Demographic and phenotypic features of 70 families segregating Barrett’s oesophagus and oesophageal adenocarcinoma


Background: Barrett’s oesophagus (BO) also termed metaplastic columnar lined oesophageal epithelium, is believed to result from a continual reparative response to chronic reflux of gastric contents. Although traditionally considered to be an acquired disorder, with several epidemiological risk factors involved, it is now recognised that there is a genetic component, and that this is likely to be autosomal dominant. Clustering of BO and of oesophageal adenocarcinoma (OAC) or oesophagogastric junctional adenocarcinoma (OGJAC) has been shown in a number of studies.

Methods: We investigated a large series of BO/OAC families to find evidence that a subset of BO/OAC is the result of a hereditary predisposition, and explored the potential of performing a linkage study with the families identified to date.

Results: Of 957 individuals in 70 families, 173 had a reported diagnosis of BO or OAC/OGJAC: 101 had BO only, 52 had OAC/OGJAC, and 20 had both BO and OAC/OGJAC. There were 133 affected males and 40 affected females, a male:female ratio of 3.3:1. In addition, 124 participants (12.9%) had a reported diagnosis of cancer other than OAC. Of these, 15 did (affected) and 109 did not (unaffected) have a diagnosis of BO or OAC. A cancer other than OAC was found in 13.9% of unaffected and 8.7% of affected individuals.

Conclusion: It is widely accepted that the majority of BO and OAC are sporadic, although familial clustering of BO and OAC has been recognised for at least three decades. Environment and lifestyle undoubtedly play a role in development of BO and OAC, and may affect the penetrance of the putative inherited factor. Although our 70 BO probands were not more likely to develop non-OAC/OGJAC cancers than normal, over a third had developed either OAC or OGJAC, and might have gone on to develop others. We intend to continue recruitment and initiate formal linkage studies to identify the causative gene(s).

Abbreviations: BO, Barrett’s oesophagus; FBO, familial BO; GORD, gastro-oesophageal reflux disease; OGJAC, oesophagogastric junctional adenocarcinoma; OAC, oesophageal adenocarcinoma; SIR, standardised incidence ratio
METHODS

Families were recruited from the following institutions: The Ohio State University Medical Center including the James Cancer Hospital; University Hospitals of Cleveland; Mercy Medical Center Northwest; Cleveland Clinic Foundation; and Vanderbilt University Medical Center and VA Tennessee Valley Health Care Authority (institutions forming a Familial Barrett Oesophagus Consortium). Eligible families were those with at least two members affected with BO and/or OAC/OGJAC, or a single affected individual diagnosed with BO under the age of 40 years, or a single affected individual diagnosed with OAC under the age of 50 years. The proband (the first patient to come to the attention of the researchers) of each family was approached through self or physician referral. Informed consent, blood samples, release of medical records and tissue samples (when available), and permission to contact relatives were obtained under the umbrella of protocols approved by the respective institutional review boards. We contacted probands who did not live locally, and interested family members, by telephone to obtain their informed consent, and we mailed release forms for medical records and tissue. Records and tissue samples of deceased individuals were released through the consent of the next of kin. Both affected and unaffected family members were recruited. Family history information was obtained through interview with the proband, and a family pedigree was created.

Pathology and/or endoscopy reports were reviewed by a study investigator and tissue samples were reviewed by a pathologist (Wendy Frankel, MD; Joseph E Willis, MD; Kay Washington, MD, PhD; Steve Schultenover, MD) when available. For probands, the diagnosis of BO was made by the presence of intestinal metaplasia with goblet cells on biopsies obtained from the esophagus. Relatives were considered affected by proband report and confirmed by pathology or tissue review when possible. Clinical and pedigree data were entered into a Progeny Enterprise database (Genetic Data Systems, LLC, South Bend, IN, USA) and exported into Microsoft Excel for descriptive statistics. In addition, the PedInfo program (SAGE: Statistical Analysis for Genetic Epidemiology, Release 4.0, 2001, Department of Epidemiology and Biostatistics, Rammelkamp Center for Education and Research, Case Western Reserve University, Cleveland, OH, USA) was used to provide descriptive statistics on pedigree size and structures.

