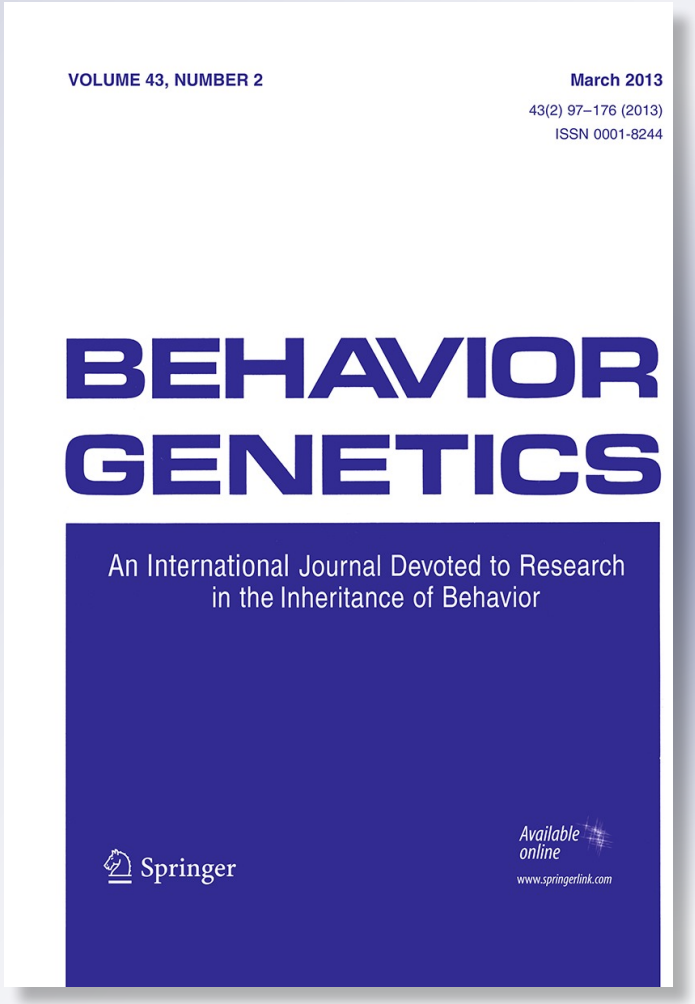
Xq27 FRAXA Locus is a Strong Candidate for Dyslexia: Evidence from a Genome-Wide Scan in French Families

M. Huc-Chabrolle, C. Charon, A. Guilmatre, P. Vourc'h, G. Tripi, M. A. Barthez, E. Sizaret, R. A. Thepault, S. Le Gallic, et al.

Behavior Genetics
An International Journal Devoted to Research in the Inheritance of Behavior

ISSN 0001-8244
Volume 43
Number 2

Behav Genet (2013) 43:132-140
DOI 10.1007/s10519-012-9575-5



Springer

Behav Genet (2013) 43:132–140
DOI 10.1007/s10519-012-9575-5

ORIGINAL RESEARCH

Xq27 FRAXA Locus is a Strong Candidate for Dyslexia: Evidence from a Genome-Wide Scan in French Families

M. Huc-Chabrolle · C. Charon · A. Guilmatre · P. Vourc'h · G. Tripi · M. A. Barthez · E. Sizaret · R. A. Thepault · S. Le Gallic · J. Hager · A. Toutain · M. Raynaud · C. Andres · D. Campion · F. Laumonnier · F. Bonnet-Brilhault

Received: 22 August 2012 / Accepted: 8 December 2012 / Published online: 11 January 2013

Abstract Dyslexia is a frequent neurodevelopmental learning disorder. To date, nine susceptibility loci have been identified, one of them being DYX9, located in Xq27. We performed the first French SNP linkage study followed by candidate gene investigation in dyslexia by studying 12 multiplex families (58 subjects) with at least two children affected, according to categorical restrictive criteria for phenotype definition. Significant results emerged on Xq27.3 within DYX9. The maximum multipoint LOD score reached 3,884 between rs12558359 and rs454992. Within this region, seven candidate genes were investigated for mutations in exonic sequences (*CXORF1*, *CXORF51*, *SLITRK2*, *FMR1*, *FMR2*, *ASFMR1*, *FMR1NB*), all having a role during brain development. We further looked for 5′UTR trinucleotide repeats in *FMR1* and *FMR2* genes. No mutation or polymorphism co-segregating with dyslexia was found. This finding in French families with Dyslexia showed significant linkage on Xq27.3 enclosing FRAXA, and consequently confirmed the DYX9 region as a robust susceptibility locus. We reduced the previously described interval from 6.8 (DXS1227–DXS8091) to 4 Mb also disclosing a higher LOD score.

Keywords Dyslexia · Linkage study · Multiplex families · Fmr1 · Dyx 9 loci

Edited by Stacey Cherny.

M. Huc-Chabrolle and C. Charon contributed equally to this study.

Electronic supplementary material The online version of this article (doi:10.1007/s10519-012-9575-5) contains supplementary material, which is available to authorized users.

M. Huc-Chabrolle · P. Vourc'h · R. A. Thepault · A. Toutain · M. Raynaud · C. Andres · F. Laumonnier · F. Bonnet-Brilhault
UMR INSERM U930, François Rabelais University, Tours, France

M. Huc-Chabrolle · G. Tripi · F. Bonnet-Brilhault
Department of Child Psychiatry, University Hospitals of Tours, Tours, France

M. Huc-Chabrolle · M. A. Barthez · E. Sizaret · F. Bonnet-Brilhault
Reference Centre for Learning Disability, University Hospitals of Tours, Tours, France

M. Huc-Chabrolle (✉)
Centre Universitaire de Pédopsychiatrie, CHRU de Tours, Hôpital Bretonneau, 2, Bd Tonnellé, 37044 Tours Cedex 9, France
e-mail: m.huc-chabrolle@chu-tours.fr

C. Charon · J. Hager
CEA/Centre National de Génotypage, Évry, France

A. Guilmatre · S. Le Gallic · D. Campion
INSERM U 1079, Rouen, France

P. Vourc'h · R. A. Thepault · C. Andres
Department of Molecular Biochemistry, University Hospitals of Tours, Tours, France

