Published in final edited form as:

Psychiatr Genet. 2014 October ; 24(5): 191–200. doi:10.1097/YPG.0000000000000045.

Association between *AVPR1A*, *DRD2*, and *ASPM* and Endophenotypes of Communication Disorders

Catherine M. Stein^{1,5}, Barbara Truitt¹, Fenghua Deng¹, Allison Avrich Ciesla¹, Feiyou Qiu¹, Peronne Joseph¹, Rekha Raghavendra¹, Jeremy Fondran¹, Robert P. Igo Jr.¹, Jessica Tag², Lisa Freebairn², H. Gerry Taylor^{2,3}, Barbara A. Lewis², and Sudha K Iyengar^{1,4}

¹Department of Epidemiology & Biostatistics, Case Western Reserve University, Cleveland, OH

²Department of Psychological Sciences, Case Western Reserve University, Cleveland, OH

³Department of Pediatrics, Case Western Reserve University, Cleveland, OH

⁴Department of Genetics, Case Western Reserve University, Cleveland, OH

⁵Center for Proteomics & Bioinformatics, Case Western Reserve University, Cleveland, OH

Abstract

Objectives—Speech sound disorder (SSD) is one of the most common communication disorders, with prevalence rates of 16% at 3 years of age, and an estimated 3.8% of children still presenting speech difficulties at 6 years of age. Several studies have identified promising associations between communication disorders and genes in brain and neuronal pathways, but there have been few studies focusing on SSD and its associated endophenotypes. Based on the hypothesis that neuronal genes may influence endophenotypes common to communication disorders, we focused on three genes related to brain and central nervous system functioning: dopamine D2 receptor (*DRD2*), arginine-vasopressin receptor 1a (*AVPR1A*), and microcephaly gene *ASPM*.

Methods—We examined the association of these genes with key endophenotypes of SSD – phonological memory measured by multisyllabic and nonword repetition, vocabulary measured by Expressive One Word Picture Vocabulary Test (EOWPVT) and Peabody Picture Vocabulary Test (PPVT), and reading decoding measured by Woodcock Reading Mastery Tests Revised – as well as the clinical phenotype of SSD. We genotyped tag SNPs in these genes and examined 498 individuals from 180 families.

Results—These data show several SNPs in all three genes were associated with phonological memory, vocabulary, and reading decoding with $p < 0.05$. Notably, associations in *AVPR1A* (rs11832266) were significant after multiple testing correction. Gene-level tests showed *DRD2* was associated with vocabulary, *ASPM* with vocabulary and reading decoding, and *AVPR1A* with all three endophenotypes.

Address correspondence to: Catherine Stein, Ph.D., Department of Epidemiology and Biostatistics, Case Western Reserve University, 2103 Cornell Rd, Wolstein Research Building, Cleveland, OH 44106, catherine.stein@case.edu. Sudha Iyengar, Ph.D., Department of Epidemiology and Biostatistics, Case Western Reserve University, 2103 Cornell Rd Wolstein Research Building, Cleveland, OH 44106, ski@case.edu.

CONFLICT OF INTEREST: The authors have no conflicts of interest to declare.

Conclusions—Endophenotypes common to SSD, language impairment, and reading disability are all associated with these neuronal pathway genes.

Keywords

neural genes; phonology; speech impairment; language development; genetics

INTRODUCTION

Speech sound disorder (SSD) is one of the most common types of communication disorders, with prevalence rates of 16% at 3 years of age, and an estimated 3.8% of children continuing to present speech delay at 6 years of age (Shriberg, Tomblin, & McSweeney 1999). SSD is characterized by difficulties in articulation and weaknesses in phonological representation. SSD is often co-morbid with language impairment (LI) (Shriberg, Tomblin, & McSweeney 1999) and may be a precursor to reading disability (RD) (Lewis, Freebairn, & Taylor 2000a). These three communication disorders are also associated with similar deficits of cognitive skills or “endophenotypes” such as phonological memory, phonological awareness, vocabulary, and appreciation of sound-symbol relationships as required for reading decoding (DeThorne et al. 2006; Lewis et al. 2011; Rvachew 2007).

Several studies have shown that SSD and other communication disorders are partially influenced by genetic factors (Lewis et al. 2006; Newbury & Monaco 2010). Because of the comorbidity among SSD, LI, and RD, and the cognitive domains they share, many studies have revealed common genetic effects among loci linked and/or associated with SSD, LI, and RD, as well as their shared endophenotypes (Bishop & Hayiou-Thomas 2008; Lewis, Shriberg, Freebairn, Hansen, Stein, Taylor, & Iyengar 2006; Pennington & Bishop 2009). Vocabulary skills are related to both reading decoding and comprehension. In addition, vocabulary is associated with the formation of phonological representations for words, as they may be stored and organized in memory according to their phonological features and articulatory gestures (Goswami 2001). Phonological Memory (PM) is the component of short term memory that holds a temporary store of phonological information, a process necessary for the formation of stable phonologic representations (Gathercole & Baddeley 1990). Deficits in PM, as measured by the repetition of multisyllabic real and nonwords, may impair an individual’s ability to learn both spoken and written new words and result in SSD, LI and/or RD (Bishop, Adams, & Norbury 2006; Shaywitz 1998). The endophenotype of reading decoding refers to the process of translating printed words into sounds. Reading decoding draws on skills such as phonological awareness that may be deficient in children with SSD, LI and/or RD (Catts et al. 1999; Nathan et al. 2004). Genetic association studies of biologically plausible candidate genes may help to elucidate how and why these phenotypes are correlated (Lewis, Shriberg, Freebairn, Hansen, Stein, Taylor, & Iyengar 2006).

Cognitive endophenotypes are more closely tied to underlying gene expression and thus have increased power to detect genetic association (Gottesman & Gould 2003; Rice, Saccone, & Rasmussen 2001). In addition, continuous traits as opposed to binary traits have increased statistical power for detecting genetic effects (Duggirala et al. 1997). As described above, the endophenotypes of phonological memory, vocabulary, and reading decoding are

key components of multiple communication disorders, and regulation of these and related executive functions occurs in the same regions of the brain (Barbey et al. 2012). Thus, our focus on these endophenotypes reveals genetic effects that may be relevant to SSD, LI, and RD. Indeed, previous linkage analyses performed by our groups and others (Cardon et al. 1994; Gayan et al. 1999; Miscimarra et al. 2007; Smith SD et al. 2005; Stein et al. 2004; Stein et al. 2006; Willcutt et al. 2002) that utilized endophenotypes have been very successful in identifying significant linkage to the same chromosomal regions that were ascertained through different communication disorders.

