Original Article

A Novel Locus and Candidate Gene for Familial Developmental Dyslexia on Chromosome 4q

Tiemo Grimm[1], Masoud Garshasbi[2], Lucia Puettmann[2], Wei Chen[2], Reinhard Ullmann[2], Bertram Müller-Myhsok[3], Eva Klopocki[1], Lina Herbst[4], Janina Haug[4], Lars R. Jensen[4], Christine Fischer[5], Markus Nöthen[6], Kerstin Ludwig[6], Andreas Warnke[7], Jürg Ott[8], Gerd Schulte-Körne[9], Hans-Hilger Ropers[2], and Andreas W. Kuss[4]

[1] Department of Human Genetics, Biozentrum, University of Würzburg, Germany
[2] Department for Human Molecular Genetics, Max Planck Institute for Molecular Genetics, Berlin, Germany
[3] Max Planck Institute of Psychiatry, Munich, Germany
[4] Interfaculty Institute for Genetics and Functional Genomics, University Medicine Greifswald, Germany
[5] Institute of Human Genetics, Heidelberg University, Germany
[6] Institute of Human Genetics, University of Bonn, Germany
[7] Department of Child and Adolescent Psychiatry and Psychotherapy, University Hospital Würzburg, Germany
[8] Laboratory of Statistical Genetics, Rockefeller University, New York, USA
[9] Department of Child and Adolescent Psychiatry and Psychotherapy, University Hospital Munich, Germany

Abstract: *Objective:* Developmental dyslexia is a highly heritable specific reading and writing disability. To identify a possible new locus and candidate gene for this disability, we investigated a four-generation pedigree where transmission of dyslexia is consistent with an autosomal dominant inheritance pattern. *Methods:* We performed genome wide array-based SNP genotyping and parametric linkage analysis and sequencing analysis of protein-coding exons, exon-intron boundaries and conserved extragenic regions within the haplotype cosegregating with dyslexia in DNA from one affected and one unaffected family member. Cosegregation was confirmed by sequencing all available family members. Additionally, we analyzed 96 dyslexic individuals who had previously shown positive LOD scores on chromosome 4q28 as well as an even larger sample (n = 2591). *Results:* We found a single prominent linkage interval on chromosome 4q, where sequence analysis revealed a nucleotide variant in the 3' UTR of brain expressed *SPRY1* in the dyslexic family member that cosegregated with dyslexia. This sequence alteration might affect the binding efficiency of the IGF2BP1 RNA-binding protein and thus influence the expression level of the *SPRY1* gene product. An analysis of 96 individuals from a cohort of dyslexic individuals revealed a second heterozygous variant in this gene, which was absent in the unaffected sister of the proband. An investigation of the region in a much larger sample further found a nominal p-value of 0.0016 for verbal short-term memory (digit span) in 2,591 individuals for a neighboring SNV. After correcting for the local number of analyzed SNVs, and after taking into account linkage disequilibrium, we found this corresponds to a p-value of 0.0678 for this phenotype. *Conclusions:* We describe a new locus for familial dyslexia and discuss the possibility that *SPRY1* might play a role in the etiology of a monogenic form of dyslexia.

Keywords: developmental dyslexia, linkage analysis, next generation sequencing, SPRY1, genetic variants

Neuer Locus und Kandidatengen für familiäre Legasthenie auf Chromosom 4q

Zusammenfassung: *Fragestellung:* Die Legasthenie ist eine spezifische Lese- und Rechtschreibschwäche mit hoher Heritabilität. Um einen neuen Genlocus mit potenziellem Kandidatengen für diese Teilleistungsstörung zu identifizieren wurde hier eine Familie mit, über vier Generationen hinweg augenscheinlich autosomal dominant vererbter, Legasthenie untersucht. *Methodik:* Basierend auf genomweiter SNP-Genotypisierung wurden parametrische Kopplungsanalysen durchgeführt. Im so identifizierten Kopplungsintervall wurden anschließend bei je einem betroffenen und nicht betroffenen Familienmitglied Sequenzuntersuchungen aller proteinkodierenden Bereiche sowie weiterer konservierter Sequenzabschnitte vorgenommen. Die Kosegregation wurde mittels Sequenzierung bei allen zugänglichen Familienmitgliedern überprüft. Darüber hinaus wurden 96 Individuen mit Legasthenie, die zuvor positive LOD Scores bezüglich Chromosom 4q28 aufgewiesen hatten, untersucht und auch mit einer noch wesentlich größeren Kohorte (n = 2591) Analysen vorgenommen. *Ergebnisse:* Es wurde ein einzelnes prominentes Kopplungsintervall auf Chromosom 4q identifiziert. Die nachfolgenden Sequenzanalysen ergaben bei den legasthenischen Probanden eine spezifische Nukleotidvariante in der 3'-UTR des im Gehirn exprimierten *SPRY1*-Gens. Diese Sequenzveränderung könnte die Bindungseffizienz des RNA-bindenden IGF2BP1 Proteins beeinträchtigen und damit Einfluss auf die Expression des *SPRY1* Genprodukts haben. Die Analyse von

96 Individuen aus einer Legasthenie-Kohorte ergab darüber hinaus eine zweite heterozygote Variante in diesem Gen, die bei der nicht betroffenen Schwester des Probanden nicht gefunden wurde. Die Untersuchung der Region in einer großen Kohorte (2591 individuen) zeigte einen nominalen *p*-Wert von 0.0016 für verbales Kurzzeitgedächtnis (Digit Span), eine benachbarte SNV. Nach Korrektur auf die Anzahl lokal untersuchter SNVs und unter Berücksichtigung des Kopplungsungleichgewichts ergab sich ein *p*-Wert von 0.0678 für diesen Phänotyp. *Schlussfolgerungen:* Diese Arbeit beschreibt einen neuen Genlocus für erbliche Legasthenie und diskutiert die Möglichkeit einer Beteiligung des *SPRY1* Gens im Rahmen der Ätiologie dieser Teilleistungsstörung.

Schlüsselwörter: Legasthenie, Kopplungsanalysen, Next Generation Sequencing, SPRY1, Genetische Varianten

Introduction

Developmental dyslexia is a specific reading and writing disability with a prevalence of around 5–10 % in school-age children (Becker et al., 2014). The highly variable phenotype is characterized not only by difficulties with phonological awareness, phonological decoding, orthographic coding, auditory short-term memory, and/or rapid naming (Peterson & Pennington, 2012), but also by sensory deficits (rapid auditory processing), attentional deficits, and perceptual deficits (the magnocellular-dorsal theory) (Stein, 2014). Developmental dyslexia has a high impact on school and psychosocial development. A compelling body of evidence suggests that developmental dyslexia has a strong genetic component, and heritability estimates range between 40 % and 80 % (Davis et al., 2001; Gayán et al., 2001; Gayán & Olson, 2003). Still, the molecular genetic basis of this disorder is largely unknown, with no gene consistently implicated across a larger number of studies (Schumacher et al., 2007). Although most dyslexia cases are probably caused by a complex interaction of several genes as well as environmental influences, the occurrence of highly loaded pedigrees with a transmission pattern consistent with autosomal dominant inheritance (such as the one we report here) suggests the possibility that monogenic subforms of dyslexia do exist. In such families, the genetic effect may be caused by variants in single genes with high penetrance. Parametric and nonparametric linkage studies have suggested several chromosomal loci, with some of them having received support from independent observations, for example, 1p36 (DYX8) (Rabin et al., 1993), 2p16 (DYX3) (Fagerheim et al., 1999), 2q22 (Raskind et al., 2005), 3q13 (DYX5) (Nopola-Hemmi et al., 2001), 4q13 (Fisher et al., 2002), 6p21 (DYX2)(Cardon et al., 1994), 7q32 (Kaminen et al., 2003), 8q12.3 (Gialluisi et al., 2019), 15q21 (DYX1) (Smith et al., 1983), 18p11 (DYX6) (Fisher et al., 2002), 18q12.2 (Gialluisi et al., 2019), 21q22 (Fisher et al., 2002), and Xq27 (DYX9) (de Kovel et al., 2004) (see also Figure 1d).

