Published in final edited form as:

Autism Res. 2008 August ; 1(4): 251–257. doi:10.1002/aur.30.

A large-scale screen for coding variants in SERT/SLC6A4 in autism spectrum disorders

Takeshi Sakurai^{1,2,3}, Jennifer Reichert^{1,2}, Ellen J. Hoffman^{1,2}, Guiqing Cai^{1,2}, Hywel B. Jones^{6,7}, Malek Faham⁶, and Joseph D. Buxbaum^{1,2,4,5}

¹Laboratory of Molecular Neuropsychiatry, the Seaver Autism Research Center, Mount Sinai School of Medicine, New York, NY

²Department of Psychiatry, Mount Sinai School of Medicine, New York, NY

³Department of Pharmacology and Systems Therapeutics, Mount Sinai School of Medicine, New York, NY

⁴Department of Genetics and Genomic Sciences, Mount Sinai School of Medicine, New York, NY

⁵Department of Neuroscience, Mount Sinai School of Medicine, New York, NY

⁶ParAllele Bioscience, Inc. (now part of Affymetrix, Inc.), South San Francisco, CA.

Abstract

In the current study we explored the hypothesis that rare variants in *SLC6A4* contribute to autism susceptibility and to rigid-compulsive behaviors in autism. We made use of a large number of unrelated cases with autism spectrum disorders (~350) and controls (~420) and screened for rare exonic variants in *SLC6A4* by a high-throughput method followed by sequencing. We observed no difference in the frequency of such variants in the two groups. Furthermore, we did not observe an association of rare coding variants in *SLC6A4* with rigid-compulsive traits scores in the cases. These results do not support a significant role for rare coding variants in *SLC6A4* in autism spectrum disorders, nor do they support a significant role for *SLC6A4* in rigid-compulsive traits in these disorders.

Introduction

Disrupted serotonin (5-hydroxytryptamine, 5-HT) signaling has been implicated in autism (OMIM 209850) and the serotonin transporter (*SLC6A4* or SERT, OMIM 182138) is a major determinant of 5-HT signaling 1. Prior studies have provided conflicting evidence for a genetic association between common *SLC6A4* variants and autism 2–8. Recently, it has been proposed that allelic heterogeneity, involving multiple rare alleles within *SLC6A4*, may contribute to susceptibility to autism 9. Moreover, it has been argued that the relevant rare variants are functional variants 10. These latter studies focused on rare variants found in families showing linkage to the 17q11.2 region that harbors the *SLC6A4* gene and that had only male cases, and further concluded that the presence of rare variants in *SLC6A4* were also associated with rigid-compulsive behaviors. In order to explore whether rare variants are overrepresented more generally in subjects with autism, we made use of data from a large multiplexed variation scan (MVS 11), in which we screened coding exons in the

Correspondence should be addressed to: Dr. Joseph D. Buxbaum Department of Psychiatry Mount Sinai School of Medicine One Gustave L. Levy Place, Box 1668 New York, New York 10029 Tel: 212-659-8862 Fax: 212-828-4221 e-mail: joseph.buxbaum@mssm.edu.

⁷Current address: NextBio Inc, Cupertino, CA

SLC6A4 gene for variant amplicons in unrelated control samples and unrelated samples from patients with autism or borderline autism (~420 controls and ~350 cases). By taking advantage of a high-throughput approach we were able to address the more general hypothesis as to whether such rare variants were associated with autism spectrum disorders in a group not selected based on possible linkage to 17q11.2 and to further examine the relationship of these variants with rigid-compulsive traits and gender.

