Title

Association of SLC6A4 variants with obsessive-compulsive disorder in a large multicenter US family study

Journal

Molecular Psychiatry, 16(1)

ISSN

1359-4184

Authors

Voyiaziakis, E
Evgrafov, O
Li, D
et al.

Publication Date

2011

DOI

10.1038/mp.2009.100

Peer reviewed

Published in final edited form as:

Mol Psychiatry. 2011 January ; 16(1): 108–120. doi:10.1038/mp.2009.100.

Association of *SLC6A4* variants with obsessive-compulsive disorder in a large multi-center US family study

Emanuel Voyiaziakis[1], **Oleg Evgrafov**[1], **Dalin Li**[2], **Hee-Jae Yoon**[3], **Paula Tabares**[4], **Jack Samuels**[5], **Ying Wang**[5], **Mark A. Riddle**[5], **Marco A. Grados**[5], **O. Joseph Bienvenu**[5], **Yin Yao Shugart**[6], **Kung-Yee Liang**[7], **Benjamin D. Greenberg**[8], **Steven A. Rasmussen**[8], **Dennis L. Murphy**[9], **Jens R. Wendland**[9], **James T. McCracken**[10], **John Piacentini**[10], **Scott L. Rauch**[11], **David L. Pauls**[12], **Gerald Nestadt**[5], **Abby J. Fyer**[4], and **James A. Knowles**[1,3]

[1]Department of Psychiatry & the Behavioral Sciences, Keck School of Medicine, University of Southern California, MC-2821, 1501 San Pablo, ZNI 401, Los Angeles, CA 90089

[2]Department of Preventive Medicine, Division of Biostatistics, Keck School of Medicine, University of Southern California, 1501 San Pablo, ZNI 437, Los Angeles, CA 90089

[3]Zilkha Neurogenetics Institute, Keck School of Medicine, University of Southern California, 1501 San Pablo, ZNI 401, Los Angeles, CA 90089

[4]Department of Psychiatry, College of Physicians and Surgeons at Columbia University and the New York State Psychiatric Institute, New York, NY 10032

[5]Department of Psychiatry and Behavioral Sciences, Johns Hopkins University, School of Medicine, Baltimore, MD 21287

[6]Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD 21287

[7]Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD 21287

[8]Department of Psychiatry and Human Behavior, Brown Medical School, Butler Hospital, Providence, RI 02906

[9]Laboratory of Clinical Science, NIMH, NIH, Bethesda, MD

[10]Department of Psychiatry and Biobehavioral Sciences, University of California, Los Angeles, School of Medicine, Los Angeles, CA 90024

[11]Department of Psychiatry, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114

[12]Psychiatric and Neurodevelopmental Genetics Unit, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114

Corresponding author: Emanuel Voyiaziakis, MD, MS; ph: 323-442-1167, fax: 323-442-2448, emanuelv@keck.usc.edu.

Supplementary figures and tables are available at the *Molecular Psychiatry* website.

Abstract

Genetic association studies of *SLC6A4* (*SERT*) and OCD have been equivocal. We genotyped 1 241 persons in 278 pedigrees from the OCD Collaborative Genetics Study for 13 SNPs, the LPR indel with molecular haplotypes at rs25531, VNTR polymorphisms in introns 2 and 7, and a 381-bp deletion 3' to the LPR. We analyzed using the Family Based Association Test (FBAT) under additive, dominant, recessive, and genotypic models, using both OCD and sex-stratified OCD as phenotype.

Two-point FBAT analysis detected association with Int2 (p= 0.0089) and Int7 (p= 0.0187) (genotypic model). Sex-stratified 2-point analysis showed strong association in females with Int2 (p< 0.0002), significant after correction for LD, multiple marker- and model testing (p_{Adj}= 0.0069). The *SLC6A4* gene is composed of two haplotype blocks (our data and the HapMap); FBAT whole-marker analysis using this structure was not significant.

Several noteworthy non-significant results emerged. We found no evidence for over-transmission of the LPR L_A allele (GRR= 1.11, 95% CI: 0.77–1.60) (see Hu et al.[1]), however rare individual haplotypes containing L_A with p<0.05 were observed. Similarly, 3 individuals (2 with OCD/OCPD) carried the rare I425V *SLC6A4* variant- but none passed it on to their 6 OCD-affected offspring, suggesting it is unlikely solely responsible for the "OCD plus syndrome" (see Ozaki et al.[2]).

In conclusion, we found evidence of genetic association at the *SLC6A4* locus with OCD. Noteworthy lack of association at the LPR, LPR-rs25531, and rare 425V variants suggests hypotheses about OCD risk need revision to accommodate these new findings, including a possible gender effect.

Keywords

genetics; OCD; family study; affected sib pair study; molecular haplotype analysis; 5-HT; serotonin transporter; heterogeneity analysis

Introduction

Twin and family studies suggest that obsessive-compulsive disorder (OCD) has a genetic basis, but a major causative gene(s) for the disorder has not yet been identified. Segregation analyses suggest that OCD is transmitted as either an autosomal dominant or codominant trait, particularly in those with disease onset before adulthood. (3–5) The first linkage study of OCD found evidence for suggestive linkage (LOD=2.25) to chromosome *9p24* in 56 individuals from 7 pedigrees at marker D9S288. (6) Additional evidence of linkage to this locus (HLOD=2.26) was observed in a sample of 50 pedigrees with markers D9S1813 and D9S1792 which are located within 0.5 cM (<350 kb) of D9S288 (7) and two of the markers tested gave modest evidence of association (p= 0.046 and 0.02). The gene encoding the neuronal glutamate transporter *SLC1A1* is located in the *9p24* linkage region and has been found to have genetic association with OCD by several groups (8–10, 69). It is likely that there are additional genomic loci for OCD as a substantially larger genome-wide linkage

scan observed suggestive linkage to chromosomes *3q, 7p, 15q, 6q,* and *1q*, although it did not strongly support linkage to *9p24*. (11)

The gene for another transporter of neurochemical importance, *SLC6A4* (*serotonin transporter*, *5-HTT,* or *SERT*), located on chromosome *17q11.1-q12*, has also been studied extensively for genetic association to OCD. *SLC6A4* is the molecular target of the selective serotonin reuptake inhibitors (SSRIs), e.g., fluoxetine, fluvoxamine, sertraline, paroxetine, and citalopram. SSRIs collectively represent the most clinically effective and widely studied pharmacological treatment for OCD; they are effective in reducing both the thought (obsessions) and behavioral (compulsions) components of OCD. (12) Imaging studies have suggested that individuals with OCD have decreased serotonin transporter availability in the midbrain/brainstem. (13–15)

A large amount of genetic variation has been observed in *SLC6A4* including both rare and common SNPs (16), repeat polymorphisms in introns 2 and 7, and an extensively studied functional 44-base pair indel polymorphism in the promoter region (5-HTTLPR). Both the 5-HTTLPR and the intron 2 polymorphism have been shown to affect transcription of *SLC6A4*. (17–18) While the initial reports suggested association between the LPR and OCD (19–20), most subsequent studies have been negative. (21–28) More recently, Dickel and colleagues (10) failed to detect association with OCD when examining select polymorphisms in *SLC6A4*, including the LPR. They carried out a meta-analysis constituted of five prior reports consisting of both positively and negatively associated findings that did not support a combined-sex association for either the long (L) or short (S) allele, after correction. However, they found nominally significant excess transmission of the L allele in females but not males. Lin (29) expanded the scope of investigation by meta-analysis of association findings of the LPR and OCD by pooling results from 13 independent case-control association studies with 3445 subjects (1 242 OCD patients and 2 203 controls). All 13 studies used an OR (odds ratio) estimator of risk, and random effects modeling showed OCD associated with the SS genotype and inversely associated with the LS genotype (OR=1.21 and p=0.04 vs. OR=0.79 and p=0.03, respectively). No association was detected for the LL genotype nor solely the allelic L variant.

Most recently, further meta-analysis (18) analyzed a larger set of case-control and, additionally, family based association findings. No evidence of association with variation at the LPR locus and OCD was detected in the overall meta-analysis. However, stratified meta-analysis demonstrated significant association between the L allele and OCD in family-based association testing and in studies involving children (child-onset OCD) and Caucasians. However, the authors note that no adjustments were made to the significance threshold for multiple comparisons.

