Published in final edited form as:

Neurobiol Dis. 2014 September ; 69: 23–31. doi:10.1016/j.nbd.2014.04.019.

A study of the role of the *FOXP2* and *CNTNAP2* genes in persistent developmental stuttering

Tae-Un Han¹, John Park², Carlos F. Domingues^{1,3}, Danilo Moretti-Ferreira³, Emily Paris¹, Eduardo Sainz¹, Joanne Guiterrez¹, and Dennis Drayna^{1,*}

¹ National Institute on Deafness and Other Communication Disorders, NIH, Bethesda, MD 20892, USA

² National Institute of Neurological Disorders and Stroke, NIH, Bethesda, MD 20892, USA

³ Department of Genetics, Institute of Bioscience São Paulo State University, Botucatu, São Paulo 18618-000, Brazil

Abstract

A number of speech disorders including stuttering have been shown to have important genetic contributions, as indicated by high heritability estimates from twin and other studies. We studied the potential contribution to stuttering from variants in the *FOXP2* gene, which have previously been associated with developmental verbal dyspraxia, and from variants in the *CNTNAP2* gene, which have been associated with specific language impairment (SLI). DNA sequence analysis of these two genes in a group of 602 unrelated cases, all with familial persistent developmental stuttering, revealed no excess of potentially deleterious coding sequence variants in the cases compared to a matched group of 487 well characterized neurologically normal controls. This was compared to the distribution of variants in the *GNPTAB*, *GNPTG*, and *NAGPA* genes which have previously been associated with persistent stuttering. Using an expanded subject data set, we again found that *NAGPA* showed significantly different mutation frequencies in North Americans of European descent ($p = 0.0091$) and a significant difference existed in the mutation frequency of *GNPTAB* in Brazilians ($p = 0.00050$). No significant differences in mutation frequency in the *FOXP2* and *CNTNAP2* genes were observed between cases and controls. To examine the pattern of expression of these five genes in human brain, real time quantitative reverse transcription PCR was performed on RNA purified from 27 different human brain regions. The expression patterns of *FOXP2* and *CNTNAP2* were generally different from those of *GNPTAB*, *GNPTG* and *NAGPA* in terms of relatively lower expression in cerebellum. This study provides an improved estimate of the contribution of mutations in the *GNPTAB*, *GNPTG* and *NAGPA* to persistent stuttering, and suggests that variants in the *FOXP2* and *CNTNAP2* are not involved in the genesis of familial

* To whom correspondence should be addressed at: NIDCD/National Institutes of Health 5 Research Court, Room 2B-46 Rockville, MD 20850 USA Tel. 301-402-4930 drayna@nidcd.nih.gov.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Author Contributions

Conceived and designed the experiments: TUH, DD. Performed experiments: all authors, Analyzed the data: TUH. Wrote the paper: TUH, DD.

persistent stuttering. This, together with the different brain expression patterns of *GNPTAB*, *GNPTG*, and *NAGPA* compared to that of *FOXP2* and *CNTNAP2*, suggests that the genetic neuropathological origins of stuttering differ from those of verbal dyspraxia and SLI.

Introduction

Developmental stuttering is a common, highly heritable speech disorder of children that can persist into adulthood (Gordon, 2002). Genetic linkage studies in families have identified a number of loci for persistent stuttering (Raza et al., 2012; Raza et al., 2010; Riaz et al., 2005), which in turn have led to the discovery of mutations in the *GNPTAB*, *GNPTG*, and *NAGPA* genes that are associated with persistent stuttering in multiple populations (Kang et al., 2010). Developmental verbal dyspraxia is another speech disorder that can be highly heritable, and mutations in the *FOXP2* gene have been associated with familial and sporadic cases (Fisher et al., 1998; Lai et al., 2001; MacDermot et al., 2005). Specific language impairment (SLI) is a third heritable communication disorder, characterized by impairment in the acquisition of speech in otherwise normal children (Bishop, 2001). Variants in the *CNTNAP2* gene have been associated with SLI in both family and population studies (Newbury et al., 2010). While stuttering, verbal dyspraxia, and SLI are traditionally viewed as three different disorders, they have several features in common. All three are highly heritable and are more common in males than in females. Several speech pathologies, such as word repetition and problems generating the desired speech sounds and syllables, are seen in both stuttering and dyspraxia (Kent, 2000). Humanization of the *FOXP2* gene in mouse affects cortico-basal ganglia, parts of the human brain proposed to be involved in stuttering (Enard et al., 2009). Finally, *FOXP2* encodes a transcription factor that has been shown to control the expression of the *CNTNAP2* gene (Vernes et al., 2008).

Identifiable mutations in the currently known genes explain only a fraction of cases for these disorders. Given the large fraction of cases that remain unexplained by identified mutations, it is not clear how much overlap in genetic factors exists for these three disorders. We sought to address this question for stuttering using two approaches. First, we searched for mutations in the *FOXP2* and *CNTNAP2* genes in a group of individuals with familial persistent stuttering, and we compared the rate of mutations and the allele frequencies of common variants in these individuals to that in a group of neurologically normal controls. Second, we compared the expression levels of the *FOXP2* and *CNTNAP2* genes in different regions of human brains of neurologically normal adults, and to the expression levels of the *GNPTAB*, *GNPTG*, and *NAGPA* genes in the same brains.

Materials and Methods

DNA samples and sequencing

Subjects with persistent developmental stuttering were enrolled with written informed consent under NIH protocol #97-DC-0057. Subjects with neurological abnormalities or speech disorders other than stuttering were excluded. All study subjects displayed stuttering beyond the age of 8 for a duration of 6 months or more, and a family history of stuttering in at least one family member (second degree relative or closer) documented by subject

