

Further Evidence of Pleiotropy Influencing Speech and Language: Analysis of the DYX8 Region

L. Miscimarra^a C. Stein^a C. Millard^a A. Kluge^a K. Cartier^a L. Freebairn^b
A. Hansen^b L. Shriberg^c H.G. Taylor^b B. Lewis^b S.K. Iyengar^a

^aDepartment of Epidemiology and Biostatistics, ^bDepartment of Pediatrics, Rainbow Babies and Childrens Hospital, Case Western Reserve University, Cleveland, Ohio, ^cWaisman Center, University of Wisconsin-Madison, Madison, Wisc., USA

Key Words

Phonology · Pleiotropic effect · Familial aggregation · Quadratic exponential model · Multivariate analysis

Abstract

Background/Aims: Genetic studies have raised the possibility of common bases for cognitive linguistic disorders such as speech sound disorder (SSD), reading disorder (RD) and language impairment (LI). Thus, some of the same genes may jointly influence cognitive components within and between these three disorders. We examined the plausibility of this theory in a sample of families ascertained on the basis of a child with SSD. **Methods:** Using the method of generalized estimating equations to solve a bivariate family predictive model we obtained measures of comorbidity and familial aggregation of SSD and LI. We then used two methods of multipoint model-free linkage analysis to evaluate SSD and LI psychometric test measures over a region previously implicated in linkage studies of RD, DYX8 region, 1p34–p36. **Results:** Bivariate phenotypic analyses show evidence of comorbidity and within family aggregation and coaggregation of SSD and LI. In addition, two regions on chromosome 1 show suggestive evidence of linkage. The first region was previously reported in dyslexia studies. Our maximum linkage signal in this region measured articulation ($p = 0.0009$)

in SSD sibling pairs. The second region is characterized by processes involved in language production, with the maximum linkage signal measuring listening comprehension ($p = 0.0019$) using all sibling pairs. **Conclusion:** We conclude that the DYX8 region could bear genes controlling pleiotropic effects on SSD, LI and RD. Copyright © 2007 S. Karger AG, Basel

Introduction

Speech Sound Disorders (SSD) are characterized by errors due to poor motor abilities associated with the production of speech sounds (articulation or phonetic structure), and/or errors in applying linguistic rules used to combine sounds to form words (phonological errors). SSD are prevalent in childhood, with rates estimated at 15.2% for pre-school age children [1] persisting in 3.8% of these affected individuals up to 6 years of age. SSD can be comorbid with reading disorder (RD), or dyslexia [1–3] and disorders of language [4].

The comorbidity with RD is hypothesized to be, at least in part, attributable to skills that underlie both written and spoken language, broadly phonological processing and more specifically phonological awareness. Dyslexic children have been shown to have core deficits in

phonological processing skills, defined as the ability to correctly use phonemes to process verbal and written information in short and long term memory by coding and decoding information. A critical component of phonological processing, phonological awareness, which is demonstrated by tasks that assess explicit knowledge of the sound system, has been shown to be particularly problematic for children with RD [5–7].

Language impairment (LI) refers to a heterogeneous group of developmental disorders or delays that are principally characterized by deficits and/or immaturities in language comprehension or production that may involve semantics, syntax or pragmatics. The reported rates of co-morbidity between SSD and LI in children with persistent speech delay are 11–15% at 6 years of age [1], with considerably higher rates of 40–60% reported for preschool children [8]. In addition, both children with comorbid SSD and LI and those with LI only often present with reading disorders at school age [9, 10].

Genetic studies of SSD and RD raise the possibility of a common genetic basis for these disorders [11–18]. Our prior findings on chromosome 3 [18] and those reported recently for chromosomes 1, 6 and 15 [19] have demonstrated that this overlapping etiology, as evidenced in part by coincident linkage signals, may be due to pleiotropy. Pleiotropy is defined as the same gene(s) influencing different disorders or influencing seemingly unrelated components of the same disorder. If coincident linkage signals were actually pleiotropy then there would also be a greater than expected level of genetic comorbidity of SSD, LI and/or RD within individuals and within families [20]. There have been reports of both comorbidity and familial aggregation of SSD and LI [1] and Specific Language Impairment and RD [21]; however, multivariate aggregation and segregation analyses have only been performed using RD samples [22–24]. To date there have been no formal multivariate analyses of the SSD and LI phenotypes. In this paper we propose such a model as a preliminary step to linkage analyses.

We therefore examine pleiotropy in two stages. First we tested for phenotypic aggregation using a bivariate analysis. This analysis elucidated the extent of familial aggregation, as well as individual and familial comorbidity of SSD and LI. These estimates are generated using information across all families in the sample. Second, using model-free linkage methods we tested for linkage of quantitative psychometric measures used to test for SSD and the binary measures of SSD and LI to the DYX8 region, corresponding to 1p34-p36, which has previously been implicated in linkage studies of dyslexia. The results

of the bivariate analyses demonstrate evidence for comorbidity and aggregation. In addition our linkage analyses of SSD phenotypes show tight overlap with both a previous SSD study using an identical test measure and tight overlap with several existing dyslexia linkage analyses in the DYX8 region at two susceptibility loci.

Methods

Family Ascertainment

The families were recruited through a proband child with a moderate to severe SSD referred from the clinical caseloads of speech language pathologists in Northeastern Ohio. All study participants were screened to ensure they met the following criteria: (1) normal speech mechanism as demonstrated by the Total Structure subscale of the Oral and Speech Motor Control Protocol [25], (2) normal hearing as demonstrated by pure tone audiometric screening test and impedance audiometry, (3) normal intelligence, defined as a performance IQ of at least 80 on the Wechsler Preschool and Primary Scale of Intelligence (Revised) [26, 27], or Wechsler Intelligence Scale for Children (Third Edition) [28], (4) absence of a history of neurological disorders or developmental delays other than speech and language as reported by the parent. The purpose of these procedures was to exclude SSD due to structural abnormalities such as cleft lip and palate, hearing loss, and general cognitive impairments, which may be under genetic influences separate from those affecting the majority of cases.

Probands and siblings were considered SSD affected if their scores on the Goldman-Fristoe Test of Articulation Sounds in Words subtest (GFTA) [29] were in the bottom tenth percentile, and they committed a minimum of four phonological process errors with severity rating 3–4 on the Khan-Lewis Phonological Analysis Test [30]. Parents were considered SSD or LI affected if they had ever been enrolled in speech or language therapy. The qualitative variable SSD therefore includes all individuals meeting these criteria, regardless of their LI affection status.

Binary affection status for language impairment (LI) was defined by a score one standard deviation or greater below the mean on two or more subtests of the age appropriate Test of Language Development, Primary (Second Edition) [31], or the Clinical Evaluation of Language Fundamentals (Revised) [32]. The qualitative variable LI therefore includes comorbid individuals. Additionally we generated the qualitative variable SL which is a composite of the two variables LI and SSD; individuals with SL have LI = 1 and/or SSD = 1.