Subjects

The cases were comprised of 377 unrelated individuals from multiplex families and 25 individuals from simplex families. Eighty-three percent of the cases were male. Three hundred ninety three (393) of the cases had an ADI-R derived diagnosis of autism, 3 had an ADI-R derived diagnosis of “not-quite-autism” (NQA, developed by Dr. C. Lord and also called borderline autism in some studies), and 6 had an autism spectrum diagnosis as defined in the AGRE website. For the purposes of the current study, we refer to the cases as having an autism spectrum disorder (ASD). Ethnicity was primarily Caucasian with 10% of unknown ethnicity, and amongst the cases 10% self-identified as Hispanic or Latino (this question was not asked of controls). The controls (n=469, all of which were Caucasian) were obtained from multiple sources including the Coriell Cell Repositories (Coriell Institute for Medical Research, Camden, NJ), from Caucasian volunteers in Texas, and from volunteers at Stanford University, Stanford, CA. Fifty two percent of the controls were male.

Methods

MVS depends on a very sensitive bacterial selection assay in which any mismatches of a test sequence (compared to a reference sequence) activate the mismatch repair detection (MRD) system in bacteria, which can be used to alter the ability of bacteria to survive on selectable media. This approach, which is described fully in Faham et al. 11, allows one to screen thousands of amplicons for variants in a fully-multiplexed manner.

Note that the overall concordance between MVS and dideoxy sequence, using >1,031 traces, was ca. 99.2% (see Faham et al. 11), which represented a concordance of ~95-96% when MVS identified a variant amplicon and ~99.6% when MVS identified a non-variant amplicon. This indicates that there are modest levels of potentially false-positive calls in MVS and negligible levels of false-negative calls. However, since there is no mechanism by which false-positive or false-negative calls would be higher or lower as a function of source of DNA and affected status, these modest rates will not affect our conclusions. In the current study, of the 40 potentially variant amplicons identified by MVS that were then sequenced, a proportion (5 samples) did not show a variant by sequencing. While higher than the previous estimates of concordance between MVS and dideoxy sequencing, this was not significant ($P>0.1$) and likely due to the relatively smaller numbers of amplicons sequenced here.

MVS was followed by direct sequencing in available samples for each rare, variant amplicon observed in subjects with autism (see Supplemental Table S1 for primers and PCR conditions). In all cases, both strands were sequenced.

Results

We first looked at the total number of variant amplicons identified by MVS directed against *SLC6A4* exons and flanking sequence (Table 1). Looking at the total number of variant amplicons observed in samples derived from controls as compare to those derived from subjects with ASD, 35 variant amplicons were detected amongst the >400 controls, while 25 variant amplicons were detected amongst the >350 patients screened (Table 1). Thus there

was no evidence for increased total numbers of rare variants in *SLC6A4* exons and flanking regions in autism ascertained solely on the basis of affected status ($P>0.5$).

In the case of one exon, exon 7, there was evidence from MVS for a nominally significant (Fisher exact test, $P=0.047$) increase in variant amplicons amongst subjects with ASD, as compared to controls (0 variant amplicons out of 452 in controls versus 4 out of 396, or 1.01%, in cases, all male). Sequencing demonstrated that all of the subjects were heterozygous for a synonymous variant in Gly308 (GGC to GGT) (Table 2). Segregation analysis did not show a clear segregation of the synonymous Gly308 variant with ASD (Figure 1A, families 1-4).

To further explore the allelic heterogeneity in *SLC6A4*, we subsequently sequenced available samples for each variant amplicon observed in subjects with ASD, along with available controls, and the results are summarized in Table 2 (some samples, predominantly controls, from the original study were not available for sequencing). In exon 3 (the first coding exon), MVS identified a variant in 10 controls and 7 patients. Sequencing of 8 samples (3 male and 1 female control and 4 male cases) demonstrated that in all 8 samples, as we expected, the variant corresponded to the known Gly56Ala variant. Three of the 8 samples, including 1 male case and 2 male controls, were found to be homozygous for the variant allele. Although the families for these 3 samples did not report a unique genetic or ethnic background, the presence of homozygotes for such a rare variant is consistent with such a unique genetic background for these samples. In the case homozygous for this variant, we confirmed that both parents were heterozygous for the variant, and determined that two brothers who were diagnosed with autism were heterozygous for the variant (Figure 1B, family 5). Note that none of cases that have the Gly56Ala variant overlapped with ~100 AGRE family samples sequenced previously by Sutcliffe et al 9, but there are likely some overlap with other AGRE samples that Sutcliffe et al genotyped for Gly56Ala.

