

ARTICLE

A family-based association study does not support *DYX1C1* on 15q21.3 as a candidate gene in developmental dyslexia

Cecilia Marino^{*1}, Roberto Giorda², Maria Luisa Lorusso³, Laura Vanzin¹, Nello Salandi¹, Maria Nobile¹, Alessandra Citterio¹, Silvana Beri², Valentina Crespi¹, Marco Battaglia^{4,5} and Massimo Molteni¹

¹Department of Child Psychiatry, Scientific Institute 'Eugenio Medea', via Don L Monza 20, 23842 Bosisio Parini (LC), Italy; ²Molecular Biology Laboratory, Scientific Institute 'Eugenio Medea', via Don L Monza 20, 23842 Bosisio Parini (LC), Italy; ³Department of Neuropsychology, Scientific Institute 'Eugenio Medea', via Don L Monza 20, 23842 Bosisio Parini (LC), Italy; ⁴Department of Psychology, Vita-Salute San Raffaele University at IRCCS 'San Raffaele', via Stamira d'Ancona 20, Milan, Italy; ⁵Division of Epidemiology, National Institutes of Public Health, P.O. Box 4404 Nydalen, Oslo, Norway

We applied a family-based association approach to investigate the role of the *DYX1C1* gene on chromosome 15q as a candidate gene for developmental dyslexia (DD) to 158 families containing at least one dyslexic child. We directly sequenced exons 2 and 10 of the *DYX1C1* gene and found eight single nucleotide polymorphism (SNPs), three of which (–3G>A, 1249 G>T, 1259 C>G) were suitable for the genetic analyses. We performed single- and multimarker association analyses with DD as a categorical trait by FBAT version 1.4 and TRANSMIT version 2.5.4 programs. Our sample had a power of at least 80% to detect an association between the selected phenotypes and the informative polymorphisms at a significance level of 5%. The results of the categorical analyses did not support the involvement of the *DYX1C1* gene variants in this sample of dyslexics and their relatives. Quantitative and multimarker analyses, which provide greater power to detect loci with a minor effect, consistently yielded nonsignificant results. While *DYX1C1* is a good candidate gene for DD, we were unable to replicate the original findings between *DYX1C1* gene and DD, perhaps due to genetic heterogeneity.

European Journal of Human Genetics (2005) 13, 491–499. doi:10.1038/sj.ejhg.5201356

Published online 26 January 2005

Keywords: developmental dyslexia; *DYX1C1*; quantitative analysis; family-based association study

Introduction

Developmental dyslexia (DD) is a heritable¹ complex disorder, which usually becomes apparent in the first school years as soon as children learn to read. While

reading difficulties are the hint that most often leads to a clinical diagnosis of DD, they probably constitute only the most visible dysfunction of a more composite picture of a defective neurocognitive profile, which is likely to predate the diagnosis and extend across different functions.^{2,3} As such, DD could be better conceptualized as a complex syndrome, whose exact etiopathogenetic pathways from genes to phenotype are still unknown in spite of considerable effort.⁴ Genetic studies have helped to unravel part of this etiologic *dilemma*, however. The linkage approach that can be applied without any *a priori* hypothesis on the

*Correspondence: Dr C Marino, Department of Child Psychiatry, Scientific Institute 'Eugenio Medea', via Don L Monza 20, 23842 Bosisio Parini (LC), Italy. Tel: +39 031 877 595; Fax: +39 031 877 356;

E-mail: cmarino@bp.lnf.it

Received 17 July 2004; revised 8 November 2004; accepted 11 November 2004

etiopathogenetic nature of the disorder has revealed several candidate chromosomal areas on chromosomes 1,^{5,6} 2,⁷⁻⁹ 3,¹⁰ 6p,¹¹⁻¹⁸ 6q,¹⁹ 15q,^{13,20-26} and 18.²⁷

Recently, *DYX1C1* has been proposed as a candidate gene in the 15q region in a family cosegregating a t(2;15)(q11;q21) and DD;²⁸ *DYX1C1* contains 10 exons spanning about 78 kb of genomic DNA and codes for a nuclear tetratricopeptide repeat domain protein. Eight SNPs were identified by direct sequencing, five of which were in the coding region (4C>T, 270G>A, 572G>A, 1249G>T, 1259C>G), whereas three resided in the 5' untranslated region (-164C>T, -3G>A, -2G>A). DD was associated with alleles -3A and 1249T in a case-control sample (corrected *P*-values 0.016 and 0.048 respectively). Furthermore, the -3A/1249T haplotype showed association with DD in the case-control sample (*P*=0.015) and in nine informative trios (*P*=0.025).²⁸ Support for *DYX1C1* as contributing to DD was recently found in a sample of 148 families identified through a proband with reading difficulties:²⁹ evidence for linkage disequilibrium with DD was found for the rs11629841 polymorphism in intron 4 and haplotypes of this polymorphism. However, quantitative analyses of reading and reading-related phenotypes were significant for the -3G allele, which contrasts with the original finding of Taipale *et al*²⁸ association with the alternative allele (-3A) of this polymorphism. Moreover, a linkage disequilibrium was found²⁹ between DD and the haplotype -3G/1249G, instead of the association with the more rare -3A/1249T haplotype originally reported by Taipale *et al*.²⁸