The simple model of the 5HTTLPR having a long allele (L) with higher expression and a lower-expressing short (S) allele has been refined with the discovery of additional nearby variation that influences transcription. Hu and colleagues (1) showed that a G allele at rs25531 located within the LPR on an L background (designated L_G) creates a functional *AP2* transcription-factor binding site and has a serotonin reuptake activity of an S allele. This group then found that the gain of function L_A/L_A genotype was twice as common in

OCD cases (n=169) as in ethnically matched controls (n=263) and the L_A allele showed a 2-fold over-transmission in OCD-affected trios (n=175), leading to a 1.8-fold increased risk of OCD.

Rare genetic variation in *SLC6A4* also has been implicated in the pathogenesis of OCD. A SNP in exon 8 alters amino acid 425 from isoleucine to valine (I425V), resulting in a protein with an increase in V_{max} and decrease in KM. (30) This variation was observed to segregate with OCD in two OCD-dense families and all six of the individuals with 425V whom could be psychiatrically assessed had either OCD or OCPD. (2) Additional studies have found a prevalence of 425V of 1.5% in OCD cases (n=530) as compared to 0.2% in controls (n=1 300), with p= 0.004 and an OR= 6.54. (31)

Given the large (n=1 241), multi-center US family sample we have collected, we investigated comprehensively whether genetic association of OCD to multiple *SLC6A4* polymorphisms and haplotypes are observed using family-based statistical approaches.

Materials and Methods

Human Subjects.

The OCD Collaborative Genetics Study (OCGS) is an ongoing, NIMH-funded collaboration among investigators at seven sites in the US (Brown University, Columbia University, University of Southern California, Johns Hopkins University, Massachusetts General Hospital, University of California at Los Angeles, and the National Institute of Mental Health). The methods of the study are described in detail elsewhere (32) and summarized below.

The OCGS targeted families containing at least one sibling pair both affected with OCD and extended these pedigrees when possible through affected first- and second-degree relatives. Family history interviews were conducted to determine that there were at least two OCD-affected siblings in the family who were willing to participate and to identify additional affected relatives. All first- and second-degree relatives were considered for inclusion, and families were extended through the first-degree relatives of all affected cases.

To be considered affected, a participant had to meet DSM-IV OCD diagnostic criteria at any time in his or her life. Probands were included if, in addition to meeting DSM-IV criteria, their first onset of obsessions and/or compulsions occurred before 18 years of age. Probands with schizophrenia, severe mental retardation, Tourette's disorder (TD), or secondary OCD (OCD occurring exclusively in the context of depression) were excluded. Participants had to be at least 7 years old to participate in the study.

SNP genotyping.

13 SNPs were genotyped using SNPlex chemistry according to manufacturer's instructions (protocol P/N 4360858 Revision B) and genotype calls were made using Applied Biosytems (ABI, Foster City, CA) GeneMapper™ v4.0 Software. A SNP set was assembled using data from phase I of the HapMap Project (hapmap.org) and a "tag" set of SNPs was generated using the Tagger server (broad.mit.edu/mpg/tagger/index.php), using parameters of pairwise

tagging, r^2 of 0.99, and a MAF threshold of 1%. Additional non-synonymous SNPs were selected with preference given to those with reported allele frequencies. Overall genotyping completion rate for the 13 SNPlex-assayed SNPs was 98.5% (96.7–99.0%). The discordance rate for 48 randomly-plated blind duplicates was 3 genotypes in 624 total genotypes (0.5%).

381-Base pair deletion genomic DNA amplification.

Amplification was performed in a final volume of 5 μL containing 20 ng genomic DNA template, 200 uM dNTPs, 1.5 M betaine (Sigma-Aldrich, St. Louis, MO), 50 mM KCl, 20 mM Tris-HCl (pH 8.4), 2.5 mM $MgCl_2$, 200 nM primers (F-CTCCAGCATTCTCCTTGCAC and R-TGAGCCCAGGAATTCAAGAC (Invitrogen, Carlsbad, CA)), and 0.25 units Platinum Taq DNA polymerase (Invitrogen), then cycled at 95°C for 1 min, followed by 10 cycles at 95°C for 20 sec, 61°–56°C for 20 sec (decreased by 0.5°C intervals per cycle), and 72°C for 1 min, followed by 35 cycles at 95°C for 20 sec, 56°C for 20 sec, and 72°C for 1 min, with a final 30 min at 72°C on a DNA Engine Dyad PTC-220 thermal cycler (MJ Research, Waltham, MA). Amplified samples were run on 2% agarose in 1XTBE with EtBr (Cambrex Bio Science, Rockland, ME) and examined under UV fluorescence.

Microsatellite and LPR genotyping.

A PCR “triplex” for *SLC6A4* LPR, intron 2, and intron 7 was performed in 2.5 μL containing 25 ng genomic DNA template, 400 μM dNTPs, 1.5 mM 10X PCR Buffer IV (Abgene, Surrey, UK), 3.0 mM $MgCl_2$, 200 nM primers for the LPR (F-FAM-ATGCCAGCACCTAACCCCTAATGT; R-GGACCGCAAGGTGGGCGGGA), 200 nM primers for intron 2 (F-FAM-GTCAGTATCACAGGCTGCGAG and R-TGTTCCTAGTCTTACGCCAGTG), 100 nM primers for intron 7 (F-HEX-ACCGCACCCCGTCTCTCTCTTT and R-ACACCTGTAAGCACAGCCACTTG) (Invitrogen), and 1.0 unit Platinum Taq DNA polymerase (Invitrogen), then cycled according to the following protocol: at 96°C for 5 m followed by 35 cycles of 96°C for 45 s, 68°C for 30 s and 72°C for 1 m, followed by a final extension at 72°C for 10 m on a DNA Engine Dyad PTC-220 thermal cycler. Samples were then subjected to capillary electrophoresis on an ABI 3100 fragment analyzer (ABI) and the following size ranges were observed: 5-HTTLPR, 375 or 419 bp; Int2 VNTR, 260–310 bp; and Int7 GAAA, 263–288 bp. Microsatellite genotyping calls were made using Gene Marker v1.5 software (SoftGenetics, LLC, State College, PA). Forty-eight randomly distributed blind duplicates revealed no discordances. This assay achieved the following genotype completion rates: LPR, 99.6%; Int2, 98.2%; and Int7, 99.6%.

LPR-rs25531 (A/G) SNP variant.

Phase-certain haplotyping of the LPR variants L_A, L_G, S_A and the rare S_G allele was performed by a 2-step protocol: I. determination of the LPR L or S allele, as above; and, II. digestion of this amplicon with *MspI* restriction endonuclease. The assay was designed to include an invariant *MspI* digest site 94 bp from the end of the LPR amplicon to provide an internal control for digestion/partial digestion (OS figure 2). After separation of the digestion products by capillary electrophoresis and analysis using Gene Marker v1.5, the following restriction fragment allele sizes were obtained: S_A (281 bp), L_A (325 bp), S_G and

L_G (151 bp). When these data are combined with the results from the undigested fragment data, unambiguous haplotypes can be deduced (see legend to OS figure 2). All of the rare S_G alleles were confirmed by capillary DNA sequencing. The 48 randomly distributed blind duplicates were 100% concordant. Completion rate for the LPR-rs25531 genotyping (step II) was 96.5%.

Statistics.

Family-based association analysis.—We used FBAT and PDT (http://www.chg.duke.edu/software/pdt.html) (33–35) to run two-point SNP analyses under additive, dominant and genotypic models and for haplotype-based association analysis.

Corrections for multiple testing.—Two-point p-values from FBAT were corrected for multiple-marker testing in the presence of linkage disequilibrium (LD), and also multiple-model assessments, but not gender. The PACT (p-values adjusted for correlated tests) approach adjusts for multiple comparisons by calculating the null distribution of the test statistics through numerical integration of their asymptotic joint distribution. This approach has been shown to be reliable and highly efficient for adjustment of multiple-comparisons. (36)

Haplotype analysis and haplotype phasing.—We used Haploview 4.1 to assess LD structure at the *SLC6A4* locus using the OCGS genotype data (17 markers). We adopted 2 LD blocks defined by the solid spine approach (extension with D'> 0.95) that encompassed our set of polymorphic markers for all haplotype analysis in FBAT (OS Figure 3). The HapMap phase III dataset supported this LD block structure (OS Figure 4). Haplotype analysis was performed on only the Caucasian portion of the sample by removal of 47 individuals who reported non-Caucasian ethnicity, which included 5 entire pedigrees (see Table 4, pedigree ethnicity summary).