In exon 5, a single variant amplicon was detected amongst the patients sampled. The variant amplicon was sequenced and observed to be heterozygous for a novel Asn211Ser change. An analysis of the patient's family (Figure 1C, family 6), demonstrated that the male patient (with a diagnosis of autism) and his affected brother (with a diagnosis of borderline autism) both carried the variant, which was paternally derived. This patient and the brother were both also heterozygous for the Lys605Asn variant that was also paternally derived (see below).

In exon 6, MVS identified a variant in 2 controls and in 2 cases. Among the cases, one was identified as being heterozygous for a novel Val274Ile change, while the other was heterozygous for a synonymous change in His235 (CAC to CAT). The synonymous His235 variant was also identified in a control sample. The other control did not show any variations by sequencing. The patient with the Val274Ile variant (male) derived the variant from his father, and had an affected sister that did not carry the variant (Figure 1, family 7).

In exon 8, sequencing did not identify a variant in the affected case, while in exon 9, sequencing of the male case identified a synonymous variant in Val382 (CTC to CTA), which was also identified in one of two controls (male) sequenced. The second control (female) carried a novel Gln359His variant. In exon 10, a case (male) carried the Ile425Leu, a variant previously reported in autism 9. At the same position, a different variant (Ile425Val) has been reported to be associated with a complex neuropsychiatric phenotype, which could include ASD 12,13. Analysis of the patient's family demonstrated that the maternally derived Ile425Leu variant was also found in 3 unaffected female siblings and in an affected brother (Figure 1, family 8). This is the same family reported by Sutcliffe et al 9. The variant identified in the control was a synonymous variant in Thr409 (ACG to ACA).

One patient sample was no longer available for exon 11. We sequenced three patients and identified a synonymous variant in Phe465 (TTC to TTT) in two of them and a novel Phe474Leu variant in one, but not the previously reported Phe465Leu variant 9. In the case of Phe474Leu variant, both the patient (male) and an affected brother have this variant, which is of maternal origin (Figure 1, family 9). The Phe465 variant was found in 2 controls, and a third control demonstrated a synonymous variant in His456 (CAC to CAT). The fourth control did not show any variations.

Both patients with exon 12 variants were found to have a change in the flanking sequence, specifically a G to C change at the 33rd nucleotide upstream of the start of exon 12, which was also observed in one of the controls. An additional control demonstrated a novel Ala511Val variant in this exon, while the other control did not show any variations by sequencing. Finally, sequencing of exon 14 identified in one male patient the Lys605Asn variant previously reported in autism 9, which was also found in 2 controls (1 male and 1 female). The patient also was heterozygous for the Asn211Ser variant (Figure 1, family 6). For the additional patient, we did not identify any variants, while in the third control, we identified a change in the noncoding, flanking sequence, specifically a C to T change at the 4th nucleotide from the end of the exon 14.

Discussion

Looking at the total number of variant amplicons observed in samples derived from controls as compared to those derived from subjects with autism, there was no evidence for increased total numbers of rare variants in *SLC6A4* exons and flanking regions in autism ascertained solely on the basis of affected status ($P>0.5$).