Mean ages of diagnosis of reported and confirmed diagnoses of BO and OAC were calculated for males and females for whom age of diagnosis data were available. Twelve individuals with a young age of diagnosis were not included in the mean age of diagnosis calculations.

A standardised incidence ratio (SIR) was calculated to compare observed non-OAC or OGJAC invasive cancer incidence to expected non-OAC/OGJAC invasive cancer incidence among the 70 FBE probands. For each of the 70 probands, expected lifetime invasive cancer risk was calculated by adding the annual sex and age specific (18 age groups) US invasive cancer incidence rates corresponding to each year of each proband’s life, until the end of 2001. US invasive cancer incidence rates were derived from the Surveillance, Epidemiology, and End Results (SEER) Program of the National Cancer Institute. For years of life between 1935 and 1973, for which there are no corresponding age specific US SEER cancer incidence rates, average annual age-specific rates (based on 5 year periods) from the Connecticut Tumor Registry were used, and for the years 2000 and 2001, 1999 SEER cancer incidence rates were used. For years of life prior to 1935, average annual Connecticut cancer incidence rates (based on the period from 1935 to 1939) were used.

Simulations (power calculations for gene mapping) were performed using Simlink based on three “hypothetical” sets of individuals. Group 1 includes all pedigree members whose DNA has already been obtained to date, and non-sampled connecting individuals, corresponding to 33 pedigrees with 104 sampled subjects. Group 2 includes individuals in Group 1 plus individuals whose DNA has not yet been collected but could readily be collected because they have personally contacted one of the study sites, corresponding to 43 pedigrees with 202 sampled subjects. Group 3 includes individuals in Group 2 plus all remaining pedigree members who were alive at the last contact with the family, and assumes all these individuals can be sampled. This group includes 52 pedigrees with 314 sampled subjects.

Individuals were considered affected if they had a reported diagnosis of BO or OAC, regardless of whether the diagnosis has been confirmed. The expected lod score was computed for each pedigree, under four realistic recessive and four realistic dominant inheritance model assumptions, in order to estimate the probability of achieving a maximum lod score greater than 2.0 or 3.0 in a parametric linkage analysis. The genetic models assume a two allele locus with susceptibility allele D and normal allele d, and are described by the susceptibility allele frequency, p, and the disease penetrances fdd, fDd, and fDD. The dominant models were as follows.

DOM1: p = 0.01, fDd = 0.2, fDD = 0.008 corresponding to \( \lambda_0 = 3.6 \) and a prevalence of 1.2%: DOM2: p = 0.001, fDd = 0.45, fDD = 0.0043 corresponding to \( \lambda_0 = 8.4 \) and a prevalence of 0.5%; DOM3: same as DOM1 for males, while for females fDd = fDD = 0.1; and DOM4: same as DOM2 for males, while for females fDd = fDD = 0.225. The recessive models were REC1: p = 0.07, fDd = 0.4, fDD = 0.008, corresponding to \( \lambda_0 = 3.15 \) and a prevalence of 1.0%; REC2: p = 0.05, fDd = 0.55, fDD = 0.0044 corresponding to \( \lambda_0 = 8.07 \) and a prevalence of 0.5%; REC3: same as REC1 for males, while for females fDd = fDD = 0.2; and REC4: same as REC2 for males, while for females fDd = fDD = 0.275. For all models, the expected lod score was computed assuming a single marker locus with four equally frequent alleles (average PIC of 0.7). This represents typical marker informativity for microsatellite markers, although the information may actually be higher in a multipoint analysis. A recombination fraction (\( \theta \)) of 0, 0.05, or 0.1 between the marker and the trait locus was used. Simulations were performed under the assumption that the probability that a particular pedigree is segregating the major locus is \( \pi \), where \( \pi \) has values of 1.0 and 0.5 corresponding to the entire sample, and that approximately half the sample is linked to the major locus specified in the genetic model, respectively. Model parameters for the disease locus, including the allele frequency and penetrance parameters, were chosen to span the range of \( \lambda_0 \) and the population prevalence consistent with the values computed in the Chak al study, including values of \( \lambda_0 \) between 3.0 and 8.0, corresponding to a population prevalence of 1% and 0.5%, respectively. As the true underlying disease model is unknown, simple models were simulated, such that the disease penetrance was not allowed to depend on age.

