Molecular Psychiatry (2006), 1–11

www.nature.com/mp

ORIGINAL ARTICLE

Tryptophan hydroxylase 2 (*TPH2*) haplotypes predict levels of *TPH2* mRNA expression in human pons

J-E Lim[1,2], J Pinsonneault[1,3], W Sadee[1,2,3,4] and D Saffen[1,2,3,4]

[1]*Department of Pharmacology, The Ohio State University, Columbus, OH, USA;* [2]*Neuroscience Graduate Studies Program, College of Medicine and Public Health, The Ohio State University, Columbus, OH, USA;* [3]*The OSU Program in Pharmacogenomics, The Ohio State University, Columbus, OH, USA and* [4]*Department of Psychiatry, The Ohio State University, Columbus, OH, USA*

Tryptophan hydroxylase isoform 2 (TPH2) is expressed in serotonergic neurons in the raphe nuclei, where it catalyzes the rate-limiting step in the synthesis of the neurotransmitter serotonin. In search for functional polymorphisms within the *TPH2* gene locus, we measured allele-specific expression of *TPH2* mRNA in sections of human pons containing the dorsal and median raphe nuclei. Differences in allelic mRNA expression – referred to as allelic expression imbalance (AEI) – are a measure of *cis*-acting regulation of gene expression and mRNA processing. Two marker SNPs, located in exons 7 and 9 of *TPH2* (rs7305115 and rs4290270, respectively), served for quantitative allelic mRNA measurements in pons RNA samples from 27 individuals heterozygous for one or both SNPs. Significant AEI (ranging from 1.2- to 2.5-fold) was detected in 19 out of the 27 samples, implying the presence of *cis*-acting polymorphisms that differentially affect *TPH2* mRNA levels in pons. For individuals heterozygous for both marker SNPs, the results correlated well ($r = 0.93$), validating the AEI analysis. AEI is tightly associated with the exon 7 marker SNP, in 17 of 18 subjects. Remarkably, expression from the minor allele exceeded that of the major allele in each case, possibly representing a gain-of-function. Genotyping of 20 additional *TPH2* SNPs identified a haplotype block of five tightly linked SNPs for which heterozygosity is highly correlated with AEI and overall expression of *TPH2* mRNA. These results reveal the presence of a functional *cis*-acting polymorphism, with high frequency in normal human subjects, resulting in increased *TPH2* expression levels. The SNPs that correlate with AEI are closely linked to *TPH2* SNPs previously shown to associate with major depression and suicide.
Molecular Psychiatry advance online publication, 12 December 2006; doi:10.1038/sj.mp.4001923

Keywords: allelic expression imbalance (AEI); tryptophan hydroxylase 2 (*TPH2*); single nucleotide polymorphism (SNP); haplotype; linkage disequilibrium (LD); serotonin

Introduction

Tryptophan hydroxylase (TPH) catalyzes the rate-limiting step in the synthesis of serotonin (5-hydroxytryptamine; 5-HT),[1] a neurotransmitter that plays an important role in the regulation of mood.[2] Dysregulation of serotonergic activity has been associated with major depression, anxiety disorders and suicidal behavior.[3] Most antidepressant drugs, including the serotonin-selective reuptake inhibitors (SSRIs) and many tricyclic antidepressants (TCAs), increase levels of extracellular serotonin by inhibiting its reuptake or blocking its metabolism. Tryptophan hydroxylase 2 (*TPH2*) is a recently discovered isoform of TPH that is specifically expressed in the brain, with particularly high expression in the serotonergic neurons of the raphe nuclei.[4–6] The dorsal and median raphe nuclei are the major source of serotonin in the forebrain,[4] including areas implicated in mood and anxiety disorders.

As *TPH2* is strategically placed to regulate serotonin levels in the brain, there is currently great interest in identifying genetic variants that affect the level of *TPH2* enzymatic activity or control the levels of expression of the *TPH2* gene. Extensive DNA sequencing of the *TPH2* gene has revealed that polymorphisms that change the amino-acid sequence of the *TPH2* protein are rare.[7–9] The focus of research has therefore, now changed to identifying genetic variants that influence the *TPH2* gene expression.

Recently, measurement of mRNA allelic expression imbalance (AEI) has emerged as a powerful method for identifying genetic variants that influence the expression of mRNAs.[10,11] In this method, relative levels of mRNA expressed from each of two alleles are measured using RNA isolated from individuals who

Correspondence: Professor W Sadee or Professor D Saffen, Department of Pharmacology, The Ohio State University, 5072 Graves Hall, 333 West 10th Avenue, Columbus, OH 43210, USA. E-mails: sadee.1@osu.edu or saffen.1@osu.edu
Received 27 June 2006; revised 25 August 2006; accepted 25 September 2006

are heterozygous for a marker single nucleotide polymorphism (SNP) within the mRNA. Using this method, it is possible to reliably detect differences in expression levels between alleles as small as 20%. As comparisons between expression levels are made using single samples of RNA isolated from specific organs or tissues, variation between individuals that arise from differences in environmental factors, physiological states, or *trans*-acting factors are minimized: the mRNA from each allele acts as the control for the other. We have previously used this technique to quantify AEI of mRNAs encoding human H^+/dipeptide transporter 2 (PEPT2),[12] *p*-glycoprotein (MDR1),[13] the μ-opiate receptor (OPRM),[14] and the serotonin transporter (SERT).[15]

The goal of this study was to determine whether allele-specific mRNA expression of *TPH2* gene occurs and, if so, identify *cis*-acting genetic elements that predict high or low levels of expression.

Materials and methods

Materials

Frozen sections of rostral pons containing the dorsal and median raphe nuclei from 48 individuals were purchased from the Brain and Tissue Bank for Developmental Disorders (University of Maryland, Baltimore, MD, USA). The demographics of this collection have been described previously.[15] Oligonucleotide primers were designed using the program Oligo 4.0 (National Biosciences Inc., Plymouth, MN, USA) and synthesized by Integrated DNA Technologies (Coralville, IA, USA).

Isolation of DNA and RNA from human pons

Isolation of DNA and RNA from the tissue samples in our collection has been described previously.[15] Briefly, frozen sections of pons were incubated in 10 volumes of RNAlater-ICE Frozen Tissue Transition solution (Ambion Inc., Austin, TX, USA) overnight at -80°C to maximize recovery of DNA and RNA. The next day, a small piece of tissue from the ventral edge of each sample was removed and homogenized in DNA lysis buffer for isolation of genomic DNA and the remaining portion of the sample homogenized in Trizol reagent (Invitrogen, Carlsbad, CA, USA) for isolation of total RNA.

Genotyping

Genotyping of TPH2 SNPs using SNaPshot primer extension assays was carried out as described previously.[15] Briefly, short (100–300 bp) segments of genomic DNA were Polymerase Chain Reaction (PCR)-amplified using pairs of synthetic oligonucleotide primers that flank each SNP. Following amplification, the unincorporated dNTPs were inactivated with antarctic alkaline phosphatase (New England Biolabs, Ipswich, MA, USA) and excess primers degraded with exonuclease I (New England Biolabs). The PCR products were used as templates in SNaPshot primer extension assays (Applied Biosystems, Foster City, CA, USA), using extension primers designed to anneal to the amplified DNA immediately adjacent to the SNP site. The resulting fluorescently labeled primers were analyzed by capillary electrophoresis using an ABI3730 DNA analysis system and Gene Mapper 3.0 software (Applied Biosystems Inc.). The TPH2 SNPs we examined are listed in Table 1. The locations of these SNPs within the TPH2 gene are

Table 1 TPH2 SNPs examined in this study

SNP no.	*DbSNP no.*	*Location on chromosome 12*	*Location within TPH2 gene*	*Allele frequencies*	*Hetero-zygosity*
01	rs4570625	70618190	Upstream	G/T = 0.72/0.28	0.403
02	rs10748185	70622122	Intron 2	A/G = 0.51/0.49	0.500
03	rs2129575	70626340	Intron 4	G/T = 0.74/0.26	0.385
04	rs1386488	70630885	Intron 5	A/C = 0.85/0.15	0.255
05	rs1843809	70634965	Intron 5	T/G = 0.83/0.17	0.282
06	rs1386495	70638589	Intron 5	T/C = 0.83/0.17	0.282
07	rs1386494	70638810	Intron 5	G/A = 0.88/0.12	0.211
08	rs6582072	70640744	Intron 5	G/A = 0.83/0.17	0.282
09	rs2171363	70646531	Intron 5	C/T = 0.53/0.47	0.498
10	rs4760815	70658496	Intron 6	T/A = 0.65/0.35	0.455
11	rs7305115	70659129	Exon 7	G/A = 0.65/0.35	0.455
12	rs6582078	70661158	Intron 7	T/G = 0.60/0.40	0.480
13	rs1023990	70668514	Intron 7	T/C = 0.79/0.21	0.332
14	rs1007023	70674641	Intron 8	T/G = 0.89/0.11	0.196
15	rs1352251	70684161	Intron 8	T/C = 0.59/0.41	0.484
16	rs1473473	70690645	Intron 8	G/A = 0.88/0.12	0.211
17	rs9325202	70693744	Intron 8	G/A = 0.65/0.35	0.455
18	rs1487275	70696559	Intron 8	T/G = 0.78/0.22	0.343
19	rs1386486	70698487	Intron 8	C/T = 0.61/0.39	0.476
20	rs4290270	70702502	Exon 9	A/T = 0.63/0.37	0.466
21	rs1872824	70716581	Intron 9	C/T = 0.64/0.36	0.461
22	rs1352252	70738308	Downstream	A/G = 0.56/0.44	0.493

shown in Figure S1 in Supplemental materials (available online). Sequences of the PCR amplification and primer extension primers and reaction conditions for each primer set used for genotyping are available upon request.

LD and haplotype analysis

D′ values for each pair of SNPs and estimated haplotype frequencies were calculated using Haploview (version 3.3; http://www.broad.mit.edu/mpg/haploview/),[16] Predicted diplotypes for each individual in our collection were calculated from the genotyping data using HelixTree^{RT} (GoldenHelix, Inc., Bozeman, MT, USA).

AEI measurements

Measurements of allele-specific mRNA expression were carried out as described previously.[15] Briefly, RNA from each sample was treated with RNase-Free DNase Set (Qiagen, Valencia, CA, USA) for 15 min and re-isolated using QIAGEN RNeasy columns. Complementary DNA (cDNA) was generated from 1 μg RNA in 20 μl reaction mixes containing 1 μl (200 U) Superscript II reverse transcriptase (Invitrogen, Carlsbad, CA, USA), 1 μl of 50 μM oligo(dT)$_{20}$ primer (Invitrogen), 1 μl of 10 mM dNTP mix (Invitrogen), 0.5 μl of 1 μM *TPH2* gene-specific primer (5′-TTAATTCTCCAATGGAGGAAAGGA-3′), 4 μl of 5 × first-strand buffer (Invitrogen), 1 μl of RNaseOUT (40 U/μl), and RNase-free water. A cDNA segment containing marker SNPs rs7305115 and rs4290270 was amplified using *Taq* DNA polymerase (Promega, Madison, WI, USA), the forward primer 5′-ACGA GACTTTCTGGCAGGACTG-3′, and the reverse primer 5′-TTAATTCTCCAATGGAGG-AAAGGA-3′ with the following cycles: (1 × (5 min at 95°C); 35 × (30 s at 95°C, 30 s at 60°C, 1 min at 72°C) 1 × (7 min at 72°C)). Following amplification, the unincorporated dNTPs were inactivated with antarctic alkaline phosphatase (New England Biolabs) and excess primers degraded with exonuclease I (New England Biolabs). SNaPshot Primer extension assays were carried out using the extension primer 5′-GATCCCCTCTACACCCC-3′ for rs7305115 and 5′-AAAGGAGTCCTGCTCCATA-3′ for rs4290270 with the following cycles: (25 × (10 s at 96°C, 5 s at 50°C, 30 s at 72°C)). Unincorporated fluorescent dNTP analogs were removed by incubation with 1.0 U of intestinal calf phosphatase (10 000 U/ml; New England Biolab) for 3 h at 37°C. The primer extension products were resolved by capillary electrophoresis using an Applied Biosystems 3730 DNA Analyzer and quantified using the Gene Mapper 3.0 software (Applied Biosystems).

