A locus on 15q15-15qter influences dyslexia: further support from a transmission/disequilibrium study in an Italian speaking population

C Marino, R Giorda, L Vanzin, M Nobile, M L Lorusso, C Baschirotto, L Riva, M Molteni, M Battaglia

D
velopmental dyslexia (dyslexia) is a heritable condition typically diagnosed in the first school years, characterised by an impairment of reading abilities in spite of normal intelligence and adequate educational opportunities.

While the exact neurobiological mechanisms underlying this condition remain obscure, the most convincing current aetiopathogenetic view of dyslexia is that impaired reading stems from a defective representation and manipulation of phonemes,—that is, the sounds that we combine to build words.1 There are at least two major complications when studying the ultimate causes of dyslexia. Firstly, the composite neuropsychological picture that often marks dyslexia may lead people to consider reading problems as a part of a more extended neurobiological syndrome whose phenotypic boundaries are blurred, and whose genetic determinants may be especially difficult to identify with certainty.2 Secondly, the leading criterion to diagnose dyslexia remains that of a reading performance below the population mean (typically, a reading score two standard deviations below the general population mean). While reading performance is distributed normally in the population,3 the prevalence of dyslexics will vary considerably across different cultures, because it depends on the complexity of orthographic rules specific to a given language to which a subject is exposed.4 Contradicting a culturally bound identity of dyslexia is a recent, functional brain imaging study of adult subjects with dyslexia from different cultures and languages (English, French, Italian) that showed the same abnormal patterns of brain activation during implicit and explicit reading.5 This suggests common neurobiological causes for dyslexia regardless of a person’s spoken language, while variation in prevalence estimates across different cultures could at least partially reflect local difficulties specific to each language, when homogenous diagnostic criteria are applied.5

While functional brain imaging findings suggest biological unity for dyslexia, evidence based on genetic analyses of common determinants of dyslexia in spite of markedly different orthographies and languages is incomplete. The evidence that dyslexia has a genetic basis is, nonetheless, very convincing.

The tendency of dyslexia to run in families has become clear since its earliest descriptions, and modern family studies indicate that a substantial majority of affected children have affected relatives, the average risk among first degree relatives being about 30%.6 Twin studies show that the role of genetic factors outweighs that of shared, non-genetic factors in explaining such familial resemblance for dyslexia.7 The complexity of the dyslexia phenotype has been taken into account by some, but not all, genetic studies. Estimates of broad heritability in twin samples vary quite widely across specific phenotypic components of dyslexia, yielding estimates that range between 0.44 and 0.75.5 Likewise, molecular genetic studies have investigated dyslexia both as a categorically defined, putatively homogenous illness, and as a composite condition, through specific and putatively independent components of phenotype with a wide array of strategies, including parametric and model free linkage analyses, as well as family based association designs.6 Overall, promising findings have been obtained for regions on chromosomes 1,8 2,10–12 3,13 6p,14–21 6q,22 15q,16 23–28 and 18.27 Robust results have come from molecular genetic research studies of dyslexia on chromosome 15. These were prompted by an initial evidence of a 3.24 lod score in a parametric linkage study of chromosome 15 in three generational pedigrees segregating dyslexia, which used heteromorphisms of the centromere region as markers.23 These results, however, were not confirmed by two similar, successive studies.9, 29 In a more recent study Smith et al46 replicated their earlier finding with markers within the 15q15-15qter region, using both a quantitative and qualitative phenotype definition of dyslexia, and a non-parametric approach (the Haseman-Elston regression model). Linkage within the 15q15-15qter region was further confirmed in a non-parametric analysis using the De Fries Fulker regression approach and a quantitative definition of dyslexia.25 Grigorenko et al39 found a lod score of 3.15 in single point linkage parametric analysis between the microsatellite D15S143 and a “single–word reading” component of reading performance. In the same study, non-parametric analyses generally yielded negative results. Defining a phenotype of spelling disability, other authors found linkage with markers...
D15S132 and D15S143, using both parametric and non-parametric methods. 26 Finally, linkage disequilibrium was recently detected in an area of 8cM within the putative region of linkage through a family based design, using both individual markers and multimarker haplotype analyses.27 While many, but not all, molecular genetic studies of chromosome 15 support an aetiological role for this area in dyslexia, they were all based on languages (English and, in one instance,26 German), that are orthographically complex and relatively similar, although German is more transparent than English. A causal role in families of dyslexics speaking a “completely transparent” language (a language where the letters, or the groups of letters, are uniquely mapped to each unit) than English. A causal role in families of dyslexics speaking a “completely transparent” language (a language where the letters, or the groups of letters, are uniquely mapped to each phoneme of the spoken language) remains to be demonstrated.

