Genes, Brain and Behavior
Official publication of the International Behavioural and Neural Genetics Society

Genes, Brain and Behavior (2015) **14:** *377–385* doi: 10.1111/gbb.12214

The DYX2 locus and neurochemical signaling genes contribute to speech sound disorder and related neurocognitive domains

J. D. Eicher[†], C. M. Stein[‡], F. Deng[‡], A. A. Ciesla[‡], N. R. Powers[†,§], R. Boada[¶], S. D. Smith[**], B. F. Pennington[††], S. K. Iyengar[‡], B. A. Lewis[‡‡] and J. R. Gruen[†,§,§§,*]

[†]Department of Genetics, Yale University School of Medicine, New Haven, CT, [‡]Department of Epidemiology & Biostatistics, Case Western Reserve University, Cleveland, OH, [§]Department of Pediatrics, Yale University School of Medicine, New Haven, CT, [¶]Department of Pediatrics, University of Colorado School of Medicine, Aurora, CO, [**]Departments of Pediatrics and Developmental Neuroscience, University of Nebraska Medical Center, Omaha, NE, [††]Department of Psychology, University of Denver, Denver, CO, [‡‡]Department of Psychological Sciences, Case Western Reserve University, Cleveland, OH, and [§§]Investigative Medicine Program, Yale Child Health Research Program, Yale University School of Medicine, New Haven, CT
**Corresponding author: J. R. Gruen, MD, Departments of Pediatrics, Genetics, and Investigative Medicine, Yale Child Health Research Center, Room 208, 464 Congress Avenue, New Haven, CT 06520-8081, USA. E-mail: jeffrey.gruen@yale.edu*

A major milestone of child development is the acquisition and use of speech and language. Communication disorders, including speech sound disorder (SSD), can impair a child's academic, social and behavioral development. Speech sound disorder is a complex, polygenic trait with a substantial genetic component. However, specific genes that contribute to SSD remain largely unknown. To identify associated genes, we assessed the association of the DYX2 dyslexia risk locus and markers in neurochemical signaling genes (e.g., nicotinic and dopaminergic) with SSD and related endophenotypes. We first performed separate primary associations in two independent samples – Cleveland SSD (210 affected and 257 unaffected individuals in 127 families) and Denver SSD (113 affected individuals and 106 unaffected individuals in 85 families) – and then combined results by meta-analysis. DYX2 markers, specifically those in the 3′ untranslated region of *DCDC2* ($P = 1.43 \times 10^{-4}$), showed the strongest associations with phonological awareness. We also observed suggestive associations of dopaminergic-related genes *ANKK1* ($P = 1.02 \times 10^{-2}$) and *DRD2* ($P = 9.22 \times 10^{-3}$) and nicotinic-related genes *CHRNA3* ($P = 2.51 \times 10^{-3}$) and *BDNF* ($P = 8.14 \times 10^{-3}$) with case–control status and articulation. Our results further implicate variation in putative regulatory regions in the DYX2 locus, particularly in *DCDC2*, influencing language and cognitive traits. The results also support previous studies implicating variation in dopaminergic and nicotinic neural signaling influencing human communication and cognitive development. Our findings expand the literature showing genetic factors (e.g., DYX2) contributing to multiple related, yet distinct neurocognitive domains (e.g., dyslexia, language impairment, and SSD). How these factors interactively yield different neurocognitive and language-related outcomes remains to be elucidated.

Keywords: *DCDC2*, dopaminergic, DYX2, nicotinic, speech sound disorder

Received 25 August 2014, revised 12 December 2014 and 5 March 2015, accepted for publication 12 March 2015

Proficiency of communication and language skills is an important aspect of child development. However, children with numerous disorders, including speech sound disorder (SSD), have impairments in these skills. Specifically, SSD is characterized by the omission or substitution of sounds and other speech production errors that prevent understanding of speech (Lewis *et al.* 2006; Newbury & Monaco 2010; Pennington & Bishop 2009). The speech deficits in children with SSD are of unknown origin and focus on sound production and articulation miscues of verbal utterances. These difficulties in articulation and sound production are common in children, with prevalence estimates ranging from 4% to 16% (Lewis *et al.* 2006; Pennington & Bishop 2009). In addition to articulation difficulties, children with SSD exhibit deficits in manipulating and understanding phonemes – the speech sound units of words (ASHA Website 2015; Bernthal *et al.* 2013; Pennington & Bishop 2009). Thus, SSD includes both errors of articulation or phonetic structure (errors due to poor motor abilities associated with the production of speech-sounds) and phonological errors (errors in applying linguistic rules to combine sounds to form words). Phonological awareness (PA) is the awareness of the sound structure of a language and includes the ability to manipulate and understand phonemes. Impairments in various facets of phonology have also been implicated in related neurodevelopmental disorders, including dyslexia and language impairment (LI) (Newbury & Monaco 2010, Pennington & Bishop 2009).

Family and twin studies have shown that there is a substantial genetic component to SSD and communication traits (Lewis *et al.* 2006; Stein *et al.* 2011). Linkage and association studies, as well as studies of rare, severe forms of SSD and/or related speech disorders, have identified several genes that influence these traits. Specifically, linkage

analyses have implicated the DYX2 locus on chromosome 6p21.3 (Peter *et al.* 2012; Smith *et al.* 2005). The DYX2 locus was first shown through linkage analysis to contribute to dyslexia, although more recent studies have also implicated the locus in LI and general cognitive abilities (Eicher *et al.* 2014; Lewis *et al.* 2006; Mascheretti *et al.* 2014; Newbury *et al.* 2011; Powers *et al.* 2013). The DYX2 locus houses two risk genes, *DCDC2* and *KIAA0319* (Becker *et al.* 2014; Cope *et al.* 2005; Konig *et al.* 2011; Ludwig *et al.* 2008; Meng *et al.* 2005; Powers *et al.* 2013; Scerri *et al.* 2011; Schumacher *et al.* 2006; Zou *et al.* 2012), although there is evidence that other DYX2 genes contribute to these traits (Eicher *et al.* 2014; Pinel *et al.* 2012; Plomin *et al.* 2004). However, no fine-mapping association studies have examined the contribution of the DYX2 locus to SSD. Therefore, it is unknown which genes are responsible for the linkage to DYX2, and more specifically, whether the risk genes *DCDC2* and *KIAA0319* contribute to SSD as well.

Recent studies have also examined the effects of prominent neurochemical signaling genes, including those in the nicotinic- and dopaminergic-related pathways, on SSD and related language processes. Our groups and others have recently reported associations of *DRD2*, which encodes a dopamine receptor, with verbal language traits and SSD (Beaver *et al.* 2010; Berman & Noble 1995; Eicher *et al.* 2013; Stein *et al.* 2014; Wong *et al.* 2013). The association of genetic markers rs4938013, rs2734849 and rs1800497 in the *DRD2/ANKK1* locus in the Cleveland SSD cohort used in this study (see below), was completed independently and before the genotyping and analyses discussed in this manuscript, and has been presented elsewhere (Stein *et al.* 2014). Dopaminergic, nicotinic and other neurochemical signaling pathways are vital to the development and execution of higher-order cognitive processes. A recent magnetic resonance spectroscopy study found that choline and glutamate levels were correlated with reading performance (Pugh *et al.* 2014). However, more studies are needed in order to further link genetic variation in genes involved in neurochemical signaling and language phenotypes.