Psychometric Measures

Articulation of speech sounds in initial, medial and final word positions, as well as consonant blends, was assessed using the GFTA. For participants over four years of age, phonological short term memory was evaluated through the repetition of nonsense words on the Nonsense Word Repetition Task (NSW) [33]. The multisyllabic word Repetition Task [33] (MSW) and difficult to articulate phrases such as ‘mixed biscuits’ or ‘blue plaid pants’ on the Speech Error Phrases (ERRORW) further assessed articulation skills in children who may not present with overt speech errors on the GFTA or in conversational speech. The Colorado Per-

ceptual Speed Test (CPST) [34] was used to assess visual speed of processing of letters and numbers using a three-part, timed task requiring the individual to rapidly scan and circle groups of matching letter-number clusters.

Verbal short term memory was assessed with the sentence imitation task of the CELF-R (CELF-Rsi) or TOLD-P:2 (TOLDP:2si) [31, 32]. Both receptive and expressive vocabulary was assessed using the Peabody Picture Vocabulary Test (Third Edition) [35], and the Expressive One Word Vocabulary Test (Revised) [36], respectively.

Reading decoding was assessed in children at least 7 years of age using the Word Attack (WRAT) and Word Identification (WRID) subtests of the Woodcock Reading Mastery Tests (Revised) [37]. Reading comprehension was evaluated using the Wechsler Individual Achievement Test Reading Comprehension subtest (WIATRC) which requires children to select the picture that matches a word or sentence as well as generate a word that matches a picture and oral description [38]. The Wechsler Individual Achievement Test Listening Comprehension subtest (WIATLC) assesses auditory language comprehension, tapping syntax, vocabulary and verbal short-term memory [39]. Spelling was evaluated using the third edition of the Test of Written Spelling (TWS-3) [40]. The TWS-3 contains both predictable words (words with regular phoneme-grapheme associations) and unpredictable words (words that do not adhere to spelling rules).

All quantitative scores were adjusted prior to analysis of the genetic data using a stepwise regression model including age, age squared and a rating of socioeconomic status using the Hollingshead Four Factor Index of Social Class [41]. Two sided t tests, accounting for significantly different variances when appropriate, were used to determine if the adjusted mean scores differed significantly between proband and SSD affected siblings and SSD unaffected siblings. The residual mean scores were used in all quantitative linkage analyses in this study.

Genotyping

High molecular weight DNA was isolated from an aliquot of blood with the Puregene Kit (Gentra Systems, Inc) or from buccal swabs using the BioRad InstaGene™ Matrix protocol (BioRad Laboratories, Inc.). Microsatellite markers on chromosome 1p34–36 were genotyped according to previously described methods [18]. The 10 markers span an approximately 94 cM region; they are indicated in bold in figure 1. Marker order, map distance, physical location and marker heterozygosity was determined using MAP-O-MATv1.1 (<http://compngen.rutgers.edu/mapomat>). Five percent blind replicates and two Centre d'Etude du Polymorphisme Humain (CEPH) controls were included on each gel to serve as internal controls. The genotype error rate, found by comparing replicates, was 4.87%.

Inconsistencies in the segregation of the genotypes within families were examined using MARKERINFO (S.A.G.E. v5.0). Individuals who demonstrated Mendelian inconsistencies at multiple markers that could not be resolved by retyping were treated as missing for the purpose of this analysis. The total of Mendelian inconsistencies and irresolvable genotypes represented 5.3% of the data. This error rate, along with the genotype error cited above, are likely due to poor DNA quality, since 82.74% of the missing data were found in buccal samples. After this data cleaning step, the marker informativity, in terms of Shannon information content, ranged from 0.74 (D1S255) to 0.86 (D1S2781). Rela-

tionships were confirmed using a likelihood-based analysis of all available molecular data, implemented in RELPAIR [42].

Bivariate Family Predictive Model

In order to examine aggregation and comorbidity we first calculated some simple odds ratios and then constructed a quadratic exponential model (QEM) [43]. All individuals with phenotypic data were used in both odds ratio calculations and the QEM; to adjust for ascertainment we did not include the affection status of the proband. Our rationale for the use of the QEM model was: (1) the parameters have clear interpretations as conditional odds and odds ratios, (2) associations within families and within individuals can be examined, (3) individual characteristics (sex, race, type of relative, etc.) can be taken into account through the use of covariates. Hudson et al. [44, 45] derived the logistic regression equations for the multivariate case detailed below.

The components of the QEM model are as follows. First we generated a vector of family member j 's binary affection status for SSD, Y_{S_j} , and LI, Y_{L_j} , written as $(Y_{S_j}, Y_{L_j})^T$, with $j = 1$ for the proband and $j > 1$ for all relatives. For each person, we then summed the number of SSD affected in the family, $S_{s,-j} = \sum Y_{S_k}$ and the number of LI affected in the family, $S_{l,-j} = \sum Y_{L_k}$, subtracting one from these totals if the individual family member was affected. The five main parameters in model 1 measure baseline risks of SSD (α_S) and LI (α_L) assuming that there are no other affected members and that the individual is not comorbid for SSD and LI, the association of SSD and LI within an individual (δ), the within family association of SSD (γ_S), the within family association of LI (γ_L) and the within family association of SSD (or LI) given LI (or SSD), (γ_{SL}). The values in the θ vectors were used to estimate covariate intercepts for a response's relative type (offspring = 0, other relative type = 1) and race (Caucasian = 0, non-Caucasian = 1). The model for $j = 1 \dots n$ is written

$$P(Y_{S_j}, Y_{L_j} | \mathbf{Y}_{S,-j}, \mathbf{Y}_{L,-j}, \mathbf{Z}_j) \propto \exp \{ \alpha_S Y_{S_j} + \alpha_L Y_{L_j} + \delta Y_{S_j} Y_{L_j} + \gamma_S Y_{S_j} S_{s,-j} + \gamma_L Y_{L_j} S_{l,-j} + \gamma_{S^*L} (Y_{S_j} S_{l,-j} + Y_{L_j} S_{s,-j}) + (\theta_S^T \mathbf{Z}_j) Y_{S_j} + (\theta_L^T \mathbf{Z}_j) Y_{L_j} \}. \quad (1)$$

This can be rewritten as two logistic regression equations with shared coefficients. One equation has SSD as an outcome, Y_{S_j} and one equation has LI, Y_{L_j} , as an outcome. For each person we summed (i) the number of SSD affected, $S_{s,-j} = \sum Y_{S_k}$ and the number of LI affected, $S_{l,-j} = \sum Y_{L_k}$, subtracting one in both cases if the individual family member was affected. In order to differentiate offspring and parents we used an indicator variable which holds a value of 1 when the pair of equations is individual that is a parent.

$$\logit P(Y_{S_j} = 1 | Y_{L_j}, \mathbf{Y}_{-j}, \mathbf{Z}_j) = \alpha_S + \delta Y_{L_j} + \gamma_S S_{s,-j} + \gamma_{S^*L} S_{l,-j} + \gamma_{S^*SL} S_{SL,-j} + \gamma_{S^*SL} Y_{L_j} S_{s,-j} + \gamma_{L^*SL} Y_{L_j} S_{l,-j} + \gamma_{SL^*SL} Y_{L_j} S_{SL,-j} + \theta_S^T \mathbf{Z}_j$$

$$\logit P(Y_{L_j} = 1 | Y_{S_j}, \mathbf{Y}_{-j}, \mathbf{Z}_j) = \alpha_L + \delta Y_{S_j} + \gamma_L S_{l,-j} + \gamma_{S^*L} S_{s,-j} + \gamma_{L^*SL} S_{SL,-j} + \gamma_{L^*SL} Y_{S_j} S_{l,-j} + \gamma_{S^*SL} Y_{S_j} S_{s,-j} + \gamma_{SL^*SL} Y_{S_j} S_{SL,-j} + \theta_L^T \mathbf{Z}_j \quad (2)$$

To adjust for ascertainment of probands on SSD(LI) status, only the proband's LI(SSD) equation was used in the estimation of the coefficients.