interview. All subjects were unrelated. Case subjects consisted of North American therapy clients ascertained at the Hollins Communications Research Institute (HCRI), Drayna et al. 1999) and single probands from a group of North American families recruited for previous linkage studies (NAF, Shugart et al. 2004). Brazilian case subjects were clients of a university-based therapy program (BRMII), and single probands from a group of Brazilian families recruited for a previous genetic study (BRPD, Canhetti-Oliveira & Richieri-Costa, 2006). Control subjects consisted of documented neurologically normal Caucasian individuals (NDPT 006, 020, 023; Coriell Cell Repository, <http://ccr.coriell.org/Sections/Collections/NINDS/DNAPanel.aspx?PgId=195&coll=ND>) and self-reported neurologically normal Brazilian subjects (BRCO) collected from the state of São Paulo through university-based speech pathology clinics (Canhetti-Oliveira and Richieri-Costa, 2006) and community ascertainment. Stuttering diagnosis was performed using the Stuttering Severity Index, 3rd Edition (SSI-3), and as previously described (Kang et al., 2010). For statistical analyses, HCRI subjects were limited to Caucasians (294 individuals) in order to match the neurologically normal Caucasian controls. Among the 294 Caucasian HCRI individuals, 196 individuals have been used for the mutation analysis of *GNPTAB*, *GNPTG* and *NAGPA* in a previous report of stuttering-associated genetic variants (Kang et al., 2010). DNA sequencing was performed on genomic DNA purified from venous blood using of standard Sanger sequencing methods. Briefly, exons and immediately flanking intronic regions, 5' upstream regions, and 3' downstream regions were PCR amplified from genomic DNA. Primer sequences for the *GNPTAB*, *GNPTG*, and *NAGPA* genes were previously published (Supplementary Appendix, Kang et al, N Engl J Med (2010) 362:677-685). The sequence of primers for *CNTNAP2* and *FOXP2* are shown in Supplementary Table 1. PCR products were treated with Shrimp Alkaline Phosphatase and Exonuclease I and subsequently sequenced in a reaction using one of the PCR primers at 200 nM final concentration according to manufacturer's instructions (Applied Biosystems). Reactions were run for 25 cycles and their products purified by ethanol/EDTA precipitation, dissolved in HiDi Formamide and run on an Applied Biosystems 3730xl analyzer. The resulting DNA sequences were analyzed using SeqBuilder™ in the DNASTAR® 9.10 package.

Brain Tissue samples

Two human whole brains from neurologically normal anonymous female donors, age 51 and 57, were provided by the National Disease Research Interchange (NDRI). The causes of death were lung cancer and breast cancer, respectively. The first brain was recovered into PBS at 5.7 hours postmortem, kept refrigerated at 4 °C in PBS for 19 hours and dissected into sub-regions at 4 °C. The second brain was recovered at 7.8 hours postmortem and kept at 4 °C in PBS for 21 hours. Both brains were dissected into the 54 functional regions listed in Supplementary Table 2. Immediately upon dissection, brain tissues were snap frozen in -80°C isopentane (Mager et al., 2007), and transferred to glass specimen bottles on dry ice. Dissection and snap freezing procedures were done in a cold room at 4°C. Tissues were kept at -80°C.

RNA extraction

Total RNA was extracted by disruption of frozen tissue using a Covaris S220 tissue homogenizer followed by purification using the Qiagen RNeasy Maxi extraction kit. Frozen

samples representing 0.5-1 gram of each tissue were transferred to a TT2 tissue tube and cryofractured with a Covaris cryoPREP impactor. Fractured frozen tissue powder was immediately transferred to a TC20 bottle tube and 5ml RNeasy buffer RLT was added. Tissues in RLT buffer were then homogenized by S220 focused ultrasonicator for 1-2 minutes (duty cycle 20%, intensity 10, 200 cycles per burst) until fully homogenized. 10ml buffer RLT was added to the homogenized solution resulting in 15ml total volume in order to adjust optimal lysate volume for the RNeasy Maxi kit. Subsequent mRNA extraction was performed according to the manufacturer's instructions using the RNeasy Maxi kit. RNase-Free Dnase I (Qiagen) was added during RNA purification procedures to remove genomic DNA as per the manufacturer's instruction. RNA solutions were stored in aliquots in RNAase-free tubes in 20 μ l volumes at -80°C . In order to maintain the integrity of RNA samples, all RNA extraction procedures were done under RNase-free conditions as per manufacturer's instruction.

RNA integrity tests and cDNA synthesis

RNA integrity was measured before synthesizing cDNA by both NanoDrop 1000 and Agilent 2100 Bioanalyzer systems according to manufacturer's instruction. RNA from the 54 human brain tissues was judged to be free of protein or chemical contaminations by A260/A280 values of 1.92-2.25 and A260/A230 values of 1.43-2.21 (Supplementary Table 2). The RNA integrity number (RIN) values measured with an Agilent 2100 Bioanalyzer and an RNA Nano 6000 kit (Agilent Technologies) were 5.6-8.1 (1 being the most degraded and 10 being the most intact, (Schroeder et al., 2006)) indicating that the RNA integrity was suitable for quantitative reverse-transcriptase real time PCR (qRT-PCR) experiments (Fleige and Pfaffl, 2006). cDNA was synthesized with 930ng of RNA as a template for 54 RNA samples using the Superscript III First-Strand Synthesis Supermix for qRT PCR from Life Technologies. 2 μ l of RT Enzyme Mix were mixed with RNA and 10 μ l of 2X Reaction Mix in a total 20 μ l reaction volume, and incubated at 25°C 10 min followed by 50°C for 50 min then 85°C 5 min. 2.5U of RNase H was then added to the mixture and incubated at 37°C for 20 minutes. Following cDNA synthesis, each reaction was diluted 3-fold by adding 40 μ l water.

Real time qPCR

For gene expression uniformity analysis, 12 candidate reference genes were selected for real time qPCR and primer and probe sets were designed using the real time PCR custom assay design tool from Integrated DNA Technologies (IDT, <http://www.idtdna.com/Scitools/Applications/RealTimePCR/Default.aspx>) and PrimerTime™. Mini qPCR assays with FAM 5'Dye produced by Integrated DNA Technologies were used. The concentration of probe in the reaction mixture was 250 nM, and that of each primer was 500 nM. In case of the main gene expression analysis using target genes, qPCR was performed using Taqman® minor groove binder (MGB) probe/primer sets designed by Primer Express™. Taqman® MGB probes were purchased from Life Technologies, and primer sets were purchased from Integrated DNA Technologies. The real time PCR reactions contained 250 nM MGB probe and 900 nM primers. All primers and probe sets were designed to recognize the same exons to control for the existence of multiple splice isoforms. All the real time PCR reactions were performed using a Viia7 real time PCR system combined with Taqman® Fast Advanced

Master mix from Applied Biosystems. 2 µl of cDNA template was used in a total volume of 10 µl. The first denaturation was done at 95°C for 20 sec after uracil-N-glycosylase (UNG) activation at 50°C for 2 min, which degrade DNA contaminant containing dUTP in it. Amplification was done with 40 cycles of 95°C for 1 sec followed by 60°C for 5 sec. The Ct values were determined by averaging two technical replicates for reference gene uniformity assay and four replicates for main gene expression analysis. Primer and probe information is listed in Supplementary Table 3.

