CrossMark

ORIGINAL INVESTIGATION

Multipoint genome-wide linkage scan for nonword repetition in a multigenerational family further supports chromosome 13q as a locus for verbal trait disorders

D. T. Truong[1] · L. D. Shriberg[2] · S. D. Smith[3] · K. L. Chapman[4] · A. R. Scheer-Cohen[5] ·
M. M. C. DeMille[1] · A. K. Adams[6] · A. Q. Nato[7] · E. M. Wijsman[7,8] · J. D. Eicher[6] ·
J. R. Gruen[1,6,9]

Abstract Verbal trait disorders encompass a wide range of conditions and are marked by deficits in five domains that impair a person's ability to communicate: speech, language, reading, spelling, and writing. Nonword repetition is a robust endophenotype for verbal trait disorders that is sensitive to cognitive processes critical to verbal development, including auditory processing, phonological working memory, and motor planning and programming. In the present study, we present a six-generation extended pedigree with a history of verbal trait disorders. Using genome-wide multipoint variance component linkage analysis of nonword repetition, we identified a region spanning chromosome 13q14–q21 with LOD = 4.45 between 52 and 55 cM, spanning approximately 5.5 Mb on chromosome 13. This region overlaps with SLI3, a locus implicated in reading disability in families with a history of specific language impairment. Our study of a large multigenerational family with verbal trait disorders further implicates the SLI3 region in verbal trait disorders. Future studies will further refine the specific causal genetic factors in this locus on chromosome 13q that contribute to language traits.

Introduction

Verbal trait disorders are comorbid, developmentally associated disorders and deficits in communication. These include clinical and subclinical disorders of speech, language, reading, spelling, and writing (Shriberg et al. 2012). Speech sound disorders (i.e., excluding dysfluency) are the most prevalent verbal trait disorders at preschool age, with an estimated population prevalence of 16 % at age 4 years (Campbell et al. 2003), decreasing to 3.8 % at age 6 years (Shriberg et al. 1999) and 3.6 % at age 8 years (Wren et al. 2012). Deficits in speech frequently co-occur with impairments in multiple domains. For example, 11–15 % of children with speech sound disorders at age 6 also have language disorder (a neurodevelopmental disorder that can affect both spoken or written language; Shriberg et al.

✉ J. R. Gruen
jeffrey.gruen@yale.edu

1 Department of Pediatrics, Yale School of Medicine, New Haven, CT 06510, USA

2 Waisman Center, University of Wisconsin-Madison, Madison, WI 53705, USA

3 Department of Pediatrics, University of Nebraska at Omaha, Omaha, NE 68182, USA

4 Department of Communication Sciences and Disorders, University of Utah, Salt Lake City, UT 84112, USA

5 Department of Speech-Language Pathology, California State University, San Marcos, CA 92096, USA

6 Department of Genetics, Yale School of Medicine, New Haven, CT 06510, USA

7 Division of Medical Genetics, Department of Medicine, University of Washington, Seattle, WA 98195, USA

8 Department of Biostatistics and Department of Genome Sciences, University of Washington, Seattle, WA 98195, USA

9 Investigative Medicine Program, Yale School of Medicine, New Haven, CT 06510, USA

 Springer

1999). Additionally, children with speech disorders are at higher risk for reading disability, with an estimated 18 % of children with speech disorders and 75 % of children with both speech and language disorders meeting criteria for reading disability at school age (Lewis et al. 2000).

Research in the genetics of verbal trait disorders was catalyzed by the seminal studies of the KE family, a large extended pedigree segregating verbal dyspraxia (also termed Childhood Apraxia of Speech; ASHA 2007; RCSLT 2011), suggesting autosomal dominant inheritance of a single gene mutation (Hurst et al. 1990). Genome-wide linkage showed a signal peak on chromosome 7q31.1 (Fisher et al. 1998). Further fine mapping identified a point mutation in *FOXP2* that resulted in a truncated protein and loss of function in all affected individuals, but not observed in unaffected individuals (Lai et al. 2001). *FOXP2* loss of function as a causal factor for verbal dyspraxia was further validated in unrelated individuals with severe speech impairments similar to those in the KE family (Lai et al. 2001; MacDermot et al. 2005) and has been cross-validated in a number of case studies (e.g., Rice et al. 2012; Shriberg et al. 2006). Although the KE family provided an example of a verbal trait disorder phenotype with a typical pattern of monogenic inheritance, their story is the exception rather than the norm. In fact, verbal trait disorders are generally multifactorial and associated with multiple genetic and environmental factors (Kang and Drayna 2011; Peterson and Pennington 2015).

Due to the behavioral and cognitive heterogeneity of verbal trait disorders, the use of endophenotypes—underlying phenotypic factors that are associated with or contribute to the manifestation of the disorder of interest because of shared genetic factors—have been critical to the genetic study of verbal trait disorders. One endophenotype is nonword repetition (NWR), which loads onto several cognitive processes critical for language-related ability including auditory processing, receptive language ability, and motor planning and programming (Dollaghan and Campbell 1998). NWR tasks examine the ability to process and temporarily store a novel series of meaningless units of phonological information in short-term memory, and then verbally repeat the stimuli. Such measures, which are sensitive to but not specific for any one disorder, may be more closely influenced by genetic variation than the verbal trait disorder itself.

NWR task performance has a strong genetic influence, with higher concordance among monozygotic twins compared to dizygotic twins, and heritability ranging from 0.64 to 1 (Bishop et al. 1996, 2004). Furthermore, an oligogenic-trait segregation analysis of NWR in nuclear families ascertained for reading disability estimated approximately 2.4 quantitative trait loci (Wijsman et al. 2000). A family-based linkage analysis on individuals with a family history

of specific language impairment (SLI) identified a linkage peak on chromosome 16q for poor performance on NWR (SLI Consortium 2002, 2004). Follow-up family-based and population-based association studies on NWR identified *CMIP* and *ATP2C2* as candidates responsible for the linkage signal (Newbury et al. 2009). *CNTNAP2* on chromosome 7q35 was also associated with NWR in families enrolled in the SLI consortium study using a candidate gene approach after it was identified as a transcriptional binding target of *FOXP2* by chromatin immunoprecipitation (Vernes et al. 2008). In addition, CNTNAP2 was identified by fine mapping a linkage analysis signal on 7q35 conditioned on language delay in the Autism Genetic Resource Exchange sample (Alarcón et al. 2002, 2005, 2008). Taken together, NWR satisfies specific testable criteria for the objective identification of endophenotypes, supporting NWR as a credible endophenotype for verbal trait disorders (Glahn et al. 2014; Lenzenweger 2013).

With the exception of the KE family, most families with a history of language impairment show a complex pattern of inheritance with subtle differences in clinical presentation within the family. In the present study, we examined an extended six-generation family with a complex pattern of inheritance for verbal trait disorders. We chose NWR in this analysis because (1) it is a robust endophenotype for verbal trait disorders (i.e., speech sound disorder, language disorder, and developmental dyslexia); (2) is highly heritable; (3) has a Mendelian model of inheritance (in at least one study; Wijsman et al. 2000); and (4) is stable throughout an individual's lifetime, even in those who are language recovered following impairment in childhood (Bishop et al. 1996; Shriberg et al. 2009). The latter attribute of NWR tasks is particularly important because subjects within this family range in age from 3 to 95 years, requiring a phenotype that can be ascertained and compared across all age groups. The present analysis provides strong support for chromosome 13q14–q21 as a locus that contributes to poor performance on NWR in this extended pedigree.

Methods and materials

Ascertainment

We studied 62 individuals from a six-generation 90-member family of European ancestry with a history of verbal trait disorders. The family was ascertained with the assistance of a family member. The 62 family members assessed included 35 females and 27 males ranging in age from 3 to 95 years. There is no evidence of consanguinity based on genealogy or unexpected high kinship coefficients within the pedigree.

Written informed consent was approved by the University of Wisconsin-Madison Institutional Review Board (IRB). All subjects were assessed by one of two experienced examiners in the participants' homes or hotel sites in five states within the continental US. All oral instructions and audio-recorded stimuli were presented at comfortable listening levels based on findings from a conventional hearing screening. The assessment protocol included the following measures and instruments: Kaufman Brief Intelligence Test-2 (Kaufman and Kaufman 2004; nonverbal and verbal IQ), Nonword Repetition Task (NRT; Dollaghan and Campbell 1998), Syllable Repetition Task (SRT; Shriberg et al. 2009), Goldman-Fristoe Test of Articulation-2 (Goldman and Fristoe 2000; speech), Clinical Evaluation of Language Fundamentals-Preschool-2 (Wiig et al. 2004; language), Clinical Evaluation of Language Fundamentals-4 Screening Test (Semel et al. 2004; language), Woodcock–Johnson Tests of Achievement, 3rd edition (Woodcock et al. 2001; reading, spelling, and writing), and questionnaires for parent-reporting or self-reporting medical and special educational histories and concerns. One individual had a composite IQ <75, but performance on the NWR tasks was unimpaired. To maximize genetic informativeness, this individual's NWR scores were retained for the analysis. All other individuals had an IQ between 84 and 126.

