# Serotonin Transporter Promoter Gain-of-Function Genotypes Are Linked to Obsessive-Compulsive Disorder

Xian-Zhang Hu,[1] Robert H. Lipsky,[1] Guanshan Zhu,[1] Longina A. Akhtar,[1] Julie Taubman,[1] Benjamin D. Greenberg,[2,3] Ke Xu,[1] Paul D. Arnold,[4] Margaret A. Richter,[4] James L. Kennedy,[4] Dennis L. Murphy,[5] and David Goldman[1]

[1]Laboratory of Neurogenetics, National Institute on Alcohol Abuse and Alcoholism, Rockville, MD; [2]Butler Hospital and [3]Department of Psychiatry, Brown University School of Medicine, Providence; [4]Centre for Addiction and Mental Health, Department of Psychiatry, University of Toronto, Toronto; and [5]Laboratory of Clinical Science, National Institute of Mental Health, Bethesda, MD

**A functional serotonin transporter promoter polymorphism, *HTTLPR*, alters the risk of disease as well as brain morphometry and function. Here, we show that *HTTLPR* is functionally triallelic. The $L_G$ allele, which is the L allele with a common G substitution, creates a functional AP2 transcription-factor binding site. Expression assays in 62 lymphoblastoid cell lines representing the six genotypes and in transfected raphe-derived cells showed codominant allele action and low, nearly equivalent expression for the S and $L_G$ alleles, accounting for more variation in *HTT* expression than previously recognized. The gain-of-function $L_AL_A$ genotype was approximately twice as common in 169 whites with obsessive-compulsive disorder (OCD) than in 253 ethnically matched controls. We performed a replication study in 175 trios consisting of probands with OCD and their parents. The $L_A$ allele was twofold overtransmitted to the patients with OCD. The *HTTLPR* $L_AL_A$ genotype exerts a moderate (1.8-fold) effect on risk of OCD, which crystallizes the evidence that the *HTT* gene has a role in OCD.**

The serotonin transporter strongly modulates serotonin function and is a major therapeutic target in several psychiatric diseases, including anxiety, depression, and obsessive-compulsive disorder (OCD [MIM 164230]). *HTTLPR*, a functional polymorphism of the 5′ flanking region of the serotonin transporter gene (called *HTT*, *SLC6A4*, or *SERT*),[1] is an intensively studied locus. More than 300 studies have investigated the role of *HTTLPR* in diverse neuropsychiatric phenotypes.[2] Recently, *HTTLPR* was shown to determine the neuroanatomical size and functional coupling of the amygdala-frontal cortical circuit that has been directly implicated in a variety of psychiatric disorders, including OCD. The focus of these imaging genetics studies was the S allele, which leads to reduced gray matter volume in limbic regions and disrupted amygdala-cingulate coupling after emotional stimuli.[3,4] Conversely, the gain-of-function L allele had the opposite effect, an effect potentially relevant to OCD for which hyperfrontality has been observed in functional neuroimaging studies.[5]

*HTTLPR* has been considered functionally biallelic, even though other genetic variations were known. *HTTLPR* consists of varying numbers of copies of a 20–23-bp imperfect repeat sequence. Among white individuals, the frequency of the L allele (16 repeats) is ~0.60, and the frequency of the S allele (14 repeats) is ~0.40.[6] Substantial interpopulation variation occurs. Rare alleles contain up to 20 copies of the repeat.[7] Single-nucleotide variants have been detected within *HTTLPR*,[8] including an A→G SNP that will be extensively discussed below. However, none of these additional alleles was found to be functional, although several (including the A→G SNP) were tested in transfected mammalian cells. *HTTLPR* was therefore treated as biallelic in linkage studies.

Concerning function, the S allele leads to lower expression of *HTT* mRNA and lower expression of serotonin transporter in membranes.[6,9] Furthermore, the S allele appears to have a dominant mode of action,[6,9] an observation that leads to the grouping of SS and SL genotypes in many but not all linkage studies.[10,11] In studies in vivo, the L allele has been associated with higher levels of HTT in platelets, postmortem brain,[12] and living brain.[12–14]

Recently, an uncommon *HTT* variant, Val425,[15] which leads to gain of function,[16] was linked to severe behavioral problems (treatment-resistant OCD, anorexia nervosa, and Asperger syndrome) in two families.[15] In particular, six patients with OCD in these families all carried the Val425 allele.[15] These results suggested that the search for functional variation in *HTT* should be extended and that gain-of-function variants might be linked to OCD.

Received December 29, 2005; accepted for publication March 1, 2006; electronically published March 28, 2006.

Address for correspondence and reprints: Dr. David Goldman, Laboratory of Neurogenetics, National Institute on Alcohol Abuse and Alcoholism, Rockville, MD 20852. E-mail: dgneuro@box-d.nih.gov

*Am. J. Hum. Genet.* 2006;78:815–826.  0002-9297/2006/7805-0009$15.00

**Table 1**

**Primers and Probes for Two-Stage 5′-Exonuclease Genotyping of *HTTLPR* Alleles**

The table is available in its entirety in the online edition of *The American Journal of Human Genetics*.

## Material and Methods

### *Human Subjects*

Genomic DNA was extracted from 2,998 individuals, including 624 African Americans (414 with drug addictions and 210 with no psychiatric diagnosis), 771 Finnish whites (480 with various psychiatric diagnoses and 291 with no psychiatric disorders), 297 U.S. whites (group 1: 177 with various psychiatric diagnoses and 120 with no psychiatric disorders), 286 U.S. whites (group 2: 154 with various psychiatric diagnoses and 132 with no psychiatric disorders), 456 American Plains Indians (a community-representative sample including 335 individuals with psychiatric diagnoses and 121 with no psychiatric diagnosis), and 564 Southwest Indians (a community-representative sample including 359 with psychiatric disorders and 205 with no psychiatric diagnosis). Allele and genotype frequencies in the patients and controls in these populations did not differ. Presumably, this is because of the clinical heterogeneity and the lack of any strong *HTTLPR* effect on risk for the common but probably general diagnostic categories, such as alcoholism and anxiety disorders. Therefore, overall frequencies were computed. Human subjects were studied under human research protocols approved by the Rutgers University School of Medicine, the National Institute on Alcohol Abuse and Alcoholism (NIAAA), and the University of Helsinki. All participants gave informed consent for genetic studies on psychiatric disease. Subjects were interviewed for psychiatric traits with either the Structured Clinical Interview for DSM-IV (SCID) or the Schedule for Affective Disorders and Schizophrenia–Lifetime (SADS-L), and diagnoses were made according to DSM-IIIR criteria.

### *NIMH OCD Case-Control Study*

Under a human research protocol approved by the National Institute of Mental Health (NIMH) Institutional Review Board (IRB), 169 patients with OCD and 253 controls without psychiatric disorders were ascertained at the NIMH. All participants gave informed consent. All participants were interviewed with the SCID, and diagnoses were made using DSM-IIIR criteria. The OCD sample was white and consisted of 97 males and 72 females, with an average age of 41.7 years. The control sample included 137 males and 116 females, with an average age of 39.9 years.

### *Toronto OCD Study of Parents-Child Trios*

A total of 581 individuals were ascertained, including 175 OCD-affected child-parents trios, at the Centre for Addiction and Mental Health, Toronto, under a human research protocol approved by that institution's IRB. All participants gave informed consent. Probands were recruited from consecutive referrals to the Anxiety Disorders Clinic. All participants were assessed using the SCID and the Yale-Brown Obsessive Compulsive Scale[17] by trained interviewers. All information obtained was reviewed by a psychiatrist (M.A.R.) experienced in the diagnosis and treatment of OCD and related conditions, to insure diagnostic accuracy in using DSM-IV criteria. Only probands with confirmed OCD were included. Exclusion criteria included lifetime history of neurological or metabolic disease, bipolar disorder, psychotic disorder, or substance dependence. The transmission/disequilibrium test (TDT) was performed using complete parents-child trios and without regard for the affection status of parents. The TDT sample included 556 whites, 11 Hispanics, and 14 Asians. The probands comprised 69 males and 96 females, with an average age of 38.2 years. There were a total of 86 transmissions from heterozygous parents.

### *Two-Stage 5′-Nuclease Genotyping of* HTTLPR*, Including S, $L_A$, and $L_G$ Alleles*

Oligonucleotide primers and dye-labeled probes were designed to optimize allele discrimination by use of Primer Express software (Applied Biosystems). The L amplicon was 182 bp, and the S amplicon was 138 bp. Genotyping was accomplished in two stages. Stage 1 distinguished S from L alleles. Stage 2 distinguished $L_A$ from $L_G$ alleles. We used a total of four fluorogenic probes, two probes for each stage. For stage 1, the allele-discriminating probe (ADP) was capable of hybridizing once, and once only, to the 43-bp L insertion, and an internal control probe (ICP) hybridized to a sequence located within the same amplicon but specific to a divergent repeat found only once in the amplicon and not involved in the insertion/deletion. For stage 2, probes were designed that were specific to the $L_A$ and $L_G$ alleles. The fluorogenic probes were labeled at the 5′ end with either FAM or VIC (table 1).