Data analysis

Variants associated with stuttering, dyspraxia, and SLI to date have often been rare coding sequence variants, and we initially selected non-synonymous coding variants with an allele frequency of less than 1% for our analysis. The number of mutant allele was counted, and the association of mutant frequency with the presence of stuttering was tested in 2×2 contingency tables with Fisher's exact test. Association was considered significant when $p \leq 0.01$ ($= 0.05/5$), and marginal when $0.01 < p < 0.05$. Additional association analyses were performed on common SNPs using identical methods. Gene expression uniformity of candidate reference genes were measured using the geNorm algorithm (Vandesompele et al., 2002) which is included in the qbase^{PLUS} software package and Normfinder software (Andersen et al., 2004). Normalized relative quantity (NRQ) of gene expression was calculated from qPCR data using efficiency-corrected comparative quantification methods. (Pfaffl, 2001). PCR efficiency of target genes and endogenous control genes were measured by plotting a standard curve and applied this to comparative Ct methods. The relative quantity (RQ) of the target genes was divided by the geometric means of the RQ of two endogenous control genes to calculate the NRQ of target genes in each brain tissue.

Results

Analysis of association between rare and common variants and persistent stuttering

The exonic regions of *FOXP2*, *CNTNAP2*, *GNPTAB*, *GNPTG*, and *NAGPA* genes were sequenced in a group with familial developmental stuttering and a similar size group of neurologically normal controls from North America and Brazil. Common and rare variants in these five genes were detected in both case and control populations. We selected all rare coding sequence variants (allele frequency < 1% in the our control cohorts (NDPT and BRCO) or in existing databases (1000 Genomes (<http://www.ncbi.nlm.nih.gov/variation/tools/1000genomes>) and the NHLBI Exome Sequencing Project (<http://evs.gs.washington.edu/EVS>)) for further analyses (Table 1) for our initial analysis. Among the total of 81 non-synonymous variants selected, 79 were single amino acid substitutions. The other mutations were an in-frame duplication of amino acid sequence (Leu5Arg7) in *GNPTG* and a 13 bp deletion at the C-terminal end of *NAGPA* that resulted in the addition of 113 nonsense amino acids (Phe513SerfsX113, (Kang et al., 2010))(Table 1). No frame shift or stop codon mutations were detected. We found a CAG repeat number variation (plus one CAG repeat) in the polyQ tract of Exon6 of *FOXP2* in one HCRI stuttering case and in one NDPT control individual. Given that the mutational mechanism and relationship to pathogenicity of polyQ tracts are different from that of amino-acid substitutions, the fact that this variant did not involve a large repeat expansion, and the fact that it was found at equal

low frequency in cases and controls, we did not include this mutation in a list of non-synonymous mutations for the further analysis. All mutation-carrying individuals were heterozygotes and had only one mutation of each gene except for a Brazilian control individual, who had two mutations in *GNPTAB*.

The frequency of mutant alleles of *FOXP2* and *CNTNAP2* in stuttering cases was not significantly different from that in normal controls in both North American and Brazilian cohorts (Table 2), with the exception of the mutation frequency in *CNTNAP2* in Brazilians in which mutations were identified in 2.2 percent of Brazilian cases compared to 0.7 percent in Brazilian controls. Among *GNPTAB*, *GNPTG* and *NAGPA*, the frequency of mutant allele in *NAGPA* was significantly higher in our North European American stuttering cohort than in normal controls ($p = 0.0091$), and the mutation frequency for *GNPTAB* was significantly higher in cases than normal controls in our Brazilian cohort ($p = 0.00050$). Rare coding variants in *GNPTG* showed a marginal association in the North American cohort ($p = 0.024$). We additionally searched for predicted functions of individual missense mutations using Polyphen-2 (<http://genetics.bwh.harvard.edu/pph2>) and SIFT (http://sift.jcvi.org/www/SIFT_enst_submit.html). Only two mutations in *GNPTAB* (Ala455Ser and Thr1066Met) were found in multiple stuttering cases and had damaging effects predicted by both Polyphen-2 and SIFT. One of these (Ala455Ser) has been previously reported as a stuttering-associated mutation (Kang et al., 2010).

We also considered common variants in these five genes. We detected 10 common single nucleotide polymorphisms (SNPs) in the coding regions of *GNPTAB*, *NAGPA*, and *CNTNAP2*. No common SNP variants were detected in the coding regions of *FOXP2* and *GNPTG* (Supplementary Table 4). Among the 10 SNPs, two were non-synonymous SNPs, rs76889468 in *GNPTAB* and rs7188856 in *NAGPA*. Neither of these two SNPs showed significant association with stuttering in our samples. Interestingly, a synonymous SNP, rs222504 (Gln9Gln) in *GNPTAB* displayed an association (7.4 times increased risk) with persistent stuttering in Brazilian cohorts ($P = 0.0001$). This variant showed marginal association in the North American cohorts ($P = 0.047$).

Selection of endogenous control genes for qPCR study in neurologically normal human brains

The identification of uniformly expressed endogenous control genes is a prerequisite for accurate qPCR experiments. We validated the uniformity of expression of 12 candidate control genes across 54 functionally different brain sub-tissues from two individuals. These 12 genes (*GAPDH*, *GPS1*, *GUSB*, *H6PD*, *HBB*, *POLR2A*, *PON1*, *TBP*, *TFRC*, *TOPI*, *UBC*, and *UBE2D2*) were selected as candidate reference genes based on previous studies of gene expression in human brain (Coulson et al., 2008; Johansson et al., 2007; Kreth et al., 2010; Penna et al., 2011) and on the Allen Brain Atlas microarray database (<http://human.brain-map.org/microarray/search>). As defined in previous studies using geNorm analysis (Vandesompele et al., 2002), a lower M value means that the expression level of the gene fluctuates comparatively less, and thus genes with lower M values are more uniformly expressed across different tissues. M values were calculated by inputting raw C_T values in qbase^{PLUS}. This analysis showed that *UBE2D2* was the most uniformly expressed gene

across brain tissue samples of the first brain ($M = 0.248$) and *GPSI* was the most uniformly expressed gene across the second brain ($M = 0.505$) (Fig. 1). The expression uniformity of *GPSI* was highly ranked in the first brain (4th, $M = 0.334$) and the uniformity of *UBE2D2* expression was also highly ranked in the second brain (4th, $M = 0.645$). In order to test reproducibility of geNorm analyses, we measured gene expression uniformity using Normfinder for these 12 candidate reference genes. Although the gene expression uniformity ranking calculated by Normfinder showed some differences from that produced by geNorm, *GPSI* was the best gene (Stability (S) value = 0.103 in the first brain) and *UBE2D2* was the best in the second brain ($S = 0.296$) in Normfinder analysis (Fig. 1). *UBE2D2* was ranked as 4th in the first brain ($S = 0.204$) and *GPSI* was ranked as 3rd in the second brain ($S = 0.332$). As a result, we chose *GPSI* and *UBE2D2* as our two endogenous control genes for our subsequent qPCR experiments.