Of the 21 patients sequenced, 9 had amino acid changes, as compared to 8 out of 19 of the controls ($P=0.93$). These results are not consistent with increased levels of *SLC6A4* coding variants in autism. Of the 9 coding variants found in cases, one (Phe474Leu) was identified as “possibly deleterious” by SIFT (a web-based program that predicts whether an amino acid substitution affects protein function based on sequence homology and the physical properties of amino acids, <http://blocks.fhcrc.org/sift/SIFT.html>), while two (Gln359His and Ala511Val) of the 8 variants identified in controls were identified as “possibly deleterious” (note that “possibly deleterious” may be associated with either a gain or loss of function of *SLC6A4*). These results are not consistent with increased levels of “possibly deleterious” *SLC6A4* coding variants in autism (see below).

It has been suggested that apparently positive findings for rare variants can be artifacts due to population stratification (e.g., Keen-Kim et al. 14). However, as the current findings are all negative this is not the issue here. At the same time, it might be argued that very significant stratification may under certain scenarios reduce power to detect rare variants. Previously, using unlinked genetic markers 15 in the exact same set of cases and controls used here, we assessed stratification between the two groups 11. We estimated the average χ^2 for such a genome control as 1.3, with the 95% upper confidence limit at 1.56. This implies only very modest evidence for stratification. In addition, as the overwhelming majority of cases were Caucasian, we also analyzed Caucasians separately, with very little loss of power. Not surprisingly, there was no evidence for an increase or decrease of *SLC6A4* variants in autism in the Caucasian subgroup (data not shown).

In the previous study by Sutcliffe et al., which first proposed multiple rare alleles in *SLC6A4* contributing to risk for autism, four findings were presented. First, the strong linkage to 17q11.2 was primarily driven by families with only male affecteds, but association with common variants of *SLC6A4* locus failed to explain the linkage. Second, from a subset of

the families in the study, chosen based on contribution to linkage at 17q11.2, several rare variants in the *SLC6A4* gene were found in coding (at highly conserved positions) and non-coding regions. Third, ADI cluster score analysis showed that the coding variants were associated with rigid-compulsive behaviors. And fourth, in this and subsequent studies, some of coding variants were determined to have functional consequences. Based on these results, the authors proposed that the *SLC6A4* locus is an autism susceptibility gene, with variants predisposing to rigid-compulsive traits 9. In the current study, we took advantage of high-throughput methods, looking at cases and controls, to ask whether rare coding variants of the *SLC6A4* gene are associated with autism, irrespective of linkage to 17q (note that ~100 AGRE families sequenced previously by Sutcliffe et al 9 were included here). The results do not provide evidence for a significant increases in numbers of exon variants in *SLC6A4* in autism, whether all exon/flanking variants, just coding variants, or just “potentially deleterious” coding variants are considered. Certain functional coding variants that have been identified previously in autism studies, including Gly56Ala and Lys605Asn, were also found in unscreened controls in the current study, consistent with public databases of controls (dbSNP heterozygosities of 0.016-0.060). It therefore is unlikely that these variants are genetically associated with autism as such. As the study by Sutcliffe et al focused on male-only families strongly linked to 17q11.2, *SLC6A4* rare variants may be specifically associated with subtypes of autism linked to 17q11.2.

The Sutcliffe et al. study showed that linkage to 17q increased when considering male-only families and that rare variants segregated with autism in males 9. For several of the rare coding variants we found in the current study, we observed non-significant apparent male biases which are of course highly confounded by the increased proportion of males affected with ASD (e.g., family 5 for Gly56Ala, family 6 for Asn211Ser and Lys605Asn, and family 9 for Phe474Leu; note that family 8 is the same family reported in Sutcliffe et al. and should not be regarded as additional evidence 9). However, we also found Gly56Ala males in unscreened controls (including homozygotes). In the case of Gly308, 1 affected and 3 unaffected females, and 4 affected and 5 unaffected males carried this variant within the families, again too small a number to draw any conclusions. As we had approximately equal numbers of males and female controls, we could explore the hypothesis as to whether female gender might be protective against detrimental effects of rare variants, which might be presented as an increased proportion of rare variants in female controls, as compared to males, but we found no evidence of any such discrepancy (Fisher's exact test, $P > 0.4$). Further studies aimed at determining whether there are *SLC6A4* genotype by gender interactions would require a detailed study with far greater number of males and females with these variants.

