Am. J. Hum. Genet. 64:146–156, 1999

A Quantitative-Trait Locus on Chromosome 6p Influences Different Aspects of Developmental Dyslexia

Simon E. Fisher,[1] Angela J. Marlow,[1] Janine Lamb,[1] Elena Maestrini,[1] Dianne F. Williams,[1] Alex J. Richardson,[2] Daniel E. Weeks,[1,3] John F. Stein,[2] and Anthony P. Monaco[1]

[1]Wellcome Trust Centre for Human Genetics and [2]Department of Physiology, University of Oxford, Oxford; and [3]Department of Human Genetics, University of Pittsburgh, Pittsburgh

Summary

Recent application of nonparametric-linkage analysis to reading disability has implicated a putative quantitative-trait locus (QTL) on the short arm of chromosome 6. In the present study, we use QTL methods to evaluate linkage to the 6p25-21.3 region in a sample of 181 sib pairs from 82 nuclear families that were selected on the basis of a dyslexic proband. We have assessed linkage directly for several quantitative measures that should correlate with different components of the phenotype, rather than using a single composite measure or employing categorical definitions of subtypes. Our measures include the traditional IQ/reading discrepancy score, as well as tests of word recognition, irregular-word reading, and nonword reading. Pointwise analysis by means of sib-pair trait differences suggests the presence, in 6p21.3, of a QTL influencing multiple components of dyslexia, in particular the reading of irregular words ($P = .0016$) and nonwords ($P = .0024$). A complementary statistical approach involving estimation of variance components supports these findings (irregular words, $P = .007$; nonwords, $P = .0004$). Multipoint analyses place the QTL within the *D6S422-D6S291* interval, with a peak around markers *D6S276* and *D6S105* consistently identified by approaches based on trait differences (irregular words, $P = .00035$; nonwords, $P = .0035$) and variance components (irregular words, $P = .007$; nonwords, $P = .0038$). Our findings indicate that the QTL affects both phonological and orthographic skills and is not specific to phoneme awareness, as has been previously suggested. Further studies will be necessary to obtain a more precise localization of this QTL, which may lead to the isolation of one of the genes involved in developmental dyslexia.

Received July 1, 1998; accepted for publication November 9, 1998; electronically published December 29, 1998.

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

0002-9297/99/6401-0020$02.00

Introduction

Children affected with developmental dyslexia (also known as "specific reading disability") have difficulty learning to read and spell, despite adequate intelligence and educational opportunity and in the absence of any profound sensory or neurological impairment (Smith et al. 1996). The prevalence of developmental dyslexia is 5%–10% of school-age children, making it the most common of the childhood learning disorders (Brown 1978; Shaywitz et al. 1990). Reading deficits can have a major influence on the cognitive, emotional, and social development of affected individuals.

Familial aggregation of developmental dyslexia has been well documented for nearly a century (Thomas 1905), and numerous segregation and twin studies have consistently supported a significant role for genetic factors in the etiology of this disorder (Hallgren 1950; Lewitter et al. 1980; DeFries et al. 1987; Pennington et al. 1991; DeFries and Gillis 1993; Lubs et al. 1993; DeFries and Alarcón 1996). Recent results of a large Colorado-based study suggest probandwise concordance rates of 68% in MZ twins, versus 38% in DZ twins (DeFries and Alarcón 1996). Using regression analysis of a composite discriminant measure of reading ability in twins, the Colorado group have estimated a heritability of >50% (DeFries et al. 1987; DeFries and Gillis 1993). However, these investigations have also indicated that any genetic basis for this disorder is likely to be complex and to involve reduced penetrance, phenocopies, heterogeneity, and oligogenic inheritance (Lewitter et al. 1980; Pennington et al. 1991; Lubs et al. 1993). In addition, the precise definition of the cognitive phenotype is crucial in the establishment of a firm basis for segregation and linkage studies (Pennington 1997). With dyslexia, as with other complex traits, there may be substantial phenotypic variability among subjects designated as affected, and there is some dispute about the nature of the core deficit of the disorder.

As a consequence of the phenotypic and genetic complexity of developmental dyslexia, there has been relatively limited success in the identification of susceptibility

loci by use of conventional parametric linkage analysis (Smith et al. 1983; Bisgaard et al. 1987; Froster et al. 1993; Rabin et al. 1993). Two studies have overcome these limitations by using alternative strategies. Cardon et al. (1994, 1995) targeted the human leukocyte antigen (HLA) region of chromosome 6, on the basis of observations of elevated risk for autoimmune disease in relatives of dyslexic probands (Pennington et al. 1987; Hugdahl et al. 1990). Using an extension of the DeFries-Fulker regression technique with a composite discriminant score of reading disability (DeFries and Fulker 1985), they examined allele sharing among the Colorado twin pairs and in sib pairs derived from an independent kindred sample. Their study suggested that a putative quantitative-trait locus (QTL) involved in reading disability maps to 6p21.3 ($P = .0094$ in twins, $P = .042$ in the kindred sample; Cardon et al. 1995).

Grigorenko et al. (1997) focused on 6p and also on a chromosome 15 region for which disputed linkage had been reported previously (Smith et al. 1983; Bisgaard et al. 1987; Rabin et al. 1993). They fractionated the overall dyslexia deficit into several partly overlapping, but partly distinct phenotypic definitions and studied six extended families, using both parametric and nonparametric techniques. Their nonparametric analysis strongly supported the existence of a 6p locus influencing dyslexia; the highest significance ($P < 10^{-6}$) was found with a phoneme-awareness phenotype, whereas the weakest results came from analysis of single-word reading. By contrast, their parametric investigation of chromosome 15 identified linkage only with single-word reading, although evidence for this was less robust than that for linkage to the chromosome 6 locus. On the basis of these results, they proposed that loci on chromosomes 6 and 15 may contribute to distinct components of the dyslexia phenotype. Like previous dyslexia linkage studies, the investigation by Grigorenko et al. (1997) used categorical definitions of deficit, but their analysis illustrates the value that a theoretically driven dissection of a complex phenotype has for identification of genetic risk factors.

In the present study, we describe an assessment of chromosome 6 linkage, using a strategy that combines and extends the approaches adopted by Cardon et al. (1994) and by Grigorenko et al. (1997). We have employed QTL methods, in order to analyze several different quantitative measures of reading deficit, which should correlate with different components of the dyslexic phenotype, in a large sample of affected-sib-pair families from the United Kingdom. Our measures include the traditionally defined IQ/reading discrepancy score, as well as tests of word recognition, irregular-word reading, and nonword reading. We have evaluated pointwise and multipoint linkage to 6p markers, using two complementary statistical methods, the first of which is based on sib-pair trait differences and the second of which involves maximum-likelihood estimation of variance components (VC).