In this study we assessed the possible association with DD and six of the SNPs previously identified within *DYX1C1*,²⁸ namely -164C>T, -3G>A, -2G>A, 4C>T, located in exon 2, and 1249G>T and 1259C>G, in exon 10, in order to further investigate the contribution of this gene to DD. We focused on exons 2 and 10 because they contain most of the known polymorphisms, including the two that showed association with DD in the original report.²⁸ We adopted the intrafamily design in order to overcome spurious associations due to population stratification effects. We implemented both single and multi-marker analyses of DD as a categorical trait. Since DD is likely to involve several different neurocognitive functions upon which subjects differ one from the other quantitatively,¹ we also ran single and multimarker quantitative analyses with neuropsychological components of the phenotypic dyslexic profile.

Methods

Sample

This study is based on a sample of children recruited for reading difficulties from the Department of Child Psychiatry and Rehabilitation Centre at the Eugenio Medea Institute, Bosisio Parini, Italy, a facility where children are

referred mainly by paediatricians and teachers from schools of the same geographical area for diagnosis and treatment of a wide range of mental disorders, including learning disorders and dyslexia.^{26,30} Proband was recruited in our sample if they met criteria for DD based on the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM IV);³¹ accordingly, exclusion criteria were the presence of neurological or sensorineural disorders. Reading abilities were investigated through a battery of tests that encompassed several reading tasks standardized on the Italian population^{32,33} and the Wechsler Intelligence Scale for Children, Revised.³⁴

Reading tests were as follows:

(1) *Text Reading*: 'Prove di rapidità e correttezza nella lettura del gruppo MT' ('Test for speed and accuracy in reading, developed by the MT group'), a text-reading task that assesses reading abilities for meaningful material. It provides separate scores for speed (seconds) and accuracy (expressed in number of errors). Texts increase in complexity with grade level. Norms are provided for each text.³³

(2) *Single word and nonword reading subtests of 'Batteria per la Valutazione della Dislessia e Disortografia Evolutiva'* (Battery for the assessment of Developmental Reading and Spelling Disorders), Subtests 4 and 5 respectively.³² These tests assess speed (s) and accuracy (expressed in number of errors) in reading word lists (four lists of 24 words) and nonword lists (three lists of 16 nonwords) and provide grade norms from the second to the last grade of junior high school.

Subjects' scores in each of these tasks are expressed in standard deviation units relative to the norm for the age-appropriate general Italian population.

The information gathered in the assessments described above is employed to decide whether each child would meet the following standardized inclusion criteria:

(A) Performance on timed text-reading tests at least 2 standard deviations below the general population mean on at least one of the following two parameters: (1) accuracy (2) speed; Or:

(B) A text reading score at least 1.5 standard deviations below the general population mean on at least one of the previous parameters, *and* an absolute score at least 2 standard deviations below the general population mean on accuracy or speed in reading single unrelated words or pronounceable nonwords;

And:

(C) IQ \geq 85.

A total of 158 probands identified as having DD and their biological parents accepted to participate in the study over a period of 37 months. Parents were also asked to allow the participation into the extensive clinical assessment of the probands' siblings (aged between 6 and 18 years) with a history of reading difficulties/probable dyslexia. Only siblings who met the above-mentioned criteria were included in the study. Blood or mouthwash samples were

then collected from parents, probands and siblings to perform genetic analyses. The sample consisted in 133 triads, 12 parent-child pairs and 13 nuclear families with two affected offsprings (for a total of 158 nuclear families with 171 affected children). Probands mean age was 10.9 ± 2.5 (min 7, max 18) and the male to female sex ratio was 3.5:1. Siblings were included in all analyses. Parents gave written informed consent.

Phenotype definition

Probands and their siblings were administered a battery of reading relevant tests. Seven different phenotypic measures were then identified:

(1) Phonological decoding (PD) defined as the ability to apply the correct grapheme/phoneme correspondence rules to the pronunciation of nonwords. It was measured by the *NonWord Reading Subtest* (ie Subtest 5) of the Battery for the Assessment of Developmental Reading and Spelling Disorders.³² Both the number of errors (PD accuracy) and the time (PD speed) required to complete the task were recorded.

(2) Orthographic coding (OC) defined as the ability to reproduce specific orthographic patterns. It was measured by recording the number of errors (OC accuracy) in writing under dictation *Sentences Containing Homophones*, Subtest 12 of the Battery for the Assessment of Developmental Reading and Spelling Disorders.³²

(3) Word reading (WR), measured by the *Single Unrelated Word Reading Subtest* (ie Subtest 4) of the Battery for the Assessment of Developmental Reading and Spelling Disorders.³² Both the time (WR speed) required to complete the list and the number of errors (WR accuracy) were recorded.