PCR was performed in a 25-$\mu$l volume, with 25–50 ng DNA, 120 nM ADP, 80 nM ICP, PCR primers (100 nM of each), dimethyl sulfoxide (DMSO) (4% by volume), and 1 × Master Mix (ABI). Amplification conditions were 2 min at 50°C, 10 min at 95°C, and then 40 cycles at 96°C for 15 s and at 62.5°C for 90 s. Genotypes were generated using ABI PRISM 7700 Sequence Detection system software. Stage 1 and stage 2 genotypes were combined to assign samples to one of six genotypes: SS, $SL_A$, $SL_G$, $L_AL_A$, $L_AL_G$, and $L_GL_G$. On each plate, previously sequenced standards were introduced. Stage 1 standards were SS, LS, and LL, and stage 2 standards were $L_AL_A$, $L_AL_G$, and $L_GL_G$.

To evaluate genotyping accuracy, one-quarter of the samples, randomly selected, were genotyped in duplicate. The error rate was <0.005, and the completion rate was >0.95.

### *Resequencing*

Genomic DNA was extracted from 106 lymphoblastoid cell lines derived from U.S. white controls. A 528-bp amplification product containing the entire *HTTLPR* region (L allele) was generated using a "touch-down" PCR procedure. Reaction mixtures contained 100 ng of genomic DNA; 0.2 mM each of dATP, dCTP, and dTTP; 0.1 mM dGTP; 0.1 mM 7-deaza-2′-dGTP; 0.4 $\mu$M of each primer; 10 mM Tris (pH 8.3); 50 mM KCl; 1.5 mM $MgCl_2$; 4% DMSO; and 1 U *Taq* DNA polymerase in a final volume of 25 $\mu$l. The reference sequence used

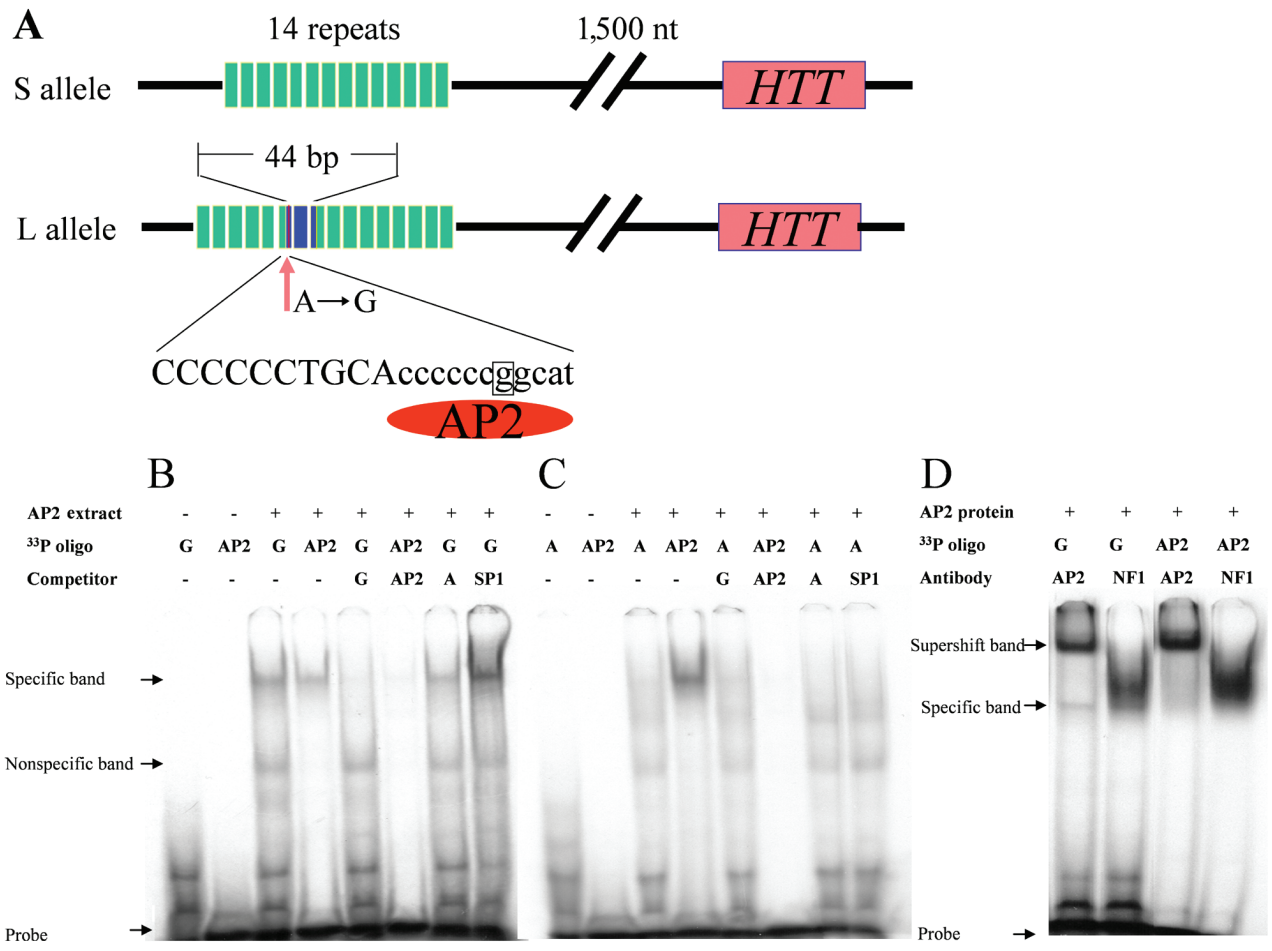


**Figure 1** *A*, Schematic of *HTTLPR* alleles. An A→G substitution (*boxed letter*) is located at nucleotide 6 within the first of two extra 22-bp repeats that characterize the L allele, located 1,629 nt 5′ of exon 1 of *HTT*.[8] This position corresponds to nucleotide 168 of the reference sequence (NCBI accession number AF126506). *B*, The $L_G$ allele creates an AP2-DNA binding site. The autoradiograph shows the presence of a specific AP2 DNA-protein complex in nuclear extracts from JEG-3 cells. The gel-shift assay used $^{33}$P-labeled $L_G$ double-stranded oligonucleotide probe (G). Addition of unlabeled competitor DNA (G) or an AP2 consensus sequence (AP2) resulted in disappearance of the specific band. Neither the $L_A$ competitor DNA (A) nor the SP1 competitor DNA (SP1) displaced $L_G$-specific binding. *C*, A specific DNA-protein complex was not seen with the $L_A$ double-stranded oligonucleotide probe. *D*, "Supershift" gel-shift assay. A combination of 5 μg anti-AP2 antibody and 5 μg recombinant AP2 further retarded the migration of the $L_G$-specific DNA-protein complex, as shown by the arrow indicating the supershift band. No supershift was seen with antibody to the NF1 transcription factor. $^{33}$P-labeled AP2 probe served as a reference.

to determine the genomic nucleotide positions and to design primers was National Center for Biotechnology Information (NCBI) accession number AF126506. The forward primer was 5′-GGCGTTGCCGCTCTGAATGC-3′, and the reverse primer was 5′-GAGGGACTGAGCTGGACAACCAC-3′. DNA was denatured at 95°C for 5 min. Annealing was at 63°C for 30 s for the first 2 cycles, at 62°C for 30 s for the next 2 cycles, and at 61°C for 30 s for another 36 cycles. Extension was consistently performed at 72°C for 1 min. PCR was terminated by extension at 72°C for 10 min. The products were purified using the MiniElute PCR Purification Kit (QIAGEN). Primers for bidirectional sequencing were the same as those used for PCR as described above. The 10-μl sequencing reaction mixture contained 4 μl BigDye Terminator RR Mix (ABI), 2.84 μl $dH_2O$, 1.6 pmol of forward or reverse primer, and 3 μl of purified *HTTLPR* PCR amplicon. Cycle conditions for sequencing were 25 cycles consisting of denaturation at 96°C for 10 s, annealing at 50°C for 5 s, and extension at 60°C for 4 min. Sequencing reaction products were purified using Performa DTR (Edge BioSystems) columns, were dried, were diluted with 25% formamide (v/v), were denatured at 95°C for 5 min, and were analyzed on a 3100 Genetic Analyzer (ABI).