Promising candidate gene associations with LI and RD have been found with genes in brain and neuronal pathways (Newbury & Monaco 2010). In our previous work, we conducted linkage analysis of regions that were previously linked to dyslexia, using endophenotypes common to RD and SSD, and found that these same chromosomal regions are linked to multiple correlated phenotypes (Lewis, Avrich, Freebairn, Hansen, Sucheston, Kuo, Taylor, Iyengar, & Stein 2011; Miscimarra, Stein, Millard, Kluge, Cartier, Freebairn, Hansen, Shriberg, Taylor, Lewis, & Iyengar 2007; Stein, Schick, Taylor, Shriberg, Millard, Kundtz-Kluge, Russo, Minich, Hansen, Freebairn, Elston, Lewis, & Iyengar 2004), it is also possible that similar phenotypes may have different underlying etiologies (Lewis, Shriberg, Freebairn, Hansen, Stein, Taylor, & Iyengar 2006). Previous studies of SSD have focused on dyslexia candidate genes, and have not explored new genes in neural pathways; thus, genetic association studies are needed in this understudied area.

Dopamine has a major role in fine motor movements (Comings, Wu, Chiu, Ring, Gade, Ahn, MacMurray, Dietz, & Muhleman 1996) due to its role in neural inhibition (Berman & Noble 1995). It is involved in language learning, procedural learning and working memory (Wong et al. 2012). In an fMRI study, the dopamine D2 receptor (*DRD2*) was associated with activity of the cortico-subcortical motor region (Fazio et al. 2011). *DRD2* is a candidate gene for attention deficit hyperactivity disorder (ADHD) (Marino et al. 2004), which is often comorbid with various communication disorders (Lewis et al. 2012); because of this comorbidity, *DRD2* has been considered a candidate gene for communication disorders. *DRD2* has been associated with dysfluent speech and stuttering in a Han Chinese population (Lan et al. 2009), verbal skills as measured by the Peabody Picture Vocabulary Test (Beaver et al. 2010), and stuttering in children with Tourette syndrome (Comings, Wu, Chiu, Ring, Gade, Ahn, MacMurray, Dietz, & Muhleman 1996) but was not associated with dyslexia in an Italian population (Marino et al. 2003). Most recently, a SNP in *DRD2* has been associated with overall language deficits captured by nonword repetition and verbal comprehension (Eicher et al. 2013).

The arginine-vasopressin (AVP) hormone affects social behavior and vocalization (Winslow & Insel 1991), and has been associated with autism and social behavior and cognition (Ebstein et al. 2012) as well as cognitive processes related to social, non-emotional information (Brunnlieb et al. 2013). AVP receptor 1a (*AVPR1A*) has been associated with LI in autistic children (Wassink et al. 2004). A small study of university students found that *AVPR1A* was associated with musical and phonological memory (Granot et al. 2007).

The Abnormal spindle-like microcephaly-associated protein (*ASPM*) gene is one of the genes that causes microcephaly, and it is involved in orientation of mitotic spindles during embryonic neurogenesis, adult neurogenesis, and decreased neural cell number (Mahmood, Ahmad, & Hassan 2011), through Wnt signaling pathways (Buchman, Durak, & Tsai 2011). Dediu and Ladd proposed that *ASPM* is associated with subtle differences in the organization of the cerebral cortex, resulting in cognitive consequences (Dediu & Ladd 2007). Dediu (Dediu 2011) hypothesized a relationship between *ASPM* and language development based on ecological analyses, and Kelsey et al. (Kelsey et al. 2007) reported an association between *ASPM* and single word reading and speech sound production proficiency. In addition, *ASPM* variants have been associated with lexical tone perception in Mandarin (Wong, Chandrasekaran, & Zheng 2012).

Though these three genes have plausible roles in brain and nervous system functioning, few studies have examined their potential contributions to communication disorders. Furthermore, most of these studies have only focused on one or a few polymorphism(s) per gene, thus overlooking other polymorphisms with unknown function and the complex linkage disequilibrium (LD) structure within these genes, and these genes do not reside on chromosomal regions previously linked to SSD, LI, and/or RD. The objective of this study was to examine the association of *DRD2*, *AVPR1A*, and *ASPM* with endophenotypes related to SSD. We focused on phonological memory, vocabulary, and reading decoding, as deficits in these domains are associated with SSD as well as with LI and RD.

METHODS

Study Design and Patient Enrollment

Data were collected as part of an ongoing longitudinal study of SSD that has been described in detail elsewhere (Lewis, Avrich, Freebairn, Hansen, Sucheston, Kuo, Taylor, Iyengar, & Stein 2011; Lewis, Freebairn, & Taylor 2000b; Stein, Schick, Taylor, Shriberg, Millard, Kundtz-Kluge, Russo, Minich, Hansen, Freebairn, Elston, Lewis, & Iyengar 2004). To summarize, probands were referred to the study from clinical caseloads of speech-language pathologists in the Greater Cleveland area. Probands were identified between the ages of 4 and 6, were enrolled in therapy for moderate to severe SSD, and had normal hearing and intelligence. Siblings were diagnosed with SSD using the same criteria as the probands, based on their performance on the Goldman-Fristoe Test of Articulation (GFTA) (Goldman & Fristoe 1986). Parents were defined as having SSD if they reported a history of speech therapy. All children and parents providing informed consent were enrolled in the study. As part of the longitudinal design, children and adults were administered age-appropriate measures (described below) at each visit; for statistical analyses, the first available assessment for each measure was utilized. Both the age range of the specific test and the cooperativeness of the individual children or parents affected the sample size available for analyses, so available sample sizes for each measure are provided below. This study was approved by the institutional review board of the Case Medical Center and University Hospitals of Cleveland, Ohio.

Measures

Phonological memory—The Multisyllabic Word Test (MSW) (Catts 1986) and Nonsense Word Repetition Test (NSW) (Kamhi & Catts 1986) were used to assess phonological short-term memory skills. Participants 4 years of age through adult were asked to repeat 20 multisyllabic real (MSW test) and 15 nonsense words (NSW test) in response to audiorecorded presentations of the words. Responses were audiorecorded and transcriptions scored to obtain the percentage of words correctly repeated. In prior studies, MSW and NSW have discriminated individuals with histories of speech and language disorders who no longer demonstrate overt speech production errors in conversational speech from individuals without such histories, even into adulthood (Lewis et al. 2007; Lewis & Freebairn 1998). Data were available for 498 individuals for MSW and NSW.

Vocabulary—Peabody Picture Vocabulary Test—Third Edition (PPVT) (Dunn & Dunn 1997) and the Expressive One Word Picture Vocabulary Test—Revised (EOWPVT) (Gardner 1990) assess receptive and expressive vocabulary, respectively. PPVT is standardized for individuals 2–80 years of age, and EOWPVT is standardized for individuals 2–18 years of age. PPVT scores were available for 450 individuals, and EOWPVT scores for 405 children.