In addition, ten candidate genes (*DYX1C1* (Taipale et al., 2003), *KIAA0319* (Francks et al., 2004), *DCDC2* (Meng et al., 2005), *ROBO1* (Hannula-Jouppi et al., 2005), *KIAA0319L* (Couto et al., 2008), *CNTNAP2* (Peter et al., 2011), *MRPL19* (Scerri et al., 2012), *C2ORF3* (Scerri et al., 2012), *CEP63* (Einarsdottir et al., 2015), and *NCAN* (Einarsdottir et al., 2017) have been reported to be associated with dyslexia.

Here, we present the findings from a study in which we investigated a well-documented four-generation pedigree, where dyslexia follows a clear pattern of autosomal dominant inheritance (Grimm et al., 2006). *SPRY1*, the gene containing the putatively causative sequence alteration we identified, was subsequently screened for variants in a cohort of 96 well-characterized dyslexic individuals from a previous study (Konig et al., 2011), revealing a second, independent *SPRY1* variant possibly associated with this phenotype.

Subjects and Methods

Subjects and Sampling

All experiments were carried out following the relevant ethical guidelines and regulations. Ethics votes were granted by the Ethics Committees of the Universities of Marburg and Würzburg for the genetic studies of dyslectic individuals on which this report is based (DFG Gr506/5-1, Re471/9-1, Schu988/2-1/2-3). Assessment included testing of literacy and cognitive skills as well as an evaluation of medical records concerning the family history of reading and writing abilities. The diagnostic criteria were applied specifically correspondent to the Multiaxial Classification of Child and Adolescent Psychiatry Disorders (Remschmidt et al., 2001; WHO, 1996):

- Reading and spelling performance significantly lower than over average IQ (significance determined by the 1.5 *SD* criterion as also described by Schulte-Körne; Schulte-Korne, 2010).
- Positive family history of developmental dyslexia, delay of acquisition of language skills in preschool years.
- Scholastic difficulties did not readily remit with increased help at school and at home.
- The specific learning deficit was not explained by mental retardation or uncorrected visual, hearing, motor, or other neurological or psychological impairments.

Table 1. In silico pathogenicity prediction results

SNP	Mutation taster	FATHMM	SNPdryad	SNAP2	Mutation assessor	VEP	CADD score	REVEL score
rs150615428	disease causing	deleterious[a]	n.a.	n.a.	n.a.	n.a.	0.105	n.a.
rs141161959	disease causing	damaging[b]	deleterious	effect	low impact	tolerated	n.a.	0.108

[a] Analysis of noncoding variant (Shihab et al., 2015); [b] analysis of coding variant (Shihab et al., 2013) (unweighted); n.a.: not analyzed, because this tool does not feature the analysis of noncoding variants; Mutation Taster: http://www.mutationtaster.org/; FATHMM: http://fathmm.biocompute.org.uk/fathmmMKL.htm
SNPdryad: http://snps.ccbr.utoronto.ca:8080/SNPdryad/; SNAP2: https://rostlab.org/services/snap2web/results.php?id=05cbf194-e928-4d96-b991-d3943b552e1c; Mutation assessor: http://mutationassessor.org/r3/; VEP (Variant Effect Predictor): http://www.ensembl.org/info/docs/tools/vep/index.html; CADD (Kircher et al., 2014), REVEL (Ioannidis et al., 2016).

- The reading and spelling impairment cannot also be attributed to socioeconomic disadvantage or lack of education.

The dyslexia phenotype of all affected individuals fulfills the criteria from the ICD-10 Diagnostic Guidelines as reviewed by, among others, Thambirajah (2010). Unaffected individuals were identified by normal school performance and the absence of any conspicuous features (see also the electronic supplementary material [ESM] 1). The pedigree of the (German) family is shown in Figure 1a. Included were all individuals from whom DNA samples could be obtained. Together with the informed consent of all living individuals, we obtained blood samples to isolate the DNA from all family members in the pedigree. The Epstein Barr virus-transformed lymphoblastoid cell lines were generated from one patient and his unaffected daughter.

Linkage Analysis

We genotyped six affected and seven healthy individuals from three generations of the family using the Illumina Linkage IV panel (Illumina, 9885 Towne Center Drive, San Diego, CA, 92121–1975, USA) and performed genetic linkage analysis on the resulting 5,695 marker variants distributed over all chromosomes. The genotyping data were converted into appropriate formats, using the ALOHOMORA software (Ruschendorf & Nurnberg 2005). To verify the interindividual relationships, we used standard quality-control routines, including the graphical representation of relationship errors (GRR) (Abecasis et al., 2002) and PedCheck (O'Connell & Weeks, 1998). Merlin (Abecasis et al., 2002) and GeneHunter (Kruglyak et al., 1996) tools were applied to perform genome-wide multipoint linkage analysis, assuming an autosomal dominant mode of inheritance and complete penetrance as also previously described (Garshasbi et al., 2006). Subsequently, we chose the pseudomarker program (Goring & Terwilliger, 2000;

Hiekkalinna et al., 2011) (updated implementation at http://www.jurgott.org/linkage/runPM.html) as it has two advantages over other approaches: (1) It treats marker allele frequencies as nuisance parameters, that is, it estimates allele frequencies separately in numerator and denominator of the likelihood ratio for the LOD score, so that results become virtually independent of marker allele frequencies. (2) It allows for the combined estimation of linkage and linkage disequilibrium between marker variant and putative disease variant, although this second advantage is not relevant in a small pedigree as analyzed here.

Array-CGH Analysis

In order to exclude insertions and deletions, we performed array-CGH using a high-resolution 1M oligonucleotide array (SurePrint G3 Human CGH Microarray Kit, 1x1M, Agilent Technologies, Santa Clara, CA, UDA) as described previously (Lohan et al., 2014), which showed no abnormalities in the tested individuals.

Variant Screening and in Silico Analysis

A total of 16,092 regions (Online Resource for coordinates [Agilent Technologies Deutschland GmbH, Herrenberger Str. 130, D-71034 Böblingen, Germany]) comprising 474,282 bp including all exons and exon-intron boundaries as well as among vertebrates conserved regions from within the interval, flanked by rs1880863 and rs906372, were enriched with custom-made Agilent SureSelect DNA capture arrays (Agilent Technologies Deutschland GmbH, Herrenberger Str. 130, D-71034 Böblingen, Germany) as previously described (Najmabadi et al., 2011) and deep sequenced by generating 50bp reads covering the enriched sequences at an average depth of 953 for sample 1890 and 2010 for sample 1893 (Online Resource Table 1 for further details) with a SOLiD 5500XL sequencer (Life Technolo-

gies GmbH, Frankfurter Straße 129B, D-64293 Darmstadt, Germany) according to the manufacturer's instructions. The total length of all enriched regions is 516,851 bp. We then aligned the reads obtained from the sequencing runs with the human reference genome (hg19) using bfast-0.6.5 (Homer et al., 2009). Variant screening was performed with samtools (H. Li et al., 2009) using reads that mapped unambiguously to the target region. Variant calls had to be supported by at least 10 nonidentical reads, a Phred-like quality score ≥ 1 and an allelic percentage of 20–80 %. We retained variants with a minor allele frequency (MAF) below 0.01 and also filtered using sample 1893 (nondyslexic family member), dbSNP132, refseq exons, silent variants, and 6503 exomes available through the Exome Variant Server (NHLBI, 2013). In addition, we performed Sanger sequencing in order to safeguard against false-positive results. We screened SPRY variants in the dyslexia cohort (Konig et al., 2011) by Sanger Sequencing. Furthermore, we used a collection of 100 samples from unrelated individuals (dyslexia status not having been previously evaluated) that matched the intellectual status of the individuals of the family investigated here (higher education entrance qualification and academic studies) as controls for identified variants.

For the in silico prediction of functional effects of the identified variants, we used a variety of prediction tools: MutationTaster2 (Schwarz et al., 2014) (http://www.mutationtaster.org), FATHMM (http://fathmm.biocompute.org.uk/fathmmMKL.htm), SNPdryad (http://snps.ccbr.utoronto.ca:8080/SNPdryad/), SNAP2 (https://rostlab.org/services/snap/). (http://mutationassessor.org/r3/), Variant Effect Predictor (http://www.ensembl.org/info/docs/tools/vep/index.html). In addition, we used the CADD (Kircher et al., 2014) and REVEL tools (Ioannidis et al., 2016).