Our data suggest that a QTL in 6p21.3 influences both phonological and orthographic processes involved in the development of dyslexia. The 6p region implicated by our study coincides with that identified in previous reports. Furthermore, our conclusions are highly consistent with those of an independent study of a new sample of twin pairs from the Colorado reading project (Gayán et al. 1999 [in this issue]). We conclude that the use of multipoint QTL linkage analysis on measures of different aspects of the dyslexia phenotype is a powerful tool for the identification of loci contributing to reading disability.

Families and Methods

Identification of Families

Families were identified through the dyslexia clinic at the Royal Berkshire Hospital, in Reading, United Kingdom. Probands were selected whose British Ability Scales (BAS) reading *t*-scores were >2 SDs below that predicted on the basis of their BAS Similarities (verbal reasoning) or BAS Matrices (nonverbal reasoning) scores (Rutter and Yule 1975; Elliot et al. 1979; Thompson 1982). Families were chosen for inclusion in the study if there was evidence (on the basis of parental reports or school history) of reading disability in one or more sibs of the proband. Eighty-two families were identified via this procedure; many of these included more than two sibs (see table 1); thus, the total number of sib pairs was substantially higher (125 independent pairs; 181 total pairs). All sibs, irrespective of their reading abilities, were assessed by use of a series of psychometric tests, as outlined below. None of the tests were used on parents; only quantitative data from sibs were needed for the methods of linkage analysis employed in this study.

Phenotypic Measures

Four quantitative phenotypes were measured for this study: word recognition, IQ-reading discrepancy, orthographic coding, and phonological decoding (see table 2). Word recognition (WR) was assessed by use of the standardized BAS test of single-word reading (Elliot et al. 1979). The discrepancy criterion (SRdisc), a traditional measure used to diagnose dyslexia, was obtained by subtracting the WR *t*-score from the score obtained in the BAS Similarities subtest, which is an oral measure of verbal reasoning (Rutter and Yule 1975; Elliot et al. 1979; Thompson 1982). Orthographic coding ability was evaluated by use of the Castles and Coltheart (CC) irregular-word reading test (CCirr), in which subjects must read aloud a series of irregular words, such as "meringue" and "yacht," which violate the standard let-

Table 1

Numbers of Families/Sib Pairs for Different Traits

Category	No. Genotyped	No. Tested by			
		CCirr	CCnon	CCcom	WR/SRdisc[a]
Family size:					
Two sibs (50)	50	50	51	50	40
Three sibs (23)	23	20	21	22	14
Four sibs (7)	7	7	6	5	3
Five sibs (2)	2	2	2	2	0
Total	82	79	80	79	57
Total no. of sib pairs:					
Completely independent[b]	125	119	119	117	77
All[c]	181	172	170	166	100

Note.—All parents/available sibs were genotyped for the study, but measures for all traits could not always be obtained from every sib. Therefore, samples analyzed for each trait are subsets of the total sample that was genotyped. This means that the size of a family varies depending on the trait under investigation; for example, for a particular trait, a family with three genotyped sibs may have phenotypic information from only two of the sibs—and thus, for that measure, would be considered as a two-sib family.

[a] Samples sizes for WR and for SRdisc were identical.

[b] A family with n sibs contributes $(n - 1)$ independent pairs.

[c] A family with n sibs contributes $[n(n - 1)/2]$ possible pairs.

ter-sound conventions of English (Castles and Coltheart 1993). Accurate reading of these irregular words cannot be achieved via the use of grapheme-phonological conversion rules but, instead, requires the recognition of a word-specific orthographic representation, followed by retrieval of the appropriate phonological form. The CC nonword reading test (CCnon), in which subjects are asked to read aloud a series of pronounceable nonwords such as "gop" and "seldent," was used to assess phonological-decoding ability (Castles and Coltheart 1993). Since these nonwords will be unfamiliar to the subject, success in this task depends on the use of the appropriate grapheme-phoneme correspondence rules. CC scores were adjusted for age, on the basis of population norms (Coltheart and Leahy 1996). In addition, a fifth measure, a composite of the CC scores (CCcom), was obtained by summation of the age-adjusted CCirr and CCnon scores. Sample sizes for the different phenotypes are given in tables 1 and 2. The BAS tests (WR and SRdisc) are standardized only for younger individuals, so sibs >16 years of age were not assessed with these measures. Thus, the sample sizes for the BAS traits were smaller than those obtained for the other traits. The psychometric test battery outlined in the preceding paragraph did not include a measure of phoneme awareness.

Semiautomated Genotyping

Sibs and both parents from each family were genotyped with 15 highly polymorphic markers from the 6p25-6p21.3 region (fig. 1). PCR primers were labeled with 6-FAM, HEX, or TET phosphoramidite (Applied Biosystems). PCR reactions were performed in 96-well Costar (THERMOWELL) plates in a 15-μl volume, with 40 ng template genomic DNA, on MJ research PTC-225 thermocyclers. Pooled products were run on a 373A sequencer (Applied Biosystems), and the results were analysed by means of GENESCAN (version 2.0) and GENOTYPER (version 1.1) software, to derive allele sizes (Reed et al. 1994).

Data Handling and Error Checking

Raw allele-size data were checked for consistent inheritance and were converted to LINKAGE format with the GAS software package, version 2.0 (A. Young). The GENBASE system, version 2.0.5 (J.-M. Sebaoun and M. Lathrop), was used for storage of genotypic and phenotypic data and for creation of the appropriate files for statistical analysis. As a further verification of genotyp-

Table 2

Distributions of Quantitative Measures in Sib Samples

Measure (Sample Size)	Mean	Variance	Minimum	Maximum
WR[a] (149)	44.9	79.8	27	73.0
SRdisc[b] (149)	15.7	93.9	−8	36.0
CCirr[c] (202)	−7.5	37.0	−29	2.6
CCnon[c] (202)	−6.0	34.8	−24	6.7
CCcom[c] (200)	−13.5	117.8	−53	9.3

Note.—Data are from all sibs tested.

[a] Data are standardized t-scores, with a normal population mean of 50; lower scores indicate greater deficit.

[b] Discrepancy between BAS verbal reasoning and BAS WR, with a normal population mean of 0; higher discrepancy scores indicate greater deficit.

[c] Residual scores after age adjustment (normal population mean is 0); the more negative the score, the greater the deficit.