The goals of this study were to complete an association scan of the DYX2 locus to characterize its relationship with SSD, and to test for association between specific variants in genes involved in neurochemical signaling with SSD. To accomplish these goals, we first performed primary genetic association testing using language and genetic data collected in two family-based SSD samples: the Cleveland SSD and Denver SSD cohorts. Then, we combined these genetic association results in a meta-analysis. Our hypotheses were that the known DYX2 risk genes *DCDC2* and *KIAA0319* also contribute to SSD, and that neurochemical-related genes are involved in SSD.

Methods

We used two family-based studies in our genetic analyses of SSD: the Cleveland SSD and Denver SSD cohorts. The details regarding subjects and recruitment methods for both cohorts are described elsewhere but are summarized below (Lewis *et al.* 2000, 2011; Smith *et al.* 2005; Stein *et al.* 2004). The Cleveland SSD cohort was collected at Case Western Reserve University as part of an on-going longitudinal study in the Greater Cleveland Metro area (Lewis *et al.* 2000, 2011; Stein *et al.* 2004). These analyses were restricted to individuals of self-reported Caucasian ethnicity. Probands between the ages of 4 and 6 years, who were enrolled in therapy for moderate-severe SSD, were referred by clinical speech-language pathologists. Siblings and parents of probands were also enrolled in the study with siblings diagnosed with SSD based on performance on the Goldman-Fristoe Test of Articulation (GFTA, Goldman & Fristoe 1986) and parents diagnosed based on history of speech therapy. Probands and siblings were required to meet the following criteria: (1) Standard American English as their first language; (2) normal hearing, as defined by passing a pure tone hearing screening; (3) a normal oral mechanism; (4) a Performance IQ score of 80 or above; (5) no history of neurological or developmental disorders according to parental report. All children and parents provided informed assent/consent with study protocols approved by the institutional review board of the Case Medical Center and University Hospitals of Cleveland, Ohio.

Similarly, the Denver SSD cohort is part of a longitudinal study examining SSD (Smith *et al.* 2005). Probands between the ages of 5 and 7 years were recruited through public and private schools in the Greater Denver Metro area, as well as through radio and newspaper advertisements. All probands were required to currently have or to have previously had SSD, as defined by prior testing by a speech-language pathologist and/or significant intelligibility problems at 3–4 years of age according to parent report. Additionally, all Denver SSD probands were currently receiving or previously received speech-language therapy and/or scored below the 30th percentile on the GFTA (Goldman & Fristoe 1986). Full siblings of the probands were recruited for inclusion if they were between 5 and 9 years old, independent of speech performance. However, if both the proband and siblings had received speech-language therapy, the child with the more extreme articulation deficit on the GFTA was designated as the proband. Participating families and children were required to be monolingual English speakers. Subjects were also excluded if they had: (1) known genetic disorders or syndromes, (2) intellectual disability (e.g., nonverbal IQ < 70), (3) pervasive developmental disorder, (4) significant birth complications, (5) acquired brain injury, (6) peripheral hearing loss, or (7) structural or functional speech mechanism impairments (e.g., cleft palate).

Subjects in both cohorts completed a wide array of neurocognitive and language measures that extensively characterized his or her neurocognitive profile that are described in detail elsewhere (Lewis *et al.* 2000, 2011; Smith *et al.* 2005; Stein *et al.* 2004). As the aim of this set of analyses was to leverage both cohorts and use them in a meta-analysis, we chose to focus only on phenotypes that were identical or measured the same neurocognitive domain. Thus, we examined three phenotypes described in further detail below: (1) case–control status, (2) performance on the GFTA, and (3) a composite quantitative trait measuring PA.

First, probands and siblings in both the Cleveland SSD and Denver SSD cohorts were clinically ascertained for SSD. There were 210 affected and 257 unaffected in 127 families in the Cleveland SSD cohort, and 113 affected and 106 unaffected individuals in 85 families in the Denver SSD cohort. Second, 187 subjects (affected and unaffected) from the Cleveland SSD cohort, and 175 subjects (affected and unaffected) from the Denver SSD cohort, were assessed by the GFTA for quantitative performance of articulation. Third, subjects completed multiple tasks that measure PA. Therefore, we constructed a composite PA measure in each cohort. In the Cleveland SSD cohort, the PA composite score was comprised of the Blending Words and Elision tasks of the CTOPP (Wagner *et al.* 1999); the raw scores were converted to a *z*-score, and a weighted composite score was derived using principal component analysis. In the Denver SSD cohort, the PA composite score was comprised of the Bird and Bishop (1992) rhyme judgment task, and the Elision, Blending Words and Sound Matching subtests of the CTOPP (Wagner *et al.* 1999). The raw scores from these four measures were age-corrected, then averaged and converted to a *z*-score to yield a PA composite score. There were 96 and 128 individuals, including both cases and controls, with PA composite scores in the Cleveland SSD and Denver SSD cohorts, respectively.

We developed a single nucleotide polymorphism (SNP) marker panel to tag common variation in the DYX2 locus and to specifically

Table 1: Genetic markers within neurochemical signaling genes genotyped in the Cleveland SSD and Denver SSD cohorts

Variant	Gene	Location	Variant	Gene	Location	Variant	Gene	Location
rs2072660	*CHRNB2*[1]	1q21.3	rs6278	*DRD2*[6,7]	11q23.2	rs1426153	*PKNOX2*[8]	11q24.2
rs2072661	*CHRNB2*[1]	1q21.3	rs11604671	*ANKK1*[6,7]	11q23.2	rs750338	*PKNOX2*[8]	11q24.2
rs12466358	*CHRND*[2]	2q37.1	rs1800497	*ANKK1*[6,7]	11q23.2	rs1051730	*CHRNA3*[9,10]	15q25.1
rs13277254	*CHRNB3*[3]	8p11.21	rs2734849	*ANKK1*[6,7]	11q23.2	rs1317266	*CHRNA3*[9,10]	15q25.1
rs4950	*CHRNB3*[3]	8p11.21	rs4938013	*ANKK1*[6,7]	11q23.2	rs578776	*CHRNA3*[9,10]	15q25.1
rs6474413	*CHRNB3*[3]	8p11.21	rs7118900	*ANKK1*[6,7]	11q23.2	rs6495308	*CHRNA3*[9,10]	15q25.1
rs4075274	*NTRK2*[4]	9q21.33	rs10893365	*PKNOX2*[8]	11q24.2	rs8034191	*AGPHD1*[11]	15q25.1
rs2030324	*BDNF*[5]	11p14.1	rs10893366	*PKNOX2*[8]	11q24.2	rs2229959	*CHRNA4*[12]	20q13.33
rs4274224	*DRD2*[6,7]	11q23.2	rs11220015	*PKNOX2*[8]	11q24.2	rs2236196	*CHRNA4*[12]	20q13.33
rs4648318	*DRD2*[6,7]	11q23.2	rs11602925	*PKNOX2*[8]	11q24.2	rs2273504	*CHRNA4*[12]	20q13.33
rs7131056	*DRD2*[6,7]	11q23.2	rs12284594	*PKNOX2*[8]	11q24.2			

References: [1]Wessel *et al.* (2010), [2]Saccone *et al.* (2009), [3]Saccone *et al.* (2010), [4]Beuten *et al.* (2007), [5]Beuten *et al.* (2005), [6]Huang *et al.* (2009), [7]Gelernter *et al.* (2006), [8]Chen *et al.* (2011), [9]Saccone *et al.* (2007), [10]Liu *et al.* (2010), [11]Li *et al.* (2005), [12]Amos *et al.* (2008).