RESULTS

Pedigrees were drawn and data entered for 70 eligible families. Four representative families are presented in fig 1. Among these families, 957 individuals for whom at least present age, age at death, year of birth, or age of diagnosis was available were used for analyses. A total of 173 individuals had a reported diagnosis of BO or OAC/OGJAC (table 1). Of these, 119 (68%) were confirmed by medical
record, pathology report, and/or tissue review. The remaining 54 affected individuals remain unconfirmed due to unavailability of the appropriate medical or pathology records. Individuals for whom medical or pathology records refuted the reported diagnosis were not considered affected. If the refuted diagnosis made the family ineligible based on the above criteria, the family was not included in the analyses (this happened in only three instances). Of the 173 affected

![Figure 1](image)

**Figure 1** Four of 70 representative familial BO/OAC families. (A) Single affected male diagnosed with OAC at 42 years of age with multiple family members with GORD; (B) proband with BO and OAC at 75 years, with maternal grandmother diagnosed with OAC at 79 years, and maternal cousin diagnosed with OGJAC at 69 years; (C) proband diagnosed with BO at 21 years, brother with BO at 40 years, father with OAC at 44 years, daughter and paternal cousin with GORD; (D) proband diagnosed with BO at 69 years, dizygotic twin brother diagnosed with BO at 67 years, sister and parents with GORD.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Number of affected and unaffected male and female subjects and average ages of diagnosis for those accrued irrespective of age of diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unaffected (n)</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td>407</td>
</tr>
<tr>
<td>Females</td>
<td>377</td>
</tr>
</tbody>
</table>
individuals, 101 had BO only, 52 had OAC/OGJAC, and 20 had both BO and OAC/OGJAC. In total, 133 affected individuals were male, and 40 were female, for a male:female ratio of 3.3:1.

DNA samples were obtained from peripheral blood lymphocytes and/or paraffin embedded or fresh frozen tissue for 174 individuals. Of these, 103 had a diagnosis of BO or OAC/OGJAC, and 71 were unaffected. An additional 64 individuals are reasonably expected to be recruited because they have expressed verbal or written interest in the study, or have already been sent recruitment materials but have not yet returned blood and/or released tissue samples.

Age of diagnosis data were available for 154 of the 173 affected individuals and these individuals were used to calculate average ages of diagnosis of BO and OAC among males and females (table 1). Twelve individuals were recruited on the basis of young age at diagnosis alone (under 40 years with a diagnosis of BO, or under 50 years with a diagnosis of OAC) and these individuals were excluded from the mean age of diagnoses calculations. Among males, the average ages of diagnosis of BO and OAC were 56.0 and 57.4 years, respectively. Among females, the average ages of diagnosis of BO and OAC were 52.1 and 63.5 years, respectively. The weighted average age for BO was 51.0 years, and for OAC was 58.2 years.

Probands were asked to report cancers other than OAC occurring in themselves or their relatives (table 2). In total, 124 (12.9%) participants had a reported diagnosis of cancer other than OAC. Of these, 109 were unaffected (did not have a diagnosis of BO or OAC) and 15 were affected (did have a diagnosis of BO or OAC). Further, 13.9% of unaffected and 8.7% of affected individuals had a cancer other than OAC.

The SIR of observed non-OAC invasive cancer incidence among the 70 FBE probands was calculated (excluding the diagnosis of non-melanoma skin cancer). Seven probands were diagnosed with a cancer other than OAC/OGJAC or non-melanoma skin cancer. The cancers reported among these seven individuals were colon, breast, pancreatic, testicular, prostate, and tongue cancer, and leukaemia. The number of diagnoses of other invasive cancers, including gastric cancer, among FBE probands was less than that expected, based on US and Connecticut invasive cancer rates (observed = 7; expected = 9.9; SIR = 0.71). However, using rates based on
than 10% in all models for group 1, less than 16% in all models for group 2, and less than 36% in all models for group 3. As an example, fig 3 illustrates the power to detect linkage for model DOM3. Therefore, although we have sufficient power assuming sample homogeneity, our goal is to continue accrual of families with BO and/or OAC for the purpose of identifying gene(s) causing susceptibility to these conditions.