### HTT *Expression in Lymphoblastoid Cell Lines*

*HTT* mRNA expression was measured in 62 lymphoblastoid cell lines derived from Finnish whites collected at the University of Helsinki as described above in "Human Subjects." The 62 cell lines were from 30 males and 32 females, with an average age of 37.6 years. Donors included individuals with various

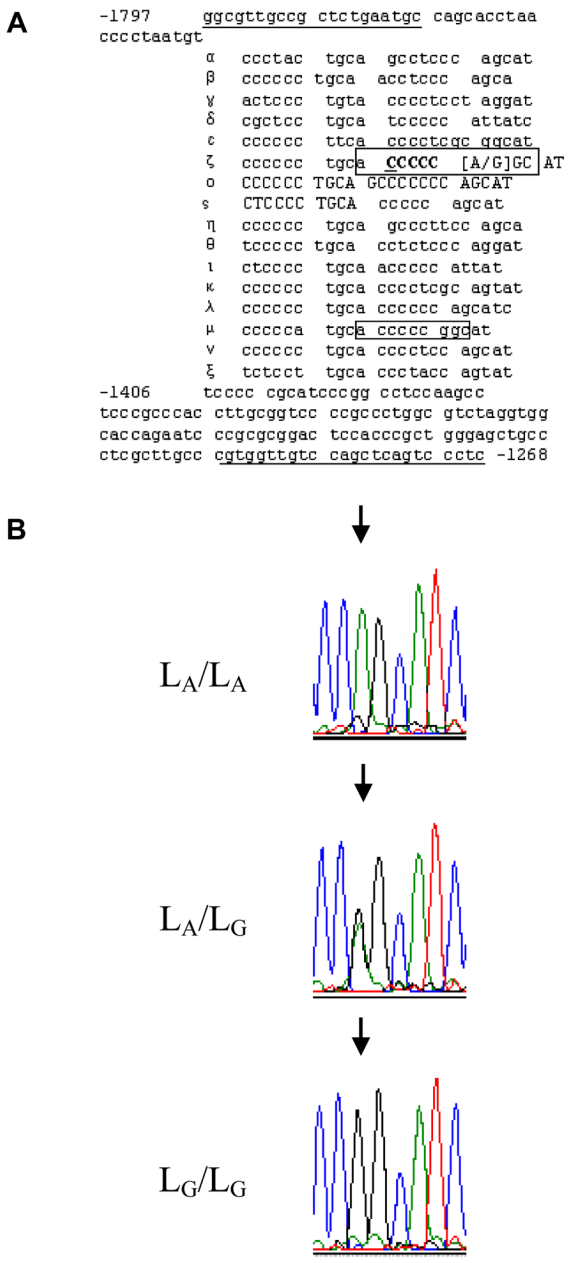


**Figure 2** *A*, Sequence and organization of *HTTLPR* repeats. Sequences from −1797 to −1268 relative to the *HTT* transcription start site are shown. This segment, encompassing the *HTTLPR* region, was inserted into the pGL4.10 expression plasmid. PCR primers used for genomic amplification are indicated in underlined lower-case letters. AP2 sites are boxed. Capital letters indicate the 43-bp insertion containing the A→G SNP, which is in brackets. Previously, this region was considered to be a 44-bp insertion/deletion, because of an additional C in the reference sequence (NCBI accession number AF116506). This region of sequence disparity is shown in bold letters (five Cs). Our sequencing results (106 samples) and those of Nakamura et al.[8] and Kraft et al.[18] confirm the sequence presented. A rare substitution within this region (C→T; frequency 0.0048 in present study and 0.007 in study by Nakamura et al.) is underlined. *B*, *HTTLPR* $L_A$ and $L_G$ alleles detected by direct sequencing. Representative electropherograms show the position of the A→G SNP (*arrows*).

psychiatric diagnoses; the most frequent were alcoholism, antisocial personality disorder, depression, and anxiety disorders. However, cell lines were selected for expression studies on the basis of genotype alone, because of substantial preexisting evidence that *HTTLPR* influences behavior and the vulnerability to psychiatric disease. Cells were cultured in 15-$cm^2$ flasks with 4 ml RPMI (Roswell Park Memorial Institute medium) containing 25% fetal bovine serum, 2% 200-mM glutamine, and 1% gentamicin (5 mg/ml) at 37°C, with 5% $CO_2$. After 48 h, 100 μM forskolin was added, and cell culture was terminated after 72 h.

### *Total RNA Extraction*

Cells were harvested by centrifugation at 2,300 rpm for 3 min at 4°C. Pellets were homogenized in 0.75 ml TRIzol (Invitrogen). Samples were incubated at room temperature for 5 min, were shaken vigorously for 1 min after 0.2 ml of chloroform was added, were held at room temperature for another 2.5 min, and then were incubated at −80°C for at least 1 h. The supernatant was transferred to 1.2-ml microfuge tubes (0.5 ml/tube) after centrifugation at 4,000 rpm for 15 min at 4°C, and a 53% volume of 100% ethanol was added. Then, samples were applied to RNeasy minicolumns (QIAGEN), and the manufacturer's protocol was followed. Total RNA concentration was measured spectrophotometrically, and samples were stored at −80°C for no more than 1 wk.

### *Quantitation of* HTT *by Real-Time PCR*

One μg of total RNA was used to synthesize first-strand cDNA by use of the Cloned AMV First-Strand Synthesis Kit (Invitrogen) and the manufacturer's protocol. *HTT* mRNA was measured on an ABI PRISM 7700 Sequence Detection System. Primers and probes were designed on the basis of NCBI *SLC6A4* cDNA reference sequence NM_001045. The forward primer was 5′-GAAACCCAATTGGCAGAAACTC-3′, and the reverse primer was 5′-GAAGATCTGAGCGGCTG-CAT-3′. The intron-spanning probe was 6FAM-ATCCACAC-CCCTGTCT-MGBNGQ. The endogenous reference was 18S rRNA (ABI). The 25-μl real-time PCR volume included 12.5 μl TaqMan Universal PCR Master Mix, 1.25 μl of 20× *HTT* mix, 1.25 μl of 20× 18S mix, 2 μl of products from the first-strand cDNA synthesis, and 8 μl RNase-free water. The thermal cycling conditions were 40 cycles at 95°C for 15 s and at 60°C for 1 min. Amplification was followed by a hold at 50°C for 2 min and at 95°C for 10 min. *HTT* mRNA concentration was estimated by the difference between *HTT* Ct (the amplification cycle at which the signal exceeded the background) and 18S Ct, which yields an *HTT*:18S rRNA ratio and a relative *HTT* concentration, by use of the equation $2^{-\Delta\Delta Ct}$, because per-cycle amplification efficiency was almost equal to 100%. Each of the 62 cell lines was assayed in quadruplicate.

### *Gel-Shift Assays*

Nuclear extracts containing AP2 were prepared from JEG-3 cells (GENEKA Biotechnology) and were stored at −80°C. For "supershift" experiments, recombinant AP2 (Promega) was used. Single-stranded, complementary, 26-bp oligonucleotide sequences corresponding to the *HTTLPR* $L_A$ and $L_G$ allele

AP2-binding motifs were annealed to form double-stranded probes. For the $L_A$ allele, the oligonucleotide was 5′-gatcgaacacCCCCCaGCAggcccgt-3′ (the upper strand, with the binding site in capital letters and the $L_A$ allele shown as a lower-case "a" within the binding site). For the $L_G$ allele, the oligonucleotide was 5′-gatcgaacacCCCCCgGCAggcccgt-3′ (with the $L_G$ allele shown as a lower-case "g" within the binding site). The "consensus" AP2 oligonucleotide was 5′-gatcgaactgaccGCCCGCGGCccgt-3′ (Promega) and was based on the consensus AP2-binding site (5′-GCCNNNGGC-3′). The AP2 antibody for the "supershift" assay was from GENEKA Biotechnology. [$\gamma$-$^{33}$P]-ATP (Perkin-Elmer Life Science) probes for gel-shift assays were generated by labeling double-stranded oligonucleotides with T4 polynucleotide kinase (Promega) after annealing the complementary oligonucleotide and purifying it through a Bio-Spin 30 chromatography column (Bio-Rad). Binding assays were performed for 20 min at room temperature by use of 2 $\mu$l nuclear extract (10 $\mu$g total protein) in 15-$\mu$l reactions containing binding buffer (10 mM HEPES [N-2-hydroxyethylpiperazine-N′-2-ethanesulfonic acid; pH 7.8], 50 mM KCl, 5 mM $MgCl_2$, 1 mM EDTA, and 5% glycerol), 1 mM dithiothreitol, 0.25 $\mu$g poly[dI:dC], and 0.2 $\mu$l radiolabeled probe (1.75 pmol; 100,000 cpm/$\mu$l). In competition experiments, unlabeled competitor (50 × molar excess) was included with the nuclear extract for 10 min before the addition of labeled probe. In the "supershift" experiments, specific antibody was incubated with AP2 for 25 min at room temperature, after probe addition. Complexes were separated on a Novex 6% DNA retardation gel (0.5 × Tris-borate-EDTA) at 300 V for 16 min at room temperature. Gels were dried and exposed to x-ray film for 12 h.