interrogate candidate markers within neurochemical signaling genes. The development and composition of this marker panel is explained in full elsewhere and described briefly below (Eicher *et al.* 2013, 2014; Powers *et al.* 2013). First, we developed a SNP marker panel to capture the common variation in the DYX2 locus. TagSNPs in the DYX2 locus were selected using the association study design server of Han *et al.* (2008). In this study, the panel contained 204 SNPs within the DYX2 locus spanning approximately 1.4 Mb. Additionally, the SNP marker panel contained 32 SNPs within 11 neurochemical signaling genes (Table 1). The candidate markers were selected due to their past implications in neurochemical function and/or association with nicotine dependence. All markers were genotyped together on the Sequenom MassARRAY platform (San Diego, CA, USA) following manufacturer's guidelines at the Yale Center for Genome Analysis (West Haven, CT, USA). Briefly, markers were genotyped in nine multiplex reactions of 30–36 markers each, totaling 300 markers. A subset of the markers in this panel was neither in the DYX2 locus nor within neurochemical-related genes, and therefore, not included in the subsequent analyses presented in this study. In addition to downstream quality control during the genetic analyses (see below), the histogram plot showing clustering of genotype groups generated by the assay for each marker was manually evaluated. Mendelian errors were detected and removed in PLINK (Purcell *et al.* 2007). Markers were also removed if not in Hardy-Weinberg equilibrium ($P < 0.0001$), had call rates <0.85, or had minor allele frequency <0.01.

First, genetic associations were performed in each cohort individually using GWAF v2.0, with gender included as a covariate for the three phenotypes of interest: (1) case–control status, (2) GFTA, and (3) PA (Chen & Yang 2010). GWAF uses both linear mixed effects (LME) and generalized estimating equation (GEE) models in continuous traits, and only GEE models in binary traits to account for familial correlations in association analyses (Chen & Yang 2010). Associations were performed only under additive and dominant models, as recessive models are less reliable in samples of smaller size. Next, we completed a meta-analysis for each phenotype using the association results for both cohorts with METAL, using sample size as the weight under default settings (Willer *et al.* 2010). Tests of sample heterogeneity, as evaluated by I^2, were performed using METAL. Associations with $P < 0.05$ in our meta-analyses are reported. To correct for multiple testing and to account for the substantial linkage disequilibrium (LD) of our markers, we used SNPSpDlite to determine the number of independent hypotheses in this study (Nyholt 2004). There were 122 DYX2 and 20 non-DYX2 independent marker loci, totaling 142 tests, in these analyses. In addition to genetic marker correlation, we have estimated the average correlation among phenotypic traits to be 0.3, meaning that 70% of the phenotypes are independent. Therefore, taking into account phenotypic and genotypic correlations, we set a threshold of $0.05/[(0.7 \times 3) \times 142] = 1.68 \times 10^{-4}$.

Results

The results of the primary genetic association studies in the two independent cohorts, Cleveland SSD and Denver SSD, under both additive and dominant models are shown in the Supporting Information (Tables S1–S3). One association, rs16889240 in *GPLD1* with PA under a dominant model ($P = 2.53 \times 10^{-6}$) in the Cleveland SSD cohort, survived correction for multiple testing within DYX2 (Table S3c). There was, however, no evidence for association with this marker in the Denver SSD cohort ($P > 0.05$, data not shown). Additionally, there was one association that approached the corrected threshold for multiple testing among the neurochemical signaling markers, rs6474413 located upstream of *CHRNB3*, with PA under an additive model ($P = 3.15 \times 10^{-4}$) (Table S3a). Again, there was no evidence of suggestive association with the marker in the Denver SSD cohort ($P > 0.05$, data not shown).

In order to leverage information from both samples together as opposed to relying on associations informed by just one, we combined the above genetic associations in both cohorts in a meta-analysis. The results of these meta-analyses are presented phenotype by phenotype, with DYX2 and neurochemical signaling genes presented together. Regional association plots of DYX2 markers for each phenotype and model are located in Figure S1. There were suggestive associations with case–control status, though none approached the multiple testing corrected threshold. These were observed mostly with markers in neurochemical signaling genes, as opposed to those in the DYX2 locus (Tables 2 and S5). Consistent with prior associations in the Cleveland SSD cohort reported independently in Stein *et al.* (2014), we found suggestive evidence for association of markers in similar direction of effects in *ANKK1* (e.g., rs4938013, $P = 0.01$; rs7118900, $P = 0.0163$) and *DRD2* (e.g., rs6278, $P = 9.22 \times 10^{-3}$; rs4274224, $P = 0.029$). Additionally, there were nominal associations with markers in *BDNF* (rs2030324, $P = 8.14 \times 10^{-3}$), *PKNOX2* (rs10893365, $P = 0.021$) and *CHRNB2* (rs2072661, $P = 0.039$). Within the DYX2 locus, nominally associated markers were limited to

Table 2: Associations with $P < 0.01$ in meta-analysis of SSD case status in the Cleveland SSD (210 Affected, 257 Unaffected) and Denver SSD (113 Affected, 106 Unaffected) under (a) additive and (b) dominant models

Marker	Chr	BP*	Gene	Allele 1†	Allele 2†	*P*-value	Direction‡	MAF Cleveland	MAF Denver	I^2	Het. *P*-value
(a)											
rs2030324	11	27726915	*BDNF*	T	c	8.14×10^{-3}	++	0.448	0.491	59.6	0.116
(b)											
rs6278	11	113280724	*DRD2*	t	G	9.22×10^{-3}	– –	0.136	0.107	0	0.847

Chr, chromosome; BP, base pair; MAF, minor allele frequency; I^2, Heterogeneity test-statistic; Het. *P*-value, *P*-value for test of heterogeneity.

*Base pair assigned by NCBI Reference Genome GRch37.p13.

†Upper case signifies major allele and lower case minor allele.

‡Direction of effect is in regards to Allele 1. '– –' denotes allele 1 is associated with decreased risk of SSD in both cohorts. '++' denotes allele 1 is associated with increased risk of SSD in both cohorts.

two markers in *FAM65B* (rs4712862, $P = 0.024$; rs640380, $P = 0.037$) and one marker in *KIAA0319* (rs9295626, $P = 0.047$) (Tables 2 and S5, Figure S1).

Similar to SSD case–control status, there were several nominal associations between neurochemical signaling markers and GFTA, which measures articulation skills (Tables 3 and S6). The strongest associations were seen with markers in *CHRNA3*, including rs1051730 ($P = 2.51 \times 10^{-3}$) and rs8034191 ($P = 4.16 \times 10^{-3}$) under a dominant model (Table 3b). Within the DYX2 locus, there were nominal associations throughout *DCDC2* (e.g., rs12192947, $P = 6.87 \times 10^{-3}$) as well as with markers in *KIAA0319*, *NRSN1*, *FAM65B* and a gene-free region located telomeric to *DCDC2* (Tables 3 and S6, Figure S1). None of these DYX2 associations approached the corrected threshold for multiple testing, however.

Overall, the strongest associations observed in the meta-analyses were between a cluster of markers in the 3′ region of *DCDC2* and performance on the PA composite measure, a distinct deficit in children with SSD (e.g., rs7764902, $P = 1.43 \times 10^{-4}$; rs9460973, $P = 1.52 \times 10^{-4}$; rs6456596, $P = 1.56 \times 10^{-4}$; rs9467062, $P = 1.66 \times 10^{-4}$) (Tables 4 and S7, Figure S1). These associations met experiment-wide correction for multiple testing of 1.68×10^{-4}. There were other associations in the DYX2 locus, including markers in the genes *GMNN*, *MRS2*, *FAM65B* and a gene-free region telomeric to *DCDC2*; however, none of these were as strong or consistent as those seen with the markers in the 3′ region of *DCDC2* (Table 4). There were few neurochemical signaling markers showing association with PA, with only one marker in *CHRNB3* (rs6474413, $P = 0.02$) and one marker in *ANKK1* (rs11604671, $P = 0.036$) showing nominal associations (Tables 4 and S7).