Quantitative RT-PCR analysis of gene expression levels in adult human brains

Because our case subjects display persistent stuttering and are adults, we studied gene expression in adult human brain. We used qPCR to measure relative expression levels of *FOXP2*, *CNTNAP2*, *GNPTAB*, *GNPTG* and *NAGPA* in 27 different brain sub-tissues that have been described as related to speech and motor functions among the 54 brain tissues we dissected. Each brain tissue was sampled from both left and right hemispheres except for the cerebellar vermis. The NRQ of target genes in two individual brains were combined to produce an average, and the values of these averages in different brain tissues were plotted (Supplementary Fig. 1). The 27 brain tissues can be categorized into 4 groups: cerebellum (anterior lobe, posterior lobe, flocculonodular lobe), basal ganglia (caudate nucleus, putamen, globus pallidus, thalamus), cerebral cortex (Broca's area, insular cortex, premotor cortex, primary motor cortex), and brain stem (medulla oblongata and pons). We calculated average NRQ in each of these four groups and compared gene expression among them. *FOXP2* and *CNTNAP2* showed lower expression in the cerebellum compared to the other three tissue groups (Fig. 2). *FOXP2* is characterized by higher expression in basal ganglia than in cerebral cortex and brain stem. In contrast, *CNTNAP2* shows lower gene expression in basal ganglia than these two tissue groups. The expression of *GNPTAB* and *GNPTG* in 27 brain tissues was more evenly distributed than that of *FOXP2* and *CNTNAP2* (Supplementary Fig. 1). The difference in NRQ values between cerebellum and the other tissue groups was smaller for *GNPTAB* (0.98 – 1.4 fold) and *GNPTG* (1.6 – 2.2 fold) than for *FOXP2* (8.4 – 11 fold) and *CNTNAP2* (3.5 – 5.0 fold). In the case of *NAGPA*, the NRQ in cerebellum is much smaller than in cerebral cortex (4.6 fold) but NRQ ratio of basal ganglia versus cerebellum (2.4 fold) and brain stem versus cerebellum (1.5 fold) is smaller than those of *FOXP2* (11 and 8.1) or *CNTNAP2* (3.5 and 5.0).

Gene expression analyses in microarray data of human brains

We also analyzed the expression of these five genes using the Allen Brain Atlas microarray data (<http://human.brain-map.org>) to compare the results of their microarray data with our qPCR data. Using raw microarray data (<http://human.brain-map.org/static/download>), we extracted expression data for these five genes in 27 different brain tissues which correspond to the brain tissues for which we had qPCR data. Among 6 donors used for the Allen Brain Atlas microarray survey, donors H0351.2001 (Male, 24 years) and H0351.2002 (Male, 39

years) were selected because they had data from both brain hemispheres, and the average of gene expression levels of these two individuals was used for our analysis. The expression levels of these five genes as measured by microarray were more uniform across brain regions than those observed in our qPCR data (Supplementary Fig. 2 and Fig 3). However, the generally lower gene expression of *FOXP2* and *CNTNAP2* in cerebellum seen our qPCR data was also observed in the Allen Brain Atlas micorarray data. In addition, a similar distribution of expression of *GNPTAB* and *GNPTG* was observed in all tissue groups and relatively high expression of *NAGPA* in cerebral cortex was observed in both our qPCR data and the microarray data.

We additionally collected microarray data from developing human brain from the BrainSpan Atlas of the Developing Human Brain database (<http://www.brainspan.org>). The age of donors for this database ranged from prenatal to 40 years. Among them, we selected data from an infant group (0-1 year old) and a childhood group (3-8 years old), and gene expression data in each group was compared with adult human brain data from the Allen Brain Atlas microarray survey. We chose five tissues (cerebellum, Broca's area, primary motor cortex, striatum, and thalamus), all of which are related to speech motor function, for our comparative analysis. The average of relative expression levels of our five genes in these five tissues were calculated in different age groups and plotted (Supplementary figure 3). The expression distributions among the three different age groups were highly consistent, suggesting that distribution of expression of these five genes in human brain change little with respect to age.

Discussion

Our results confirm the presence of a statistically significant increase of rare coding sequence variants of the *GNPTAB* and *NAGPA* genes in persistent developmental stuttering and similar trend of *GNPTG* even if it was not statistically significant. Our case-control association results also show no such significant increase of rare variants of *FOXP2* and *CNTNAP2* in stuttering cases. In addition, gene expression data shows different expression patterns of *FOXP2* and *CNTNAP2* compared to *GNPTAB*, *GNPTG* and *NAGPA* with lower expression in cerebellum in neurologically normal adult human brain. Thus two types of molecular genetic information, DNA coding sequence variants and RNA expression patterns in different anatomic regions of the brain, suggest that verbal dyspraxia and SLI are etiologically distinct from stuttering.

Approximately 5% of all children exhibit developmental stuttering, however the majority of these recover, often at a young age. Such recovered stutterers are therefore likely to exist in any population of neurologically normal adults. Without detailed speech fluency history in our otherwise normal controls we cannot test for an association with a previous history of stuttering. However, variants in the *GNPTAB*, *GNPTG*, and *NAGPA* genes may also contribute to stuttering for which recovery occurs. This may explain our findings, in which 5.3% of neurologically normal adults carry rare non-synonymous coding sequence variants in the *GNPTAB*, *GNPTG*, or *NAGPA* genes previously shown to be associated with stuttering (Kang et al., 2010). Because mutations also exist in normal individuals, we used case-control association tests to compare genetic contributions of different genes to

persistent stuttering without any selective criteria for putative causal mutations. These results are consistent with our previous results with *GNPTAB*, *GNPTG* and *NAGPA* that used selective criteria for determining stuttering-causal mutations (Kang et al., 2010). Although the rare variants in *FOXP2* and *CNTNAP2* did not show any association with persistent stuttering in our case-control association tests, we cannot rule out the possibility that individual mutations such as Ala615Thr in *FOXP2*, which are predicted to have damaging effects by Poyphen2 and SIFT, are functionally pathogenic. Because *FOXP2* variants are generally rare (MacDermot et al., 2005), genetic association evidence may be limited, and further functional investigations may be necessary in order to confirm pathogenicity of these mutations.

Conventional endogenous control genes for qPCR typically include *ACTB*, encoding beta actin or *GAPDH*, encoding glyceraldehyde-3-phosphate dehydrogenase. Because the expression levels of conventional or other house-keeping genes have been shown to vary in a number of circumstances (Bustin, 2002; Dheda et al., 2004; Spanakis, 1993), identification of proper endogenous control genes is critical for obtaining accurate data from qPCR using specific tissues or phenotypes. There have been several reports regarding the selection of reference genes for qPCR using human brain (Coulson et al., 2008; Johansson et al., 2007; Kreth et al., 2010; Penna et al., 2011), however all of these were done using individuals with brain disorders or a specific phenotype. We validated gene expression uniformity testing a large series of candidate reference genes, and *UBE2D2*, *GPS1* and *TBP* were found to be the best reference genes when comparing expression between sub-tissues of two normal human brains. In our study, we surveyed a range of potential reference genes in normal human brain for normalization of expression values. Previous studies that generated the Allen Brain Atlas microarray data did not employ a normalization strategy using endogenous control genes (Johnson et al., 2007), and thus our studies add additional information on the relative advantages of different housekeeping genes previously used for normalization of expression data in human neural tissue. Our results suggest that *UBE2D2*, *GPS1* and *TBP* represent superior normalization controls for expression studies in human brain.