**DISCUSSION**

It is widely accepted that the majority of BO and OAC are sporadic, although familial clustering of BO and OAC has been recognised for at least 3 decades. In a hospital based clinical epidemiological series, Chak et al showed that up to 20% of all BO presentations were familial and therefore might have a genetic component, although a causative gene has yet to be identified. Based on the reported familial patterns of BO/OAC (table 3) as well as the patterns of inheritance of the families in our series, it is likely that inheritance of BO/OAC predisposition follows an autosomal dominant model of inheritance. The function of the putative susceptibility gene(s) for BO/OAC may be broad, ranging from that of a tumour suppressor gene to that of a controller of muscle tone of the lower oesophageal sphincter. In the tumour suppressor gene model, germline mutations in the gene predispose to neoplasia, but cancer does not develop until the second allele is lost, mutated or silenced by environmental factors playing a larger role in the development of BO and OAC, and may affect the penetrance of the putative inherited factor. With strictly autosomal dominant inheritance with 100% penetrance and no gender differences, the sum of gender ratios across all reported families is 4:3 male:female.

Although it would be tempting to perform a segregation analysis to help determine a genetic aetiology for familial BO, the fact that nearly all of the affected individuals in this study were identified as part of the pedigree ascertainment means that little information for a segregation analysis is available from these families once the ascertainment scheme has been taken into account. Our consortium plans to systematically and prospectively recruit consecutive persons identified with BO and OAC at participating institutions to enable a segregation analysis in the future.

The reported average ages of diagnosis of sporadic BO and OAC are estimated to be approximately 63 and 64 years, respectively. The average ages of diagnosis of BO seen in our series and in some of the previously reported families are considerably younger than these (tables 1 and 3). Younger ages of onset are consistent with an inherited germline predisposition and the two hit hypothesis. Individuals born with one “hit” are one step closer to the development of precancer and cancer and are likely to develop their second “hit” earlier in life than an individual without an inherited predisposition.

Environment and lifestyle undoubtedly play a role in development of BO and OAC, and may affect the penetrance of the putative inherited factor. With strictly autosomal dominant inheritance with 100% penetrance and no gender differences, the sum of gender ratios across all reported families is 4:3 male:female, which is similar to the gender ratios reported with BO/OAC in the general population. This may represent biased ascertainment of male probands in our series of families, dilution of our familial population with truly sporadic cases of BO/OAC, or additional genetic and environmental factors playing a larger role in the development of BO/OAC in males.

An association between Barrett’s oesophagus and colorectal cancer has been proposed. Sontag et al reported that of 65 patients with sporadic BO, 29 (44.6%) were also diagnosed with colonic tumours (19 benign and 10 malignant). This number was significantly higher than their control populations in which only 7.7% of symptomatic individuals seen for colonoscopy and 11.3% of symptom free patients with occult blood in the stool had been diagnosed with colonic tumours. Howden and Horhnung performed a systematic review of the literature and concluded that patients with BO have an increased risk of developing colon cancer with an odds ratio of 5.19 (p<0.0001), but other researchers have not found an association. If such an association does exist, it may be due to shared risk factors in the development of BO and colon