Discussion

In this study of the genetics of SSD, we assessed the association of markers within the DYX2 locus and in genes important in neurochemical signaling with SSD and related endophenotypes. In order to increase sample size and to gain insight from more than one study population, we used data from two family-based SSD cohorts to first perform genetic associations in each individual cohort, and then combined these results in a meta-analysis. The strongest, most consistent associations were with markers spanning the 3′ region of *DCDC2* in the DYX2 locus. These associations were consistent across both studies, as indicated by similar direction of effect and non-significant heterogeneity. There were various other associations of markers in the DYX2 locus, notably *GPLD1*, and within multiple previously characterized neurochemical signaling genes, including *ANKK1*, *DRD2*, *CHRNA3* and *BDNF*. In our associations with neurochemical signaling and DYX2 markers, there was a consistent difference in the nature of these associations, as neurochemical signaling markers tended to associate with articulation traits (i.e., case–control status and GFTA), while DYX2 markers associated with phonology (i.e., the PA composite measure). The results of these analyses suggest that the DYX2 gene *DCDC2* contributes to SSD and related neurocognitive domains, as well as provides limited evidence of dopaminergic and nicotinic signaling pathways influencing SSD.

Our DYX2 results centered on the associations between markers in the 3′ region of *DCDC2* and scores on a PA composite measure in children with SSD. *DCDC2* has been associated with both dyslexia and LI, disorders where deficits in phonology play a major role (Pennington & Bishop 2009). Children with SSD often exhibit impairments in phonology in addition to the primary deficits seen in articulation (ASHA Website 2015; Pennington & Bishop 2009). *DCDC2* appears to influence phonological deficits in children with SSD as well as phonological impairments in other neurodevelopmental disorders. Many studies of *DCDC2*, including those from our group, have focused on the complex tandem repeat READ1 (GenBank Accession Number BV677278) and a naturally occurring microdeletion encompassing it, located in intron 2 of *DCDC2* (Ludwig *et al.* 2008, Meng *et al.* 2005; Powers *et al.* 2013). There is evidence suggesting that READ1 may influence the transcriptional activity of *DCDC2* (Meng *et al.* 2011; Powers *et al.* 2013). However, our associations did not center near intron 2 of *DCDC2*. In fact, association results of READ1 in these cohorts did not show evidence of association ($P > 0.05$, data not shown). Instead, we found consistent associations with the 3′ untranslated region (UTR) of *DCDC2*. The 3′ UTR of genes typically contains regulatory

Table 3: Associations with $P < 0.01$ in meta-analysis of GFTA performance in the Cleveland SSD ($n = 187$) and Denver SSD ($n = 175$) under (a) additive and (b) dominant models

Marker	Chr	BP*	Gene	Allele 1†	Allele 2†	*P*-value	Direction‡	MAF Cleveland	MAF Denver	I^2	Het. *P*-value
(a)											
rs1051730	15	78894339	*CHRNA3*	t	C	2.82×10^{-3}	– –	0.324	0.332	0	0.414
rs8034191	15	78806023	*CHRNA3*	T	c	6.14×10^{-3}	++	0.327	0.325	0	0.945
rs9379651	6	24314900	*DCDC2*	a	G	7.18×10^{-3}	++	0.140	0.128	59.1	0.118
rs12192947	6	24318570	*DCDC2*	A	g	7.48×10^{-3}	++	0.150	0.146	0	0.637
rs3789221	6	24189538	*DCDC2*	a	G	8.99×10^{-3}	++	0.125	0.123	75.1	0.0451
(b)											
rs1051730	15	78894339	*CHRNA3*	t	C	2.51×10^{-3}	– –	0.324	0.332	0	0.913
rs8034191	15	78806023	*CHRNA3*	T	c	4.16×10^{-3}	++	0.327	0.325	0	0.377
rs9356927	6	24101847		t	C	4.63×10^{-3}	++	0.093	0.125	67.4	0.0801
rs12192947	6	24318570	*DCDC2*	A	g	6.87×10^{-3}	++	0.150	0.146	0	0.507
rs3789221	6	24189538	*DCDC2*	a	G	7.03×10^{-3}	++	0.125	0.123	73	0.0544

BP, base pair; Chr, chromosome; Het. *P*-value, *P*-value for test of heterogeneity; I^2, Heterogeneity test-statistic; MAF, minor allele frequency.

*Base pair assigned by NCBI Reference Genome GRch37.p13.

†Upper case signifies major allele and lower case minor allele.

‡Direction of effect is in regards to Allele 1. '– –' denotes allele 1 is associated with decreased performance in both cohorts. '++' denotes allele 1 is associated with increased performance in both cohorts.

Table 4: Associations with $P < 0.01$ in meta-analysis of composite PA score in the Cleveland SSD ($n = 96$) and Denver SSD ($n = 128$) under (a) additive and (b) dominant models

Marker	Chr	BP†	Gene	Allele 1‡	Allele 2‡	*P*-value	Direction§	MAF Cleveland	MAF Denver	I^2	Het. *P*-value
(a)											
rs7764902	6	24174834	*DCDC2*	a	T	3.00×10^{-4}	– –	0.022	0.049	0	0.863
rs9460973	6	24175021	*DCDC2*	a	T	3.05×10^{-4}	– –	0.025	0.047	0	0.912
rs6456596	6	24174390	*DCDC2*	a	C	3.26×10^{-4}	– –	0.023	0.047	0	0.897
rs9467062	6	24173382	*DCDC2*	T	c	3.32×10^{-4}	++	0.025	0.048	0	0.922
rs554400	6	23998157		T	c	8.03×10^{-3}	++	0.105	0.063	18	0.269
rs9467046	6	24120269		A	g	9.06×10^{-3}	– –	0.464	0.446	0	0.387
(b)											
rs7764902*	6	24174834	*DCDC2*	a	T	1.43×10^{-4}	– –	0.022	0.049	0	0.965
rs9460973*	6	24175021	*DCDC2*	a	T	1.52×10^{-4}	– –	0.025	0.047	0	0.927
rs6456596*	6	24174390	*DCDC2*	a	C	1.56×10^{-4}	– –	0.023	0.047	0	0.932
rs9467062*	6	24173382	*DCDC2*	T	c	1.66×10^{-4}	++	0.025	0.048	0	0.917
rs9467046	6	24120269		A	g	1.29×10^{-3}	– –	0.464	0.446	0	0.350
rs554400	6	23998157		T	c	5.57×10^{-3}	++	0.105	0.063	35	0.215

BP, base pair; Chr, chromosome; Het. *P*-value, *P*-value for test of heterogeneity; I^2, Heterogeneity test-statistic; MAF, minor allele frequency.

*Survives correction for multiple testing ($P < 1.68\,0 \times 10^{-4}$).

†Base pair assigned by NCBI Reference Genome GRch37.p13.

‡Upper case signifies major allele and lower case minor allele.