### *Luciferase Reporter Constructs, Transfection, and Expression Assays*

Genomic DNA samples with S, $L_G$, or $L_A$ alleles were selected for PCR-based cloning of *HTTLPR*. PCR primers and conditions were the same as for resequencing. The PCR products (529 bp for $L_G$ and $L_A$; 486 bp for S) were cloned into pCR-TOPO vector (Invitrogen) and then were subcloned into pGL4.10 (Promega) by standard molecular techniques. Sequences of inserts were confirmed by direct sequence analysis using M13 forward and reverse primers. The RN46A raphe-derived cell line was generously provided by Scott R. Whittemore (University of Louisville School of Medicine) and was maintained in Dulbecco's modified Eagle medium (DMEM) and Ham's F12 (1:1 v/v) with 5% fetal calf serum at 33°C in a humidified incubator (95% air, 5% $CO_2$). RN46A cells were seeded at a density of $2 \times 10^5$ cells per well into a 12-well tissue-culture plate. When the cells were 90% confluent, the *HTTLPR*-pGL4.10 (firefly luciferase) and pRL-SV40 (*Renilla* luciferase) constructs were cotransfected using Lipofectamine 2000 (Invitrogen). For AP2 transcription-factor competition experiments, a double-stranded decoy DNA was also cotransfected. Sequences used to prepare the $L_G$ and $L_A$ decoy DNAs were identical to those used for gel-shift assay probes. The cells were harvested for the Dual-Luciferase Reporter Assay System (Promega) by passive lysis buffer 24 h after transfection. The activities of both firefly luciferase and *Renilla* luciferase were quantified as relative light units (RLU). Each of the three *HTTLPR*-pGL4.10 constructs or an empty pGL4.10 plasmid was transfected in triplicate. The ratio of firefly RLU to *Renilla* RLU in a sample from each well was determined, and the values were normalized to the S allele.

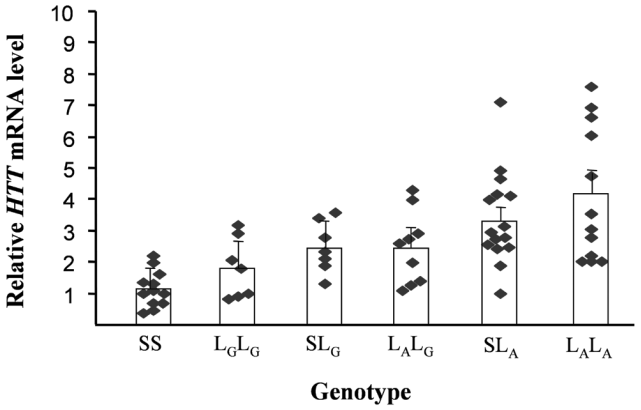


**Figure 3** Effect of *HTTLPR* S, $L_A$, and $L_G$ alleles on *HTT*-specific mRNA expression in lymphoblastoid cell lines. Cell lines with genotypes SS ($n$ = 12), $SL_A$ ($n$ = 15), $SL_G$ ($n$ = 7), $L_AL_A$ ($n$ = 12), $L_AL_G$ ($n$ = 9), and $L_GL_G$ ($n$ = 7) were identified. Mean *HTT* mRNA expression values were measured by real-time PCR in triplicate cultures and were normalized to 18S rRNA levels, as described in the "Material and Methods" section.

### *Data Analysis*

*HTT* allele and genotype expression values were computed by multiple regression using data from 62 lymphoblastoid cell

**Table 2**

**Distribution of *HTTLPR* Genotype and Allele Frequencies in Several Populations**

| | | Frequency of Genotype | | | | | | Frequency of Allele | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Population | *N* | SS | $SL_A$ | $SL_G$ | $L_AL_A$ | $L_AL_G$ | $L_GL_G$ | S | $L_A$ | $L_G$ |
| Finnish whites | 771 | .15 | .44 | .08 | .24 | .09 | .01 | .40 | .51 | .09 |
| U.S. whites group 1 | 297 | .16 | .33 | .09 | .26 | .14 | .03 | .37 | .49 | .14 |
| U.S. whites group 2 | 286 | .12 | .37 | .08 | .22 | .18 | .02 | .35 | .50 | .15 |
| Plains Indians | 456 | .42 | .47 | .01 | .09 | .01 | .00 | .66 | .33 | .01 |
| Southwest Indians | 564 | .42 | .43 | .01 | .13 | .01 | .00 | .64 | .35 | .01 |
| African Americans | 624 | .07 | .25 | .12 | .27 | .23 | .06 | .25 | .51 | .24 |

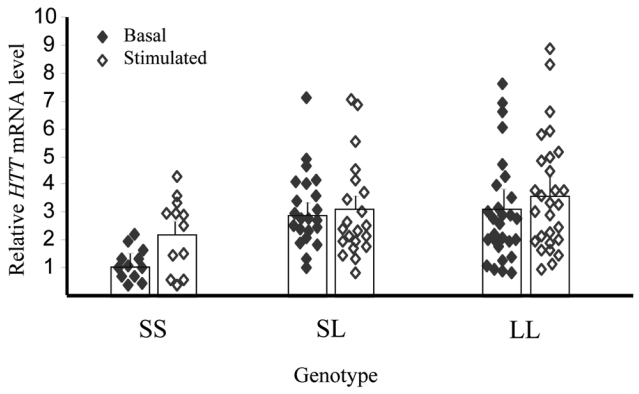


**Figure 4** Relative *HTT* mRNA levels in 62 human lymphoblastoid cell lines with *HTTLPR* genotypes SS ($n = 12$), SL ($n = 22$), and LL ($n = 28$), before and after treatment with 100 $\mu$M forskolin.

lines (7–15 cell lines per genotype). Genotype-specific expression was estimated assuming codominance (i.e., additive effects of alleles in a genotype). To test for dominance and/or recessivity, these predicted genotype expression values were compared, by linear regression, with the expression observed in the cell lines. Analysis of variance was used to calculate genotype-attributable variance.

Association with OCD was evaluated using $\chi^2$ and logistic regression analyses. First, association was tested using the information that $L_A$ is a high-expressing allele and by combining the S and $L_G$ alleles, which are low and closely equivalent in expression. Thus, there were again three genotypes. Second, linkage was evaluated using *HTT* expression scores predicted by genotype.

## Results

### *Single-Nucleotide Variants in* HTTLPR

A common single-base substitution (A→G) occurs at the sixth nucleotide within the first of two extra 20–23-bp repeats in the L allele (fig. 1). The two L alleles $L_A$ and $L_G$, together with the S allele, comprise a triallelic locus. The A→G substitution is located 1,629 bp upstream of the transcription start site (fig. 2*A*) and was genotyped in the second step of a two-stage 5′-nuclease assay for the triallelic locus. $L_G$ corresponds to 1 of 10 SNPs detected within *HTTLPR* by Nakamura et al.[8] (fig. 2*A* and 2*B*). The alleles corresponding to $L_A$ and $L_G$ were designated "$\zeta$" and "$\mu$" by those investigators. In several populations, all three alleles are abundant or highly abundant, as will be described below. Other novel alleles detected by Nakamura et al.[8] were rare, and we did not observe them, even though we resequenced the *HTTLPR* region in 106 clinically and ethnically diverse individuals.

### HTTLPR *Allele and Genotype Distributions in Populations*

Using a two-stage 5′-nuclease assay, we determined triallelic *HTTLPR* genotypes and allele frequencies in five populations, comprising 2,998 total individuals: 624 African Americans, 771 Finnish whites, 583 U.S. whites, 456 Plains Indians, and 564 Southwest Indians. The Southwest Indians, Plains Indians, and Finnish whites are populations that are relatively nonadmixed. The frequency of S was lowest in individuals of African descent (0.25), was intermediate in whites (0.35–0.40), and was highest in American Indians (0.64–0.66). The $S:L_A:L_G$ ratio was 4:5:1 in whites, 2.5:5:2.5 in African Americans, and 2:1:0 in American Indians (i.e., the $L_G$ allele was nearly absent in both American Indian populations) (table 2).

### HTTLPR *Genotype–Attributable Variance*

To estimate the proportion of variance in *HTT* expression that is attributable to *HTTLPR* in natural populations, simulation analyses were performed with an $N$ of 100,000 to represent various populations. Genotypes were assigned on the basis of allele frequencies (e.g., frequency of the $L_GL_G$ genotype was set at 0.0001 in American Indians but at 0.0586 in African Americans). Expression values were randomly assigned for each individual on the basis of genotype, expression value of each genotype, and SD of genotype-specific expression. Genotype-attributable variance was then recalculated. With knowledge of S and L alleles alone, genotype-attributable variance was 0.22 in whites, 0.09 in African Americans, and 0.44 in American Indians. In American Indians, the $L_G$ allele is rare. Knowledge of the $L_G$ allele substantially improved prediction of *HTT* expression in other populations, increasing to 0.30 in whites, 0.25 in African Americans, and 0.45 in American Indians.