G. Tripi
University of Palermo, Palermo, Italy

A. Toutain · M. Raynaud · F. Laumonnier
Department of Human Genetics, University Hospitals of Tours, Tours, France

Introduction

Dyslexia is a complex neurodevelopmental disorder that consists of a specific learning disability with a neurological

origin (Habib 2000). It affects 5–10 % of school age children (Shaywitz 1998) and boys are more often affected than girls (sex ratio around 1.6) (Flannery et al. 2000; Rutter et al. 2004). These learning difficulties are unexpected in view of other cognitive abilities and the provision of effective classroom instruction. Neurobiological investigations using post-mortem brain specimens (Galaburda et al. 1985; Galaburda et al. 1994) and, more recently, brain morphometry (Eliez et al. 2000; Brown et al. 2001), diffusion tensor MRI imaging (Klingberg et al. 2000; Vandermosten et al. 2012) and functional MRI (Shaywitz et al. 2002) suggest that disruption of parieto-temporal and, in particular, occipito-temporal systems in dyslexic readers underlie a failure to develop skilled reading. The report that dyslexia is both a familial and heritable disorder was published early on (Hallgren 1950), and was confirmed by twin studies (DeFries and Alarcon 1996). The risk of dyslexia in siblings of affected children is increased more than 3.5 fold (Ziegler et al. 2005).

Since the first linkage study, extensive effort has been invested in understanding the mechanisms of heredity in dyslexia. Several groups have therefore investigated heritability, models of transmission, and genetic background of specific difficulties in reading. The most relevant genetic model seems to be a complex polygenic model. Nine loci have been linked to dyslexia on chromosomes 15q21 (DYX1), 6p21-p22 (DYX2), 2p15-16 (DYX3), 6q11-q12 (DYX4), 3p11-q13 (DYX5), 18p11 (DYX6), 11p15 (DYX7), 1p34-p36 (DYX8), and Xq26-q27 (DYX9) (for review Scerri and Schulte-Körne 2010; Poelmans et al. 2011). The first candidate gene proposed was *DYX1C1* (15q21), identified in a kindred with a t(2; 15) (q11; q21) translocation co-segregating with the disorder (Taipale et al. 2003). However, further studies on other populations did not confirm the role of *DYX1C1* in common forms of reading disability. The *ROBO1* gene (3p12-q13) was also proposed as a candidate, after the discovery of a t(3; 8) (p12; q11) translocation in an affected individual (Hannula-Jouppi et al. 2005). Study of the most replicated linkage region, the 6p22 region, suggested the implication of two other genes, i.e., *DCDC2* and *KIAA0319* (Meng et al. 2005; Cope et al. 2005; Schumacher et al. 2006). Most of these putative genes contribute to neuron migration (Pechansky et al. 2010; Massinen et al. 2011; Currier et al. 2011), consistent with anatomical studies that show structural cortical anomalies in dyslexic individuals (Galaburda et al. 1985). Association studies have revealed correlations with genetic variants, but results concerning specific risk factors remain mainly sparse or inconsistent in the different samples (for review Scerri and Schulte-Körne 2010). This may be due to the genetic heterogeneity of the disorder and/or to the different methods of investigation. Indeed, depending on the study design, sporadic or familial cases may be investigated, but the genetic factors underpinning familial dyslexia may not be the same as the genetic factors in sporadic cases. Additionally, despite extensive efforts to propose a consensual definition (Lyon et al. 2003), the criteria for the diagnosis remain confusing, particularly in adults. Furthermore, test results of the different cognitive tasks necessary in reading prove to be highly inter-correlated, and thus probably share most of their genetic factors (Bates et al. 2007), and genetic weight seems to increase for more severe phenotypes (Cope et al. 2005; Schumacher et al. 2006; Francks et al. 2004). The choice of diagnostic criteria in order to select severe and well-characterized phenotypes is therefore essential to increase the power of genetic studies.

No such investigation has yet been performed with French patients, and no linkage study report has explored French families to date. We therefore present the first linkage analyses in a French multiplex family sample with categorical restrictive criteria.

Materials and methods

Subjects

Twelve families with two or more dyslexic children attending the Reference Centre for Language Disorder at the University Hospital of Tours were contacted and asked to participate. Inclusion criteria comprised normal intelligence, normal hearing tests and no associated neurological or psychiatric disorder (including ADHD).

Full pedigrees represented a total of 78 subjects (33 females and 45 males) with 30 founders and 48 non-founders. Average family size was 6.5 (range 3–16) with a generation average size of 2.33 in the following proportions 2 (75 %), 3 (16.7 %), and 4 (8.3 %).

Clinical data and DNA were finally available for 58 subjects 42 of whom were dyslexic, which constituted the effective sample for the study with mostly nuclear families (Fig. 1). Average age for proband generation when recruited was 15.6, 31 children out of 35 were dyslexics 23 males, eight females; sex ratio 2.9; four children were normal readers two males and two females. One family had four affected children, six families had three affected children, four families had two affected children, and one family had only one dyslexic child but father and grandfather were affected too. Among the 26 participants from the parents and grand-parents generations, 10 were assigned the dyslexic status (eight males, two females, sex ratio 4), and two subjects had an ambiguous phenotype with only persistent spelling difficulties reported (one male, one female).