Reading measures—We administered the Woodcock Reading Mastery Tests—Revised (Woodcock 1987) Word Identification (WI) and Word Attack (WA) subtests to assess reading decoding. The WI task includes reading and identification of real words, and the WA test assesses reading decoding of nonwords. These reading tests are appropriate for children aged 7–18 years. WI scores were available for 467 children and WA scores for 466 children.

Because these individual measures may tap different cognitive domains (eg, expressive versus receptive language, identification of known words versus application of phonetic and structural analysis to read unknown words) (Bishop & Hayiou-Thomas 2008; Conti-Ramsden & Botting 1999; Woodcock 1987), we analyzed them individually rather than creating composite scores.

Binary traits—In addition to examining history of SSD (as defined in the Patient Enrollment section above) as a binary trait, we also created a binary trait using the GFTA. Analysis of GFTA as a continuous variable is problematic because it has a steep trajectory; typical children do not have overt speech errors beyond 9 years of age (Shriberg 2010; Smit et al. 1990; Wren, Roulstone, & Miller 2012). Thus, we created a binary variable capturing scores above and below the 25th percentile; this represents a severe articulation phenotype. The first version of the GFTA is normed through 16 years of age and the second version is normed through age 21, so parents do not get tested on this measure. This variable was available for 355 individuals.

Genetic methods

DNA was collected from blood or saliva samples and prepared as previously reported (Stein, Schick, Taylor, Shriberg, Millard, Kundtz-Kluge, Russo, Minich, Hansen, Freebairn, Elston,

Lewis, & Iyengar 2004). We genotyped single nucleotide polymorphisms (SNPs) within *DRD2*, *AVPR1A*, and *ASPM* using a KASP platform at LCG Genomics (Teddington, Middlesex, UK) and called via the KASP SNP Genotyping System. SNP functions were looked up on Genome Variation Server (<http://gvs.gs.washington.edu/GVS137/>) and SNPDOC (<https://wakegen.phs.wakehealth.edu/public/snpdoc/index.cfm>). We used a haplotype-tagging approach to capitalize on the correlation (LD) structure within the genes while minimizing genotyping efforts (Gabriel et al. 2002). TagSNPs were selected with an R-squared threshold of 80% and 5% minor allele frequency (MAF) in the CEPH reference population, which is most appropriate for a predominantly Caucasian study population, using Haploview (Barrett et al. 2005). SNPs were selected 20 kb up- and downstream of the gene. Both parents and offspring were genotyped. There were 4 SNPs in *ASPM*, 7 in *AVPR1A*, and 27 in *DRD2*, with a range of MAF from 3.3% to 49.9% in our study sample.

Statistical analysis

Prior to further analysis, z-scores were computed for each of the measures described above after adjusting for age (at assessment) and age². The distribution of endophenotypes in cases vs. controls was assessed using a t-test in SAS version 9.3. Genotype distributions and MAF were estimated in founders and non-founders using PLINK v1.07 (<http://pngu.mgh.harvard.edu/purcell/plink/>) (Purcell et al. 2007). SNP genotypes were also tested for Hardy-Weinberg proportions in founders and non-founders, both affected and unaffected with SSD, using PLINK. Standard quality control (QC) filters were applied (Laurie et al. 2010), resulting in exclusions of SNPs with Hardy-Weinberg equilibrium (HWE) χ^2 P-value $<1.0 \times 10^{-6}$, MAF <0.01 , or genotype call rates $<95\%$. Call rates and MAF in our data met these criteria. Only one SNP was in significant departure from Hardy-Weinberg proportions: rs4938014 in *DRD2*, with $p=3 \times 10^{-10}$ in unaffected founders, and $p=3.7 \times 10^{-20}$ in all unaffected individuals, though this latter test is biased because it does not properly account for relationships. This SNP was discarded from further analysis. Analyses were conducted in the entire sample as a whole and repeated in the Caucasian subsample. In general, the results were consistent among these two analyses, though slightly less significantly (within the same order of magnitude) so in the Caucasian subsample, as would be expected in a smaller sample size. Thus, only the results from the entire sample are presented for brevity.

Genetic association analyses were conducted using a method that models familial correlations as a polygenic random effect and estimates polygenic variance, as implemented in GWAf version 1.2 (Chen & Yang 2010). Both parents and offspring are included in this analysis when their phenotypes are available. For binary traits, GWAf estimates the correlation structure among relatives via a generalized estimating equation (GEE) model. Sex was included as a covariate in the model. SNPs were coded according to an additive or dominant model with respect to the minor allele. If there were fewer than 10 homozygotes for the rare allele in either the affected or unaffected group (Supplemental Table 1), the results from the additive model were not considered. Similarly, if there were fewer than 10 rare homozygotes plus heterozygotes, results from the dominant model were not considered. Adjustment for multiple testing of SNPs was performed with a correction based on the effective number of independent tests accounting for LD structure among SNPs in the same gene, as determined by the program SNPSpDlite (<http://gump.qimr.edu.au/general/daleN/>

SNPSpDlite) using the estimate proposed by Li and Ji (Li & Ji 2005). This analysis estimated that there are 38 independent tests, resulting in a corrected significance level of $\alpha=0.0013$.

Additionally, because of the multiple SNPs and multiple traits analyzed, we examined a gene-level statistic for each trait. Gene-level tests have two primary advantages. First, by providing a single test of association for a given gene and endophenotype, interpretation for the reader is simpler and reflects the gene's role in that endophenotype and not necessarily that of a specific genetic variant. Second, a single genetic association test per gene reduces the multiple testing problem, and thus a traditional $\alpha=0.05$ level may be used for declaring statistical significance. These tests are essentially burden tests for common variants. For this, we used VEGAS (<http://gump.qimr.edu.au/VEGAS/>), which implements the method of Liu et al. (Liu et al. 2010) to derive a summary statistic for all SNPs within a gene (or pathway) while accounting for LD among the SNPs. This method was used for *DRD2* and *ASPM*. Because *AVPR1A* had a SNP with low numbers of individuals with the minor allele (see Supplemental Table 1), we used a burden test that is appropriate for rare and common alleles in family data (Feng, Elston, & Zhu 2011). This method computes a “genetic score” that can be used as a fixed effect within a regression model.

RESULTS

This analysis included 498 individuals from 180 families, including both children and parents (Table 1). Of the individuals in this study, 50.2% had SSD. Most of them were males (67.2%, data not shown), which is typical of SSD because it is more prevalent in males than females. Most of the endophenotypes had significantly different means in individuals with and without SSD with $p < 0.05$, with the exception of EOWPVT ($p=0.0665$) (Table 2). This reaffirms our previous work demonstrating these endophenotypes significantly differ in individuals with and without SSD. In addition, as we have shown in our previous analyses, these endophenotypes are highly correlated (Table 3); measures within the same domain (phonological memory, vocabulary, and reading) were highly correlated ($r > 0.6$), and measures across domains were also significantly correlated ($r > 0.3$).