Phenotype

The NRT is a NWR task that consists of 16 nonwords. To reduce the articulatory burden, the NRT does not contain the most phonetically complex consonants (the "late-8" consonants; Shriberg 1993; Dollaghan and Campbell 1998). Nonwords ranged in length from one to four syllables (four each) with the shortest nonwords presented first and the longest last. Each repeated consonant and vowel/diphthong (totaling 20 different phonemes) was later transcribed as correct or incorrect by two research speech pathologists. NRT scores were calculated by dividing the total phonemes correctly repeated by the total phoneme targets. Ratios were then converted to age–sex standardized scores for downstream analyses using a reference database of 200 typical speakers, ages 3–80 years (Potter et al. 2012; Scheer-Cohen et al. 2013) that included descriptive statistics for NRT and SRT scores.

The SRT is another NWR task comprised of 18 nonwords that include only four of the "early-8" consonants (/b/,/d/,/m/, and n/) and the vowel/ɑ/(Shriberg 1993). This NWR task was designed to accommodate individuals who have incomplete phonetic inventories and/or articulatory impairments. Items range in length from two to four syllables with the shortest presented first and the longest last. The consonant responses to each recorded syllable were transcribed as correct or incorrect. The number of correctly repeated consonants was divided by the total number of target consonants. The ratio was then converted to a standard score using the reference database.

Studies of speech-language disorders using the NRT have supported its validity and reliability (e.g., Archibald and Gathercole 2006; Moore et al. 2010), including a reference sample of 95 children with typical speech and 63 children with speech delay, described in a technical report on the NRT and SRT (Shriberg and Lohmeier 2008). Findings from this reference sample include psychometric data supporting the distributional characteristics of scores for parametric statistical analyses, and analyses supporting the construct validity, concurrent validity, interjudge transcription reliability, and internal reliabilities of both tasks.

In the present study, point-to-point percentage of agreement estimates ranged from 75.6 to 88 % across nonword task and phoneme class; other validity and reliability estimates were generally in the 0.70–0.85 range. The Pearson r coefficient between standardized scores on the two nonword tasks in the present data was 0.66, consistent with the coefficient of 0.73 reported in Shriberg and Lohmeier (2008). Thus, consistent with discussion elsewhere, there is only moderate collinearity between the two measures of NWR (Shriberg et al. 2009).

Sensitivity and specificity for identifying speech and language disorders using the SRT were further supported by a second reference sample of 550 speakers, including speakers with typical speech and typical language, speech delay and typical language, language impairment and typical speech, and speech delay and language impairment (Lohmeier and Shriberg 2011). Additional construct validity support for the SRT was presented in Shriberg et al. (2009), followed by a series describing SRT procedures to explicate encoding, memorial, and transcoding processes underlying performance on nonword imitation tasks (Shriberg et al. 2012).

Because there is no battery of speech, language, reading, spelling, and writing tests appropriate for the lifespan ages of the present extended family, we used standardized scores from either the NRT or SRT to assign a categorical phenotype. Verbal trait impaired (Verbal Trait+) was defined as performing greater than one standard deviation below the mean on either the NRT or SRT. Preliminary studies indicated that a cutoff below one standard deviation on either the NRT or the SRT was maximally sensitive and specific to subjects with only mild, subclinical difficulty in one or more of the five verbal traits based on parent- and self-reported histories of children and adults. Of the 41.9 % of participants in the present study who met the nonword criteria for a verbal trait disorder (see Table 1), 19.2 % met criteria on the NRT only, 23.1 % met criteria on the SRT only, and 57.7 % met criteria on both nonword tasks.

Table 1 Percentages[a] of affected (Verbal Trait+) and not affected (Verbal Trait−) participants in an extended family of 62 members and tests of two proportions results for each variable. Participant age was divided into four lifespan cohorts

Variable	Total n	Verbal trait+ (VT+)		Verbal trait− (VT−)		Tests of two proportions[b]			
		n	%	n	%	Z	p	Confidence interval	Sig.[c]
Participants	62	26	41.9	36	58.1	−1.82	0.069	−0.335, 0.012	
Gender									
Female	35	11	31.4	24	68.6	−3.35	0.001	−0.589, −0.154	*
Male	27	15	55.6	12	44.4	0.82	0.411	−0.154, 0.376	
Age									
Preschool (3–5)	3	1	33.3	2	66.7	−0.87	0.386	−1.000, 0.421	
School age (6–18)	21	7	33.3	14	66.7	−2.29	0.022	−0.618, −0.048	*
Adult (19–64)	30	14	46.7	16	53.3	−0.52	0.605	−0.319, 0.186	
Senior (65–84)	8	4	50.0	4	50.0	0.00	1.000	−0.490, 0.490	
Verbal trait history									
Speech	15	11	73.3	4	26.7	2.89	0.004	0.150, 0.783	*
Language	17	8	47.1	9	52.9	−0.34	0.731	−0.394, 0.277	
Reading	24	13	54.2	11	45.8	0.58	0.562	−0.199, 0.365	
Spelling	18	10	55.6	8	44.4	0.67	0.502	−0.214, 0.436	
Writing	3	1	33.3	2	66.7	−0.87	0.386	−1.000, 0.421	
Participants scoring more than one SD below the mean in one or more verbal trait domains	37	19	73.1	18	50.0	1.92	0.055	−0.005, 0.469	

[a] The row-wise percentages use the Total n in the second column as the denominator. The denominators for each percentage in the last row are 26 and 36, respectively

[b] Minitab 17 Statistical Software (2010). [Computer software]. State College, PA: Minitab, Inc. (www.minitab.com)

[c] * $p < 0.05$

Participants

Tables 1, 2 describe demographic and phenotype variables for participants meeting NWR task criteria for affected (Verbal Trait+; VT+) and not affected (Verbal Trait−; VT−), including tests for significant differences between the proportions of each classification. The difference in the percentages of VT+ (41.9 %) compared to VT− (58.1 %) participants was non-significant ($Z = −1.82$). Significantly fewer females met criteria for VT+ (31.4 %) than VT− (68.6 %; $Z = −3.35$), but the proportion of males who met criteria for VT+ (55.6 %) compared to the proportion who met criteria for VT− (44.4 %) was non-significant ($Z = 0.82$). Among the four age groups, the only age group within which affection status differed significantly was the school-aged participants, who had a significantly lower percentage of participants who met criteria for VT+ (33.3 %) than VT− (66.7 %; $Z = −2.29$).

Last, Verbal Trait History for problems in verbal trait domains of speech, language, reading, spelling, and/or writing were determined by test scores in any of the relevant

Table 2 Descriptives for VT+ and VT− individuals in the family across the syllable repetition and nonword repetition tasks

	Verbal trait+ (VT+)	Verbal trait− (VT−)
SRT		
Mean (SD)	−3.01 (3)	0.27 (0.7)
Skewness	−0.96	0.25
Kurtosis	−0.05	0.53
NRT		
Mean (SD)	−2.07 (1.67)	0.42 (0.86)
Skewness	−0.22	0.04
Kurtosis	−0.16	−0.77

domains lower than one standard deviation below standardized means, or any self- or parent-reported difficulty in any of the five domains (Supplemental Table 1). Of the Verbal Trait History variables in Table 1, only one verbal trait domain was associated with significant between group proportions. A significantly greater percentage of participants with test scores, self-reported, or parental-reported histories

of speech disorders met the nonword criterion for VT+ (73.3 %) compared to the percentage who met criterion for VT− (26.7 %; $Z = 2.89$). Using conventional criteria for statistical significance, the percentage of VT+ participants who had at least one test score or questionnaire entry indicating a concern with any one of the five verbal traits (73.1 %) was not significantly larger than the percentage of VT− participants with such histories (50.0 %; $Z = 1.92$; $p = 0.055$; CI −0.005, 0.469).

DNA collection and genotyping

DNA was extracted from whole blood using the Gentra Puregene Blood Kit (Qiagen) at the University of Nebraska Medical Center. Genotyping across 551,839 single nucleotide polymorphism (SNP) markers was performed using the Illumina Infinium HumanCoreExome-24-v.1 at the Yale Center for Genome Analysis (Orange, CT). Genotypes were called using Illumina GenomeStudio with a total of 547,644 (99.24 %) passing quality control (QC). One individual failed QC due to low genotyping call rate and was excluded from the analysis.

PBAP: marker sub-selection, pedigree structure validation, and IBD computation for linkage analysis

Reference map files for the HumanCoreExome dense marker panel were obtained from the Rutgers maps (Matise et al. 2007), with integrated linkage-physical maps in sex averaged Haldane genetic distances (cM). Reference genotype data for Europeans were extracted from the main European (EUR) population data from the 1000 genomes project (The 1000 Genomes Project Consortium 2010) to determine linkage disequilibrium (LD) and minor allele frequencies (MAF) between markers for marker sub-selection.