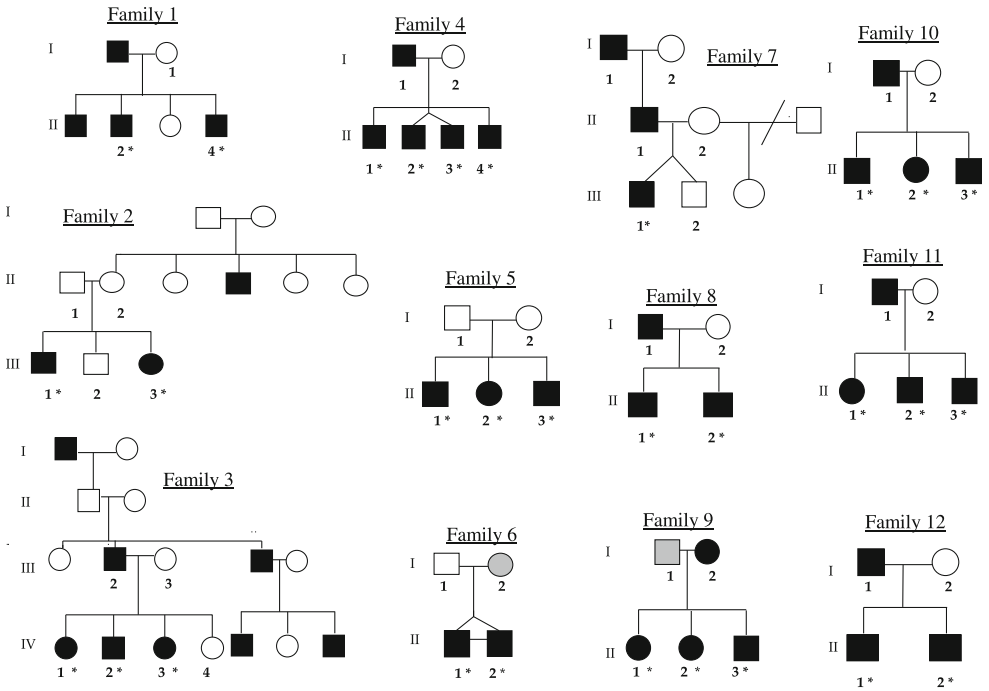


Fig. 1 Families included in the sample. Individuals for which DNA was available are numbered; probands are identified with a *star*. In *white* are depicted non dyslexic individuals. In *black* are depicted dyslexic individuals. In *grey* are depicted ambiguous phenotype individuals with only spelling problems

The study was approved by the local ethics committee and all family members gave informed written consent.

Clinical assessment

Proband generation

The clinical data were collected from the regular diagnostic evaluations in the Reference Centre for Learning and Language Disorders that included oral and written language assessment with French standardized age matched tests (supplementary data) and non verbal skills assessment with WISC III or K-ABC subtests.

Characterisation of the dyslexia categorical phenotype was based on the results in 5 cognitive reading abilities (non-word reading, irregular word reading, reading speed, spelling and phonological skills) obtained with age appropriate French standardized tests (supplementary data). Affected dyslexic status was attributed to individuals scoring at least 2 standard deviations below the norm in at least two of the following tests: spelling or phonological skills and non-word or irregular word reading. Reading speed results were used if the phenotype remained ambiguous using these criteria. In terms of non-verbal abilities, dyslexia was diagnosed if children scored above the 25th percentile in their age group.

Parents and grand-parents generations

Subject status for the grand-parent and parent generations (generation I and II, respectively) was determined retrospectively. They were interviewed at the time of inclusion by two experienced psychiatrists about their history of reading skills and their possible experience of specific education. Affected status was attributed to subjects reporting a diagnosis of reading learning disability in childhood, sometimes with no reading acquisition or need for special reading education, and persistent difficulties in reading or spelling in adulthood, contrasting with proper education and good professional achievement. Ambiguous status was attributed to individuals who reported former difficulties in learning to read but who had achieved a good literacy level, and to the individuals who only reported spelling difficulties. Individuals were classified as unaffected if they reported no former or current difficulties in reading or spelling.

Genotyping panel and quality control

DNA was extracted from peripheral blood sample using a standard protocol.Genome wide scan was carried out on 250 ng of DNA at the French Genotyping Centre using GeneChip® 250 K Affymetrix NSP set. Genotype calls were analysed with the BRLMM algorithm implemented in

Affymetrix GeneChip® Genotyping software (GTYPE 4.0).

Among the 58 DNA samples, 56 could be hybridised on the chip. Most of the genotyped subjects had both parents available except for pedigree 1 and 12, for which one of the founders was missing. The number of offsprings genotyped per family was 2.5 in average (range 1–4). Analyses were therefore performed mostly on nuclear families. Family links were verified by IBS. Genotypes in one sibling was not used for analyses as his brother was a twin, also one individual was eliminated due to low success rate (95 %). The markers were filtered and verified (Wigginton and Abecasis 2005; Purcell et al. 2007; R Development Core Team 2008) for Hardy–Weinberg equilibrium (HWE $\geq$ 0.0001), MAF ($\geq$0.01) and mendelian errors. The mean success rate of the 54 remaining subjects was 0.9909 (SD = 0.007446) and the average success rate of the SNPs was 0.9953 (with SD = 0.01001).

Statistical analyses

Both parametric and non-parametric analyses on all affected (all) were performed with Merlin (multipoint-engine-for rapid likelihood-interference) (Abecasis et al. 2002) on the 54 genotyped individuals, corresponding mostly on the nuclear families of the probands.

We also performed the NPL pairs statistics multipoint analyses as recommended and based on the fact that T pairs has greater power than Tall to detect linkage for recessive than additive and dominant diseases in nuclear families (Kruglyak et al. 1996; Davis and Weeks 1997; Feingold et al. 2000).

Parametric dominant models were investigated as suggested by Muller-Myohsok and Grimm (1999) using a disease allele frequency of 4 % with 0–0.9–0.99 for penetrance of homozygous wild types, heterozygotes and homozygous carriers, and also as previously mentioned by de Kovel et al. (2004) (i.e., 2 % frequency for the disease allele, penetrances 0.02–0.95–1 for males while 0.005–0.85–0.95 for females). Recessive model based on that of de Kovel et al. (2004), was adopted by us with the same allele frequency of 2 %, and 0.02–0.001–1 penetrances for aa, Aa, AA genotypes for the males, and with 0.005, 0.001 and 0.95 for the females, respectively (Fig. 2). The male X-allele locus having only two genotypes, we considered the penetrances of de Kovel (0.02–0.95) with Merlin-in-x algorithm (Abecasis et al. 2002), males being hemizygous at this chromosome. HLOD scores are reported in the results (Pal and Greenberg 2002; Cavalli-Sforza and King 1986). Single point analyses Merlin and Pseudomarker (Terwilliger and Goring 2000) were also performed on the region showing significant linkage.