The Allen Brain Atlas project includes genome-wide gene expression data in normal human brains (Hawrylycz et al., 2012; Jones et al., 2009). In both the Allen Brain Atlas microarray data and our current qPCR data, lower expression of both *FOXP2* and *CNTNAP2* in cerebellum was consistently detected, although the expression differences between tissues in Allen Brain Atlas data are not as high as those seen in our qPCR data. This difference between the two methods may arise from the low dynamic range and sensitivity of microarray compared to qPCR methods (Morey et al., 2006). While our qRT-PCR method has the strengths of high accuracy and wide dynamic range, it used tissue samples from different brain regions that are composed of multiple neuronal cellular layers and cell types. We note that a previous study showed that the *FOXP2* and *CNTNAP2* are expressed in different layers in human fetal cortex (Vernes et al., 2008). Thus, our data does not show gene expression information in neuronal subpopulations within our anatomic locations, which is a potential shortcoming of this study. Additional gene expression studies in individual neuronal cells types will be able to complement the shortcomings of the current study.

The expression patterns of *GNPTAB*, *GNPTG* and *NAGPA* do not show an obvious common feature in human brain, in agreement with previously reported studies (Bohland et al., 2013). Thus at the current time, an assignment of a functional roll for specific brain tissues in the pathogenesis of stuttering based on only gene expression data is difficult. Interestingly, the relative expression patterns of *GNPTAB*, *GNPTG*, and *NAGPA* were markedly different and those of *FOXP2* and *CNTNAP2* only in cerebellum. This would suggest that the cerebellum may have a specific functional roll in the pathogenesis of stuttering. Most of the previous studies on the learning and production of vocalizations in animal models using *Foxp2* gene have focused on motor circuit between motor cortex and basal ganglia (Bolhuis et al., 2010; Enard et al., 2009; French et al., 2012; Haesler et al., 2007; Reimers-Kipping et al., 2011; Schulz et al., 2010; Teramitsu et al., 2010). Our data suggest that motor circuits related to the cerebellum may be another target for the study of the molecular mechanisms of stuttering pathogenesis caused by mutation of *GNPTAB*, *GNPTG*, or *NAGPA*.

Speech production involves complex coordination of brain motor circuits, and multiple genes must be related to this process. *FOXP2* regulates a broad range of genes (Spiteri et al., 2007; Vernes et al., 2008) and these target genes have been implicated in other neurodevelopmental disorders. For example, variants of *CNTNAP2* have been also been strongly implicated in autism, a disorder characterized by deficits in spoken language (Alarcon et al., 2008; Penagarikano et al., 2011; Sampath et al., 2013). Thus, while studies of *FOXP2* and its targets have emphasized common neurodevelopmental pathways in speech and language disorders, our results suggest differentiation of stuttering from these previously studied pathways. Causal mutations for different speech disorders may also be found in different genes that may function independently from each other, and in this study, we found that mutations in genes related to other speech-language disorders are not associated with stuttering. However we note that mutations in the genes studied here have been shown to be causative in only a small fraction of cases of their respective disorders. While our results suggest that stuttering has a different genetic etiology than that of verbal dyspraxia and SLI, it seems likely that additional causative genes for all three disorders remain to be discovered. Thus mutations in additional genes may identify pathways or neuropathology common to both stuttering and other speech disorders.

Our results also support the view that a focus on lysosomal targeting genes will be useful for studies of the pathogenesis of persistent stuttering. Because gene expression data itself does not currently suggest a function of these genes specific to a specific brain region, functional studies of lysosomal targeting proteins, such as protein activity assays or investigation of impact on signaling pathways may reveal functional contributions of specific neuronal cells or brain tissues to the pathogenesis of stuttering.

Supplementary Fig. 1. Relative expression levels of five genes using qPCR in 27 human brain subtissues. The normalized relative quantity (NRQ) of target gene expression levels measured in 27 brain tissues was plotted. NRQ of each tissue represents the average of two individuals. The 27 tissues can be categorized into four general locations; brain stem (blue), cerebellum (red), cerebral cortex (green), basal ganglia (yellow).

Supplementary Fig. 2. Relative expression levels of five genes in Allen Brain Atlas human microarray data. Raw microarray data were downloaded from <http://human.brain-map.org/microarray/search>. Gene expression data for five genes were extracted for the 27 different brain tissues in two individuals (H0351.2001 and H0351.2002). The 27 tissues were categorized into four groups; brain stem (blue), cerebellum (red), cerebral cortex (green), basal ganglia (yellow). In cases where a functional brain region was comprised of multiple sub-tissues, the average was calculated and plotted.

Supplementary Fig. 3. Comparison of expression profiles of human brain in different age groups using the Allen Brain Atlas database. Microarray data of infant group (0~1year old) and a children group (3~8 years old) were collected from the BrainSpan Atlas of the Developing Human Brain database (<http://www.brainspan.org>) and data from two adult human brain was collected from Allen Brain Atlas microarray survey (<http://human.brain-map.org/microarray/search>). The average of relative expression of different individuals was calculated in each age group and plotted.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the National Disease Research Interchange for provision of normal human brain samples, and the stuttering subjects who participated in this study. We thank Rachel Rahn for assistance with DNA sequencing, and Drs. Thomas Friedman and Robert Morell for helpful comments on the manuscript. This research was supported by NIDCD intramural grant # Z01-000046-13 and by the NINDS.