Multipoint Linkage Analysis

Genotypes from all family members were used to calculate both single point and multipoint identity by descent (IBD) allele

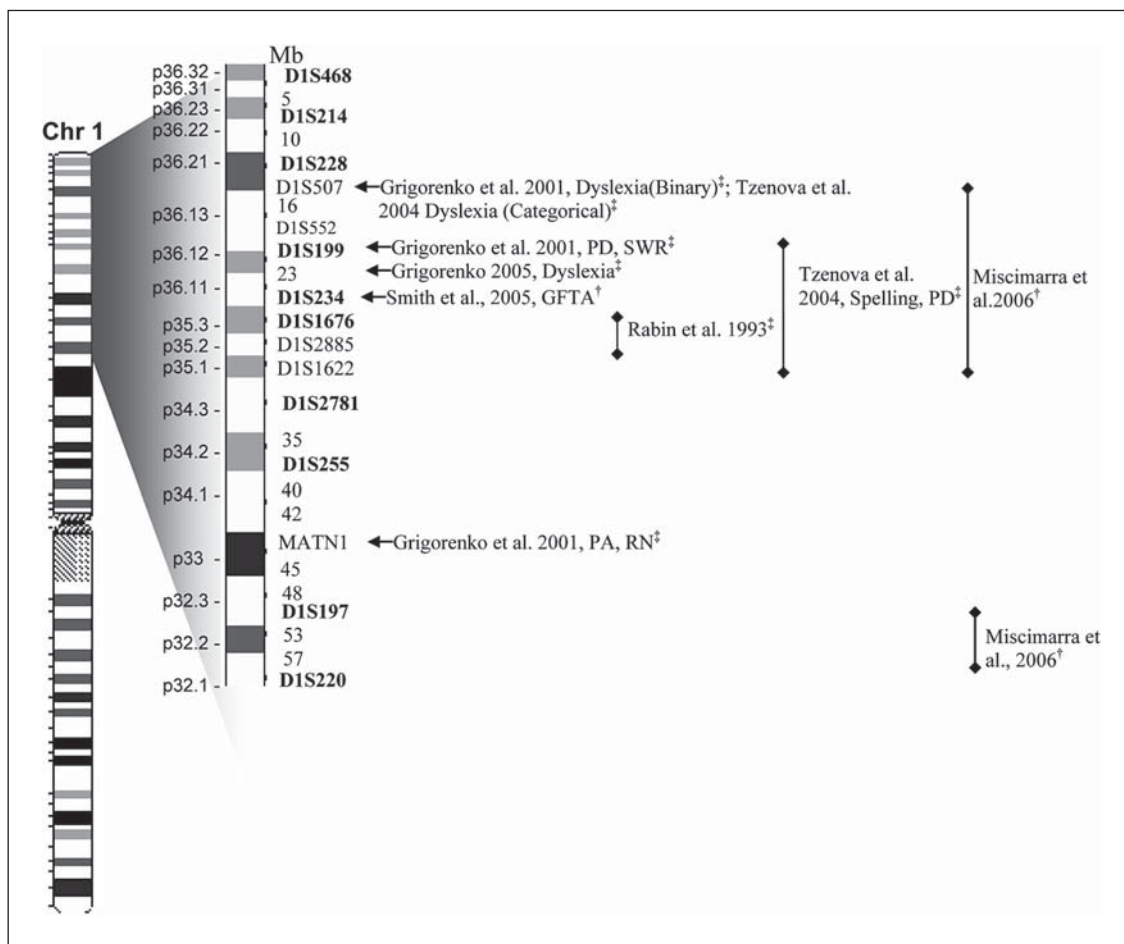


Fig. 1. Genetic and cytogenetic locations with evidence for linkage to RD, SSD and related phenotypes. ← = Closest marker to most significant linkage signal. ‡ Linkage analyses using population ascertained on the basis of an individual with RD. † Linkage analysis using population ascertained on the basis of an individual with SSD. GFTA = Goldman Fristoe Test of Articulation; PA = phonemic awareness; PD = phonological decoding; SWR = single word reading; RN = rapid naming.

sharing distributions using the GENIBD program of S.A.G.E (v5.0). To allow for covariate-related locus heterogeneity, we applied a covariate-based affected relative pair LOD score method as implemented in the LODPAL program of S.A.G.E (v5.0). This model is a 1 parameter modification of the conditional logistic parameterization of the Affected Sib Pair (ASP) LOD score method [47]. An optimal mode of inheritance parameter is specified that allows one to fit only a single additional parameter per covariate. The model is parameterized in terms of offspring recurrence risk ratio, λ_1 , conditional on K covariates:

$$\lambda_1(x) = \exp(\beta + \sum_{k=1}^K \gamma_k x_k) \quad (3)$$

The covariates are centered around the sib pair sample mean before inclusion in the model in order to improve numeric stability, simplify specification of constraints on parameter estimates and so that β reflects 'average' allele sharing in the sample. The γ_k are covariate-specific parameters that measure the change in

allele sharing as a function of the covariates. While the actual β and γ_k values are coding dependent, conclusions about locus heterogeneity are not dependent on the estimated value of β .