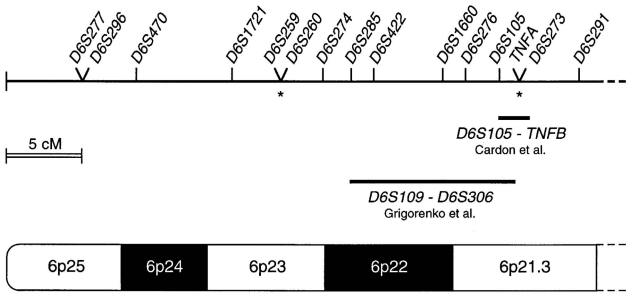

Figure 1 Linkage map of markers used in the present study, derived from other reports (Martin et al. 1995; Dib et al. 1996; Feder at al. 1996; Malfroy et al. 1997) and verified in our sample. The chromosomal bands are aligned beneath the map. The approximate regions implicated by two previous studies of 6p in developmental dyslexia are indicated. Asterisks (*) denote marker intervals that were estimated on the basis of our sample only.

ing quality, marker haplotypes were generated from the data with the SIMWALK2 computer program (Sobel and Lange 1996), to identify any chromosomes showing an excessive number of recombination events.

Marker Map

Since multipoint analysis programs were used in this study, it was important to establish an accurate and reliable marker map for the 6p region (fig. 1). The most likely order of markers was derived from the latest linkage and physical maps of 6p (Martin et al. 1995; Dib et al. 1996; Feder at al. 1996; Malfroy et al. 1997). The 6p21.3 region just distal of the HLA complex is well characterized at the physical level, since it has been studied intensively in recent years by groups investigating hereditary hemochromatosis (Feder at al. 1996). An independent prediction of marker order on the basis of our own data was entirely consistent with this consensus map. Intermarker distances were derived from recent Généthon data (Dib et al. 1996) and from studies of recombination around the HLA region (Martin et al. 1995; Malfroy et al. 1997) and were confirmed by use of data from our own family sample.

Statistical Analysis Based on Sib-Pair Trait Differences

One approach to investigation of linkage is to determine the difference in trait values for each sib pair and to test whether the variance of this difference becomes smaller as identity-by-descent (IBD) sharing increases (i.e., whether $\sigma_0^2 > \sigma_1^2 > \sigma_2^2$, where σ_i^2 represents the variance of the difference when i alleles are shared IBD; Kruglyak and Lander 1995). For pointwise analyses, we used SIBPAL (version 2.6) to perform traditional Haseman-Elston (H-E) regression of sib-pair squared trait differences against estimated IBD sharing for each genetic marker (Haseman and Elston 1972). This program accommodates multiple sibships by employing reduced df, based on the number of sibs (n) minus one (i.e., $n - 1$) in each sibship, summed over all the families (Wilson and Elston 1993). For multipoint analyses, the complete probability distribution of IBD status across the 6p25–6p21.3 region was inferred from all the genotypic information, with the computer package MAPMAKER/SIBS (version 2.0; Kruglyak and Lander 1995). Multipoint linkage could then be investigated by means of three methods: H-E regression; direct maximum-likelihood estimation of σ_0^2, σ_1^2, and σ_2^2 (MLvar); and a nonparametric technique involving the ranking of sib pairs on the basis of phenotypic difference (NP).

Weighting Schemes for MAPMAKER/SIBS Analysis

MAPMAKER/SIBS offers two alternative weighting schemes, neither of which is ideal, for analysis of all possible sib pairs from a multiple sibship. One method weights linkage evidence from all pairings equally, regardless of the size of the sibship from which these pairings are formed. However, this method may lead to a false inflation of significance, as a result of statistical dependence between pairs (Hodge 1984). The other method, termed "strict," weights pairs from a family of n sibs by a factor of $(2/n)$. Such a weighting scheme makes the implicit assumption that the linkage information obtained from fully independent sib pairings of a multiple sibship will determine completely the linkage information obtained from the remaining possible pairings, which is not necessarily true in the case of QTL mapping. For example, the three possible sib pairs formed from a trio are treated as though they are equivalent to only two pairs under strict weighting, even though the third pairing may provide some additional independent linkage information. Therefore, strict weighting sometimes can be overconservative. In the present study, we have analyzed families by each of these weighting schemes, and we present here the results from both. Note that MAPMAKER/SIBS also offers the option of using either the first pair only or independent pairs only. However, these options do not make full use of the genotypic and phenotypic information from each family, and they give results that depend on the order in which the sibs are specified; thus, these options were not used for the present study.

Statistical Analysis Based on VC

Evidence for linkage was also assessed by use of a sibling-based VC approach, as implemented in the program MULTIC (version 2.3.2) from the ACT package (Amos 1994). This method involves maximum-likelihood estimations of major-gene (σ_a^2), polygenic (σ_G^2), and environmental (σ_e^2) VC contributing to trait variability

between siblings, for each locus under investigation. The likelihood under the assumption that there is a major-gene effect due to this locus is compared with the likelihood under the null hypothesis that there is no major gene effect (i.e., when σ_a^2 is constrained to be 0), and significance can be estimated by means of standard χ^2 tests. This approach treats each family as a unit, explicitly allowing for statistical nonindependence among sibs, and is less liable to type I error than are statistics based on pairs of relatives (Wijsman and Amos 1997). The IBD probability distribution obtained from MAPMAKER/SIBS (see the preceding discussion) was used to perform multipoint VC analysis with MULTIC. For all the statistical analyses, we have followed the recommendation of Elston (1997), by reporting precise *P* values without adjustment for multiple comparisons, so that they can be properly interpreted by the reader. Note that a Bonferroni correction would assume that all tests are independent, which is not the case here, because some of our results are determined on the basis of correlated phenotypes (in particular, CCcom is the sum of CCirr and CCnon). Therefore, such a correction would be too conservative.

Results

QTL Analysis Using Sib-Pair Trait Differences

Pointwise H-E analysis with the different phenotypic measures (table 3) suggested the presence, in the *D6S285-D6S291* region, of a QTL influencing the CC scores. The most compelling evidence for this was found for CCirr, with all markers in the *D6S1660-D6S291* interval giving *P* values <.026 (table 3). For each of the three CC measures, the highest significance was observed for the same marker, *D6S276* (CCirr, $P = .0016$; CCnon, $P = .0024$; CCcom, $P = .0017$). WR gave some suggestion of linkage, with *P* values of .009 and .047 for the markers *D6S285* and *D6S276*, respectively. The weakest results were found for SRdisc, with all *P* values >.05, although *D6S285* ($P = .054$) and *D6S276* ($P = .07$) did approach this significance level.