### Table 3 Summary of published familial BO/OAC cases

<table>
<thead>
<tr>
<th>Author [ref]</th>
<th>Case(s)</th>
<th>BO [n]</th>
<th>OAC [n]</th>
<th>BO</th>
<th>OAC</th>
<th>Gender ratio</th>
<th>AD/AR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drovetic et al, 2003</td>
<td>70 families</td>
<td>121</td>
<td>62</td>
<td>51</td>
<td>58.2</td>
<td>3.3:1</td>
<td>Y/YN</td>
</tr>
<tr>
<td>Everhart et al, 1978</td>
<td>1 family</td>
<td>3</td>
<td>0</td>
<td>23.6</td>
<td>—</td>
<td>All M</td>
<td>Y/N</td>
</tr>
<tr>
<td>Gelfand et al, 1983'</td>
<td>56 pair</td>
<td>2</td>
<td>0</td>
<td>67</td>
<td>—</td>
<td>Bath F</td>
<td>Y/Y</td>
</tr>
<tr>
<td>Crabbe et al, 1985'</td>
<td>1 family</td>
<td>2</td>
<td>0</td>
<td>59.5</td>
<td>—</td>
<td>1:1</td>
<td>Y/N</td>
</tr>
<tr>
<td>Prior &amp; Whorwell, 1986</td>
<td>56 pair</td>
<td>2</td>
<td>0</td>
<td>66</td>
<td>—</td>
<td>Bath F</td>
<td>Y/Y</td>
</tr>
<tr>
<td>Jochum et al, 1992'</td>
<td>1 family</td>
<td>6</td>
<td>3</td>
<td>43.6</td>
<td>74</td>
<td>All F</td>
<td>Y/N</td>
</tr>
<tr>
<td>Fahmy &amp; King, 1993'</td>
<td>4 families</td>
<td>8</td>
<td>2</td>
<td>62</td>
<td>72</td>
<td>3:2</td>
<td>Y/N</td>
</tr>
<tr>
<td>Eng et al, 1993'</td>
<td>1 family</td>
<td>7</td>
<td>2</td>
<td>54</td>
<td>68</td>
<td>2.5</td>
<td>Y/N</td>
</tr>
<tr>
<td>Payton et al, 1996'</td>
<td>3 families</td>
<td>6</td>
<td>8</td>
<td>62</td>
<td>60</td>
<td>5:3</td>
<td>Y/N</td>
</tr>
</tbody>
</table>

AD, autosomal dominant; AR, autosomal recessive; F, female; M, male; *identical twins.

**Note:** The summary of published familial BO/OAC cases (table 3) as well as the patterns of aggregation of BO and/or OAC for the purpose of identifying gene(s) causing susceptibility to these conditions.
cancer such as increasing age, excessive alcohol consumption, smoking, obesity, and dietary factors. However, oesophageal and colorectal adenocarcinoma also share somatic genetic changes, such as inactivating mutations in the APC and TP53 tumour suppressor genes and activating mutations in the SRC oncogene, which lead to increased expression of the cyclo-oxygenase 2 (COX-2) enzyme. In our series of families, 13 (1.3%) were diagnosed with colon cancer. However, the SIR of cancers other than OAC or OGJAC among 70 FBE probands was 0.71, indicating that these probands were not more likely to develop non-OAC/OGJAC cancers than expected. The sample size of this analysis is relatively small, making this a crude comparison, but it suggests that other component cancers may not be a part of a hereditary BO/OAC syndrome and that the putative correlation with colon cancer may be due primarily to environmental factors. However, 26 of the 70 probands were diagnosed with OAC/OGJAC, and may not have lived long enough to develop a second cancer. In addition, the expected number of cancers for this analysis may be falsely inflated by using average annual Connecticut cancer incidence rates (based on the period from 1935 to 1939 for the time period prior to 1935, which may have decreased the resulting SIR. It is also possible that gastroenterologists perform screening colonoscopies on persons affected with BO or OAC, and that these persons consequently may be at decreased risk for colon cancer compared to the general population. Therefore, rejection of the null hypothesis may be premature, and more study in this area is warranted.

Recruitment of a series of families with familial BO/OAC is the first step in identifying the causative genetic predisposition. One large family with multiple affected individuals may be sufficient to identify a causative gene through linkage analysis, but most recruited families are small with few affected members. Our Familial Barrett’s Oesophagus Consortium has the potential to perform informative linkage analysis with the recruited families, if genetic homogeneity is assumed. However, because familial Barrett oesophagus is more likely to be a complex disease with genetic heterogeneity, our aim is to continue recruitment until nearly double the present accrual before we initiate formal linkage studies. The results of this study show that the pedigree structures we have identified have sufficient power for future success in this goal. Following the paradigm of other well characterised hereditary cancer syndromes, identifying a gene or genes that predispose to the development of these conditions will enable cancer genetics professionals to offer genetic counseling and testing for families and individuals at risk. Asymptomatic predisposition testing is a powerful tool to use in making decisions about lifestyle and screening behaviours, and would allow prevention to be targeted to those at high risk.