Merlin modeled linkage disequilibrium (LD) by creating marker clusters using a specified r2, appointed in our analyses as 0.1. We then evaluated every 5 Mb if there was evidence of linkage in the region of interest with our parametric model taking LD into account (Abecasis and Wigginton 2005).

To verify our findings, original phenotypes with family structure were also used with new data set created by permutations taking into consideration the map with authentic order of markers with their original genotypes and alleles' frequencies. We ran 10,000 and 100,000 simulated pedigrees under the assumption that none was linked, and we calculated the empirical *p* value for our recessive model in the region that had shown positive.

In the positive region of linkage, family association studies was investigated by LD mapping with pseudomarker (Ilink, FASTLINK 4.1) computing 2 point LOD score linkage analysis and likelihood ratio tests for linkage and/or LD under our model-based analysis (Terwilliger and Goring 2000; Cottingham et al. 1993; Shugart et al. 2007). Pseudomarker enabled for the conjoint analyses of nuclear pedigrees and singletons. Unaffected singletons were included in the joint statistic for family based association to obtain allele frequencies and estimate the maximised likelihoods. PCA analyses performed with eigenstrat (Price et al. 2006 and Patterson et al. 2006) confirmed that our control population of 273 subjects and the families studied were of the same ethnic origin.

Investigation of chromosomal micro-rearrangements and mutation analyses of candidate genes

To assess the copy number variations (CNV) associated with dyslexia, the genome of one dyslexic proband in every family (12 genomes) has been investigated with the Agilent Sureprint G3 1 M markers chip, providing with a median resolution of 1 Kb. The data were extracted using the Feature Extraction software and analysed with Cytogenomics (Agilent).

Mutation analyses of seven candidate genes were performed by forward and reverse sequencing of the exons, including the exons-introns junctions (ABI 3130xl, Applied). The 5′UTR region were also investigated for the FMR1 and FMR2 genes (primers sequences and PCR conditions available upon request).

Results

Linkage

Linkage analyses under dominant, recessive and non-parametric analyses were not significant on all autosomes.

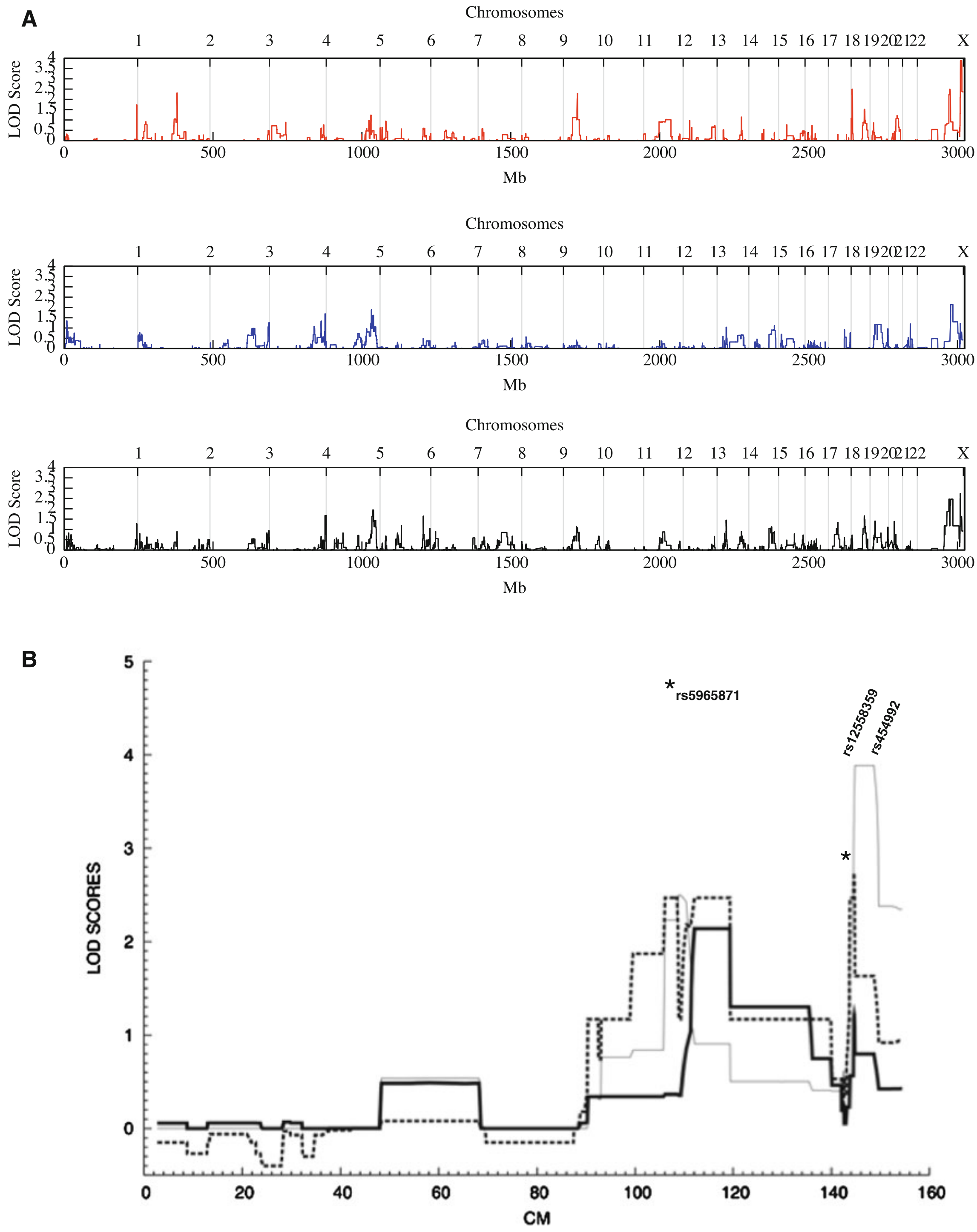


Fig. 2 Results of linkage analysis. **a** Genome wide linkage: parametric dominant model I and recessive model with NPL pairs free model. *Red line* recessive model with sex dependent penetrance: 2 % disease allele frequency and penetrances on autosomes for females: 0.005–0.001–0.95 and males: 0.02–0.001–1 on chromosome X males: 0.02–0.95. *Blue line* parametric dominant model I: 4 % disease allele frequency with 0–0.9–0.99 penetrances in males and females. *Black line* NPL statistics. **b** Chromosome X linkage: parametric dominant model I and recessive model with NPL pairs free model. *Solid black line* parametric dominant model I: 4 % disease allele frequency with 0–0.9–0.99 penetrances in males and females. *Solid fine line* recessive model with sex dependent penetrance: 2 % disease allele frequency and penetrances on chromosome X for females: 0.005–0.001–0.95, males: 0.02–0.95, *dotted line* non parametric linkage (NPL) pairs results

On chromosome X, two notable regions around 112–119 Mb (Xq23–q24) and at 144.7 Mb (Xq27.3) were observed. Only the second region Xq27.3 reached significative LOD score above 3 (Fig. 2a).