We identified several novel additional coding variants, including Asn211Ser, Val274Ile, and Phe474Leu in patients, and Gln359His and Ala511Val in controls. Of the three we identified in patients, two are found in both affected individuals in the sibship (Asn211Ser and Phe474Leu, all males), while the other was not (Val274Ile). The significance of these variants awaits further analyses, including functional studies. Furthermore, we identified a single, synonymous variant Gly308, which showed nominally significant increases in cases, as compared to controls. Whether this variant regulates expression of *SLC6A4* remains to be determined.

As it was reported that rare variants in *SLC6A4* are associated with increased rigid-compulsive behaviors 9·16·17 13·18, we analyzed our cases with rare variants to see if they have increased rigid-compulsive behaviors compared to cases without such variants, using the same factors used previously (Table 3). Two analyses were carried out. First, using a non-parametric ranking test (as our data for the rigid-compulsive factor was not normally distributed), we determined the P-value associated with each patient with a rare variant. All

such individuals showed rigid-compulsive traits consistent with the study population as a whole. We also compare the means of the two groups (those with and those without rare variants), and observed no difference in the means (one-tailed Mann-Whitney U-test, $P=0.44$; two-tailed $P=0.88$). No significant differences in the means were observed for other factors (Table 3).

In summary, we screened for rare variants in *SLC6A4* in a large sample ($n\sim 770$) using MVS, followed by sequencing. This represents a comprehensive study, with 0.95 power to detect any variant with a frequency of $> 1/250$. Even if we postulate that there are rare variants only found in patients, we have 0.95 power to detect any variant with a frequency $> 1/125$ in the patients, and 0.80 power to detect any variant with a frequency of $1/230$. Given the power of the current study and the negative results, we conclude that the results do not support a significant role for rare coding variants in *SLC6A4* in autism irrespective of whether one considers exon/flanking variants, just coding variants, or just “potentially deleterious” coding variants.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

TS is a Beatrice and Samuel A. Seaver Fellow and is supported by a New York State Spinal Cord Injury grant and a grant from the Stanley Medical Research Institute. This study was supported by the Beatrice and Samuel A. Seaver Foundation, by the National Institutes of Health through a Studies To Advance Autism Research and Treatment grant (MH-066673) and (NS-042165), and by the Autism Genome Project through Autism Speaks (JDB). We gratefully acknowledge the resources provided by the Autism Genetic Resource Exchange (AGRE) Consortium and the participating AGRE families. The Autism Genetic Resource Exchange is a program of Autism Speaks and is supported, in part, by grant 1U24MH081810 from the National Institute of Mental Health to Clara M. Lajonchere (PI).

References

1. Cook EH, Leventhal BL. The serotonin system in autism. *Curr Opin Pediatr*. 1996; 8:348–354. [PubMed: 9053096]
2. Kim SJ, Cox N, Courchesne R, Lord C, Corsello C, Akshoomoff N, Guter S, Leventhal BL, Courchesne E, Cook EH Jr. Transmission disequilibrium mapping at the serotonin transporter gene (*SLC6A4*) region in autistic disorder. *Mol Psychiatry*. 2002; 7:278–288. [PubMed: 11920155]
3. Conroy J, Meally E, Kearney G, Fitzgerald M, Gill M, Gallagher L. Serotonin transporter gene and autism: a haplotype analysis in an Irish autistic population. *Mol Psychiatry*. 2004; 9:587–593. [PubMed: 14708029]
4. McCauley JL, Olson LM, Dowd M, Amin T, Steele A, Blakely RD, Folstein SE, Haines JL, Sutcliffe JS. Linkage and association analysis at the serotonin transporter (*SLC6A4*) locus in a rigid-compulsive subset of autism. *Am J Med Genet B Neuropsychiatr Genet*. 2004; 127:104–112. [PubMed: 15108191]
5. Devlin B, Cook EH Jr, Coon H, Dawson G, Grigorenko EL, McMahon W, Minshew N, Pauls D, Smith M, Spence MA, Rodier PM, Stodgell C, Schellenberg GD. Autism and the serotonin transporter: the long and short of it. *Mol Psychiatry*. 2005; 10:1110–1116. [PubMed: 16103890]
6. Ramoz N, Reichert JG, Corwin TE, Smith CJ, Silverman JM, Hollander E, Buxbaum JD. Lack of evidence for association of the serotonin transporter gene *SLC6A4* with autism. *Biol Psychiatry*. 2006; 60:186–191. [PubMed: 16616719]
7. Yonan AL, Palmer AA, Gilliam TC. Hardy-Weinberg disequilibrium identified genotyping error of the serotonin transporter (*SLC6A4*) promoter polymorphism. *Psychiatr Genet*. 2006; 16:31–34. [PubMed: 16395127]