### *Creation of an AP2 Site by the $L_G$ Allele*

The $L_G$ allele creates a potential binding site for transcription factor AP2 (fig. 1*A*). These nucleotides are ab-

**Table 3**

***P* Values for Significance of Effects of *HTTLPR* Genotypes on *HTT* mRNA Expression**

| *HTTLPR* Genotype | *P* for Genotype | | | | |
|---|---|---|---|---|---|
| | $SL_G$ | $L_GL_G$ | $L_AL_G$ | $SL_A$ | $L_AL_A$ |
| SS | … | … | .034 | .00007 | .000002 |
| $SL_G$ | | … | … | … | .0157 |
| $L_GL_G$ | | | … | .012 | .0009 |
| $L_AL_G$ | | | | … | .009 |
| $SL_A$ | | | | | … |

Note.—Results are for *HTT* mRNA expression in 62 lymphoblastoid cell lines as shown in figure 3.

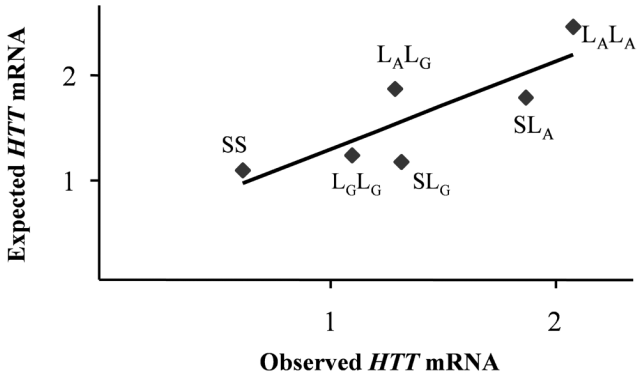


**Figure 5** Codominance of *HTT* mRNA expression. The codominance of *HTT* mRNA expression normalized to 18S mRNA was observed in 62 lymphoblastoid cell lines representing the six *HTTLPR* genotypes. Expression values represent the means of 7–15 cell lines per genotype. Observed expression is plotted versus the expression expected on the basis of the codominant model and of computed allele-expression values: 0.50 for S, 0.57 for $L_A$, and 1.16 for $L_G$. For observed vs. expected, $r = 0.84$ and $P = .038$.

sent in the S allele. To determine whether AP2 binds to $L_A$ or $L_G$, we performed gel-shift assays, using nuclear extracts enriched for AP2, prepared from JEG-3 cells. Only the $L_G$ double-stranded oligonucleotide specifically bound AP2 (fig. 1*B* and 1*C*). AP2 was part of the protein-DNA complex, as confirmed by a gel "supershift" experiment using an antibody specific to AP2 (fig. 1*D*).

### *Quantitation of* HTT *mRNA Levels by Real-Time PCR*

*HTTLPR* predicts *HTT* expression in lymphoblasts, reporter constructs,[6] and antemortem[14] and postmortem[12] human brain. We found that *HTT* mRNA levels varied across the six *HTTLPR* genotypes in 62 lymphoblastoid cell lines (figs. 3 and 4). The lowest-expressing genotype was SS, whereas $L_AL_A$ was the highest expressing, differing significantly from each of the five genotypes other than $SL_A$ (fig. 3 and table 3). The normalized (to SS genotype) expression value of each allele was calculated using all the data, and these values were 0.50 for the S allele, 0.57 for the $L_G$ allele, and 1.16 for the $L_A$ allele.

The *HTTLPR* alleles are codominant in action, as shown in figure 5, in which the expression value expected (on the basis of codominance) for each genotype is plotted versus the observed level. Expression was predicted by the codominant model ($r = 0.84$). In addition, the two heterozygous genotypes that contain one copy of the S allele did not exhibit lower-than-expected expression, as predicted by the S-dominant model.

Residual variation in *HTT* expression that is visible within each genotype group may be attributable to a second polymorphism at *HTT* or elsewhere. A potential origin is an intron 2 17-bp VNTR.[19] However, the mean levels of relative *HTT* mRNA for the VNTR genotypes were 2.81 ± 1.29 for genotype 9/10, 3.23 ± 0.81 for 10/10, 2.88 ± 0.40 for 10/12, and 2.53 ± 0.33 for 12/12 ($P > .46$) in the 62 lymphoblastoid cell lines (fig. 6). These findings are consistent with a previous report.[19]

### *Functional Equivalence of $L_G$ and S in Transfected Cells*

We introduced *HTTLPR*-luciferase plasmids (*HTTLPR*-pGL4.10) into RN46A cells derived from rat dorsal raphe neurons[20] and determined the effect of each allele on basal transcription. The 2.8-fold difference produced by the S and $L_A$ allele constructs is similar to the allele-specific effect seen by Mortensen et al.,[21] who used a different and larger reporter construct but the same recipient cell line. Under basal conditions, the $L_G$ construct had half the relative luciferase activity as did the $L_A$ allele (fig. 7*A*). The S and $L_G$ constructs did not differ in basal expression. All three reporter constructs were equally activated by 50 $\mu$M forskolin, indicating that the *HTTLPR* region was activated normally (fig. 7*A*).

### *Equalizing of $L_A$- and $L_G$-Driven Transcription by an $L_G$-Based "Decoy" DNA*

$L_G$ had reduced transcriptional efficiency relative to that of $L_A$ and $L_G$, as a result of the creation of an AP2-binding site. Therefore, we tested whether this AP2 site acts as a transcriptional suppressor. We depleted cells of AP2 by introducing into RN46A cells carrying S, G, or A allele constructs a double-stranded DNA "decoy" containing the AP2-binding sequence from the $L_G$ allele (Oligo-G). A decoy DNA having the same sequence but carrying the $L_A$ allele (Oligo-A) was used for comparison. Twenty-four hours after transfection, luciferase activity was measured. The Oligo-G treatment equalized $L_G$ and $L_A$ allele reporter expression (fig. 7*B*), whereas

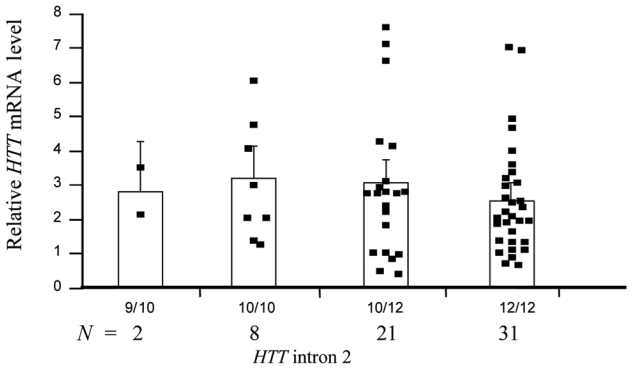


**Figure 6** Relative *HTT* mRNA levels in lymphoblastoid cell lines with various intron 2 VNTR genotypes.

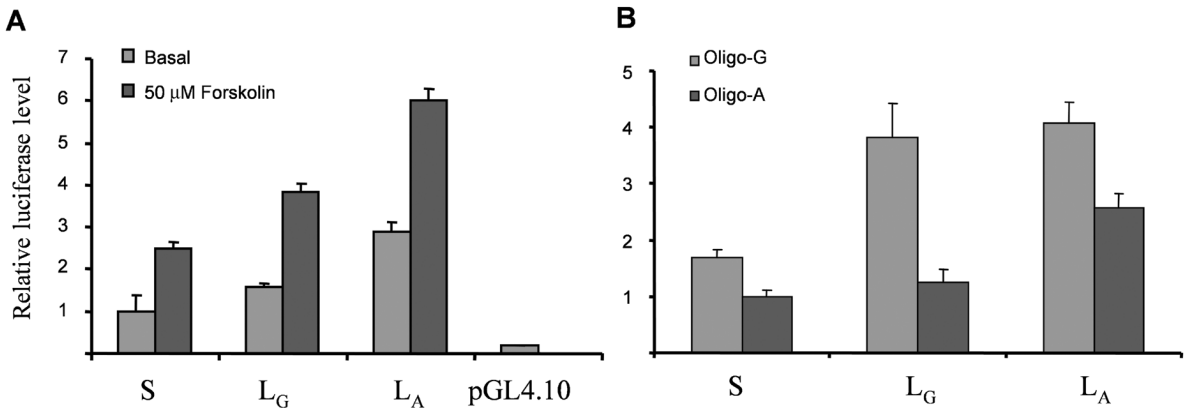


**Figure 7** Reduction of basal *HTTLPR* transcriptional activity by the $L_G$ allele relative to the $L_A$ allele. *A*, Effect of three *HTTLPR* alleles on reporter transcription. RN46A cells were transfected with *HTTLPR*-luciferase plasmids (pGL4.10). The effect of each allele on basal transcription was measured as relative luciferase level (ratio of firefly luciferase to *Renilla* luciferase) and was normalized to basal S allele expression. There was no statistically significant difference between S and $L_G$. However, $L_A$ was significantly different from both S and $L_G$ ($P < .01$). Transfected cells carrying each of the three different reporter plasmids and treated with 50 $\mu$M forskolin showed an approximately twofold increase in relative luciferase level, compared with the basal condition. *B*, Effect of an $L_G$ decoy on expression of three *HTTLPR* reporters. The double-stranded $L_G$-specific decoy DNA (Oligo-G) containing the AP2-binding sequence was transfected into RN46A cells carrying the S, $L_G$, or $L_A$ allele in *HTTLPR*-pGL4.10 reporter plasmids. An $L_A$-specific decoy DNA (Oligo-A) was used for the control. Twenty-four hours after transfection, luciferase activity was measured and normalized for transfection efficiency. Results are presented relative to expression determined for Oligo-A–treated cells carrying the S allele reporter plasmid. Treatment with Oligo-G decoy DNA equalized $L_G$ and $L_A$ allele reporter expression, whereas the control decoy DNA had little effect (compare basal expression in panel A with expression of Oligo-A–treated cells in panel B).

the Oligo-A decoy DNA had little effect (compare basal expression in panel A with expression of Oligo-A–treated cells in panel B in fig. 7).