NPL all LOD max 2.36 was attained between 112.185 (rs5929497) and 119.477 Mb (rs5910841) with a p value of 0.0005. This region was also revealed with parametric dominant statistics when applying the deleterious allele frequency of 0.04, as the multipoint HLOD reached 2.140, within the same interval of 7.29 Mb, while under parametric recessive model HLODmax 2.5 was achieved between 108.974 (rs5943427) and 109.388 Mb (rs946820) (Fig. 1). The second region depicted by NPL all statistics had a LOD increased above 2.3 from 144.673 (rs5965871) until 144.881 Mb (rs9698110), reaching its maximum of 2.60 (p value of 0.0003) within this interval at 144.7 Mb on locus rs5919606–rs5965631.

The parametric conditions on chromosome X produced a max HLOD score of only 1.2 within 144.734–144.8277 Mb under the first dominant model.

In our multipoint parametric recessive model, HLOD max 3.884 was bounded by the markers rs12558359 (144.969 Mb) and rs454992 (148.885 Mb).

Both Merlin and Pseudomarker single point parametric studies with the recessive model generated identical LOD scores $\geq$3. The results spread from rs12558359 (144.969 Mb) disclosing LOD score of 3.175 ($p = 0.000066$) as far as rs2536561 (147.401 Mb) yielding LOD 3.02 ($p = 0.000097$). Two tightly close SNP, firstly rs1072149 at 145.980 Mb revealed LODmax 3.305 ($p = 0.000048$) and secondly SNP rs2392669 at 146.028 Mb showed LOD 3.004 ($p = 0.000101$).

Our recessive model being established in our families we investigated the NPL pairs statistics, providing it was previously reported to have greater power to detect linkage for recessive forms of diseases compared with NPL all statistics. We found increased NPL LOD scores. NPL pairs LOD upholding a value of 2.47 was also obtained within the same borders 112 and 119 Mb ($p = 0.0004$). The same LOD was also shown between 106.162 (rs5917027) and 108.607 Mb (rs615213), but the highest LOD of 2.74 was found for three consecutive markers including the furthest rs5965631 ($p = 0.0002$) at 144.746 Mb (Fig. 2b). On chromosome X with the dominant model per de Kovel et al. 2004 (35) HLOD did not exceed 1.83.

To attest for linkage at the location that had shown peak significance (HLOD max 3.884) on chromosome X, we computed our recessive model assuming the null hypothesis—(of no linkage)—measured against our observed threshold (3.884) and obtained the empirical p value for 100,000 simulations $p = 0.00002$. HLOD ranged 3.9–4.129 (mean 4.0145) with 0.002 % probability of observing a false hit. Concurrently, as per a threshold of 3 the p value reached 0.00012 on chromosome X (HLOD range 3–4.129, mean: 3.51). Under 10,000 simulations the HLOD max found of 2.6 was well under our peak significance. On the basis of simulation studies with our pedigree structure, we could confirm that our findings were a true linkage in the region surrounding rs12558359.

We also investigated LD in this region. The positive signal with a LODmax was still present within the characterized region, as merlin-LD yielded equivalent LOD scores to the original linkage.

As a French control population was used to test our families association against, we searched for any SNP association. Where linkage was shown to be significative at 144.969 Mb with multipoint (LOD 3.884) and single point parametric studies with LOD score 3.175 ($p = 0.000066$), no signal could be demonstrated. The best results were obtained in the very close vicinity of this SNP rs12558359 (144.969 Mb) with its three consecutive markers but without any strong evidence of association. The LD|Linkage statistics (HRR) of those three markers rs12862591, rs12861185 and rs905089 placed at 144.99 Mb for the first two and 145 Mb for the third one, attained p values of 0.0076–0.007025 and 0.008248, respectively. For the test of linkage allowing for LD (similar to a TDT) at rs12558359 (144.969 Mb) and the three markers previously mentioned the p values achieved 1^{E-4}, 9^{E-5}, 8.8^{E-5} and 1^{E-4}, respectively. We therefore could not demonstrate with the set of markers we had the presence of an association signal.

Nothing significant was found either between 105 and 120 Mb, as only rs7891927 and rs5974267 both a 111.763 Mb displaying NPL pairs LOD 2.29 and HLOD 1.09 revealed LD|Linkage p values of just 0.001047 and 0.0005, whilst in Linkage|LD their respective p values were at most 0.043–0.046.

Array competitive genomic hybridization (CGH) and candidate genes analyses

High resolution CGH array was used to analyse the genome of one proband in every family, providing results for 12 genomes. No copy number variation was found to be associated with dyslexia.

Mutation analysis was investigated on seven candidate genes identified in the region with elevated LOD scores and participating in brain development: *Cxorf1*, *FMR2*, *FMR1*, *SLITRK2*, *ASFMR1*, *FMR1NB*, *Cxorf51* (Fig. 3). Several silent polymorphisms on the exons were found in the population but none of them segregated with dyslexia in the families. *FMR1* and *FMR2* 5′UTR trinucleotide repeats were also studied by PCR but no abnormal amplification was identified.