Multipoint analysis with MAPMAKER/SIBS was used to facilitate a more precise localization of the putative QTL (fig. 2). As was expected on the basis of the pointwise results, multipoint analyses with each of the CC measures supported a putative QTL in the *D6S285-D6S291* region, with a peak in the region of markers *D6S276* and *D6S105*. This was consistent for all the different methods of quantitative-trait analysis (H-E, MLvar, and NP) implemented in MAPMAKER/SIBS and for both weighting schemes. Figure 2 shows the LOD-score curves obtained by use of MLvar, which gave the most significant results (for a summary of all methods, see table 4). Although the use of strict weighting results in lower LOD scores, the shapes of the curves are very similar to those found when there is no weighting. There is considerable overlap between the region implicated by CCirr and that implicated by CCnon, with the former giving higher LOD scores. The position of the peak for CCirr is at marker *D6S105*, within 2 cM of the peak, at *D6S276*, for CCnon. Furthermore, the CCcom peak is also at *D6S105*, and the results are more significant than those obtained for CCirr alone (fig. 2).

For WR and SRdisc, a main peak in the *D6S276-D6S105* region was identified consistently by the different types of MAPMAKER/SIBS analysis (and by both weighting schemes), but, in general, significance levels were much lower. In addition, residual peaks in the more

Table 3

Pointwise Results of Haseman-Elston Regression Analysis, by SIBPAL, with Different Traits

	P (df)[b]				
Marker (Location [cM][a])	WR	SRdisc	CCirr	CCnon	CCcom
D6S259 (18.0)	NS (71)	NS (69)	.047 (115)	NS (115)	NS (113)
D6S260 (18.0)	NS (71)	NS (69)	.041 (113)	.041 (113)	.044 (111)
D6S274 (20.8)	NS (65)	NS (63)	NS (101)	NS (100)	NS (99)
D6S285 (22.6)	.009 (69)	NS (68)	.029 (113)	.0097 (113)	.0089 (111)
D6S422 (24.1)	NS (68)	NS (66)	NS (102)	NS (101)	.043 (100)
D6S1660 (28.6)	NS (72)	NS (70)	.0037 (114)	.024 (114)	.0055 (112)
D6S276 (30.1)	.047 (73)	NS (71)	.0016 (115)	.0024 (115)	.0017 (113)
D6S105 (32.3)	NS (71)	NS (69)	.0076 (115)	NS (115)	.033 (113)
TNFA (33.6)	NS (70)	NS (68)	.026 (110)	NS (110)	.032 (108)
D6S273 (33.6)	NS (70)	NS (69)	.0018 (107)	.048 (107)	.011 (105)
D6S291 (37.5)	NS (71)	NS (69)	.016 (111)	NS (111)	NS (109)

[a] Distance from 6pter.

[b] The test statistic is the estimated regression coefficient divided by its standard error, modified to allow for nonindependence of sib pairs. The df for the test are based on the effective sample size. Significance of the test is assessed by use of a *t*-distribution; NS = not significant (i.e., $P > .05$).

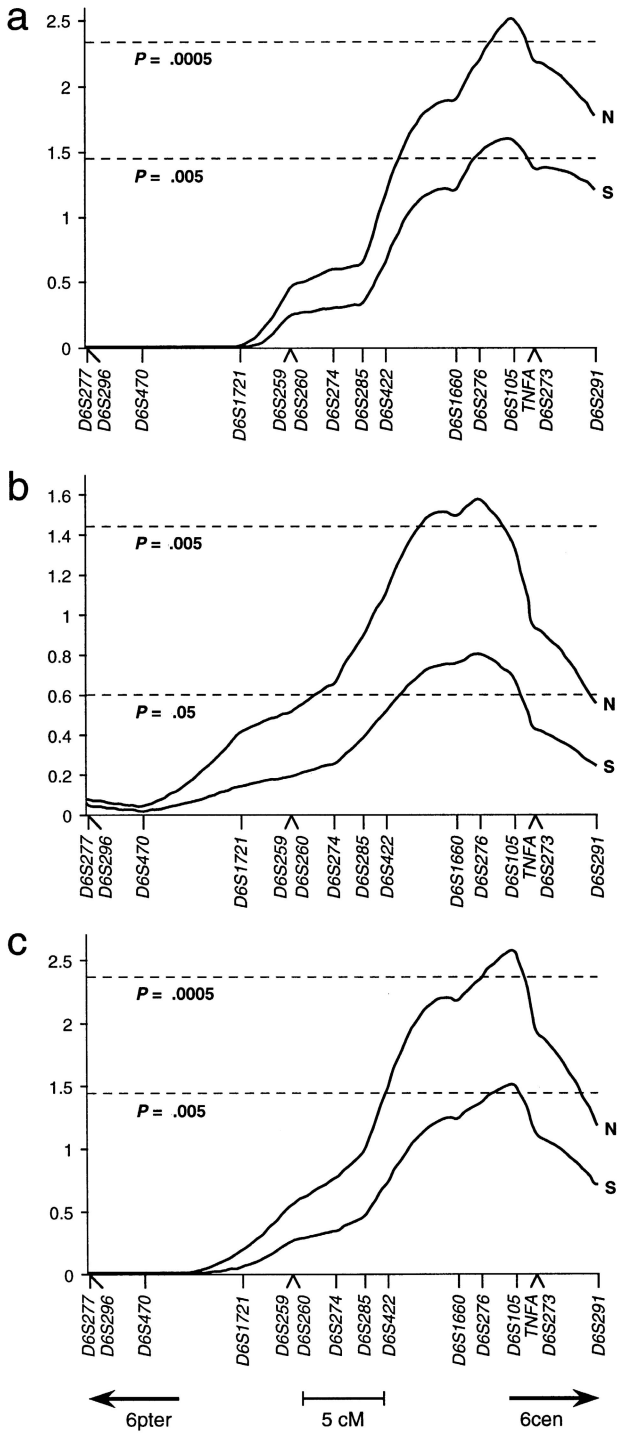

Figure 2 Multipoint sib-pair trait-difference analysis for CCirr (*a*), CCnon (*b*), and CCcom (*c*). The graphs show LOD-score curves obtained by use of the MLvar method (if no dominance variance is assumed) in MAPMAKER/SIBS, with use of strict weighting (S) or of no weighting (N). The results were almost identical when the analysis was done without the assumption of no dominance variance. Broken lines indicate LOD scores corresponding to significance levels (P = .05, P = .005, and P = .0005). The orientation of markers relative to chromosome 6 is given.

distal parts of 6p were almost as significant as the main peak. All the P values for WR and SRdisc that were obtained with the H-E and MLvar options were >.05, but the NP analysis of the WR phenotype was more convincing (P = .03 with strict weighting, P = .0064 without weighting)

QTL Analysis Using VC

Pointwise sibling-based VC analysis (table 5) supported the findings of the trait-differences approach, again indicating that a QTL on 6p influences performance in the CC tests but that it has much less effect on the WR and SRdisc measures in the families that we studied. (However, as with the pointwise H-E analysis, there was weak evidence for linkage of WR to *D6S285:* P = .021.) When VC were used, the most significant results were observed for CCnon and CCcom, with all markers in the *D6S260-D6S276* region giving pointwise P values <.029 (table 5). The highest significance for each CC measure was found with the same marker, *D6S276* (CCirr, P = .007; CCnon, P = .0004; CCcom, P = .0010), in agreement with the H-E regression.