References

- Alarcon M, et al. Linkage, association, and gene-expression analyses identify CNTNAP2 as an autism-susceptibility gene. *Am J Hum Genet.* 2008; 82:150–9. [PubMed: 18179893]
- Andersen CL, et al. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res.* 2004; 64:5245–50. [PubMed: 15289330]
- Bishop DV. Genetic and environmental risks for specific language impairment in children. *Philos Trans R Soc Lond B Biol Sci.* 2001; 356:369–80. [PubMed: 11316485]
- Bohland JW, et al. An Informatics Approach to Integrating Genetic and Neurological Data in Speech and Language Neuroscience. *Neuroinformatics.* 2013
- Bolhuis JJ, et al. Twitter evolution: converging mechanisms in birdsong and human speech. *Nat Rev Neurosci.* 2010; 11:747–59. [PubMed: 20959859]
- Bustin SA. Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems. *J Mol Endocrinol.* 2002; 29:23–39. [PubMed: 12200227]
- Canhetti-Oliveira CM, Richieri-Costa A. A study of familial stuttering. *Am J Med Genet A.* 2006; 140:2139–41. [PubMed: 16752384]
- Coulson DT, et al. Identification of valid reference genes for the normalization of RT qPCR gene expression data in human brain tissue. *BMC Mol Biol.* 2008; 9:46. [PubMed: 18460208]
- Dheda K, et al. Validation of housekeeping genes for normalizing RNA expression in real-time PCR. *Biotechniques.* 2004; 37:112–4. 116, 118–9. [PubMed: 15283208]
- Enard W, et al. A humanized version of Foxp2 affects cortico-basal ganglia circuits in mice. *Cell.* 2009; 137:961–71. [PubMed: 19490899]

- Fisher SE, et al. Localisation of a gene implicated in a severe speech and language disorder. *Nat Genet.* 1998; 18:168–70. [PubMed: 9462748]
- Fleige S, Pfaffl MW. RNA integrity and the effect on the real-time qRT-PCR performance. *Mol Aspects Med.* 2006; 27:126–39. [PubMed: 16469371]
- French CA, et al. An aetiological Foxp2 mutation causes aberrant striatal activity and alters plasticity during skill learning. *Mol Psychiatry.* 2012; 17:1077–85. [PubMed: 21876543]
- Gordon N. Stuttering: incidence and causes. *Dev Med Child Neurol.* 2002; 44:278–81. [PubMed: 11995897]
- Haesler S, et al. Incomplete and inaccurate vocal imitation after knockdown of FoxP2 in songbird basal ganglia nucleus Area X. *PLoS Biol.* 2007; 5:e321. [PubMed: 18052609]
- Hawrylycz MJ, et al. An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature.* 2012; 489:391–9. [PubMed: 22996553]
- Johansson S, et al. Validation of endogenous controls for quantitative gene expression analysis: application on brain cortices of human chronic alcoholics. *Brain Res.* 2007; 1132:20–8. [PubMed: 17188656]
- Johnson WE, et al. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics.* 2007; 8:118–27. [PubMed: 16632515]
- Jones AR, et al. The Allen Brain Atlas: 5 years and beyond. *Nat Rev Neurosci.* 2009; 10:821–8. [PubMed: 19826436]
- Kang C, et al. Mutations in the lysosomal enzyme-targeting pathway and persistent stuttering. *N Engl J Med.* 2010; 362:677–85. [PubMed: 20147709]
- Kent RD. Research on speech motor control and its disorders: a review and prospective. *J Commun Disord.* 2000; 33:391–427. quiz 428. [PubMed: 11081787]
- Kreth S, et al. Identification of valid endogenous control genes for determining gene expression in human glioma. *Neuro Oncol.* 2010; 12:570–9. [PubMed: 20511187]
- Lai CS, et al. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature.* 2001; 413:519–23. [PubMed: 11586359]
- MacDermot KD, et al. Identification of FOXP2 truncation as a novel cause of developmental speech and language deficits. *Am J Hum Genet.* 2005; 76:1074–80. [PubMed: 15877281]
- Mager SR, et al. Standard operating procedure for the collection of fresh frozen tissue samples. *Eur J Cancer.* 2007; 43:828–34. [PubMed: 17329097]
- Morey JS, et al. Microarray validation: factors influencing correlation between oligonucleotide microarrays and real-time PCR. *Biol Proced Online.* 2006; 8:175–93. [PubMed: 17242735]
- Newbury DF, et al. Recent advances in the genetics of language impairment. *Genome Med.* 2010; 2:6. [PubMed: 20193051]
- Penagarikano O, et al. Absence of CNTNAP2 leads to epilepsy, neuronal migration abnormalities, and core autism-related deficits. *Cell.* 2011; 147:235–46. [PubMed: 21962519]
- Penna I, et al. Selection of candidate housekeeping genes for normalization in human postmortem brain samples. *Int J Mol Sci.* 2011; 12:5461–70. [PubMed: 22016602]
- Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* 2001; 29:e45. [PubMed: 11328886]
- Raza MH, et al. Studies in a consanguineous family reveal a novel locus for stuttering on chromosome 16q. *Hum Genet.* 2012; 131:311–3. [PubMed: 22205390]
- Raza MH, et al. Identification of an autosomal recessive stuttering locus on chromosome 3q13.2-3q13.33. *Hum Genet.* 2010; 128:461–3. [PubMed: 20706738]
- Reimers-Kipping S, et al. Humanized Foxp2 specifically affects cortico-basal ganglia circuits. *Neuroscience.* 2011; 175:75–84. [PubMed: 21111790]
- Riaz N, et al. Genomewide significant linkage to stuttering on chromosome 12. *Am J Hum Genet.* 2005; 76:647–51. [PubMed: 15714404]
- Sampath S, et al. Defining the contribution of CNTNAP2 to autism susceptibility. *PLoS One.* 2013; 8:e77906. [PubMed: 24147096]
- Schroeder A, et al. The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Mol Biol.* 2006; 7:3. [PubMed: 16448564]

- Schulz SB, et al. Knockdown of FoxP2 alters spine density in Area X of the zebra finch. *Genes Brain Behav.* 2010; 9:732–40. [PubMed: 20528955]
- Spanakis E. Problems related to the interpretation of autoradiographic data on gene expression using common constitutive transcripts as controls. *Nucleic Acids Res.* 1993; 21:3809–19. [PubMed: 8367299]
- Spiteri E, et al. Identification of the transcriptional targets of FOXP2, a gene linked to speech and language, in developing human brain. *Am J Hum Genet.* 2007; 81:1144–57. [PubMed: 17999357]
- Teramitsu I, et al. Striatal FoxP2 is actively regulated during songbird sensorimotor learning. *PLoS One.* 2010; 5:e8548. [PubMed: 20062527]
- Vandesompele J, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* 2002; 3:RESEARCH0034. [PubMed: 12184808]
- Vernes SC, et al. A functional genetic link between distinct developmental language disorders. *N Engl J Med.* 2008; 359:2337–45. [PubMed: 18987363]

Highlights

- We tested the hypothesis that the *FOXP2* and *CNTNAP2* genes play a role in persistent stuttering
- Coding sequence mutations in these genes occur at the same rates in stuttering cases and controls
- The pattern of expression of these genes in brain differs from that of known stuttering genes
- The genetic origins of dyspraxia and specific language impairment differ from those of stuttering

Fig. 1.

Gene expression uniformity tests of 12 candidate reference genes in 54 sub-tissues from two human brains. The M value of geNorm and the S value of Normfinder were calculated using C_T values of 12 candidate genes in 54 different human brain tissues. Lower M values or S values correspond to more uniform gene expression. Gene expression uniformity is displayed in increasing order from left to right.