We used the pedigree based analysis pipeline (PBAP) to sub-select genetic markers for pedigree quality control (QC) and for interfacing with MORGAN (Thompson 2011) to calculate inheritance vectors (IV) used for linkage analysis (Nato et al. 2015). Use of MORGAN allowed multipoint analysis on the complete pedigree. Generation of genome-wide SNP marker sub-panels from the dense marker panel was conducted to (1) reduce LD between markers and minimize type 1 error, (2) reduce computational time (while maximizing genotypic informativeness within the pedigree), and (3) perform QC on pedigree structure (i.e., parent–offspring swaps). PBAP marker sub-selection and pedigree structure validation are described in detail elsewhere (Nato et al. 2015). Briefly, three non-overlapping marker sub-panels from the original dense marker panel (Illumina HumanCoreExome-24-v.1) were generated based on the following criteria: (1) maximum LD (r^2) threshold equal to 0.04 in the EUR reference population;

(2) MAF 0.2–0.5 in the EUR reference population; (3) non-monomorphic marker within the pedigree;(4) minimum intermarker distance of 0.5 cM; and (5) restricted to the 22 autosomes. A separate marker sub-panel was generated for pedigree structure validation using similar criteria as above except maximum LD threshold was equal to 0.25 and MAF from 0.3 to 0.5 in the EUR reference population. For genome-wide linkage analysis (excluding sex chromosomes), a total of 5448, 5493, and 5498 markers for sub-panels 1, 2, and 3, respectively, were created. A sub-panel of 5454 genome-wide markers was created for pedigree structure validation.

QC for appropriate parent–offspring relationships within the larger pedigree was assessed by comparing expected kinship coefficients (based on pedigree structure) and estimated coefficients computed by maximizing the likelihood from available genotype data across the 5454 marker sub-panel (Choi et al. 2009). Individual relationship pairs were flagged if the estimated kinship coefficient fell outside a 99.5 % confidence interval from expected. No sample swaps or incorrect parent–offspring relationships were observed within the larger pedigree.

From each marker sub-panel created for linkage analysis, PBAP prepared data files to generate IVs that described the flow of genetic information through a pedigree for an individual using gl_auto in the MORGAN suite of programs (Thompson 2011). The gl_auto program uses a combination of exact and Markov Chain Monte Carlo (MCMC) based estimations to sample IVs for each individual. For the current analysis, IVs were sampled for each marker subpanel using the following parameters: 15,000 MCMC burn-in iterations, sampling by scan and 100,000 MCMC iterations with progress checked every 20,000 iterations (L-Sampler = 0.2), saving 2000 realizations for IV sampling. Sampled IVs were then converted to Sequential Oligonucleotide Linkage Analysis Routines (SOLAR) (Almasy and Blangero 1998) compatible multipoint identity-by-descent (MIBD) matrices using custom scripts written by the Wijsman lab, and imported into SOLAR for downstream linkage analysis.

Statistical analysis

All statistical analyses were conducted using the SOLAR software package (version 7.3.9; Almasy and Blangero 1998). SOLAR utilizes a maximum likelihood variance decomposition approach to estimate the influence of genetic and environmental effects on a phenotypic trait by modeling the covariance among family members relative to genetic kinship (identity by descent). A liability threshold model was used to handle discrete traits under the assumption that the affection status of an individual was determined by their underlying genetic risk exceeding a certain

threshold for the phenotype (i.e., VT+; Duggirala et al. 1997). Using maximum likelihood techniques, initial models were screened for the covariate effects of age, age^2, sex, age × sex, age^2 × sex, and IQ. After covariate screening non-significant covariates ($p > 0.1$) were removed from the final model. In addition, a variance component for household random effects that further controlled for shared environment among nuclear families within the larger pedigree was included. The final model representing the log likelihood when the additive genetic variance was equal to 0 (no linkage elements) and covarying for household and age*sex effects, was used as the null model for hypothesis testing during linkage analysis.

Genome-wide multipoint variance component linkage analyses were conducted to examine linkage between VT+ and MIBDs. Multipoint linkage analysis considers recombination along a chromosome to determine the probability that a trait locus is located within a genomic region. Maximum likelihood estimates for linkage were calculated at approximately 0.5 cM intervals across the 22 autosomes and compared against the null model (no linkage) using a likelihood ratio test ($df = 1$).

Empirical p values were computed using a simulation that generated a distribution of LOD scores under a null model of no linkage. 1,000,000 simulations were conducted, each generating a random informative marker that was tested for linkage with VT+ status. This distribution of observed LOD scores in the simulation was then used to determine the empirical p value of the experimentally observed LOD scores.

Haplotypes were assigned using MERLIN (Abecasis et al. 2002). The large pedigree exceeded the bit limit MERLIN could handle, thus the family was split into six smaller subpedigrees for haplotype assignment and then manually reconfigured to confirm consistency of haplotypes called across the lineages. Five of the six subpedigrees were assigned based on distinct sublineages that originated from Generation II, and included all individuals within the last four generations (Generations III–VI) of the family. The sixth subpedigree consisted of all individuals in the first two generations (Generations I and II) and select individuals in Generations III and IV to confirm the transmission of the haplotypes observed in the aforementioned five subpedigrees.

Results

Genome-wide multipoint linkage analyses for VT+ revealed a peak LOD score of 4.35 (empirical p value $<1 × 10^{-6}$) between 52 and 55 cM on chromosome 13q14.2–q14.3 (Fig. 1 and Supplemental Table 2) with marker subpanel 1. This region spans base pair positions 48–53.5 Mb across 5

linkage (SNP) markers on chromosome 13, encoding a total of 41 genes (build GRCh37/hg19). To determine whether the linkage signal was due to an effect of pseudorandom marker sub sampling, multipoint analyses on chromosome 13 were conducted again using non-overlapping marker subpanels 2 and 3, which were generated at the same time as marker subpanel 1 in PBAP. Findings were recapitulated with peak LOD scores of 4.24 and 3.96 using marker subpanels 2 and 3, respectively, again spanning the same region of chromosome 13q14.2–14.3 (Supplemental Table 3). Of the 41 genes within base pair positions 48–53.5 Mb, 26 genes show moderate expression in the developing brain (BrainSpan 2011), but only 8 have known neurological or cognitive function (Carrozzo et al. 2007; de Bie et al. 2007; Dening et al. 1989; Elpeleg et al. 2005; Hilschmann et al. 2002; Jaberi et al. 2013; Kind et al. 2014; La Piana et al. 2016; Maas et al. 2015; Morris et al. 2012; Ocklenburg et al. 2015; Ostergaard et al. 2007; Rice et al. 2007; Spiechowicz et al. 2006; Vidal et al. 1999, 2000; Wei and Hemmings 2005; Xu et al. 2010; Yamagata et al. 1999; Yasuda et al. 2007; Zhang et al. 2006; Supplemental Table 4). When considering LOD scores greater than 3 (empirical p value <0.0002), the linkage signal expands to 46–61 cM across 22 linkage markers spanning a 23.2 Mb region on chromosome 13q14.11-21.32, encompassing approximately 77 genes.