As shown in figure 3, multipoint VC analysis of the CC measures gave likelihood-ratio test (LRT) curves that were similar in shape to the LOD curves obtained by means of MAPMAKER/SIBS options (compare fig. 2). Furthermore, the positions of peaks indicated by the use of VC approach were identical to those identified by the trait-differences analyses (table 4). WR and SRdisc gave no evidence for linkage to this region, on the basis of multipoint VC analysis.

Discussion

Our study of quantitative measures of reading deficit in affected-sib-pair families suggests that a locus on 6p21.3 influences components of developmental dyslexia. This conclusion is supported by two independent complementary statistical approaches, one done on the basis of sib-pair trait differences and the other involving estimation of VC. The most likely location for this QTL, as indicated by multipoint analysis, is in the vicinity of markers *D6S276* and *D6S105,* which are ~2 cM-apart and are just distal to the HLA region. The 1-LOD-unit support interval identified by the most significant results (MLvar analysis of CCcom) spans ~11 cM in the *D6S422-D6S291* interval (see fig. 2). This position of the putative QTL is remarkably consistent with that suggested by other studies of the role of chromosome 6p in developmental dyslexia (Cardon et al. 1994; Grigorenko et al. 1997; Gayán et al. 1999). The study by Cardon and colleagues placed the QTL in the *D6S105-TNFB* region, on the basis of an interval-mapping procedure (Cardon et al. 1994), and recent multipoint analysis of

Table 4

Summary of Peaks Identified using Multipoint Analysis of CC Measures

Approach and Program	*P* (Location [a] [Marker[b]])		
	CCirr	CCnon	CCcom
Trait differences:[c]			
H-E (S)	.0066 (32.3 [*D6S105*])	.035 (30.1 [*D6S276*])	.01 (32.3 [*D6S105*])
H-E (N)	.00062 (32.3 [*D6S105*])	.007 (30.1 [*D6S276*])	.0013 (32.3 [*D6S105*])
MLvar (S)	.0035 (32.3 [*D6S105*])	.027 (30.1 [*D6S276*])	.0044 (32.3 [*D6S105*])
MLvar (N)	.00035 (32.3 [*D6S105*])	.0035 (30.1 [*D6S276*])	.00030 (32.3 [*D6S105*])
NP (N)	.018 (31.8 [*D6S105*])	.0086 (30.1 [*D6S276*])	.015 (32.3 [*D6S105*)
VC:[d]			
MULTIC	.007 (32.3 [*D6S105*])	.0038 (30.1 [*D6S276*])	.0036 (32.3 [*D6S105*])

[a] Distance (in cM) from 6pter.

[b] Nearest to peak.

[c] (S) = strict weighting; (N) = no weighting. Peak LOD scores for H-E and MLvar were converted into *P* values by multiplication by $2\log_e 10$ and subsequent determination of significance by use of χ^2 tables; the one-sided nature of the linkage test was taken into account (Lander and Kruglyak 1995). Significance for the peak Z scores for NP was determined by use of a standard normal distribution. With use of the NP(S) option, $P > .05$ but implicated the same region as was implied by the other methods of analysis.

[d] Peak LRT scores were converted into *P* values by use of χ^2 tables. These values take into account the one-sided nature of the linkage test.

a new set of Colorado twin pairs has identified peak LOD scores between *D6S276* and *D6S105* (Gayán et al. 1999). Using allele-sharing methods, Grigorenko et al. (1997) mapped the chromosome 6 susceptibility locus to *D6S109-D6S306,* a region that shows substantial overlap with the *D6S422-D6S291* interval implicated by our study (see fig. 1). Their multipoint analysis gave only weak evidence of linkage to *D6S276* and *D6S105,* which they placed on the centromeric side of the *D6S109-D6S306* interval; thus, they argued for a QTL location that is more telomeric than that suggested by Cardon et al. (1994). However, it should be noted that the marker order used by Grigorenko et al. (1997) disagrees with the consensus map derived from recent linkage and physical mapping studies, which unambiguously place *D6S276* and *D6S105* on the telomeric side of *D6S306,* within the *D6S109-D6S306* interval (Feder at al. 1996; Malfroy et al. 1997). Therefore, reevaluation of the Grigorenko et al. (1997) data with this consensus map order may provide stronger evidence for linkage to *D6S276* and *D6S105* in the pedigrees that they studied.

A recent report by Field and Kaplan (1998) failed to find significant evidence of linkage to the 6p23-p21.3 region in 79 families with at least two siblings affected with dyslexia. However, their study used pointwise parametric and nonparametric linkage analyses of a single dichotomous qualitative phenotype, on the basis of impaired phonological-coding skills, in contrast to the quantitative-trait approaches adopted here and in the studies by Cardon et al. (1994, 1995) and Gayán et al. (1999). Therefore, it remains to be determined whether these results represent a nonreplication of the positive QTL-based findings.

For our study, we identified families on the basis of a proband defined as dyslexic (by traditional, strict diagnostic criteria), with one or more sibs also showing evidence of reading disability, but we made no adjustment for our method of selection. When the VC approach is used, lack of correction for ascertainment tends to lead to loss of power rather than to inflation of the rate of type I error (C. Amos, personal communication). Furthermore, a recent QTL analysis of a simulated data set by means of the MAPMAKER/SIBS package, using an affected-sib-pair ascertainment scheme with no adjustment for ascertainment bias, found that false-positive results verifiable across all three alternative linkage statistics did not occur (Iyengar et al. 1997). Therefore, the observation of a consistent peak for several different analysis programs (H-E, MLvar, NP, and MULTIC) and across multiple studies indicates that the 6p21.3 QTL is unlikely to be a false positive.