(4) Word spelling (WS), measured by recording the number of errors (WS accuracy) in the *Single Unrelated Word Writing Subtest* (ie Subtest 10) of the Battery for the Assessment of Developmental Reading and Spelling Disorders.³²

(5) Nonword spelling (NWS), measured by recording the number of errors (NWS accuracy) in the *NonWord Writing Subtest* (ie Subtest 11) of the Battery for the Assessment of Developmental Reading and Spelling Disorders.³²

Raw scores were converted into age-adjusted standard deviation units from the norm by application of the standard norms in the test protocol.

It should be noted that, because of transparent orthography, Italian words can generally be read both via the indirect phonological route and/or via the direct lexical route, the latter probably being more employed in the case of highly familiar, high-frequency words. In contrast, word spelling requires more detailed orthographic knowledge as phonologically plausible but orthographically incorrect alternatives are sometimes possible in the written form. Therefore, only word spelling and sentence writing under dictation have been taken to represent the use of lexical orthographic knowledge in the present study. Both non-word reading and spelling on the other hand require and measure, as it is usually assumed, access to phonological conversion rules (the phonological route).

Tables 1a and b show, respectively, the descriptive statistics and the correlation coefficients for the seven phenotypes: PD accuracy, PD speed, OC accuracy, WR accuracy, WR speed, WS accuracy and NWS accuracy in our sample. Information on psychometric tests was not available for a minority of subjects, leading to slightly different sample sizes across reading measures. Since the distributions of reading measures deviated from normality, log-transformation of the absolute scores was applied where necessary to generate normal distributions. We alternatively used the original and normalized scores of the phenotypic measures in the genetic analyses since normalizing transformation might decrease sensitivity to the genetic signal.

Laboratory procedure

Genomic DNA was extracted from 3 ml samples of blood³⁵ or, in a minority of cases, from mouth wash samples collected in 4% sucrose using DNAzol Genomic DNA isolation reagent (MRC, Cincinnati, OH, USA).

Exons 2 and 10 of the *DYX1C1* gene were amplified from genomic DNA of all subjects (primer sequences and amplification protocols are available from the authors on request). We decided that direct sequencing of both exons was the best approach. A 0.5 µl aliquot of each amplified

Table 1a Descriptive analyses for the phenotypes' measures of DD

Phenotype	N	Mean	Min	Max	Skewness	Kurtosis	Skewness after transformation	Kurtosis after transformation
PD accuracy	131	-3.9	-40.7	0.84	-4.51	30.75	-0.30	0.89
PD speed	134	-2.58	-15.76	1.14	-1.43	-4.8	-1.29	0.94
OC accuracy	109	-5.08	-25.3	1.4	-1.44	1.8	-0.42	-0.18
WR accuracy	131	-4.30	-18.6	1.23	-1.73	3.31	-0.68	0.71
WR speed	134	-3.34	-16.0	0.77	-1.73	2.94	-0.68	0.41
WS accuracy	118	-3.36	-18.9	0.93	-1.9	3.7	-0.53	0.4
NWS accuracy	120	-0.68	-15.56	1.38	-4.15	27.79	-0.30	-0.39

Measures are in standard deviation units, relative to the age-appropriate Italian population norm; the absolute scores were log-transformed, to achieve suitable skewness and kurtosis values.

Table 1b Bivariate correlations for the phenotype measures

	<i>WR accuracy</i>	<i>WR speed</i>	<i>PD accuracy</i>	<i>PD speed</i>	<i>WS accuracy</i>	<i>NWS accuracy</i>	<i>OC accuracy</i>
WR accuracy		0.44	0.72	0.32	0.39	0.25	0.47
WR speed	0.000*		0.32	0.67	0.28	0.28	0.46
PD accuracy	0.000*	0.000*		0.24	0.19	0.13	0.29
PD speed	0.000*	0.000*	0.006		0.30	0.45	0.28
WS accuracy	0.000*	0.002*	0.047	0.001*		0.54	0.52
NWS accuracy	0.007	0.002*	0.158	0.000*	0.000*		0.37
OC accuracy	0.000*	0.000*	0.003	0.003	0.000*	0.000*	

The upper right triangle shows the Spearman's ρ correlation coefficients; the lower left triangle shows the P -values for two-tailed tests and not adjusted for multiple comparisons; corrected nominal alpha was set at 0.002 (corrected for 21 tests). Flag * marks significant correlations.

DNA sample was labelled with a BigDye Terminator cycle sequencing kit (Applied Biosystems, Monza, Italy) and sequenced on an ABI3100 Avant Genetic Analyzer (Applied Biosystems). Sequences were aligned with Autoassembler (Applied Biosystems) and scored for known and new polymorphisms.