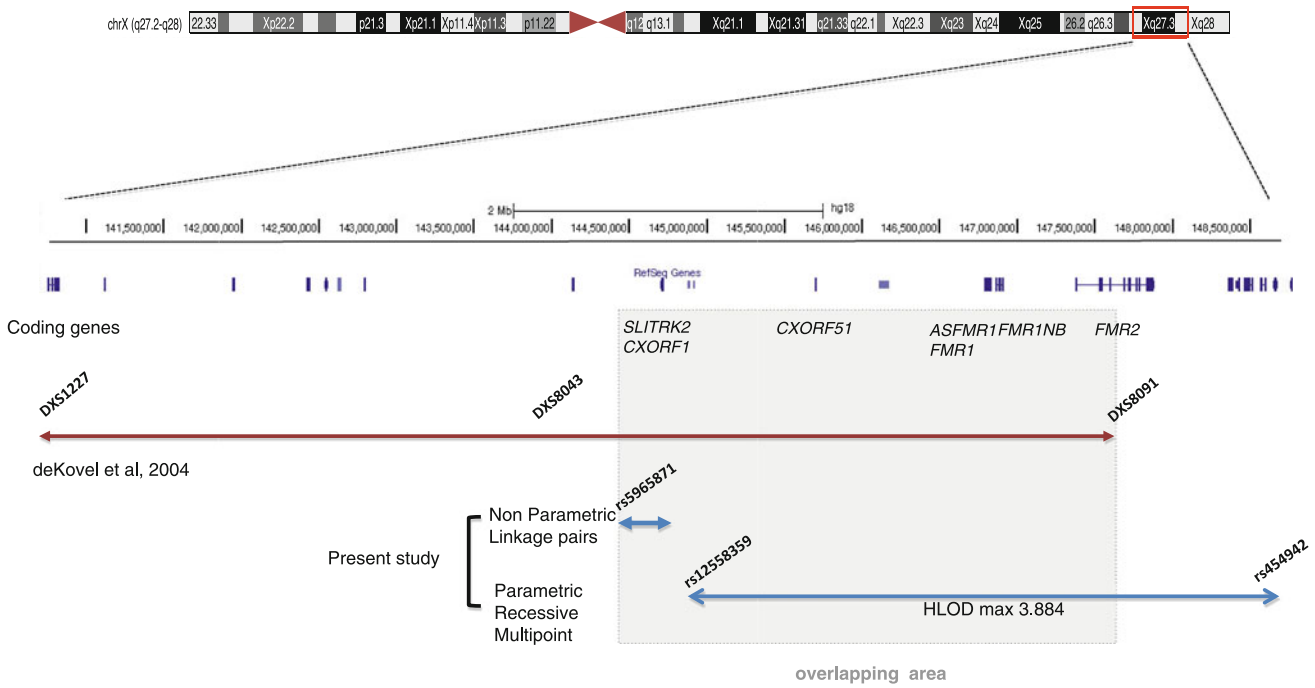


Fig. 3 Details of the Xq27 genomic region identified by linkage analysis and position of the candidate genes surrounding the *FMR1* locus

Discussion

We report here the first genome wide linkage on a French population of families with developmental dyslexia, with highest LOD scores (HLOD 3.884). Our results on several families provide consistent evidence of a genetic susceptibility locus to dyslexia in the Xq27 within DYX9 (Fig. 3).

The first suggestive linkage near this locus uncovered on Xq26 was reported by Fisher et al. (2002) with QTL analyses.Another independent microsatellite genome scan in an extended Dutch family (de Kovel et al. 2004), revealed high LOD score (multipoint LOD 3.68) at marker DXS8043 (144.028 Mb), around 700–900 kb distant from our candidate region. Their linkage extended between DXS1227 (140.802 Mb) and DXS8091 (147.603 Mb). A replication study in twins near Brisbane in Australia, based on categorical phenotype on Xq 27.3 only suggested linkage for the marker DXS9908 within the DXS1227 and DXS8091 interval (Bates et al. 2007). The difference between our parametric results and others findings suggested that the dominant effect was not prominent in our sample set on chromosome X. Only under a recessive mode of inheritance using sex specific parameter, we could demonstrate better evidence of linkage and narrowed down the region with the dense genetic map of SNPs.

The max multipoint HLOD score was outstanding 3.884 within 3.916 Mb extending from 144.969 to 148.885 Mb, and highest single point LOD reached 3.30 ($p = 0.000048$) at 146.028 Mb. To our knowledge this is the first report to withstand this level of statistical significance on linkage studies with a small number of nuclear families. The LOD max NPL pairs 2.74 in our study was marginally significative ($p = 0.0002$) whilst model free, NPL on all affected, used by de Kovel et al. (2004) resulted in LOD score 1.98 ($p = 0.0014$). Also the authors could not explain the skewed sex ratio towards male bias with their dominant model.

Previous reports described that hemizygous males were affected by the disease (Raskind 2001), which was possibly caused by an X-linked recessive allele (de Kovel et al. 2004) in agreement with our model. Fisher et al. (2002) previously studied 89 families with multiple sib-ships in the UK and 119 twin and non-twin pairs of US origin where they found a positive region on chromosome Xq26 near DXS1047.

Few studies to date have implicated the X chromosome in dyslexia, but most of the investigations were performed on sporadic cases. It is striking that multiplex family samples have been investigated in studies pointing chromosome X. For example, Fisher's study reported linkage on the X chromosome near DXS1047 only for the UK multiplex sample and not for the US twin pair sample, and in de Kovel's study, the sample was an extended family. X-linked genetic factors may thus be more involved in familial dyslexia.