§Direction of effect is in regards to Allele 1. '– –' denotes allele 1 is associated with decreased performance in both cohorts. '++' denotes allele 1 is associated with increased performance in both cohorts.

sequence with microRNA (miRNA) binding sites that influence gene and protein expression. Here, the associations we observe of the 3′ UTR in *DCDC2*, in addition to the literature implicating READ1 in *DCDC2*, suggests that pre- and post-transcriptional regulation of *DCDC2* expression levels may contribute to phonology and language traits. Biologically, the implications of changes in *DCDC2* expression still remain to be elucidated, although knockout and knockdown experiments suggest that neuronal migration, ciliary function and *N*-methyl-D-aspartate receptor (NMDAR) activity are strong candidates (Che *et al.* 2013; Ivliev *et al.* 2012; Massinen *et al.* 2011; Meng *et al.* 2005).

In our analyses, we found evidence suggesting that markers in dopaminergic-related genes, *ANKK1* and *DRD2*, as

well as nicotinic-related genes, *CHRNA3* and *BDNF*, influence SSD and articulation skills. Our groups and others have recently shown the association of markers within *ANKK1* and *DRD2* with endophenotypes of communication and language skills (Beaver *et al.* 2010; Berman & Noble 1995; Eicher *et al.* 2013; Stein *et al.* 2014; Wong *et al.* 2013). *ANKK1* and *DRD2* have previously been implicated in many neurocognitive processes, including working memory, reinforcement learning and executive function, as well as neuropsychiatric disorders, including schizophrenia and bipolar disorder (Bertolino *et al.* 2010; Bolton *et al.* 2010; London *et al.* 2009; McAllister *et al.* 2008; Ripke *et al.* 2014; Yang *et al.* 2007). Variation in *DRD2* and other dopaminergic signaling genes may alter these processes, contributing to the risk of impairment in articulation and speech skills. In addition to dopaminergic factors, the results also showed suggestive evidence for *CHRNA3* and *BDNF*, which encode components of nicotinic pathways. *CHRNA3* is a neuronal nicotinic acetylcholine receptor, most widely known for its relationship with nicotine dependence (Berrettini & Doyle 2012). Its relationship with language and communication traits has been far less studied, although markers implicated in nicotine dependence, including ones examined in this study, have been previously associated with neurocognitive performance (Winterer *et al.* 2010, Zhang *et al.* 2010). *BDNF* is one of the most widely studied genes in neuropsychiatric and neurobehavioral traits, including Alzheimer's disease, depression and bipolar disorder. From these studies, evidence has emerged that markers in *BDNF* are also associated with neurocognitive traits (Bath & Lee 2006; Goldberg & Weinberger 2004). Functionally, studies have demonstrated that BDNF can modulate nicotinic receptor activity (Fernandes *et al.* 2008; Massey *et al.* 2006; Zhou *et al* 2004). Our results, in addition to these studies, suggest that changes in nicotinic signaling may modulate articulation and other communicative processes, and thus contribute to SSD.

Here, we implicate genetic factors in SSD that had previously been associated with related neurodevelopmental disorders, dyslexia and LI. The associations we observe between *ANKK1*, *DRD2*, and *DCDC2* in SSD and related endophenotypes add to the growing body of evidence implicating genes in multiple related traits. As opposed to specific associations between genes and individual disorders, it appears that genes are more generally associated with language and phonological processes. The complex, polygenic nature of these neurodevelopmental disorders suggests that many interacting genetic and environmental factors account for risk of impairment. Therefore, in order to discern how genetic predisposition leads to clinical presentation, future studies need to thoroughly examine the combinatorial effects of genes and environment on the development of communication, language and other neurobehavioral traits.

This study is subject to several limitations. First, the size of the individual Cleveland SSD and Denver SSD cohorts is relatively small. Smaller sample size can reduce statistical power and hamper our ability to detect genetic associations. However, the use of a meta-analysis strategy allowed us to increase our statistical power and gain confidence in the fidelity of our associations. Second, subjects in both the Cleveland SSD and Denver SSD cohorts completed a wide variety of neurocognitive assessments. The number of measures shared between the two cohorts was limited; therefore, some key neurobehavioral domains such as reading and working memory could not be specifically examined in our meta-analyses. Third, we selected specific markers in neurochemical signaling genes of interest as well as a candidate locus of interest, DYX2. However, our DYX2 marker panel was constructed to tag as much common variation in the locus as possible and to allow for an unbiased association scan of the locus. In the future, it would be of interest to similarly saturate these neurochemical signaling genes to fine map putative associations. Fourth, the minor allele frequencies of associated markers in *DCDC2*, ranging from 2.2% to 2.5% in Cleveland SSD and 4.7% to 4.9% in Denver SSD, were relatively low (Table 4). However, the consistency of associations with multiple markers in this region, suggestive associations of other markers in the locus (e.g., rs3789221and rs9460974 with GFTA in Table 3), and the same direction of effect in both cohorts, increases the confidence we have in these associations. In addition, some of the individual SNP-phenotype associations in the meta-analysis showed significant heterogeneity based on the I^2 parameter. However, I^2 should be interpreted with caution when using only two small studies as we have here (Thorlund *et al.* 2012). Furthermore, our most significant findings did not show heterogeneity, showing these results are indeed robust.

In conclusion, we associated markers in the 3′ UTR of *DCDC2* within the DYX2 locus, as well as provide suggestive evidence in dopaminergic- (*ANKK1* and *DRD2*) and nicotinic-related genes (*CHRNA3* and *BDNF*), with SSD, and related neurocognitive domains. We expand the role of the DYX2 locus, especially *DCDC2*, to also influence phonological skills in children with SSD. These results also further implicate genetic variation in important neurochemical signaling pathways in higher-order cognitive and communication skills, and their respective disorders. Future studies should further examine shared genetic risk factors among communication and language disorders, as well as elucidate the common mechanisms among these disorders. Discerning the underlying genetic and neurological mechanisms, whether they reside in neurochemical signaling, neuronal migration, and/or ciliary function or a combination of these pathways, is vital for improving treatment of affected individuals.

References

Amos, C.I., Wu, X., Broderick, P., GOrloc, I.P., Gu, J., Eisen, T., Dong, Q., Zhang, Q., Gu, X., Vijayakrishnan, J., Sullivan, K., Matakidou, A., Wang, Y., Mills, G., Doheny, K., Tsai, Y.Y., Chen, W.V., Shete, S., Spitz, M.R. & Houlston, R.S. (2008) Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. *Nat Genet* **40**, 616–622.

ASHA Website. (2015) URL http://www.asha.org/public/speech/disorders/SpeechSoundDisorders/

Bath, K.G. & Lee, F.S. (2006) Variant BDNF (Val66Met) on brain structure and function. *Cogn Affect Behav Neurosci* **6**, 79–85.

Beaver, K.M., Delisi, M., Vaughn, M.G. & Wright, J.P. (2010) Association between the A1 allele of the DRD2 gene and reduced verbal abilities in adolescence and early adulthood. *J Neural Transm* **117**, 827–830.

Becker, J., Czamara, D., Scerri, T.S. *et al.* (2014) Genetic analysis of dyslexia candidate genes in the European cross-linguistic NeuroDys cohort. *Eur J Hum Genet* **22**, 675–680.

Berman, S.M. & Noble, E.P. (1995) Reduced visuospatial performance in children with the D2 dopamine receptor A1 allele. *Behav Genet* **25**, 45–58.

Bernthal, J.E., Bankson, N.W. & Flipsen, P. (2013) *Articulation and Phonological Disorders: Speech Sound Disorders in Children*, 7th edn. Upper Saddle River, NJ, Pearson Education Inc.