Estimation of haplotypes in the extended family provides evidence of at least 3 distinct haplotypes on chromosome 13 segregating with VT+ (Fig. 2). A recombination in Haplotype 1 (Haplotype 1-recombined) in affected individual 3 between SNPs rs7337528 and rs2981 defines the centromeric boundary of LOD = 4.35 at 52 cM (Figs. 2, 3). This particular segment of Haplotype 1, defined by alleles shared IBD at rs2981, rs2812219, rs6561602, and rs997687 (52.92–55.06 cM), segregates with the phenotype in the lineage originated by a founder in the oldest generation. For the remaining information, data are presented without a conventional haplotype pedigree graphic to preserve anonymity of the family. Haplotype 2 originates from a married-in founder in Generation II and segregates to two siblings in Generation III, four descendants in Generation IV and four descendants in Generation V. Seven of ten descendants with Haplotype 2 are VT+. Last, Haplotype 3 originates from a married-in founder in Generation III, segregates to two siblings in generation IV, and four descendants in Generation V. Six of seven pedigree members with Haplotype 3 are VT+. There are no recombinations within Haplotypes 2 or 3 that would define the broader shoulders of the linkage signal from 45 to 62 cM (Fig. 3). Overall, the presence of at least three distinct haplotypes that cosegregate with VT+—including married-in Haplotypes 2 and 3—implicate multiple contributing variants segregating through this family.
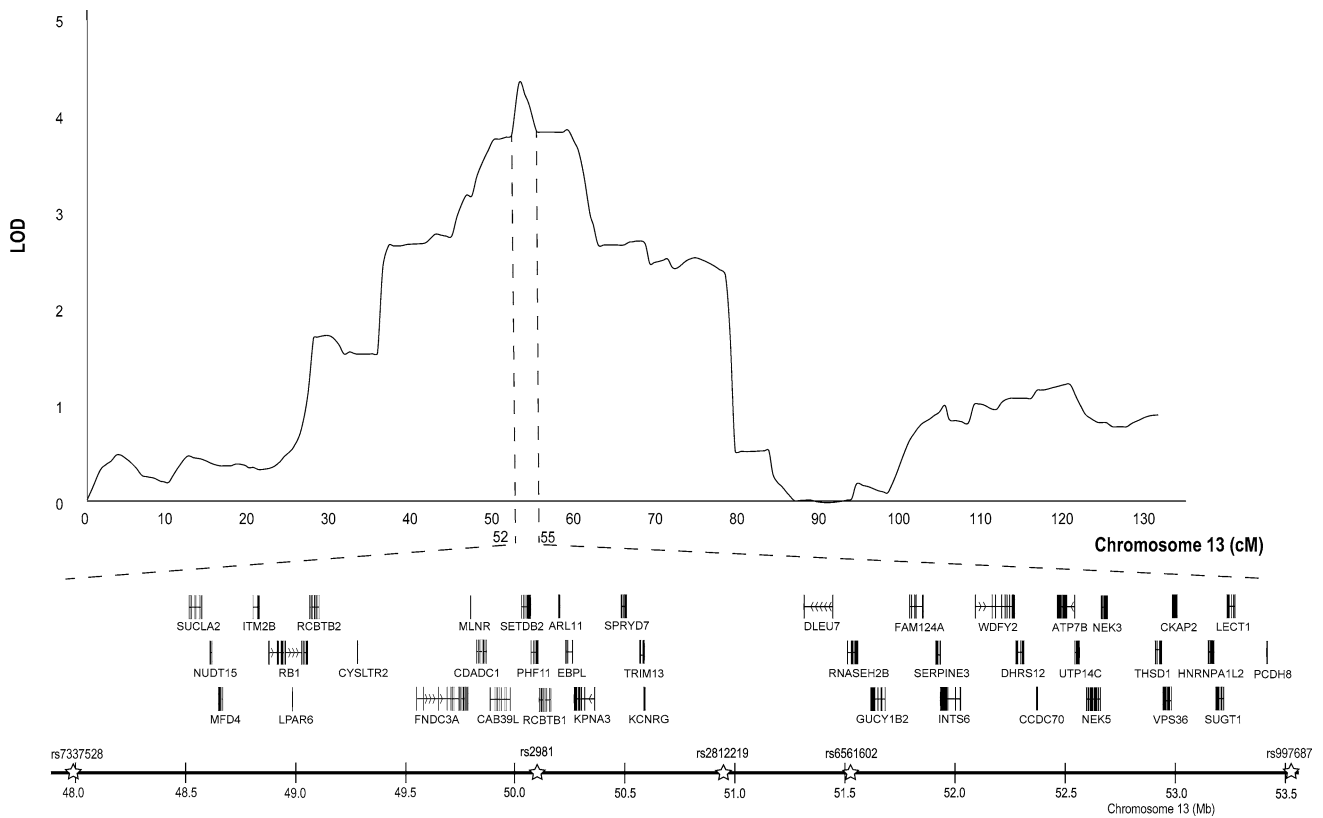
Fig. 1 Multipoint linkage results conditioned on impaired NWR at chromosome 13. Genes and associated SNPs under the highest linkage peak of LOD = 4.35, between 52 and 55 cM spanning physical positions 48–53.5 Mb on reference genome assembly build GRCh37/hg19

Fig. 2 Haplotypes spanning genomic location 45–62 cM on chromosome 13 segregating with Verbal Trait+ (affected) status in the family. The centromeric and telomeric boundaries of Haplotype 1-Recombined are defined by a recombinatorial events within individual 3 (Fig. 3). No recombinatorial events in the family offer clear centromeric and telomeric boundaries for Haplotypes 2 and 3

| SNP	Chr13 Haldane map position (cM)	Haplotype 1-Recombined	Haplotype 2	Haplotype 3
exm-rs1324015	45.41		3	3
rs9525816	46.44		1	1
rs7326184	47.13		2	2
rs7335530	47.76		2	4
rs9595337	48.46		2	4
rs2573263	49.14		2	2
rs912975	49.80		1	1
rs9534442	50.54		4	4
rs1328671	51.35		4	2
rs7337528	52.02		1	1
rs2981	52.92	3	3	3
rs2812219	53.67	2	2	2
rs6561602	54.19	2	4	4
rs997687	55.05	3	1	1
rs2898060	55.59		2	2
rs9538051	56.37		3	3
rs1410270	56.89		2	3
rs9317116	57.65		2	4
rs9598276	58.29		4	1
rs9592236	58.93		4	4
rs9540208	59.68		2	3
rs7988964	60.22		4	1
rs9317622	60.84		3	3
rs287398	61.61		1	1
rs2784843	62.02		2	2
rs9542142	62.62		3	1
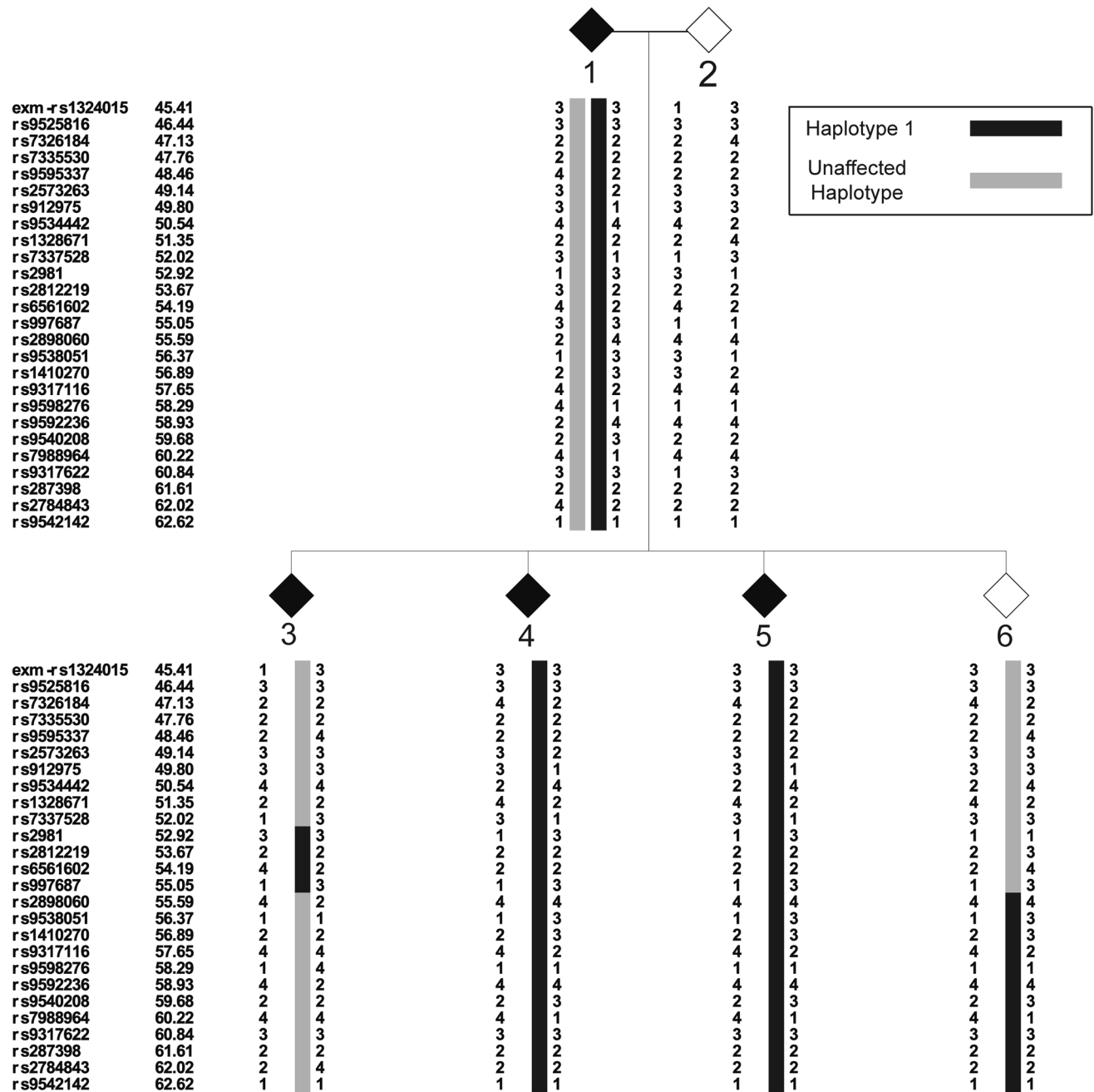
Fig. 3 Haplotype assignments spanning genomic location 45–62 cM on chromosome 13. The pedigree depicted is truncated to reflect the recombinatorial event observed in affected individual 3 that outlines the centromeric and telomeric border of the linkage signal spanning 52–55 cM. Affected individuals are black diamonds, while unaffected individuals are gray diamonds

Other suggestive linkage signals (LOD >1.5; empirical *p* value <0.01) were observed on chromosome 2q37.1, 4q12–13.2, 4q25, 7q22.3–31.2, 8q24.3, and 12p13.33 (Supplemental Table 2). Most notably, the linkage peak spanning 7q22.3–q31.2 has a max LOD = 2.06 and contains the gene FOXP2—a causal gene for Childhood Apraxia of Speech (Fisher et al. 1998; Lai et al. 2001).