In our haplotype analysis we assigned the most likely estimated haplotype to each individual. Given that the haplotypes were estimated within haplotype blocks with tight LD structure for SNPs within a block, haplotype diversity was low with estimated haplotype posterior probabilities within each block greater than 92%. Thus, the method of assigning the most likely haplotype yields similar results to an approach using the expectation or haplotype dosage (37–39) The permutation method implemented in FBAT was used to compute haplotype whole-marker (minimal) p-values (biostat.harvard.edu/~fbat).

Genotype Relative Risk (GRR).—Conditional logistic analysis was carried out using the STATA package (http://www-gene.cimr.cam.ac.uk/clayton/software/stata).

Power calculations.—In separate calculations assessed in PBAT (http://www.goldenhelix.com/SNP_Variation/PBAT/pbat_methodology.html), the power to detect association assuming an additive model of inheritance with a disease allele frequency of 0.05 and a population prevalence of 0.03 using the 278 families (459 nuclear families) already collected is roughly 0.60 (if the D' between the disease locus and candidate SNP is 1.0).

Sex-stratified Analyses.—We stratified the dataset by the gender status of male or female OCD-affecteds in each family. We coded the affection status of either all female OCD-affecteds as “unknown”, or all male affecteds as “unknown,” and analyzed with FBAT. P-values for sex-stratified analyses were not corrected for the three hypotheses tested (all, male-only, female-only).

Results

We genotyped 278 pedigrees containing 459 nuclear families and 1 241 persons (>96% Caucasian pedigrees; see Table 4) with 14 SNPs, the LPR indel and two VNTR polymorphisms in introns 2 and 7 within *SLC6A4* to determine if they are in linkage disequilibrium (LD) with OCD (Table 1, OS Figure 1). Three SNPs, rs6353 (MAF= 0), I425V and rs140699 (MAFs= 0.003) with low minor allele frequencies (MAF) did not have a sufficient number of families to calculate 2-point results with FBAT and were also excluded from the haplotype analysis. In addition, we also failed to detect a 381-bp deletion variant immediately 3’ of the LPR (40) in 354 randomly selected samples from 102 families. All other SNP data were confirmed to be in Hardy-Weinberg equilibrium (HWE) and were used to determine the haplotype structure in the region, which was in agreement with the haplotype structure of *SLC6A4* observed with HapMap Project phase III data (OS Figures 3 & 4).

We performed two-point analyses using FBAT under additive, dominant, recessive and genotypic models and the results for the additive model are presented in Table 2, parts A and C. All allele-model combinations for the whole sample that yielded nominal p-values <0.05 are presented in Table 3. Specifically, under an additive model we observed nominal evidence of positive association with allele 5 of the microsatellite repeat in intron 7 (p= 0.0330; 14 informative families). Under a genotypic model, the intron 2 genotype containing alleles 9/10 showed positive association (p-value= 0.0089; 13 families), and the intron 7 genotype containing alleles 3/7 was inversely associated with OCD (i.e., reduced risk) (p= 0.0187; 105 families). Given that a number of correlated tests were performed and that significant LD existed within the set of markers (see OS Figure 3), we employed the method of Conneely and Boehnke (36) to correct jointly for multiple marker comparisons, in the setting of LD. On the full sample, none of the resulting 2-point analysis adjusted P-values (P_{Adj}) were <0.05 (Table 3). The results did not change appreciably when analyzed with the PDT.

The most notable negative results were at the LPR locus. We found no evidence of association to either the LPR polymorphism, or to the molecular haplotype of LPR-rs25531 with any of the genetic models (additive model data shown in Table 2, OS Table 1). We did observe 10 individuals with the S_G variant, all of which were confirmed by DNA sequencing. Two of these individuals had a LPR variant similar to that as reported previously (see variant 14B in (41)) (data not shown). We investigated whether these results were influenced by either genotypic (e.g., LPR allele frequencies) or phenotypic heterogeneity (e.g., % females, ethnicity, mean age of onset of symptoms) across the different clinical sites. No evidence of either genotypic or phenotypic heterogeneity was observed (Table 4).

Furthermore, we did not observe association of the LPR-rs25531 haplotype when each site was analyzed individually.

Three individuals with the rare 425V variant were detected and we examined their pedigrees in greater detail. In all three, the individual with 425V was the biological parent of an OCD affected sib-pair. Two of the carriers were fathers: one with OCD (1808–3), one with OCPD (2113–69); the third was an unaffected mother (2947–68). However, none of these individuals passed the 425V variant to their OCD-affected sibpair offspring (confirmed by DNA sequencing). This non-segregation of 425V to the 6 OCD-affected offspring was unexpected; such segregation by chance alone should only occur 1/64th of the time. Two of the 425V carriers were diagnosed with depressive disorders (1808–3 probable MDD and 2947–68 recurrent MDD/dysthymia), two had skin picking (1808–3 (probable) and 2947–68 (definite)); and two had probable alcohol dependence (1808–3 and 2113–69) (see OS Table 2).

We also analyzed the data stratified by sex by FBAT (similar PDT results), under the hypothesis that genetic loci for OCD may be sex-specific (Tables 2 & 3). One SNP, rs2020930 located 5’ to the LPR had a nominal p-value of <0.05 in males, but this did not withstand correction for multiple testing (P_{Adj}= 0.3379). The most statistically significant 2-point findings were obtained using a genotypic model and showed association in females with intron 2 alleles 9/10 (p-value< 0.0002; 10 informative families). With P_{Adj}= 0.0054, this finding withstood the joint correction for multiple marker comparison and LD. We then further adjusted the statistical significance for the testing of multiple genetic models, and the result remained significant (P_{MAdj}= 0.0069) (Table 3). Even using a conservative Bonferroni correction for the testing of 3 sex-based models (all, male-only, female-only) this result is still significant after correction for all tests done (p< 0.017).

Lastly, we carried out a haplotype analysis with FBAT for the two observed haplotype blocks using the entire and sex-stratified Caucasian sample. The 5’ block (rs2020930:rs4392119:LPR-rs25531:rs25533:rs2020933) contained the LPR and the first exon while the 3’ block (rs6355:Int2:rs140700:Int7:rs4583306:rs7224199) includes the rest of the gene (Table 2, OS Figures 3 & 4). Neither block achieved whole-marker significance using a permutation test. In the 5’ block, four individual haplotypes had p<0.05 of the 18 observed haplotypes. Interestingly, both individual haplotypes containing a high expressing L_A allele conferred risk (ACL_ATA and GTL_ATT, albeit in different sexes), while the two individual haplotypes containing low expressing alleles (GTS_ATT and GCL_GTT, principally in males) were protective. We note that the individual haplotypes GTL_ATT and GTS_ATT are identical except at LPR-rs25531, yet have opposite effects on OCD risk. In the 3’ block of 6 markers (including 2 microsatellites) and 25 observed individual haplotypes, three haplotypes with p< 0.05 were observed. Two of these contain the 267 bp allele at intron 2, which is part of the 250/267 genotype that was significant after correction. The set of 10 female families (250/267 genotype, Table 3) overlapped with the set of 8 female families (267 bp allele- containing haplotype, Table 4) in only a single female. We note that individual haplotype p-values are uncorrected for multiple testing.

Discussion

The serotonin transporter, *SLC6A4*, is the molecular target of the selective serotonin reuptake inhibitors (SSRIs), which are the current medications of choice for the treatment of OCD. There have been multiple genetic association studies of OCD and *SLC6A4* and while some of these have been positive, overall the results have been equivocal. In an effort to determine whether variation in *SLC6A4* alters the probability of an individual developing OCD, we genotyped 1,241 DNA samples from the OCGS OCD-affected sib pair study (278 pedigrees, 459 nuclear families, 1 598 total persons) with 17 marker loci.

After correction for multiple testing with several models and marker loci, none of the two-point results using the entire sample remained statistically significant. Given the suggestion of gender differences that have been observed in genetic association studies of OCD and *HTR2A, COMT* and *MAOA* (42–44) and also in segregation studies (3), we also analyzed our data stratified by sex, under the hypothesis that genetic loci for OCD may be sex-specific. In these analyses, genotype 9/10 of the Int2 VNTR was associated with OCD in females both before ($p < 0.0002$), and after correction for LD structure and multiple marker comparisons ($p_{Adj} = 0.0054$). This result remained significant after we further corrected for the testing of additive, dominant, recessive and genotypic genetic models ($P_{MAdj} = 0.0069$). There have been suggestions of sex-specific OCD findings for *SLC6A4*, for the LPR, in a Dutch sample (45) and a recent meta-analysis (10); the latter did not withstand multiple-marker nor multiple-analysis corrections, while the former may have. While we approached sex differences in the conferring of OCD risk as separate hypotheses, we note that even with additional correction, findings remain significant with $p < 0.0167$.