ACKNOWLEDGEMENTS

This is a Familial Barrett Oesophagus Consortium study. We thank Shannon Edwards BS and Jill Griesbach BS for technical assistance. This study is partially funded by US National Institutes of Health Grants R03DK61426 (to AC, WMG, GE, and CF), HG01577 (to KABG), and K24DK02800 (to AC); National Cancer Institute Grants R21CA130172 (to CE and WMG), Vadim V. Poynton AR48041/02A1/02-5 (to CE), and P50CA16058 (to The Ohio State University Comprehensive Cancer Center). Some of the results of this paper were obtained by using the program package SAGE, which is supported by a U.S. Public Health Service Resource Grant (1 P41 RR03655) from the National Center for Research Resources. CE is a Doris Duke Distinguished Clinical Scientist.

Authors’ affiliations

C M Drovdlic, C Eng, Clinical Cancer Genetics Program, The Ohio State University, Columbus, OH, USA; James Cancer Hospital, Columbus, OH, USA; Solove Research Institute and Comprehensive Cancer Center; Division of Human Genetics, Department of Internal Medicine, Cleveland, OH, USA

J L Fisher, S Lemeshow, Center for Biostatistics, The Ohio State University, Columbus, OH, USA

A Chak, W Brock, Division of Gastroenterology, University Hospitals of Cleveland, Cleveland, OH, USA

K A B Goddard, Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, OH, USA

L Chessler, J F King, Department of Gastroenterology, Mercy Medical Center, Canton, OH, USA

J Richter, G W Falk, Department of Gastroenterology and Hepatology, Cleveland Clinic Foundation, Cleveland, OH, USA

D K Johnston, W M Grady, Division of Gastroenterology, Department of Internal Medicine, Vanderbilt University, Nashville, TN, USA

C Eng, Cancer Research UK Human Cancer Genetics Research Group, University of Cambridge, Cambridge, UK

REFERENCES


Demographic and phenotypic features of 70 families segregating Barrett's oesophagus and oesophageal adenocarcinoma


*J Med Genet* 2003 40: 651-656
doi: 10.1136/jmg.40.9.651

Updated information and services can be found at:
http://jmg.bmj.com/content/40/9/651

*These include:*

**Supplementary Material**
Supplementary material can be found at:
http://jmg.bmj.com/content/suppl/2003/09/10/40.9.651.DC1

**References**
This article cites 19 articles, 3 of which you can access for free at:
http://jmg.bmj.com/content/40/9/651#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.

**Errata**
An erratum has been published regarding this article. Please see next page or:
/content/41/1/46.full.pdf

**Topic Collections**
Articles on similar topics can be found in the following collections

- Oesophageal cancer (5)
- Epidemiology (630)

**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/
Developmental 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. 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. 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, 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. Contradicting a culturally bound identity orthographic rules specific to a given language, when homogenous diagnostic approaches, including parametric and model free linkage analyses, as well as family based association designs. Overall, promising findings have been obtained for regions on chromosomes 1, 2, 3, 6, 15 and 18. 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. These results, however, were not confirmed by two similar, successive studies. In a more recent study Smith et al 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 Fulk regression approach and a quantitative definition of dyslexia. Grigorenko et al 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

**Key points**

- We reproduced previous findings of an involvement of the 15q15-15qter chromosomal region in developmental dyslexia in an Italian speaking population by a transmission/disequilibrium approach, implementing analyses of single marker and multimarker haplotypes in a sample of 121 parent-offspring families. We found evidence for linkage disequilibrium with the combination D15S214/D15S508/D15S182 (global $\chi^2 = 25.451$, 9 df, two tailed p = 0.005) using TRANSMIT version 2.5.4; neither single marker nor two marker analyses yielded significant results. Our finding provides further support for the hypothesis of a heritable basis for a unitary, neurobiological process leading to dyslexia, independent of linguistic differences.