As can be seen by the genotype counts in affected and unaffected individuals (Supplemental Table 1), many SNPs had fewer than 10 rare homozygotes. For this reason, only results from the dominant model were considered further. The LD patterns for each gene are shown in Figures 1a–1c. The SNPs in *ASPM* were in very high LD, and the SNPs in *AVPR1A* and *DRD2* were in moderate to high LD.

Results significant at the nominal $\alpha=0.05$ level from the dominant model are summarized in Tables 4 (continuous traits), full results of the dominant model are provided in Supplemental Table 2, and results of the binary trait analyses significant at the $\alpha=0.05$ level are provided in Supplemental Table 3. In *AVPR1A*, one SNP was significantly associated with multiple traits after multiple testing correction; rs1182266 was associated with PPVT ($p=4.27 \times 10^{-4}$), WI ($p=4.46 \times 10^{-4}$), and WA ($p=4.81 \times 10^{-4}$), and nominally associated with MSW ($p=0.027$) and EOWPVT ($p=0.010$). Other SNPs were associated with all traits at the $\alpha=0.05$ level.

Because one SNP (rs11174810) had low frequencies for both the rare homozygote and heterozygote, we used a gene burden test appropriate for rare alleles (Table 5). This analysis revealed associations between *AVPR1A*, including EOWPVT ($p=0.015$), PPVT ($p=0.002$), NSW ($p=0.002$), MSW ($p=0.011$), WI ($p=0.006$) and WA ($p=0.001$). Thus, *AVPR1A* appears to be associated with receptive and expressive vocabulary, phonological memory, and reading decoding.

Though there were no SNPs in *DRD2* or *ASPM* that were associated with these traits after multiple testing correction, there were several SNPs associated with multiple traits at the nominal $\alpha=0.05$ level, suggesting these associations are robust. Shared associations at the uncorrected $\alpha=0.05$ level were evident at rs4938013 with EOWPVT, PPVT, MSW, WI and WA, as well as rs2734948 with NSW, MSW, WI and WA, and rs10891549 with NSW, WI and WA. In *DRD2*, 4 SNPs with PPVT, 3 SNPs with NSW, 2 SNPs with MSW, 4 SNPs with WI, and 4 SNPs with WA at the nominal $\alpha=0.05$ level. Only PPVT was associated with *DRD2* in the gene-level test ($p=0.043$) (Table 5). Thus, receptive vocabulary was associated with *DRD2* and there was suggestive evidence of association with phonological memory and reading.

Fewer SNPs within *ASPM* were associated with traits at the $\alpha=0.05$ level; 1 SNP was associated with PPVT, 1 SNP with MSW, 1 SNP with WI, and 1 SNPs with WA. Shared associations were again evident; rs10754215 was associated with PPVT, WI, and WA at the uncorrected $\alpha=0.05$ level. In the gene-level test, both PPVT ($p=0.024$) and WA ($p=0.04$) were associated with *ASPM*. Thus, *ASPM* appears to be associated with both receptive language and reading decoding.

DISCUSSION

This is the first study to examine associations of polymorphisms in *DRD2*, *AVPR1A*, and *ASPM* with endophenotypes of communication disorders and SSD and is to our knowledge the most rigorous study of these genes with respect to LD coverage. Previous studies of the association of these genes with speech and language phenotypes generally only focused on one polymorphism per gene. We found that all three genes were associated with measures of vocabulary, phonological memory, and reading, though all three genes were most significantly associated with vocabulary. These common association effects are similar to genetic linkage effects observed between other dyslexia and SSD loci (Lewis, Shriberg, Freebairn, Hansen, Stein, Taylor, & Iyengar 2006).

The endophenotypes of phonological memory, phonological awareness, vocabulary and processing speed underlie both early SSD and LI and later reading disorder (RD), spelling disorders and written expression disorders, thus accounting for the high co-morbidity of these disorders (Pennington BF 1997; Pennington & Bishop 2009). Genes act on specific biologic process which can span a range of domains and brain regions, therefore association signals can dissolve distinctions between seemingly different communication and learning disorders (Haworth & Plomin 2010). These endophenotypes are highly correlated, so shared association effects are not surprising and do not represent true pleiotropy, as occurs when

common associations are observed between the same SNP and phenotypes that are not known to be related (Pendergrass et al. 2011; Sivakumaran et al. 2011).

AVP is involved in social behavior and vocalization (Winslow & Insel 1991) and *AVPR1A* has been associated with autism (Ebstein, Knafo, Mankuta, Chew, & Lai 2012), which in turn has been related to decreased language ability in autistic children (Wassink, Piven, Vieland, Pietila, Goedken, Folstein, & Sheffield 2004). We found gene-level and significant SNP-specific association between *AVPR1A* and both expressive and receptive vocabulary, which reflects this gene's role in language. Granot et al. (Granot, Frankel, Gritsenko, Lerer, Gritsenko, Bachner-Melman, Israel, & Ebstein 2007) provisionally reported an association between *AVPR1A* and phonological memory as measured by auditory discrimination and nonword repetition, and we replicated the finding for nonword repetition in a much larger cohort. Lastly, we observed association between *AVPR1A* and reading, which reflects this gene's role in language and memory abilities that contribute to reading skills. Our most significant finding was at rs11832266, ($p < 10^{-3}$); this SNP is in a methylated region (<http://www.genome.ucsc.edu/cgi-bin/hgTracks?hgsid=356406641>) and is a clinical variant associated with unexplained developmental and intellectual disability (<http://www.ncbi.nlm.nih.gov/clinvar/RCV000051959/>)

The dopaminergic system is involved in motor control, endocrine function, cognition, language learning, procedural learning, and working memory (Frank & Fossella 2011; Hoenicka et al. 2007; Wong, Morgan-Short, Ettlinger, & Zheng 2012). Berman and Noble found that variants of *DRD2* have been associated with both visuospatial performance and vocabulary (Beaver, Delisi, Vaughn, & Wright 2010; Berman & Noble 1995). Recently, a variant in *DRD2* (rs6278) has been associated with a language variable that combines nonword repetition and verbal comprehension (Eicher, Powers, Cho, Miller, Mueller, Ring, Tomblin, & Gruen 2013); this SNP is tagged by Block #2 (Figure 1). In our study, we found that variants in *DRD2* were associated primarily with receptive vocabulary but additionally with measures of expressive vocabulary, phonological memory and reading. This suggests that *DRD2*'s role in cognition, learning, visuospatial skills, and especially working memory are also important for phonological memory. A study of Italian dyslexics did not detect a statistically significant association with *DRD2* (Marino, Giorda, Vanzin, Molteni, Lorusso, Nobile, Baschirotto, Alda, & Battaglia 2003), possibly because they only analyzed one polymorphism, while our SNP coverage was extensive. Alternatively, we may have had increased power because we analyzed quantitative traits.