Addition of different fluorescently labeled dideoxynucleotides onto the 3′-end of the primers produces oligonucleotides with slightly different electrophoretic mobilities and distinct fluorescence spectra. As different fluorophores differentially affect the efficiency of nucleotide incorporation and have different fluorescence yields, peak area ratios of genomic DNA diverge from the theoretical ratio of 1.0. The measured ratios for genomic DNA were therefore normalized to 1.0 by multiplying each measured ratio by the inverse of the mean of the genomic DNA ratios (correction factor = 1/(mean of measured genomic DNA ratios)). Two tissue samples (no. 1230 and no. 1609) yielded allelic DNA ratios significantly different from the mean (>4 s.d., indicating the presence of a gene dosage effect), and were excluded from the calculated mean DNA ratios. RNA (i.e., cDNA) ratios from heterozygous samples were multiplied by the same correction factor. SNaPshot assays were performed 3 × with genomic DNA and 3 × with three independent cDNA preparations per sample.

Real-time PCR

TPH2 and glyceraldehydes-3-phosphate dehydrogenase *(GAPDH*) mRNA levels were measured by real-time PCR using an ABI 7000 DNA sequence detection system (Applied Biosystems, Foster City, CA, USA) as previously described.[15] Briefly, *TPH2* or *GAPDH* cDNA was synthesized from 1 μg total pons RNA using reverse-transcriptase and the primers: 5′-TTAATTCTCCAATGGAGGAAAGGA-3′ (*TPH2*) or 5′-GTGTGGTGGGGGACTGAGTGTG-3′ (*GAPDH*). Segments of *TPH2* or *GAPDH* cDNAs were amplified using *TPH2*- or *GAPDH*-specific primer sets and heat-activated *Taq* DNA polymerase in reaction mixes containing dNTPs, buffer, SYBR-Green and a reference dye (Applied Biosystems, Foster City, CA, USA). The *TPH2* amplification primers were: 5′-ACGA GACTTTCTGGCAGGACTG-3′ (forward) and 5′-TT AATTCTCCAATGGAGGAAAGGA-3′ (reverse) and the *GAPDH* amplification primers were: 5′-CAGCAA GAGCACAAGAGGAAGAGAGA-3′ (forward) and 5′-GTGTGGTGGGGGACT-GAGTGTG-3′ (reverse). Amplification conditions consisted of a 10-min preincubation at 95°C to activate the *Taq* DNA polymerase, followed by 40 cycles of denaturation at 95°C for 15 s and primer annealing and extension for 1 min at 60°C. PCR product melting curves were examined to confirm the homogeneity of PCR products. *TPH2* mRNA measurements were expressed as cycle thresholds (C_T) and normalized by subtracting C_T values obtained with *GAPDH* mRNA.

Statistics

Differences between corrected genomic and mRNA (cDNA) ratios were tested for statistical significance using the General Linear Model (GLM) procedure in SAS (SAS Institute Inc., Cary, NC, USA). Agreement between AEI measurements using the marker SNP rs7305115 or rs4290270 was assessed by calculating the Pearson correlation coefficient for mean AEI values for individuals heterozygous for both SNPs (n = 13). Correlations between heterozygosity of *TPH2* SNPs and AEI of THP2 mRNA were examined by calculating Kappa-coefficients using SPSS (SPSS Inc., Chicago, IL, USA). Agreement was defined to be either heterozygous and *TPH2* AEI >1.2, or homozygous with *TPH2* AEI <1.2. Exact two-sided

P-values for the significance of the kappa estimate were computed.

Results

To identify samples suitable for *TPH2* mRNA AEI measurements, we genotyped chromosomal DNA from each of our samples for two marker SNPs: rs7305115 (exon 7) and rs4290270 (exon 9). (See Figure S1 in Supplemental materials for the locations of these and additional *TPH2* SNPs.) Among the 48 individuals in our collection, 18 were heterozygous for rs7305115 (G/A) and 22 heterozygous for rs4290270 (A/T). Five individuals were heterozygous only for rs7305115 (no. 1027, 1230, 1540, 1551, 1609), nine individuals were heterozygous only for rs4290270 (no. 1054, 1104, 1169, 1430, 1442, 1486, 1546, 1613, 1614), and 13 individuals were heterozygous for both SNPs (no. 813, 879, 914, 917, 1078, 1101, 1103, 1105, 1112, 1135, 1279, 1489, and 1607). Alleles of both marker SNPs were in Hardy–Weinberg equilibrium within the complete collection of 48 individuals (not shown).

Figure 1 shows the results of mRNA AEI measurements for the 18 individuals heterozygous for rs7305115. Seventeen of the samples (94%) showed higher expression of mRNA for the *A*-allele compared to the *G*-allele, with ratios ranging from 1.2 to 2.5 (Table 2). The *G*-allele represents the reference sample (wild-type), while the *A*-allele is a minor, albeit frequent, variant. Sample 1540 showed no significant AEI. All but two of the samples yielded allelic ratios for genomic DNA close to the expected value of 1.0. Two samples (no. 1230 and no. 1609) consistently yielded ratios significantly below 1.0. These low ratios suggest a possible duplication in the TPH2 locus containing the *G*-allele.

Figure 2 shows the results of AEI assays for the 22 individuals heterozygous for rs4290270. There was significant AEI in 13 RNA samples, with higher expression of the *T*-allele (again the frequent minor variant). Ratios ranged from 1.2 to 2.5 (Table 2). Thirteen of the 22 samples were heterozygous for both marker SNPs, affording the opportunity to validate the results obtained with the marker SNP rs7305115. Figure 3 shows that there is an excellent correlation between AEI measurements made using the two marker SNPs.

The results in Figures 1–3 show that heterozygosity of rs7305115 is highly correlated with *TPH2* mRNA AEI (17/18 = 94%), while heterozygosity of rs4290270 is less highly correlated (13/22 = 59%). These results raise the possibility that rs7305115 is tightly linked to the 'functional' polymorphism that controls levels of *TPH2* mRNA expression, or is itself a functional polymorphism.

To determine whether additional SNPs correlate with *TPH2* mRNA AEI, we genotyped 20 additional common *TPH2* SNPs. (See Table 1 and Figure S1 in Supplemental materials for allele frequencies and locations of these SNPs.) Alleles of each of the SNPs were in Hardy–Weinberg equilibrium in our population (not shown). Figure S2A shows a linkage disequilibrium (LD) (D') plot for each pair of SNPs, which was constructed from the genotyping data for the 36 Caucasians in our sample. These data show that *TPH2* comprises four haplotype blocks: the first contains SNPs rs4570625 to rs2129575, the second rs1386488 to rs1352251, the third rs1473473 to rs9325202, and the fourth rs1487275 to rs1352252. These results are in close agreement with the haplotype structure determined from previous studies of Caucasian subjects: the HapMap CEU collection (http://www.hapmap.org/; Figure S1) and US and

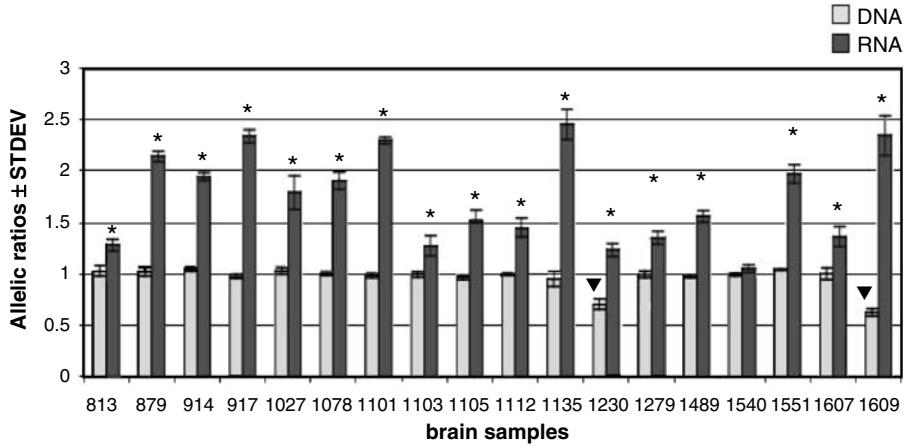


Figure 1 Comparison of genomic DNA and mRNA (cDNA) ratios assayed using the marker SNP rs7305115. Data are expressed as ratios of *A:G* alleles, corrected as described in Materials and methods. The lightly shaded bars represent the average of three DNA ratio measurements using three independent preparations of pons genomic DNA. The darkly shaded bars represent the average of three mRNA ratio measurements using three independent cDNA preparations from a single preparation of pons total RNA. The error bars indicate (±) s.d. for each set of measurements. Samples where the mRNA ratios are statistically different from 1.0 ($P < 0.001$) using the GLM procedure in SAS are marked with an asterisk (*). Two genomic DNA samples (no. 1230 and no. 1609) that yielded AEI ratio significantly <1.0 are marked with arrowheads.

Table 2 Measurements of AEI using the marker SNPs rs7305115 and rs4290270

Sample no.	*rs7305115 (A/G)*			*rs4290270 (T/A)*		
	DNA	*RNA*		*DNA*	*RNA*	
	Mean±*s.d.*	*Mean*±*s.d.*	P-*value*	*Mean*±*s.d.*	*Mean*±*s.d.*	P-*value*
813	1.03±0.05	1.28±0.06	<0.0001*	1.04±0.06	0.99±0.14	0.8467
879	1.02±0.04	2.14±0.05	<0.0001*	1.13±0.07	1.91±0.07	<0.0001*
914	1.04±0.02	1.95±0.04	<0.0001*	1.01±0.03	1.94±0.06	<0.0001*
917	0.98±0.02	2.34±0.06	<0.0001*	1.02±0.05	2.55±0.38	<0.0001*
1027	1.03±0.03	1.79±0.16	<0.0001*	—	—	—
1054	—	—	—	0.88±0.06	0.97±0.06	0.6504
1078	1.00±.001	1.90±0.08	<0.0001*	0.96±0.06	1.76±0.05	<0.0001*
1101	0.99±0.02	2.30±0.03	<0.0001*	0.97±0.02	2.41±0.032	<0.0001*
1103	1.00±0.03	1.28±0.100	<0.0001*	0.97±0.02	1.22±0.112	0.0029*
1104	—	—	—	0.98±0.02	1.09±0.04	0.212
1105	0.97±0.02	1.52±0.02	<0.0001*	0.98±0.04	1.25±0.05	0.0009*
1112	1.00±0.01	1.50±0.08	<0.0001*	0.97±0.03	1.22±0.02	0.0035*
1135	0.95±0.08	2.46±0.15	<0.0001*	1.04±0.08	2.52±0.16	<0.0001*
1169	—	—	—	1.02±0.02	1.14±0.05	0.0442
1230	0.71±0.05	1.23±0.06	0.0006*			
1279	1.00±0.03	1.35±0.06	<0.0001*	0.95±0.02	0.94±0.03	0.394
1430	—	—	—	1.06±0.02	1.25±0.12	<0.0001*
1442	—	—	—	1.08±0.19	0.92±0.11	0.251
1486	—	—	—	0.97±0.06	0.95±0.03	0.5221
1489	0.98±0.01	1.57±0.06	<0.0001*	0.99±0.07	1.62±0.07	<0.0001*
1540	1.00±0.02	1.05±0.04	0.0943	—	—	—
1546	—	—	—	0.98±0.03	1.1±0.08	0.1467
1551	1.04±0.00	1.97±0.09	<0.0001*	—	—	—
1607	1.00±0.06	1.37±0.10	<0.0001*	1.11±0.04	1.72±0.23	<0.0001*
1609	0.63±0.04	2.35±0.20	<0.0001*			
1613	—	—	—	0.89±0.06	1.06±0.06	0.3639
1614	—	—	—	1.0±0.07	1.29±0.07	0.0002*

* = statistically significant.
Bold characters indicate samples showing statistically significant AEI for the marker SNPs rs7305115 and/or rs4290270.