### *Linkage of Gain-of-Function* HTTLPR *Genotype to OCD in Case-Control Sample*

*HTTLPR* genotype and allele frequencies were compared in 169 patients with OCD and 253 controls. The higher expressing $L_A$ allele was associated with OCD ($\chi^2 = 6.64$; $P = .036$) (table 4). We also took advantage of knowledge of *HTTLPR* allele function, grouping the $L_G$ and S alleles—which are low and closely equivalent in expression—and comparing them with the high-expressing $L_A$ allele ($P < .021$ for allele-based association with OCD). Patients with OCD were twice as likely to have the highest-expressing genotype $L_AL_A$ (table 4). Finally, linkage was performed using genotype-predicted expression scores. The mean ($\pm$SD) predicted expression score in patients with OCD was higher than in controls ($2.76 \pm 0.80$ vs. $2.59 \pm 0.59$; $P = .024$, by two-tailed test).

### *Replication of Linkage in TDT Data Set*

TDT analysis of 86 informative OCD-affected child-parents trios derived from 175 pedigrees (total $N = 581$) in the Toronto OCD Study revealed that the $L_A$ allele was twofold overtransmitted relative to the lower-expressing alleles (S and $L_G$) ($P < .023$). The strongest result was again obtained after function-based grouping of the S and $L_G$ alleles ($P = .010$) (table 5).

## Discussion

The *HTTLPR* promoter polymorphism is linked to behavior. However, there are many contradictory findings, and the bulk of the findings involve the loss-of-function S allele. Contradictions arise from differences in source populations, methods of phenotypic assessment, and chance. In some studies, ethnic variation could mask or create associations. In our study, it is unlikely that the linkages of OCD to the gain-of-function allele and genotypes are biased by population stratification, because TDT allele-based linkage in complete parents-child trios is invulnerable to these effects[22] and because, in the case-control study, all subjects were white.

Our results point to an important additional origin of variability—namely, unrecognized functional allelic variation. Because $L_G$ drives expression nearly equivalently as S, studies that include many $L_G$ alleles within SL and LL genotypes may underestimate the effect of *HTTLPR*. $L_G$ creates a binding site for AP2, one of many nuclear factors that function as transcriptional activators or repressors.[23–26] This is based on three lines of evidence: basal mRNA levels in lymphoblasts, allele-specific expression in the transfected neuronal cell line, and the

**Table 4**

**Genotype- and Allele-Based Association of *HTTLPR* with OCD**

| Group, Frequency Measured, and Genotype or Allele | OCD Cases ($n$ = 169) | Controls ($n$ = 253) |
|---|---|---|
| S, $L_A$, and $L_G$ alleles: | | |
| Genotype frequency[a]: | | |
| SS | .21 | .16 |
| $SL_A$ | .34 | .47 |
| $SL_G$ | .03 | .08 |
| $L_GL_G$ | .01 | .02 |
| $L_AL_G$ | .07 | .08 |
| $L_AL_A$ | .34 | .19 |
| Allele frequency[b]: | | |
| S | .38 | .44 |
| $L_G$ | .06 | .10 |
| $L_A$ | .56 | .47 |
| Allele expression grouping: | | |
| Genotype frequency[c]: | | |
| SS, $SL_G$, $L_GL_G$ | .25 | .26 |
| $SL_A$, $L_AL_G$ | .41 | .55 |
| $L_AL_A$ | .34 | .19 |
| Allele frequency[d]: | | |
| $SL_G$ | .45 | .53 |
| $L_A$ | .55 | .47 |

NOTE.—The *HTTLPR* $L_A$ (high expressing) allele is more abundant in OCD cases, as is the $L_AL_A$ (most highly expressing) genotype. Because $L_G$ and S alleles are closely equivalent in *HTT* expression, these alleles were grouped and were compared with the $L_A$ allele, also revealing significant genotype-based and allele-based linkage.

[a] $\chi^2$ = 24.7 (5 df); $P$ = .0001.
[b] $\chi^2$ = 6.69 (2 df); $P$ = .035.
[c] $\chi^2$ = 13.6 (2 df); $P$ = .0011.
[d] $\chi^2$ = 5.32 (1 df); $P$ = .021.

AP2 decoy-DNA effect in transfected cells, with the $L_G$-specific AP2 site acting as a suppressor. However, this is not the only AP2 site within *HTTLPR*. The S and $L_A$ reporter constructs also respond to AP2 decoy DNA (fig. 7*B*). Nonetheless, the responses for those constructs were weaker than those for the $L_G$ construct. A second, untested AP2 site is predicted in all three alleles, located 135 bp 3′ of the 43-bp insertion that creates the L allele. Nakamura et al.[8] tested the function of the $L_A$ and $L_G$ alleles, also transfecting them into RN46A cells. In their hands, the $\zeta$ and $\mu$ alleles (as they designated them) were indistinguishable, although both alleles acted as transcriptional suppressors.[27] Some functional analyses of the A→G SNP by Sakai et al.[27] were performed on the S allele background rather than on the L allele sequence context we consistently observed. Probably for this reason, the existence of the A→G substitution was subsequently ignored. Recently, the G allele was reported on the S allele background in a largely white population. The frequency of "$S_G$" in that study was 5/200 chromosomes.[18] We did not observe "$S_G$," although the two-stage genotyping method we used should have detected it. Furthermore, we sequenced 106 whites and never observed it. Nakamura et al.[8] observed it in one Japanese individual and in none of the 148 whites they sequenced. Therefore, we refer to the A and G alleles as $L_A$ and $L_G$, emphasizing that G is a sequence variation that almost always occurs in the context of the L allele, in which it alters AP2 binding and suppresses transcription. If the G substitution is observed in the context of the S allele, we suggest that it be called "$S_G$" or "14-B" (Nakamura et al.'s designation).

The story of the influence of *HTTLPR* S, $L_A$, and $L_G$ alleles on *HTT* expression is one of incremental progress in the genetic characterization of a single neurochemical function. In lymphoblasts, substantial variation in *HTT* expression is explained by *HTTLPR,* but a large amount of variation has some other origin. Although *HTTLPR* predicts variation in *HTT* expression in the brain, there was a similar amount of within-genotype variability.[28] We predict that *HTTLPR* genotypes incorporating $L_G$ will also better predict *HTT* expression in the brain, but, again, there will be substantial residual variance. Assays performed in triplicate seem to indicate that <8% of variation in expression is attributable to measurement error. The remaining variance may be due to differences in expression associated with cell culture (which can be considered another type of measurement error) and cell line–specific factors, including other genetic loci. Although an intron 2 VNTR domain has allele-dependent enhancer activity,[29,30] this locus does not independently predict *HTT* expression,[19] as we confirmed in the present study.

Unrecognized $L_G$ alleles in SL and LL genotypes can create the appearance of S allele dominance because certain of these $L_G$-containing genotypes tend not to be high expressing. The codominant allele action that we observed is, on the other hand, consistent with the general pattern of *cis*-action of promoter alleles. Probably the main result of not scoring the $L_G$ allele is to obscure effects of *HTTLPR* on phenotype, especially for phe-

**Table 5**

**TDT for Association of *HTTLPR* with OCD**

| | Triallelic Analysis | | | Low/High Analysis[a] | |
|---|---|---|---|---|---|
| Outcome | S | $L_G$ | $L_A$ | S, $L_G$ | $L_A$ |
| Transmitted | 27 | 11 | 48 | 20 | 41 |
| Untransmitted | 44 | 16 | 26 | 41 | 20 |

NOTE.—Results are from 86 informative parents-child trios derived from 175 pedigrees. Triallelic analysis $P$ = .023; low/high analysis $P$ = .010.