We first examined evidence for linkage of regions on chromosome 1 to SSD, LI and SL using only concordantly affected sibling pairs. We then incorporated the phenotypic indicators for each sib pair into the conditional logistic regression as covariates. To do this we generated LOD scores at all marker positions for ASP analyses of SSD, LI and SL. Using three models, we added the 13 test measures listed one at a time. In total, 3 ASP and 39 single covariate analyses were run. The likelihood ratio statistics (LRSs) for all models were calculated by dividing the LOD scores at each location by $2\log_e 10$. The probability values for the ASP LRSs using only the binary traits were obtained using the chi-square distribution with one degree of freedom. To calculate the significance of the additional covariate to each of the three ASP models the LOD score at each marker location was subtracted from the covariate

Table 1. General description of SSD population

Characteristics of SSD sample	Caucasian families	Non-Caucasian families
<i>Description of families</i>		
Pedigrees	126	25
<i>Sibships</i>		
Mean sibship size (range)	2.51 (1–7)	1.91 (1–4)
Full (half) sibling pairs	316 (10)	42 (8)
Concordantly affected full (half) sibling pairs	99 (1)	14 (2)
SES: Hollingshead Four-Factor Index (% of total number of families classified by ethnicity)		
Category 1 (low)	2.6	21.4
Category 2	6.9	17.9
Category 3	16.4	25.0
Category 4	46.5	21.4
Category 5 (high)	27.6	35.5
<i>Description of individuals</i>		
Individuals (male/female)		
Siblings	191 (1.22)	31 (1.07)
Half-siblings	5 (4.0)	10 (1.33)
Parents	215 (0.84)	36 (0.50)
Mean age (range), years		
Proband	6.02 (3–12)	6.26 (3–10)
Affected siblings	7.05 (3–12)	8.86 (4–11)
Unaffected siblings	7.70 (2–17)	8.22 (3–11)
<i>Description of relative pairs</i>		
Affected relative pairs, siblings, parents		
SSD	71, 53	30, 5
With LI	28, 8	11, 2
With no LI	43, 45	0, 3
LI	36, 13	19, 4
With no SSD	8, 5	8, 2
Neither disorder	114, 152	14, 29

based LOD score at the corresponding marker location. The resulting LOD score difference was then converted to a LRS and evaluated for significance using a chi-square distribution with one degree of freedom. Lastly, to calculate the overall significance of the one covariate models a 50:50 mixed chi-square distribution of 0 and 1 degree of freedom was used [46].

The genetic contribution to pleiotropy was also assessed using Haseman-Elston regression analysis [47], which regresses a value for the sibling pair trait on the proportion of alleles shared IBD using iterative Generalized Estimating Equations (GEE) to account for the non-independent sibpairs. The weighted combination of the mean corrected sibpair squared trait sum and squared trait difference (option w4), is asymptotically most powerful and was therefore used for all regressions [48]. As we are confirming linkage and not establishing new evidence for linkage, the threshold for significance was set to $\alpha = 0.05$. A non-parametric permutation-based procedure for controlling the false discovery rate was used to correct multiple testing [49]. All p values reported for ASP linkage analyses and Haseman-Elston regression analyses are p values generated from this procedure. Linkage signals within 20 cM of one another were considered attributable to the same

underlying genetic susceptibility. If the maximum LOD score for different phenotypes was greater than 20 cM apart these peaks were considered to be generated by different genetic mechanisms [50].

Results

Description of the Sample

Table 1 describes the characteristics of the 151 families (358 full sibling pairs) with genetic data available. Caucasian and non-Caucasian groups were considered separately in both the bivariate and genetic analyses due to differences in the sample sizes and the distributions of family socioeconomic status, SSD and LI. Because of the small number (<10) of non-Caucasian affected sibling pairs on which genetic and phenotypic information are available the maximum likelihood estimates did not con-

Table 2. Mean proband and sibling test scores (Caucasian only)

Test (range)	Sibling mean trait values			Probands mean trait value		Test of means ^a		
	SSD = 0 LI = 0 (n = 114)	LI = 1 (n = 36)	SSD = 1 (n = 71)	SSD = 1 (n = 126)	SSD = 1 LI = 0 (n = 50)	SSD = 1 vs. SSD = 0	proband vs. SSD = 1 sibling	proband vs. SSD = 0 sibling
GFTA (0–100) ⁺	81.88	60.17	61.78	24.22	25.86	*	*	*
NSW (0–100) [‡]	84.91	36.55	42.89	67.48	55.14	*	*	*
MSW (0–100) [‡]	66.85	43.75	47.87	51.57	52.28	<0.01	NS	
ERRORW (0–100) [‡]	86.76	56.61	61.67	52.38	46.58	*	*	*
CPST (–60 to 60) ^{**}	28.96	23.18	24.81	16.60	16.37	<0.05	*	*
CELF-Rsi ^b (3–17) ^{***}	11.26	7.76	9.36	6.46	13	<0.01	*	*
TOLD-P:2si ^b (1–20) ^{***}	11.02	6.71	9.68	7.69	10.38			
PPVT (40–160) ^{***}	107.67	95.70	102.53	106.60	108.00	NS	*	*
EOWPVT (55–145) ^{***}	135.16	104.51	110.30	110.27	113.32	<0.05	NS	*
WRAT (10–175) ^{***}	105.39	82.82	93.12	81.21	99.33	*	NS	*
WRID (10–175) ^{***}	105.88	88.70	97.12	84.72	105.33	*	<0.05	*
WIATRC (68–134) ^{***}	121.26	90.94	119.95	87.65	104.50	<0.01	<0.01	*
WIATLC (61–135) ^{***}	122.63	95.71	128.55	95.62	110.33	NS	<0.05	*
TWS-3 (60–141) ^{***}	111.80	84.30	89.51	107.50	86.33	<0.01	*	*

^a Means tests were conducted with residuals of scores adjusted for age, age², sex, SES (when found significant at $p \leq 0.05$) using a generalized linear model.

^b These measures were combined to form the metric SI.

* Significant at ≤ 0.001 level; NS = not significant at ≤ 0.05 level.

** Raw score.

*** Standard score.

⁺ Percentile.

[‡] Percent correct.

verge. Additionally, neither familial aggregation nor binary or quantitative trait regression results for non-Caucasian had greater than 15 sibling pairs, thus the results reported are limited to the Caucasian sample.

While few studies have specifically examined families of children with SSD, frequency of SSD has been reported at approximately 26% with higher frequency in brothers (40.3%) than in sisters (19.4%) [51]. This was similar to our sample in which the frequency of SSD (either alone or comorbid with LI), including parents and siblings of probands, was 29.8% with the ratio of SSD brothers to sisters 2:1. The frequency of SSD (either alone or comorbid with LI) of only the siblings of probands was 37.17%, with approximately 1/3 of the SSD affected siblings comorbid for LI. In contrast, approximately 2/3 of the SSD probands were comorbid for LI.

Cognitive Test Scores

Table 2 shows the mean scores and ranges for all cognitive test scores for all Caucasian probands and siblings with genetic data. The p values displayed in table 2 are

generated from tests of means using adjusted values. To construct the adjusted values a Generalized Linear Model with variables age, age-squared, sex and SES was generated for each test score. Significant variables ($p < 0.05$) were retained in the model and the best fit model was selected using Akaike Information Criteria (results not shown). The residual values after adjustment for all significant covariates were generated and used in all subsequent analyses. For all residual mean test scores, proband and unaffected sibling scores differed significantly, indicating that all adjusted scores were sensitive to severity of SSD affection. The SSD proband and SSD affected sibling mean scores also differed significantly for all but three measures, specifically, MSW, EOWPVT and WRAT. However, for these three scores SSD affected and unaffected siblings were discernible on the basis of the residual mean values, indicating better score sensitivity within the continuum of SSD. In contrast, two scores, PPVT and WIATLC could not be used to discern SSD affected siblings and unaffected siblings. However, for both scores proband values significantly differed from SSD affected

sibling means, indicating better test sensitivity to extreme score values (as determined by the SSD probands).