Statistical analyses

We used the transmission disequilibrium test (TDT) to assess whether the SNPs within the *DYC1X1* candidate gene and the hypothetical disease locus were in linkage disequilibrium.³⁶ The TDT for single markers and haplotypes for both quantitative and qualitative analyses was performed using the FBAT program, version 1.4^{37,38} (available from 'the FBAT Web page' at <http://biosun1.harvard.edu/~fbat/fbat.htm>). The general 'FBAT' statistic (S) is based on a linear combination of offspring genotypes and traits; it is calculated under the null hypothesis of no association, conditioning on traits and on parental genotypes.³⁷ The biallelic option was used, because all SNPs were diallelic. We set at 10 the minimum number of informative families necessary to perform the analyses.³⁷ Since the familial transmission of DD is complex, and evidence from twin studies is compatible with a multifactorial condition with an additive genetic component,³⁹ analyses were conducted under the assumption of an additive pattern of inheritance. This is a viable choice to analyse family-based associations when the true mechanism of transmission is unknown.³⁷ For comparison purposes, we performed the same haplotype analysis for the categorical trait using the program TRANSMIT version 2.5.4⁴⁰ (available from the 'David Clayton's Genetic Programs' Web page at <http://www.gene.cimr.cam.ac.uk/clayton/software/>). TRANSMIT was run with the robust variance estimator option, which allows for the inclusion of more than one affected offspring per family, even in the presence of linkage.⁴¹ The $-c$ flag was used to omit haplotypes with frequencies $\leq 5\%$ from the analyses. Statistical analyses such as bivariate correlations and analyses of variance (ANOVAs) were run with the SPSS software, version 7.5 for Windows. We applied the

Bonferroni correction, by which the nominal alpha is adjusted upon the number of tests performed for each set of analyses; some authors argued that Bonferroni correction is an extremely stringent approach to deal with when variables tested are highly correlated, as it is the case for both phenotypes and SNPs involved in the present association analyses.⁴² Linkage disequilibrium between the markers was estimated using the program EMLD (available at <https://epi.mdanderson.org/~qhuang/Software/pub.htm>). The protocol of this study had received the approval of the ethical committee of the Eugenio Medea Institute.

Results

We found eight SNPs ($-164C>T$, $-159A>G$, $-129A>C$, $-13C>T>G$, $-3G>A$, $4C>T$, $1249G>T$, $1259C>G$), five of which ($-164C>T$, $-3G>A$, $4C>T$, $1249G>T$, $1259C>G$) have been previously described²⁸ (Table 2a). Three new SNPs were identified ($-159A>G$, $-129A>C$, $-13C>T>G$); five SNPs ($-159A>G$, $-129A>C$, $-13C>T>G$, $4C>T$, $-164C>T$) were not suitable for FBAT analyses due to low informativeness. Only three SNPs were included in all analyses, $-3G>A$, $1249G>T$, $1259C>G$. The previously described SNP $-2G>A$ ²⁸ was not found in our population.

The informative SNPs ($-3G>A$, $1249G>T$, $1259C>G$) were in Hardy-Weinberg equilibrium both among probands and parents. The degree of linkage disequilibrium, here measured as D' , between the polymorphisms is significant (Table 2b).

We performed ANOVAs to test if the means of phenotype's measures were different among probands' genotypes for the informative SNPs. The dependent variables were thus the seven original phenotypic scores and the independent variable was the number of the minor alleles (ie, 0, 1, or 2) for $-3G>A$, $1249G>T$ and $1259C>G$, inclusively. We set the corrected nominal alpha level at 0.002 (7×3 tests). None of the ANOVAs yielded significant results (Table 3).

Table 2a Number of informative families in the single-marker analyses of DD as a diagnostic category by FBAT and SNPs allele frequencies in the present study and in the Taipale *et al*²⁸ and Wigg *et al*²⁹ studies

Polymorphism	Allele	Number of informative families	Allele frequencies in the present study	Allele frequencies in Taipale <i>et al</i> ²⁸	Allele frequencies in Wigg <i>et al</i> ²⁹
-164C>T	C	9	0.983	0.962	—
-159A>G	A	1	0.998	—	—
-129A>C	A	1	0.998	—	—
-13C>T>G	C	4	0.993	—	—
-3G>A	G	35	0.924	0.915	0.922
4C>T	C	1	0.998	0.994	—
1249G>T	G	61	0.874	0.883	0.901
1259C>G	C	53	0.883	0.903	—

Table 2b Linkage disequilibrium between the informative markers

Polymorphism	-3G>A	1249G>T	1259C>G
-3G>A		0.818	0.591
1249G>T	0.385		1
1259C>G	0.004	0.021	

D' values are shown in the upper half of the table, r² values in the lower half of the table.

Table 3 Unadjusted P-values for phenotypic ANOVA's

	-3G>A	1249G>T	1259C>G
WR accuracy	0.341	0.944	0.398
WR speed	0.510	0.475	0.220
PD accuracy	0.956	0.344	0.458
PD speed	0.698	0.064	0.440
WS accuracy	0.944	0.420	0.995
NWS accuracy	0.270	0.056	0.616
OC accuracy	0.650	0.342	0.976

Table 4 shows the results of the single-marker analyses by FBAT. No single SNP showed significant association with either DD as a diagnostic category or the normalized quantitative phenotypes at the 0.017 (adjusted for three markers) or 0.002 (adjusted for 3 markers × 7 phenotypes) significance levels respectively.

Table 5 shows the results of both qualitative and quantitative analyses by haplotype FBAT for the most common haplotypes of the combination -3G>A/1249G>T/1259C>G. No haplotype showed significant association with either DD as a diagnostic category or the normalized measures of the neuropsychological phenotypes at the 0.013 (adjusted for four haplotypes) and 0.002 (adjusted for 4 haplotypes × 7 phenotypes) significance levels, respectively.