8. Brune CW, Kim SJ, Salt J, Leventhal BL, Lord C, Cook EH Jr. 5-HTTLPR Genotype-Specific Phenotype in Children and Adolescents With Autism. *Am J Psychiatry*. 2006; 163:2148–2156. [PubMed: 17151167]
9. Sutcliffe JS, Delahanty RJ, Prasad HC, McCauley JL, Han Q, Jiang L, Li C, Folstein SE, Blakely RD. Allelic heterogeneity at the serotonin transporter locus (SLC6A4) confers susceptibility to autism and rigid-compulsive behaviors. *Am J Hum Genet*. 2005; 77:265–279. [PubMed: 15995945]
10. Prasad HC, Zhu CB, McCauley JL, Samuvel DJ, Ramamoorthy S, Shelton RC, Hewlett WA, Sutcliffe JS, Blakely RD. Human serotonin transporter variants display altered sensitivity to protein kinase G and p38 mitogen-activated protein kinase. *Proc Natl Acad Sci U S A*. 2005; 102:11545–11550. [PubMed: 16055563]
11. Faham M, Zheng J, Moorhead M, Fakhrai-Rad H, Namsaraev E, Wong K, Wang Z, Chow SG, Lee L, Suyenaga K, Reichert J, Boudreau A, Eberle J, Bruckner C, Jain M, Karlin-Neumann G, Jones HB, Willis TD, Buxbaum JD, Davis RW. Multiplexed variation scanning for 1,000 amplicons in hundreds of patients using mismatch repair detection (MRD) on tag arrays. *Proc Natl Acad Sci U S A*. 2005; 102:14717–14722. [PubMed: 16203980]
12. Ozaki N, Goldman D, Kaye WH, Plotnicov K, Greenberg BD, Lappalainen J, Rudnick G, Murphy DL. Serotonin transporter missense mutation associated with a complex neuropsychiatric phenotype. *Mol Psychiatry*. 2003; 8:933–936. [PubMed: 14593431]
13. Delorme R, Betancur C, Wagner M, Krebs MO, Gorwood P, Pearl P, Nygren G, Durand CM, Buhtz F, Pickering P, Melke J, Ruhrmann S, Anckarsater H, Chabane N, Kipman A, Reck C, Millet B, Roy I, Mouren-Simeoni MC, Maier W, Rastam M, Gillberg C, Leboyer M, Bourgeron T. Support for the association between the rare functional variant I425V of the serotonin transporter gene and susceptibility to obsessive compulsive disorder. *Mol Psychiatry*. 2005; 10:1059–1061. [PubMed: 16088327]
14. Keen-Kim D, Mathews CA, Reus VI, Lowe TL, Herrera LD, Budman CL, Gross-Tsur V, Pulver AE, Bruun RD, Erenberg G, Naarden A, Sabatti C, Freimer NB. Overrepresentation of rare variants in a specific ethnic group may confuse interpretation of association analyses. *Hum Mol Genet*. 2006; 15:3324–3328. [PubMed: 17035247]
15. Reich DE, Goldstein DB. Detecting association in a case-control study while correcting for population stratification. *Genet Epidemiol*. 2001; 20:4–16. [PubMed: 11119293]
16. Wendland JR, Kruse MR, Murphy DL. Functional SLITRK1 var321, varCDfs and SLC6A4 G56A variants and susceptibility to obsessive-compulsive disorder. *Mol Psychiatry*. 2006; 11:802–804. [PubMed: 16936762]
17. Grados MA, Samuels J, Shugart YY, Willour VL, Wang Y, Cullen B, Bienvenu OJ, Hoehn-Saric R, Valle D, Liang KY, Riddle MA, Wendland JR, Murphy DL, Nestadt G, Detera-Wadleigh S. Rare plus common SERT variants in obsessive-compulsive disorder. *Mol Psychiatry*. 2007; 12:422–423. [PubMed: 17453059]
18. Wendland JR, DeGuzman TB, McMahon F, Rudnick G, Detera-Wadleigh SD, Murphy DL. SERT Ileu425Val in autism, Asperger syndrome and obsessive-compulsive disorder. *Psychiatr Genet*. 2008; 18:31–39. [PubMed: 18197083]