Odds Ratios and Bivariate Family Predictive Model

Prior to constructing the QEM model we calculated simple odds ratios of SSD and LI in the study population without regard to family structure. Table 3 shows the counts of individuals by SSD and LI affection status; to adjust for ascertainment we did not include probands. These numbers reflect all individuals with phenotypic data available.

The odds of finding SSD (or LI) in an individuals with LI(or SSD) was 10 times greater than finding SSD(or LI) alone, indicating that in our sample an individual is significantly more often comorbid than affected with only SSD or LI only. The significant odds justified further modeling to determine the effect of at both the individual and family level. Table 4 describes all parameters in the family predictive model and table 5 provides the values of the model coefficients, the odds ratios and confidence intervals (CI) and significance levels for all aggregation parameters. CI are robust and adjusted for correlations within persons and between persons within families. A

very significant within person association (δ) between SSD and LI was evident as seen in table 5. This is not surprising due to the high proportion of comorbid siblings (approximately 1/3) in the sample and was indicated in our simple odds calculation. The QEM also showed the odds of SSD and LI attributable to one additional SSD and LI affected family member are approximately 2 and 5, respectively. The coaggregation within the family, γ_{SL} , was also found significant, indicating that an additional person with LI (or SSD) increases the odds of SSD (or LI) within a family member. This result is taken into account when interpreting γ_{S*SL} and γ_{L*SL} . The odds ratios for these two terms are both less than one. This indicates that

Table 3. SSD and LI individuals*

N (parent and offspring)	LI = 1	LI = 0
SSD = 1	98	174
SSD = 0	43	679

* Proband is not included.

Table 4. Interpretation of the main parameters of aggregation in family predictive model in the absence of interactions

Parameter	Description	Interpretation
δ	Within person association of SSD and LI	Log odds ratio measuring the increase in log odds of SSD (or LI) in a person with LI (or SSD) compared with a person LI (or SSD)
γ_S	Aggregation of SSD within families	Increase in log odds of SSD in a person with $k+1$ relatives with SSD compared with a person with k relatives with SSD
γ_L	Aggregation of LI within families	Increase in log odds of LI in a person with $k+1$ relatives with LI compared with a person with k relatives with SSD
γ_{SL}	Coaggregation of SSD and LI within different family members	Increase in log odds of SSD (or LI) in a person with $k+1$ relatives with LI (or SSD) compared with a person with k relatives with LI (or SSD)
γ_{S*SL}	Coaggregation of SSD and LI with SSD only in different family members	Increase in log odds of SSD in a person with $k+1$ relatives with both disorders compared to a person with only k relatives with both disorders
γ_{L*SL}	Coaggregation of SSD and LI with LI only in different family members	Increase in log odds of LI in a person with $k+1$ relatives with both disorders compared to a person with only k relatives with both disorders
γ_{SL*SL}	Coaggregation of SSD and LI	Increase in log odds of both disorders in a person $k+1$ relatives with both disorders compared to a person with only k relatives with both disorders

Table 5. Model coefficients and the corresponding association measures for best fitting multivariate familial aggregation model of binary traits SSD and LI

Parameter	Coefficient	Odds (CI) ^a	Z
α_1^*	-2.819	0.059 (0.013, 0.281)	-7.67 ¹
α_2^{**}	-6.041	0.002 (0.001, 0.005)	-15.01 ¹
δ	4.096	60.09 (32.88, 109.80)	13.32 ¹
γ_S	0.720	2.05 (1.60, 2.635)	5.68 ¹
γ_L	1.637	5.14 (3.49, 7.54)	8.28 ¹
γ_{S^*L}	0.343	1.41 (1.09, 1.82)	2.62 ²
γ_{S^*SL}	-0.183	0.833 (0.780, 0.890)	-5.08 ¹
γ_{L^*SL}	-0.313	0.731 (0.608, 0.879)	-3.32 ¹
γ_{SL^*SL}	0.110	1.16 (0.916, 1.36)	1.09
θ_1^b	0.018	1.02 (0.312, 4.40)	1.10
θ_2^c	0.839	2.32 (1.44, 3.73)	-2.36 ¹

^a Odds ratios were calculated by exponentiating the parameter coefficient. The confidence intervals for the odds ratios were calculated by exponentiating the parameter coefficients confidence intervals.

^b The coefficient of the product of two indicator variables: one designating if the response equation is for SSD and the second if the individual is an offspring (1=offspring, 0=other relative type). This value is then added to the intercept for offspring (either proband or proband's sibling) with LI.

^c Intercept for offspring (either proband or proband's sibling) with LI.

¹ $p < 0.001$, ² $p < 0.025$.

* Sum of the model intercept and coefficient of the indicator variable designating the response equation.

** Model intercept.

while an individual with one disorder increases the risk of the other disorder, someone who is comorbid for SSD and LI does not further increase the risk of SSD or LI. Put another way, individuals who are comorbid do not confer additional risk for SSD or LI to other family members by being affected with a second disorder. This is again evident when we examine the odds ratio generated by the coaggregation coefficient γ_{SL^*SL} , which although above one is not significant. This confirms that comorbid individuals do not increase the risk of comorbidity in a family member. In summary, the odds ratios of γ_{S^*SL} , γ_{L^*SL} and γ_{SL^*SL} show that risk of SSD or LI in an individual does not increase multiplicatively with each additional comorbid family member.

Multipoint Linkage Analyses

We examined evidence for linkage using ASP in a conditional logistic regression as well as Haseman-Elston regression of all sibling pairs. The results of these multi-

point linkage analyses for $p < 0.05$ significance are reported in tables 6 and 7.

Tables 6 and 7 show the p-values of the ASP analyses and Haseman-Elston trait regressions, which after adjustment for multiple testing are suggestive of linkage of two regions on chromosome 1. The first region spans slightly over 30 cM, beginning ~20 cM from our first marker, D1S468. ASP analyses of LI, SSD and SL affected sibling pairs' GFTA score differences all obtained suggestive evidence for linkage between 20.3 and 24.3 cM from marker D1S468. Haseman-Elston trait regression of GFTA scores and an indicator variable for discordant sibling pairs corroborate this result with suggestive evidence for linkage ($p = 0.002$) maximizing at D1S228, approximately 20 cM from D1S468. The regression coefficients for GFTA and SSD indicate that this linkage signal is best characterized by SSD discordant pairs with a small difference in GFTA scores. In close proximity, the linkage signal using Haseman-Elston regression of binary traits LI and SL demonstrated suggestive evidence for linkage ($p = 0.0042$ and $p = 0.0098$, respectively) and SSD approached significance ($p = 0.058$) at ~20.3 cM from our first marker. Less than 20 cM from this region there is evidence for linkage of ASP with test measures for articulation and verbal short term memory.