In the OCGS sample, there appears to be little difference in the OCD symptomatology between males and females. There are no sex differences in YBOCS severity scores, numbers of different categories of obsessions or compulsions, or in the proportions treated for their OCD symptoms. There are sex differences in prevalence of several Axis I disorders, females showing more recurrent major depression (MD), hypomania, dysthymia, hypochondriasis, skin picking, panic disorder, agoraphobia, social phobia (SP), generalized anxiety disorder (GAD), anorexia and bulimia and less TD, alcohol dependence and Asperger's Syndrome. (32) Serotonin has been hypothesized to play a role in many of these disorders. Given that the majority of OCD-affected individuals in the OCGS sample are female (66%), it is possible that the observation of the association to alleles 9/10 of the Int2 VNTR in the OCGS females may be due to presence of co-morbid Axis I disorders, rather than OCD.

There have been previous observations of association of alleles at the Int2 marker with a number of neuropsychiatric disorders. Allele 9 has been associated with both unipolar depression (46–47) and depression in the context of bipolar disorder (BPD). (47) The Int2 allele 12 has been associated with BPD (48), schizophrenia (meta-analysis) (49), autism (50), and anxiety disorders with and without OCD, in Japanese subjects, but not Caucasians. (51) Additionally, less favorable SSRI treatment response was observed in Asians with major depression and the Int2 10/12 genotype in a meta-analysis. (52) More recently, study

of OCD in a Spanish Caucasian sample found an excess of 12/12 and 12/10 genotypes in cases. (53)

The functional mechanism driving the associations of these Int2 alleles *in vivo* may be complex and may involve combined effects with the linked polymorphic region (LPR) polymorphism. (54) VNTR copy number variation has been shown to have effects on transcription (55), as have VNTR sequence variants on differential reporter gene expression. (56) Recent investigation of the Int2 VNTR as a target for mediating a transcriptional response to LiCl via the transcription factors *CTCF* and *YB-1*, found that *in vivo* transcriptional variation was correlated with differential binding of both these transcription factors to the three distinct VNTR variants after exposure to LiCl, suggesting differential allelic gene expression. (17)

Although there is evidence that the Int2 VNTR can influence the expression of *SLC6A4*, it is the LPR indel that has been most extensively studied. Early studies of the LPR by Lesch and colleagues observed that the transcriptional activity of the L allele was 2- to 3 fold higher than that of the S allele in lymphoblastoid cell cultures. (57) The L allele was associated with higher platelet serotonin uptake. (70–71) Consistent with many prior reports, we failed to detect association of the LPR with OCD in the OCGS sample, either in the combined-sex data or when analyzing by sex.

It appears that additional variation within the LPR affects the strength of the *SLC6A4* promoter. Hu and colleagues showed that the A to G base change at rs25531 within the LPR forms a haplotype (L_G) that has comparable low activity to the S allele. (1) They also found that the L_A allele was associated with OCD in both case-control and trio samples and conferred a 1.8-fold increased risk (OR) of developing the disorder. One other published report of the haplotype found equivocal replication with L_A/L_A, or L_A alone, in a case-control study of 347 OCD-affecteds vs. 749 population-matched controls, and results did not withstand correction for multiple testing. (58) We generated phase-certain haplotypes of the LPR-rs25531, but failed to observe any evidence that the L_A/L_A genotype conferred risk for OCD. In addition, we did not observe an association to either L_A or L_G under additive, dominant models, recessive or genotypic models, either in the total data set or when stratified by sex. In our sample the GRR for the L_A/L_A genotype vs. all others was 1.11 (95% CI: 0.77–1.60, p= 0.58). While positive, this risk estimate is markedly less than that reported by Hu et al (see also OS table 1).

Phenotypic and/or genotypic heterogeneity may have contributed to lack of replication in this multi-site association study. Factors such as ethnicity, method of ascertainment, site of patient recruitment, mean age at time of patient recruitment, and mean age at onset of OCD symptoms all may play a contributory role. We did not observe significant differences when analyzing these parameters by site of patient recruitment. We then assessed for genetic heterogeneity in our sample by examining LPR L and S allele-, and LPR-rs25531 haplotype frequencies in the OCGS sample by site. We did not detect significant differences.

Alternatively, it is possible that our results differ from those of Hu and colleagues (1) due to additional unmeasured genetic variation in the *SLC6A4* gene that has a functional effect on

the promoter of the gene. Two SNPs in the first intron of *SLC6A4*, rs16965628 and rs2020933 were shown correlated with allelic expression imbalances (59) and rs25532 has been suggested to affect expression. (58) The latter report described a functional C to T single nucleotide polymorphism (rs25532) located near the LPR; its minor allele significantly decreased luciferase reporter gene expression levels. Haplotype-based testing of rs25532 and all other known non-coding functional *SLC6A4* variants revealed significant overrepresentation in probands who had the higher-expressing allele at each locus, supporting the notion of increased serotonin transporter functioning being pathogenetically involved in OCD. Conditional haplotype analyses revealed that this association was driven by the LPR, rs25532 and rs16965628 in concert. Finally, our results may differ from those of Hu and colleagues due to the characteristics of the respective samples. The OCGS sample is composed entirely of familial OCD cases (33), whereas case-control and trio samples will contain a higher proportion of sporadic cases that may have a different genetic basis.

We also investigated several rare variants within *SLC6A4* to determine if they were associated with OCD in our sample. We found no individuals with either rs6353 (T439T) or the 381-bp deletion variant immediately 3' of the LPR that has been previously shown to contain several canonical transcription factor binding sites. Rs6353 is thought to be functional (56,16) and has been associated with both autism and depression. (61,62) We also investigated the rare SNP that causes a non-synonymous change from isoleucine to valine at position 425 of *SLC6A4* that has been suggested to contribute to an uncommon familial form of OCD in two unrelated pedigrees with OCD and other serotonin-related comorbidities. (2) We observed three individuals from three pedigrees with 425V, all of whom are parents of OCD affected sib-pairs. Collectively, these 3 individuals had 6 OCD affected offspring, none of whom received the 425V allele from these parents, a result that should occur in only $1/64^{th}$ of such segregations. These genetic data are not supportive of the hypothesis that 425V is solely responsible for the "OCD plus syndrome". (2,63) We also tested the "double-hit" hypothesis proposed by Ozaki and colleagues (2), in which homozygosity of the L-allele plus the 425V variant, which increase both transcription and functional activity, and hence increases the risk for OCD. Consistent with this hypothesis, the single 425V carrier who was affected with OCD is homozygous for the LPR L allele. Haplotype analysis of LPR-rs25531 revealed that this individual is L_G/L_A, and hence a functional heterozygote. Haplotype analysis of the three pedigrees segregating the 425V variant found that this rare variant was found on an LPR L_A background (the other two 425V carriers were S/L_A), suggesting one mutation event, possibly as a shared ancestral event in these pedigrees. The 425V allele has been observed in controls (60,63) and in another study in families with pervasive developmental disorders, eating disorders and OCD, but other large studies on various types of patients and controls have failed to detect it (64,65). Non-transmittal of the 425V allele to 6 definite OCD cases proves that those OCD cases in high-density OCD families have some other cause of their OCD.

Ozaki and colleagues have suggested the 425V variant as a possible mediator of SSRI treatment response or as a mediator/exacerbator of neuropsychiatric comorbidity. (2) They hypothesized an "atypical OCD plus", a syndrome incorporating OCD plus AS/autism/social phobia and eating disorders. We failed to find support for either of these hypotheses, as we observed no instances of 425V in the 32 self-reported SSRI treatment non-responders, and

while each of the three 425V carriers do have a complex neuropsychiatric phenotype, it differs from “atypical OCD plus”. We suggest that 425V might predispose to a spectrum of complex serotonergic phenotypes that is wider than that observed by Ozaki et al, and that OCD may not be a necessary component of this spectrum.