In this study we investigated by a family based linkage disequilibrium approach, implementing single and multimarker haplotypes’ analyses, 6 microsatellite markers within the 15q15-15qter region in a sample of 121 parent-offspring Italian families. The aims were: 1) to replicate previous findings of an involvement of this chromosomal region to influence the risk for dyslexia, and 2) to support the hypothesis of a heritable basis for an unitary, neurobiological process leading to dyslexia, independent of linguistic differences.

**MATERIALS AND METHODS**

**Subjects**

Subjects were recruited consecutively 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.

To be eligible for the study, children had to have a diagnosis of reading difficulties based on the criteria in the Diagnostic and statistical manual of mental disorders, 4th ed,24 confirmed by an extensive clinical investigation that encompassed a careful medical assessment and a battery of tests. The medical assessment included a thorough neurological examination, an ophthalmological examination, and an audiometric test. The battery of tests was typically organised into two sessions, each lasting approximately 1.5 hours, and included several reading tasks standardised on the Italian population,25 and the revised Wechsler intelligence scale for children.26

Reading tests were as follows:

- **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 meant to assess reading abilities for meaningful material. It provides separate scores for speed and accuracy. Texts increase in complexity with grade level. Norms are provided for each text.24

- **Single word/non-word reading:** “Batteria per la Valutazione della Dislessia e Disortografia Evolutiva” (Battery for the assessment of developmental reading and spelling disorders).27 This test assesses speed and accuracy (expressed as the number of errors) in reading word lists (four lists of 24 words) and non-word lists (three lists of 16 non-words) and provides grade norms from the second to the last grade of junior high school.

The information gathered in the assessments described above was employed to decide whether each subject would meet the following, standardised inclusion criteria:

- performance on timed text reading tests of a reading score 2 standard deviations below the general population mean on at least one of accuracy and speed;
- a reading score 1.5 standard deviations below the general population mean on at least one of the previous parameters, and an absolute score 2 standard deviations below the general population mean on accuracy or speed in reading single unrelated words or pronounceable non-words; and
- IQ>85.

Subjects’ scores in each of these tasks were appropriately age regressed and expressed in standard deviation units relative to the average score for the normal Italian population.

One hundred and twenty one subjects identified as having dyslexia by these procedures were accepted over a period of 36 months to participate in the study after their parents were informed about the scope of the research and had provided their written informed consent to participate in the study. Parents were also asked to permit probands’ siblings to participate in the extensive clinical assessment if the siblings were between 6 and 18 years old, and if a suggestive history of reading difficulties or probable dyslexia was evident from their academic history. Only siblings who were shown to be affected according to the inclusion criteria were then included in the study. In our pool of 121 subjects, there were 12 families with at least one affected sibling in addition to the index proband.

**Laboratory procedure**

Blood samples, anticoagulated with EDTA, were taken from all probands (n = 121), affected siblings (n = 14) and their biological parents (n = 242); DNA was extracted from samples of 3 ml of blood.28 In a minority of cases DNA was extracted using the Isoquick Nucleic Acid extraction kit (ORCA Research Inc, Bothell, WA) from mouth wash samples collected in 4% sucrose.

The chromosomal interval was selected on the basis of the recent studies on suggestive linkage of dyslexia to the 15q region and markers were chosen to cover the putative region:26 27 29 30: six microsatellite markers defining allelic variation across a 9 Mb region were then chosen: D15S214 (37979 kb); D15S399 (38161 kb), D15S508 (41157 kb), D15S182 (42315 kb), D15S132 (44780 kb), and D15S1028 (46576 kb) (position kb from the April 2003 Human Genome Freeze). Locus information and primer sequences were obtained from the Genome Database (www.gdb.org). Forward primers, labelled with 5-Fam, Hex or Tet ABI dyes, and unlabelled reverse primers were obtained from MWG Biotech (Ebersberg, Germany). PCR reactions were carried out as described elsewhere.28 Samples were analysed on an ABI 310 genetic analyser (Applied Biosystems, Foster City, CA), using GeneScan version 3.1.2. Allele sizes were verified using CEPH controls.