We also examined the haplotype association of the combination -3G>A/1249G>T/1259C>G by TRANSMIT. Since data from all families can be used by TRANSMIT³⁷

158 families were informative for this test, with no evidence towards association (global $\chi^2 = 5.924$, 4 df, *P*-value = 0.20). Table 6 shows the χ^2 test statistics for the common haplotypes for combination -3G>A/1249G>T/1259C>G. No haplotype showed a significant linkage disequilibrium with DD as a diagnostic category at a corrected nominal alpha of 0.013 (corrected for four tests). We also ran the haplotype analyses for the combination -3G>A/1249G>T with no evidence towards association (global $\chi^2 = 2.9$, 3 df, *P*-value = 0.41). Quantitative data analyses are not possible with TRANSMIT yet, therefore a comparison on quantitative results by haplotype FBAT was not possible at this time.

Genetic analyses for quantitative traits were also run using the original, non-normalized scores for the phenotypic measures without any significant results.

Discussion

We applied a within-family association approach using tightly linked markers (-3G>A, 1249G>T, 1259C>G) within the *DYX1C1* gene as a powerful tool to explore the causal relationships of this candidate gene with DD. We ran both single- and multi-marker analyses with both DD as a diagnostic category and quantitative measures of the dyslexic phenotype, none of which yielded significant results at rigorously stringent, corrected nominal alpha levels.

Thus, there appears to be a discrepancy between our study and the original report²⁸ regarding the role of *DYX1C1* in DD. While this study did not replicate the Taipale *et al*²⁸ work, a direct comparison between our and the Taipale *et al* data would be hazardous, due to non complete overlapping of the statistical designs implemented in the two studies.

Our results are also in contrast with the recent report by Wigg *et al*.²⁹ We have not replicated the association with -3G in quantitative analyses, even though phenotypes of reading and reading-related measures may not be completely comparable, due to language-related differences.

Table 4 Single-marker analyses between SNPs $-3G>A$, $1249G>T$, $1259C>G$ and both DD as a diagnostic category and phenotypes' measures of DD by FBAT

SNP	Phenotype	Allele	Number of informative families	S	$E(S)$	$Var(S)$	Z	P
	DD							
	Diagnostic category							
$-3G>A$		G	35	53	54.5	10.75	-0.46	0.65
		A		23	21.5	10.75	0.46	0.65
$1249G>T$		G	61	89	93.5	17.75	-1.07	0.29
		T		41	36.5	17.75	1.07	0.29
$1259C>G$		C	53	75	79.5	16.25	-1.12	0.26
		G		37	32.5	16.25	1.12	0.26
	WR accuracy							
$-3G>A$		G	28	44.63	47.63	15.67	-0.75	0.44
		A		21.15	18.14	15.67	0.75	0.44
$1249G>T$		G	46	67.21	73.12	22.11	-1.25	0.20
		T		33.31	27.4	22.11	1.25	0.20
$1259C>G$		C	38	63.98	67.88	30.88	-0.7	0.48
		G		33.96	30.05	30.88	0.7	0.48
	WR speed							
$-3G>A$		G	29	24.08	25.64	14.16	-0.41	0.78
		A		12.52	10.96	14.16	0.41	0.78
$1249G>T$		G	47	30.54	33.33	21.17	-0.6	0.62
		T		17.7	14.91	21.17	0.6	0.62
$1259C>G$		C	40	31.6	34.93	29.04	-0.61	0.66
		G		22.02	18.69	29.04	0.61	0.66
	OC accuracy							
$-3G>A$		G	22	39.34	39.80	15.44	-0.11	0.90
		A		15.24	14.77	15.44	0.11	0.90
$1249G>T$		G	35	58.58	54.4	20.61	0.92	0.35
		T		17.38	21.55	20.61	-0.92	0.35
$1259C>G$		C	31	44.95	48.8	25.16	-0.76	0.44
		G		26.81	22.96	25.16	0.76	0.44
	PD accuracy							
$-3G>A$		G	28	29.93	31.77	11.63	-0.54	0.58
		A		13.31	11.46	11.63	0.54	0.58
$1249G>T$		G	46	51.34	51.25	16.87	0.02	0.98
		T		20.22	20.30	16.87	-0.02	0.98
$1259C>G$		C	38	51.47	54.87	26.99	-0.65	0.51
		G		29.03	25.62	26.99	0.65	0.51
	PD speed							
$-3G>A$		G	28	-3.77	1.78	21.56	-1.19	0.23
		A		6.47	0.93	21.56	1.19	0.23
$1249G>T$		G	47	-15.34	-16.43	37.21	0.18	0.86
		T		-4.62	-3.54	37.21	-0.18	0.86
$1259C>G$		C	40	17.09	12.38	40.6	0.74	0.46
		G		3.91	8.63	40.6	-0.74	0.46
	WS accuracy							
$-3G>A$		G	24	15.49	20.68	13.24	-1.42	0.15
		A		13.85	8.66	13.24	1.42	0.15
$1249G>T$		G	39	26.64	27.46	22.55	-0.17	0.86
		T		9.2	8.37	22.55	0.17	0.86
$1259C>G$		C	36	42.95	40.80	22.33	0.45	0.64
		G		18.23	20.37	22.33	-0.45	0.64
	NWS accuracy							
$-3G>A$		G	24	-27.65	-28.94	11.04	0.39	0.69
		A		-11.49	-10.19	11.04	-0.39	0.69
$1249G>T$		G	40	-43.86	-46.04	18.27	0.51	0.6
		T		-19.8	-17.61	18.27	-0.51	0.6
$1259C>G$		C	38	-20.74	-20.76	13.54	0.007	0.99
		G		-6.86	-6.83	13.54	-0.007	0.99

