3 SD would be expected to have only a moderately elevated frequency (18%) in individuals with average phenotypic values 2 SD below the mean. This may explain why phenotypic selection was necessary to detect the QTL effect

Table 5**Minor Allele Frequency for SNP *rs2143340* within Phenotypic Subcategories in the U.K. (Samples 1 and 2) and U.S. (Sample 3) Siblings**

PHENOTYPIC GROUP	MINOR ALLELE FREQUENCY (%)				
	<-1.5	<-1	<-0.5	<0	ALL
U.K. ^a	28	19	16	16	16
U.S. ^b	23	16	16	15	15

^a The phenotypic groups were calculated on the average of IQ-adjusted OC-irreg and PD.^b The phenotypic groups were calculated on the average of IQ-adjusted READ and PD.

in our samples. Alternatively, the phenotypic effect of the risk haplotype may be contingent either on the genotype at another unidentified locus (i.e., epistasis) or on environmental influences, such that only a subset of risk haplotype carriers experience a detectable phenotypic effect. No other replicated QTLs for RD have yet been identified that can be analyzed in a joint epistatic association model. A third possibility is that the risk haplotype that we have identified carries functionally relevant variants only on a subset of its overall diversity.

In contrast to the pleiotropic effect of this QTL for many reading-related measures, there is no association of the QTL with IQ in either the U.K. or U.S. samples (table 3). Our data therefore show directly, for the first time, that reading-related cognition can be influenced by at least one relatively specific genetic-developmental deficit, while other cognitive abilities are left intact. The QTL may influence a pathogenic process that is mostly restricted to regions of the brain involved in reading and language ability (Paulesu et al. 2001). Alternatively, the pathogenic process may occur more broadly throughout the developing brain, but the resulting damage may have its most detrimental effect on certain neuronal circuits—or a class of neuronal operations—that are necessary for reading-related cognition.

This QTL is the first identified genetic influence to be shown by repeated replication to be of relevance to many individuals with specific RD. The functions of *TTRAP*, *KIAA0319*, and *THEM2* must now be studied in detail to determine which gene is responsible for the QTL effect and to identify the mechanism of pathogenesis.

Acknowledgments

We are grateful to all the families who participated in this study. We thank K. Taylor, for assistance with handling the U.K. phenotype data, and K. King, for helping with mutation screening. The Colorado Twin Study was supported in part by program project and center grants from National Institute of Child Health and Human Development (HD-11681 and HD-27802). I.L.M. was funded by the British Council and Natural Sciences and Research Council, S.E.F. is a Royal Society Research Fellow, and L.R.C. and A.P.M. are Wellcome Trust Principal Research Fellows. This work was funded by the Wellcome Trust.

Electronic-Database Information

The URLs for data presented herein are as follows:

dbSNP Home Page, <http://www.ncbi.nlm.nih.gov/SNP/>
 GNF SymAtlas, <http://symatlas.gnf.org/SymAtlas/>
 LocusLink, <http://www.ncbi.nlm.nih.gov/LocusLink/>
 Online Mendelian Inheritance in Man (OMIM), <http://www.ncbi.nlm.nih.gov/Omim/> (for RD and the 6p23-21.3 region)

University of California–Santa Cruz Genome Bioinformatics, <http://genome.ucsc.edu/>
 Vega Human Genome Browser, http://vega.sanger.ac.uk/Homo_sapiens/