Dyslexia Cohort

For additional analysis, we employed 96 dyslexic individuals, phenotyped according to the same criteria applied for the pedigree described, who had previously shown positive LOD scores (LOD > 0) for chromosome 4q28, from a German cohort of 248 dyslexia families described by Konig et al. (2011). These had been recruited through a single-proband sib-pair study design with a detailed assessment of dyslexia and related cognitive traits (Konig et al., 2011). The power to detect a locus explaining 1% of trait variance at this locus employing a p-value criterion of .05 is 10 %, reflecting the comparatively low power of the sibship design in these settings. Furthermore, we performed analyses in a previously described multinational dataset (Gialluisi et al., 2019). This dataset consists of individuals from nine cohorts of reading-impaired and typically developing children of European ancestry with a total sample size of between 2,562 and 3,468, depending on the phenotype considered. The power to detect a genome-wide signal (p < 5e-08), explaining 1.5 % of the phenotypic variance for a specific trait, in this dataset is beyond 75 %. In the regional replication scenario employed here, the power for explaining 1% of the variance is beyond 99 % (setting a p-value threshold of .01). For a more detailed discussion of power in this cohort, see Gialluisi et al. (2019). All power calculations were done using the Genetic Power Calculator website (Purcell et al., 2003).

Gene Expression Analysis

Commercially available RNA from various adult brain tissues (cerebellum, cerebral cortex, corpus callosum, diencephalon, frontal lobe, hippocampus, medulla oblongata, occipital lobe, parietal lobe, temporal lobe, thalamus) was analyzed by RNA-Sequencing on a 5500xl SOLiD System (Life Technologies, Carlsbad, CA, USA) as previously described (Esmaeeli-Nieh et al., 2016). Raw sequencing data were submitted to read mapping, using the Lifescope software analysis tool (Life Technologies, Carlsbad, CA, USA) and reference genome GRCh37/hg19. Statistical analysis was carried out using R (R Core Team 2020). Normalization and expression analysis was performed using the R packages EDASeq (Risso et al., 2011) and DESeq (Anders & Huber 2010), respectively.

Results

In the pedigree investigated in our study, 8 out of a total of 15 offspring of affected parents were affected with dyslexia. Assuming simple dominant inheritance, a theoretical number of 7.5 would have been expected. Comparing these numbers using a chi^2 test as suggested by Emery (Emery, 1976), we found no significant difference between the expected and the observed values. Still, linkage analysis under a fully penetrant dominant, affected-only analysis showed only minimal lod scores, so we allowed for incomplete penetrance of 0.96 for susceptibility genotypes, a penetrance for phenocopies of 0.01, and specified a disease allele frequency of 0.005 (the exact parameter values are not too relevant as long as incomplete penetrance is specified). A maximum LOD score of 2.61 occurred at basepair position 130,292,021 (rs318539) on chromosome 4q27-28.3 (Figure 1). If linkage analysis is carried out allowing for linkage disequilibrium (LD) between the marker and the putative disease locus, the result is somewhat stronger, and the p-value of .000220 translates into a LOD score of
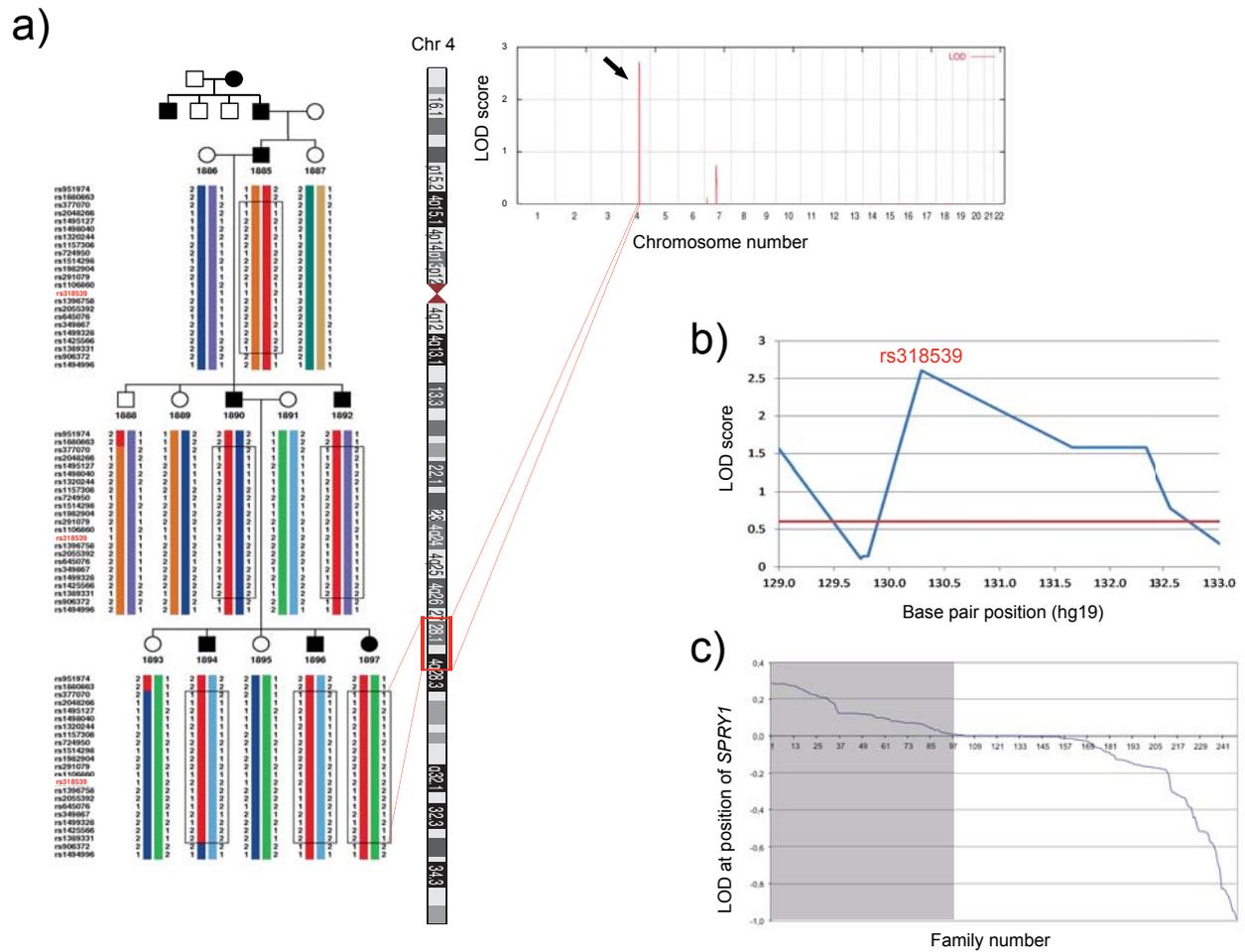
Figure 1. (a) Family pedigree and haplotypes (box) within linkage interval (SNP IDs indicated on the left) and genome-wide LOD score plot showing a single peak (arrow) on chromosome 4 exceeding any other peak by more than 1 unit. The corresponding linkage interval between rs1880863 and rs906372 is indicated (red box) on the ideogram of chromosome 4. **(b)** Curve of LOD scores (*y*-axis, red line) at variants surrounding the peak LOD score, graphed against base pair positions (*x*-axis, in units of MB) on chromosome 4. The red line reflects a LOD score drawn 2 units below the maximum LOD score, whose intersections with the LOD score curve provide the endpoints of a 95 % confidence interval (obtained as previously recommended by Conneally et al., 1985). **(c)** LOD scores at the position of *SPRY1* in dyslexia cohort (Konig et al., 2011). 97 families had positive LOD scores (indicated by grey shading). **(d)** Known dyslexia loci. The newly identified locus on chromosome 4 is highlighted (black circle).

approximately 2.68. While this result falls short of the traditional threshold of a LOD score of 3 for genome-wide significance, it is nonetheless suggestive (Lander & Kruglyak, 1995), also because it belongs to a single peak exceeding any other peak by more than 1 LOD unit (Figure 1a). Finally, it should also be noted that the maximum LOD scores of 2.61 and 2.68 are very close to the theoretical maximum LOD score obtainable in this family, which is 2.71.