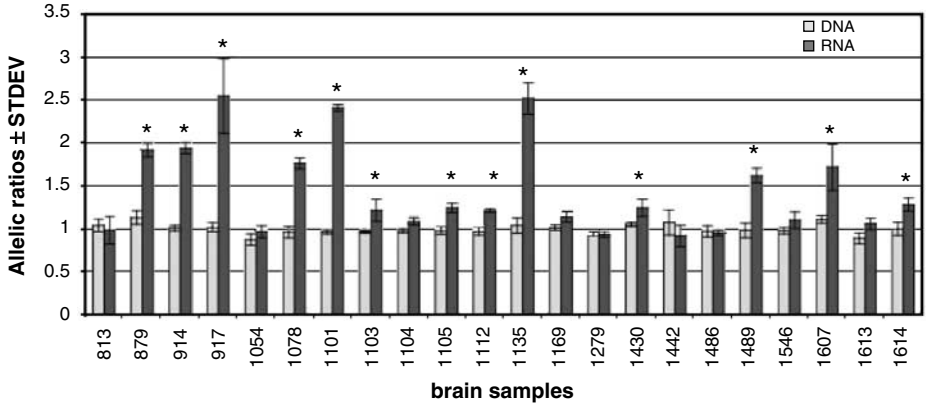


Figure 2 Comparison of corrected genomic DNA and mRNA (cDNA) ratios assayed using the marker SNP rs4290270. Data are expressed as ratios of *T:A* alleles, as described in Materials and methods. The lightly shaded bars represent the average of three DNA ratio measurements using three independent preparation of pons genomic DNA. The darkly shaded bars represent the average of three mRNA ratio measurements using three independent cDNA preparations from a single preparation of pons total RNA. The error bars indicate (±) s.d. for each set of measurements Samples where the mRNA ratios are statistically different from 1.0 ($P<0.001$) using the GLM procedure in SAS are marked with an asterisk (*).

Finnish populations.[17] The frequencies of haplotypes within each block are listed in Figure S2B, and the predicted diplotypes for each individual in our collection are listed in Table S1 in Supplemental materials.

The possible contribution of each SNP to *TPH2* mRNA AEI was evaluated by looking for correlations between heterozygosity/homozygosity of the SNP and the presence/absence of AEI for *TPH2* mRNA within the 27 samples where AEI measurements were made. A tabulation of these results is shown in Table S2 in Supplemental materials. The strength of each correlation was assessed using the Kappa-statistic.[18] As shown in Figure 4, five closely linked SNPs, rs2171363 (*C/T*), rs4760815 (*T/A*), rs7305115 (*G/A*), rs6582078 (*T/G*), and rs9325202 (*G/A*), showed statistically significant correlations with *TPH2* mRNA AEI (Kappa-coefficients >0.66). Heterozygosity of rs1352251 (*T/C*) also correlated with *TPH2* mRNA AEI (Kappa-coefficient = 0.534). An independent test using a decision-tree-based algorithm (Helix-Tree^{RT}) found significant statistically correlations between SNP heterozygosity and AEI ($P<0.01$) for rs2171363, rs4760815, rs7305115, rs6582078 and rs9325202 (data not shown).

As mentioned above, AEI measurements revealed that *TPH2* mRNA containing the rs7305115 *A*-allele is expressed at higher levels than mRNA containing the *G*-allele. Among 18 samples showing AEI for *TPH2* mRNA, 17 were heterozygous for rs7305115 (Table S1). Fifteen of the 18 samples were heterozygous for the exact complementary (i.e., 'yin' and 'yang') haplotypes CTGTG and TAAGA, comprising the SNPs rs2171363, rs4760815, rs7305115, rs6582078, and rs9325202, respectively. Table 3 lists the frequencies for haplotypes containing the rs7305115 *G*-allele or *A*-allele within the Caucasian subset of our sample. These data show that *G*-allele haplotypes, which are associated with low *TPH2* mRNA expression, are more common (0.6) than *A*-allele haplotypes (0.4),

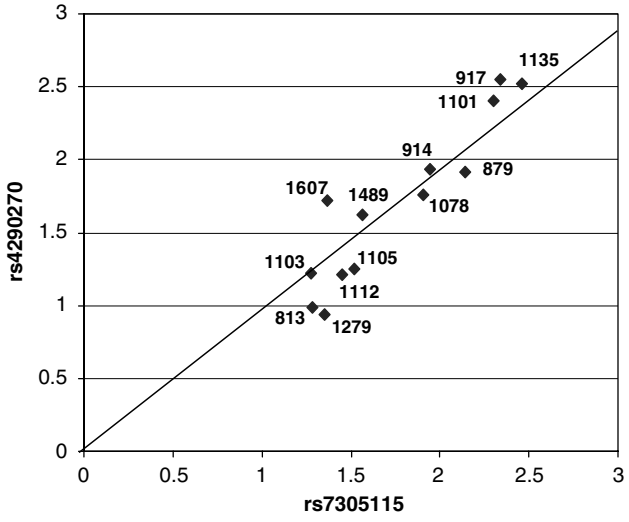


Figure 3 Comparison of mRNA allelic expression ratios determined using the marker SNPs rs7305115 and rs4290270. The solid line represents the best fit for the data determined by linear regression, with the added requirement that the line pass through the origin, 0.0 ($R=0.93$; $r^2=0.86$).

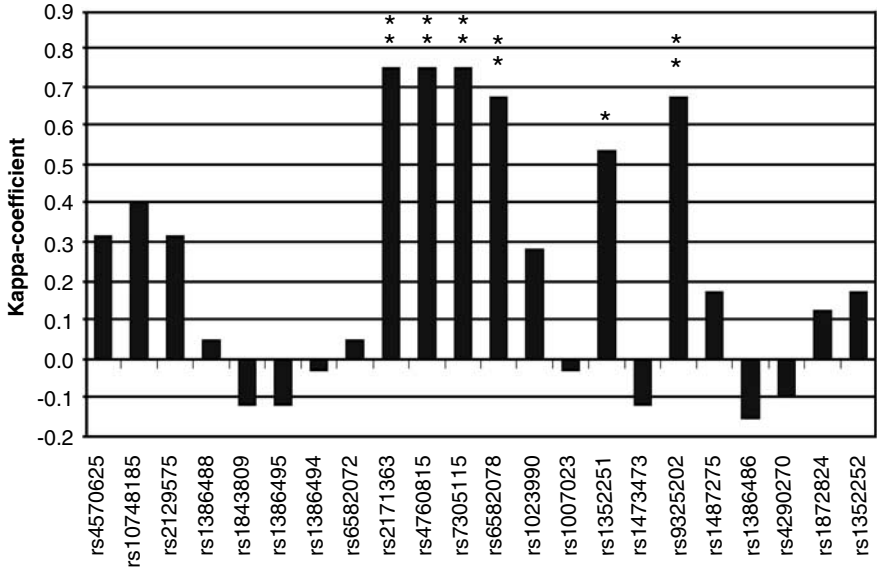


Figure 4 Correlations between heterozygosity of individual TPH2 SNPs and AEI of TPH2 mRNA. *Y*-axis: Kappa-coefficients were calculated from the data in Table S4 using SPSS. The values of Kappa-coefficients range from 1.0 for perfect correlation between heterozygosity and AEI (i.e., all samples heterozygous for the SNP show AEI and all homozygous samples show no AEI) and −1.0 for perfect anti-correlation (i.e., no samples heterozygous for the SNP show AEI and all homozygous samples show AEI). A SNP showing random correlations with AEI (i.e., 50% of heterozygous and homozygous samples show AEI) would have a Kappa value of 0.0 ((**): $P<0.001$; (*): $P=0.003$).

Table 3 Haplotype frequencies for 38 Caucasians in sample (76 chromosomes)

rs7305115 G-allele haplotypes	*rs7305115 A-allele haplotypes*	*Frequency*
C T **G** T G		0.553 (42/76)
	T A **A** G A	0.316 (24/76)
	T A **A** G G	0.053 (4/76)
	T T **A** G G	0.026 (2/76)
T T **G** T A		0.026 (2/76)
C T **G** T A		0.013 (1/76)
C T **G** G G		0.013 (1/76)
Total *G*-allele haplotypes		0.605 (46/76)
	Total *A*-allele haplotypes	0.395 (30/76)

Listed haplotypes comprise the following SNPs: rs2171363 (C/T), rs4760815 (T/A), rs7305115 (**G/A**), rs6582078 (T/G) and rs9325202 (G/A).
* = statistically significant.
Bold characters denote alleles of rs7305115.

which are associated with high TPH2 mRNA expression. The population frequencies of the rs7305115 *G*- and *A*-alleles are similar to those previously reported for Caucasian populations ((http://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?rs = 7305115) and Zhm *et al*.[17]).

To test our ability to predict levels of *TPH2* mRNA expression based upon genotype, we compared levels of *TPH2* mRNA in pons samples from individuals who are heterozygous (*G/A*) or homozygous (*G/G* or *A/A*) for rs7305115 alleles. Real-time reverse transcriptase (RT)-PCR measurements of *TPH2* mRNA were carried out using RNA isolated from 18 (*G/G*), 21 (*G/A*) and nine (*A/A*) samples. *TPH2* mRNA measurements (expressed as C_T) were normalized by subtracting C_T values for GAPDH mRNA, which is ubiquitously expressed. Pairwise comparisons between groups showed that the *A/A* sample contained statistically higher levels of *TPH2* mRNA compared to the G/G sample ($P = 0.024$) or the G/A sample ($P = 0.04$). There was no statistical difference in levels of *TPH2* mRNA expression between the *G/G* and *G/A* samples ($P = 0.659$). Figure 5 shows the distribution of *TPH2* mRNA measurements for combined *G/G* and *G/A* samples compared to *A/A* samples. Although the spread of the data is large for both sets of samples, the *A/A* samples contain statistically significant higher levels of *TPH2* mRNA compared to the combined *G/G* and *G/A* samples ($P = 0.0075$). *C* for GAPDH varied from 15 to 18.4.

To address the question whether mRNA levels in the pons tissue sections reflect specific expression in serotonergic neurons, rather than nonspecific background expression, we compared *TPH2* mRNA levels in pons with levels in cerebellum and cortex and lymphoblasts. Again, GAPDH mRNA was used as a reference. As shown in Figure 6, *TPH2* mRNA levels were significantly higher in pons compared to cerebellum, occipital, frontal, parietal or temporal cortex and much higher than levels in lymphoblasts (ANOVA; $P < 0.0001$).