In our sample of families, we have demonstrated that the QTL influences performance in tests of nonword and irregular-word reading, which suggests that this locus affects both phonological decoding and orthographic coding. Much weaker support for linkage to this region was found with the standard measures of word recognition or IQ-reading discrepancy. A recent study has found that the measured IQ of children with reading disabilities is related strongly to socioeconomic status and tends to decrease with increasing age, suggesting that, as an index of dyslexia, the traditional IQ/achievement discrepancy may be flawed (Siegel and Himel 1998). In the present study, qualitative nonparametric analysis using a dichotomous classification of affection status based on IQ-reading discrepancy gave no evidence for linkage (data not shown).

Our study agrees well with the recent report by Gayán

Table 5

Pointwise Results of VC Analysis, by MULTIC, with Different Traits

Marker (Location[a])	P^b (LRT)				
	WR	SRdisc	CCirr	CCnon	CCcom
D6S259 (18.0)	NS (.30)	NS (.01)	NS (1.67)	.04 (3.06)	NS (2.33)
D6S260 (18.0)	NS (1.90)	NS (.25)	.046 (2.85)	.007 (6.16)	.017 (4.53)
D6S274 (20.8)	NS (.62)	NS (0)	NS (2.49)	.017 (4.54)	.029 (3.57)
D6S285 (22.6)	.021 (4.11)	NS (.42)	.008 (5.82)	.0012 (9.18)	.0012 (9.19)
D6S422 (24.1)	NS (1.28)	NS (0)	NS (1.97)	.022 (4.02)	.017 (4.5)
D6S1660 (28.6)	NS (.68)	NS (.27)	.02 (4.22)	.015 (4.69)	.014 (4.77)
D6S276 (30.1)	NS (1.77)	NS (.42)	.007 (6.03)	.0004 (11.28)	.001 (9.54)
D6S105 (32.3)	NS (1.91)	NS (.11)	.036 (3.22)	NS (1.59)	.043 (2.94)
TNFA (33.6)	NS (.19)	NS (.02)	NS (1.26)	NS (1.87)	NS (2.25)
D6S273 (33.6)	NS (.10)	NS (.41)	.02 (4.2)	.042 (2.97)	.038 (3.13)
D6S291 (37.5)	NS (.14)	NS (0)	.045 (2.89)	NS (1.4)	NS (1.46)

[a] Distance (in cM) from 6pter.

[b] Determined on the basis of χ^2 tables (1 df), as in table 4. NS = not significant (i.e., $P > .05$).

et al. (1999), whose investigation of DZ twins with several psychometric measures similarly supports a role for the 6p QTL in both phonological and orthographic skills. They also found significant effects on phoneme awareness (defined as the ability to reflect explicitly on the individual speech sounds that make up a word), but we are still in the process of gathering phenotypic information relating to this trait in our set of families.

Using regression analysis on a large sample of the Colorado twin pairs, Olson et al. (1994) have demonstrated substantial heritability for phonological decoding ($h_g^2 = .59 \pm .12$) and for orthographic coding ($h_g^2 = .56 \pm .13$), which indicates a significant genetic etiology for deficits in each of these skills. Performance on phonological tests and performance on orthographic tests usually is found to be moderately correlated; in our study, for example, $r = .64$ for CCnon and CCirr. It is important to note that the observation of similar heritabilities for these different correlated components of reading disability does not necessarily imply that there is a common genetic influence. However, Olson et al. (1994) have also performed bivariate genetic analyses, which indicate that a large proportion (~66%) of the correlated variance between deficits in phonological and orthographic processing is due to the same genetic factors. Thus, the heritability studies would predict the presence of a locus (or loci) affecting both of these components of reading disability, as has been identified in the present study.

Grigorenko et al. (1997) have proposed that their analysis of extended pedigrees with five different categorical reading-related phenotypes could indicate a dissociation of genetic effects, such that phoneme awareness maps to chromosome 6 whereas single-word reading maps to chromosome 15. However, there are several points that should be considered when their results are compared with ours. First, both we and they find effects of the chromosome 6 locus on multiple aspects of reading disability; although Grigorenko et al.'s (1997) strongest result was for phoneme awareness, their investigation did in fact yield significant linkage to 6p21.3 for all phenotypes tested, including phonological decoding, rapid automized naming, single-word reading, and IQ-reading discrepancy. (They did not include a test of orthographic coding.) Second, Grigorenko et al. (1997) observed reduced evidence of chromosome 6 linkage for phenotypes such as single-word reading, as compared with phoneme awareness. However, inferences about the relative effects of a locus on different phenotypes should be made cautiously, because significance levels may differ as a consequence of variation of methodological factors, such as sample size and the reliability of phenotypic testing. In the Grigorenko et al. (1997) study, linkage to chromosome 6 was found by use of the affected-pedigree–member method, a nonparametric technique in which only those individuals classified as affected are included within the analysis (Weeks and Lange 1988). A substantial proportion of dyslexics show compensation in their single-word reading as they get older but still manifest deficits in phonological skills. Thus, their reduced significance for phenotypes such as single-word reading and IQ-reading discrepancy could partly reflect a loss of power, resulting from the smaller number of individuals classified as affected on the basis of these criteria. Last, Grigorenko et al. (1997) found linkage to chromosome 6 only with nonparametric techniques and to chromosome 15 only with traditional parametric analysis. Since alternative statistical approaches may respond differentially to the use of different phenotypic information for analysis of the same family sample, additional data are required, in order to allow a detailed comparison of the relative effects of different phenotypes at these two loci.

The issues highlighted here are also relevant to the data that we present in the current study. Our lack of evidence for linkage with WR and SRdisc may partly be