The marker rs318539 is part of a haplotype between the markers rs1880863 and rs906372 which is shared among all affected members (see also supplemental Table 1 for phenotype information) of the three genotyped genera-

tions (Figure 1a), spans 11,653,660 bp on the long arm of chromosome 4, and contains 31 REFSEQ genes. Array CGH analysis did not show any copy number variants in a DNA sample from the index patient at this locus. The exonic regions of the genes as well as sequences conserved among vertebrates within this region were screened for variants using targeted enrichment and next generation sequencing. After we filtered out known frequent variants, a heterozygous variant (hg19 chr4:g.124324233C>T) in the 3' UTR of *SPRY1*, the human ortholog of the Drosophila sprouty gene 1, remained (Figure 2), which cosegregated with dyslexia in the investigated family, as verified by se-
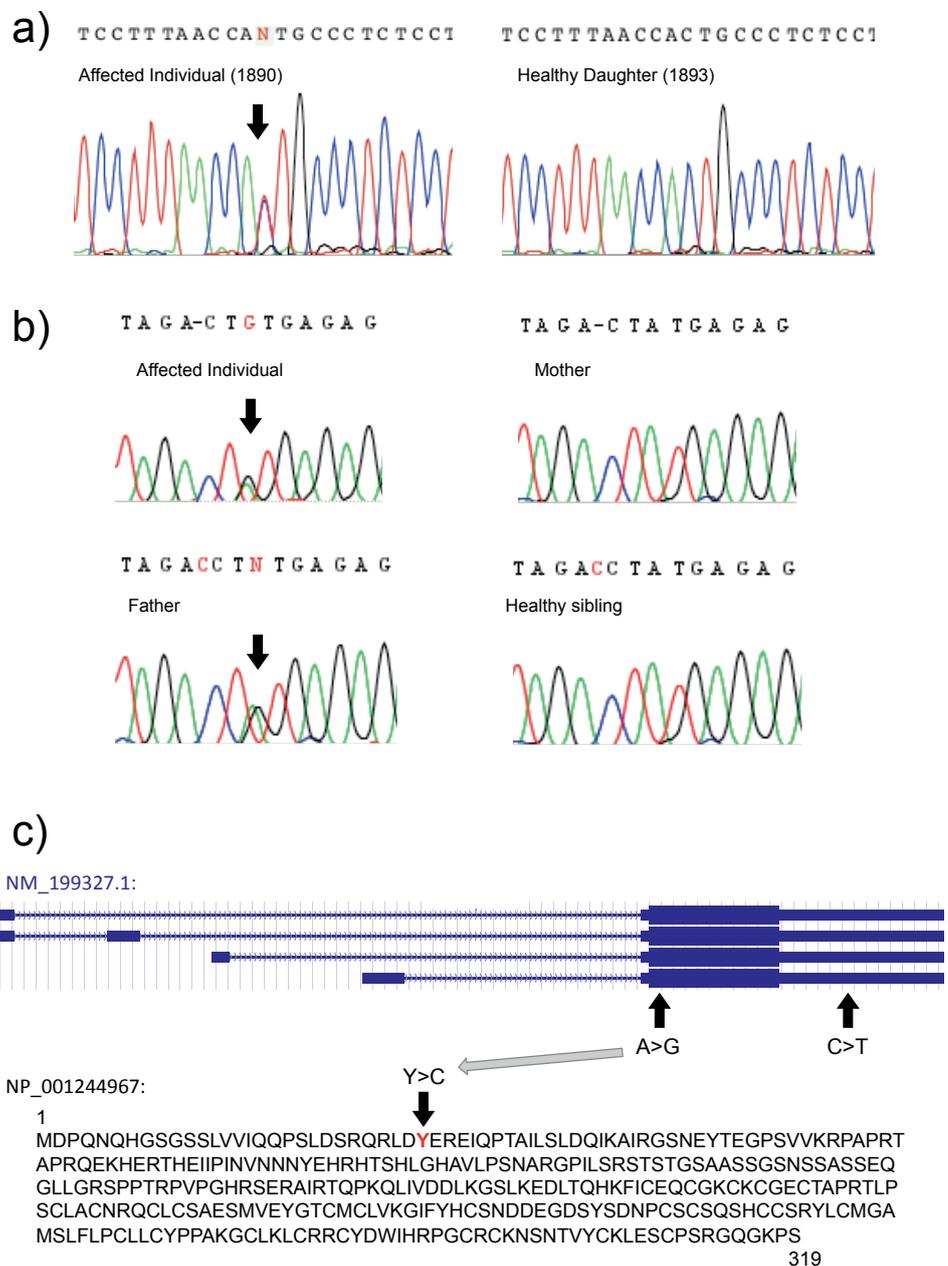

Figure 2. Sequence alterations affecting *SPRY1* **(a)** in the 3'UTR or **(b)** the terminal exon. **(c)** Schematic representation of *SPRY1* (above) and protein sequence of the gene product (below). Arrows indicate the positions of the respective sequence alterations.

quencing in all available family members. This variant can be found in dbSNP (Sherry et al., 2001) build 134 and above as rs150615428. The MAF (T-allele) based on 2,504 individuals from the 1,000 genomes project (Genomes Project et al., 2015) is 0.005, and in all carriers, it was found in a heterozygous state. Data from the available exome studies or repositories (200 Danish exomes; Y. Li et al., 2009), Exome Variant Server (NHLBI, 2013), and ExAC data (Consortium, 2015) did not contain this alteration, though the investigated genomic regions did not include the position of rs150615428, so that these data sets could not be used for determining an allele frequency.

Interestingly, when screening *SPRY1* in a collection of 100 samples from unrelated individuals (dyslexia status was not previously evaluated) who match the intellectual status of the individuals of the family investigated here (higher education entrance qualification and academic studies) as a more matching set of controls, we also detected one case of heterozygous for rs150615428. Because the samples were anonymized, the phenotype of the carrier of this variant could unfortunately not be determined.

Thus, taking into account the reported 24 cases from the 1000 Genome Project and the one case from our control cohort, we found the allele carrying the DNA alteration to be present in 25 out of at least 2,604 individuals (i.e. 5,208 chromosomes) with a resulting theoretical MAF of 0.0048, which is consistent with the MAF reported in dbSNP (http://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?rs=rs 150615428, accessed 08/2016).

The rs150615428 variant is predicted by MutationTaster2 (Schwarz et al., 2014) to be disease-causing (http://www.mutationtaster.org). This was supported by the FATHMM tool (http://fathmm.biocompute.org.uk/fathmmMKL.htm), which predicted rs150615428 to be deleterious (see Table 1 for pathogenicity prediction results from various software tools). Moreover, according to the doRiNA database (Anders et al., 2012), g.124324233C>T is situated in the center of an IGF2BP1 binding site and therefore most likely affects the interaction of IGF2BP1, an RNA-binding protein that controls the cytoplasmic fate of its target mRNAs (Stohr & Huttelmaier, 2012), with *SPRY1* mRNA. Thus, it could affect the expression level of *SPRY1*. In order to ascertain whether *SPRY1* is expressed in relevant parts of the human brain, we carried out RNA-Seq experiments with RNA obtained from different parts of the human brain. This analysis showed that *SPRY1* is expressed in all investigated brain regions (cerebellum, cerebral cortex, corpus callosum, diencephalon, frontal lobe, hippocampus, medulla oblongata, occipital lobe, parietal lobe, temporal lobe, thalamus; see also Table 2).