Family-based haplotype reconstruction has demonstrated higher reliability when compared to haplotype reconstruction from unrelated individuals. (66) When analyzing haplotypes, we excluded the small number of non-Caucasian pedigrees that may have contributed to heterogeneity. Several less common individual haplotypes do show the L_A allele associated with OCD risk, and the S_G and S_A alleles comparably associated with inverse risk. In contrast, 2-pt analysis at the LPR showed no association with OCD. Investigation of the COMT gene recently showed that while single variants might fail to show association (e.g., COMT V158M), functional haplotypes containing these alleles can show strong association (e.g., TMJ disorder and pain sensitivity). (67) The selection of haplotype blocking algorithm in the context of degree of LD at the given locus is known to affect association findings and type 1-error rates. (68) We noted that several of the blocking algorithms (e.g., Gabriel et al., Patil et al.) gave marginally significant findings after correction for the number of LD blocks tested in the 3’ haplotype block, notably when a significant portion of the marker set was excluded (e.g., Gabriel algorithm) or when haplotype “windows” approached small numbers of markers (e.g., Patil algorithm). Hence, the full significance of these individual haplotypes has yet to be determined; some may confer appreciable risk in a genotypic subset of this sample.

In conclusion, we found evidence of genetic association at the *SLC6A4* locus with OCD. Noteworthy was the lack of association at the LPR, the LPR-rs25531 haplotype, and the rare 425V variant. Current hypotheses about the risk of specific variants need revision to accommodate these new findings, including a possible gender effect.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The authors wish to thank the many families who have participated in the OCGS; the OCF; Ann Pulver, PhD, Kathleen Merikangas, PhD, David Houseman, MD, and Alec Wilson, PhD, for consultation; and clinicians and coordinators at each OCGS site: Providence (Maria Mancebo, PhD, Richard Marsland, RN, and Shirley Yen, PhD); New York (Renee Goodwin, PhD, Joshua Lipsitz, PhD, and Jessica Page, PsyD); Baltimore (Laura Eisen, BS, Karan Lamb, PsyD, Tracey Lichner, PhD, Yung-mei Leong, PhD, and Krista Vermillion, BA); Boston (Dan Geller, MD, Anne Chosak, PhD, Michelle Wedig, BS, Evelyn Stewart, MD, Michael Jenike, MD, Beth Gershuny, PhD, and Sabine Wilhelm, PhD); Bethesda (Lucy Jestement, Diane Kazuba, V. Holland LaSalle-Ricci, and Theresa B. DeGuzman); and Los Angeles (R. Lindsey Bergman, PhD, Susanna Chang, PhD, Audra Langley, PhD, and Amanda Pearlman, BA). In addition, we wish to thank Jason Briggs at ABI for his technical assistance during set-up of the SNPlex assay in a high-throughput, 384-well robotized format, Ming-Chen Chien, MS, at the Columbia University Genome Center, Dave Conti, USC, Keck School, Biostatistics, and Dr. Jonathan Liu & colleagues, Soft Genetics, Inc. for early design versions of *Gene Marker* & *Mutation Surveyor* softwares.

Funding Source

This study is supported by NIMH R01 MH50214 and NIH/NCRR/OPD-GCRC RR00052 (GN). EV is a recipient of a 2006 grant award from the Obsessive Compulsive Foundation (OCF) and a departmental fellowship award (USC, Keck School, Psychiatry). Other awards: NIMH K23-MH066284 (MG) & K23-MH64543 (OJB).

References

1. Hu XZ, Lipsky RH, Zhu G, Akhtar LA, Taubman J, Greenberg BD, et al. Serotonin transporter promoter gain-of-function genotypes are linked to OCD. Am. J. Hum. Genet 2006; 78: 815–826. [PubMed: 16642437]
2. Ozaki N, Goldman D, Kaye WH, Plotnicov K, Greenberg BD, et al. Serotonin transporter missense mutation associated with a complex neuropsychiatric phenotype. Mol. Psychiatry 2003; 8: 933–936. [PubMed: 14593431]
3. Nestadt G, Lan T, Samuels J, Riddle M, Bienvenu OJ III, et al. Complex segregation analysis provides compelling evidence for a major gene underlying OCD and for heterogeneity by sex. AJHG 2000; 67: 1611–1616.
4. Nestadt G, Samuels J, Riddle M, Bienvenu OJ III, Liang KY, Labuda M, et al. A family study of OCD. Arch. Gen. Psychiatry 2000; 57: 358–363. [PubMed: 10768697]
5. Hanna GL, Fingerlin TE, Himle JA, and Boehnke M Complex segregation analysis of OCD in families with pediatric probands. Hum. Hered 2005; 60: 1–9. [PubMed: 16088199]
6. Hanna GL, Veenstra-VanderWeele J, Cox NJ, Boehnke M, Himle JA, et al. Genome-wide linkage analysis of families with OCD ascertained through pediatric probands. Am. J. Med. Genet 2002; 114: 541–552. [PubMed: 12116192]
7. Willour VL, Yao SY, Samuels J, Grados M, Cullen B, Bienvenu OJ III, Wang Y, et al. Replication study supports evidence for linkage to 9p24 in OCD. Am. J. Hum. Genet 2004; 75: 508–513. [PubMed: 15272418]
8. Arnold PD, Sicard T, Burroughs E, Richter MA, and Kennedy JL Glutamate transporter gene SLC1A1 associated with OCD. Arch. Gen. Psychiatry 2006; 63: 769–776. [PubMed: 16818866]
9. Stewart SE, Fagerness JA, Platko J, Smoller JW, Scharf JM, et al. Association of the SLC1A1 glutamate transporter gene and OCD. Am. J. Med. Genet. B Neuropsychiatr. Genet 2007; 144B: 1027–1033. [PubMed: 17894418]
10. Dickel DE, Veenstra-VanderWeele J, Bivens NC, Wu X, Fischer DJ, Van Etten-Lee M, et al. Association studies of serotonin system candidate genes in early-onset OCD. Biol. Psychiatry 2007; 61: 322–329. [PubMed: 17241828]
11. Shugart YY, Samuels J, Willour VL, Grados MA, Greenberg BD, et al. Genomewide linkage scan for OCD: evidence for susceptibility loci on chromosomes 3q, 7p, 1q, 15q & 6q. Mol. Psychiatry 2006; 11: 763–770. [PubMed: 16755275]
12. Goddard AW, Shekhar A, Whiteman AF, and McDougle CJ Serotoninergic mechanisms in the treatment of OCD. Drug Discov. Today 2008; 13: 325–332. [PubMed: 18405845]
13. Stengler-Wenzke K, Muller U, Angermeyer MC, Sabri O, and Hesse S Reduced serotonin transporter-availability in OCD. Eur. Arch. Psychiatry Clin. Neurosci 2004; 254: 252–255.
14. Hesse S, Müller U, Lincke T, Barthel H, Villmann T, Angermeyer MC, et al. Serotonin and dopamine transporter imaging in patients with obsessive-compulsive disorder. Psychiatry Res. 2005; 140(1): 63–72. [PubMed: 16213689]
15. Hasselbalch SG, Hansen ES, Jakobsen TB, Pinborg LH, Lønborg JH, and Bolwig TG Reduced midbrain-pons serotonin transporter binding in patients with OCD. Acta Psychiatr Scand. 2007 ; 115: 388–94. [PubMed: 17430417]
16. Prasad HC, Zhu CB, McCauley JL, Samuvel DJ, et al. Human serotonin transporter variants display altered sensitivity to protein kinase G and p38 mitogen-activated protein kinase. PNAS 2005; 102: 11545–11550. [PubMed: 16055563]
17. Roberts J, Scott AC, Howard MR, Breen G, Bubb VJ, et al. Differential regulation of the serotonin transporter gene by lithium is mediated by transcription factors, CCCTC binding protein and Y-box binding protein 1, through the polymorphic intron 2 variable number tandem repeat. J. Neurosci 2007; 27: 2793–2801. [PubMed: 17360901]
18. Bloch MH, Landeros-Weisenberger A, Sen S, Dombrowski P, et al. Association of the SERT polymorphism and OCD: Systematic review. Am. J. Med. Genet. B Neuropsych. Genet 2008; 147B: 850–8.