**Statistical analyses**

We used the transmission disequilibrium test (TDT) to assess if the marker loci and the hypothetical disease locus were in linkage disequilibrium.30 The TDT for individual markers with multiple alleles was performed using the FBAT program.31 (available at http://biosun1.harvard.edu/~bat/batdoc.htm) with the empirical variance option, as appropriate in the presence of a linkage when more than one sibling per family is included in the analyses.32 The multi-allelic option was used, which takes into account multiple alleles at a one marker locus and estimates p values to indicate the degree of significance of the association of the disease with all allelic variation at each individual locus. The transmission pattern
of inheritance of dyslexia is complex, and evidence from twin studies is compatible with a multifactorial condition with an additive genetic component; therefore analyses were conducted under the assumption of an additive pattern of inheritance. This is a viable choice to analyse family based on an additive genetic component; therefore analyses were performed only for those haplotypes which frequencies exceed five. This would correspond to a value of 2.5 for Var(O–E) (see manual for TRANSMIT, by David Clayton, available at www.gene.cimr.cam.ac.uk/clayton/software/), with the robust variance estimator option, which allows for the inclusion of more than one affected offspring per family, even in the presence of a linkage. TRANSMIT’s main advantage is its ability to deal with the transmission of multilocus haplotypes, even if phase is unknown and when parental genotypes may be missing. TRANSMIT requires parameters to be entered to handle rare haplotypes and legitimately use \( \chi^2 \) statistics. For the present analysis the –c flag option, which deals with the minimum haplotype frequencies, was set at 3, which means that restricted analyses were performed only for those haplotypes which had frequencies higher than 3%. The “global” \( \chi^2 \) value represents the overall significance when the observed against expected transmissions of all haplotypes are considered together. A widely used guideline for the applicability of the \( \chi^2 \) test is that it should only be used when all expected frequencies exceed five. This would correspond to a value of 2.5 for Var(O–E); thus, \( \chi^2 \) tests were considered on restricted analyses only. Two tailed \( \chi^2 \) values were adopted. We applied the Bonferroni correction, by which the nominal alpha is adjusted upon the number of tests performed for each set of analyses. In our case the statistical significance levels were then set at 0.008, 0.003 and 0.0025, respectively for the TDT for single, two, and three markers. However, the issue of the adequate nominal alpha level of significance for these results is under debate in the scientific community. The Bonferroni correction is considered by some as unduly conservative, since there are dependencies between these multiple tests, and thus the \( \chi^2 \) values are correlated.

The protocol of this study had received the approval of the ethical committee of the Eugenio Medea Institute.

RESULTS

Allele frequencies in the sample population were comparable to available frequency data (CEPH genotype database at http://www.cephb.fr/cephdb/). The results of FBAT analyses for the six microsatellite markers taken individually are shown in table 1. Of the six markers, the D15S214 was the only one to approach significance for linkage disequilibrium with dyslexia.

The results of TRANSMIT analyses with the two markers haplotype are shown in table 2. The pattern of significance varies across the 15 possible combinations with mild trends toward linkage disequilibrium for combinations D15S214/D15S508, and D15S214/D15S182. We also examined the rate of transmission from parents to affected offspring of the 20 combinations of the three markers haplotypes, the maximum number tolerated by TRANSMIT. Evidence toward association was obtained with the combination D15S214/D15S508/D15S182 (global \( \chi^2 = 25.451, \ df = 0.005 \); table 3 shows the \( \chi^2 \) test statistics for the nine haplotypes for combination D15S214/D15S508/D15S182: haplotype 6-1-7 showed a trend toward linkage disequilibrium (\( \chi^2 = 6.99, \ 1 \ df, \ p = 0.016, \ p \ corrected \ for \ 9 \ tests = 0.14 \)). 13 dyslexics were carriers of this haplotype.

DISCUSSION

This study was primarily designed to verify the reproducibility (by allelic association or allelic linkage disequilibrium) of previous findings of an involvement of the 15q15-15qter chromosomal region in dyslexia in an Italian speaking population. When linkage results for complex diseases are obtained, supplementing evidence by linkage disequilibrium methodology is important to specify more precisely the area of the genome that is probably involved in the aetio-pathogenesis of a given disorder.

We found evidence toward linkage disequilibrium with the combination D15S214/D15S508/D15S182, which encompasses a region spanning around 4336 kb, approximately 4 cm; this result is supported mainly by the distorted transmission to the affected offspring of haplotype 6/1/7, which could point to an aetiological factor itself or to another, yet unidentified, haplotype superimposed on the 6/1/7 haplotype. In particular, the 6/1/7 haplotype had significant decreased transmission to offspring, thus leaving the possibility that it is protective against the disorder.