In order to find additional evidence for the relevance of alterations that affect the expression, cellular localization,

Table 2. *SPRY1* expression in the human brain

Tissue	Gene expression [RPKM][a]
Cerebellum	4.4
Cerebral cortex	2.6
Corpus callosum	1.6
Diencephalon	2.9
Frontal lobe	4.4
Hippocampus	2.8
Medulla oblongata	5.9
Occipital lobe	3.3
Parietal lobe	2.6
Temporal lobe	2.8
Thalamus	3.7

[a]Reads per kilobase per million reads.

and/or function of this putative candidate gene for dyslexia, we analyzed *SPRY1* in 96 dyslexic individuals from a German cohort of 248 dyslexia families (Konig et al., 2011). These families had been recruited through a single-proband sib-pair study design with detailed assessment of dyslexia and related cognitive traits. All 96 investigated families had previously provided positive LOD scores for chromosome 4q28 (Konig et al., 2011). The sequence analysis of *SPRY1* in these probands, however, revealed only one sequence variant in this gene in a single individual (LMR-2022-1; LOD score 0.015): A missense variant Y29C (rs141161959; hg19:g.124322832A>G, see also Figure 2b and 2c) with reported frequencies of 0.001-0.002 for the G allele. This variant was also determined to be disease-causing by a variety of prediction tools including MutationTaster2 (Schwarz et al., 2014), FATHMM, and SNP-dryad. Also, the mutation assessor tool predicted a possible impact and only the Variant Effect Predictor software predicted this variant to be tolerated (see also Table 1). In addition, we checked the variants using the CADD (Kircher et al., 2014) and REVEL tools (Ioannidis et al., 2016), both of which yielded scores that do not exclude pathogenicity for the described variants (Table 1).

No individuals homozygous for the G allele were described (dbSNP build 138). Subsequent sequence analysis of two additional family members showed that this sequence variant was inherited from the father, whose dyslexia phenotype is unknown. Still, both the unaffected sibling and the proband's mother are homozygous for the unaltered sequence.

Both variants described here were submitted to the public Leiden Open Variation Database 3.0 (http://www.lovd.nl/) and have the following identifiers: #00100429 (g.124324233C>T) and 00100431 (g.124322832A>G).

By investigating the region in a much larger sample, we further found that, for the verbal short-term memory phenotype digit span in n = 2,591 individuals, a nominal p-value of 0.0016 was observed for the SNV rs159864 (Gialluisi et al., 2019) neighboring the *SPRY1* gene. After correcting for the investigated local number of SNVs and taking linkage disequilibrium into account, this corresponds to a p-value of 0.0678 for digit span. Correcting for the possible tests of the other phenotypes would add another five tests as noted in the originating manuscript (Gialluisi et al., 2019). In the original genome-wide testing of linkage in the extended family, a correction for the family is implicit in the LOD score threshold of 3 for significance. This threshold would correspond to a much stricter p-value than 0.05 required to declare genome-wide significance at 0.05. For the lookup in the family-based sample, no correction for genome-wide analysis needs to be done as the lookup is strictly only on that region and, because of the high general serial autocorrelation of LOD score curves, amounts to a single test. Finally, for the lookup in the data from Gilaluisi et al. we took into account the number of SNVs, linkage disequilibrium, and the correlation between phenotypes, thus implementing a correction of the family-wise error rate by taking these factors into account.

Discussion

We report here on a family with 8 out of 15 members affected by developmental dyslexia. In line with previous studies reporting autosomal dominant inheritance in dyslexia and related phenotypes (Chapman et al., 2003; Gilger et al., 1994; Wijsman et al., 2000), we assumed a comparable mode of inheritance and thus performed genetic linkage analysis leading to the identification of a previously undescribed locus on chromosome 4q28, which is supported by reaching the near maximum LOD score obtainable in this family. This does not overlap with published loci and contains a putative new candidate gene. The only limitation in this context is the fact that we could not obtain DNA samples from the first generation of the investigated pedigree, as they were not available for genotyping and linkage analysis. The inclusion of more generations would have made it possible to reach even more significant LOD scores. Our experimental strategy is in keeping with previous studies, where the authors used a very similar combination of linkage analysis and next generation sequencing for the successful identification of pathogenic sequence alterations with an autosomal dominant mode of inheritance in linkage regions with comparable LOD scores (Bowen et al., 2011; Sobreira et al., 2010).

Interestingly, several other loci on chromosome 4q (e.g., 4q13.1, 4q22.1, and 4q32.3) have been associated with ADHD. This is of note, because developmental dyslexia and ADHD were observed to occur in both clinical and community samples side by side in individuals more often than would be expected by chance. However, these loci do not overlap with the linkage interval we investigated here, and, more importantly, ADHD was not observed in any of the affected individuals in the pedigree we studied.

A variant screen in all known exonic regions as well as conserved extragenic sequences within the cosegregating haplotype yielded a DNA variant (rs150615428) that affects the *SPRY1* gene. Interestingly, a second *SPRY1* variant (rs141161959) was found in one out of 96 families who had previously shown positive LOD scores for chromosome 4q28. Both variants are predicted to be disease-causing by most prediction tools. In the family we report on here, the dyslexia phenotype appears to follow an autosomal dominant mode of inheritance, with a penetrance that appears to be 100 %; that could, however, result from a genetic background favoring the development of dyslexia, while in other carriers of this variant, a different genetic background could prevent the occurrence of the phenotype. A similar situation has been observed concerning variants in *BRCA1* and *BRCA2*, for which penetrance estimates vary substantially between different studies, and for which it has been shown that these differences are caused by genetic and environmental modifiers (Milne & Antoniou, 2016).

SPRY1 belongs to a family of four genes that respond to FGF signaling and whose gene products seem to function as negative feedback regulators (Faedo et al., 2010; Hanafusa et al., 2002; Klein et al., 2006; Mahoney Rogers et al., 2011; Minowada et al., 1999). In mice, Spry proteins are required for the development of several tissues, including the hind- and midbrain (Basson et al., 2005; Klein et al., 2006; Mahoney Rogers et al., 2011; Shim et al., 2005). Interestingly, the repression of Spry1 can lead to an increase in the activation of the Erk (Ras-extracellular signal-regulated kinase) signaling pathway (Thum et al., 2008). This highly conserved pathway is also essential for different aspects of neuronal signaling and is thus involved in biochemical events connected with learning and the formation of memory (Adams & Sweatt, 2002; Mazzucchelli & Brambilla, 2000; Orban et al., 1999).

Among the 60,706 exomes investigated by the Exome Aggregation Consortium (ExAC) (Consortium, 2015), only six loss-of-function variants are reported in this gene, none of which was observed in a homozygous state. Also,

the vast majority of single nucleotide alterations were found only in a heterozygous state. This observation could tempt one to speculate that a complete loss of functional SPRY1 might not be compatible with life, and that changes in the dosage of the gene product can already have an effect. Such dosage effects would most likely become apparent in evolutionarily young features with a therefore comparatively low degree of mechanistic redundancy. This could be assumed for the processes involved in the development and maintenance of human cognitive features, let alone for the complex integration of several such features, as is required for reading and writing abilities.

On the molecular level, SPRY proteins can inhibit the Ras-MAPK pathway (Hanafusa et al., 2002; Kim & Bar-Sagi, 2004; Mason et al., 2006). Variants in members of this signaling pathway can result in a variety of congenital disorders, the so-called RASopathies (for a review, see, e.g., Rauen, 2013). Among the hallmarks shared by these disorders is the impairment of cognitive functions. One could therefore speculate that a mild disturbance in the regulation of this pathway might well interfere with the acquirement of scholastic skills. Such a presumably mild effect on RAS-MAPK signaling would be in line with the predicted interference of the 3'UTR variant we report here with an IGF2BP1 binding site. IGF2BP1 is an RNA-binding protein, which plays a role in nuclear export, subcellular localization, stability, and translation of its target mRNAs. It could be speculated that impaired IGF2BP1 binding to SPRY1 transcripts affects the amounts of available SPRY1 gene products in the cell, which in turn might lead to the downstream disturbances that are eventually responsible for the dyslexia phenotype observed in the carriers of this variant.

Still, bearing in mind that a causative variant could also be located elsewhere in this locus and would cosegregate equally well, we also screened among vertebrates conserved extragenic regions in the linkage interval, which, however, failed to lead to the discovery of any further interpretable potentially causative variants.

Taken together, our results suggest a new locus for developmental dyslexia and provide tentative evidence that could be seen to point to the possible involvement of SPRY1 in the etiology of this condition.