Berrettini, W.H. & Doyle, G.A. (2012) The CHRNA5-A3-B4 gene cluster in nicotine addiction. *Mol Psychiatry* **17**, 856–866.

Bertolino, A., Taurisano, P., Pisciotta, N.M., Blasi, G., Fazio, L., Romano, R., GElao, B., LoBianco, L., Lozupone, M., DiGiorgio, A., Caforio, G., Sambataro, F., Niccoli-Asabella, A., Papp, A., Ursini, G., Sinibaldi, L., Popolizio, T., Sadee, W. & Rubini, G. (2010) Genetically determined measures of striatal D2 signaling predict prefrontal activity during working memory performance. *PLoS One* **5**, e9348.

Beuten, J., Ma, J.Z., Payne, T.J., Dupont, R.T., Quezada, P., Huang, W., Crews, K.M. & Li, M.D. (2005) Significant association of BDNF haplotypes in European-American male smokers but not in European-American female of African-American smokers. *Am J Med Genet B Neuropsychiatr Genet* **139B**, 73–80.

Beuten, J., Ma, J.Z., Payne, T.J., Dupont, R.T., Lou, X.Y., Crews, K.M., Elston, R.C. & Li, M.D. (2007) Association of specific haplotypes of neurotropic tyrosine kinase receptor 2 gene (NTRK2) with vulnerability to nicotine dependence in African-Americans and European-Americans. *Biol Psychiatry* **61**, 48–55.

Bird, J. & Bishop, D. (1992) Perception and awareness of phonemes in phonological impaired children. *Eur J Disord Commun* **27**, 289–311.

Bolton, J.L., Marioni, R.E., Deary, I.J., Harris, S.E., Stewart, M.C., Murray, G.D., Fowkes, F.G. & Price, J.F. (2010) Association between polymorphisms of the dopamine receptor D2 and catechol-o-methyl transferase genes and cognitive function. *Behav Genet* **40**, 630–638.

Che, A., Girgenti, M.J. & LoTurco, J. (2013) The dyslexia-associated gene Dcdc2 is required for spike-timing precision in mouse neocortex. *Biol Psychiatry* **76**, 387–396.

Chen, M.H. & Yang, Q. (2010) GWAF: an R package for genome-wide association analyses with family data. *Bioinformatics* **26**, 580–581.

Chen, X., Cho, K., Singer, B.H. & Zhang, H. (2011) The nuclear transcription factor PKNOX2 is a candidate gene for substance dependence in European-origin women. *PLoS One* **6**, e16002.

Cope, N.A., Harold, D., Hill, G., Moskvina, V., Stevenson, J., Holmans, P., Owen, M.J., O'Donovan, M.C. & Williams, J. (2005) Strong evidence that KIAA0319 on chromosome 6p is a susceptibility gene for developmental dyslexia. *Am J Hum Genet* **76**, 581–591.

Eicher, J.D., Powers, N.R., Cho, K., Miller, L.L., Mueller, K.L., Ring, S.M., Tomblin, J.B. & Gruen, J.R. (2013) Associations of prenatal nicotine exposure and the dopamine related genes ANKK1 and DRD2 to verbal language. *PLoS One* **8**, e63762.

Eicher, J.D., Powers, N.R., Miller, L.L., Mueller, K.L., Mascheretti, S., Marino, C., Willcutt, E.G., DeFries, J.C., Olson, R.K., Smith, S.D., Pennington, B.F., Tomblin, J.B., Ring, S.M. & Gruen, J.R. (2014) Characterization of the DYX2 locus on chromosome 6p22 with reading disability, language impairment, and IQ. *Hum Genet* **133**, 869–881.

Fernandes, C.C., Pinto-Duarte, A., Ribeiro, J.A. & Sebastiao, A.M. (2008) Postsynaptic action of brain-derived neurotrophic factor attenuates alpha7 nicotinic acetylcholine receptor-mediated responses in hippocampal interneurons. *J Neurosci* **28**, 5611–5618.

Gelernter, J., Yu, Y., Weiss, R., Brady, K., Panhuysen, C., Tang, B.Z., Kranzler, H.R. & Farrer, L. (2006) Haplotype spanning TTC12 and ANKK1, flanked by the DRD2 and NCAM1 loci, is strongly associated to nicotine dependence in two distinct American populations. *Hum Mol Genet* **15**, 3498–3507.

Goldberg, T.E. & Weinberger, D.R. (2004) Genes and the parsing of cognitive processes. *Trends Cogn Sci* **8**, 325–335.

Goldman, R. & Fristoe, M. (1986) *The Goldman-Fristoe Test of Articulation.* American Guidance Service, Circle Pines, MN.

Han, B., Kang, H.M., Seo, M.S., Zaitlen, N. & Eskin, E. (2008) Efficient association study design via power-optimized tag SNP selection. *Ann Hum Genet* **72** (Pt 6), 834–847.

Huang, W., Payne, T.J., Ma, J.Z., Beuten, J., Dupont, R.T., Inohara, N. & Li, M.D. (2009) Significant association of ANKK1 and detection of a functional polymorphism with nicotine dependence in an African-American sample. *Neuropsychopharmacology* **34**, 319–330.

Ivliev, A.E., 't Hoen, P.A., van Roon-Mom, W.M., Peters, D.J. & Sergeeva, M.G. (2012) Exploring the transcriptome of ciliated cells using in silico dissection of human tissues. *PLoS One* **7**, e35618.

Konig, I.R., Schumacher, J., Hoffmann, P., Kleensang, A., Ludwig, K.U., Grimm, T., Neuhoff, N., Preis, M., Roeske, D., Warnke, A., Propping, P., Remschmidt, H., Nothen, M.M., Ziegler, A., Muller-Myhsok, B. & Schulte-Korne, G. (2011) Mapping for dyslexia and related cognitive trait loci provides strong evidence for further risk genes on chromosome 6p21. *Am J Med Genet B Neuropsychiatr Genet* **156B**, 36–43.

Lewis, B.A., Freebairn, L.A. & Taylor, H.G. (2000) Follow-up of children with early expressive phonology disorders. *J Learn Disabil* **33**, 433–444.

Lewis, B.A., Shriberg, L.D., Freebaim, L.A., Hansen, A.J., Stein, C.M., Taylor, H.G. & Iyengar, S.K. (2006) The genetic basis of speech sound disorders: evidence from spoken and written language. *J Speech Lang Hear Res* **49**, 1294–1312.

Lewis, B.A., Avrich, A.A., Freebairn, L.A., Hansen, A.J., Sucheston, L.E., Kuo, I., Taylor, H.G., Iyengar, S.K. & Stein, C.M. (2011) Literacy outcomes of children with early childhood speech sound disorders: impact of endophenotypes. *J Speech Lang Hear Res* **54**, 1626–1643.

Li, M.D., Beuten, J., Ma, J.Z., Payne, T.J., Lou, X.Y., Garcia, V., Duenes, A.S., Crews, K.M. & Elston, R.C. (2005) Ethnic- and gender-specific association of the nicotinic acetylcholine receptor alpha4 subunit gene (CHRNA4) with nicotine dependence. *Hum Mol Genet* **14**, 1211–1219.

Liu, J.Z., Tozzi, F., Waterworth, D.M. *et al.* (2010) Meta-analysis and imputation refines the association of 15q25 with smoking quantity. *Nat Genet* **42**, 436–440.