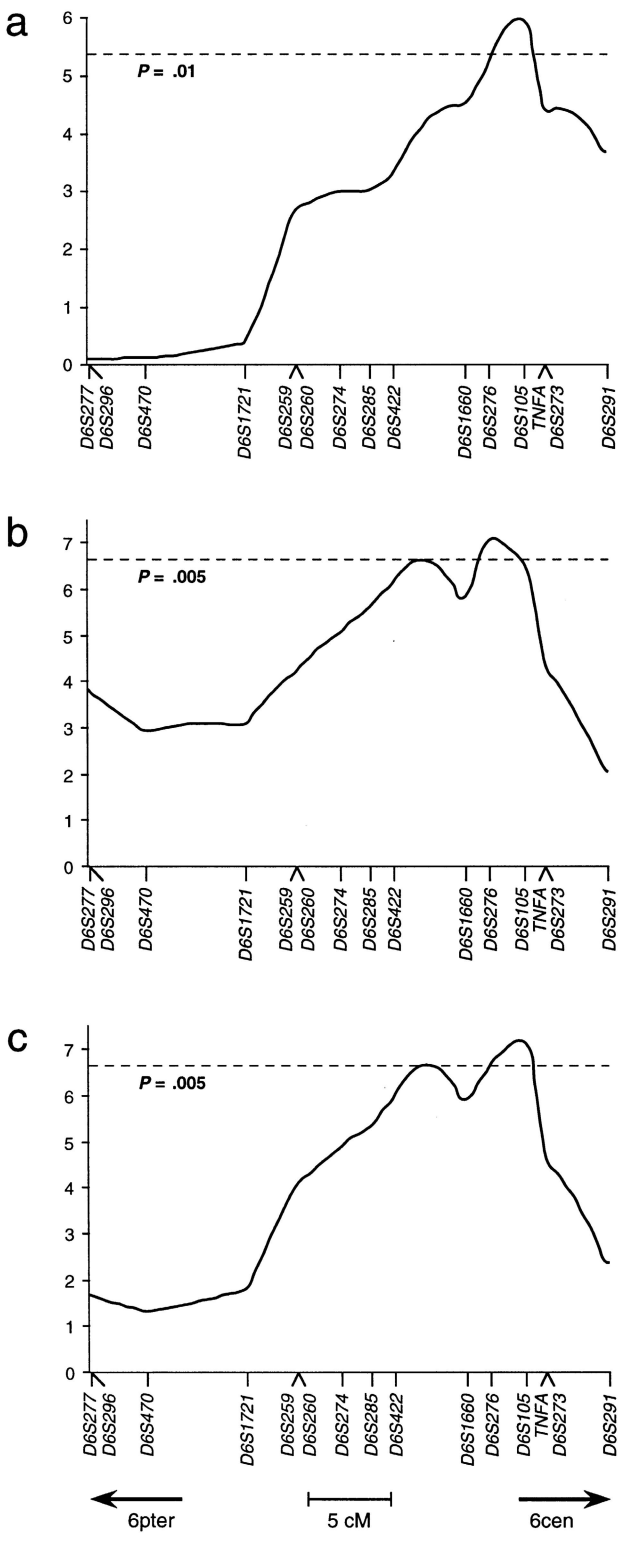

Figure 3 Multipoint VC analysis for CCirr (*a*), CCnon (*b*), and CCcom (*c*). The graphs show LRT curves obtained by use of MULTIC. Broken lines indicate LRT scores corresponding to significance levels ($P = .01$ and $P = .005$).

a result of reduced power, given the smaller sample size available for these measures (table 1). We are attempting to resolve this problem by assessing older sibs with tests of word recognition and intelligence that are standardized for adults. (However, it is possible that some of the older individuals may show compensation for the WR and SRdisc phenotypes.) Furthermore, it is interesting to note that sib-pair trait differences indicated a larger effect of the 6p QTL on CCirr than on CCnon, whereas the VC approach suggested the converse. It is therefore difficult, at this stage, to quantify the relative effects that this locus has on phonological-decoding and orthographic-coding skills. Nevertheless, both the high level of similarity between the shapes of the multipoint curves obtained by use of sib-pair trait differences versus those obtained by use of VC (compare figs. 2 and 3) and the consistent position of the peak for the different phenotypes constitute strong evidence that the 6p locus is affecting both aspects of reading performance. We currently are collecting additional phenotypic data from the families in the present study—including measures of phoneme awareness and spelling ability—to facilitate a more detailed investigation of the putative QTL.

We note that all linkage studies for dyslexia that have been reported thus far (including this one) have targeted specific chromosomal regions for investigation. Of these, the most consistently replicable locus is that on 6p21.3, but it is yet to be determined whether there are other QTLs that have similar or stronger effects in other parts of the genome. We are in the process of conducting a systematic genomewide search in our family sample, using highly polymorphic markers spaced at ~10-cM intervals. Such a genome scan, which uses the same families, sets of measures, and statistical approaches for all the genetic markers examined, will be important in addressing properly the question of how particular loci affect different components of the phenotype. Although the heritability studies by Olson et al. (1994, and in press) have supported the existence of genes, such as that on 6p, that contribute to the shared variance among word recognition, phoneme awareness, phonological decoding, and orthographic coding, they also have suggested the presence of independent genetic effects that influence separately the different components. Therefore, QTL analysis with several different measures appears to be the most appropriate approach for these investigations.

To conclude, there is now converging evidence from several independent studies that a QTL on 6p21.3 contributes to multiple components of developmental dyslexia. At this stage, it is difficult to evaluate rigorously the relative effects that the locus has on these different aspects of the phenotype. However, when investigating this locus, each study found a lower significance for word recognition and/or IQ-reading discrepancy criteria

than for measures of component skills, such as phonological decoding and orthographic coding. The observation that both phonological and orthographic measures appear to be affected by this locus suggests that the latter is involved in an underlying mechanism that is common to the development of both types of skills. Further analyses with larger sample sizes and powerful statistical methods will be necessary to confirm this observation and to refine the map position of this QTL. These investigations offer the hope that we eventually may be able to identify and study one of the genes involved in this complex behavioral trait.

Acknowledgments

We are very grateful to all the families who participated in this study. We thank Sue Fowler, Janet Walter, and Pam Southcott, for collection of blood samples and for testing of subjects; Diana Wylie, for conducting the DNA extractions; and Chris Amos and Meg Gelder-Ehm, for providing advice and software for the VC analyses. This research was funded by the Wellcome Trust. A.P.M. is a Wellcome Trust Principal Research Fellow. Some of the results presented here were obtained by use of the program package SAGE from the Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, which is supported by U.S. Public Health Service resource grant 1 P41 RR03655 from the National Center for Research Resources.