In the 4 Mb region we identified seven interesting genes for mutation analysis, based on location and relevant function: *FMR1*, *ASFMR1*, *CXorf51*, *FMR1NB*, *CXorf1*,

SLITRK2 and *FMR2*. All play a role during brain development, which makes them potential candidates for developmental dyslexia (Ladd et al. 2007; Aruga and Mikoshiba 2003). In particular, *FMR1* and *FMR2* are involved in cognitive disability. Neither mutation nor polymorphisms co-segregating with dyslexia were identified in the coding sequence of these seven genes in the families. However, we can hypothesize that their implication might be mediated by a non-coding regulation sequence. Interestingly, a 46–47 CGG allele (upper limit of normal alleles, below premutation threshold) segregates in one family in two dyslexic males. Fragile X syndrome is caused by the expansion of the CGG repeat >200 in the 5′UTR, resulting in transcriptional silencing of the *FMR1* gene, whereas premutation alleles demonstrate an increase in *FMR1* mRNA level and normal or reduced amounts of FMRP. The *ASFMR1* gene, which is overlapping the *FMR1* CGG repeat region and is transcribed in the antisense orientation, is thought to contribute to the variable phenotype associated with the CGG expansion (Ladd et al. 2007). Premutation CGG repeat expansion has been proved to impair embryonic neocortical development in mice, causing migration defects in the neocortex (Cunningham et al. 2011), and premutation carriers have demonstrated impairment of verbal working memory, which is an important cognitive ability for verbal and written language (Cornish et al. 2009). The CGG repeat threshold for specific cognitive disabilities remains to be established and we can hypothesize a link between *FMR1*, *ASFMR1*, or genetic factors that modulate *FMR1* expression, and dyslexia. Lastly, a recent study proposed a model of biological pathways implicated in dyslexia where nine putative candidate genes, including *FMR1* through its interaction with RAC1 protein and *FLNA* genes, interact in neurons to perform essential functions such as neuritogenesis and neuronal migration (Poelmans et al. 2011).

The 4 SNPs rs12558359–rs12862591–rs12861185–rs905089 found to be very modestly associated did not fully explain the linkage signal peak on chromosome X. As we did not find any causal SNP, it is plausible that more than one SNP i.e., multiple genetic variants and/or with the conjunction of mutation or frameshift variant at the susceptibility locus may be responsible for the disease. We also convey that other SNP could be responsible for the disease, stating that our favorite candidate gene was not covered by any SNPs and could therefore not be directly tested. To unravel the responsible SNP it would be necessary to sequence the introns and promoters.

Also, the co-existence of at least two related genes could be necessary such as the interaction of FMR1 on chromosome X with the CYFIP2 gene located on chromosome 5, or the interaction of SLITRK2 with ROBO1 on chromosome 3 (ENCODE project consortium 2012). We therefore suggest that the *FMR1* region and its complex local regulation can be convincing candidate in familial cases of dyslexia.

Acknowledgments This work was supported by a grant of the French Ministry of Education and Research. We thank D Raine, AS Galloux, AG Piller, G Turlotte, E Schweitzer, S Heath, and C Blanc-Tailleur for their valuable contribution. We are also very grateful to the families for their participation to the study.

Conflict of interest All authors report no biomedical financial disclosures or potential conflicts of interest.

References

Abecasis GR, Wigginton JE (2005) Handling marker–marker linkage disequilibrium: pedigree analysis with clustered markers. Am J Hum Genet 77:754–767

Abecasis GR, Cherny SS, Cookson WO, Cardon LR (2002) Rapid-analysis of dense genetic maps using sparse gene flow trees. Nat Genet 30:97–101

Aruga J, Mikoshiba K (2003) Identification and characterization of Slitrk, a novel neuronal transmembrane protein family controlling neurite outgrowth. Mol Cell Neurosci 24:117–129

Bates T, Luciano M, Castles A, Coltheart M, Wright M, Martin N (2007) Replication of reported linkages for dyslexia and spelling and suggestive evidence for novel regions on chromosomes 4 and 17. Eur J Hum Genet 15:194–203

Brown WE, Eliez S, Menon V, Rumsey JM, White CD, Reiss AL (2001) Preliminary evidence of widespread morphological variations of the brain in dyslexia. Neurology 56:781–783

Cavalli-Sforza LL, King MC (1986) Detecting linkage for genetically heterogeneous diseases and detecting heterogeneity with linkage data. Am J Hum Genet 38:599–616

Cope N, Harold D, Hill G, Moskvina V, Stevenson J, Holmans P et al (2005) Strong evidence that KIAA0319 on chromosome 6p is a susceptibility gene for developmental dyslexia. Am J Hum Genet 76:59–581

Cornish KM, Kogan CS, Li L, Turk J, Jacquemont S, Hagerman RJ (2009) Lifespan changes in working memory in fragile X permutation males. Brain Cogn 69:551–558

Cottingham RW Jr., Idury RM, Schaffer AA (1993) Faster sequential genetic computations. Am J Hum Genet 53:252–263 A version of 4.1P Fastlink

Cunningham CL, Martinez Cerdeno V, Navarro E, Prakash A, Angelastro JM, Willemsen R et al (2011) Premutation CGG-repeat expansion of the Fmr1 gene impairs mouse neocortical development. Hum Mol Genet 20:64–79

Currier TA, Etchegaray MA, Haight JL, Galaburda AM, Rosen GD (2011) The effect of embryonic knockdown of the candidate dyslexia susceptibility gene DYX1C1 on the distribution of GABAergic neurons in the cerebral cortex. Neuroscience 172:535–546

Davis S, Weeks DE (1997) Comparison of nonparametric statistics for detection of linkage in nuclear families: single-marker evaluation. Am J Hum Genet 61:1431–1444

de Kovel GG, Hol FA, Heister JG, Willemen JJ, Sandkuijl LA, Franke B, Padberg GW (2004) Genomewide scan identifies susceptibility locus for dyslexia on Xq27 in an extended Dutch family. J Med Genet 41:652–657

DeFries JC, Alarcon M (1996) Genetic of specific reading disability. Ment Retard Dev Disabil Res Rev 2:39–47

Eliez S, Rumsey JM, Giedd JN, Schmitt JE, Patwardhan AJ, Reiss AL (2000) Morphological alteration of temporal lobe gray matter in dyslexia: an MRI study. J Child Psychol Psychiatry 41:637–644

ENCODE project consortium (2012) An integrated encyclopedia of DNA elements in the human genome. Nature 489:57–74

Feingold E, Song KK, Weeks DE (2000) Comparison of allele-sharing statistics for general pedigrees. Genet Epidemiol Suppl 1:92–98

Fisher SE, Francks C, Marlow AJ, MacPhie IL, Newbury DF, Cardon LR et al (2002) Independent genome-wide scans identify a chromosome 18 quantitative-trait locus influencing dyslexia. Nat Genet 30:86–91