ASPM is involved in mitotic spindle formation, neurogenesis, cognition, and is associated with intracranial volume (Mahmood, Ahmad, & Hassan 2011; Rimol et al. 2010). In addition, examination of adaptive changes occurring with signatures of natural selection in *ASPM* and their co-evolution with language have led to the hypothesis that *ASPM* may be related to language development (Dediu 2011). We found that SNPs in *ASPM* were associated with receptive vocabulary and, reading. This replicates a report of association between a SNP in *ASPM* and single word reading and speech sound production proficiency (Kelsey, Tomblin, Bjork, Iyengar, Sucheston, Samelson, & Murrar 2007); though their association was with a different SNP than we genotyped, we likely captured it via tagging because of the strong LD in this gene. Wong et al. (2012) found that *ASPM* was associated

with lexical tone perception (recognition of Mandarin vowel syllable and pitch) but not with the mental manipulation of phonemes as measured by sound blending, suggesting *ASPM* has a role in the memory component of human communication. Together, these findings indicate that neuronal processes regulated by *ASPM* influence speech, language, and reading abilities.

The results must be considered in light of several study limitations. Though parental phenotypes were available for SSD, MSW, and NSW, they were missing for the other endophenotypes, as some measures are only age-appropriate through adolescence and some parents refused testing. Although to our knowledge this is the largest ongoing longitudinal study of genetic influences on SSD and our sample was of moderate size (N=498), these factors limited the sample sizes available for these genetic association analyses. A further limitation was reliance on a history of therapy as the basis for the clinical diagnosis of SSD in parents and older siblings. As some individuals may have been undiagnosed, the binary SSD trait may have misclassified in some instances. Because we do not have genome-wide data available, we are unable to conduct analyses to evaluate the potential impact of population substructure. Finally, many of our findings did not remain significant after correction for multiple testing. However, we believe the findings are robust given associations across multiple SNPs and multiple phenotypes, and we draw our strongest conclusions from results of gene-level tests.

In conclusion, we observed significant association between a SNP in *AVRR1A* and receptive vocabulary and reading, and gene-level associations between *AVPR1A* and all endophenotypes, *ASPM* and receptive vocabulary and reading, and *DRD2* and receptive vocabulary. Additional studies of other neuronal pathway genes are warranted to see if they may also be related to communication disorders. The comorbidity between SSD and LI and RD make these strong candidate genes for those communication disorders and justify independent replication. Multivariate analyses will also be needed to disentangle these correlated genetic effects in order to construct a model of how these genes work collectively.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

FUNDING: This research was supported by the National Institutes of Health, National Institute on Deafness and Other Communication Disorders, Grant DC00528 awarded to Barbara A. Lewis, and grant DC012380 awarded to Sudha Iyengar.

We would like to thank the families who have so generously participated in this study for many years. Also, we would like to thank Drs. Fung and Zhu with assistance in using their genetic burden test.

References

- Barbey AK, Colom R, Solomon J, Krueger F, Forbes C, Grafman J. An integrative architecture for general intelligence and executive function revealed by lesion mapping. *Brain*. 2012; 135(Pt 4): 1154–1164. [PubMed: 22396393]

- Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*. 2005; 21(2):263–265. [PubMed: 15297300]
- Beaver KM, Delisi M, Vaughn MG, Wright JP. Association between the A1 allele of the DRD2 gene and reduced verbal abilities in adolescence and early adulthood. *J Neural Transm*. 2010; 117(7): 827–830. [PubMed: 20532925]
- Berman SM, Noble EP. Reduced visuospatial performance in children with the D2 dopamine receptor A1 allele. *Behav Genet*. 1995; 25(1):45–58. [PubMed: 7755518]
- Bishop DV, Adams CV, Norbury CF. Distinct genetic influences on grammar and phonological short-term memory deficits: evidence from 6-year-old twins. *Genes Brain Behav*. 2006; 5(2):158–169. [PubMed: 16507007]
- Bishop DV, Hayiou-Thomas ME. Heritability of specific language impairment depends on diagnostic criteria. *Genes Brain Behav*. 2008; 7(3):365–372. [PubMed: 17919296]
- Brunnlieb C, Munte TF, Tempelmann C, Heldmann M. Vasopressin modulates neural responses related to emotional stimuli in the right amygdala. *Brain Res*. 2013; (13):10.
- Buchman JJ, Durak O, Tsai LH. ASPM regulates Wnt signaling pathway activity in the developing brain. *Genes Dev*. 2011; 25(18):1909–1914. [PubMed: 21937711]
- Cardon L, Smith S, Fulker D, Kimberling W, Pennington B, DeFries J. Quantitative trait locus for reading disability on chromosome 6. *Science*. 1994; 266:276–279. [PubMed: 7939663]
- Catts HW, Fey ME, Zhang X, Tomblin J. Language basis of reading and reading disabilities: Evidence from a longitudinal investigation. *Scientific Studies of Reading*. 1999; 3:331–362.
- Catts H. Speech production/phonological deficits in reading disordered children. *Journal of Learning Disabilities*. 1986; 19:504–508. [PubMed: 3760705]
- Chen MH, Yang Q. GWAFA: an R package for genome-wide association analyses with family data. *Bioinformatics*. 2010; 26(4):580–581. [PubMed: 20040588]
- Comings DE, Wu S, Chiu C, Ring RH, Gade R, Ahn C, MacMurray JP, Dietz G, Muhleman D. Polygenic inheritance of Tourette syndrome, stuttering, attention deficit hyperactivity, conduct, and oppositional defiant disorder: the additive and subtractive effect of the three dopaminergic genes--DRD2, D beta H, and DAT1. *Am J Med Genet*. 1996; 67(3):264–288. [PubMed: 8725745]
- Conti-Ramsden G, Botting N. Classification of children with specific language impairment: Longitudinal considerations. *Journal of Speech, Language, and Hearing Research*. 1999; 42:1195–1204.
- Dediu D. Are languages really independent from genes? If not, what would a genetic bias affecting language diversity look like? *Hum Biol*. 2011; 83(2):279–296. [PubMed: 21615290]
- Dediu D, Ladd DR. Linguistic tone is related to the population frequency of the adaptive haplogroups of two brain size genes, ASPM and Microcephalin. *Proc Natl Acad Sci USA*. 2007; 104(26): 10944–10949. [PubMed: 17537923]
- DeThorne LS, Hart SA, Petrill SA, ater-Deckard K, Thompson LA, Schatschneider C, Davison MD. Children's history of speech-language difficulties: genetic influences and associations with reading-related measures. *J Speech Lang Hear Res*. 2006; 49(6):1280–1293. [PubMed: 17197496]
- Duggirala R, Williams J, Williams-Blangero S, Blangero J. A variance components approach to dichotomous trait linkage using a threshold model. *Genet Epidemiol*. 1997; 14:987–992. [PubMed: 9433612]
- Dunn, L.; Dunn, L. Peabody Picture Vocabulary Test. 3. American Guidance Service, Inc; Circle Pines, MN: 1997.
- Ebstein RP, Knafo A, Mankuta D, Chew SH, Lai PS. The contributions of oxytocin and vasopressin pathway genes to human behavior. *Horm Behav*. 2012; 61(3):359–379. [PubMed: 22245314]
- Eicher JD, Powers NR, Cho K, Miller LL, Mueller KL, Ring SM, Tomblin JB, Gruen JR. Associations of prenatal nicotine exposure and the dopamine related genes ANKK1 and DRD2 to verbal language. *PLoS ONE*. 2013; 8(5):e63762. [PubMed: 23691092]
- Fazio L, Blasi G, Taurisano P, Papazacharias A, Romano R, Gelao B, Ursini G, Quarto T, Lo BL, Di GA, Mancini M, Papolizio T, Rubini G, Bertolino A. D2 receptor genotype and striatal dopamine signaling predict motor cortical activity and behavior in humans. *Neuroimage*. 2011; 54(4):2915–2921. [PubMed: 21087673]