Electronic supplementary material

The electronic supplementary material (ESM) is available with the online version of the article at https://doi.org/10.1024/1422-4917/a000758.
ESM 1. Phenotypes and diagnostics for the individuals from the pedigree (Table).

References

Abecasis, G.R., Cherny, S.S., Cookson, W.O., & Cardon, L.R. (2002). Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. *Nature genetics, 30*(1), 97–101. https://doi.org/10.1038/ng786

Adams, J.P., & Sweatt, J.D. (2002). Molecular psychology: Roles for the ERK MAP kinase cascade in memory. *Annual Review of Pharmacology and Toxicology, 42*, 135–163. https://doi.org/10.1146/annurev.pharmtox.42.082701.145401

Anders, G., Mackowiak, S.D., Jens, M., Maaskola, J., Kuntzagk, A., Rajewsky, N., … Dieterich, C. (2012). doRiNA: A database of RNA interactions in post-transcriptional regulation. *Nucleic Acids Research, 40*(Database issue), D180–D186. https://doi.org/10.1093/nar/gkr1007

Anders, S., Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biology 11*, R106. doi: 10.1186/gb-2010-11-10-r106, http://genomebiology.com/2010/11/10/R106/.

Basson, M.A., Akbulut, S., Watson-Johnson, J., Simon, R., Carroll, T.J., Shakya, R., … Licht, J.D. (2005). Sprouty1 is a critical regulator of GDNF/RET-mediated kidney induction. *Developmental Cell, 8*(2), 229–239. https://doi.org/10.1016/j.devcel.2004.12.004

Becker, J., Czamara, D., Scerri, T.S., Ramus, F., Csepe, V., Talcott, J.B., … Schumacher, J. (2014). Genetic analysis of dyslexia candidate genes in the European cross-linguistic NeuroDys cohort. *European Journal of Human Genetics, 22*(5), 675–680. https://doi.org/10.1038/ejhg.2013.199

Bowen, M.E., Boyden, E.D., Holm, I.A., Campos-Xavier, B., Bonafe, L., Superti-Furga, A., … Warman, M.L. (2011). Loss-of-function mutations in PTPN11 cause metachondromatosis, but not Ollier disease or Maffucci syndrome. *PLoS Genetics, 7*(4), e1002050. https://doi.org/10.1371/journal.pgen.1002050

Cardon, L.R., Smith, S.D., Fulker, D.W., Kimberling, W.J., Pennington, B.F., & DeFries, J.C. (1994). Quantitative trait locus for reading disability on chromosome 6. *Science, 266*(5183), 276–279.

Chapman, N.H., Raskind, W.H., Thomson, J.B., Berninger, V.W., & Wijsman, E.M. (2003). Segregation analysis of phenotypic components of learning disabilities. II. Phonological decoding. *American Journal of Medical Genetics, Part B: Neuropsychiatric Genetics, 121B*(1), 60–70. https://doi.org/10.1002/ajmg.b.20068

Conneally, P.M., Edwards, J.H., Kidd, K.K., Lalouel, J.M., Morton, N.E., Ott, J., & White, R. (1985). Report of the Committee on Methods of Linkage Analysis and Reporting. *Cytogenetics and Cell Genetics, 40*(1–4), 356–359.

Consortium, E.A. (2015). Analysis of protein-coding genetic variation in 60,706 humans. *bioRxiv.* https://doi.org/http://dx.doi.org/10.1101/030338

Couto, J.M., Gomez, L., Wigg, K., Cate-Carter, T., Archibald, J., Anderson, B., … Barr, C.L. (2008). The KIAA0319-like (KIAA0319L) gene on chromosome 1p34 as a candidate for reading disabilities. *Journal of Neurogenetics, 22*(4), 295–313. https://doi.org/10.1080/01677060802354328

Davis, C.J., Gayan, J., Knopik, V.S., Smith, S.D., Cardon, L.R., Pennington, B.F., Olson, R.K., DeFries, J.C. (2001) Etiology of reading difficulties and rapid naming: the Colorado Twin Study of Reading Disability. *Behav Genet 31*(6), 625–635.

de Kovel, C.G., Hol, F.A., Heister, J.G., Willemen, J.J., Sandkuijl, L.A., Franke, B., & Padberg, G.W. (2004). Genomewide scan identifies susceptibility locus for dyslexia on Xq27 in an extended Dutch family. *Journal of Medical Genetics, 41*(9), 652–657. https://doi.org/10.1136/jmg.2003.012294

Einarsdottir, E., Peyrard-Janvid, M., Darki, F., Tuulari, J.J., Merisaari, H., Karlsson, L., … Kere, J. (2017). Identification of NCAN as a candidate gene for developmental dyslexia. *Scientific Reports, 7*(1), 9294. https://doi.org/10.1038/s41598-017-10175-7

Einarsdottir, E., Svensson, I., Darki, F., Peyrard-Janvid, M., Lindvall, J.M., Ameur, A., … Matsson, H. (2015). Mutation in CEP63 cosegregating with developmental dyslexia in a Swedish family. *Human Genetics, 134*(11–12), 1239–1248. https://doi.org/10.1007/s00439-015-1602-1

Emery, A.E.H. (1976). *Methodology in medical genetics.* Churchill Livingstone.

Esmaeeli-Nieh, S., Fenckova, M., Porter, I.M., Motazacker, M.M., Nijhof, B., Castells-Nobau, A., … Kuss, A.W. (2016). BOD1 is required for cognitive function in humans and Drosophila. *PLoS Genetics, 12*(5), e1006022. https://doi.org/10.1371/journal.pgen.1006022

Faedo, A., Borello, U., & Rubenstein, J.L. (2010). Repression of Fgf signaling by sprouty1-2 regulates cortical patterning in two distinct regions and times. *The Journal of Neuroscience, 30*(11), 4015–4023. https://doi.org/10.1523/JNEUROSCI.0307-10.2010

Fagerheim, T., Raeymaekers, P., Tonnessen, F.E., Pedersen, M., Tranebjaerg, L., & Lubs, H.A. (1999). A new gene (DYX3) for dyslexia is located on chromosome 2. *Journal of Medical Genetics, 36*(9), 664–669.

Fisher, S.E., Francks, C., Marlow, A.J., MacPhie, I.L., Newbury, D.F., Cardon, L.R., … Monaco, A.P. (2002). Independent genome-wide scans identify a chromosome 18 quantitative-trait locus influencing dyslexia. *Nature Genetics, 30*(1), 86–91. https://doi.org/10.1038/ng792

Francks, C., Paracchini, S., Smith, S.D., Richardson, A.J., Scerri, T.S., Cardon, L.R., … Monaco, A.P. (2004). A 77-kilobase region of chromosome 6p22.2 is associated with dyslexia in families from the United Kingdom and from the United States. *American Journal of Human Genetics, 75*(6), 1046–1058. https://doi.org/10.1086/426404

Garshasbi, M., Motazacker, M.M., Kahrizi, K., Behjati, F., Abedini, S.S., Nieh, S.E., … Najmabadi, H. (2006). SNP array-based homozygosity mapping reveals MCPH1 deletion in family with autosomal recessive mental retardation and mild microcephaly. *Human Genetics, 118*(6), 708–715. https://doi.org/10.1007/s00439-005-0104-y

Gayan, J., Olson, R.K. (2001) Genetic and environmental influences on orthographic and phonological skills in children with reading disabilities. *Dev Neuropsychol 20*(2), 483–507.

Gayán, J., Olson, R.K. (2003) Genetic and environmental influences on individual differences in printed word recognition. *J Exp Child Psychol 84*(2), 97–123. doi:10.1016/s0022-0965(02)00181-9

Genomes Project, C., Auton, A., Brooks, L.D., Durbin, R.M., Garrison, E.P., Kang, H.M., … Abecasis, G.R. (2015). A global reference for human genetic variation. *Nature, 526*(7571), 68–74. https://doi.org/10.1038/nature15393

Gialluisi, A., Andlauer, T.F.M., Mirza-Schreiber, N., Moll, K., Becker, J., Hoffmann, P., … Schulte-Korne, G. (2019). Genome-wide association scan identifies new variants associated with a cognitive predictor of dyslexia. *Translational Psychiatry, 9*(1), 77–91. https://doi.org/10.1038/s41398-019-0402-0

Gilger, J.W., Borecki, I.B., DeFries, J.C., & Pennington, B.F. (1994). Commingling and segregation analysis of reading performance in families of normal reading probands. *Behavior Genetics, 24*(4), 345–355.