Discussion

The present study identified a linkage signal spanning chromosome 13q14–q21 using a categorical phenotype (VT+ or VT−) derived from performance on NWR in an extended pedigree with a history of verbal trait disorders. This region encompasses SLI3 on chromosome 13q21,

a SLI locus previously identified by Bartlett et al. (2002), using a family-based linkage analysis in five Canadian families of Celtic ancestry with a history of specific language impairment (SLI). Their analysis was conditioned on a categorical reading-IQ discrepancy phenotype (nonword reading score at least one standard deviation below performance IQ), which they replicated in a larger independent US sample using the same phenotype (Bartlett et al. 2004). This region has also been implicated in autism spectrum disorder (ASD), a neurodevelopmental disorder with a core language component in combination with other core abnormalities in social and repetitive behaviors. A linkage signal at chromosome 13q21 was observed with a language delay phenotype in the Collaborative Linkage Study of Autism (CLSA; Bradford et al. 2001). Furthermore, deletions at 13q12 through 13q21 have been reported in three subjects with ASD and poor receptive and expressive vocabulary (but normal speech), and in a subject with ASD with auditory processing deficits (Mitter et al. 2011; Smith et al. 2002; Steele et al. 2001). These deleted segments partially overlap the linkage peak in our present study, where LOD >1.5 (Supplemental Tables 2 and 3). The convergence of these findings associated with 13q14–q21 with related language phenotypes that underlie verbal trait disorders provides compelling support for this locus.

Within this pedigree, there are at least three distinct haplotypes segregating with VT+, of which, only Haplotype 1 originated with a founder in the oldest generation—the other two are more recently married into, consistent with assortative mating. Within the EUR reference population of 1000 genomes project, Haplotype 1-Recombined (Fig. 2) is common with a frequency of 0.046. In a clinical context, verbal trait disorders such as developmental dyslexia and specific language impairment (SLI) have a high prevalence in the United States. The prevalence of developmental dyslexia is 7 % in the general population, and the prevalence of specific language impairment (SLI) is 5–8 % among preschool children (Peterson and Pennington 2015; Tomblin et al. 1997). Different haplotypes segregating within the family could indicate a single gene with different causal variants segregating within the family. It is also possible that different genes at the same locus, or at different loci, are mediating NWR performance. Further fine mapping and sequencing of the region is necessary to disentangle these possibilities and elucidate the potential variants driving the signal observed in the present study.

Underneath the peak linkage signal spanning the 52–55 cM region of chromosome 13 with LOD >4, there are interesting gene candidates with known function in neuropsychiatric disorders and neurodevelopment. *ITM2B* encodes a transmembrane protein that helps to inhibit the accumulation of beta-amyloid, but mutations have been implicated in Familial British Dementia and Familial Danish Dementia with similar pathology to Alzheimer disease (Vidal et al. 1999, 2000). *Setdb2*, the zebrafish ortholog of *SETDB2*, is known to regulate left–right asymmetry in the zebrafish central nervous system (Xu et al. 2010). Human epidemiological research has also associated *SETDB2* with handedness (left versus right preference) with specific variants linked to reduction in laterality (Ocklenburg et al. 2015). This provides an interesting parallel to previous evidence that suggests language and reading disability are linked to atypical cerebral laterality (asymmetry) since language-related behavior is typically left lateralized (Leonard and Eckert 2008; Scerri et al. 2011). However, it is important to note that *SETDB2* has not yet been directly implicated in reading or language disability. *ATP7B* is a known gene associated with Wilson disease, which is a disorder characterized by the deposition of copper in the liver, brain, and other tissues, leading to neurological and cognitive deterioration including memory loss, tremors, and emotional changes (de Bie et al. 2007; Dening Tr 1989). *PCDH8* is part of the protocadherin family of CNS-specific cell adhesion molecules that plays a role in the development of neural circuitry (Hilschmann et al. 2002; Yamagata et al. 1999). Interestingly, the rat ortholog of *PCDH8*, *Arcadlin*, has been implicated in synaptic function and is dynamically expressed upon activation of hippocampal circuitry—a neural network necessary for learning and memory (Yasuda et al. 2007).

By examining the extended linkage peak spanning 45–62 cM with LOD scores >3, we identified another three genes in the protocadherin family located telemetric to *PCDH8*—*PCDH17*, *PCDH20*, and *PCDH*—each of which encodes cell–cell adhesion molecules that are primarily expressed in the brain (Kim et al. 2010). Variation in *PCDH9* has been linked to ASD and SLI, while *PCDH17* is highly expressed in the prefrontal and anterior regions of the temporal cortex and subcortical structures such as the thalamus, ventral striatum, and anterior cingulate—an expression pattern highlighting an overlap with corticostriatothalamic circuitry critical for higher order cognitive function and language development (Abrahams et al. 2007; Marshall et al. 2008).

Our linkage findings on chromosome 13q do not correspond to other genome-wide linkage scans conditioned on NWR. A family-based linkage analysis conducted by the SLI consortium localized to chromosome 16q with further fine mapping identifying *CMIP* and *ATP2C2* as potential gene candidates mediating NWR in their sample (SLI Consortium 2002, 2004; Newbury et al. 2009). Another study performed by Brkanac et al. (2008), observed linkage signals on chromosomes 4p12, 12p, and 17q in families with a history of dyslexia. Discrepancies in genomic regions associated with NWR, in part, may be due to differences in the particular NWR test used to evaluate respective

subjects. Although the measures used in these studies do ostensibly assess NWR, there are differences in each that may more heavily tap into different combinations of underlying cognitive and/or behavioral abilities, such as phonological working memory, long-term lexical knowledge, and articulatory difficulty with nonsense words (Estes et al. 2007; Gathercole 1995). Ascertainment differences and differences in age ranges between the present study and others could also contribute to the observed discrepancies. The SLI consortium used a family-based linkage analysis examining 186 nuclear families affected with SLI (SLI Consortium 2002, 2004). Brkanac et al. (2008), used a family-based design examining 144 families with a history of dyslexia, whereas the present study examined one extended family with verbal trait disorder that could be derived from one or more rare variants. The small number of subjects in each of these studies would significantly limit the power to detect rare and uncommon variants. Ultimately, these findings may also reflect locus heterogeneity and highlight different molecular and biological mechanisms associated with NWR.

Due to the complex inheritance pattern of impaired NWR performance, a nonparametric analysis using variance components was used so that pre-specified values for parameters defining the genetic model would not be required (Bailey-Wilson 2004). This is in contrast to a parametric analysis that requires the specification of a genetic model, with the concern that a poorly specified model could lead to suboptimal results. An advantage to using a variance components approach is that it tends to be more powerful relative to other trait mapping methods (Kleensang et al. 2010). However, a limitation is that variance components provides poorer localization of the trait locus compared to a parametric analysis, and generally requires additional fine mapping to isolate the region (Amos and de Andrade 2001; Williams et al. 1997). An additional limitation is that we used a composite variable across two different NWR tests—performing more than one standard deviation below the age–sex standardized mean on either the NRT or SRT—to derive VT+ or − status. Although both measures evaluate NWR, the individual test items differ. As described previously, the SRT focuses on the repetition of syllables that comprise only four "early-8" consonants and can be used to examine NWR ability in young children and speakers of any age with limited phonetic inventories or speech sound disorder (Shriberg et al. 2009). In comparison, although the NRT was designed to exclude late developing English consonants (the 'late-8'), younger children and speakers of any age with speech sound disorder can have articulation errors repeating the 9 different vowels and diphthongs, and 11 different consonants in the nonsense words, thus confounding test performance and reducing

transcription reliability (Shriberg et al. 2009). Because individuals tested in the extended pedigree ranged in age from 3 to 95 years old, it would not be optimal to test all individuals on only the NRT or SRT, as they may differ in their sensitivity to persistent types and levels of NWR deficits across the lifespan.