S is the test statistic; $E(S)$ and $Var(S)$ are the expected value and variance of the test statistic under H_0 ; Z and P are the Z statistic (S normalized using $E(S)$ and $Var(S)$) and its corresponding P -value. The sign of the Z statistic (+/-) indicates whether the allele/haplotype is associated with better or poorer scores. P -values are two-tailed and not adjusted for multiple comparisons.

Table 5 Estimated Z statistics for common haplotypes of combination –3G>A/1249G>T/1259C>G for both qualitative and quantitative analyses by haplotype FBAT

Haplotype	Frequency	Informative families	S	E(S)	Var(S)	Z	P
<i>DD diagnostic category</i>							
G/G/C	0.75	83	115.82	124.57	27.03	-1.68	0.09
G/G/G	0.11	52	37	31.75	16.15	1.31	0.19
G/T/C	0.07	37	22.18	20.18	10.88	0.61	0.54
A/T/C	0.06	26	16.82	14.32	6.63	0.97	0.33
<i>PD speed</i>							
G/G/C	0.75	61	16.07	9.85	65.16	0.77	0.44
G/G/G	0.11	38	7.9	12.44	36.49	-0.75	0.45
G/T/C	0.07	24	-5.2	-1.13	18.07	-0.96	0.34
A/T/C	0.06	20	2.8	2.18	13.54	0.17	0.87
<i>PD accuracy</i>							
G/G/C	0.75	61	80.17	84.49	40.44	-0.68	0.50
G/G/G	0.11	37	29.03	25.37	26.94	0.71	0.48
G/T/C	0.07	26	8.77	9.96	7.42	-0.44	0.66
A/T/C	0.06	20	9.57	8.47	9.0	0.37	0.71
<i>OC accuracy</i>							
G/G/C	0.75	48	80.24	81.36	40.28	-0.18	0.86
G/G/G	0.11	31	26.81	22.55	24.65	0.86	0.39
G/T/C	0.07	20	7.21	10.82	10.29	-1.13	0.26
A/T/C	0.06	16	9.57	10.14	11.02	-0.17	0.86
<i>WR speed</i>							
G/G/C	0.75	63	47.85	51.95	44.66	-0.61	0.54
G/G/G	0.11	39	22.02	18.72	28.63	0.62	0.54
G/T/C	0.07	27	3.98	4.74	13.76	-0.20	0.84
A/T/C	0.06	21	12.67	9.12	8.59	1.21	0.23
<i>WR accuracy</i>							
G/G/C	0.75	61	93.34	105.11	47.41	-1.71	0.87
G/G/G	0.11	36	33.96	29.38	30.35	0.83	0.41
G/T/C	0.06	26	17.77	13.58	10.13	1.32	0.19
A/T/C	0.06	20	13.2	11.47	10.56	0.53	0.59
<i>WS accuracy</i>							
G/G/C	0.75	55	49.70	51.14	38.38	-0.23	0.82
G/G/G	0.11	35	18.23	19.56	21.48	-0.29	0.78
G/T/C	0.06	23	0.95	3.38	12.37	-0.69	0.49
A/T/C	0.06	16	7.6	4.35	8.52	1.12	0.26
<i>NWS accuracy</i>							
G/G/C	0.75	57	-44.88	-45.9	22.97	0.22	0.83
G/G/G	0.11	37	-6.86	-6.88	13.48	0.01	0.99
G/T/C	0.06	24	-8.47	-8.72	8.82	0.08	0.93
A/T/C	0.06	17	-10.99	-8.56	9.44	-0.79	0.43

S is the test statistic; E(S) and Var(S) are the expected value and variance of the test statistic under H0; Z and P are the Z statistic (S normalized using E(S) and Var(S)) and its corresponding P-value. The sign of the Z statistic (+/-) indicates whether the allele/haplotype is associated with better or poorer scores. P-values are two-tailed and not adjusted for multiple comparisons.