London, E.D., Berman, S.M., Mohammadian, P., Ritchie, R., Mandelkern, M.A., Susselman, M.K., Schlagenhauf, F. & Noble, E.P. (2009) Effect of the TaqIA polymorphism on ethanol response in the brain. *Psychiatry Res* **174**, 163–170.

Ludwig, K.U., Schumacher, J., Schulte-Korne, G., Konig, I.R., Warnke, A., Plume, E., ANthoni, H., Peyrard-Janvid, M., Meng, H., Ziegler, A., Remschmidt, H., Kere, J., Gruen, J.R., Muller-Myshok, B., Nothen, M.M. & Hoffmann, P. (2008) Investigation of the DCDC2 intron 2 deletion/compound short tandem repeat polymorphism in a large German dyslexia sample. *Psychiatr Genet* **18**, 310–312.

Mascheretti, S., Riva, V., Giorda, R., Beri, S., Lanzoni, L.F., Cellino, M.E. & Marino, C. (2014) KIAA0319 and ROBO1: evidence on association with reading and pleiotropic effects on language and mathematics abilities in developmental dyslexia. *J Hum Genet* **59**, 189–197.

Massey, K.A., Zago, W.M. & Berg, D.K. (2006) BDNF up-regulates alpha7 nicotinic receptor levels on subpopulations of hippocampal interneurons. *Mol Cell Neurosci* **33**, 381–388.

Massinen, S., Hokkanen, M.E., Matsson, H., Tammimies, K., Tapia-Páez, I., Dahlström-Heuser, V., Kuja-Panula, J., Burghoorn, J., Jeppsson, K.E., Swoboda, P., Peyrard-Janvid, M., Toftgård, R., Castrén, E. & Kere, J. (2011) Increased expression of the dyslexia candidate gene DCDC2 affects length and signaling of primary cilia in neurons. *PLoS One* **6**, e20580.

McAllister, T.W., Flashman, L.A., Harker Rhodes, C., Tyler, A.L., Moore, J.H., Saykin, A.J., McDonald, B.C., Tosteson, T.D. & Tsongalis, G.J. (2008) Single nucleotide polymorphisms in ANKK1 and the dopamine D2 receptor gene affect cognitive outcome shortly after traumatic brain injury: a replication and extension study. *Brain Inj* **22**, 705–714.

Meng, H., Smith, S.D., Hager, K., Held, M., Liu, J., Olson, R.K., Pennington, B.F., DeFries, J.C., Gelernter, J., O'Reilly-Pol, T., Somlo, S., Skudlarski, P., Shaywitz, S.E., Shaywitz, B.A., Marchione, K., Wang, Y., Paramasivam, M., LoTurco, J.J., Page, G.P. & Gruen, J.R. (2005) DCDC2 is associated with reading disability and modulates neuronal development in the brain. *Proc Natl Acad Sci U S A* **102**, 17053–17058.

Meng, H., Powers, N.R., Tang, L., Cope, N.A., Zhang, P.X., Fuleihan, R., Gibson, C., Page, G.P. & Gruen, J.R. (2011) A dyslexia-associated variant in DCDC2 changes gene expression. *Behav Genet* **41**, 58–66.

Newbury, D.F. & Monaco, A.P. (2010) Genetic advances in the study of speech and language disorders. *Neuron* **68**, 309–320.

Newbury, D.F., Paracchini, S., Scerri, T.S., Winchester, L., Addis, L., Richarson, A.J., Walter, J., Stein, J.F., Talcott, J.B. & Monaco, A.P. (2011) Investigation of dyslexia and SLI risk variants in reading- and language-impaired subjects. *Behav Genet* **41**, 90–104.

Nyholt, D.R. (2004) A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. *Am J Hum Genet* **74**, 765–769.

Pennington, B.F. & Bishop, D.V.M. (2009) Relations among speech, language, and reading disorders. *Annu Rev Psychol* **60**, 283–306.

Peter, B., Matsushita, M. & Raskind, W.H. (2012) Motor sequencing deficit as an endophenotype of speech sound disorder: a genome-wide linkage analysis in a multigenerational family. *Psychiatr Genet* **22**, 226–234.

Pinel, P., Fauchereau, F., Moreno, A., Barbot, A., Lathrop, M., Zelenika, D., Le Bihan, D., Poline, J.B., Bourgeron, T. & Dehaene, S. (2012) Genetics variants of *FOXP2* and *KIAA0319/TTRAP/THEM2* locus are associated with altered brain activation in distinct language-related regions. *J Neurosci* **32**, 817–825.

Plomin, R., Turic, D.M., Hill, L., Turic, D.E., Stephens, M., Williams, J., Owen, M.J. & O'Donovan, M.C. (2004) A functional polymorphism in the succinate-semialdehyde dehydrogenase (aldehyde dehydrogenase 5 family member A1) gene is associated with cognitive ability. *Mol Psychiatry* **9**, 582–586.

Powers, N.R., Eicher, J.D., Butter, F., Kong, Y., Miller, L.L., Ring, S.M., Mann, M. & Gruen, J.R. (2013) Alleles of a polymorphic ETV6 binding site in DCDC2 confer risk of reading and language impairment. *Am J Hum Genet* **93**, 19–28.

Pugh, K.R., Frost, S.J., Rothman, D.L., Hoeft, F., Del Tufo, S.N., Mason, G.F., Molfese, P.J., Mencl, W.E., Grigorenko, E.L., Landi, N., Preston, J.L., Jacobsen, L., Seindenberg, M.S. & Fulbright, R.K. (2014) Glutamate and choline levels predict individual differences in reading ability in emergent readers. *J Neurosci* **34**, 4082–4089.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J. & Sham, P.C. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559–575.

Ripke, S.R. & Schizophrenia Working Group of the Psychiatric Genomics Consortium (2014) Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421–427.

Saccone, S.F., Hinrichs, A.L., Saccone, N.L. *et al.* (2007) Cholinergic nicotinic receptor genes implicated in a nicotine dependence association study targeting 348 candidate genes with 3713 SNPs. *Hum Mol Genet* **16**, 36–49.

Saccone, N.L., Saccone, S.F., Hinrichs, A.L., Stitzel, J.A., Duan, W., Pergadia, M.L., Agrawal, A., Breslau, N., Grucza, R.A., Hatsukami, D., Johnson, E.O., Madden, P.A., Swan, G.E., Wang, J.C., Goate, A.M., Rice, J.P. & Bierut, L.J. (2009) Multiple distinct risk loci for nicotine dependence identified by dense coverage of the complete family of nicotinic receptor subunit (CHRN) genes. *Am J Med Genet B Neuropsychiatr Genet* **150B**, 453–466.

Saccone, N.L., Schwantes-An, T.H., Wang, J.C., Grucza, R.A., Breslau, N., Hatsukami, D., Johnson, E.O., Rice, J.P., Goate, A.M. & Bierut, L.J. (2010) Multiple cholinergic nicotinic receptor genes affect nicotine dependence risk in African and European Americans. *Genes Brain Behav* **9**, 741–750.

Scerri, T.S., Morris, A.P., Buckingham, L.L., Newbury, D.F., Miller, L.L., Monaco, A.P., Bishop, D.V. & Paracchini, S. (2011) DCDC2, KIAA0319 and CMIP are associated with reading-related traits. *Biol Psychiatry* **70**, 237–245.