The second region on chromosome 1 that is suggestive of linkage spans from 69.3–82.3 cM from D1S468. Analyses of GFTA score differences between concordantly affected LI and SSD siblings showed evidence for linkage, as measured from the first marker, at 74.4 and 81.9 cM, respectively and at ~71.3 cM. Haseman-Elston analyses were suggestive of evidence for linkage of vocabulary (PPVT; $p = 0.0068$). In this second region, both types of linkage analyses were significant for phenotypic indicators of verbal short term memory (Sentence Imitation and WIATLC) maximizing between at 69.4 and 81.9 cM from D1S468.

Discussion

We first examined evidence for comorbidity and aggregation within our SSD sample using a bivariate family predictive model. Secondly, using model-free linkage methods we tested for the presence of linkage of SSD and LI to the DYX8 region, corresponding to 1p34-p36, which has previously been implicated in linkage studies of dyslexia [13, 17, 52, 53] and speech sound disorder [19]. The results of the bivariate analyses demonstrate evidence for comorbidity, familial aggregation and co-

Table 6. Relative position of Multipoint LOD score maximal linkage for binary phenotypes and trait covariates using affected sibling pairs (ASP)

Linkage peak	Affected sibling pairs	Covariate	Location of peak (cM) from D1S468	Marker	Z [†]	Z [‡]	Z [‡] - Z [†]	p [‡]	p [†]
1	LI	GFTA	24.4	D1S228	0.1	1.72	1.62	0.024	0.012
			44.1	D1S234	0	1.26	1.26	0.055	0.036
	SSD	GFTA	20.3	D1S228	0.06	2.85	2.79	0.002	0.0009
			44.1	D1S234	0	1.43	1.43	0.037	0.0103
			22.3	D1S228	0.1	1.32	1.22	0.060	0.031
			8.5	D1S214	0.1	1.81	1.71	0.049	0.028
	WIATLC	44.4	D1S1676	0	2.52	2.52	0.009	0.006	
2	SSD	GFTA	81.9	D1S220	0.04	1.12	1.08	0.083	0.014
			92.6	D1S220	0	1.12	1.12	0.076	0.023
			81.9	D1S220	0	2.52	2.52	0.003	0.0019
	SL	WIATLC	81.9	D1S220	0	2.85	2.85	0.0044	0.003

Z[†] Baseline LOD score, no covariate.

Z[‡] LOD Score with covariate.

Z[‡] - Z[†] Change in LOD score attributable to covariate.

p[‡] Covariate p value, adjusted for multiple testing.

p[†] Model p value, adjusted for multiple testing.

Table 7. Relative position and empirical p values, p < 0.10 (adjusted for multiple testing) of maximal linkage for binary and quantitative SSD measures using all Caucasian sibling pairs in a Haseman-Elston regression

Linkage peak	SSD measure	Trait	Covariate	Location of peak(s) in cM from D1S468	Nearest marker	Empirical p value
1	Binary	LI	-	12.3	D1S214	0.0042
		SSD	-	20.3	D1S228	0.058
		SL	-	12.3	D1S214	0.0098
	Quantitative	GFTA	-	16.4	D1S228	0.069
			SSD sibpair affection status (discordant = 1)	20.3	D1S228	0.002
			EOWPVT	-	2.4	D1S214
2		SI	-	71.4	D1S197	0.038
		PPVT	-	71.4	D1S197	0.0068
		WIATLC	-	71.4	D1S197	0.081

aggregation of SSD and LI. Our two maximum linkage signals, while not meeting genome-wide criteria, are suggestive of linkage and being over 20 cM apart are suggestive of evidence for linkage to markers in the DYX8 region at two locations [50]. Thus, the phenotypic and genetic analyses provide preliminary evidence that this region likely bears genes with pleiotropic effects.

Figure 1 maps three published results of linkage analyses of RD and one of SSD to a cytogenetic and physical map of chromosome 1 [13, 17, 19, 52, 53]. Three independent studies of RD show evidence for linkage to chromosome 1 from ~15–30 Mb (~24–48 cM from D1S468) and a recent meta-analysis of these studies found that this region exhibits some of the strongest evidence for the presence of a dyslexia gene [52]. Multipoint results from Grigorenko et al., 2001 [13] and Tzenova et al., 2004 [53]

indicate localization of a dyslexia gene at ~20 Mb. Both studies found Phonological Decoding achieving maximum significance closest to marker D1S199, with suggestive evidence for linkage of the categorical variable RD presenting proximal to the estimated gene location. Additionally, slightly distal to D1S199, Grigorenko et al., 2001 [13] and Tzenova et al., 2004 [53] reported evidence of linkage to Single Word Reading (orally, of printed words) and Tzenova et al., 2004 [53] reported a linkage signal to Spelling spanning ~10 Mb. As figure 1 shows, this signal encompasses the marker loci linked to a dyslexia phenotype [17] and a measure of articulation [19].

The location of our first linkage peak spans approximately 30 cM beginning distal to marker D1S228 and encompasses the linkage signals reported in all three dyslexia studies. In addition, Grigorenko's localization of the dyslexia gene places it within ~1 Mb of two of our maximum suggestive linkage signals measuring verbal short term memory ($p = 0.0009$) and articulation ($p = 0.002$). This gives some evidence that a gene in this region is potentially contributing to the linkage signals of both RD and SSD phenotypes. The validity of an SSD linkage signal in this region is strengthened by the agreement between the results of our analyses and those of Smith et al., 2005 [19] in a different SSD population. Using GFTA to measure articulation and the S.A.G.E. software program SIBPAL to perform regression based linkage analysis, our linkage signal, $p = 0.036$, and that of Smith et al., 2005 [19], $p = 0.064$, were within ~1 Mb (see fig. 1). Although we did not find evidence for linkage of any of our phonological psychometric measures used in the dyslexia studies to this chromosomal region, this could be attributable to the difference between the age distribution of our SSD participants and the participant age distribution of the dyslexia studies. Several of our study participants were not yet of reading age (3/4 of the Caucasian probands were less than 6 years of age) and therefore could not be scored on the reading based metrics used in the dyslexia studies. As these children reach reading age they will be tested on reading and spelling measures so that the RD and SSD populations can be compared more thoroughly.

Our first linkage peak also demonstrates evidence for a common underlying genetic factor between SSD and LI when both the binary and quantitative traits are analyzed together. This peak is best characterized by three groups of siblings: concordant LI, concordant SSD and discordant SSD. When GFTA was included as a covariate in LODPAL analyses, the beta and gamma coefficients indicated that as the evidence for linkage maximizes in this

region, the sibling pair GFTA score difference decreases. This means that siblings with similar GFTA scores, regardless of SSD binary affection status, comprise the linkage signal. This conclusion is supported by significant evidence for linkage shown using Haseman-Elston regressions. Including SSD affection status (discordant = 1) with GFTA sibling pair scores resulted in an overall increase in significance (from $p = 0.069$ to $p = 0.002$). In order to determine if the linked SSD discordant pairs were phenotypically similar, the SSD unaffected sibling GFTA measure was compared with that of their SSD affected sibling. Discordant sibling pairs with the strongest evidence for linkage to this region were found to have a GFTA sibling pair score difference that was less than the average concordantly SSD affected sibling pair GFTA score difference. While use of the binary linkage model gave similar results, dichotomizing phenotypes imposes an artificial threshold on these disorders [54], making the model less sensitive than direct use of the quantitative data. Using both results ensured that neither was in direct contradiction to the other and we concluded that in this case, some of the SSD discordant pairs are more accurately described as concordant sibling pairs based on the GFTA measure. Examining the data using these different approaches and measurements provided valuable information about the true phenotypic and genetic state of nature.