The loss of significance after correction for multiple comparisons that we observed among single and two marker analyses should be considered cautiously, and viewed as a

### Table 1

<table>
<thead>
<tr>
<th>Markers D15S-</th>
<th>Allele</th>
<th>( \chi^2 ) (df)</th>
<th>( p ) values</th>
</tr>
</thead>
<tbody>
<tr>
<td>D15S214</td>
<td>214</td>
<td>12.35 (5)</td>
<td>0.03</td>
</tr>
<tr>
<td>D15S994</td>
<td>994</td>
<td>7.83 (6)</td>
<td>0.25</td>
</tr>
<tr>
<td>D15S508</td>
<td>508</td>
<td>0.78 (1)</td>
<td>0.38</td>
</tr>
<tr>
<td>D15S182</td>
<td>182</td>
<td>7.96 (5)</td>
<td>0.15</td>
</tr>
<tr>
<td>D15S129</td>
<td>129</td>
<td>5.81 (5)</td>
<td>0.33</td>
</tr>
<tr>
<td>D15S1028</td>
<td>1028</td>
<td>3.83 (7)</td>
<td>0.8</td>
</tr>
</tbody>
</table>

### Table 2

<table>
<thead>
<tr>
<th>Interval length</th>
<th>Markers: D15S-</th>
<th>D15S214</th>
<th>D15S994</th>
<th>D15S508</th>
<th>D15S182</th>
<th>D15S129</th>
<th>D15S1028</th>
</tr>
</thead>
<tbody>
<tr>
<td>37979</td>
<td>214</td>
<td>–</td>
<td>0.38</td>
<td>0.028</td>
<td>0.034</td>
<td>0.14</td>
<td>0.18</td>
</tr>
<tr>
<td>38161</td>
<td>994</td>
<td>9.89 (7)</td>
<td>–</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>41157</td>
<td>508</td>
<td>14.34 (5)</td>
<td>8.16 (9)</td>
<td>–</td>
<td>0.2</td>
<td>0.82</td>
<td>ns</td>
</tr>
<tr>
<td>42315</td>
<td>182</td>
<td>18.56 (8)</td>
<td>7.23 (8)</td>
<td>10.6 (6)</td>
<td>–</td>
<td>0.068</td>
<td>0.76</td>
</tr>
<tr>
<td>44780</td>
<td>129</td>
<td>11.79 (6)</td>
<td>9.23 (11)</td>
<td>8.25 (8)</td>
<td>18.07 (9)</td>
<td>0.36</td>
<td>–</td>
</tr>
<tr>
<td>46576</td>
<td>1028</td>
<td>15.1 (9)</td>
<td>6.49 (11)</td>
<td>8.51 (10)</td>
<td>10.71 (10)</td>
<td>11.35 (8)</td>
<td>–</td>
</tr>
</tbody>
</table>

Interval length: length of interval of markers from p-tel (in kb). Lower diagonal: \( \chi^2 \) (degrees of freedom); upper diagonal: \( p \) values are two sided; corrected nominal alpha was set at 0.008.
decrease of p values around the area in linkage disequilibrium rather than a complete loss.44 Tverrilliger et al45 suggested that seeing a region of markers with p values nearing significance for association is an important indicator of the presence of a disease gene which could help distinguishing true positive from false positive, although this was mainly referred to linkage analysis.

Allelic variation at marker D15S994, which appeared to be in strong linkage disequilibrium with dyslexia in the TDT analyses of Morris et al46 did not show any significant association in our sample, either in single or multiple marker analyses. While these results appear puzzling, they are not necessarily contradictory. Such a discrepancy might be due to the phenomenon of population stratification, which accounts for great variability of marker alleles’ frequencies among populations with different founders. The chance of detecting linkage disequilibrium between a marker allele and a disease depends not only upon their reciprocal physical distance, but also upon the frequency of the marker allele, the chance being higher if the marker allele is relatively infrequent, because otherwise the increase of risk would be harder to detect. In other words, in two different populations the same marker may have a different potential for informativeness for linkage disequilibrium, depending on its alleles’ frequency.47 It appears to be particularly relevant that marker D15S994 lies within the region covered by the haplotype that displayed association with dyslexia in our sample, thus yielding further support of the boundaries of the area showing linkage disequilibrium with dyslexia.