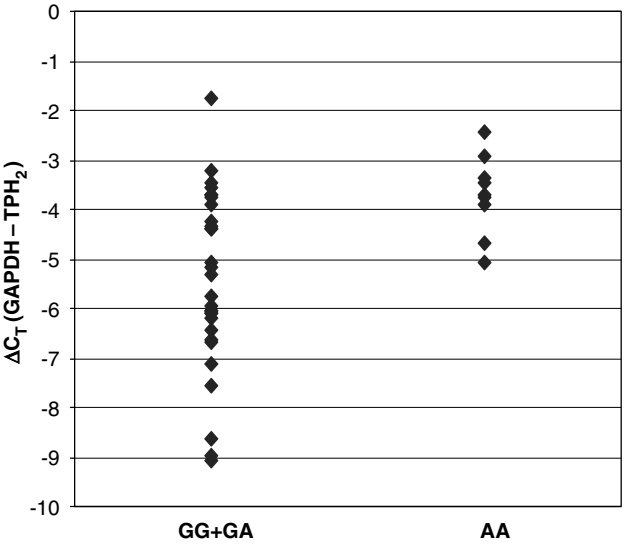


Figure 5 TPH2 mRNA levels in pons measured using real-time PCR. The *Y*-axis plots the difference between C_T determined for *GAPDH* and *TPH2* mRNAs. Individuals where grouped according to their genotype for the marker SNP rs7305115: (*G/G* or *G/A*) (left) or (*A/A*) (right). Statistical significance was evaluated by the two-tailed *t*-test ($P = 0.0075$).

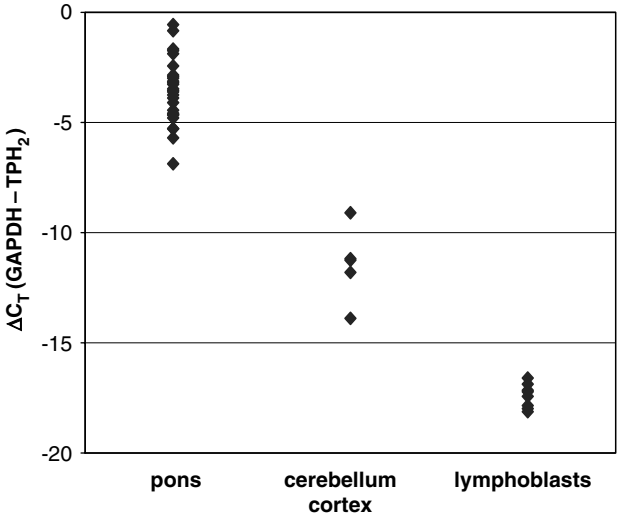


Figure 6 Comparison of *TPH2* mRNA expression levels in different tissues. The *Y*-axis plots the difference between C_T for *GAPDH* and *TPH2* mRNAs. Results obtained from 27 pons samples, five non pons brain regions (cerebellum and occipital, frontal, parietal and temporal cortexes) and eight lymphoblast cell lines are shown. The pons sample set comprised individuals homozygous (*A/A* or *G/G*) for rs7305115 alleles (One-way ANOVA; $P < 0.0001$).

Discussion

This study is the first to reveal the presence of a frequent, functional, *cis*-acting polymorphism in the *TPH2* gene that significantly affects mRNA expression. To detect allelic differences in *TPH2* mRNA expression, we developed and validated an accurate assay of AEI applicable to human autopsy brain tissues. Importantly, the functional analysis was performed in human pons, the physiologically relevant target tissue. Allelic differences in *TPH2* mRNA levels likely reflect expression in serotonergic neurons in the dorsal and median raphe nuclei, which are the primary source of serotonin in forebrain. Genotyping SNPs located within the *TPH2* gene identified individual SNPs and haplotypes that predict high or low levels of *TPH2* mRNA expression in human pons (Figure 4). Specifically, low levels of *TPH2* mRNA expression are associated with the CTGTG combination of alleles and high levels of expression with the TAAGA combination of alleles for the SNPs rs2171363, rs4760815, rs7305115, rs6582078 and rs9325202.

As these SNPs are tightly linked (Figure S1), it is not evident which SNP or SNP-combination is the 'functional' element that controls *TPH2* mRNA levels. Four of these SNPs (rs2171363, rs4760815, rs6582078 and rs9325202) are located within introns and one (rs7305115) within a coding exon. Analysis of predicted changes in mRNA structure for each of these SNPs using Mfold[19] showed only small differences between alleles (A Johnson, data not shown). To investigate possible functional effects of the exonic SNP rs7305115, we exogenously expressed *TPH2* mRNA of both alleles in CHO cells using cDNA expression vectors. No prominent differences in allelic expression or mRNA degradation rates were detectable between exogenously expressed *TPH2* mRNAs containing the rs7305115 *A*- or *G*-allele (Figure S3). This result, however, does not address possible differences in mRNA processing and maturation occurring at the level of hnRNA, since introns were absent from the cDNA constructs.

Analysis of possible effects of *TPH2* SNPs on mRNA transcription and processing using the webtool PupaSNP (http://pupasnp.bioinfo.ochoa.fib.es; Conde *et al*.[20]), showed that the *A*-allele of rs7305115 (the minor allele) generates a consensus binding site for the serine-arginine (SR)-proteins SR35 and SRP40, splicing factors that bind exonic splicing enhancers (ESEs).[21] Exons containing a nonfunctional or partially functional ESE are often skipped during RNA splicing,[21] possibly accounting for the lower yield of mRNA from the *G*-allele, which appears to be the main ancestral allele (see below). Skipping of exon 7 of the *TPH2* gene would result in a modified mRNA that encodes a truncated form of *TPH2* due to the insertion of an in-frame stop codon (data not shown). Recent studies have shown that mRNAs containing a premature translation termination signal often undergo preferential degradation via a poorly understood mechanism termed nonsense-mediated mRNA decay.[21] Thus, the *G*-allele of rs7305115 might be expected to produce lower levels of full length *TPH2* mRNA by increasing the frequency of exon skipping. This mechanism could account for the observed AEI of *TPH2* mRNA in *A/G* heterozygotes (Figures 1–3) and lower levels of *TPH2* mRNA expression in *G/A* heterozygotes and *G/G* homozygotes compared to *A/A* homozygotes (Figure 5).

To determine if aberrant *TPH2* mRNAs lacking exon 7 are expressed in pons, we carried out RT-PCR amplification of *TPH2* cDNA using sets of synthetic oligonucleotide primers that specifically amplify cDNA segments that contain or lack exon 7, respectively. These measurements produced two PCR products, with sizes corresponding to exon 7-containg and exon 7-deleted cDNAs, in each of the 48 samples in our collection. The predicted structures of both PCR products were confirmed by DNA sequencing. Real-time PCR measurements using primer sets specific for each mRNA showed that relative levels of the full-length and exon 7-deleted forms of *TPH2* mRNA varied widely between samples (data not shown). Exon 7-deleted mRNA appears to be present at very low levels, impeding a quantitative analysis. Nevertheless, these experiments provide evidence for aberrant splicing of the *TPH2* gene in the pons and suggest a possible mechanism by which the rs7305115 *A*-allele increases and the *G*-allele decreases levels of *TPH2* mRNA. Our results suggest that the *A*-allele may yield higher mRNA levels by enhancing the efficiency of proper mRNA splicing, representing a gain-of-function.

The rs7305115 *G*-allele appears to be the ancestral allele, since sequences from a rhesus monkey (http://www.hgsc.bcm.tmc.edu/projects/rmacaque/) and a chimpanzee (http://www.hgsc.bcm.tmc.edu/projects/chimpanzee/) have *G* at this position. The *G*-allele is also present in the mouse and rat. The high frequency of the *A*-allele in Caucasian populations (0.33–0.41) could have resulted from a population bottleneck or random genetic drift, or by positive selection. Since the *A*-allele is also present at high frequency (0.29–0.39) in African populations, it dates back to early human evolution. The high accumulation of a gain-of-function polymorphism is unusual and points towards positive selection, or balanced selection.[22] The existence of positive selection would indicate that *TPH2* variants significantly affect reproduction, possibly through a positive effect on mood or mental activity.

Even before the functional element(s) that control levels of *TPH2* mRNA expression are identified, knowledge of marker SNPs and haplotypes that strongly predict high or low levels of *TPH2* mRNA expression should be useful for association studies seeking to establish a role for *TPH2* in human disease. As *TPH2* encodes the enzyme that catalyzes the rate-limiting step in the synthesis of serotonin, it is plausible that differences in *TPH2* mRNA expression in the range of 1.2- to 2.5-fold could contribute to disorders in which serotonin plays a role. Moreover,

the high frequencies of the implicated SNPs and haplotypes suggest a possible role in brain disorders that affect a significant portion of the population, such as major depression, which has a life-time prevalence of about 16%.[23]

Since the discovery of the *TPH2* gene in 2003,[4] 14 published studies have examined possible associations between *TPH2* SNPs and various mental disorders including major depression,[17,24,25] bipolar disorder,[26] anxiety disorders,[27,28] attention-deficit/hyperactivity disorder (ADHD),[29,30] autism[31] or suicidal behavior.[17,32–36] The results of these studies have been mixed, with nine studies showing weak, but statistically significant associations between one or more *TPH2* SNP and a specific mental disorder,[17,24–27,29–32] and five showing no significant associations.[28,33–36] Most of the studies reporting negative results failed to detect statistically significant associations for SNPs in the putative promoter region or intronic SNPs in the 5′-end of the gene. These regions are also not associated with allelic mRNA expression observed in our study.

One of the nine positive studies reported an association between a rare loss-of-function mutation (G1463A; R441H) in *TPH2* exon 9 and treatment-resistant depression,[25] but this was not replicated in five subsequent studies.[8,37–40] Among the remaining eight studies reporting an association between *TPH2* SNP alleles and a mental disorder, four found statistically significant associations for SNPs within the region comprising introns 5–8.

Zill *et al*.[24] reported a statistically significant ($P = 0.012$, after Bonferroni correction) difference in allele frequencies for an intron 5 SNP (rs1386494) between Caucasian patients ($n = 300$) with major depression and ethnically matched controls ($n = 265$), with lower frequency of the *A*-allele in patients (*A*/*G* = 0.14/0.86) compared to controls (*A*/*G* = 0.21/0.79). (See Figure S1 in Supplementary materials for the locations of *TPH2* SNPs discussed in this section.) Statistically significant associations were also demonstrated between major depression and three haplotypes comprising alleles of 10 SNPs located within introns 5 and 6. In a second study,[32] the same group reported a statistically significant association between rs1386494 and completed suicides ($n = 263$) vs ethnically matched controls ($n = 266$), again finding higher levels of the *G*-allele in suicide victims (*A*/*G* = 0.14/0.86) and the *A*-allele in controls (*A*/*G* = 0.21/0.79). Four 10-SNP haplotypes (different from those identified in the depression study) showed correlations with suicide, but were not statistically significant after correction for multiple testing.