- Feng T, Elston RC, Zhu X. Detecting rare and common variants for complex traits: sibpair and odds ratio weighted sum statistics (SPWSS, ORWSS). *Genet Epidemiol.* 2011; 35(5):398–409. [PubMed: 21594893]
- Frank MJ, Fossella JA. Neurogenetics and pharmacology of learning, motivation, and cognition. *Neuropsychopharmacology.* 2011; 36(1):133–152. [PubMed: 20631684]
- Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, Defelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ, Altshuler D. The structure of haplotype blocks in the human genome. *Science.* 2002; 296(5576):2225–2229. [PubMed: 12029063]
- Gardner, M. Expressive One Word Picture Vocabulary Test-Revised. Academic Therapy Publications; Novato, CA: 1990.
- Gathercole SE, Baddeley AD. Phonological memory deficits in language disordered children - Is there a causal connection? *Journal of Memory and Language.* 1990; 29:336–360.
- Gayan J, Smith S, Cherny S, Cardon L, Fulker D, Brower A, Olson R, Pennington B, DeFries J. Quantitative-trait locus for specific language and reading deficits on chromosome 6p. *Am J Hum Genet.* 1999; 64:157–164. [PubMed: 9915954]
- Goldman, R.; Fristoe, M. The Goldman-Fristoe test of articulation. American Guidance Service; Circle Pines MN: 1986.
- Goswami, U. Early phonological development and the acquisition of literacy. Guilford Press; New York: 2001.
- Gottesman II, Gould TD. The endophenotype concept in psychiatry: etymology and strategic intentions. *Am J Psychiatry.* 2003; 160(4):636–645. [PubMed: 12668349]
- Granot RY, Frankel Y, Gritsenko V, Lerer E, Gritsenko I, Bachner-Melman R, Israel S, Ebstein RP. Provisional evidence that the arginine vasopressin 1a receptor gene is associated with musical memory. *Evolution and Human Behavior.* 2007; 28(5):313–318.
- Haworth CM, Plomin R. Quantitative genetics in the era of molecular genetics: learning abilities and disabilities as an example. *J Am Acad Child Adolesc Psychiatry.* 2010; 49(8):783–793. [PubMed: 20643312]
- Hoenicka J, Aragues M, Ponce G, Rodriguez-Jimenez R, Jimenez-Arriero MA, Palomo T. From dopaminergic genes to psychiatric disorders. *Neurotox Res.* 2007; 11(1):61–72. [PubMed: 17449449]
- Kamhi A, Catts H. Toward an understanding of developmental language disorders and reading disorders. *Journal of Speech and Hearing Disorders.* 1986; 51:337–347. [PubMed: 3773490]
- Kelsey, K.; Tomblin, J.; Bjork, J.; Iyengar, SK.; Sucheston, L.; Samelson, B.; Murrar, JC. Genes showing accelerated evolutionary properties associate with communication abilities in humans. American Society of Human Genetics annual meeting; San Diego, CA.
- Lan J, Song M, Pan C, Zhuang G, Wang Y, Ma W, Chu Q, Lai Q, Xu F, Li Y, Liu L, Wang W. Association between dopaminergic genes (SLC6A3 and DRD2) and stuttering among Han Chinese. *J Hum Genet.* 2009; 54(8):457–460. [PubMed: 19590515]
- Laurie CC, Doheny KF, Mirel DB, Pugh EW, Bierut LJ, Bhangale T, Boehm F, Caporaso NE, Cornelis MC, Edenberg HJ, Gabriel SB, Harris EL, Hu FB, Jacobs KB, Kraft P, Landi MT, Lumley T, Manolio TA, McHugh C, Painter I, Paschall J, Rice JP, Rice KM, Zheng X, Weir BS. Quality control and quality assurance in genotypic data for genome-wide association studies. *Genet Epidemiol.* 2010; 34(6):591–602. [PubMed: 20718045]
- Lewis BA, Avrich AA, Freebairn LA, Hansen AJ, Sucheston LE, Kuo I, Taylor HG, Iyengar SK, Stein CM. Literacy outcomes of children with early childhood speech sound disorders: impact of endophenotypes. *J Speech Lang Hear Res.* 2011; 54(6):1628–1643. [PubMed: 21930616]
- Lewis BA, Freebairn LA, Hansen AJ, Miscimarra L, Iyengar SK, Taylor HG. Speech and language skills of parents of children with speech sound disorders. *Am J Speech Lang Pathol.* 2007; 16(2):108–118. [PubMed: 17456889]
- Lewis BA, Freebairn LA, Taylor HG. Academic outcomes in children with histories of speech sound disorders. *J Commun Disord.* 2000a; 33(1):11–30. [PubMed: 10665511]