In conclusion, we found a statistically significant genome-wide multipoint linkage signal on chromosome 13q14–q21 using a NWR phenotype in an extended pedigree with a family history of verbal trait disorder. We hypothesize that the region of 13q14–q21 is a susceptibility locus for verbal trait disorders, but additional work must be conducted to (1) identify the gene(s) in this region contributing to the linkage signals observed in the present study and others that have been identified this same region, and (2) elucidate the complex genetic and environmental interactions that may increase susceptibility.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Abecasis GR, Cherny SS, Cookson WO, Cardon LR (2002) Merlin rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet 30:97. doi:10.1038/ng786

Abrahams BS, Tentler D, Perederiy JV, Oldham MC, Coppola G, Geschwind DH (2007) Genome-wide analyses of human perisylvian cerebral cortical patterning. Proc Natl Acad Sci USA 104:17849–17854. doi:10.1073/pnas.0706128104

Alarcón M, Cantor RM, Liu J, Gilliam TC, Geschwind DH (2002) Evidence for a language quantitative trait locus on chromosome 7q in multiplex autism families. Am J Hum Genet 70:60–71. doi:10.1086/338241

Alarcón M, Yonan AL, Gilliam TC, Cantor RM, Geschwind DH (2005) Quantitative genome scan and ordered-subsets analysis of autism endophenotypes support language QTLs. Mol Psychiatry 10:747–757. doi:10.1038/sj.mp.4001666

Alarcón M, Abrahams BS, Stone JL, Duvall JA, Perederiy JV, Bomar JM, Sebat J, Wigler M, Martin CL, Ledbetter DH, Nelson SF, Cantor RM, Geschwind DH (2008) Linkage, association, and gene-expression analyses identify CNTNAP2 as an autism-susceptibility gene. Am J Hum Genet 82:150–159. doi:10.1016/j.ajhg.2007.09.005

Almasy L, Blangero J (1998) Multipoint quantitative-trait linkage analysis in general pedigrees. Am J Hum Genet 62:1198–1211. doi:10.1086/301844

American Speech-Language-Hearing Association (2007) Ad Hoc Committee on Childhood Apraxia of Speech: Childhood Apraxia of Speech [Position Statement]. ASHA, Rockville

Amos CI, de Andrade M (2001) Genetic linkage methods for quantitative traits. Stat Methods Med Res 10:3–25. doi:10.1177/096228020101000102

Archibald LMD, Gathercole SE (2006) Nonword repetition: a comparison of tests. J Speech Lang Hear Res 49:970–983. doi:10.1044/1092-4388(2006/070)

Bailey-Wilson JE (2004) Parametric versus nonparametric and two-point versus multipoint: controversies in gene mapping. Encyclopedia of Genetics, Genomics, Proteomics and Bioinformatics. Wiley, Hoboken

Bartlett CW, Flax JF, Logue MW, Vieland VJ, Bassett AS, Tallal P, Brzustowicz LM (2002) A major susceptibility locus for specific language impairment is located on 13q21. Am J Hum Genet 71:45–55. doi:10.1086/341095

Bartlett CW, Flax JF, Logue MW, Smith BJ, Vieland VJ, Tallal P, Brzustowicz LM (2004) Examination of potential overlap in autism and language loci on chromosomes 2, 7, and 13 in two independent samples ascertained for specific language impairment. Hum Hered 57:10–20. doi:10.1159/000077385

Bishop DVM, North T, Donlan C (1996) Nonword repetition as a behavioural marker for inherited language impairment: evidence from a Twin Study. J Child Psychol Psychiatry 37:391–403. doi:10.1111/j.1469-7610.1996.tb01420.x

Bishop DVM, Adams CV, Norbury CF (2004) Using nonword repetition to distinguish genetic and environmental influences on early literacy development: a study of 6-year-old twins. Am J Med Genet B Neuropsychiatr Genet 129B:94–96. doi:10.1002/ajmg.b.30065

Bradford Y, Haines J, Hutcheson H, Gardiner M, Braun T, Sheffield V, Cassavant T, Huang W, Wang K, Vieland V, Folstein S, Santangelo S, Piven J (2001) Incorporating language phenotypes strengthens evidence of linkage to autism. Am J Med Genet 105:539–547. doi:10.1002/ajmg.1497

BrainSpan: Atlas of the Developing Human Brain [Internet]. Funded by ARRA Awards 1RC2MH089921-01, 1RC2MH090047-01, and 1RC2MH089929-01. © 2011. http://developinghumanbrain.org

Brkanac Z, Chapman NH, Igo RP Jr, Matsushita MM, Nielsen K, Berninger VW, Wijsman EM, Raskind WH (2008) Genome scan of a nonword repetition phenotype in families with dyslexia: evidence for multiple loci. Behav Genet 38:462–475. doi:10.1007/s10519-008-9215-2

Campbell TF, Dollaghan CA, Rockette HE, Paradise JL, Feldman HM, Shriberg LD, Sabo DL, Kurs-Lasky M (2003) Risk factors for speech delay of unknown origin in 3-year-old children. Child Dev 74:346–357. doi:10.1111/1467-8624.7402002

Carrozzo R, Dionisi-Vici C, Steuerwald U, Lucioli S, Deodato F, Di Giandomenico S, Bertini E, Franke B, Kluijtmans LAJ, Meschini MC, Rizzo C, Piemonte F, Rodenburg R, Santer R, Santorelli FM, van Rooij A, Vermunt-de Koning D, Morava E, Wevers RA (2007) SUCLA2 mutations are associated with mild methylmalonic aciduria, Leigh-like encephalomyopathy, dystonia and deafness. Brain 130:862–874

Choi Y, Wijsman EM, Weir BS (2009) Case-control association testing in the presence of unknown relationships. Genet Epidemiol 33:668–678. doi:10.1002/gepi.20418

de Bie P, Muller P, Wijmenga C, Klomp LWJ (2007) Molecular pathogenesis of Wilson and Menkes disease: correlation of mutations with molecular defects and disease phenotypes. J Med Genet 44:673–688. doi:10.1136/jmg.2007.052746

Dening Tr BGE (1989) Wilson's disease: psychiatric symptoms in 195 cases. Arch Gen Psychiatry 46:1126–1134. doi:10.1001/archpsyc.1989.01810120068011

Dollaghan C, Campbell TF (1998) Nonword repetition and child language impairment. J Speech Lang Hear Res 41:1136–1146. doi:10.1044/jslhr.4105.1136

Duggirala R, Williams JT, Williams-Blangero S, Blangero J (1997) A variance component approach to dichotomous trait linkage analysis using a threshold model. Genet Epidemiol 14:987–992. doi:10.1002/(SICI)1098-2272(1997)14:6<987:AID-GEPI71>3.0.CO;2-G

Elpeleg O, Miller C, Hershkovitz E, Bitner-Glindzicz M, Bondi-Rubinstein G, Rahman S, Pagnamenta A, Eshhar S, Saada A (2005) Deficiency of the ADP-forming succinyl-CoA synthase activity is associated with encephalomyopathy and mitochondrial DNA depletion. Am J Hum Genet 76:1081–1086. doi:10.1086/430843

Estes KG, Evans JL, Else-Quest N (2007) Differences in the nonword repetition performance of children with and without specific language impairment: a meta-analysis. J Speech Lang Hear Res 50:177–195. doi:10.1044/1092-4388(2007/015)

Fisher SE, Vargha-Khadem F, Watkins KE, Monaco AP, Pembrey ME (1998) Localisation of a gene implicated in a severe speech and language disorder. Nat Genet 18:168–170. doi:10.1038/ng0298-168

Gathercole SE (1995) Is nonword repetition a test of phonological memory or long-term knowledge–it all depends on the nonwords. Mem Cognit 23:83–94. doi:10.3758/Bf03210559

Glahn DC, Knowles EE, McKay DR, Sprooten E, Raventós H, Blangero J, Gottesman II, Almasy L (2014) Arguments for the sake of endophenotypes: examining common misconceptions about the use of endophenotypes in psychiatric genetics. Am J Med Genet B Neuropsychiatr Genet. doi:10.1002/ajmg.b.32221

Goldman R, Fristoe M (2000) Goldman-Fristoe test of articulation, 2nd edn. American Guidance Service Inc, Circle Pines

Hilschmann N, Barnikol HU, Barnikol-Watanabe S, Götz H, Kratzin H, Thinnes FP (2002) The immunoglobulin-like genetic predetermination of the brain: the protocadherins, blueprint of the neuronal network. Naturwissenschaften 88:2–12. doi:10.1007/s001140000190

Hurst JA, Baraitser M, Auger E, Graham F, Norell S (1990) An extended family with a dominantly inherited speech disorder. Dev Med Child Neurol 32:352–355. doi:10.1111/j.1469-8749.1990.tb16948.x

Jaberi E, Chitsazian F, Ali Shahidi G, Rohani M, Sina F, Safari I, Malakouti Nejad M, Houshmand M, Klotzle B, Elahi E (2013) The novel mutation p.Asp251Asn in the [beta]-subunit of succinate-CoA ligase causes encephalomyopathy and elevated succinylcarnitine. J Hum Genet 58:526–530. doi:10.1038/jhg.2013.45