A study by Zhou *et al*.[17] examined associations between 15 *TPH2* SNPs and: (1) anxiety/depression, (2) suicide attempt, and (3) major depression in four populations. Weak associations between these disorders and individual SNPs located within the introns 5–8 segment of *TPH2* were observed. The SNPs showing associations, however, varied between disorders and between populations, and none remained significant after correction for multiple testing. Haplotype analysis revealed the presence of high-frequency 'yin' and 'yang' haplotypes, with complementary patterns of major and minor alleles. Again, weak associations (significant only in the absence of corrections for multiple testing) were observed, with a trend towards association of the yin-haplotype with anxiety/depression and suicide, and possible protection from these disorders by the yang-haplotype. The yin-haplotype was also associated with lower cerebral spinal fluid levels of the serotonin metabolite 5-hydroxyindolacetic acid in nonmedicated controls who were free of psychiatric disorders. Significantly, the yin-haplotype includes the *G*-allele of rs7305115, which we showed in this study to associate with low levels of *TPH2* mRNA expression. The yang-haplotype includes the *A*-allele of rs7305115.

A study by Harvey *et al*.[26] uncovered a weak association between bipolar disorder and haplotypes comprising alleles of SNPs located within the exons 7–9 segment of the TPH2 gene. Mossner *et al*.[27] described an association between obsessive-compulsive disorder (OCD) and the *G-C* haplotype for rs4570625 and rs4565946, SNPs located in the putative regulatory region and intron 2, respectively. Walitza *et al*.[29] described a weak association between regulatory region SNPs (rs4570625 and rs11178997) and attention-deficit hyperactivity disorder (ADHD). Sheehan *et al*.[30] detected a statistically significant association between the *T*-allele of rs1843809 (intron 5) and ADHD in transmission disequilibrium analysis of 179 Irish families. Finally, a recent study by Coon *et al*.[31] reported statistically significant associations between autism and *TPH*2 SNPs in introns 1 and 4 (rs4341581 and rs11179000).

Three recent studies have detected associations between a SNP in the putative TPH2 regulatory promoter region (rs4570625; G/T) and amygdala activity[41,42] or 'emotional processing'.[43] Brown *et al*.[41] used functional magnetic resonance imaging (fMRI) to detect greater bilateral dorsal amygdala reactivity to fearful stimuli in individuals carrying the T-allele (T/T or T/G) compared to G/G homozygotes. An independent fMRI study by Canli *et al*.[42] observed increased responses in both the right and left amydala of T-allele carriers viewing fearful, happy or sad faces compared to faces with neutral expression. Hermann *et al*.[43] detected a tendency in T-allele carriers towards increased event-related potentials (ERPs) in electroencephalograms recorded 240 ms after viewing pictures with high emotional content. An additive effect was detected in individuals carrying both the T-allele and the 'short' promoter allele[44,45] of the serotonin transporter (*SERT*) gene.

As described above, we observed a weak, but positive, correlation between rs4570625 heterozygosity and *TPH2* mRNA AEI in adult pons (Kappa-coefficient = 0.311; $P = 0.053$; Figure 4 and Table S2). These data suggest that rs4570625 does not control TPH2 mRNA expression, but may be in partial LD

with a functional polymorphism that does. In fact, our genotyping results (Figure S2) predict that rs4570625 is in partial LD with SNPs (rs2171363, rs4760815, rs7305115, and rs6582078) that highly correlate with TPH2 mRNA AEI. These observations suggest that reanalysis of the imaging and electroencephalography data in the above studies might show stronger correlations with rs7305115 compared to rs4570625. Alternatively, it is possible that rs4570625 (or a closely linked polymorphism in the promoter region) directly regulates TPH2 mRNA expression specifically during times of emotional stress and/or during brain development. As serotonin has been shown to play a role in the development of the brain,[46,47] it is possible that differential expression of *THP2* at specific stages of brain development may differentially influence the development of neuronal circuits that control amygdala activity in the adult. This interesting possibility remains to be examined.

Taken together, the studies described above provide preliminary evidence for a role for *TPH2* alleles in several mental disorders and processing of emotional stimuli. None of the studies, however, identified functional alleles, and thus do not provide mechanistic explanations for the observed associations. The fact that many of the above studies identified different associating SNPs suggests that the studies may lack sufficient power to reliably detect associations for SNPs that are in partial linkage with a functional polymorphism within the *TPH2* gene. We suspect that larger studies would show stronger associations for most SNPs in the region, with the strongest association observed for the functional polymorphism.

Future studies examining the potential associations between TPH2 and mental disorders should consider the following points: (1) The contribution of *TPH2* to a complex disease may be small, and therefore large numbers of individuals may need to be examined to observe contributions of specific alleles. (2) As previously shown for *SERT* promoter polymorphisms,[48] stronger associations may be detected when 'environmental' factors, such as a history of stressful life-events, are taken into consideration. (3) Stronger associations may also be observed with endophenotypes of a mental illness compared to the illness *per se*. For example, meta-analysis has shown *SERT* promoter polymorphisms to correlate more highly with 'neuroticism,' a personality trait highly associated with depression, than with depression itself.[49] (4) Additional *cis*-acting elements may need to be taken into consideration. In this study, we scored heterozygous SNPs as being positively correlated with AEI, if the measured AEI was >1.2. Perhaps stronger associations with mental illness could be detected using combinations of SNPs that predict higher levels of AEI, for example, >2. (5) If the gain-of-function we have observed for the rs7305115 *A*-allele indeed were to have phenotypic penetrance in mental disorders (in this case, possibly a protective effect), this may only become apparent in combination with variants in one or more additional genes that functionally interact with *TPH2*. AEI measurements, as described here for *TPH2,* have already revealed the presence of frequent functional polymorphisms in other genes previously implicated in mental disorders, including the μ-opiate (*OPRM1*),[14] monoamine oxidase A (*MAOA*)[50] and the type 2-dopamine receptor (*DRD2*) (manuscript in preparation), so that accounting for interactions among multiple genes could reveal significant impact on mental disorders, or variation in normal human behavior.

Clearly we are only at the beginning of the process of elucidating the genetic basis of mental illness. As for other complex diseases, multiple genes are likely to play a role. Identifying genetic variants that modify, or strongly predict, levels of mRNA expression for candidate genes provides a rich source of markers with high 'prior-probability' for association studies. In particular, using allele-specific mRNA expression as an intermediate phenotype is an efficient method for identifying 'functional' polymorphisms that contribute to the complex phenotypes associated with mental illness or response to therapeutic drugs.

Acknowledgments

This project was supported by NIDA 1R21 DA108744 (to WS), the Department of Pharmacology and a grant from the Psychiatric Research Foundation, Columbus Ohio (to DS). We thank Greg Young, Center for Biostatistics, Ohio State University for statistical analyses, Audrey Papp for advice on genotyping and AEI measurements, Andrew Johnson for help with computer program-related issues, and Gloria Smith for technical assistance with many of the assays.

References

1 Fitzpatrick PF. Tetrahydropterin-dependent amino acid hydroxylases. *Annu Rev Biochem* 1999; **68**: 355–381.
2 Lucki I. The spectrum of behaviors influenced by serotonin. *Biol Psychiatry* 1998; **44**: 151–162.
3 Ressler KJ, Nemeroff CB. Role of serotonergic and noradrenergic systems in the pathophysiology of depression and anxiety disorders. *Depress Anxiety* 2000; **12**(Suppl 1): 2–19.
4 Walther DJ, Peter JU, Bashammakh S, Hortnagl H, Voits M, Fink H *et al.* Synthesis of serotonin by a second tryptophan hydroxylase isoform. *Science* 2003; **299**: 76.
5 Zill P, Buttner A, Eisenmenger W, Moller HJ, Ackenheil M, Bondy B. Analysis of tryptophan hydroxylase I and II mRNA expression in the human brain: A post-mortem study. *J Psychiatr Res* 2007; **41**: 168–173.
6 Bach-Mizrachi H, Underwood MD, Kassir SA, Bakalian MJ, Sibille E, Tamir H *et al.* Neuronal tryptophan hydroxylase mrna expression in the human dorsal and median raphe nuclei: major depression and suicide. *Neuropsychopharmacology* 2006; **31**: 814–824.
7 Breidenthal SE, White DJ, Glatt CE. Identification of genetic variants in the neuronal form of tryptophan hydroxylase (TPH2). *Psychiatr Genet* 2004; **14**: 69–72.
8 Zhou Z, Peters EJ, Hamilton SP, McMahon F, Thomas C, McGrath PJ *et al.* Response to Zhang *et al.* (2005): loss-of-function mutation in tryptophan hydroxylase-2 identified in unipolar major depression. *Neuron* 2005; **45**: 11–16; **48**: 702–703; author reply 705–706.
9 Zhang X, Beaulieu JM, Gainetdinov RR, Caron MG. Functional polymorphisms of the brain serotonin synthesizing enzyme tryptophan hydroxylase-2. *Cell Mol Life Sci* 2006; **63**: 6–11.
10 Yan H, Yuan W, Velculescu VE, Vogelstein B, Kinzler KW. Allelic variation in human gene expression. *Science* 2002; **297**: 1143.