[a] The low/high analysis grouped the S and $L_G$ alleles because they have low and nearly equivalent effect on *HTT* expression.

notypes (such as OCD) for which the $L_AL_A$ (highest expressing) genotype is crucial and perhaps less so for phenotypes (such as anxiety and dysphoria) for which associations to the low-expressing S allele have emerged.[6,10,11] However, although S allele linkages to anxiety and dysphoria are coherent with genotype effects on brain morphometry and brain function after emotional challenges, and even though these findings have been robust in relatively small imaging data sets, the relationship of *HTT* to anxiety and dysphoria is more ambiguous.[2] The contradictory findings for complex behavior, even in large populations, may be attributable to a small effect size. A meta-analysis revealed that each copy of the S allele leads to an increment of only 0.106 SD in dimensional anxiety.[31] Gene-environment interaction is among the confounding factors. In the absence of stress, the SS homozygous genotype does not appear to lead to depressive symptoms in humans,[11] and somewhat parallel findings have been observed in studies of the *Rhesus macaque* monkey, which has an orthologous polymorphism (*rh-HTTLPR*).[32] In the human, changes in brain functional activation and interregional coupling predicted by the S allele were seen after cognitive fear challenge but not at baseline.[10] Therefore, unstressed populations may not so readily reveal effects of *HTTLPR* on dysphoria and anxiety.

OCD is a chronic and disabling disorder characterized by recurrent, intrusive thoughts that cause distress and interfere with function and by repetitive behaviors or mental acts performed in response to obsession. OCD occurs with a frequency of ~2% in the U.S. population and is therefore the fourth-most-prevalent psychiatric disorder.[15,33] Often of early onset, OCD affects patients throughout their lives, leading to diminished quality of life for patients and families, reduced productivity, and high health care costs. Cognitive behavioral therapy and selective serotonin reuptake inhibitor (SSRI) drugs are partially effective[34] and have led to an interest in refining knowledge of the etiology, which is largely unknown. Although OCD is moderately heritable, replicated genes for OCD are unknown, with the possible exception of the uncommon 425Val allele in *HTT*. Gene discovery could have profound implications for understanding the neurobiology of OCD and, potentially, for treatment.

The neuroanatomical areas thought to be involved in the pathogenesis of OCD are the frontocortical-subcortical pathways, including orbitofrontal cortex, basal ganglia, and thalamus.[33] The hypothesis of a serotonergic dysfunction in OCD derives largely from the clinical observation that 5-HT reuptake–inhibiting drugs are partially effective in treating OCD.[35] Recently, Pogarell et al.[36] applied single-photon emission computed tomography to a small group of unmedicated patients with OCD and control subjects and observed a 25% increase in the specific binding of [$^{123}$I]$\beta$-CIT (which accesses the serotonin transporter density) in the midbrain-pons of the patients with OCD. These unreplicated results suggest that HTT level is increased in OCD and are consistent with the linkage of a common high-expressing *HTT* allele to OCD that we report here and with linkage of a rare higher-functioning *HTT* allele, Ile425Val, to treatment-resistant OCD and other severe psychopathologies.[15]

It is interesting, and a seeming paradox, that the loss-of-function S allele has been linked to anxiety and dysphoria, which, like OCD, are treated with SSRI drugs, and that the gain-of-function $L_A$ allele is shown here to have a role in OCD. OCD is frequently classified as an anxiety disorder; however, anxiety has many origins. Recent studies of *HTTLPR* in anxiety and depression indicate that the S allele leads to poorer coupling of the amygdala to regions of the brain, such as the cingulate cortex, that modulate the amygdala and thereby modulate emotional responses.[3,4] People ordinarily have the ability to regulate emotional responses, but this regulation is relatively impaired in certain individuals; in fact, the degree of cingulate/amygdala decoupling correlated with depression in one of these studies.[4] Thus, anxiety and dysphoria attributable to low serotonin transporter expression appear to be a consequence of release of the amygdala from cortical regulation and especially inhibition. It has never been clear why, if the low-expressing S allele causes anxiety and dysphoria, there would be a positive therapeutic response to an SSRI drug, which of course further inhibits serotonin transporter function.[37] However, there is some evidence that depressed patients with the loss-of-function S genotypes respond less well to SSRI drugs, indicating again that serotonin transporter inhibition is a key to the efficacy of SSRI drugs in these diseases.[38]

OCD is not a typical anxiety disorder. In OCD, emotional distress appears to be a consequence of recurrent, intrusive thoughts. Several neurobiological models of OCD postulate a primary role for dysfunction of the anterior cingulate gyrus, whose hyperactivity could play a role in abnormal conflict detection as part of an overactive action-monitory system in OCD. Other studies reported hypermetabolism in the cingulate gyrus,[39–41] with decreased activation after treatment with SSRIs,[42,43] whereas increased anterior cingulate gray matter volume was found in patients with OCD.[44–46] Therefore, the gain-of-function *HTTLPR* genotype may predict hyperactivity and more gray matter in anterior cingulate gyrus in patients with OCD. These proposed HTT-neurobiology relationships need to be further tested in the specific context of OCD through combined genetic and neuroimaging studies.

Although the difference between the predicted *HTTLPR* expression scores for patients with OCD and controls is modest, our data support a role for HTT in

OCD, with linkage between the disorder and high HTT function. The *HTTLPR* genotype has a 1.8-fold effect on relative risk. This is substantially less than the risk attributable to being a first-degree relative of a patient with OCD. Clearly, other loci are involved in the heritability of OCD, and inheritance of the $L_AL_A$ genotype alone is insufficient to produce OCD. We predict that studies of other loci controlling the still-unaccounted-for variance in *HTT* expression will further refine our understanding of the neurobiology of OCD and the relationship of HTT to this disease. This could even assist in prevention and treatment of this severe disease. In this regard, it could be critical to combine information from genotype, clinical features of OCD (including treatment response), and neurobiologic measures (i.e., frontal/amygdala connectivity) in attempts to trace the origin of this psychiatric disease to neurobiological determinants.

## Acknowledgments

We acknowledge Matti Virkkunen (University of Helsinki), Alec Roy (Rutgers University), Rob Robin (Sitka, AK), and Mary-Anne Enoch (NIAAA), for providing population samples; Zhifeng Zhou (NIAAA), for assistance with sequencing; and Joseph Trakalo (University of Toronto), for data management and coordination in the Toronto OCD Study. Support was provided for the Toronto site (to P.D.A., M.A.R., and J.L.K.) by the Ontario Mental Health Foundation (type B grant) and by the Canadian Institutes for Health Research (grant MOP-38077).

## Web Resources

Accession numbers and URLs for data presented herein are as follows:

NCBI, http://www.ncbi.nlm.nih.gov/ (for reference sequences [accession numbers AF126506 and NM_001045])

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for OCD)