References

- Abecasis GR, Cardon LR, Cookson WO (2000) A general test of association for quantitative traits in nuclear families. *Am J Hum Genet* 66:279–292
- Abecasis GR, Cherny SS, Cookson WO, Cardon LR (2002) Merlin—rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 30:97–101
- Abecasis GR, Cookson WO (2000) GOLD—graphical overview of linkage disequilibrium. *Bioinformatics* 16:182–183
- Aggarwal BB (2003) Signalling pathways of the TNF superfamily: a double-edged sword. *Nat Rev Immunol* 3:745–756
- Cardon LR, Smith SD, Fulker DW, Kimberling WJ, Pennington BF, DeFries JC (1994) Quantitative trait locus for reading disability on chromosome 6. *Science* 266:276–279
- (1995) Quantitative trait locus for reading disability: correction. *Science* 268:1553
- Castles A, Coltheart M (1993) Varieties of developmental dyslexia. *Cognition* 47:149–180
- Deffenbacher KE, Kenyon JB, Hoover DM, Olson RK, Pennington BF, DeFries JC, Smith SD (2004) Refinement of the 6p21.3 quantitative trait locus influencing dyslexia: linkage and association analyses. *Hum Genet* 115:128–138
- DeFries JC, Fulker DW, LaBuda MC (1987) Evidence for a genetic aetiology in reading disability of twins. *Nature* 329:537–539
- Fisher SE. Tangled webs: tracing the connections between genes and cognition. *Cognition* (in press)
- Fisher SE, DeFries JC (2002) Developmental dyslexia: genetic dissection of a complex cognitive trait. *Nat Rev Neurosci* 3:767–780
- Fisher SE, Francks C, Marlow AJ, MacPhie IL, Newbury DF, Cardon LR, Ishikawa-Brush Y, Richardson AJ, Talcott JB, Gayan J, Olson RK, Pennington BF, Smith SD, DeFries JC, Stein JF, Monaco AP (2002) Independent genome-wide scans identify a chromosome 18 quantitative-trait locus influencing dyslexia. *Nat Genet* 30:86–91
- Fisher SE, Marlow AJ, Lamb J, Maestrini E, Williams DF, Richardson AJ, Weeks DE, Stein JF, Monaco AP (1999) A quantitative-trait locus on chromosome 6p influences different aspects of developmental dyslexia. *Am J Hum Genet* 64:146–156
- Francks C, MacPhie IL, Monaco AP (2002) The genetic basis of dyslexia. *Lancet Neurol* 1:483–490
- Fridmacher V, Kaltschmidt B, Goudeau B, Ndiaye D, Rossi FM, Pfeiffer J, Kaltschmidt C, Israel A, Memet S (2003) Forebrain-specific neuronal inhibition of nuclear factor- κ B activity leads to loss of neuroprotection. *J Neurosci* 23:9403–9408
- Gayan J, Olson RK (2001) Genetic and environmental influences on orthographic and phonological skills in children with reading disabilities. *Dev Neuropsychol* 20:483–507
- (2003) Genetic and environmental influences on individual differences in printed word recognition. *J Exp Child Psychol* 84:97–123