11 Bray NJ, Buckland PR, Owen MJ, O'Donovan MC. Cis-acting variation in the expression of a high proportion of genes in human brain. *Hum Genet* 2003; **113**: 149–153.
12 Pinsonneault J, Nielsen CU, Sadee W. Genetic variants of the human H+/dipeptide transporter PEPT2: analysis of haplotype functions. *J Pharmacol Exp Ther* 2004; **311**: 1088–1096.
13 Wang D, Johnson AD, Papp AC, Kroetz DL, Sadee W. Multidrug resistance polypeptide 1 (MDR1, ABCB1) variant 3435C>T affects mRNA stability. *Pharmacogenet Genom* 2005; **15**: 693–704.
14 Zhang Y, Wang D, Johnson AD, Papp AC, Sadee W. Allelic expression imbalance of human mu opioid receptor (OPRM1) caused by variant A118G. *J Biol Chem* 2005; **280**: 32618–32624.
15 Lim JE, Papp A, Pinsonneault J, Sadee W, Saffen D. Allelic expression of serotonin transporter (SERT) mRNA in human pons: lack of correlation with the polymorphism SERTLPR. *Mol Psychiatry* 2006; **11**: 649–662.
16 Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005; **21**: 263–265.
17 Zhou Z, Roy A, Lipsky R, Kuchipudi K, Zhu G, Taubman J *et al.* Haplotype-based linkage of tryptophan hydroxylase 2 to suicide attempt, major depression, and cerebrospinal fluid 5-hydroxyindoleacetic acid in 4 populations. *Arch Gen Psychiatry* 2005; **62**: 1109–1118.
18 Saffen D, Mieda M, Okamura M, Haga T. Control elements of muscarinic receptor gene expression. *Life Sci* 1999; **64**: 479–486.
19 Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* 2003; **31**: 3406–3415.
20 Conde L, Vaquerizas JM, Santoyo J, Al-Shahrour F, Ruiz-Llorente S, Robledo M *et al.* PupaSNP Finder: a web tool for finding SNPs with putative effect at transcriptional level. *Nucleic Acids Res* 2004; **32**: W242–W248.
21 Cartegni L, Chew SL, Krainer AR. Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat Rev Genet* 2002; **3**: 285–298.
22 Wang ET, Kodama G, Baldi P, Moyzis RK. Global landscape of recent inferred Darwinian selection for *Homo sapiens*. *Proc Natl Acad Sci USA* 2006; **103**: 135–140.
23 Ebmeier KP, Donaghey C, Steele JD. Recent developments and current controversies in depression. *Lancet* 2006; **367**: 153–167.
24 Zill P, Baghai TC, Zwanzger P, Schule C, Eser D, Rupprecht R *et al.* SNP and haplotype analysis of a novel tryptophan hydroxylase isoform (TPH2) gene provide evidence for association with major depression. *Mol Psychiatry* 2004; **9**: 1030–1036.
25 Zhang X, Gainetdinov RR, Beaulieu JM, Sotnikova TD, Burch LH, Williams RB *et al.* Loss-of-function mutation in tryptophan hydroxylase-2 identified in unipolar major depression. *Neuron* 2005; **45**: 11–16.
26 Harvey M, Shink E, Tremblay M, Gagne B, Raymond C, Labbe M *et al.* Support for the involvement of TPH2 gene in affective disorders. *Mol Psychiatry* 2004; **9**: 980–981.
27 Mossner R, Walitza S, Geller F, Scherag A, Gutknecht L, Jacob C *et al.* Transmission disequilibrium of polymorphic variants in the tryptophan hydroxylase-2 gene in children and adolescents with obsessive-compulsive disorder. *Int J Neuropsychopharmacol* 2006; **9**: 437–442.
28 Mossner R, Freitag CM, Gutknecht L, Reif A, Tauber R, Franke P *et al.* The novel brain-specific tryptophan hydroxylase-2 gene in panic disorder. *J Psychopharmacol* 2006; **20**: 547–552.
29 Walitza S, Renner TJ, Dempfle A, Konrad K, Wewetzer C, Halbach A *et al.* Transmission disequilibrium of polymorphic variants in the tryptophan hydroxylase-2 gene in attention-deficit/hyperactivity disorder. *Mol Psychiatry* 2005; **10**: 1126–1132.
30 Sheehan K, Lowe N, Kirley A, Mullins C, Fitzgerald M, Gill M *et al.* Tryptophan hydroxylase 2 (TPH2) gene variants associated with ADHD. *Mol Psychiatry* 2005; **10**: 944–949.
31 Coon H, Dunn D, Lainhart J, Miller J, Hamil C, Battaglia A *et al.* Possible association between autism and variants in the brain-expressed tryptophan hydroxylase gene (TPH2). *Am J Med Genet B Neuropsychiatr Genet* 2005; **135**: 42–46.
32 Zill P, Buttner A, Eisenmenger W, Moller HJ, Bondy B, Ackenheil M. Single nucleotide polymorphism and haplotype analysis of a novel tryptophan hydroxylase isoform (TPH2) gene in suicide victims. *Biol Psychiatry* 2004; **56**: 581–586.
33 De Luca V, Mueller DJ, Tharmalingam S, King N, Kennedy JL. Analysis of the novel TPH2 gene in bipolar disorder and suicidality. *Mol Psychiatry* 2004; **9**: 896–897.
34 De Luca V, Voineskos D, Wong GW, Shinkai T, Rothe C, Strauss J *et al.* Promoter polymorphism of second tryptophan hydroxylase isoform (TPH2) in schizophrenia and suicidality. *Psychiatry Res* 2005; **134**: 195–198.
35 De Luca V, Hlousek D, Likhodi O, Van Tol HH, Kennedy JL, Wong AH. The interaction between TPH2 promoter haplotypes and clinical-demographic risk factors in suicide victims with major psychoses. *Genes Brain Behav* 2006; **5**: 107–110.
36 Mergen H, Demirel B, Akar T, Senol E. Lack of association between the serotonin transporter and tryptophan hydroxylase gene polymorphisms and completed suicide. *Psychiatr Genet* 2006; **16**: 53.
37 Garriock HA, Allen JJ, Delgado P, Nahaz Z, Kling MA, Carpenter L *et al.* Lack of association of TPH2 exon XI polymorphisms with major depression and treatment resistance. *Mol Psychiatry* 2005; **10**: 976–977.
38 Van Den Bogaert A, De Zutter S, Heyrman L, Mendlewicz J, Adolfsson R, Van Broeckhoven C *et al.* Response to Zhang *et al.* (2005): loss-of-function mutation in tryptophan hydroxylase-2 identified in unipolar major depression. *Neuron* 2005; **45**: 11–16; **48**: 704; author reply 705–706.
39 Glatt CE, Carlson E, Taylor TR, Risch N, Reus VI, Schaefer CA. Response to Zhang *et al.* (2005): loss-of-function mutation in tryptophan hydroxylase-2 identified in unipolar major depression. *Neuron* 2005; **45**: 11–16; **48**: 704–705; author reply 705–706.
40 Delorme R, Durand CM, Betancur C, Wagner M, Ruhrmann S, Grabe HJ *et al.* No human tryptophan hydroxylase-2 gene r441h mutation in a large cohort of psychiatric patients and control subjects. *Biol Psychiatry* 2006; **60**: 202–203.
41 Brown SM, Peet E, Manuck SB, Williamson DE, Dahl RE, Ferrell RE *et al.* A regulatory variant of the human tryptophan hydroxylase-2 gene biases amygdala reactivity. *Mol Psychiatry* 2005; **10**: 884–888, 805.
42 Canli T, Congdon E, Gutknecht L, Constable RT, Lesch KP. Amygdala responsiveness is modulated by tryptophan hydroxylase-2 gene variation. *J Neural Transm* 2005; **112**: 1479–1485.
43 Herrmann MJ, Huter T, Muller F, Muhlberger A, Pauli P, Reif A *et al.* Additive effects of serotonin transporter and tryptophan hydroxylase-2 gene variation on emotional processing. *Cereb Cortex* 2006; 26 June [Epub ahead of print].
44 Heils A, Teufel A, Petri S, Seemann M, Bengel D, Balling U *et al.* Functional promoter and polyadenylation site mapping of the human serotonin (5-HT) transporter gene. *J Neural Transm Gen Sect* 1995; **102**: 247–254.
45 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.
46 Whitaker-Azmitia PM. Serotonin and brain development: role in human developmental diseases. *Brain Res Bull* 2001; **56**: 479–485.
47 Gaspar P, Cases O, Maroteaux L. The developmental role of serotonin: news from mouse molecular genetics. *Nat Rev Neurosci* 2003; **4**: 1002–1012.
48 Caspi A, Sugden K, Moffitt TE, Taylor A, Craig IW, Harrington H *et al.* Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science* 2003; **301**: 386–389.
49 Munafo MR, Clark TG, Roberts KH, Johnstone EC. Neuroticism mediates the association of the serotonin transporter gene with lifetime major depression. *Neuropsychobiology* 2006; **53**: 1–8.
50 Pinsonneault JK, Papp AC, Sadee W. Allelic mRNA expression of X-linked monoamine oxidase a (MAOA) in human brain: dissection of epigenetic and genetic factors. *Hum Mol Genet* 2006; **15**: 2636–2649.

Supplementary Information accompanies the paper on the Molecular Psychiatry website (http://www.nature.com/mp)

Fig. S1.

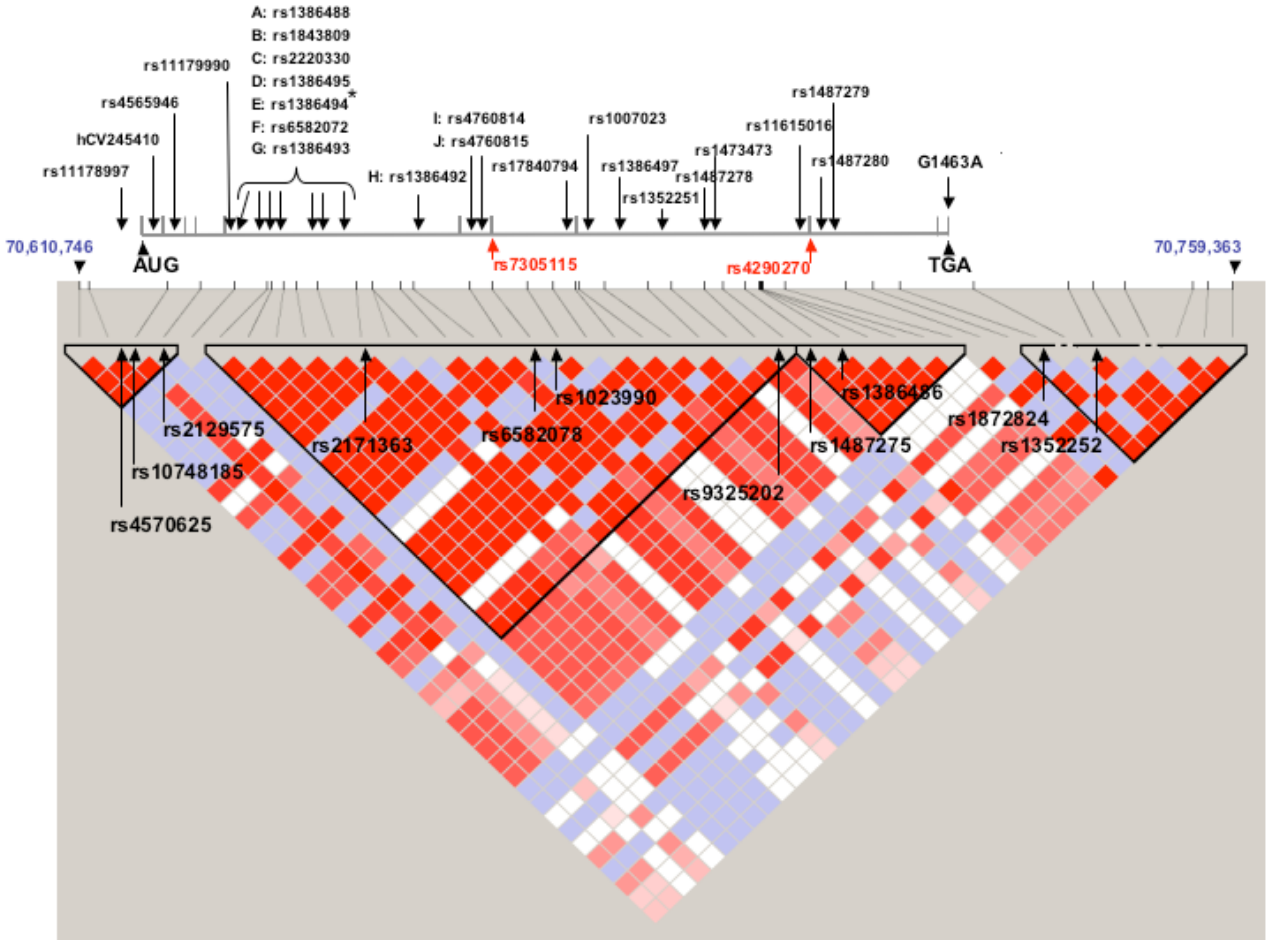

A: rs1386488
B: rs1843809
C: rs2220330
D: rs1386495
E: rs1386494*
F: rs6582072
G: rs1386493
rs11179990
rs4565946
hCV245410
rs11178997
H: rs1386492
I: rs4760814
J: rs4760815
rs17840794
rs1007023
rs1386497
rs1352251
rs1473473
rs1487278
rs1487279
rs11615016
rs1487280
G1463A
70,610,746
AUG
rs7305115
rs4290270
TGA
70,759,363
rs2129575
rs10748185
rs4570625
rs2171363
rs6582078
rs1023990
rs9325202
rs1386486
rs1487275
rs1872824
rs1352252

Fig. S1. Haplotype structure of the human TPH2 gene and locations of key SNPs. The grey bar in the center of the figure represents the transcribed region of the TPH2 gene. Exons (1-11) are represented by vertical grey bars. The open bar below the transcribed region represents the segment of chromosome 12 (12q21) containing the TPH2 gene. The exact chromosomal location of this segment is indicated by the numbers at the beginning and end of the open bar. The vertical lines within the open bar denote the positions of the HapMap SNPs that were used for the determination of the haplotype structure of the TPH2 gene. The rs numbers for 11 HapMap SNPs examined in this study are listed below the open bar. The marker SNPs (rs7305115 and rs4290270) examined in this study are indicated in red type. The location of a rare missense mutation that reduces tryptophan hydroxylase activity (G1463A) is also indicated. The set of SNPs examined by Zill and coworkers in association studies of TPH2 and depression or suicide are annotated with the letters A though J. A SNP showing a statistically significant association with major depression (E: rs1386494) is marked with an asterisk (*). The triangular plot in the bottom half of the figure depicts estimated pairwise linkage disequilibrium (D') values for HapMap SNPs. The plot was generated using the Haploview version 3.2 program with genotyping data from the CEU (Utah residents with ancestry from northern and western Europe) sample. Both the program and data set were downloaded from the International HapMap Project website (http://www.hapmap.org). Red boxes indicate high estimated linkage disequilibrium (D') between pairs of SNPs. Blue, pink and white boxes indicate lower estimated linkage disequilibrium (bright red: D' = 1, LOD ≥ 2; blue: D' = 1, LOD< 2; pink: D' <1, LOD≥2; white: D' < 1, LOD < 2).