There are two more general elements that suggest a cautious approach to these results. First, the 2.5 value as the threshold for informative transmissions (Var 0–E) can be considered a lenient criterion by some. On the other hand, there is no established minimum value that has been collectively accepted as a common reference for this point, and more work is admittedly needed in this direction (see manual for TRANSMIT, by David Clayton, available at http://www-gene.cimr.cam.ac.uk/clayton/software/). Second, the choice of a 3% frequency as the threshold for inclusion of haplotypes can also be considered too lenient, and some may feel that only haplotypes with greater frequencies should be included in the analyses. However, it is difficult to have a good a priori idea of what would be an ideal frequency threshold for dyslexia. Other investigations on several different multifactorial illnesses (including schizophrenia, bipolar disorder, and inflammatory bowel disease) have used threshold of 3% or even 2% for haplotype frequencies.48–50

The second aim of this study was to explore the hypothesis of a heritable basis for a unitary, neurobiological process leading to dyslexia,5 independent of linguistic differences. In as much as these results support the same genetic findings in a population with a shallow orthography as are reported for populations with deeper orthography, we further suggest a unitary hypothesis of the biological basis of dyslexia.


table 3

<table>
<thead>
<tr>
<th>H</th>
<th>D15S214</th>
<th>D15S508</th>
<th>D15S182</th>
<th>Obs</th>
<th>Exp</th>
<th>Var (Obs - Exp)</th>
<th>χ² 1 df</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>11.26</td>
<td>11.61</td>
<td>0.505</td>
<td>0.02</td>
<td>ns</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>7.74</td>
<td>8.76</td>
<td>4.71</td>
<td>0.22</td>
<td>ns</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>1</td>
<td>6</td>
<td>45.85</td>
<td>51.68</td>
<td>24.87</td>
<td>1.37</td>
<td>0.48</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>1</td>
<td>6</td>
<td>15.99</td>
<td>15.67</td>
<td>3.60</td>
<td>0.02</td>
<td>ns</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>2</td>
<td>7</td>
<td>31.84</td>
<td>32.01</td>
<td>10.08</td>
<td>0</td>
<td>ns</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>6</td>
<td>2</td>
<td>4.25</td>
<td>9.13</td>
<td>3.66</td>
<td>6.50</td>
<td>0.02</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>1</td>
<td>7</td>
<td>26.56</td>
<td>24.35</td>
<td>9.44</td>
<td>0.51</td>
<td>0.96</td>
</tr>
<tr>
<td>8</td>
<td>6</td>
<td>1</td>
<td>7</td>
<td>3.58</td>
<td>8.73</td>
<td>3.37</td>
<td>6.99</td>
<td>0.016</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>2</td>
<td>7</td>
<td>11.75</td>
<td>7.66</td>
<td>3.0</td>
<td>5.56</td>
<td>0.04</td>
</tr>
</tbody>
</table>

Markers are shown in order. H, haplotypes for the combination D15S214/D15S508/D15S182; only haplotypes with frequencies above 0.03 are included. 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 two sided; corrected nominal alpha was set at 0.006 (corrected for nine tests)

*Allele 3 = 262 bp; allele 6 = 268 bp
*Allele 1 = 159 bp; allele 2 = 161 bp
*Allele 4 = 283 bp; allele 6 = 287 bp; allele 7 = 289 bp

REFERENCE

1 Snowling MJ. From language to reading and dyslexia. Dyslexia 2001;7:37–46.
Background: Based on reported familial patterns, inheritance of a predisposition of developing Barrett’s oesophagus (BO) and oesophageal adenocarcinoma (OAC) likely follows an autosomal dominant model of inheritance.

Aims: We analysed the phenotypic features of 70 familial BO/OAC families accrued for the purpose of initiating a linkage study to search for genes that contribute to susceptibility for BO/OAC.

Methods: Families with young or familial BO/OAC were recruited from participating institutions and self-referral from advertisement.

Results: A total of 70 families (173 affected and 784 unaffected individuals) were recruited into this study. Mean ages of diagnosis of BO and OAC among males were 50.6 and 57.4 years, respectively; among females, 52.1 and 63.5 years, respectively. The standardised incidence ratio (SIR) of cancers other than OAC/OGJAC, among probands was 0.71.

Conclusions: This is the largest series of families with BO/OAC yet reported, features of which are consistent with inherited germline predisposition. Further, the SIR of cancers other than OAC/OGJAC was 0.71 among 70 probands, indicating these individuals were not more likely to develop non-OAC cancers.
A locus on 15q15-15qter influences dyslexia: further support from a transmission/disequilibrium study in an Italian speaking population

C Marino, R Giorda, L Vanzin, M Nobile, M L Lorusso, C Baschirotto, L Riva, M Molteni and M Battaglia

*J Med Genet* 2004 41: 42-46
doi: 10.1136/jmg.2003.010603

Updated information and services can be found at:
http://jmg.bmj.com/content/41/1/42

**References**

This article cites 37 articles, 8 of which you can access for free at:
http://jmg.bmj.com/content/41/1/42#BIBL

**Email alerting service**

Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

**Topic Collections**

Articles on similar topics can be found in the following collections

- Epidemiology (630)
- Genetic screening / counselling (886)

**Notes**

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/