Fig. 2.

Relative expression levels of five genes by qRT-PCR in two human brains. The normalized relative quantity (NRQ) of target gene expression levels measured in 27 brain tissues was averaged within 4 groups, cerebellum (anterior lobe, posterior lobe, flocculonodular lobe), basal ganglia (caudate nucleus, putamen, globus pallidus, thalamus), cerebral cortex (Broca's area, insular cortex, premotor cortex, primary motor cortex), Brain stem (medulla oblongata, pons). The Y axis of each graph is the NRQ.

Fig. 3.

Relative expression levels of five genes in Allen Brain Atlas microarray data. Raw gene expression data of two individuals (H0351.2001 and H0351.2002) was downloaded from Allen Brain Atlas microarray database (<http://human.brain-map.org/static/download>) and gene expression data for these five genes was extracted for the 27 different brain tissues which correspond to the brain tissues for which we had qPCR data. As in the qPCR analyses, the 27 brain tissues were categorized into 4 groups, and the average of normalized gene expression levels in each group was plotted. Y axis: normalized gene expression levels.

Table 1

Non-synonymous mutations found in this study

	Mutation	Amino acid change	HCRI ^a	NAF	BRMII	BRPD	NDPT	BRCO	POLYPEN_2	SIFT
FOXP2	NM_014491.3		n = 294	n = 144	n = 108	n = 51	n = 276	n = 211		
EXON 16	c.2217 G>g/a	Ala615Thr	1	0	0	0	0	0	POSS. DAMAGING	DAMAGING
	c.2310 A>a/g	Asn646Asp	0	0	0	0	0	1	PROB. DAMAGING	DAMAGING
CNTNAP2	NM_014141.5		n = 294	n = 144	n = 108	n = 51	n = 276	n = 211		
EXON 1	c.73 G>g/a	Ala25Thr	0	1	0	0	0	0	BENIGN	DAMAGING
EXON 3	c.282 G>g/t	Gln94His	0	0	1	0	0	0	PROB. DAMAGING	DAMAGING
EXON 6	c.854 G>g/c	Gly285Ala	1	0	0	0	1	1	POSS. DAMAGING	TOLERATED
EXON 8	c.1135 A>a/g	Asn379Asp	0	0	0	0	0	1	BENIGN	TOLERATED
	c.1145 G>g/a	Ser382Asn	0	0	1	0	0	0	PROB. DAMAGING	TOLERATED
	c.1220 A>a/g	Asn407Ser	0	2	0	0	0	0	BENIGN	TOLERATED
	c.1247 C>c/t	Ala416Val	1	0	1	2	0	1	BENIGN	TOLERATED
EXON 9	c.1406 T>t/c	Phe469Ser	1	0	0	0	0	0	BENIGN	TOLERATED
	c.1423 G>g/a	Asp475Asn	0	0	0	0	1	0	PROB. DAMAGING	TOLERATED
	c.1485 G>g/t	Lys495Asn	0	0	0	0	1	0	BENIGN	TOLERATED
EXON 10	c.1579 G>g/c	Asp527His	0	0	0	0	1	0	PROB. DAMAGING	DAMAGING
EXON 12	c.1879 G>g/a	Val627Ile	0	0	1	0	0	0	PROB. DAMAGING	DAMAGING
EXON 13	c.2038 G>g/a	Glu680Lys	0	0	0	0	1	0	POSS. DAMAGING	TOLERATED
EXON 17	c.2580 T>t/a	Phe860Leu	0	0	0	0	1	0	BENIGN	TOLERATED
	c.2606 T>t/c	Ile869Thr	1	0	0	0	0	0	BENIGN	TOLERATED
	c.2743 C>c/t	Arg915Cys	1	0	0	0	0	0	PROB. DAMAGING	TOLERATED
EXON 18	c.2903 C>c/t	Thr968Ile	0	0	0	0	1	0	BENIGN	TOLERATED
EXON 19	c.3095 G>g/a	Arg1032Lys	0	0	1	0	0	0	BENIGN	TOLERATED
	c.3116 A>a/c	Gln1039Pro	1	0	0	0	0	0	BENIGN	TOLERATED
	c.3148 G>g/a	Glu1050Lys	1	0	0	0	0	0	BENIGN	TOLERATED
	c.3179 C>c/t	Ala1060Val	1	0	0	0	0	0	PROB. DAMAGING	TOLERATED
EXON 24	c.3932 T>t/c	Ile1311Thr	1	0	0	0	0	0	PROB. DAMAGING	TOLERATED
	c.3950 A>a/g	Asn1317Ser	1	0	0	0	0	0	BENIGN	TOLERATED
GNPTAB	NM_024312.4		n = 293	n = 149	n = 108	n = 51	n = 276	n = 211		

	Mutation	Amino acid change	HCR ^a	NAF	BRMII	BRPD	NDPT	BRCO	POLYPEN_2	SIFT
EXON 1	c.70 T>T/g	Phe24Val	1	0	0	0	0	0	POSS. DAMAGING	DAMAGING
EXON 2	c.137 G>g/a	Arg46Gln	5	4	2	0	3	0	BENIGN	TOLERATED
EXON 4	c.337 A>a/g	Lys113Gln	0	0	1	1	0	0	BENIGN	TOLERATED
EXON 5	c.346 A>a/g	Thr116Ala	0	0	0	0	1	0	POSS. DAMAGING	DAMAGING
	c.500 T>t/a	Ile167Asn	0	0	0	0	0	1	BENIGN	DAMAGING
	c.561 G>g/c	Ser187Thr	1	0	0	0	0	0	BENIGN	TOLERATED
EXON 6	c.580 G>g/a	Ala194Thr	0	0	0	0	1	0	POSS. DAMAGING	DAMAGING
EXON 8	c.886 A>a/g	Thr296Ala	0	0	0	0	0	1	BENIGN	TOLERATED
EXON 9	c.1042 A>a/c	Ile348Leu	0	1	3	2	0	1	POSS. DAMAGING	TOLERATED
EXON 10	c.1253 A>a/g	Asp418Gly	0	0	0	1	0	0	POSS. DAMAGING	DAMAGING
	c.1283 A>a/g	Lys428Arg	0	0	0	0	0	0	POSS. DAMAGING	DAMAGING
EXON 11	c.1363 G>g/t	Ala455Ser	2	0	0	0	0	0	PROB. DAMAGING	DAMAGING
EXON 12	c.1429 T>t/g	Tyr477Asp	0	0	0	0	0	1	POSS. DAMAGING	TOLERATED
EXON 13	c.1669 A>a/c	Ile557Leu	0	0	0	0	0	0	BENIGN	TOLERATED
	c.1803 A>a/g	Ile601Met	0	0	0	0	1	0	POSS. DAMAGING	DAMAGING
	c.1943 G>g/a	Gly648Asp	0	0	1	1	0	0	BENIGN	TOLERATED
	c.1948 G>g/a	Glu650Lys	1	0	0	0	0	0	BENIGN	TOLERATED
	c.2341 G>g/a	Val781Met	0	0	1	0	0	0	BENIGN	TOLERATED
	c.2582 G>g/a	Arg861Lys	0	0	0	0	0	0	BENIGN	TOLERATED
	c.2611 G>g/a	Gly871Ser	1	4	2	2	0	1	BENIGN	TOLERATED
EXON 16	c.3197 C>c/t	Thr1066Met	0	2	0	0	0	0	PROB. DAMAGING	DAMAGING
EXON 19	c.3572 G>g/a	Arg1191His	0	0	0	0	1	0	PROB. DAMAGING	DAMAGING
GNPTG	NM_032520.4		n = 293	n = 149	n = 108	n = 51	n = 276	n = 211		
EXON 1	c.11_19dup	Leu5Arg7	1	0	0	0	0	0	N/A	N/A
EXON 2	c.74 C>c/a	Ala25Glu	2	0	0	0	0	0	BENIGN	TOLERATED
	c.74 C>c/g	Ala25Gly	1	0	0	0	0	0	BENIGN	TOLERATED
EXON 3	c.161 A>a/t	Asp54Val	0	1	0	0	0	0	POSS. DAMAGING	TOLERATED
EXON 5	c.316 G>g/a	Gly106Ser	0	0	0	0	0	1	PROB. DAMAGING	DAMAGING
EXON 6	c.388 C>c/t	Arg130Cys	1	0	0	0	0	0	BENIGN	TOLERATED
	c.401 G>g/a	Arg134Gln	0	0	0	0	0	1	PROB. DAMAGING	DAMAGING
EXON 7	c.422 C>c/t	Ala141Val	0	0	1	0	0	0	BENIGN	TOLERATED