Fig. S2A

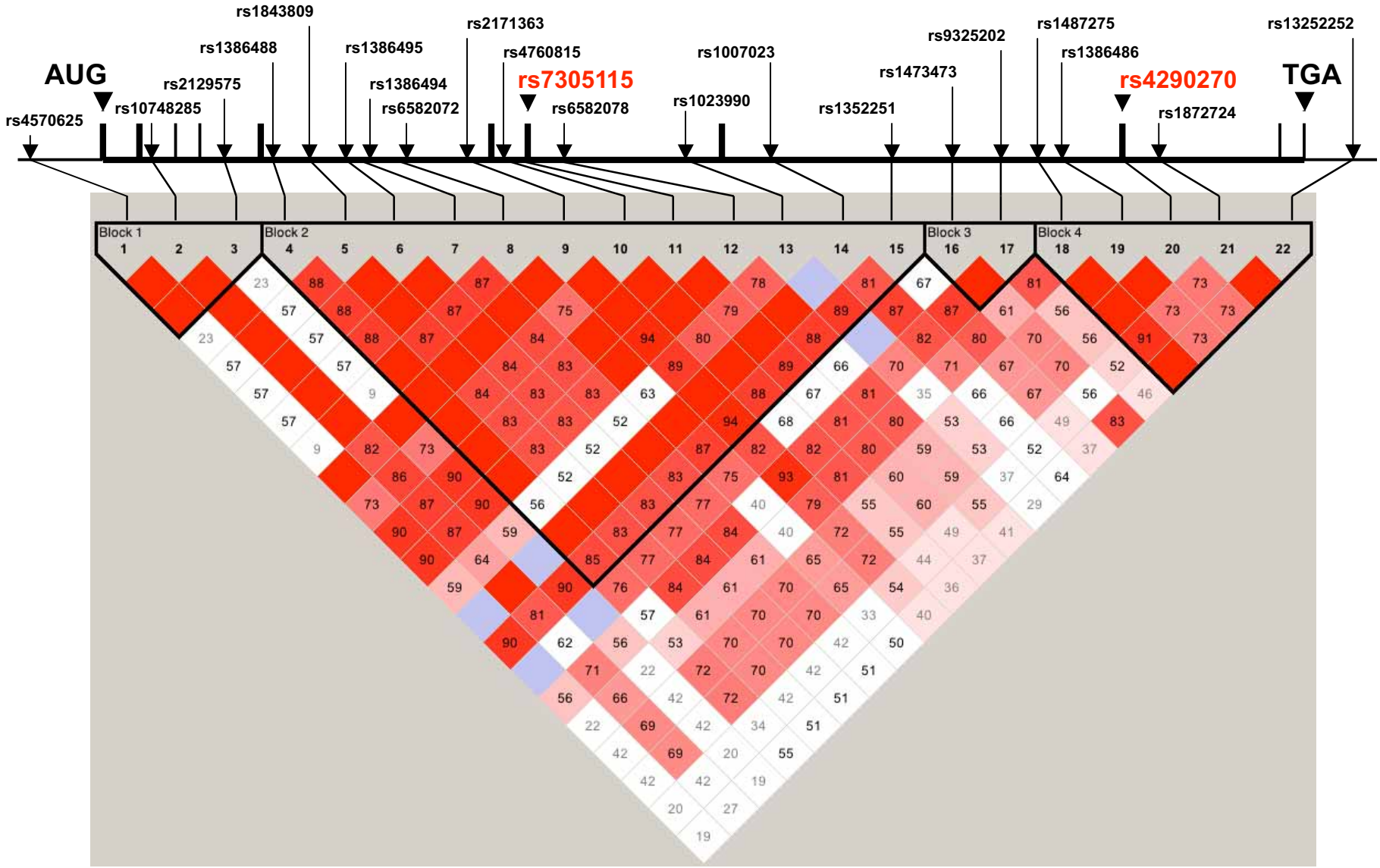

AUG
TGA
rs4570625
rs10748285
rs2129575
rs1386488
rs1843809
rs1386495
rs1386494
rs6582072
rs2171363
rs4760815
rs7305115
rs6582078
rs1007023
rs1023990
rs1352251
rs1473473
rs9325202
rs1487275
rs1386486
rs4290270
rs1872724
rs13252252
Block 1
Block 2
Block 3
Block 4

Fig. S2B

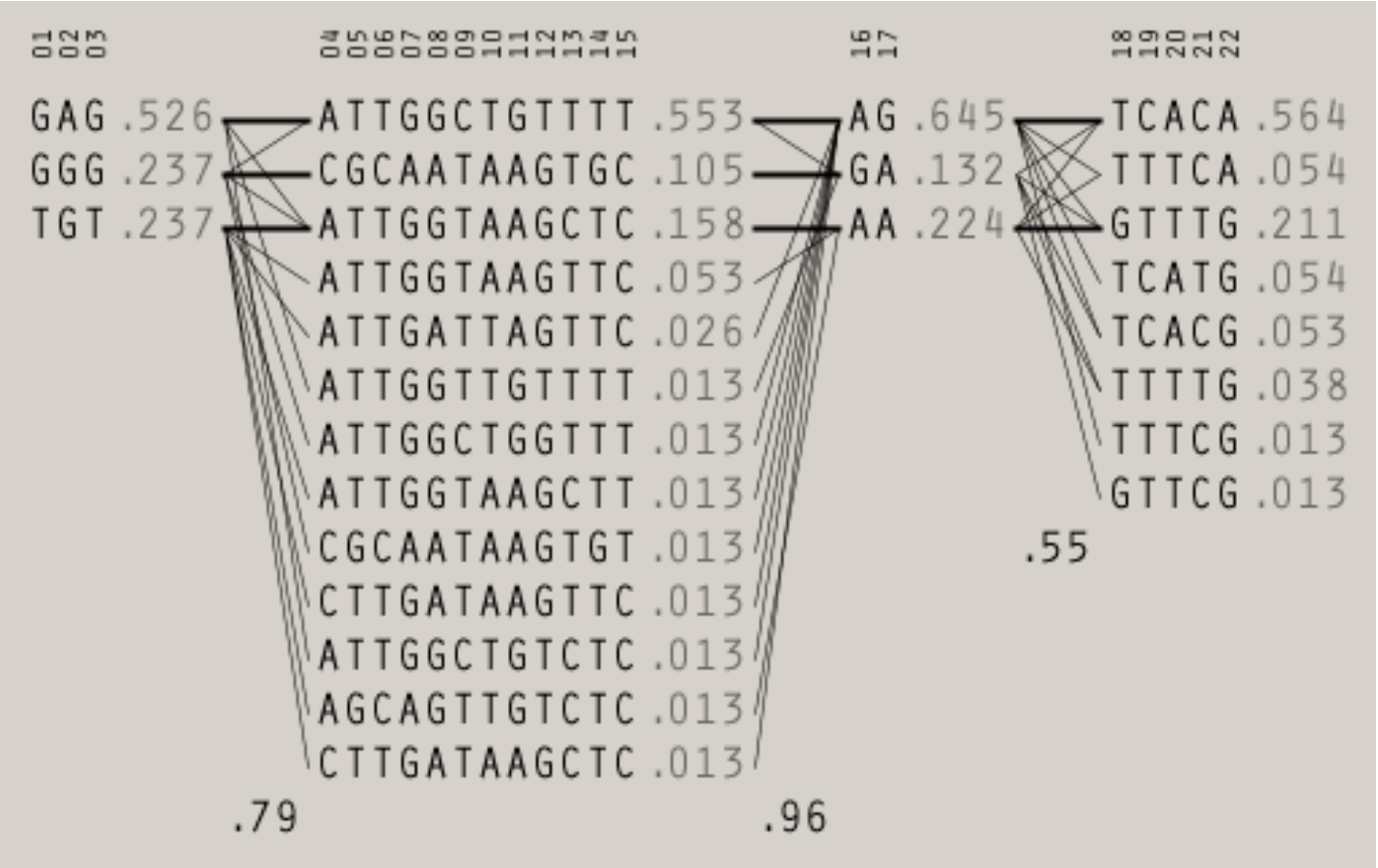

01 02 03
04 05 06 07 08 09 10 11 12 13 14 15
16 17
18 19 20 21 22
GAG .526
GGG .237
TGT .237
ATTGGCTGTTTT .553
CGCAATAAGTGC .105
ATTGGTAAGCTC .158
ATTGGTAAGTTC .053
ATTGATTAGTTC .026
ATTGGTTGTTTT .013
ATTGGCTGGTTT .013
ATTGGTAAGCTT .013
CGCAATAAGTGT .013
CTTGATAAGTTC .013
ATTGGCTGTCTC .013
AGCAGTTGTCTC .013
CTTGATAAGCTC .013
AG .645
GA .132
AA .224
TCACA .564
TTTCA .054
GTTTG .211
TCATG .054
TCACG .053
TTTTG .038
TTTCG .013
GTTCG .013
.79
.96
.55

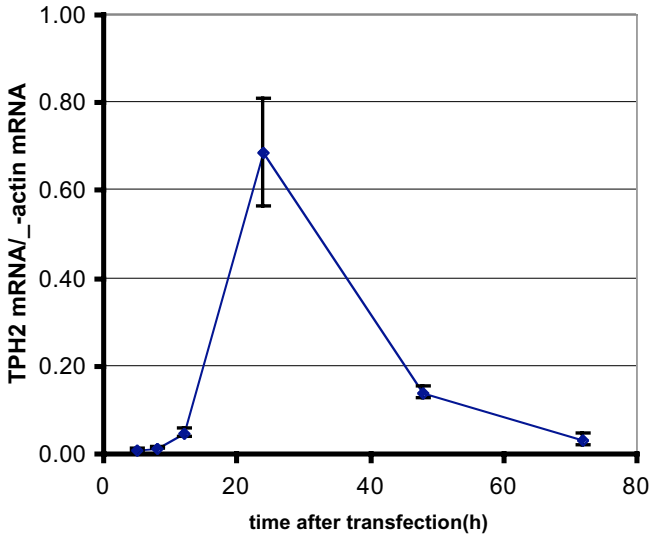
Fig. S2

A. D' plot for the 22 SNPs listed in Table 1 (main text) based upon genotyping data from 36 Caucasian individuals in our collection. The plot was generated using Haploview (version 3.3; LD plot>Analysis>Solid Spine of LD, where the LD spine was extended if D' > 0.7). Red boxes indicate high estimated linkage disequilibrium (D') between pairs of SNPs. Blue, pink and white boxes indicate lower estimated linkage disequilibrium (bright red: D' = 1, LOD ≥ 2; blue: D' = 1, LOD< 2; pink: D' <1, LOD≥2; white: D' < 1, LOD < 2). Haplotype blocks demarcate segments of high linkage disequilibrium. Number within each square = D' x 100.