19. McDougle CJ, Epperson CN, Price LH, and Gelernter J Evidence for linkage disequilibrium between serotonin transporter protein gene (SLC6A4) and OCD. Mol. Psychiatry 1998; 3: 270–273. [PubMed: 9672904]
20. Bengel D, Greenberg BD, Cora-Locatelli G, Altemus M, Heils A, Li Q, et al. Association of the serotonin transporter promoter regulatory region polymorphism and OCD. Mol. Psychiatry 1999; 4: 463–466. [PubMed: 10523819]
21. Camarena B, Rinetti G, Cruz C, Hernandez S, de la Fuente JR, and Nicolini H Association study of the serotonin transporter gene polymorphism in OCD. Int. J. Neuropsychopharmacol 2001; 4: 269–272. [PubMed: 11602033]
22. Cavallini MC, Di Bella D, Siliprandi F, Malchiodi F, et al. Exploratory factor analysis of obsessive-compulsive patients and association with 5-HTTLPR polymorphism. Am. J. Med. Genet 2002; 114: 347–353. [PubMed: 11920862]
23. Di Bella D, Erzegovesi S, Cavallini MC, and Bellodi L OCD, 5-HTTLPR polymorphism and treatment response. Pharmacogenomics J. 2002; 2: 176–181. [PubMed: 12082589]
24. Kim SJ, Lee HS, and Kim CH OCD, factor-analyzed symptom dimensions and serotonin transporter polymorphism. Neuropsychobiology 2005; 52: 176–182. [PubMed: 16220023]
25. Kinnear CJ, Niehaus DJ, Moolman-Smook JC, du Toit PL, van Kradenberg J, Weyers JB, et al. OCD and the promoter region polymorphism (5- HTTLPR) in the serotonin transporter gene (SLC6A4): a negative association study in the Afrikaner population. Int. J. Neuropsychopharmacol 2000; 3: 327–331. [PubMed: 11343612]
26. Frisch A, Finkel B, Michaelovsky E, Sigal M, Laor N, and Weizman R A rare short allele of the serotonin transporter promoter region (5- HTTLPR) found in an aggressive schizophrenic patient of Jewish Libyan origin. Psychiatr. Genet 2000; 10: 179–183. [PubMed: 11324943]
27. Walitza S, Wewetzer C, Gerlach M, Klampfl K, Geller F, Barth N, et al. Transmission disequilibrium studies in children and adolescents with OCD pertaining to polymorphisms of genes of the serotonergic pathway. J. Neural Transm 2004; 111: 817–825. [PubMed: 15206000]
28. Chabane N, Millet B, Delorme R, Lichtermann D, Mathieu F, Laplanche JL, et al. Lack of evidence for association between serotonin transporter gene (5-HTTLPR) and OCD by case control and family association study in humans. Neurosci. Lett 2004; 363: 154–156. [PubMed: 15172105]
29. Lin PY Meta-analysis of the association of serotonin transporter gene polymorphism with OCD. Prog. Neuropsychopharmacol. Biol. Psychiatry 2007; 31: 683–689. [PubMed: 17291658]
30. Kilic F, Murphy DL, and Rudnick G A human serotonin transporter mutation causes constitutive activation of transport activity. Mol. Pharmacol 2003; 64: 440–446. [PubMed: 12869649]
31. Wendland JR, DeGuzman TB, McMahon F, Rudnick G, Tera-Wadleigh SD, and Murphy DL SERT Ileu425Val in autism, Asperger syndrome and OCD. Psychiatr. Genet 2008; 18: 31–39. [PubMed: 18197083]
32. Samuels JF, Riddle MA, Greenberg BD, Fyer AJ, McCracken JT, et al. The OCD collaborative genetics study: methods and sample description. Am J Med Genet B Neuropsychiatr. Genet 2006; 141: 201–207.
33. Rabinowitz D and Laird N A unified approach to adjusting association tests for population admixture with arbitrary pedigree structure and arbitrary missing marker information. Hum. Hered 2000; 50: 211–223. [PubMed: 10782012]
34. Horvath S, Xu X, and Laird NM The family based association test method: strategies for studying general genotype-phenotype associations. Eur. J Hum. Genet 2001; 9: 301–306. [PubMed: 11313775]
35. Martin ER, Monks SA, Warren LL and Kaplan NL A test for linkage and association in general pedigrees: the pedigree disequilibrium test. Am. J. Hum. Genet 2001; 68: 1065–7. [PubMed: 11254459]
36. Conneely KN and Boehnke M So Many Correlated Tests, So Little Time! Rapid Adjustment of P Values for Multiple Correlated Tests. Am. J. Hum. Genet 2007; 10 26; 81(6). [Epub ahead of print]
37. Kraft P and Stram DO Re: the use of inferred haplotypes in downstream analysis. AJHG 2007; 81: 863–5.

38. Sinnwell JP, Schaid DJ and Yu Z R Development Core Team. R: A language and environment for statistical computing. 2008; ISBN 3-900051-07-0, URL http://www.R-project.org
39. Statistical Analysis of Haplotypes with Traits and Covariates when Linkage Phase is Ambiguous. R package version 1.3.8. URL http://mayoresearch.mayo.edu/mayo/research/schaid_lab/software.cfm
40. Flattem NL and Blakely RD Modified structure of the human serotonin transporter promoter. Mol. Psychiatry 2000; 5: 110–115. [PubMed: 10673778]
41. Nakamura M, Ueno S, Sano A, and Tanabe H The human serotonin transporter gene linked polymorphism (5-HTTLPR) shows ten novel allelic variants. Mol. Psychiatry 2000; 5: 32–38. [PubMed: 10673766]
42. Enoch MA, Greenberg BD, Murphy DL, and Goldman D Sexually dimorphic relationship of a 5-HT2A promoter polymorphism with OCD. Biol. Psychiatry 2001; 49: 385–388. [PubMed: 11239910]
43. Karayiorgou M, Sobin C, Blundell ML, Galke BL, et al. Family-based association studies support a sexually dimorphic effect of COMT and MAOA on genetic susceptibility to OCD. Biol. Psych 1999; 45: 1178–89.
44. Alsobrook JP, Zohar AH, Leboyer M, Chabane N, Ebstein RP, and Pauls DL Association between the COMT locus and OCD in females but not males. Am. J. Med. Genet 2002; 114: 116–120. [PubMed: 11840516]
45. Denys D, Van NF, Deforce D, and Westenberg HG Association between serotonergic candidate genes and specific phenotypes of OCD. J. Affect. Disord 2006; 91: 39–44. [PubMed: 16443280]
46. Ogilvie AD, Battersby S, Bubb VJ, Fink G, Harmar AJ, Goodwim GM, et al. Polymorphism in serotonin transporter gene associated with susceptibility to major depression. Lancet 1996; 347: 731–733. [PubMed: 8602004]
47. Battersby S, Ogilvie AD, Smith CA, Blackwood DH, Muir WJ, et al. Structure of a variable number tandem repeat of the SERT gene and association with affective disorder. Psychiatr. Genet 1996; 6: 177–181. [PubMed: 9149321]
48. Collier DA, Arranz MJ, Sham P, Battersby S, Vallada H, Gill P, et al. The serotonin transporter is a potential susceptibility factor for bipolar affective disorder. Neuroreport 1996; 7: 1675–1679. [PubMed: 8904780]
49. Fan JB and Sklar P Meta-analysis reveals association between serotonin transporter gene STin2 VNTR polymorphism and schizophrenia. Mol. Psychiatry 2005; 10: 928–38, 891. [PubMed: 15940296]
50. Conroy J, Meally E, Kearney G, Fitzgerald M, Gill M, and Gallagher L Serotonin transporter gene and autism: a haplotype analysis in an Irish autistic population. Mol. Psychiatry 2004; 9: 587–593. [PubMed: 14708029]
51. Ohara K, Suzuki Y, Ochiai M, Tsukamoto T, Tani K, and Ohara K A variable-number-tandem-repeat of the serotonin transporter gene and anxiety disorders. Prog. Neuropsychopharm. Biol. Psych 1999; 23: 55–65.
52. Smits KM, Smits LJ, Schouten JS, Stelma FF, et al. Influence of SERTPR and STin2 on the effect of selective serotonin reuptake inhibitors in depression: a systematic review. Mol. Psychiatry 2004; 9: 433–441. [PubMed: 15037864]
53. Baca-Garcia E, Vaquero-Lorenzo C, Diaz-Hernandez M, Rodriguez-Salgado B, Dolengevich-Segal H, Arrojo-Romero M, et al. Association between OCD and a variable number of tandem repeats polymorphism in intron 2 of the serotonin transporter gene. Prog. Neuropsychopharmacol. Biol. Psychiatry 2007; 31: 416–420. [PubMed: 17174018]
54. Hranilovic D, Stefulj J, Schwab S, Borrmann-Hassenbach M, et al. SERT promoter and intron 2 polymorphisms: relationship between allelic variants and gene expression. Biol. Psych 2004; 55: 1090–1094.
55. MacKenzie A and Quinn J A SERT gene intron 2 polymorphic region, correlated with affective disorders, has allele-dependent differential enhancer-like properties in the mouse embryo. PNAS 1999; 96: 15251–15255. [PubMed: 10611371]
56. Lovejoy EA, Scott AC, Fiskerstrand CE, Bubb VJ, and Quinn JP The serotonin transporter intronic VNTR enhancer correlated with a predisposition to affective disorders has distinct regulatory

elements within the domain based on the primary DNA sequence of the repeat unit. Eur. J. Neurosci 2003; 17: 417–420. [PubMed: 12542679]