Furthermore, our haplotype analyses involving the –3G/1249G combination showed no clear evidence of linkage disequilibrium with DD. Our sample has not been genotyped for the rs11629841 polymorphism located in intron 4 since our main aim was to replicate the Taipale *et al*²⁸ work. We examined the possibility that our failure to detect an association between the *DYX1C1* locus and DD susceptibility could be due to insufficient power using the

PBAT Power Calculator (<http://www.biostat.harvard.edu/~fbat/pbat.htm>). We assessed the conditional power to detect the association with the informative polymorphisms for both the dichotomous and continuous traits. The Taipale *et al*²⁸ original report was used as a guide to estimate the parameters for the calculations. Calculations were run for a genetic additive model, with an allele frequency of the disease gene of 0.06 and an Attributable

Table 6 Estimated χ^2 statistics for common haplotypes of combination -3G>A/1249G>T/1259C>G by TRANSMIT

Haplotype	Obs	Exp	Var (Obs-Exp)	χ^2	P
G/G/C	233.77	245.2	28.85	4.53	0.03
G/T/C	22.11	20.56	11.03	0.22	0.64
A/T/C	24.96	21.07	6.8	2.2	0.14
G/G/G	45.12	38.03	17.32	2.9	0.09

Markers are shown in order -3G>A/1249G>T/1259C>G. Obs=observed transmissions of haplotype to affected offspring. Exp=expected transmission under Mendelian inheritance. Var(Obs-Exp)=variance of the difference between observed and expected transmissions. P-values are not adjusted for multiple comparisons; corrected nominal alpha was set at 0.013 (corrected for four tests).

Fraction of 0.13. According to our power calculations, for disease status the conditional power at a significance level of 5% is 80%. Quantitative analyses enhanced the statistical power. Assuming a conservative estimate of heritability for each phenotype of 0.4, for a sample of 120 individuals (the average of subjects in our sample for whom the phenotype was described) the conditional power is 87%.

The implementation of haplotype analyses further increased the number of informative families, at least for the most common haplotypes, and this led to much more reliable results. Furthermore, haplotype analyses by TRANSMIT use data from all the nuclear families of the sample, an attractive feature for an association study when the frequencies of markers are quite low.³⁷

In conclusion, there is strong evidence of the presence of genes of small effects that contribute to the DD phenotype on 15q^{13,20-26} and *DYX1C1* is a good putative candidate gene,^{28,29} although our data clearly provide evidence against *DYX1C1* being a susceptibility gene for both dyslexia and its reading-related phonological phenotypes.

However, a better understanding of the component processes of the dyslexic phenotype would help to understand which functions are specifically affected by this gene. Perhaps, and more reliably, our negative results deal with the heterogeneity of DD, and *DYX1C1* might be considered a candidate gene for special dyslexic subgroups, or within some isolated populations.

References

- Fisher SE, De Fries JC: Developmental dyslexia: genetic dissection of a complex cognitive trait. *Nat Rev Neurosci* 2002; **3**: 767-780.
- Facoetti A: Reading and selective spatial attention: evidence from behavioral studies in dyslexic children; in *Progress in dyslexia research*. Hauppauge, NY: Nova Science Publishers, 2004, in press.
- Stein J: Visual motion sensitivity and reading. *Neuropsychologia* 2003; **41**: 1785-1793.
- Habib M: The neurological basis of developmental dyslexia: an overview and working hypothesis. *Brain* 2000; **12**: 2373-2399.