**Abbreviations:** TDT, transmission disequilibrium test
D15S132 and D15S143, using both parametric and non-parametric methods.\textsuperscript{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.\textsuperscript{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,\textsuperscript{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 of the phonemes 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\textsuperscript{‘} 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,\textsuperscript{2} 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,\textsuperscript{14} 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,\textsuperscript{12,13} and the revised Wechsler intelligence scale for children.\textsuperscript{14} 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.\textsuperscript{11}

- Single word/non-word reading: “Batteria per la Valutazione della Dislessia e Disortografia Evolutiva” (Battery for the assessment of developmental reading and spelling disorders).\textsuperscript{23} 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; or
- 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.\textsuperscript{33} 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\textsuperscript{18,20,21}; six microsatellite markers defining allelic variation across a 9 Mb region were then chosen: D15S214 (37979 kb), D15S5994 (38161 kb), D15S308 (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.\textsuperscript{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.\textsuperscript{26,27} The TDT for individual markers with multiple alleles was performed using the FBAT program.\textsuperscript{37} (available number [http://biosun1.harvard.edu/~bat/)]. The empirical variance option, as appropriate in the presence of a linkage when more than one sibling per family is included in the analyses.\textsuperscript{28} 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 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 associations when the true mechanism of transmission is unknown. For haplotype transmission of multiple markers we used the program TRANSMIT version 2.5.4. (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 absence of a linkag. 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 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 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/); thus, $\chi^2$ tests were considered on restricted analyses only. Two tailed p 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 p 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/cephpd/). 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 = 9, p = 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 15q15-15pter 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
decrease of p values around the area in linkage disequilibrium rather than a complete loss.\textsuperscript{39} Terwilliger et al\textsuperscript{32} 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 al\textsuperscript{32} 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.\textsuperscript{39} 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 more general elements that suggest a cautious approach to these results. First, the 2.5 value as the threshold for informative transmissions (Var O–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.\textsuperscript{39–41}

The second aim of this study was to explore the hypothesis of a heritable basis for a unitary, neurobiological process leading to dyslexia,\textsuperscript{4} 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.

\textbf{References}


\textbf{Table 3} Estimated $\chi^2$ test for the most significant three marker combinations using TRANSMIT

<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>$\chi^2$</th>
<th>df</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 3 2</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>11.26</td>
<td>11.61</td>
<td>0.55</td>
<td>0.2</td>
<td></td>
<td>ns</td>
</tr>
<tr>
<td>2 3 2</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>7.74</td>
<td>8.76</td>
<td>4.71</td>
<td>0.22</td>
<td></td>
<td>ns</td>
</tr>
<tr>
<td>3 3 2</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>
<td></td>
</tr>
<tr>
<td>4 6 1</td>
<td>1</td>
<td>6</td>
<td>15.99</td>
<td>15.67</td>
<td>5.60</td>
<td>0.02</td>
<td>0.02</td>
<td></td>
<td>ns</td>
</tr>
<tr>
<td>5 3 2</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>31.84</td>
<td>32.01</td>
<td>10.08</td>
<td>0</td>
<td></td>
<td>ns</td>
</tr>
<tr>
<td>6 2 2</td>
<td>3</td>
<td>6</td>
<td>4.25</td>
<td>9.13</td>
<td>3.66</td>
<td>0.02</td>
<td>0.02</td>
<td></td>
<td>ns</td>
</tr>
<tr>
<td>7 3 1</td>
<td>7</td>
<td>6</td>
<td>26.56</td>
<td>24.35</td>
<td>9.44</td>
<td>0.016</td>
<td>0.016</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 6 1</td>
<td>1</td>
<td>7</td>
<td>3.88</td>
<td>8.73</td>
<td>3.37</td>
<td>6.99</td>
<td>0.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 3 2</td>
<td>7</td>
<td>3</td>
<td>11.75</td>
<td>7.66</td>
<td>3.0</td>
<td>5.56</td>
<td>0.04</td>
<td></td>
<td></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 = 289 bp
†Allele 1 = 159 bp; allele 2 = 161 bp
‡Allele 4 = 283 bp; allele 6 = 287 bp; allele 7 = 289 bp

www.jmedgenet.com