A

Family1, Gly308

Family 2, Gly308

Family 3, Gly308

Family 4, Gly308

B

Family 5, Gly56Ala

C

Figure 1. Segregation of *SLCA4* variants in families

Families 1-4 (A) carry the synonymous Gly308 variant, family 5 (B) includes a Gly56Ala/Gly56Ala homozygote, while (C) family 6 has Asn211Ser and Lys605Asn variants, family 7 has the Val274Ile variant, family 8 has the Ile425Leu variants, and family 9 has the Phe474Leu variant, respectively. Filled symbols represent affected individuals, open symbols represent unaffected individuals with the wild-type allele, and open symbols with dots inside represent unaffected individuals with variant alleles. The open symbol with thick borders (family 5) represents a subject with borderline autism. Squares and circles represent males and females, respectively.

Table 1

Frequency of common and variant amplicons in coding exons of the SLC6A4 gene in controls and subjects with ASD.

Exon	Control		ASD	
	Count of individuals with common amplicons	Count of individuals with variant amplicons	Count of individuals with common amplicons	Count of individuals with variant amplicons
3	441	10 (4)	389	7 (4)
4	421	4	366	0
5	451	0	397	1 (1)
6	417	2 (2)	377	2 (2)
7	452	0	392	4 (4)
8	424	1	333	1 (1)
9	448	4 (2)	397	1 (1)
10	452	1 (1)	397	1 (1)
11	433	5 (4)	374	4 (3)
12	430	3 (3)	386	2 (2)
13	450	2	397	0
14	444	3 (3)	392	2 (2)
Total		35 (19)		25 (21)

We use exon numbering according to NCBI. Exon 1 and 2 are non-coding exons. Note that exon 15, which codes for the last 24 amino acids, was not successfully screened. The numbers in parentheses show the number of DNA samples available and sequenced in the further aspects of this study.

Table 2

Characterization of rare variants of the SLC6A4 gene identified by MVS.