Goring, H.H., & Terwilliger, J.D. (2000). Linkage analysis in the presence of errors IV: Joint pseudomarker analysis of linkage and/or linkage disequilibrium on a mixture of pedigrees and singletons when the mode of inheritance cannot be accurately specified. *American Journal of Human Genetics, 66*(4), 1310–1327. https://doi.org/10.1086/302845

Grimm, T. (2006). Schicksal Legasthenie. *Medizinische Genetik 18*, 145–150.

Hanafusa, H., Torii, S., Yasunaga, T., & Nishida, E. (2002). Sprouty1 and Sprouty2 provide a control mechanism for the Ras/MAPK signalling pathway. *Nature Cell Biology, 4*(11), 850–858. https://doi.org/10.1038/ncb867

Hannula-Jouppi, K., Kaminen-Ahola, N., Taipale, M., Eklund, R., Nopola-Hemmi, J., Kaariainen, H., & Kere, J. (2005). The axon guidance receptor gene ROBO1 is a candidate gene for developmental dyslexia. *PLoS Genetics, 1*(4), e50. https://doi.org/10.1371/journal.pgen.0010050

Hiekkalinna, T., Schaffer, A.A., Lambert, B., Norrgrann, P., Goring, H.H., & Terwilliger, J.D. (2011). PSEUDOMARKER: A powerful program for joint linkage and/or linkage disequilibrium analysis on mixtures of singletons and related individuals. *Human Heredity, 71*(4), 256–266. https://doi.org/10.1159/000329467

Homer, N., Merriman, B., & Nelson, S.F. (2009). BFAST: An alignment tool for large scale genome resequencing. *PloS One, 4*(11), e7767. https://doi.org/10.1371/journal.pone.0007767

Ioannidis, N.M., Rothstein, J.H., Pejaver, V., Middha, S., McDonnell, S.K., Baheti, S., … Sieh, W. (2016). REVEL: An ensemble method for predicting the pathogenicity of rare missense variants. *American Journal of Human Genetics, 99*(4), 877–885. https://doi.org/10.1016/j.ajhg.2016.08.016

Kaminen, N., Hannula-Jouppi, K., Kestila, M., Lahermo, P., Muller, K., Kaaranen, M., Myllyluoma, B., Voutilainen, A., Lyytinen, H., Nopola-Hemmi, J., Kere, J. (2003) A genome scan for developmental dyslexia confirms linkage to chromosome 2p11 and suggests a new locus on 7q32. *J Med Genet, 40*(5), 340–345.

Kim, H.J., & Bar-Sagi, D. (2004). Modulation of signalling by Sprouty: A developing story. Nature Reviews. *Molecular Cell Biology, 5*(6), 441–450. https://doi.org/10.1038/nrm1400

Kircher, M., Witten, D.M., Jain, P., O'Roak, B.J., Cooper, G.M., & Shendure, J. (2014). A general framework for estimating the relative pathogenicity of human genetic variants. *Nature Genetics, 46*(3), 310–315. https://doi.org/10.1038/ng.2892

Klein, O.D., Minowada, G., Peterkova, R., Kangas, A., Yu, B.D., Lesot, H., … Martin, G.R. (2006). Sprouty genes control diastema tooth development via bidirectional antagonism of epithelial-mesenchymal FGF signaling. *Developmental Cell, 11*(2), 181–190. https://doi.org/10.1016/j.devcel.2006.05.014

Konig, I.R., Schumacher, J., Hoffmann, P., Kleensang, A., Ludwig, K.U., Grimm, T., … Schulte-Korne, G. (2011). Mapping for dyslexia and related cognitive trait loci provides strong evidence for further risk genes on chromosome 6p21. *American Journal of Medical Genetics, Part B: Neuropsychiatric Genetics, 156B*(1), 36–43. https://doi.org/10.1002/ajmg.b.31135

Kruglyak, L., Daly, M.J., Reeve-Daly, M.P., Lander, E.S. (1996) Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet, 58*, 1347–1363.

Lander, E., & Kruglyak, L. (1995). Genetic dissection of complex traits: Guidelines for interpreting and reporting linkage results. *Nature Genetics, 11*(3), 241–247. https://doi.org/10.1038/ng1195–241

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics, 25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Li, Y., Vinckenbosch, N., Tian, G., Huerta-Sanchez, E., Jiang, T., Jiang, H., … Nielsen, R. (2010) Resequencing of 200 human exomes identifies an excess of low-frequency nonsynonymous coding variants. *Nature Genetics*, *42*(11), 969–972. https://doi.org/10.1038/ng.680

Lohan, S., Spielmann, M., Doelken, S.C., Flottmann, R., Muhammad, F., Baig, S.M., … Klopocki, E. (2014). Microduplications encompassing the Sonic hedgehog limb enhancer ZRS are associated with Haas-type polysyndactyly and Laurin-Sand-

row syndrome. *Clinical Genetics, 86*(4), 318–325. https://doi.org/10.1111/cge.12352

Mahoney Rogers, A.A., Zhang, J., & Shim, K. (2011). Sprouty1 and Sprouty2 limit both the size of the otic placode and hindbrain Wnt8a by antagonizing FGF signaling. *Developmental Biology, 353*(1), 94–104. https://doi.org/10.1016/j.ydbio.2011.02.022

Mason, J.M., Morrison, D.J., Basson, M.A., & Licht, J.D. (2006). Sprouty proteins: Multifaceted negative-feedback regulators of receptor tyrosine kinase signaling. *Trends in Cell Biology, 16*(1), 45–54. https://doi.org/10.1016/j.tcb.2005.11.004

Mazzucchelli, C., & Brambilla, R. (2000). Ras-related and MAPK signalling in neuronal plasticity and memory formation. *Cellular and Molecular Life Sciences, 57*(4), 604-611. https://doi.org/10.1007/PL00000722

Meng, H., Smith, S.D., Hager, K., Held, M., Liu, J., Olson, R.K., … Gruen, J.R. (2005). DCDC2 is associated with reading disability and modulates neuronal development in the brain. *Proceedings of the National Academy of Sciences of the United States of America, 102*(47), 17053–17058. https://doi.org/10.1073/pnas.0508591102

Milne, R.L., & Antoniou, A. (2016). Modifiers of breast and ovarian cancer risks for BRCA1 and BRCA2 mutation carriers. *Endocrine-Related Cancer*, *23*(10), T69–T84. https://doi.org/10.1530/ERC-16-0277

Minowada, G., Jarvis, L.A., Chi, C.L., Neubuser, A., Sun, X., Hacohen, N., … Martin, G.R. (1999). Vertebrate Sprouty genes are induced by FGF signaling and can cause chondrodysplasia when overexpressed. *Development, 126*(20), 4465–4475.

Najmabadi, H., Hu, H., Garshasbi, M., Zemojtel, T., Abedini, S.S., Chen, W., … Ropers, H.H. (2011). Deep sequencing reveals 50 novel genes for recessive cognitive disorders. *Nature, 478*(7367), 57–63. https://doi.org/10.1038/nature10423

NHLBI. (2013). Exome variant server. Retrieved January 2013 from http://evs.gs.washington.edu/EVS/

Nopola-Hemmi, J., Myllyluoma, B., Haltia, T., Taipale, M., Ollikainen, V., Ahonen, T., … Widen, E. (2001). A dominant gene for developmental dyslexia on chromosome 3. *Journal of Medical Genetics, 38*(10), 658–664.

O'Connell, J.R., Weeks, D.E. (1998) PedCheck: a program for identification of genotype incompatibilities in linkage analysis. *Am J Hum Genet, 63*(1), 259–266. doi:10.1086/301904

Orban, P.C., Chapman, P.F., & Brambilla, R. (1999). Is the Ras-MAPK signalling pathway necessary for long-term memory formation? *Trends in Neurosciences, 22*(1), 38–44.