- Gayan J, Smith SD, Cherny SS, Cardon LR, Fulker DW, Brower AM, Olson RK, Pennington BF, DeFries JC (1999) Quantitative-trait locus for specific language and reading deficits on chromosome 6p. *Am J Hum Genet* 64:157–164
- Goldberg TE, Weinberger DR (2004) Genes and the parsing of cognitive processes. *Trends Cogn Sci* 8:325–335
- Grigorenko EL, Wood FB, Meyer MS, Hart LA, Speed WC, Shuster A, Pauls DL (1997) Susceptibility loci for distinct components of developmental dyslexia on chromosomes 6 and 15. *Am J Hum Genet* 60:27–39
- Grigorenko EL, Wood FB, Meyer MS, Pauls DL (2000) Chromosome 6p influences on different dyslexia-related cognitive processes: further confirmation. *Am J Hum Genet* 66:715–723
- Kaplan DE, Gayan J, Ahn J, Won TW, Pauls D, Olson RK, DeFries JC, Wood F, Pennington BF, Page GP, Smith SD, Gruen JR (2002) Evidence for linkage and association with reading disability on 6p21.3-22. *Am J Hum Genet* 70:1287–1298
- Kong A, Gudbjartsson DF, Sainz J, Jonsson GM, Gudjonsson SA, Richardson B, Sigurdardottir S, Barnard J, Hallbeck B, Masson G, Shlien A, Palsson ST, Frigge ML, Thorgeirsson TE, Gulcher JR, Stefansson K (2002) A high-resolution recombination map of the human genome. *Nat Genet* 31:241–247
- Lee BH, Yoshimatsu K, Maeda A, Ochiai K, Morimatsu M, Araki K, Ogino M, Morikawa S, Arikawa J (2003) Association of the nucleocapsid protein of the Seoul and Hantaan hantaviruses with small ubiquitin-like modifier-1-related molecules. *Virus Res* 98:83–91
- Londin ER, Meng H, Gruen JR (2003) A transcription map of the 6p22.3 reading disability locus identifying candidate genes. *BMC Genomics* 4:25
- Marlow AJ, Fisher SE, Francks C, MacPhie IL, Cherny SS, Richardson AJ, Talcott JB, Stein JF, Monaco AP, Cardon LR (2003) Use of multivariate linkage analysis for dissection of a complex cognitive trait. *Am J Hum Genet* 72:561–570
- Marlow AJ, Fisher SE, Richardson AJ, Francks C, Talcott JB, Monaco AP, Stein JF, Cardon LR (2001) Investigation of quantitative measures related to reading disability in a large sample of sib-pairs from the UK. *Behav Genet* 31:219–230
- Meffert MK, Chang JM, Wiltgen BJ, Fanselow MS, Baltimore D (2003) NF- κ B functions in synaptic signaling and behavior. *Nat Neurosci* 6:1072–1078
- Paterson SJ, Brown JH, Gsodl MK, Johnson MH, Karmiloff-Smith A (1999) Cognitive modularity and genetic disorders. *Science* 286:2355–2358
- Paulesu E, Demonet JF, Fazio F, McCrory E, Chanoine V, Brunswick N, Cappa SF, Cossu G, Habib M, Frith CD, Frith U (2001) Dyslexia: cultural diversity and biological unity. *Science* 291:2165–2167
- Pei H, Yordy JS, Leng Q, Zhao Q, Watson DK, Li R (2003) EAPII interacts with ETS1 and modulates its transcriptional function. *Oncogene* 22:2699–2709
- Pennington BF (1997) Using genetics to dissect cognition. *Am J Hum Genet* 60:27–39
- Pettersson F, Jonsson O, Cardon LR. GOLDSurfer: three dimensional display of linkage disequilibrium. *Bioinformatics* (in press)
- Pype S, Declercq W, Ibrahim A, Michiels C, Van Rietschoten JG, Dewulf N, de Boer M, Vandenaabeele P, Huylebroeck D, Remacle JE (2000) TTRAP, a novel protein that associates with CD40, tumor necrosis factor (TNF) receptor-75 and TNF receptor-associated factors (TRAFs), and that inhibits nuclear factor- κ B activation. *J Biol Chem* 275:18586–18593
- Rioux JD, Daly MJ, Silverberg MS, Lindblad K, Steinhart H, Cohen Z, Delmonte T, et al (2001) Genetic variation in the 5q31 cytokine gene cluster confers susceptibility to Crohn disease. *Nat Genet* 29:223–228
- Rohde K, Fürst R (2001) Haplotyping and estimation of haplotype frequencies for closely linked biallelic multilocus genetic phenotypes including nuclear family information. *Hum Mutat* 17:289–295
- Scerri TS, Fisher SE, Francks C, MacPhie IL, Paracchini S, Richardson AJ, Stein JF, Monaco AP. Putative functional alleles of DYX1C1 are not associated with dyslexia susceptibility in a large sample of sibling pairs from the UK. *J Med Genet* (in press)
- Smith SD, Gilger JW, Pennington BF (1996) Dyslexia and other specific learning disorders. Churchill Livingstone, New York
- Streets AJ, Newby LJ, O'Hare MJ, Bukanov NO, Ibraghimov-Beskrovnaya O, Ong AC (2003) Functional analysis of PKD1 transgenic lines reveals a direct role for polycystin-1 in mediating cell-cell adhesion. *J Am Soc Nephrol* 14:1804–1815
- Taipale M, Kaminen N, Nopola-Hemmi J, Haltia T, Myllyluoma B, Lyytinen H, Muller K, Kaaranen M, Lindsberg PJ, Hannula-Jouppi K, Kere J (2003) A candidate gene for developmental dyslexia encodes a nuclear tetratricopeptide repeat domain protein dynamically regulated in brain. *Proc Natl Acad Sci USA* 100:11553–11558
- Thompson ME (1982) Assessment of children with specific reading difficulties using the British Ability Scales. *Brit J Psychol* 73:461–478
- 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 (2003) Linkage disequilibrium mapping provides further evidence of a gene for reading disability on chromosome 6p21.3-22. *Mol Psychiatry* 8:176–185
- Wigg KG, Couto JM, Feng Y, Anderson B, Cate-Carter TD, Macciardi F, Tannock R, Lovett MW, Humphries TW, Barr CL (2004) Support for EKN1 as the susceptibility locus for dyslexia on 15q21. *Mol Psychiatry* (in press)
- Zhang Y, Leaves NI, Anderson GG, Ponting CP, Broxholme J, Holt R, Edser P, Bhattacharyya S, Dunham A, Adcock IM, Pulleyn L, Barnes PJ, Harper JI, Abecasis G, Cardon L, White M, Burton J, Matthews L, Mott R, Ross M, Cox R, Moffatt ME, Cookson WO (2003) Positional cloning of a quantitative trait locus on chromosome 13q14 that influences immunoglobulin E levels and asthma. *Nat Genet* 34:181–186