- Lewis BA, Freebairn LA, Taylor HG. Predicting school-age outcomes for children with histories of speech sound disorders. *Journal of Communication Disorders*. 2000b; 33:11–30. [PubMed: 10665511]
- Lewis BA, Short E, Iyengar SK, Taylor H, Freebairn L, Tag J, Avrich AA, Stein CM. Speech Sound Disorders and Attention-Deficit/Hyperactivity Disorder Symptoms. *Top Lang Dis*. 2012; 32:247–263.
- Lewis BA, Shriberg LD, Freebairn LA, Hansen AJ, Stein CM, Taylor HG, Iyengar SK. The Genetic Bases of Speech Sound Disorders: Evidence From Spoken and Written Language. *J Speech Lang Hear Res*. 2006; 49(6):1294–1312. [PubMed: 17197497]
- Lewis B, Freebairn L. Speech production skills of nuclear family members of children with phonology disorders. *Speech and Language*. 1998; 41:45–61.
- Li J, Ji L. Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity (Edinb)*. 2005; 95(3):221–227. [PubMed: 16077740]
- Liu JZ, McRae AF, Nyholt DR, Medland SE, Wray NR, Brown KM, Hayward NK, Montgomery GW, Visscher PM, Martin NG, Macgregor S. A versatile gene-based test for genome-wide association studies. *Am J Hum Genet*. 2010; 87(1):139–145. [PubMed: 20598278]
- Mahmood S, Ahmad W, Hassan MJ. Autosomal Recessive Primary Microcephaly (MCPH): clinical manifestations, genetic heterogeneity and mutation continuum. *Orphanet J Rare Dis*. 2011; 6:39. [PubMed: 21668957]
- Marino C, Giorda R, Vanzin L, Molteni M, Lorusso ML, Nobile M, Baschirotto C, Alda M, Battaglia M. No evidence for association and linkage disequilibrium between dyslexia and markers of four dopamine-related genes. *Eur Child Adolesc Psychiatry*. 2003; 12(4):198–202. [PubMed: 14505070]
- Marino C, Vanzin L, Giorda R, Frigerio A, Lorusso ML, Nobile M, Molteni M, Battaglia M. An assessment of transmission disequilibrium between quantitative measures of childhood problem behaviors and DRD2/TaqI and DRD4/48bp-repeat polymorphisms. *Behav Genet*. 2004; 34(5):495–502. [PubMed: 15319572]
- Miscimarra L, Stein C, Millard C, Kluge A, Cartier K, Freebairn L, Hansen A, Shriberg L, Taylor HG, Lewis B, Iyengar SK. Further evidence of pleiotropy influencing speech and language: analysis of the DYX8 region. *Hum Hered*. 2007; 63(1):47–58. [PubMed: 17230025]
- Nathan L, Stackhouse J, Goulandris N, Snowling MJ. The development of early literacy skills among children with speech difficulties: a test of the “critical age hypothesis””. *J Speech Lang Hear Res*. 2004; 47(2):377–391. [PubMed: 15157138]
- Newbury DF, Monaco AP. Genetic advances in the study of speech and language disorders. *Neuron*. 2010; 68(2):309–320. [PubMed: 20955937]
- Pendergrass SA, Brown-Gentry K, Dudek SM, Torstenson ES, Ambite JL, Avery CL, Buyske S, Cai C, Fesinmeyer MD, Haiman C, Heiss G, Hindorff LA, Hsu CN, Jackson RD, Kooperberg C, Le ML, Lin Y, Matise TC, Moreland L, Monroe K, Reiner AP, Wallace R, Wilkens LR, Crawford DC, Ritchie MD. The use of phenome-wide association studies (PheWAS) for exploration of novel genotype-phenotype relationships and pleiotropy discovery. *Genet Epidemiol*. 2011; 35(5):410–422. [PubMed: 21594894]
- Pennington BF. Using genetics to dissect cognition. *Am J Hum Genet*. 1997; 60:13–16. [PubMed: 8981941]
- Pennington BF, Bishop DV. Relations among speech, language, and reading disorders. *Annu Rev Psychol*. 2009; 60:283–306. 10.1146/annurev.psych.60.110707.163548 [PubMed: 18652545]
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007; 81(3):559–575. [PubMed: 17701901]
- Rice J, Saccone N, Rasmussen E. Definition of the phenotype. *Adv Genet*. 2001; 42:69–76. [PubMed: 11037314]
- Rimol LM, Agartz I, Djurovic S, Brown AA, Roddey JC, Kahler AK, Mattingdal M, Athanasiu L, Joyner AH, Schork NJ, Halgren E, Sundet K, Melle I, Dale AM, Andreassen OA. Sex-dependent association of common variants of microcephaly genes with brain structure. *Proc Natl Acad Sci USA*. 2010; 107(1):384–388. [PubMed: 20080800]

- Rvachew S. Phonological processing and reading in children with speech sound disorders. *Am J Speech Lang Pathol*. 2007; 16(3):260–270. [PubMed: 17666551]
- Shaywitz SE. Dyslexia. *N Engl J Med*. 1998; 338(5):307–312. [PubMed: 9445412]
- Shriberg, LD. Childhood speech sound disorders: From postbehaviorism to the postgenomic era. In: Paul, R.; Flipsen, P., Jr, editors. *Speech sound disorders in children*. Plural Publishing Inc; San Diego: 2010. p. 1-33.
- Shriberg L, Tomblin J, McSweeney J. Prevalence of speech delay in 6-year-old children and comorbidity with language impairment. *Journal of Speech, Language, and Hearing Research*. 1999; 42:1461–1481.
- Sivakumaran S, Agakov F, Theodoratou E, Prendergast JG, Zgaga L, Manolio T, Rudan I, McKeigue P, Wilson JF, Campbell H. Abundant pleiotropy in human complex diseases and traits. *Am J Hum Genet*. 2011; 89(5):607–618. [PubMed: 22077970]
- Smit A, Hand L, Freilinger J, Bernthal J, Bird A. The Iowa articulation norms project and its Nebraska replication. *J Speech Hear Disorders*. 1990; 55:779–798.
- Smith SD, Pennington BF, Boada R, Shriberg LD. Linkage of speech sound disorder to reading disability loci. *Journal of Child Psychology and Psychiatry*. 2005; 46:1057–1066. [PubMed: 16178929]
- Stein CM, Millard C, Kluge A, Miscimarra LE, Cartier KC, Freebairn LA, Hansen AJ, Shriberg LD, Taylor HG, Lewis BA, Iyengar SK. Speech sound disorder influenced by a locus in 15q14 region. *Behav Genet*. 2006; 36(6):858–868. [PubMed: 16786424]
- Stein CM, Schick JH, Taylor HG, 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(2):283–297. [PubMed: 14740317]
- Wassink TH, Piven J, Veland VJ, Pietila J, Goedken RJ, Folstein SE, Sheffield VC. Examination of AVPR1a as an autism susceptibility gene. *Mol Psychiatry*. 2004; 9(10):968–972. [PubMed: 15098001]
- Willcutt EG, Pennington BF, Smith SD, Cardon LR, Gayan J, Knopik VS, Olson RK, DeFries JC. Quantitative trait locus for reading disability on chromosome 6p is pleiotropic for attention-deficit/hyperactivity disorder. *Am J Med Genet*. 2002; 114(3):260–268. [PubMed: 11920845]
- Winslow J, Insel TR. Vasopressin modulates male squirrel monkeys' behavior during social separation. *Eur J Pharmacol*. 1991; 200(1):95–101. [PubMed: 1769376]
- Wong PC, Chandrasekaran B, Zheng J. The derived allele of ASPM is associated with lexical tone perception. *PLoS One*. 2012; 7(4):e34243. [PubMed: 22529908]
- Wong PC, Morgan-Short K, Ettlinger M, Zheng J. Linking neurogenetics and individual differences in language learning: the dopamine hypothesis. *Cortex*. 2012; 48(9):1091–1102. [PubMed: 22565204]
- Woodcock, R. Woodcock Reading Mastery Tests-Revised. American Guidance Service; Circle Pines, MN: 1987.
- Wren YE, Roulstone SE, Miller LL. Distinguishing groups of children with persistent speech disorder: findings from a prospective population study. *Logoped Phoniatr Vocol*. 2012; 37(1):1–10. [PubMed: 22059376]