Exon	Nucleotide changes	Molecular nature (conserved among species)	Controls identified (males, females)	Cases identified (males, females)	rs#	Heterozygosity reported in dbSNP, build 127
3	G/C	Gly56Ala (hs and mm only)	4 (3, 1)	4 (4, 0)	rs6355	0.016
5	A/G	Asn211Ser (no)	0 (0, 0)	1 (1, 0)		
6	C/T	Synonymous His235	1 (1, 0)	1 (1, 0)	rs41274284	Not determined
	G/A	Val274Ile (yes)	0 (0, 0)	1 (1, 0)		
7	C/T	Synonymous Gly308	0 (0, 0)	4 (4, 0)		
9	A/C	Gln359His (yes)	1 (0, 1)	0 (0, 0)		
	C/T	Synonymous Val382	1 (1, 0)	1 (1, 0)		
10	G/A	Synonymous Thr409	1 (0, 1)	0 (0, 0)		
	C/A	Ile425Leu (yes)	0 (0, 0)	1 (1, 0)	rs3275367	0.017
11	C/T	Synonymous His456	1 (1, 0)	0 (0, 0)		
	C/T	Synonymous Phe465	2 (1, 1)	2 (2, 0)	rs41274280	Not determined
	C/T	Phe474Leu (yes)	0 (0, 0)	1 (1, 0)		
12	G/C	Flanking region	1 (0, 1)	2 (1, 1)		
	C/T	Ala511Val (no)	1 (1, 0)	0 (0, 0)		
14	C/A	Lys605Asn (yes)	2 (1, 1)	1 (1, 0)	rs6352	0.060
	C/T	Flanking region	1 (0, 1)	0 (0, 0)		

We compared human (hs), mouse (mm), chick, worm and fly for conservation of amino acid sequence among species. Whether the sequence is conserved across human, mouse, and chick, is noted (and some of them are also conserved in worm and/or fly). If the sequence is conserved only in human and mouse it is noted as well. Note that: 1 case and 2 controls are 56Ala homozygotes; 1 case and 1 control have both His235 and Phe465 variants; 2 case and 1 control have both Phe465 and exon 11 flanking region variants; and 1 case has both Asn211Ser and Lys605Asn variants (ethnicity unknown). AGRE ID for cases are listed in Supplemental Table S2.

Table 3

ADI cluster Scores for individuals with SLC6A4 coding variant

Exon	Variant	Gender	diagnosis	Spoken Language	Social Intent	Compulsions	Ranking P=	Developmental Milestones	Savant Skills	Sensory Aversion
3	Gly56Ala homo	male	Autism	0.27	0.32	0.35	0.30	0.74	0.13	0.17
3	Gly56Ala	male	Autism	0.83	0.65	0.36	0.28	0.68	0.00	0.50
3	Gly56Ala	male	Autism	0.84	0.29	0.11	0.81	0.79	0.11	0.69
3	Gly56Ala	male	Autism/S spectrum	0.12	0.61	0.10	0.83	0.71	0.17	0.73
6	Val274Ile	male	Autism	0.17	0.44	0.26	0.48	0.78	0.30	0.28
7	Gly308	male	Autism/S spectrum	0.81	0.38	0.25	0.52	0.74	0.22	0.25
7	Gly308	male	Autism evaluated by 2003ADIR which uses different algorithm							
7	Gly308	male	Autism	0.88	0.86	0.41	0.21	0.73	0.11	0.58
7	Gly308	male	Autism	0.84	0.75	0.00	0.99	0.67	0.00	0.52
9	Val382	male	Autism	0.43	0.52	0.35	0.30	0.76	0.25	0.00
10	Ile425Leu	male	Autism	0.80	0.64	0.32	0.34	0.78	0.00	0.14
11	Phe474Leu	male	Autism	0.32	0.35	0.39	0.24	0.75	0.00	0.43
	G/C/exon11 Phe465/exo n6 His235	male	ADI data removed from database at families request							
12	G/C/exon11 Phe465	male	Autism	0.22	0.26	0.37	0.27	0.74	0.07	0.27
	Lys605Asn/exon5									
14/5	Asn211Ser	male	Autism	0.14	0.34	0.07	0.89	0.70	0.11	0.29
	U-test P value			0.96	0.63	0.90		0.59	0.66	0.49

A ranking P value for the Compulsions factor was determined for each case by the non-parametric ranking test. In addition, a U-test P value was derived for each factor, comparing groups with and those without variants.