Flannery K, Liederman J, Daly L, Schultz J (2000) Male prevalence for reading disability is found in a large sample of black and white children free for ascertainment bias. J Int Neuropsychol Soc 6(4):433–444

Francks C, Paracchini S, Smith S, Richardson A, Scerri T, Cardon L et al (2004) A 77-kilobase region of chromosome 6p22.2 is associated with dyslexia in families from the United Kingdom and from the United States. Am J Hum Genet 75:1046–1058

Galaburda AM, Sherman GF, Rosen GD, Aboitiz F, Geschwind N (1985) Four consecutive patients with cortical anomalies. Ann Neurol 18:222–233

Galaburda AM, Menard M, Rosen GD (1994) Evidence for aberrant auditory anatomy in developmental dyslexia. Proc Natl Acad Sci USA 91:80110–80113

Habib M (2000) The neurological basis of developmental dyslexia: an overview and working hypothesis. Brain 123:2373–2399

Hallgren B (1950) Specific dyslexia (congenital word-blindness); a clinical and genetic study. Acta Psychiatr Neurol Suppl 65:1–287

Hannula-Jouppi K, Kaminen-Ahola N, Taipale M, Eklund R, Nopola-Hemmo J, Kaariainen H, Kere J (2005) The axon guidance receptor gene ROBO1 is a candidate gene for developmental dyslexia. PLoS Genet 1:467–474

Klingberg T, Hedehus M, Temple E, Salz T, Gabrieli JD, Moseley ME et al (2000) Microstructure of temporo-parietal white matter as a basis for reading ability: evidence from diffusion tensor magnetic resonance imaging. Neuron 25:493–500

Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES (1996) Parametric and nonparametric linkage analysis: a unified multipoint approach. Am J Hum Genet 58:1347–1363

Ladd P, Smith L, Rabaia N, Moore J, Georges S, Hansen S et al (2007) An antisense transcript spanning the CGG repeat region of FMR1 is up regulated in permutation carriers but silenced in full mutation individuals. Hum Mol Genet 16:3174–3187

Lyon G, Shaywitt SE, Shaywitz BA (2003) A definition of dyslexia. Ann Dyslexia 53:1–14

Massinen S, Hokkanen M-E, Matsson H, Tammimies K, Tapia-Páez I et al (2011) Increased expression of the dyslexia candidate gene DCDC2 affects length and signaling of primary cilia in neurons. PLoS One 6:e20580

Meng HY, Smith SD, Hager K, Held M, Liu J, Olson RK et al (2005) DCDC2 is associated with reading disability and modulates neuronal development in the brain. Proc Natl Acad Sci USA 102:17053–17058

Muller-Myohsok B, Grimm T (1999) Linkage analysis and genetic models in dyslexia—considerations pertaining to discrete trait analysis and quantitative trait analyses. Eur Child Adolesc Psychiatry 8(Suppl 3):40–42

Pal DK, Greenberg BA (2002) Evaluating genetic heterogeneity in complex disorders. Hum Hered 53:216–226

Patterson N, Pricie AL, Reich D (2006) Population structure and gene analysis. PLoS Genet 2:2190

Pechansky VJ, Burbridge TJ, Volz AJ, Fiodella C, Wissner-Gross Z, Galaburda AM et al (2010) The effect of variation in expression of the candidate dyslexia susceptibility gene homolog Kiaa0319 on neuronal migration and dendritic morphology in the rat. Cereb Cortex 20:884–897

Poelmans G, Buitelaar J, Pauls D, Francke B (2011) A theoretical network for dyslexia: integrating available genetic findings. Mol Psychiatry 16:365–382

Price AL, Patterson NJ, Plenge RM et al (2006) Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 3:904–909

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, Sham PC (2007) PLINK: a toolset for whole-genome association and population-based linkage analysis. Am J Hum Genet 81:559–575

Raskind WH (2001) Current understanding of the genetic basis of reading and spelling disability. Learn Disabil Q 21:141–157

Rutter M, Caspi A, Fergussin D, Horwood LJ, Goodman R, Maughan B et al (2004) Sex differences in developmental reading disability: new findings from 4 epidemiological studies. JAMA 291:2007–2012

Scerri T, Schulte-Körne G (2010) Genetics of developmental dyslexia. Eur Child Adolesc Psychiatry 19:179–197

Schumacher J, Anthoni H, Dahdouh F, König IR, Hillmer AM, Kluck N et al (2006) Strong genetic evidence of DCDC2 as a susceptibility gene for dyslexia. Am J Hum Genet 78:52–62

Shaywitz SE (1998) Current concepts: dyslexia. N Engl J Med 338:307–312

Shaywitz BA, Shaywitz SE, Pugh KR, Mencl WE, Fullbright RK, Skudlarski P et al (2002) Disruption of posterior brain systems for reading in children with developmental dyslexia. Biol Psychiatry 52:101–110

Shugart YY, Chen L, Li R, Beaty T (2007) Family-based linkage disequilibrium tests using general pedigrees. Methods Mol Biol 376:141–149

Taipale M, Kaminen M, Nopola-Hemii J, Haltia T, Myllyluoma B, Lyytinen H et al (2003) A candidate gene for developmental dyslexia encodes a nuclear tetratricopeptide repeat domain protein dynamically regulated in brain. Proc Natl Acad Sci USA 100:11553–11558

R Development Core Team (2012) R: a language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria, ISBN 3-900051-07-0. http://www.R-project.org. Accessed 9 Feb 2012

Terwilliger JD, Goring HH (2000) Gene mapping in the 20th and 21st centuries: statistical methods data analysis and experimental design. Hum Biol 72(1):63–132

Vandermosten M, Boets B, Poelmans H, Sunaert S, Wouters J, Ghesquière P (2012) A tractography study in dyslexia: neuro-anatomic correlates of orthographic, phonological and speech processing. Brain 135:935–948

Wigginton JE, Abecasis GR (2005) PEDSTATS: descriptive statistics, graphics and quality assessment for gene mapping data. Bioinformatics 21:3445–3447

Ziegler A, König I, Deimel W, Plume E, Nöthen M, Propping P et al (2005) Developmental dyslexia—recurrence risk estimates from a German bi-center study using the single proband sib pair design. Hum Hered 59:136–143