## References

1. Heils A, Teufel A, Petri S, Seemann M, Bengel D, Balling U, Riederer P, Lesch KP (1995) Functional promoter and polyadenylation site mapping of the human serotonin (5-HT) transporter gene. J Neural Transm Gen Sect 102:247–254
2. Glatt CE, Freimer NB (2002) Association analysis of candidate genes for neuropsychiatric disease: the perpetual campaign. Trends Genet 18:307–312
3. Heinz A, Braus DF, Smolka MN, Wrase J, Puls I, Hermann D, Klein S, Grusser SM, Flor H, Schumann G, Mann K, Buchel C (2005) Amygdala-prefrontal coupling depends on a genetic variation of the serotonin transporter. Nat Neurosci 8:20–21
4. Pezawas L, Meyer-Lindenberg A, Drabant EM, Verchinski BA, Munoz KE, Kolachana BS, Egan MF, Mattay VS, Hariri AR, Weinberger DR (2005) 5-HTTLPR polymorphism impacts human cingulate-amygdala interactions: a genetic susceptibility mechanism for depression. Nat Neurosci 8:828–834
5. Ursu S, Stenger VA, Shear MK, Jones MR, Carter CS (2003) Overactive action monitoring in obsessive-compulsive disorder: evidence from functional magnetic resonance imaging. Psychol Sci 14:347–353
6. Lesch KP, Bengel D, Heils A, Sabol SZ, Greenberg BD, Petri S, Benjamin J, Muller CR, Hamer DH, Murphy DL (1996) Association of anxiety-related traits with a polymorphism in the serotonin transporter gene regulatory region. Science 274:1527–1531
7. Delbruck SJ, Wendel B, Grunewald I, Sander T, Morris-Rosendahl D, Crocq MA, Berrettini WH, Hoehe MR (1997) A novel allelic variant of the human serotonin transporter gene regulatory polymorphism. Cytogenet Cell Genet 79:214–220
8. Nakamura M, Ueno S, Sano A, Tanabe H (2000) The human serotonin transporter gene linked polymorphism (5-HTTLPR) shows ten novel allelic variants. Mol Psychiatry 5:32–38
9. Heils A, Teufel A, Petri S, Stober G, Riederer P, Bengel D, Lesch KP (1996) Allelic variation of human serotonin transporter gene expression. J Neurochem 66:2621–2624
10. Caspi A, Sugden K, Moffitt TE, Taylor A, Craig IW, Harrington H, McClay J, Mill J, Martin J, Braithwaite A, Poulton R (2003) Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. Science 301:386–389
11. Hariri AR, Mattay VS, Tessitore A, Kolachana B, Fera F, Goldman D, Egan MF, Weinberger DR (2002) Serotonin transporter genetic variation and the response of the human amygdala. Science 297: 400–403
12. Little KY, McLaughlin DP, Zhang L, Livermore CS, Dalack GW, McFinton PR, DelProposto ZS, Hill E, Cassin BJ, Watson SJ, Cook EH (1998) Cocaine, ethanol, and genotype effects on human midbrain serotonin transporter binding sites and mRNA levels. Am J Psychiatry 155:207–213
13. Heinz A, Jones DW, Mazzanti C, Goldman D, Ragan P, Hommer D, Linnoila M, Weinberger DR (2000) A relationship between serotonin transporter genotype and in vivo protein expression and alcohol neurotoxicity. Biol Psychiatry 47:643–649
14. Greenberg BD, Tolliver TJ, Huang SJ, Li Q, Bengel D, Murphy DL (1999) Genetic variation in the serotonin transporter promoter region affects serotonin uptake in human blood platelets. Am J Med Genet 88:83–87
15. Ozaki N, Goldman D, Kaye WH, Plotnicov K, Greenberg BD, Lappalainen J, Rudnick G, Murphy DL (2003) Serotonin transporter missense mutation associated with a complex neuropsychiatric phenotype. Mol Psychiatry 8:895, 933–936
16. Kilic F, Murphy DL, Rudnick G (2003) A human serotonin transporter mutation causes constitutive activation of transport activity. Mol Pharmacol 64:440–446
17. Goodman WK, Price LH, Rasmussen SA, Mazure C, Fleischmann RL, Hill CL, Heninger GR, Charney DS (1989) The Yale-Brown Obsessive Compulsive Scale. I. Development, use, and reliability. Arch Gen Psychiatry 46:1006–1011
18. Kraft JB, Slager SL, McGrath PJ, Hamilton SP (2005) Sequence analysis of the serotonin transporter and associations with antidepressant response. Biol Psychiatry 58:374–381
19. Hranilovic D, Stefulj J, Schwab S, Borrmann-Hassenbach M, Albus M, Jernej B, Wildenauer D (2004) Serotonin transporter promoter and intron 2 polymorphisms: relationship between allelic variants and gene expression. Biol Psychiatry 55:1090–1094
20. Eaton MJ, Whittemore SR (1995) Adrenocorticotropic hormone activation of adenylate cyclase in raphe neurons: multiple regulatory pathways control serotonergic neuronal differentiation. J Neurobiol 28:465–481
21. Mortensen OV, Thomassen M, Larsen MB, Whittemore SR, Wiborg O (1999) Functional analysis of a novel human serotonin transporter gene promoter in immortalized raphe cells. Brain Res Mol Brain Res 68:141–148

22. Ewens WJ, Spielman RS (2004) The TDT is a statistically valid test: comments on Wittkowski and Liu. Hum Hered 58:59–60
23. Jiang JG, DeFrances MC, Machen J, Johnson C, Zarnegar R (2000) The repressive function of AP2 transcription factor on the hepatocyte growth factor gene promoter. Biochem Biophys Res Commun 272:882–886
24. Yen FC, Hong CJ, Hou SJ, Wang JK, Tsai SJ (2003) Association study of serotonin transporter gene VNTR polymorphism and mood disorders, onset age and suicide attempts in a Chinese sample. Neuropsychobiology 48:5–9
25. Eloranta JJ, Hurst HC (2002) Transcription factor AP-2 interacts with the SUMO-conjugating enzyme UBC9 and is sumolated in vivo. J Biol Chem 277:30798–30804
26. Braganca J, Eloranta JJ, Bamforth SD, Ibbitt JC, Hurst HC, Bhattacharya S (2003) Physical and functional interactions among AP-2 transcription factors, p300/CREB-binding protein, and CITED2. J Biol Chem 278:16021–16029
27. Sakai K, Nakamura M, Ueno S, Sano A, Sakai N, Shirai Y, Saito N (2002) The silencer activity of the novel human serotonin transporter linked polymorphic regions. Neurosci Lett 327:13–16
28. Little KY, McLauglin DP, Ranc J, Gilmore J, Lopez JF, Watson SJ, Carroll FI, Butts JD (1997) Serotonin transporter binding sites and mRNA levels in depressed persons committing suicide. Biol Psychiatry 41:1156–1164
29. Fiskerstrand CE, Lovejoy EA, Quinn JP (1999) An intronic polymorphic domain often associated with susceptibility to affective disorders has allele dependent differential enhancer activity in embryonic stem cells. FEBS Lett 458:171–174
30. MacKenzie A, Quinn J (1999) A serotonin transporter gene intron 2 polymorphic region, correlated with affective disorders, has allele-dependent differential enhancer-like properties in the mouse embryo. Proc Natl Acad Sci USA 96:15251–15255
31. Sen S, Burmeister M, Ghosh D (2004) Meta-analysis of the association between a serotonin transporter promoter polymorphism (5-HTTLPR) and anxiety-related personality traits. Am J Med Genet B Neuropsychiatr Genet 127:85–89
32. Barr CS, Newman TK, Lindell S, Shannon C, Champoux M, Lesch KP, Suomi SJ, Goldman D, Higley JD (2004) Interaction between serotonin transporter gene variation and rearing condition in alcohol preference and consumption in female primates. Arch Gen Psychiatry 61:1146–1152
33. Stein DJ (2002) Obsessive-compulsive disorder. Lancet 360:397–405
34. Hollander E, Bienstock CA, Koran LM, Pallanti S, Marazziti D, Rasmussen SA, Ravizza L, Benkelfat C, Saxena S, Greenberg BD, Sasson Y, Zohar J (2002) Refractory obsessive-compulsive disorder: state-of-the-art treatment. J Clin Psychiatry Suppl 6 63:20–29
35. Jenike MA (2004) Clinical practice: obsessive-compulsive disorder. N Engl J Med 350:259–265
36. Pogarell O, Hamann C, Popperl G, Juckel G, Chouker M, Zaudig M, Riedel M, Moller HJ, Hegerl U, Tatsch K (2003) Elevated brain serotonin transporter availability in patients with obsessive-compulsive disorder. Biol Psychiatry 54:1406–1413
37. Goldman D (1996) High anxiety. Science 274:1483
38. Seretti A, Cusin C, Lattuada E, Di Bella D, Catalano M, Smeraldi E (1999) Serotonin transporter gene (5-HTTLPR) is not associated with depressive symptomatology in mood disorders. Mol Psychiatry 4:280–283
39. Rauch SL, Jenike MA, Alpert NM, Baer L, Breiter HC, Savage CR, Fischman AJ (1994) Regional cerebral blood flow measured during symptom provocation in obsessive-compulsive disorder using oxygen 15-labeled carbon dioxide and positron emission tomography. Arch Gen Psychiatry 51:62–70
40. Breiter HC, Rauch SL, Kwong KK, Baker JR, Weisskoff RM, Kennedy DN, Kendrick AD, Davis TL, Jiang A, Cohen MS, Stern CE, Belliveau JW, Baer L, O'Sullivan RL, Savage CR, Jenike MA, Rosen BR (1996) Functional magnetic resonance imaging of symptom provocation in obsessive-compulsive disorder. Arch Gen Psychiatry 53:595–606
41. Adler CM, McDonough-Ryan P, Sax KW, Holland SK, Arndt S, Strakowski SM (2000) fMRI of neuronal activation with symptom provocation in unmedicated patients with obsessive compulsive disorder. J Psychiatr Res 34:317–324
42. Swedo SE, Pietrini P, Leonard HL, Schapiro MB, Rettew DC, Goldberger EL, Rapoport SI, Rapoport JL, Grady CL (1992) Cerebral glucose metabolism in childhood-onset obsessive-compulsive disorder: revisualization during pharmacotherapy. Arch Gen Psychiatry 49:690–694
43. Carey PD, Warwick J, Niehaus DJ, van der Linden G, van Heerden BB, Harvey BH, Seedat S, Stein DJ (2004) Single photon emission computed tomography (SPECT) of anxiety disorders before and after treatment with citalopram. BMC Psychiatry 4:30
44. Rosenberg DR, Keshavan MS (1998) A.E. Bennett Research Award. Toward a neurodevelopmental model of obsessive-compulsive disorder. Biol Psychiatry 43:623–640
45. Szeszko PR, MacMillan S, McMeniman M, Lorch E, Madden R, Ivey J, Banerjee SP, Moore GJ, Rosenberg DR (2004) Amygdala volume reductions in pediatric patients with obsessive-compulsive disorder treated with paroxetine: preliminary findings. Neuropsychopharmacology 29:826–832
46. Szeszko PR, Ardekani BA, Ashtari M, Malhotra AK, Robinson DG, Bilder RM, Lim KO (2005) White matter abnormalities in obsessive-compulsive disorder: a diffusion tensor imaging study. Arch Gen Psychiatry 62:782–790