- Rabin M, Wen XL, Hepburn M: Suggestive linkage of developmental dyslexia to chromosome 1p34-p36. *Lancet* 1993; **342**: 178.
- Grigorenko EL, Wood FB, Meyer MS *et al*: Linkage studies suggest a possible locus for developmental dyslexia on chromosome 1p. *Am J Med Genet* 2001; **105**: 120-129.
- Fagerheim T, Raeymaekers P, Tonnessen FE *et al*: A new gene (*DYX3*) for dyslexia is located on chromosome 2. *J Med Genet* 1999; **36**: 664-669.
- Francks C, Fisher SE, Olson RK *et al*: Fine mapping of the chromosome 2p12-16 dyslexia susceptibility locus: quantitative association analysis and positional candidate genes SEMA4F and OTX1. *Psychiatry Genet* 2002; **12**: 35-41.
- Petryshen TL, Kaplan BJ, Hughes ML *et al*: Supportive evidence for the *DYX3* dyslexia susceptibility gene in Canadian families. *J Med Genet* 2002; **39**: 125-126.
- Nopola-Hemmi J, Myllyluoma B, Haltia T *et al*: A dominant gene for developmental dyslexia on chromosome 3. *J Med Genet* 2001; **38**: 658-664.
- Cardon LR, Smith SD, Fulker DW: Quantitative trait locus for reading disability on chromosome 6. *Science* 1994; **266**: 2276-2279.
- Cardon LR, Smith SD, Fulker DW: Quantitative trait locus for reading disability: a correction. *Science* 1995; **268**: 5217.
- Grigorenko EL, Wood FB, Meyer SB *et al*: Susceptibility loci for distinct components of developmental dyslexia on chromosome 6 and 15. *Am J Hum Genet* 1997; **60**: 27-39.
- Fisher SE, Marlow AJ, Lamb J *et al*: A quantitative trait locus on chromosome 6 influences different aspects of developmental dyslexia. *Am J Hum Genet* 1999; **64**: 146-156.
- Gayan J, Smith SD, Cherny SS *et al*: Quantitative trait locus for specific language and reading deficits in chromosome 6p. *Am J Hum Genet* 1999; **64**: 157-164.
- Grigorenko EL, Wood FB, Meyer MS *et al*: Chromosome 6p influences on different dyslexia-related cognitive processes: further confirmation. *Am J Hum Genet* 2000; **66**: 715-723.
- Kaplan DE, Gayan J, Ahn J *et al*: Evidence for linkage and association with reading disability on 6p21.3-22. *Am J Hum Genet* 2002; **70**: 1287-1298.
- Turic D, Robinson L, Duke M *et al*: Linkage disequilibrium mapping provides further evidence of a gene for reading disability on chromosome 6p21.3-22. *Mol Psychiatry* 2003; **8**: 176-185.
- Petryshen TL, Kaplan BJ, Fu Liu M *et al*: Evidence for a susceptibility locus on chromosome 6q influencing phonological coding dyslexia. *Am J Med Genet* 2001; **105**: 507-517.
- Smith SD, Kimberling WJ, Pennington BF *et al*: Specific reading disability-identification of an inherited form through linkage analysis. *Science* 1983; **219**: 1345.
- Smith SD, Kimberling WJ, Pennington BF: Screening for multiple genes influencing dyslexia. *Read Writ* 1991; **3**: 285-298.
- Fulker DW: Multiple regression of sib-pair data on reading to detect quantitative trait loci. *Read Writ* 1991; **3**: 299-313.
- Schulte-Korne G, Grimm T, Nothen MM *et al*: Evidence for linkage of spelling disability to chromosome 15. *Am J Hum Genet* 1998; **63**: 279-282.
- Morris DW, Robinson L, Turic D *et al*: Family-based association mapping provides evidence for a gene for reading disability on chromosome 15q. *Hum Mol Genet* 2000; **9**: 843-848.
- Nopola-Hemmi J, Taipale M, Haltia T *et al*: Two translocations of chromosome 15q associated with dyslexia. *J Med Genet* 2000; **37**: 771-775.
- Marino C, Giorda R, Vanzin L *et al*: A locus on 15q15-15qter influences dyslexia: further support from a transmission/disequilibrium study in an Italian speaking population. *J Med Genet* 2004; **41**: 42-46.
- Fisher SE, Francks C, Marlow AJ *et al*: Independent genome-wide scans identify a chromosome 18 quantitative-trait locus influencing dyslexia. *Nat Genet* 2002; **30**: 86-91.
- Taipale M, Kaminen N, Nopola-Hemmi J *et al*: A candidate gene for developmental dyslexia encodes a nuclear tetratricopeptide

- repeat domain protein dynamically regulated in brain. *Proc Natl Acad Sci USA* 2003; **20**: 11553–11558.
- 29 Wigg KG, Couto JM, Feng Y *et al*: Support for EKN1 as the susceptibility locus for dyslexia on 15q21. *Mol Psychiatry* 2004; July 13; [E-pub ahead of print].
- 30 Marino C, Giorda R, Vanzin L *et al*: No evidence for association and linkage disequilibrium between dyslexia and markers of four dopamine-related genes. *Eur Child Adolesc Psychiatry* 2003; **4**: 198–202.
- 31 American Psychiatric Association: *Diagnostic and statistical manual of mental disorders*. Fourth Edition. Washington DC: American Psychiatric Association, 1994.
- 32 Sartori G, Job R, Tressoldi PE: *Batteria per la valutazione della dislessia e della disortografia evolutiva*. Firenze: Organizzazioni Speciali, 1995.
- 33 Cornoldi C, Colpo G, Gruppo MT: *Prove di lettura*. Firenze: Organizzazioni Speciali, 1986.
- 34 Wechsler D: *Examiner's manual. Wechsler Intelligence Scale for Children, Revised*. New York: Psychological Corp., 1981.
- 35 Dracopoli NC, Haines JL, Korf BR *et al*: *Current Protocols in Human Genetics* 2004, Unit 9.7. New York: Wiley.
- 36 Spielman RS, McGinnis RE, Ewens WJ: Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet* 1993; **52**: 506–516.
- 37 Horvath S, Xu X, Laird N: The family based association test method: strategies for studying general genotype–phenotype associations. *Eur J Hum Genet* 2001; **9**: 301–306.
- 38 Horvath S, Xu X, Lake SL *et al*: Family-based tests for associating haplotypes with general phenotype data: application to asthma genetics. *Genet Epidemiol* 2004; **26**: 61–69.
- 39 Pennington BF: Genetics of learning disabilities. *J Child Neurol* 1995; **10**: S69–S77.
- 40 Clayton D, Jones H: Transmission/disequilibrium tests for extended marker haplotypes. *Am J Hum Genet* 1999; **65**: 1161–1169.
- 41 Clayton D: A generalization of the transmission/disequilibrium test for uncertain haplotype transmission. *Am J Hum Genet* 1999; **65**: 1170–1177.
- 42 Zhao H, Zhang S, Merikangas KR *et al*: Transmission disequilibrium test using multiple tightly linked markers. *Am J Hum Genet* 2000; **67**: 936–946.