Mutation	Amino acid change	HCR1 ^a	NAF	BRMII	BRPD	NDPT	BRCO	POLYPEN_2	SIFT
EXON 8	c.472 G>g/a	2	2	0	0	1	0	POSS. DAMAGING	TOLERATED
	c.557 g>g/a	0	1	0	0	0	0	BENIGN	TOLERATED
	c.814 A>a/g	0	1	0	1	1	0	BENIGN	TOLERATED
	c.878 A>a/c	0	2	0	0	0	0	BENIGN	TOLERATED
EXON 11	c.887 G>g/a	1	0	0	0	0	0	PROB. DAMAGING	DAMAGING
	c.904 C>c/t	0	0	0	0	0	1	PROB. DAMAGING	DAMAGING
	c.910 A>a/g	0	0	4	0	0	0	BENIGN	TOLERATED
	NM_016256.3	n = 293	n = 149	n = 108	n = 51	n = 276	n = 211		
EXON 1	c.83 C>c/t	0	0	1	0	0	1	BENIGN	TOLERATED
EXON 2	c.136 C>c/t	0	0	1	0	0	0	PROB. DAMAGING	DAMAGING
	c.220 G>g/c	0	0	0	0	0	1	BENIGN	TOLERATED
	c.236 G>g/a	0	1	0	0	0	0	BENIGN	TOLERATED
	c.252 C>c/g	2	0	0	0	0	0	BENIGN	TOLERATED
EXON 3	c.257 G>g/a	1	0	0	0	0	0	BENIGN	TOLERATED
	c.562 G>g/a	0	0	0	1	0	0	PROB. DAMAGING	DAMAGING
	c.591 G>g/c	0	1	0	0	0	0	PROB. DAMAGING	DAMAGING
	c.641 A>a/g	0	0	1	0	0	1	BENIGN	TOLERATED
EXON 4	c.667 G>g/a	0	1	0	0	0	0	BENIGN	TOLERATED
	c.676 G>g/c	0	1	0	0	0	0	PROB. DAMAGING	DAMAGING
	c.785 A>a/g	0	0	0	0	0	1	POSS. DAMAGING	TOLERATED
	c.982 C>c/t	7	1	0	0	0	1	PROB. DAMAGING	TOLERATED
EXON 7	c.1129 G>g/a	0	1	0	0	0	0	PROB. DAMAGING	TOLERATED
	c.1187 G>g/c	1	0	0	0	0	0	PROB. DAMAGING	TOLERATED
	c.1243 G>g/c	0	0	0	0	1	0	BENIGN	TOLERATED
	c.1264 A>a/t	0	0	0	0	0	1	PROB. DAMAGING	TOLERATED
EXON 9	c.1319 G>g/c	0	0	0	1	1	1	BENIGN	TOLERATED
EXON 10	c.1538_1553Del16bp	1	0	0	0	0	0	N/A	N/A

^aHCR1 = North American unrelated stutterers; NAF = North American stuttering probands; BRMII = Brazilian unrelated stuttering cases; BRPD = Brazilian stuttering probands; NDPT = North American normal controls; BRCO = Brazilian normal controls

Table 2

Association analysis of non-synonymous rare variants in *FOXP2*, *CNTNAP2*, *GNPTAB*, *GNPTG* and *NAGPA* with persistent stuttering.

		Stuttering case ^a	Normal control	p	OR (95% CI)
North American	<i>FOXP2</i>	1/875 (0.0011)	0/552 (0)	0.39	N/A
	<i>CNTNAP2</i>	13/863 (0.015)	7/545 (0.013)	0.82	1.17 (0.47 - 2.96)
	<i>GNPTAB</i>	22/862 (0.025)	7/545 (0.013)	0.13	1.99 (0.84 - 4.68)
	<i>GNPTG</i>	15/869 (0.017)	2/550 (0.0036)	0.024	4.75 (1.08 - 20.84)
	<i>NAGPA</i>	18/866 (0.020)	2/550 (0.0036)	0.0091	5.72 (1.32 - 24.73)
Brazilian	<i>FOXP2</i>	0/318 (0)	1/421 (0.0024)	1	N/A
	<i>CNTNAP2</i>	7/311 (0.022)	3/419 (0.0071)	0.082	3.14 (0.81 - 12.25)
	<i>GNPTAB</i>	17/301 (0.053)	4/418 (0.0095) ^c	0.00050	5.90 (1.97 - 17.72)
	<i>GNPTG</i>	6/312 (0.019)	3/419 (0.0071)	0.18	2.69 (0.67 - 10.82)
	<i>NAGPA</i>	5/313 (0.016)	7/415 (0.017)	1	0.95 (0.30 - 3.01)

^bThe total number of mutant alleles does not equal 5 because one Brazilian individual carries two mutations.

^a# of mutant allele /# of wild type allele (frequency of mutant allele)