Kang C, Drayna D (2011) Genetics of speech and language disorders. Annu Rev Genomics Hum Genet 12:145–164. doi:10.1146/annurev-genom-090810-183119

Kaufman AS, Kaufman NL (2004) Kaufman brief intelligence test, Second edn (KBIT-2), Pearson Assessments, San Antonio

Kim S-Y, Yasuda S, Tanaka H, Yamagata K, Kim H (2010) Nonclustered protocadherin. Cell Adh Migr 5:97–105. doi:10.4161/cam.5.2.14374

Kind B, Muster B, Staroske W, Herce HD, Sachse R, Rapp A, Schmidt F, Koss S, Cardoso MC, Lee-Kirsch MA (2014) Altered

spatio-temporal dynamics of RNase H2 complex assembly at replication and repair sites in Aicardi-Goutières syndrome. Hum Mol Genet 23:5950–5960

Kleensang A, Franke D, Alcaïs A, Abel L, Müller-Myhsok B, Ziegler A (2010) An extensive comparison of quantitative trait loci mapping methods. Hum Hered 69:202–211. doi:10.1159/000289596

La Piana R, Uggetti C, Roncarolo F, Vanderver A, Olivieri I, Tonduti D, Helman G, Balottin U, Fazzi E, Crow YJ, Livingston J, Orcesi S (2016) Neuroradiologic patterns and novel imaging findings in Aicardi-Goutières syndrome. Neurology 86:28–35. doi:10.1212/WNL.0000000000002228

Lai CSL, Fisher SE, Hurst JA, Vargha-Khadem F, Monaco AP (2001) A forkhead-domain gene is mutated in a severe speech and language disorder. Nature 413:519–523. doi:10.1038/35097076

Lenzenweger MF (2013) Endophenotype, intermediate phenotype, biomarker: definitions, concept comparisons, clarifications. Depress Anxiety 30(3):185–189

Leonard CM, Eckert MA (2008) Asymmetry and dyslexia. Dev Neuropsychol 33:663–681. doi:10.1080/87565640802418597

Lewis BA, Freebairn LA, Taylor HG (2000) Academic outcomes in children with histories of speech sound disorders. J Commun Disord 33:11–30. doi:10.1016/S0021-9924(99)00023-4

Lohmeier HL, Shriberg LD (2011) Reference data for the syllable repetition task (SRT). Technical Report No. 17, Phonology Project, Waisman Center, University of Wisconsin-Madison

Maas RR, Marina AD, Brouwer APM, Wevers RA, Rodenburg RJ, Wortmann SB (2015) SUCLA2 Deficiency: a deafness-dystonia syndrome with distinctive metabolic findings (Report of a New Patient and Review of the Literature). Springer Berlin Heidelberg, p 1–6. doi: 10.1007/8904_2015_464

MacDermot KD, Bonora E, Sykes N, Coupe A-M, Lai CSL, Vernes SC, Vargha-Khadem F, McKenzie F, Smith RL, Monaco AP, Fisher SE (2005) Identification of FOXP2 truncation as a novel cause of developmental speech and language deficits. Am J Hum Genet 76:1074–1080. doi:10.1086/430841

Marshall CR, Noor A, Vincent JB, Lionel AC, Feuk L, Skaug J, Shago M, Moessner R, Pinto D, Ren Y, Thiruvahindrapduram B, Fiebig A, Schreiber S, Friedman J, Ketelaars CE, Vos YJ, Ficicioglu C, Kirkpatrick S, Nicolson R, Sloman L, Summers A, Gibbons CA, Teebi A, Chitayat D, Weksberg R, Thompson A, Vardy C, Crosbie V, Luscombe S, Baatjes R, Zwaigenbaum L, Roberts W, Fernandez B, Szatmari P, Scherer SW (2008) Structural variation of chromosomes in autism spectrum disorder. Am J Hum Genet 82:477–488. doi:10.1016/j.ajhg.2007.12.009

Matise TC, Chen F, Chen W, De La Vega FM, Hansen M, He C, Hyland FCL, Kennedy GC, Kong X, Murray SS, Ziegle JS, Stewart WCL, Buyske S (2007) A second-generation combined linkage–physical map of the human genome. Genome Res 17:1783–1786. doi:10.1101/gr.7156307

Mitter D, Ullmann R, Muradyan A, Klein-Hitpas L, Kanber D, Ounap K, Kaulisch M, Lohmann D (2011) Genotype-phenotype correlations in patients with retinoblastoma and interstitial 13q deletions. Eur J Hum Genet 19:947–958. doi:10.1038/ejhg.2011.58

Moore MW, Tompkins CA, Dollaghan CA (2010) Manipulating articulatory demands in non-word repetition: a 'late-8' non-word repetition task. Clin Linguist Phon 24:997–1008. doi:10.3109/02699206.2010.510917

Morris CP, Baune BT, Domschke K, Arolt V, Swagell CD, Hughes IP, Lawford BR, McD Young R, Voisey J (2012) KPNA3 variation is associated with schizophrenia, major depression, opiate dependence and alcohol dependence. Dis Markers. doi:10.3233/dma-2012-0921

Nato AQ, Chapman NH, Sohi HK, Nguyen HD, Brkanac Z, Wijsman EM (2015) PBAP: a pipeline for file processing and quality control of pedigree data with dense genetic markers. Bioinformatics 31:3790–3798. doi:10.1093/bioinformatics/btv444

Newbury DF, Winchester L, Addis L, Paracchini S, Buckingham L-L, Clark A, Cohen W, Cowie H, Dworzynski K, Everitt A, Goodyer IM, Hennessy E, Kindley AD, Miller LL, Nasir J, O'Hare A, Shaw D, Simkin Z, Simonoff E, Slonims V, Watson J, Ragoussis J, Fisher SE, Seckl JR, Helms PJ, Bolton PF, Pickles A, Conti-Ramsden G, Baird G, Bishop DVM, Monaco AP (2009) CMIP and ATP2C2 modulate phonological short-term memory in language impairment. Am J Hum Genet 85:264–272. doi:10.1016/j.ajhg.2009.07.004

Ocklenburg S, Arning L, Gerding WM, Hengstler JG, Epplen JT, Güntürkün O, Beste C, Akkad DA (2015) Left-right axis differentiation and functional lateralization: a haplotype in the methyl-transferase encoding gene SETDB2 might mediate handedness in healthy adults. Mol Neurobiol. doi:10.1007/s12035-015-9534-2

Ostergaard E, Hansen FJ, Sorensen N, Duno M, Vissing J, Larsen PL, Faeroe O, Thorgrimsson S, Wibrand F, Christensen E, Schwartz M (2007) Mitochondrial encephalomyopathy with elevated methylmalonic acid is caused by SUCLA2 mutations. Brain 130:853–861

Peterson RL, Pennington BF (2015) Developmental dyslexia. Annu Rev Clin Psychol 11:283–307. doi:10.1146/annurev-clinpsy-032814-112842

Potter NL, Hall S, Karlsson HB, Fourakis M, Lohmeier HL, McSweeny JL, Tilkens CM, Wilson DL, Shriberg LD (2012) Reference data for the Madison Speech Assessment Protocol (MSAP): a database of 150 participants 3-to-18 years of age with typical speech. Technical Report No. 18, Phonology Project, Waisman Center, University of Wisconsin-Madison

Rice G, Patrick T, Parmar R, Taylor CF, Aeby A, Aicardi J, Artuch R, Montalto SA, Bacino CA, Barroso B, Baxter P, Benko WS, Bergmann C, Bertini E, Biancheri R, Blair EM, Blau N, Bonthron DT, Briggs T, Brueton LA, Brunner HG, Burke CJ, Carr IM, Carvalho DR, Chandler KE, Christen H-J, Corry PC, Cowan FM, Cox H, D'Arrigo S, Dean J, De Laet C, De Praeter C, Déry C, Ferrie CD, Flintoff K, Frints SGM, Garcia-Cazorla A, Gener B, Goizet C, Goutières F, Green AJ, Guët A, Hamel BCJ, Hayward BE, Heiberg A, Hennekam RC, Husson M, Jackson AP, Jayatunga R, Jiang Y-H, Kant SG, Kao A, King MD, Kingston HM, Klepper J, van der Knaap MS, Kornberg AJ, Kotzot D, Kratzer W, Lacombe D, Lagae L, Landrieu PG, Lanzi G, Leitch A, Lim MJ, Livingston JH, Lourenco CM, Lyall EGH, Lynch SA, Lyons MJ, Marom D, McClure JP, McWilliam R, Melancon SB, Mewasingh LD, Moutard M-L, Nischal KK, Østergaard JR, Prendiville J, Rasmussen M, Rogers RC, Roland D, Rosser EM, Rostasy K, Roubertie A, Sanchis A, Schiffmann R, Scholl-Bürgi S, Seal S, Shalev SA, Corcoles CS, Sinha GP, Soler D, Spiegel R, Stephenson JBP, Tacke U, Tan TY, Till M, Tolmie JL et al (2007) Clinical and molecular phenotype of Aicardi-Goutières syndrome. Am J Hum Genet 81:713–725. doi:10.1086/521373