Our second peak, with suggestive evidence of linkage spanning from slightly distal of D1S2781 to 2 cM proximal of D1S197 can be characterized by processes that are involved more with language rather than in speech production. We draw this conclusion based on the following. The addition of LI sibling pairs to SSD sibling pairs to form a composite measure, SL, results in WIATLC showing additional evidence for linkage at ~92 cM. However, ASP analysis of WIATLC shows no evidence for linkage at this location using only SSD concordantly affected siblings (table 6). Thus, the addition of concordantly affected LI pairs produces an additional linkage signal. Also the results of our Haseman-Elston regression help corroborate that the second locus is more involved with language, as SI, WIATLC and PPVT show maximum evidence for linkage slightly proximal D1S197 (table 7).

The locational similarity of our linkage signals but phenotypic differences with those of dyslexia studies could be attributable to the study participants' age distribution differences. A validated characterized protein-coding gene, KIAA1836, lies within ~1 Mb of the D1S197 marker locus and is predominantly expressed in the brain of juveniles. If the effect of a shared allelic variant on cog-

nitive abilities changed over time this could manifest as linkage to multiple phenotypes at different ages, which is exactly what is seen in our results. Additional evidence this region contains a gene linked to language comes from comparative genomics studies. Because language is believed to have evolved in the human lineage under the influence of natural selection, genes were chosen for their ability to incrementally improve multiple components of language [55, 56], we would expect to see chromosomal regions with evidence for natural selection and language to consistently overlap. Our region on chromosome 1, specifically a 4 Mb interval, spanning from 48.58 Mb (containing KIAA1836) as well as a 7.6 Mb interval on chromosome 7q, containing FOXP2, a gene associated with articulation and speech production, demonstrated significant evidence for natural selection, defined as a reduction in overall allele diversity and an excess of derived alleles with high frequency [55, 57]. As we have suggestive evidence of linkage to a language phenotype future research will be directed towards testing for allelic association in this area.

We examined and found evidence for individual comorbidity, familial aggregation and coaggregation of SSD and LI. Additionally, quantitative SSD phenotypes showed linkage to the DYX8 region, providing further evidence of a common genetic basis for SSD and RD. This area on chromosome 1 merits further exploration using fine mapping, sequencing and mutation detection in order to elucidate the extent of genetic effects on these various cognitive disorders.

Acknowledgements

This research was supported by the NIDCD grants DC00528 and DC004005 and NHLBI training grant HL0756. Some of the results of this paper were obtained by using the program package S.A.G.E. Release 5.0, which is supported by U.S. Public Health Resource Grant RR03655 from the National Center for Research Resources. We are grateful to the families who participated in this research. We would also like to thank Mrs. Paula Wedig, who rendered technical assistance on the manuscript.

References

- Shriberg LD, Tomblin JB, McSweeney JL: Prevalence of speech delay in 6-year-old children and comorbidity with language impairment. *J Speech Lang Hear Res* 1999;42:1461–1481.
- Bishop DV, Adams C: A prospective study of the relationship between specific language impairment, phonological disorders and reading retardation. *J Child Psychol Psychiatry* 1990;31:1027–1050.
- Lewis BA, Ekelman BL, Aram DM: A familial study of severe phonological disorders. *J Speech Hear Res* 1989;32:713–724.
- Lewis BA, Thompson L: A Study of developmental speech and language disorders in twins. *J Speech Hear Res* 1992;35:1086–1094.
- Manis FR, Bride-Chang C, Seidenberg MS, Keating P, Doi LM, Munson B, Petersen A: Are speech perception deficits associated with developmental dyslexia? *J Exp Child Psychol* 1997;66:211–235.
- Elbro C: Early linguistic abilities and reading development: A review and a hypothesis. *Reading and Writing: An Interdisciplinary Journal* 1996;8:453–483.
- Rack JP: Dyslexia: The Phonological Deficit Hypothesis; in Fawcett AJ, Nicholson RI (eds): *Dyslexia in Children: Multidisciplinary perspectives*. London, Harvester, 1994.
- Shriberg LD, Austin D, Lewis BA, McSweeney JL, Wilson DL: The percentage of consonants correct (PCC) metric: extensions and reliability data. *J Speech Lang Hear Res* 1997;40:708–722.
- Bishop DV: Genetic influences on language impairment and literacy problems in children: same or different? *J Child Psychol Psychiatry* 2001;42:189–198.
- Flax JF, Realpe-Bonilla T, Hirsch LS, Brzustowicz LM, Bartlett CW, Tallal P: Specific language impairment in families: evidence for co-occurrence with reading impairments. *J Speech Lang Hear Res* 2003;46:530–543.
- Lewis BA, Freebairn L: A clinical tool for evaluating the familial basis of speech and language disorders. *Amer J of Speech Language Pathology* 1993;2:38–43.
- Felsenfeld S, McGue M, Broen PA: Familial aggregation of phonological disorders: results from a 28-year follow-up. *J Speech Hear Res* 1995;38:1091–1107.
- Grigorenko EL, Wood FB, Meyer MS, Pauls JE, Hart LA, Pauls DL: Linkage studies suggest a possible locus for developmental dyslexia on chromosome 1p. *Am J Med Genet* 2001;105:120–129.
- Grigorenko EL, Wood FB, Meyer MS, Hart LA, Speed WC, Schuster A, Pauls DL: Susceptibility loci for distinct components of developmental dyslexia on chromosomes 6 and 15. *Am J Hum Genet* 1997;60:27–39.
- Nopola-Hemmi J, Myllyluoma B, Haltia T, Taipale M, Ollikainen V, Ahonen T, Voutilainen A, Kere J, Widen E: A dominant gene for developmental dyslexia on chromosome 3. *J Med Genet* 2001;38:658–664.
- Pennington B: Annotation: the genetics of dyslexia. *J Child Psychol Psychiatry* 1991;31:193–201.
- Rabin M, Wen XL, Hepburn M, Lubs HA, Feldman E, Duara R: Suggestive linkage of developmental dyslexia to chromosome 1p34-p36. *Lancet* 7–17–1993;342:178.
- Stein CM, Schick JH, Gerry TH, Shriberg LD, Millard C, Kundtz-Kluge A, Russo K, Minich N, Hansen A, Freebairn LA, Elston RC, Lewis BA, Iyengar SK: Pleiotropic effects of a chromosome 3 locus on speech-sound disorder and reading. *Am J Hum Genet* 2004;74:283–297.
- Smith SD, Pennington BF, Boada R, Shriberg LD: Linkage of speech sound disorder to reading disability loci. *J Child Psychol Psychiatry* 2005;46:1057–1066.
- Plomin R, Kovas Y: Generalist genes and learning disabilities. *Psychol Bull* 2005;131:592–617.
- Weismer S, Plante E, Jones M, Tomblin JB: A functional magnetic resonance imaging investigation of verbal working memory in adolescents with specific language impairment. *Journal of Speech, Language, and Hearing Research* 2005;48:405–425.