Schumacher, J., Anthoni, H., Dahdouh, F., Konig, I.R., Hillmer, A.M., Kluck, N., Manthey, M., Plume, E., Warnke, A., Remschimdt, H., Hulsmann, J., Cichon, S., Lindgren, C.M., Propping, P., Zucchelli, M., Ziegler, A., Peyrard-Janvid, M., Schulte-Korne, H., Nothen, M.M. & Kere, J. (2006) Strong evidence of DCDC2 as a susceptibility gene for dyslexia. *Am J Hum Genet* **78**, 52–62.

Smith, S.D., Pennington, B.F., Boada, R. & Shriberg, L.D. (2005) Linkage of speech sound disorder to reading disability loci. *J Child Psychol Psychiatry* **46**, 1057–1066.

Stein, C.M., Schick, J.H., Gerry-Taylor, H., Shriberg, L.D., Millard, C., Kundtz-Kluge, A., Russo, K., Minich, N., Hansen, A., Freebaim, L.A., Elston, R.C., Lewis, B.A. & Iyengar, S.K. (2004) Pleiotropic effects of a chromosome 3 locus on speech-sound disorder and reading. *Am J Hum Genet* **74**, 283–297.

Stein, C.M., Lu, Q., Elston, R.C., Freebaim, L.A., Hansen, A.J., Shriberg, L.D., Taylor, H.G., Lewis, B.A. & Iyengar, S.K. (2011) Heritability estimation for speech-sound traits with developmental trajectories. *Behav Genet* **41**, 184–191.

Stein, C.M., Truitt, B., Deng, F., Ciesla, A.A., Qiu, F., Joseph, P., Raghavendra, R., Fondran, J., Igo, R.P. Jr., Tag, J., Freebairn, L., Taylor, H.G., Lewis, B.A. & Iyengar, S.K. (2014) Association between AVPR1A, DRD2, and ASPM and endophenotypes of communication disorders. *Psychiatr Genet* **24**, 191–200.

Thorlund, K., Imberger, G., Johnston, B.C., Walsh, M., Awad, T., Thabane, L., Gluud, C., Devereaux, P.J. & Wetterslev, J. (2012) Evolution of heterogeneity (I^2) estimates of their 95% confidence intervals in large meta-analyses. *PLoS One* **7**, e39471.

Wagner, R.K., Torgesen, J.K. & Rashotte, C.A. (1999) *Comprehensive Test of Phonological Process*. Pro-ED, Inc, Austin, TX.

Wessel, J., McDonald, S.M., Hinds, D.A., Stokowski, R.P., Javitz, H.S., Kennemer, M., Krasnow, R., Dirks, W., Hardin, J., Pitts, S.L., Michel, M., Jack, L., Ballinger, D.G., McClure, J.B., Swan, G.E. & Bergen, A.W. (2010) Resequencing of nicotinic acetylcholine receptor genes and association of common and rare variants with the Fagerstrom test for nicotine dependence. *Neuropsychopharmacology* **35**, 2392–2402.

Willer, C.J., Li, Y. & Abecassis, G.R. (2010) METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191.

Winterer, G., Mittelstrass, K., Giegling, I. *et al.* (2010) Risk gene variants for nicotine dependence in the CHRNA5-CHRNA3-CHRNB4 cluster are associated with cognitive performance. *Am J Med Genet B Neuropsychiatr Genet* **153B**, 1448–1458.

Wong, P.C., Ettinger, M. & Zheng, J. (2013) Linguistic grammar learning and DRD2-TAQ-IA polymorphism. *PLoS One* **8**, e64983.

Yang, B.Z., Kranzler, H.R., Zhao, H., Gruen, J.R., Luo, X. & Gelernter, J. (2007) Association of haplotypic variants in DRD2, ANKK1, TTC12, and NCAM1 to alcohol dependence in independent case control and family samples. *Hum Mol Genet* **16**, 2844–2853.

Zhang, H., Kranzler, H.R., Poling, J. & Gelernter, J. (2010) Variation in the nicotinic acetylcholine receptor gene cluster CHRNA5-CHRNA3-CHRNB4 and its interaction with recent tobacco use influence cognitive flexibility. *Neuropsychopharmacology* **35**, 2211–2224.

Zhou, X., Nai, Q., Chen, M., Dittus, J.D., Howard, M.J. & Margiotta, J.F. (2004) Brain-derived neurotrophic factor and trkB signaling in parasympathetic neurons : relevance to regulating alpha7-containing nicotinic receptors and synaptic function. *J Neurosci* **24**, 4340–4350.

Zou, L., Chen, W., Shao, S., Sun, Z., Zhong, R., Shi, J., Miao, X. & Song, R. (2012) Genetic variant in KIAA0319, but not in DYX1C1, is associated with risk of dyslexia: an integrated meta-analysis. *Am J Med Genet B Neuropsychiatr Genet* **159B**, 970–976.

Acknowledgments

We wish to thank all the families and subjects for their participation in these studies. Additional, we would like to acknowledge Irina Tikhonova, Anna Rogers and the Yale Center for Genome Analysis for genotyping services. This research was funded by the National Institutes of Health [Grant ref: R01 NS043550 (J.D.E., N.R.P., J.R.G.), P50 HD027802 (S.D.S., B.F.P., J.R.G.), F31 DC012270 (J.D.E.), DC000528 (B.A.L., S.K.I., C.M.S., A.A.C., F.D.), DC012380 (S.K.I., B.A.L., C.M.S.) and R01 HD049027 (R.B., S.D.S., B.F.P.)] and the Manton Family Foundation (J.D.E., N.R.P., J.R.G.).

Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's web-site:

Table S1: Associations of DYX2 and neurochemical signaling genetic markers with SSD case status in the individual Cleveland SSD (Table S1a and Table S1c) and Denver SSD (Table S1b and Table S1d) cohorts.

Table S2: Associations of DYX2 and neurochemical signaling genetic markers with GFTA in the individual Cleveland SSD (Table S2a and Table S2c) and Denver SSD (Table S2b and Table S2d) cohorts.

Table S3: Associations of DYX2 and neurochemical signaling genetic markers with PA in the individual Cleveland SSD (Table S3a and Table S3c) and Denver SSD (Table S3b and Table S3d) cohorts.

Table S4: (a) Genotype breakdown of markers with associations $P < 0.05$ in meta-analysis of SSD case status under an additive model. (b) Genotype breakdown of markers with associations $P < 0.05$ in meta-analysis of SSD case status under a dominant model.

Table S5: (a) Associations with $P < 0.05$ in meta-analysis of SSD case status in the Cleveland SSD (210 Affected, 257 Unaffected) and Denver SSD (113 Affected, 106 Unaffected) under an additive model. (b) Associations with $P < 0.05$ in meta-analysis of SSD case status in the Cleveland SSD (210 Affected, 257 Unaffected) and Denver SSD (113 Affected, 106 Unaffected) under a dominant model.

Table S6: (a) Associations with $P < 0.05$ in meta-analysis of GFTA performance in the Cleveland SSD ($n = 187$) and Denver SSD ($n = 175$) under an additive model. (b) Associations with $P < 0.05$ in meta-analysis of GFTA performance in the Cleveland SSD ($n = 187$) and Denver SSD ($n = 175$) under a dominant model.

Table S7: (a) Associations with $P < 0.05$ in meta-analysis of composite PA score in the Cleveland SSD ($n = 96$) and Denver SSD ($n = 128$) under an additive model. (b) Associations with $P < 0.05$ in meta-analysis of composite PA score in the Cleveland SSD ($n = 96$) and Denver SSD ($n = 128$) under a dominant model.

Figure S1: DYX2 association results on case-control status. DYX2 association results on GFTA performance. DYX2 association results on PA performance.