Rice GM, Raca G, Jakielski KJ, Laffin JJ, Iyama-Kurtycz CM, Hartley SL, Sprague RE, Heintzelman AT, Shriberg LD (2012) Phenotype of FOXP2 haploinsufficiency in a mother and son. Am J Med Genet A 158A:174–181. doi:10.1002/ajmg.a.34354

Royal College of Speech and Language Therapists (2011) Developmental verbal dyspraxia policy statement. The Royal College of Speech and Language Therapists, London

Scerri TS, Brandler WM, Paracchini S, Morris AP, Ring SM, Richardson AJ, Talcott JB, Stein J, Monaco AP (2011) PCSK6 is associated with handedness in individuals with dyslexia. Hum Mol Genet 20:608–614. doi:10.1093/hmg/ddq475

Scheer-Cohen AR, Holt AS, Karlsson HB, Mabie HL, McSweeny JL, Tilkens CM, Shriberg LD (2013) Reference data for the Madison Speech Assessment Protocol (MSAP): a database of fifty 20-to-80 year old participants with typical speech, Technical Report No. 20, Phonology Project, Waisman Center, University of Wisconsin-Madison

Semel E, Wiig EH, Secord WA (2004) Clinical evaluation of language fundamentals, Fourth Edition—Screening test (CELF-4 screening test). The Psychological Corporation, Toronto

Shriberg LD (1993) Four new speech and prosody-voice measures for genetics research and other studies in developmental phonological disorders. J Speech Hear Res 36:105. doi:10.1044/jshr.3601.105

Shriberg LD, Lohmeier HL (2008) The Syllable Repetition Task (SRT). Technical Report No. 14, Phonology Project, Waisman Center, University of Wisconsin-Madison

Shriberg LD, Tomblin JB, McSweeny JL (1999) Prevalence of speech delay in 6-year-old children and comorbidity with language impairment. J Speech Lang Hear Res 42:1461–1481. doi:10.1044/jslhr.4206.1461

Shriberg LD, Ballard KJ, Duffy JR, Odell KH, Williams CA (2006) Speech, prosody, and voice characteristics of a mother and daughter with a 7;13 translocation affecting FOXP2. J Speech Lang Hear Res 49:500–525. doi:10.1044/1092-4388(2006/038)

Shriberg LD, Lohmeier HL, Campbell TF, Dollaghan CA, Green JR, Moore CA (2009) A nonword repetition task for speakers with misarticulations: the syllable repetition task (SRT). J Speech Lang Hear Res 52:1189–1212. doi:10.1044/1092-4388(2009/08-0047)

Shriberg LD, Lohmeier HL, Strand EA, Jakielski KJ (2012) Encoding, memory, and transcoding deficits in childhood apraxia of speech. Clin Linguist Phon 26:445–482. doi:10.3109/02699206.2012.655841

SLI Consortium (2002) A genome-wide scan identifies two novel loci involved in specific language impairment*. Am J Hum Genet 70:384–398. doi:10.1086/338649

SLI Consortium (2004) Highly significant linkage to the SLI1 locus in an expanded sample of individuals affected by specific language impairment. Am J Hum Genet 74:1225–1238. doi:10.1086/421529

Smith M, Woodroffe A, Smith R, Holguin S, Martinez J, Filipek PA, Modahl C, Moore B, Bocian ME, Mays L, Laulhere T, Flodman P, Spence MA (2002) Molecular genetic delineation of a deletion of chromosome 13q12 → q13 in a patient with autism and auditory processing deficits. Cytogenet Genome Res 98:233–239. doi:10.1159/000071040

Spiechowicz M, Bernstein H-G, Dobrowolny H, Leśniak W, Mawrin C, Bogerts B, Kuźnicki J, Filipek A (2006) Density of Sgt1-immunopositive neurons is decreased in the cerebral cortex of Alzheimer's disease brain. Neurochem Int 49:487–493. doi:10.1016/j.neuint.2006.03.009

Steele MM, Al-Adeimi M, Siu VM, Fan Y-S (2001) Brief report: a case of autism with interstitial deletion of chromosome 13. J Autism Dev Disord 31:231–234. doi:10.1023/A:1010759401344

The 1000 Genomes Project Consortium (2010) A map of human genome variation from population-scale sequencing. Nature 467:1061–1073. doi:10.1038/nature09534

Thompson E (2011) The structure of genetic linkage data: from LIPED to 1 M SNPs. Hum Hered 71:86–96. doi:10.1159/000313555

Tomblin JB, Records NL, Buckwalter P, Zhang X, Smith E, O'Brien M (1997) Prevalence of specific language impairment in kindergarten children. J Speech Lang Hear Res 40:1245. doi:10.1044/jslhr.4006.1245

Vernes SC, Newbury DF, Abrahams BS, Winchester L, Nicod J, Groszer M, Alarcón M, Oliver PL, Davies KE, Geschwind DH, Monaco AP, Fisher SE (2008) A functional genetic link between distinct developmental language disorders. N Engl J Med 359:2337–2345. doi:10.1056/NEJMoa0802828

Vidal R, Frangione B, Rostango A, Mead S, Revesz T, Plant G, Ghiso J (1999) A stop-codon mutation in the BRI gene associated with familial British dementia. Nature 399:776. doi:10.1038/21637

Vidal R, Révész T, Rostagno A, Kim E, Holton JL, Bek T, Bojsen-Møller M, Braendgaard H, Plant G, Ghiso J, Frangione B (2000) A decamer duplication in the 3' region of the BRI gene originates an amyloid peptide that is associated with dementia in a Danish kindred. Proc Natl Acad Sci USA 97:4920–4925. doi:10.1073/pnas.080076097

Wei J, Hemmings GP (2005) The KPNA3 gene may be a susceptibility candidate for schizophrenia. Neurosci Res 52:342–346. doi:10.1016/j.neures.2005.04.005

Wiig EH, Secord WA, Semel E (2004) Clinical evaluation of language fundamentals—Preschool, Second edn (CELF Preschool-2). The Psychological Corporation, Toronto

Wijsman EM, Peterson D, Leutenegger A-L, Thomson JB, Goddard KAB, Hsu L, Berninger VW, Raskind WH (2000) Segregation analysis of phenotypic components of learning disabilities. I. nonword memory and digit span. Am J Hum Genet 67:631–646. doi:10.1086/303044

Williams JT, Duggirala R, Blangero J (1997) Statistical properties of a variance components method for quantitative trait linkage analysis in nuclear families and extended pedigrees. Genet Epidemiol 14:1065–1070. doi:10.1002/(SICI)1098-2272(1997)14:6<1065:AID-GEPI84>3.0.CO;2-F

Woodcock RW, McGrew K, Mather N (2001) Woodcock-Johnson tests of achievement, 3rd edition (WJ-III). Riverside Publishing, Rolling Meadows

Wren YE, Roulstone SE, Miller LL (2012) Distinguishing groups of children with persistent speech disorder: findings from a prospective population study. Logoped Phoniatr Vocol 37:1–10. doi:10.3109/14015439.2011.625973

Xu P-F, Zhu K-Y, Jin Y, Chen Y, Sun X-J, Deng M, Chen S-J, Chen Z, Liu TX (2010) Setdb2 restricts dorsal organizer territory and regulates left–right asymmetry through suppressing fgf8 activity. Proc Natl Acad Sci USA 107:2521–2526. doi:10.1073/pnas.0914396107

Yamagata K, Andreasson KI, Sugiura H, Maru E, Dominique M, Irie Y, Miki N, Hayashi Y, Yoshioka M, Kaneko K, Kato H, Worley PF (1999) Arcadlin is a neural activity-regulated cadherin involved in long term potentiation. J Biol Chem 274:19473–19479. doi:10.1074/jbc.274.27.19473

Yasuda S, Tanaka H, Sugiura H, Okamura K, Sakaguchi T, Tran U, Takemiya T, Mizoguchi A, Yagita Y, Sakurai T, De Robertis EM, Yamagata K (2007) Activity-induced protocadherin arcadlin regulates dendritic spine number by triggering N-cadherin endocytosis via TAO2β and p38 MAP kinases. Neuron 56:456–471. doi:10.1016/j.neuron.2007.08.020

Zhang H, Ju G, Wei J, Hu Y, Liu L, Xu Q, Chen Y, Sun Z, Liu S, Yu Y, Guo Y, Shen Y (2006) A combined effect of the KPNA3 and KPNB3 genes on susceptibility to schizophrenia. Neurosci Lett 402:173–175. doi:10.1016/j.neulet.2006.03.073