57. Lesch KP, Bengel D, Heils A, Sabol SZ, Greenberg BD, Petri S, et al. Association of anxiety-related traits with a polymorphism in the serotonin transporter gene regulatory region. Science 1996; 274: 1527–1531. [PubMed: 8929413]
58. Wendland JR, Moya PR, Kruse MR, Ren-Patterson RF, Jensen CL, Timpano KR, et al. A novel, putative gain-of-function haplotype at SLC6A4 associates with OCD. Hum. Mol. Genet 2008; 17: 717–723. [PubMed: 18055562]
59. Martin J, Cleak J, Willis-Owen SA, Flint J, and Shifman S Mapping regulatory variants for the serotonin transporter gene based on allelic expression imbalance. Mol. Psychiatry 2007; 12: 421–422. [PubMed: 17453058]
60. Glatt CE, DeYoung JA, Delgado S, Service SK, et al. Screening a large reference sample to identify very low frequency sequence variants: comparisons between two genes. Nat. Genet 2001; 27: 435–438. [PubMed: 11279528]
61. Kim SJ, Herzing LB, Veenstra-Vanderweele J, Lord C, Courchesne R, Leventhal BL, et al. Mutation screening and transmission disequilibrium study of ATP10C in autism. Am. J. Med. Genet 2002; 114: 137–143. [PubMed: 11857573]
62. Kraft JB, Slager SL, McGrath PJ and Hamilton SP Sequence analysis of the serotonin transporter and associations with antidepressant response. Biol. Psychiatry 2005; 58: 374–81. [PubMed: 15993855]
63. Delorme R, Betancur C, Wagner M, Krebs MO, Gorwood P, Pearl P, et al. Support for association between the rare functional variant I425V of the serotonin transporter gene and susceptibility to OCD. Mol. Psychiatry 2005; 10: 1059–1061. [PubMed: 16088327]
64. Cross S, Kim SJ, Weiss LA, Delahanty RJ, Sutcliffe JS, et al. Molecular genetics of the platelet serotonin system in first-degree relatives of patients with autism. Neuropsychopharm. 2008; 33: 353–60.
65. Ikeda M, Iwata N, Suzuki T, Kitajima T, Yamanouchi Y, Kinoshita Y and Ozaki N No association of serotonin transporter gene (SLC6A4) with schizophrenia and bipolar disorder in Japanese patients: association analysis based on linkage disequilibrium. J. Neural Transm 2006; 113: 899–905. [PubMed: 16082508]
66. Levenstien MA, Ott J and Gordon D 2006 Are Molecular Haplotypes Worth the Time and Expense? A Cost-Effective Method for Applying Molecular Haplotypes. PLoS Genet. 2(8): e127. [PubMed: 16933998]
67. Nackley AG, Shabalina SA, Tchivileva IE, Satterfield K, Korchynskyi O, et al. Human COMT haplotypes modulate protein expression by altering mRNA secondary structure. Science. 2006; 314: 1930–3. [PubMed: 17185601]
68. Nicodemus KK, Liu W, Chase GA, Tsai Y-Y and Fallin, M.D. Comparison of type I error for multiple test corrections in large single-nucleotide polymorphism studies using principal components versus haplotype blocking algorithms. BMC Genetics. 2005; 6(Suppl 1): S78. [PubMed: 16451692]
69. Wendland JR, Moya PR, Timpano KR, Anavitarte AP, et al. A haplotype containing quantitative trait loci for SLC1A1 gene expression and its association with OC. Arch. Gen. Psych 2009; 66: 408–16.
70. Greenberg BD, Tolliver TJ, Huang SJ, Li Q, Bengel D and Murphy DL Genetic variation in the SERT promoter region affects serotonin uptake in human blood platelets. Am. J. Med. Genet 1999; 88: 83–7. [PubMed: 10050973]
71. Altemus M, Murphy DL, Greenberg B and Lesch KP Intact coding region of the serotonin transporter gene in obsessive-compulsive disorder. Am. J. Med. Genet 1996; 67: 409–11. [PubMed: 8837711]

Table 1:

FBAT analysis of 17 markers spanning the *SLC6A4* locus

Marker	Notes (bp)	NCBI Build 126 bp	GT comp. %[3]
rs1487971	5′ near *BLMH*	25,596,879	98.6
rs2020930	5′	25,590,167	98.7
rs4392119	5′	25,589,489	96.7
LPR	$(42–44)_{\text{usu. 7–9}}$	1.2 kb 5’	99.6
LPR-rs25531[1]	See OS fig 2.	25,588,472	96.5
rs25533	5′	25,587,018	97.5
rs2020933	Intron 1	25,585,881	98.7
rs2066713	Intron 1	25,575,791	99.0
rs6355	G56A; exon 1	25,572,936	98.6
Intron 2	$(16–17)_{\text{9, 10 or 12}}$	25,572,759	99.2
rs140699[2]	Intron 4	25,569,251	98.1
rs140700	Intron 5	25,567,515	98.8
Intron 7	$(GAAA)_{1–7}$	25,566,563	99.6
rs4583306	Intron 8	25,562,841	98.5
I425V[2]	Exon 8	25,562,500	98.9
rs6353[2]	T439T; exon 8	25,562,456	98.7
rs7224199	3′	25,547,852	99.0

[1] LPR-rs25531 indicates molecular haplotype of the LPR: L vs. S allele and rs25531 A vs. G, yielding 4 haplotypes.

[2] Nonpolymorphic SNPs excluded from all haplotype analysis.

[3] Completion % of respective genotyping assay.

Mol Psychiatry. Author manuscript; available in PMC 2019 May 21.

Table 2:

FBAT analysis of *SLC6A4* locus: 2-pt & haplotype analysis (additive model)

A. 2-pt analysis by SNP												B. Haplotypes (Caucasians)[1]
Marker	**Base[2]/ Allele**	**OCGS MAF**	**Fam. n all[3]**	**Z_{all}[4]**	**Fam. n m[3]**	**Z_m**	**Fam. n f[3]**	**Z_f**	**P_{all}[4]**	**P_m**	**P_f**	**P_{all} (P_m, P_f)**
rs1487971	G/**A**	0.379	157	1.009	100	0.582	133	0.780	0.3132	0.5607	0.4357	
rs2020930	G/**A**	0.049	38	0.616	20	1.877	31	−0.609	0.538	0.0606	0.5428	5’ Block: 0.1322 (0.1716, 0.0944)
rs4392119	C/**T**	0.028	23	−0.897	9	**	25	0.278	0.3696	**	0.7809	
LPR-rs25531[5]	See C. below											
rs25533	T/**C**	0.053	38	1	20	0	32	1.206	0.3173	1	0.2278	
rs2020933	T/**A**	0.079	54	−0.017	28	1.635	46	−1.198	0.9863	0.1020	0.2308	
rs2066713	C/**T**	0.393	149	0.822	99	0.468	129	0.619	0.4111	0.6399	0.5359	
rs6355	G/**C**	0.016	18	0.478	6	**	17	−0.084	0.633	**	0.9330	3’ Block: 0.1956 (0.1634, 0.1353)
Intron 2	See C. below											
rs140700	G/**A**	0.078	66	−0.595	36	−0.295	57	−0.491	0.5517	0.7681	0.6234	
Intron 7	See C. below											
rs4583306	A/**G**	0.426	154	0.018	100	0.069	131	−0.033	0.9855	0.9450	0.9734	
rs7224199	G/**T**	0.452	154	−0.144	99	0.716	132	−0.762	0.8857	0.4743	0.4462	
C. 2-pt analysis for multi-allelic markers												
LPR	L/**S**	0.425	169	0.732	114	0.862	150	0.181	0.4642	0.3889	0.8567	
LPR-rs25531	L_A	0.504	177	0.493	114	1.591	148	−0.717	0.6217	0.1116	0.4731	
	L_G	0.065	63	0.439	31	−1.321	53	1.474	0.6606	0.1865	0.1404	
	S_A	0.426	165	−0.806	109	−1.117	142	−0.059	0.4204	0.2639	0.9532	
	S_G	0.004	4	**	2	**	3	**				
Intron 2	$9_{250\ bp}$	0.017	17	0.750	11	−0.277	15	1.209	0.4531	0.7815	0.2265	
	$10_{267\ bp}$	0.376	161	1.017	104	0.976	137	0.428	0.3092	0.3292	0.6687	
	$12_{300\ bp}$	0.607	158	−1.215	102	−0.912	136	−0.735	0.2244	0.3619	0.4623	
Intron 7	$1_{263\ bp}$	0.003	7	**	4	**	3	**				
	$2_{267\ bp}$	0.016	13	0.408	8	**	15	−1.118	0.6835	**	0.2635	