References

Amos CI (1994) Robust variance-components approach for assessing genetic linkage in pedigrees. Am J Hum Genet 54: 535–543

Bisgaard ML, Eiberg H, Moller N, Niebuhr E, Mohr J (1987) Dyslexia and chromosome 15 heteromorphism: negative LOD score in a Danish sample. Clin Genet 32:118–119

Brown BS (1978) In: Benton AL, Pearl D (eds) Dyslexia: an appraisal of current knowledge, Oxford University Press, Oxford, pp v–vii

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

Coltheart M, Leahy J (1996) Assessment of lexical and nonlexical reading abilities in children: some normative data. Aust J Psychol 48:136–140

DeFries JC, Alarcón M (1996) Genetics of specific reading disability. Ment Retard Dev Disabilities Res Rev 2:39–47

DeFries JC, Fulker DW (1985) Multiple regression analysis of twin data. Behav Genet 15:467–473

DeFries JC, Fulker DW, LaBuda MC (1987) Evidence for a genetic aetiology in reading disability of twins. Nature 329: 537–539

DeFries JC, Gillis JJ (1993) Genetics of reading disability. In: Plomin R, McClearn G (eds) Nature, nurture, and psychology. APA Press, Washington, DC, pp 121–145

Dib C, Faure S, Fizames C, Samson D, Drouot N, Vignal A, Millasseau P, et al (1996) A comprehensive genetic map of the human genome based on 5,264 microsatellites. Nature 380:152–154

Elliot CD, Murray DJ, Pearson LS (1979) The British ability scales. NFER, Slough, UK

Elston RC (1997) Algorithms and inferences: the challenge of multifactorial diseases. Am J Hum Genet 60:255–262

Feder JN, Gnirke A, Thomas W, Tsuchihashi Z, Ruddy DA, Basava A, Dormishian F, et al (1996) A novel MHC class I-like gene is mutated in patients with hereditary haemochromatosis. Nat Genet 13:399–408

Field LL, Kaplan BJ (1998) Absence of linkage of phonological coding dyslexia to chromosome 6p23-p21.3 in a large family data set. Am J Hum Genet 63:1448–1456

Froster U, Schulte-Körne G, Hebebrand J, Remschmidt H (1993) Cosegregation of balanced translocation (1;2) with retarded speech development and dyslexia. Lancet 342: 178–179

Gayán J, Smith SD, Cherny SS, Cardon LR, Fulker DW, Brower AM, Olson RK, et al (1999) Quantitative-trait locus for specific language and reading deficits on chromosome 6p. Am J Hum Genet 64:157–164 (in this issue)

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

Hallgren B (1950) Specific dyslexia (congenital word blindness): a clinical and genetic study. Acta Psychiatr Neurol 65: 2–289

Haseman JK, Elston RC (1972) The investigation of linkage between a quantitative trait and a marker locus. Behav Genet 2:3–19

Hodge SE (1984) The information contained in multiple sibling pairs. Genet Epidemiol 1:109–122

Hugdahl K, Synnevag B, Satz P (1990) Immune and autoimmune disorders in dyslexic children. Neuropsychologia 28: 673–679

Iyengar S, Calafell F, Kidd KK (1997) Detection of major genes underlying several quantitative traits associated with a common disease using different ascertainment schemes. Genet Epidemiol 14:809–814

Kruglyak L, Lander ES (1995) Complete multipoint sib-pair analysis of qualitative and quantitative traits. Am J Hum Genet 57:439–454

Lander E, Kruglyak L (1995) Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. Nat Genet 11:241–247

Lewitter FI, DeFries JC, Elston RC (1980) Genetic models of reading disabilities. Behav Genet 10:9–30

Lubs HA, Rabin M, Feldman E, Jallad BJ, Kushch A, Gross-Glenn K, Duara R, et al (1993) Familial dyslexia: genetic and medical findings in eleven three-generation families. Ann Dyslexia 43:44–60

Malfroy L, Roth MP, Carrington M, Borot N, Volz A, Ziegler A, Coppin H (1997) Heterogeneity in rates of recombination in the 6-Mb region telomeric to the human major histocompatibility complex. Genomics 43:226–231

Martin M, Mann D, Carrington M (1995) Recombination rates across the HLA complex: use of microsatellites as a rapid screen for recombinant chromosomes. Hum Mol Genet 4:423–428

Olson RK, Datta H, Gayán J, DeFries JC. A behavioral-genetic analysis of reading disabilities and component processes. In: Klein RM, McMullen PA (eds) Converging methods for understanding reading and dyslexia. MIT Press, Cambridge, MA (in press)

Olson RK, Forsberg H, Wise B (1994) Genes, environment, and the development of orthographic skills. In: Berninger VW (ed) The varieties of orthographic knowledge. I. Theoretical and developmental issues. Kluwer, Dordrecht, pp 27–71

Pennington BF (1997) Using genetics to dissect cognition. Am J Hum Genet 60:13–16

Pennington BF, Gilger JW, Pauls D, Smith SA, Smith SD, DeFries JC (1991) Evidence for major gene transmission of developmental dyslexia. JAMA 266:1527–1534

Pennington BF, Smith S, Kimberling W, Green P, Haith M (1987) Left-handedness and immune disorders in familial dyslexics. Arch Neurol 44:634–639

Rabin M, Wen XL, Hepburn M, Lubs HA, Feldman E, Duara R (1993) Suggestive linkage of developmental dyslexia to chromosome 1p34-p36. Lancet 342:178

Reed PW, Davies JL, Copeman JB, Bennett ST, Palmer SM, Pritchard LE, Gough SC, et al (1994) Chromosome-specific microsatellite sets for fluorescence-based, semi-automated genome mapping. Nat Genet 7:390–395

Rutter M, Yule W (1975) The concept of specific reading retardation. J Child Psychol Psychiatry 16:181–197

Shaywitz SE, Shaywitz BA, Fletcher JM, Escobar MD (1990) Prevalence of reading disability in boys and girls. JAMA 264:998–1002

Siegel LS, Himel N (1998) Socioeconomic status, age and the classification of dyslexics and poor readers: the dangers of using IQ scores in the definition of reading disability. Dyslexia 4:90–103

Smith SD, Gilger JW, Pennington BF (1996) Dyslexia and other specific learning disorders. In: Rimoin DL, Connor JM, Pyeritz RE (eds) Principles and practice of medical genetics, Churchill Livingston, New York, pp 1767–1789

Smith SD, Kimberling WJ, Pennington BF, Lubs HA (1983) Specific reading disability: identification of an inherited form through linkage analysis. Science 219:1345–1347

Sobel E, Lange K (1996) Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker-sharing statistics. Am J Hum Genet 58:1323–1337

Thomas CJ (1905) Congenital "word-blindness" and its treatment. Ophthalmoscope 3:380–385

Thompson ME (1982) The assessment of children with specific reading disabilities (dyslexia) using the British ability scales. Br J Psychol 73:461–478

Weeks DE, Lange K (1988) The affected-pedigree-member method of linkage analysis. Am J Hum Genet 42:315–326

Wijsman EM, Amos CI (1997) Genetic analysis of simulated oligogenic traits in nuclear and extended pedigrees: summary of GAW10 contributions. Genet Epidemiol 14: 719–735

Wilson AF, Elston RC (1993) Statistical validity of the Haseman-Elston sib-pair test in small samples. Genet Epidemiol 10:593–598