Peter, B., Raskind, W.H., Matsushita, M., Lisowski, M., Vu, T., Berninger, V.W., … Brkanac, Z. (2011). Replication of CNTNAP2 association with nonword repetition and support for FOXP2 association with timed reading and motor activities in a dyslexia family sample. *Journal of Neurodevelopmental Disorders, 3*(1), 39–49. https://doi.org/10.1007/s11689-010-9065-0

Peterson, R.L., & Pennington, B.F. (2012). Developmental dyslexia. *Lancet, 379*(9830), 1997–2007. https://doi.org/10.1016/S0140-6736(12)60198-6

Purcell, S., Cherny, S.S., & Sham, P.C. (2003). Genetic power calculator: Design of linkage and association genetic mapping studies of complex traits. *Bioinformatics, 19*(1), 149–150. https://doi.org/10.1093/bioinformatics/19.1.149

R Core Team (2020). R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*. URL https://www.R-project.org.

Rabin, M., Wen, X.L., Hepburn, M., Lubs, H.A., Feldman, E., & Duara, R. (1993). Suggestive linkage of developmental dyslexia to chromosome 1p34–p36. *Lancet, 342*(8864), 178.

Raskind, W.H., Igo, R.P., Chapman, N.H., Berninger, V.W., Thomson, J.A., Matsushita, M., … Wijsman, E.M. (2005). A genome scan in multigenerational families with dyslexia: Identification

of a novel locus on chromosome 2q that contributes to phonological decoding efficiency. *Molecular Psychiatry, 10*(7), 699–711. https://doi.org/10.1038/sj.mp.4001657

Rauen, K.A. (2013). The RASopathies. *Annual Review of Genomics and Human Genetics, 14*, 355–369. https://doi.org/10.1146/annurev-genom-091212-153523

Remschmidt, H., Schmidt, M.H., & Poustka, F. (2001). *Multiaxiales Klassifikationsschema für psychische Störungen des Kindes – und Jugendalters nach ICD-10 der WHO* (4th ed.). Hans Huber.

Risso, D., Schwartz, K., Sherlock, G., & Dudoit, S. (2011). GC-content normalization for RNA-Seq data. *BMC Bioinformatics, 12*, 480–496. https://doi.org/10.1186/1471-2105-12-480

Rüschendorf, F., Nürnberg, P. (2005) ALOHOMORA: a tool for linkage analysis using 10K SNP array data. *Bioinformatics*, *21*(9), 2123–2125. doi:10.1093/bioinformatics/bti264

Scerri, T.S., Darki, F., Newbury, D.F., Whitehouse, A.J., Peyrard-Janvid, M., Matsson, H., … Paracchini, S. (2012). The dyslexia candidate locus on 2p12 is associated with general cognitive ability and white matter structure. *PLoS One, 7*(11), e50321. https://doi.org/10.1371/journal.pone.0050321

Schulte-Korne, G. (2010). The prevention, diagnosis, and treatment of dyslexia. *Deutsches Ärzteblatt International, 107*(41), 718–726; quiz 727. https://doi.org/10.3238/arztebl.2010.0718

Schumacher, J., Hoffmann, P., Schmäl, C., Schulte-Körne, G., Nöthen, M.M. (2007) Genetics of dyslexia: the evolving landscape. *J Med Genet, 44*(5), 289–297.

Schwarz, J.M., Cooper, D.N., Schuelke, M., & Seelow, D. (2014). MutationTaster2: Mutation prediction for the deep-sequencing age. *Nature Methods, 11*(4), 361–362. https://doi.org/10.1038/nmeth.2890

Sherry, S.T., Ward, M.H., Kholodov, M., Baker, J., Phan, L., Smigielski, E.M., & Sirotkin, K. (2001). dbSNP: The NCBI database of genetic variation. *Nucleic Acids Research, 29*(1), 308–311.

Shihab, H.A., Gough, J., Cooper, D.N., Stenson, P.D., Barker, G.L., Edwards, K.J., Day, I.N., & Gaunt, T.R. (2013) Predicting the functional, molecular, and phenotypic consequences of amino acid substitutions using hidden Markov models. *Human mutation*, *34*(1), 57–65. https://doi.org/10.1002/humu.22225

Shihab, H.A., Rogers, M.F., Gough, J., Mort, M., Cooper, D.N., Day, I.N., Gaunt, T.R., & Campbell, C. (2015) An integrative approach to predicting the functional effects of non-coding and coding sequence variation. *Bioinformatics* (Oxford, England), *31*(10), 1536–1543. https://doi.org/10.1093/bioinformatics/btv009t

Shim, K., Minowada, G., Coling, D.E., & Martin, G.R. (2005). Sprouty2, a mouse deafness gene, regulates cell fate decisions in the auditory sensory epithelium by antagonizing FGF signaling. *Developmental Cell, 8*(4), 553–564. https://doi.org/10.1016/j.devcel.2005.02.009

Smith, S.D., Kimberling, W.J., Pennington, B.F., & Lubs, H.A. (1983). Specific reading disability: Identification of an inherited form through linkage analysis. *Science, 219*(4590), 1345–1347.

Sobreira, N.L., Cirulli, E.T., Avramopoulos, D., Wohler, E., Oswald, G.L., Stevens, E.L., … Goldstein, D.B. (2010). Whole-genome sequencing of a single proband together with linkage analysis identifies a Mendelian disease gene. *PLoS Genetics, 6*(6), e1000991. https://doi.org/10.1371/journal.pgen.1000991

Stein, J. (2014). Dyslexia: The role of vision and visual attention. *Current Developmental Disorders Reports, 1*(4), 267–280. https://doi.org/10.1007/s40474-014-0030-6

Stohr, N., & Huttelmaier, S. (2012). IGF2BP1: A post-transcriptional "driver" of tumor cell migration. *Cell Adhesion and Migration, 6*(4), 312–318. https://doi.org/10.4161/cam.20628

Taipale, M., Kaminen, N., Nopola-Hemmi, J., Haltia, T., Myllyluoma, B., Lyytinen, H., … Kere, J. (2003). A candidate gene for developmental dyslexia encodes a nuclear tetratricopeptide repeat do-

main protein dynamically regulated in brain. *Proceedings of the National Academy of Sciences of the United States of America, 100*(20), 11553–11558. https://doi.org/10.1073/pnas.1833911100

Thambirajah, M.S. (2010). Developmental dyslexia: An overview. *BJPsych Advances, 16*, 299–307. https://doi.org/10.1192/apt.bp.108.006072

Thum, T., Gross, C., Fiedler, J., Fischer, T., Kissler, S., Bussen, M., ... Engelhardt, S. (2008). MicroRNA-21 contributes to myocardial disease by stimulating MAP kinase signalling in fibroblasts. *Nature, 456*(7224), 980–984. https://doi.org/10.1038/nature07511

WHO. (1996). *Multiaxial classification of child and adolescent psychiatric disorders: The ICD-10 classification of mental and behavioural disorders in children and adolescents.* Cambridge University Press.

Wijsman, E.M., Peterson, D., Leutenegger, A.L., Thomson, J.B., Goddard, K.A., Hsu, L., ... Raskind, W.H. (2000). Segregation analysis of phenotypic components of learning disabilities. I. Nonword memory and digit span. *American Journal of Human Genetics, 67*(3), 631–646. https://doi.org/10.1086/303044

Authors' Note

Masoud Garshasbi is currently at the Department of Medical Genetics, Faculty of Medical Sciences, Tarbiat Modares University, Tehran, Iran.

Wei Chen is currently at the Department of Biology, Southern University of Science and Technology, 1088 Xueyuan Rd, Nanshan District, Shenzhen, Guangdong, China.

Reinhard Ullmann is currently at the Bundeswehr Institute of Radiobiology, Munich, Germany.

Conflict of interest

The authors declare that they have no conflict of interest.

Andreas W. Kuss
Interfaculty Institute for Genetics and Functional Genomics, University Medicine Greifswald, Greifswald, Germany
Felix-Hausdorff-Str. 8
17475 Greifswald
Germany

kussa@uni-greifswald.de