A. 2-pt analysis by SNP											B. Haplotypes (Caucasians)[1]
Marker	**Base[2] / Allele**	**OCGS MAF**	**Fam. n all[3]**	**Z_{all}[4]**	**Fam. n m[3]**	**Z_m**	**Fam. n f[3]**	**Z_f**	**P_{all}[4]**	**P_m**	**P_f**	**P_{all} (P_m, P_f)**
	$3_{271\ bp}$	0.148	107	−1.64	61	−1.762	91	−0.673	0.1009	0.0781	0.5010	
	$4_{280\ bp}$	0	0	**	0	**	0	**				
	$5_{284\ bp}$	0.018	14	2.132	8	**	12	2.50	0.0330	**	0.0124	
	$6_{285\ bp}$	0.001	1	**	0	**	1	**				
	$7_{288\ bp}$	0.810	118	0.817	73	0.981	101	0.235	0.4137	0.3265	0.8141	

[1] Haploview v4.1 analysis of LD structure at the *SLC6A4* locus: 2 haplotype blocks are indicated by grey shading (leftmost column). Haplotype most significant whole-marker permutation test p-values (minimal p) are shown for blocks I & II in Caucasian pedigrees only, uncorrected (rightmost column). We analyzed using the solid spine algorithm, with extension of the LD blocks when D'>0.95. Non-polymorphic SNPs excluded from analysis. See Table 5 for best individual haplotypes (p_i).

[2] Minor allele is shown in bold.

[3] Throughout tabular data, fam. n indicates number of transmission-informative families;

"**" Indicates FBAT family number threshold not achieved for analysis.

[4] $Z_{all, m, and f}$ and $P_{all, m, and f}$, show the FBAT test statistic and p-value for the combined data set, males, and females, respectively.

[5] LPR-rs25531 indicates molecular haplotyping of the LPR locus: L vs. S allele and rs25531 A vs. G, yielding the 4 haplotypes in Part C, column 2.

Table 3:

FBAT best models, markers and alleles with corrections for LD, multiple marker- & multiple model testing

Marker	Allele(s) (bp)	Sex	Model	Allele freq.	Fam. n	Z[1]	P	P_{Adj}[2]	P_{MAdj}[3]
rs2020930	G/A	M	Genotypic	0.099	20	2.294	0.0219	0.3379	0.3772
	G/G			0.901	19	−2.132	0.0330	0.4528	0.4978
Intron 2	9/10 (250/267)	All	Genotypic	0.019	13	2.617	0.0089	0.2053	0.2355
		F			**10**	**3.738**	**<0.0002**	**0.0054**	**0.0069**
Intron 7	3/7 (271/288)	All	Genotypic	0.241	105	−2.352	0.0187	0.3590	0.4033
		F			90	−1.903	0.0571	0.7155	0.7557
	5 (284)	All	Additive, Dominant	0.018	14	2.132	0.0330	0.3719	0.5762
		F			12	2.5	0.0124	0.1662	0.2974

[1] Z statistic sign indicates positive vs. negatively directed association, ie risk-conferring vs. protective allele(s).

[2] P_{Adj} indicates p-value correction for multiple-marker testing in the context of LD within the marker set (see Methods).

[3] P_{MAdj} indicates p-value correction across multiple models, i.e., global correction (see Methods).

Mol Psychiatry. Author manuscript; available in PMC 2019 May 21.

Table 4:

OCGS sample heterogeneity analysis

A. Phenotypes															
	Providence., RI		***NY, NY***		***Baltimore., MD***		***Boston., MA***		***Bethesda., MD***		***Los Angeles, CA***		***Total***		***Chi²*** ***(p value)***
	n	***%***	***n***	***%***	***n***	***%***	***n***	***%***	***n***	***%***	***n***	***%***	***n***	***%***	
Subjects with DNA	119		73		906		45		19		64		1226		
Family n	32		16		194		12		7		17		278		
Male n	45	37.8	28	38.4	358	39.5	19	42.2	5	26.3	29	45.3	484	39.5	2.61 (0.760)
Female n	74	62.2	45	61.6	548	60.5	26	57.8	14	73.7	35	54.7	742	60.5	
Pedigree ethnicity (by proband)															
Caucasian	31	96.9	16	100.0	187	96.4	11	91.7	7	100.0	15	88.2	267	96.0	
Black	1	0	0	0	0	0	0	0	0	0	0	0	1	0.4	
Hispanic	0	0	0	0	0	0	0	0	0	0	1	5.9	1	0.4	
Other	0	0	0	0	2	1.0	0	0	0	0	1	0	3	1.1	
Unknown	0	0	0	0	1	0.5	1	8.3	0	0	0	0	2	0.7	
OCD n	84	70.6	44	60.3	603	66.6	34	75.6	15	78.9	44	68.8	824	67.2	5.06 (0.408)[1]
B. Genotypes															
LPR Allele frequencies															
L	122	51.3	79	54.1	1047	57.8	49	54.4	16	42.1	66	51.6	1379	56.2	6.37 (0.272)
S	104	43.7	63	43.2	729	40.2	41	45.6	16	42.1	64	50	1017	41.5	
L_A	111	46.6	76	52.1	929	51.3	38	42.2	15	39.5	64	50	1233	50.3	24.4 (0.059)
L_G	11	4.6	3	2.1	118	6.5	11	12.2	1	2.6	2	1.6	146	6.0	
S_A	103	43.3	63	43.2	719	39.7	41	45.6	16	42.1	64	50	1006	41.0	
S_G	1	0.4	0	0	10	0.6	0	0	0	0	0	0	11	0.4	

Notes

[1] OCD n vs. non-OCD n.

Table 5:

FBAT significant individual haplotypes, Caucasian pedigrees only (additive model)

Multi-marker haplotype block: markers[1]	Sex	Haplo-type freq.	# Fam.	Z	P_i[2]	Alleles [bp][3]
5' Block rs2020930:rs4392119:LPR-rs25531[4]:rs25533:rs2020933	All	0.004	6	2.111	0.0348	GTL_ATT
	F			2.53	0.0114	
	M	0.017	7	−2.178	0.0294	GCL_GTT
	All	0.021	17	−2.46	0.0138	GTS_ATT
	M		8	−2.255	0.0241	
	M	0.035	18	1.964	0.0495	ACL_ATA
3' Block rs6355:Int2:rs140700:Int7:rs4583306:rs7224199	All	0.011	8	2.381	0.0172	G[267]G[284]AG
	F			2.429	0.0151	
	M	0.015	8	2.333	0.0196	G[300]A[267]AG
		0.029	12	1.982	0.0474	G[267]G[288]AT

[1] Haplotype blocks shown graphically in OS Figure 3.

[2] Best individual haplotype p-value, uncorrected.

[3] Int2 alleles 9–12, and Int7 alleles 1–7, correspond to the following amplicon sizes [bp], respectively: 9=250, 10=267, and 12=300; and 1=263, 2=267, 3=271, 4=280, 5=284, 6=285, and 7=288.

[4] LPR-rs25531 indicates functional phase-known haplotype at the LPR locus, e.g., L_A, L_G, S_A or S_G.

Mol Psychiatry. Author manuscript; available in PMC 2019 May 21.