- 22 Hsu L, Wijsman EM, Berninger VW, Thomson JB, Raskind WH: Familial aggregation of dyslexia phenotypes II: paired correlated measures. *Am J Med Genet (Neuropsych Genet)* 2002;114:471–478.
- 23 Raskind WH, Hsu L, Berninger VW, Thomson JB, Wijsman EM: Familial aggregation of phenotypic subtypes in dyslexia. *Behavior Genetics* 2000;30:385–399.
- 24 Vogler GP, DeFries JC: Bivariate path analysis of familial resemblance for reading ability and symbol processing speed. *Behav Genet* 1985;15:111–121.
- 25 Robbins J, Klee T: Clinical assessment of oropharyngeal motor development in young children. *J Speech Hear Disord* 1987;52:271–277.
- 26 Wechsler D: Wechsler Preschool and Primary Scale of Intelligence (WPPSI). Revised. San Antonio, TX, Psychological Corporation, 1989.
- 27 Wechsler D: Intelligence Scale for Children Third Edition (WISC), revised. San Antonio, TX, Psychological Corporation, 1989.
- 28 Wechsler D: Intelligence Scale for Children Third Edition (WISC), revised. San Antonio, TX, Psychological Corporation, 1991.
- 29 Goldman R, Fristoe M: The Goldman-Fristoe Test of Articulation. Circle Pines, MN, American Guidance Service, 1986.
- 30 Khan ML, Lewis NP: Khan-Lewis Phonological Analysis. Circle Pines, Mn, American Guidance Service, 1986.
- 31 Newcomer P, Hammill D: Test of Language Development – Primary, ed 2 (TOLD: P-2), ed 2. Austin, TX, PRO-ED, 1988.
- 32 Semel E, Wiig EH, Secord WA: Clinical Evaluation of Language Fundamentals, ed 3. San Antonio TX, The Psychological Corporation, 1997.
- 33 Kamhi AG, Catts HW: Toward an understanding of developmental language and reading disorders. *J Speech Hear Disord* 1986;51:337–347.
- 34 Decker SN, DeFries JC: Cognitive ability profiles in families of reading-disabled children. *Dev Med Child Neurol* 1981;23:217–227.
- 35 Dunn LM, Dunn LM: Peabody Picture Vocabulary Test, ed 3. Circle Pines, MN, American Guidance Service, 1997.
- 36 Gardner MF: Expressive One Word Picture Vocabulary Test-Revised. Novato, CA, Academic Therapy Press, 1990.
- 37 Woodcock RN: Woodcock Reading Mastery Tests-Revised, ed Revised. Circle Pines, MN, American Guidance Service, 1987.
- 38 Wechsler D: Wechsler Individual Achievement Test (WIAT) Reading Comprehension subtest. San Antonio, TX, Psychological Corporation, 1992.
- 39 Wechsler D: The Wechsler Individual Achievement Test (WIAT) Listening Comprehension subtest. San Antonio, TX, Psychological Corporation, 1992.
- 40 Larsen S, Hammill D: Test of Written Spelling-Third Edition (TWS-3), ed 3. San Antonio, TX, The Psychological Corporation, 1994.
- 41 Hollingshead AB: Four factor index of social status. 1975.
- 42 Epstein MP, Duren WL, Boehnke M: Improved inference of relationship for pairs of individuals. *Am J Hum Genet* 2000;67:1219–1231.
- 43 Zhao LP, Prentice RL: Correlated binary regression using a quadratic exponential model. *Biometrika* 1990;77:642–648.
- 44 Hudson JI, Laird NM, Betensky RA: Multivariate logistic regression for familial aggregation of two disorders. I. Development of models and methods. *Am J Epidemiol* 2001;153:500–505.
- 45 Hudson JI, Laird NM, Betensky RA, Keck PE Jr, Pope HG Jr: Multivariate logistic regression for familial aggregation of two disorders. II. Analysis of studies of eating and mood disorders. *Am J Epidemiol* 2001;153:506–514.
- 46 Goddard KA, Witte JS, Suarez BK, Catalona WJ, Olson JM: Model-free linkage analysis with covariates confirms linkage of prostate cancer to chromosomes 1 and 4. *Am J Hum Genet* 2001;68:1197–1206.
- 47 Haseman JK, Elston RC: The investigation of linkage between a quantitative trait and a marker locus. *Behav Genet* 1972;2:3–19.
- 48 Shete S, Jacobs KB, Elston RC: Adding further power to the Haseman and Elston method for detecting linkage in larger sibships: weighting sums and differences. *Hum Hered* 2003;55:79–85.
- 49 Benjamini Y, Yekutieli D: The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics* 2001;29:1165–1188.
- 50 Cordell HJ: Sample size requirements to control for stochastic variation in magnitude and location of allele-sharing linkage statistics in affected sibling pairs. *Ann Hum Genet* 2001;65:491–502.
- 51 Lewis BA: Pedigree analysis of children with phonology disorders. *J Learn Disabil* 1992;25:586–597.
- 52 Grigorenko EL: A Conservative Meta-Analysis of Linkage and Linkage-Association Studies of Developmental Dyslexia. *Scientific Studies of Reading* 2005;9:285–316.
- 53 Tzenova J, Kaplan BJ, Petryshen TL, Field LL: Confirmation of a dyslexia susceptibility locus on chromosome 1p34-p36 in a set of 100 Canadian families. *Am J Med Genet B Neuropsychiatr Genet* 2004;127:117–124.
- 54 Shaywitz SE: Dyslexia. *N Engl J Med* 1998;338:307–312.
- 55 Enard W, Przeworski M, Fisher SE, Lai CS, Wiebe V, Kitano T, Monaco AP, Paabo S: Molecular evolution of FOXP2, a gene involved in speech and language. *Nature* 2002;418:869–872.
- 56 Hannula-Jouppi K, Kaminen-Ahola N, Taipale M, Eklund R, Nopola-Hemmi J, Kaariainen H, Kere J: The axon guidance receptor gene *ROBO1* is a candidate gene for developmental dyslexia. *PLoS Genet* 2005;1:e50.
- 57 Clark AG, Glanowski S, Nielsen R, Thomas PD, Kejariwal A, Todd MA, Tanenbaum DM, Civello D, Lu F, Murphy B, Ferreira S, Wang G, Zheng X, White TJ, Sninsky JJ, Adams MD, Cargill M: Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios. *Science* 2003;302:1960–1963.