Figure 1.
LD plots for a) *DRD2*, b) *AVPR1A*, and c) *ASPM*

Table 1

Sample description

Number of individuals	498
Number of families	180
Males / females (includes individuals with and without SSD)	271 / 227
Number of parents / children	105 / 366
Number of individuals with SSD	250 (50.2%)
<i>Ethnic distribution</i>	
Caucasian	427 (85.7%)
African American	40 (8.0%)
Hispanic	4 (0.8%)
Middle Eastern	1 (0.2%)
East Asian	5 (1.0%)
Native American	1 (0.2%)
Mixed	20 (4.0%)

Table 2

Descriptive statistics for measures for individuals with and without speech sound disorder (SSD)

	Individuals with SSD Mean (SD)	Individuals without SSD Mean (SD)	P-Value*
EOWPVT	106.25 (18.10)	114.47 (16.18)	0.0665
PPVT	100.74 (16.51)	108.45 (14.58)	0.0427
NSW	36.47 (27.50)	64.62 (22.26)	0.0037
MSW	39.28 (30.81)	77.99 (23.84)	<.0001
WI	95.52 (17.22)	102.69 (13.28)	<.0001
WA	95.22 (18.44)	104.87 (13.69)	<.0001

* T-test comparing individuals with and without SSD, after adjustment for age and age-squared.

EOWPVT = Expressive One Word Picture Vocabulary Test, PPVT = Peabody Picture Vocabulary Test, NSW = nonsense word repetition task, MSW = multisyllabic word repetition task, WI = single word reading task (Word Identification), WA = nonsense word reading task (Word Attack). All scores are standard scores except NSW and MSW, which are percent correct.

Table 3

Correlation between measures

	EOWPVT	PPVT	NSW	MSW	WI	WA
EOWPVT	1					
PPVT	0.61**	1				
NSW	0.35**	0.35**	1			
MSW	0.38**	0.39**	0.75**	1		
WI	0.37**	0.49**	0.47**	0.46**	1	
WA	0.34**	0.43**	0.43**	0.46**	0.81**	1

** significant at p < 0.001.

EOWPVT = Expressive One Word Picture Vocabulary test, PPVT = Peabody Picture Vocabulary test, NSW = nonsense word repetition task, MSW = multisyllabic word repetition task, WI = single word reading task (Word Identification), WA = nonsense word reading task (Word Attack).

Table 4

Continuous trait association results significant at $p < 0.05$

SNP	EOWPVT			PPVT			NSW			MSW			WI			WA		
	beta	p-value		beta	p-value		beta	p-value		beta	p-value		beta	p-value		beta	p-value	
<i>ASPM</i>	rs6700180	0.173	0.179	0.248	0.038	0.233	0.061	0.288	0.040	0.200	0.110	0.170	0.186					
	rs10754215	-0.189	0.170	-0.358	0.005	-0.209	0.121	-0.220	0.145	-0.322	0.016	-0.328	0.018					
<i>DRD2</i>	rs4938013	0.249	0.024	0.312	0.002	0.180	0.097	0.265	0.031	0.254	0.019	0.294	0.009					
	rs2734849	0.192	0.127	0.160	0.175	0.366	0.003	0.283	0.043	0.343	0.006	0.283	0.026					
	rs1800497	0.175	0.140	0.244	0.027	0.257	0.026	0.244	0.062	0.226	0.053	0.279	0.020					
	rs10891549	0.106	0.370	0.127	0.253	0.241	0.038	0.254	0.054	0.342	0.003	0.269	0.024					
	rs2440390	0.269	0.049	0.262	0.043	0.034	0.799	0.221	0.138	0.199	0.130	0.199	0.142					
	rs2734831	0.097	0.400	0.100	0.352	0.166	0.141	0.192	0.134	0.292	0.010	0.230	0.048					
<i>AVPR1A</i>	rs4581480	-0.145	0.293	-0.301	0.019	-0.137	0.296	-0.198	0.181	0.003	0.980	-0.006	0.961					
	rs4483623	0.059	0.753	0.365	0.039	0.011	0.954	0.146	0.489	0.167	0.372	0.003	0.986					
	rs1587097	0.107	0.475	0.113	0.422	0.337	0.022	0.190	0.247	0.347	0.019	0.286	0.060					
	rs11174811	-0.047	0.712	0.039	0.748	0.223	0.073	0.038	0.786	0.204	0.105	0.152	0.238					
	rs1042615	0.183	0.118	0.277	0.011	0.188	0.100	0.322	0.013	0.220	0.055	0.206	0.081					
	rs10784342	-0.272	0.025	-0.272	0.020	-0.011	0.929	-0.128	0.351	-0.153	0.216	-0.265	0.037					
					4.272×10⁻⁴*						4.46×10⁻⁵*		4.81×10⁻⁶*					
	rs11832266	-0.448	0.010	-0.568		-0.309	0.076	-0.428	0.027	-0.691		-0.787						

Uncorrected p-values < 0.05 are indicated in boldface font. Results significant after multiple testing correction (38 independent tests based on SNPSpDlite) are represented with *. EOWPVT = Expressive One Word Picture Vocabulary test, PPVT = Peabody Picture Vocabulary test, NSW = nonsense word repetition task, MSW = multisyllabic word repetition task, WI = single word reading task (Word Identification), WA = nonsense word reading task (Word Attack). SNP location and function can be found in Supplemental Table 1.

Table 5

Gene-level association results

Gene	EOWPVT	PPVT	NSW	MSW	WI	WA	SSD
<i>ASPM</i> ^a	0.332	0.024	0.086	0.131	0.055	0.040	0.148
<i>DRD2</i> ^a	0.207	0.043	0.270	0.198	0.118	0.116	0.157
<i>AVPR1A</i> ^b	0.015	0.002	0.002	0.011	0.006	0.001	0.649

^aUsing VEGAS method by Liu et al. (2010)

^bUsing rare variant method by Feng et al. (2011)

Uncorrected p-values shown.

EOWPVT = Expressive One Word Picture Vocabulary test, PPVT = Peabody Picture Vocabulary test, NSW = nonsense word repetition task, MSW = multisyllabic word repetition task, WI = single word reading task (Word Identification), WA = nonsense word reading task (Word Attack), SSD= binary trait for speech sound disorder.