B. Estimated haplotypes and population frequencies for each haplotype block. Multiblock haplotypes are indicated by the lines between the blocks, with frequencies corresponding to the thickness of the lines. Observed frequencies of haplotypes within each block are listed in grey type. The numbers in black type are Hendrick multiallelic D's, which estimate linkage disequilibrium between blocks by treating each block as an individual "allele."

Fig. 3S.

A.


1.00
0.80
0.60
0.40
0.20
0.00
TPH2 mRNA/_-actin mRNA
0
20
40
60
80
time after transfection(h)

B.

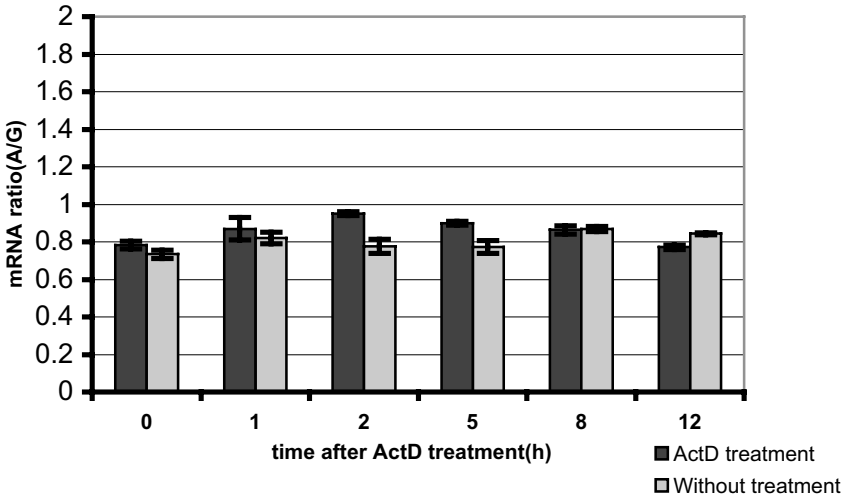

2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
mRNA ratio(A/G)
0
1
2
5
8
12
time after ActD treatment(h)
ActD treatment
Without treatment

Fig. S3. Comparison of *TPH2* mRNA stability for rs7305115 *A*- and *G*-alleles.

A. Levels of *TPH2* mRNA were quantified by real-time PCR at the indicated times (h) following transfection of CHO cells with an expression vector encoding human *TPH2* (rs7305115 *A*-allele) at t = 0. As indicated, highest levels of TPH2 *A*-allele mRNA were detected 24 h after transfection. Similar results were obtained following transfection of CHO cells with an expression vector encoding the TPH2 *G*-allele (*data not shown*). B. Allelic expression imbalance (AEI) assays for *TPH2 A*- and *G*-alleles were carried out using RNA isolated from CHO cells transfected with equal-molar amounts of expression vector encoding the *TPH2 A*- and *TPH2 G*-alleles. RNA was isolated at the indicated times following addition of 10 μg/ml actinomycin D (added 24 h after transfection). As indicted, AEI ratios did not change with time in either cells treated with actinomycin D (black bars) or not treated with actinomycin (grey bars). These data indicate that the rate of mRNA decay is the same for the *TPH2 A*- and *G*-alleles, both in the presence or absence of actinomycin D.

Methods.

1) Expression vectors: Reverse transcriptase was used to synthesize cDNA from RNA isolated from an individual homozygous for the TPH2 A-allele of rs7305115. An expression vector encoding the *TPH2 A*-allele was constructed by subcloning this cDNA in the *Bam*H I / *Xba* I site of pcDNA3.1. An expression vector encoding the *TPH2 G*-allele was produced by using site-directed mutagenesis to convert the *A*-allele to a *G*. DNA sequencing of the *TPH2* coding regions confirmed that the only difference between the expression vectors was the presence of the *A*- or *G*-allele.

2) Transfections: CHO cells were cultured at 37ºC in a humidified incubator at 5% CO_2 in Ham's F-12 Medium plus 10% fetal bovine serum, 100 U/ml penicillin and 100 mg/ml streptomycin. The day before transfection, cells were re-plated into 6-well plates at approximately 50% confluency. Transfection of *TPH2* expression constructs was performed using lipofectamine 2000 reagent according to the manufacturer's protocol. To determine the time course of TPH2 expression, CHO cells were transfected with 4mg TPH2-A expression vector. Total RNA was isolated at 5, 8 12, 24 48 72 h after transfection to determine peak levels of TPH2 mRNA expression. For mRNA stability studies, CHO cells were co-transfected with 2 mg (each) of *TPH2-A* and *TPH2-G*. Twenty-hour hours after transfection, the cells were treated with vehicle or 10 mg/ml actinomycin D for 0, 1, 2, 5, 8, and 12 hrs. At these time points, cell cultures were either trypsinized and collected for plasmid DNA preparation using QIAGEN mini prep kits, or lysed with 1 ml Trizol, followed by RNA purification with QIAGEN easy RNA mini prep kits. Contaminating DNA in the RNA samples was eliminated by DNase I treatment prior to column purification. The amplification primers did not amplify cDNA prepared from untransfected CHO cells, indicating that the primers used in this study specifically detected TPH2 mRNA produced from the expression vectors.

3) mRNA quantification: *TPH2* mRNA levels were measured in transfected CHO cells by reverse transcription followed by real-time PCR analysis. Endogenous β-actin mRNA was also measured using primers specific for hamster β-actin. The expression of TPH2 was expressed as the ratio of TPH2 mRNA /β-actin mRNA. To ensure absence of genomic DNA in RNA samples, control tubes containing the same amounts of RNA without reverse transcriptase were also assayed. Real-time PCR analysis showed the

cycle thresholds from these control samples were higher than 30 cycles, similar to blank controls, showing that genomic DNA levels were undetectable.

Table S1. Predicted diplotypes for individuals in sample

sample#	Haplotype 1	Haplotype 2	EM-p
602	T G T C T T G A **T A A G** T T C A **G** T T T C G	G G G C G C A A **T A A G** T G C G **A** G T T T G	1
813	G A G A T T G G **C T G T** T T T A **G** T C A C A	T G T A T T G G **T A A G** C T C A **A** G T T T G	0.99
879	G A G A T T G G **C T G T** T T T A **G** T C A C A	G G G C G C A A **T A A G** T G C G **A** G T T T G	0.98
914	G A G A T T G G **C T G T** T T T A **G** T C A C G	G G G A T T G G **T A A G** C T C A **A** T T T T G	0.99
917	G A G A T T G G **C T G T** T T T A **G** T C A T G	T G T A T T G G **T A A G** C T C A **A** G T T T G	1
1025	G A G A T T G G **C T G T** T T T A **G** T C A C G	G A G A T T G G **C T G T** T T T A **G** T C A T G	1
1027	G A G A T T G G **C T G T** T T T A **G** T C A C A	T G T A T T G G **T A A G** C T T A **G** T C A C A	1
1029	G A G A T T G G **C T G T** T T T A **G** T C A C A	G A G A T T G G **C T G T** T T T A **G** T C A C A	1
1054	G A G A T T G G **C T G T** T T T A **G** T C A C G	G A G A T T G G **C T G T** T T T A **G** G T T T G	1
1065	G G G C G C A A **T A A G** T G C G **A** G T T C G	T G T A T T G G **T A A G** C T C A **A** G T T T G	0.99
1078	G A G A T T G G **C T G T** T T T A **G** T C A C A	G A G A T T G G **T A A G** C T C A **A** G T T T G	0.96
1078	G A G A T T G G **C T G T** T T T A **G** G T T T G	G A G A T T G G **T A A G** C T C A **A** T C A C A	0.04
1101	G A G A T T G G **C T G T** T T T A **G** T C A C A	T G T A T T G G **T A A G** C T C A **A** G T T T G	0.99
1103	G A G A T T G G **C T G T** T T T A **G** T C A C A	T G T A T T G G **T A A G** C T C A **A** G T T T G	0.99
1104	G G G A T T G G **C T G T** T T T A **G** T C A C A	G G G A T T G G **C T G T** T T T A **G** T T T C A	1
1105	G A G A T T G G **C T G T** T T T A **G** T C A C A	T G T A T T G G **T A A G** C T C A **A** G T T T G	0.99
1112	G A G A T T G G **C T G T** T T T A **G** T C A T G	T G T A T T G G **T A A G** C T C A **A** G T T T G	1
1115	T G T A T T G G **T A A G** T T C A **A** T C A C A	T G T A T T G G **T A A G** T T C A **A** T C A C A	1
1135	G A G A T T G G **C T G T** T T T A **G** T C A C A	G G G C G C A A **T A A G** T G C G **A** G T T T G	0.98
1169	G G G A G C G G **T A A G** C T C G **A** T T T T G	T G T A T T G G **T A A G** T T C A **A** T C A C A	1
1209	G A G A T T G G **C T G T** T T T A **G** T C A C A	G A G A T T G G **C T G T** T T T A **G** T C A C G	1
1230	G A G A T T G G **C T G T** T T T A **G** T C A C A	T G T A T T G G **T A A G** C T C A **A** T C A C A	1
1257	G A G A T T G G **C T G T** T T T A **G** T C A C A	G A G A T T G G **C T G T** T T T A **G** T C A T G	1
1269	G A G A T T G G **C T G T** T T T A **G** T C A C A	G G G A T T G G **C T G T** T T T A **G** T C A C A	1
1279	G A G A T T G G **C T G G** T T T A **G** T C A C A	T G T A T T G G **T A A G** C T C A **A** G T T T G	0.99
1297	G G G A T T G G **C T G T** T T T A **G** T T T C A	T A T A T T G G **T T G T** T T T A **G** T T T T G	0.99
1347	G A G A T T G G **C T G T** T T T A **G** T C A C A	G G G A T T G G **C T G G** T T T A **G** T C A T G	1
1365	T G T A G C A G **T T G T** C T C A **A** T T T C A	G A G A T T G G **T T G T** T T T G **A** T T T T G	1

1407	T G T A T T G G **T T G T** C T C A **A** T T T T G	T G T A G C G G **T T G T** T T C A **G** T T T T G	1
1409	G A G A T T G G **C T G T** T T T A **G** T C A C A	G A G A T T G G **C T G T** T T T G **A** T C A C G	1
1429	G A G A T T G G **C T G T** T T T A **G** T C A C A	G A G A T T G G **C T G T** T T T A **G** T C A C A	1
1430	G A G A T T G G **T T G T** T T T A **G** T T A C A	G A G A T T G G **T T G T** T T T A **G** T T T T G	1
1442	G A G A T T G G **C T G T** T T T A **G** T C A C A	G G G A T T G G **C T G T** T T T A **G** T T T C A	1
1486	T G T A T T G A **T T A G** T T C A **G** T X Y Z A	T G T A T T G A **T T A G** T T C A **G** T X Y Z G	1
1489	T G T A T T G G **C T G T** T T T A **G** T T T T G	T G T C T T G A **T A A G** T T T A **G** T C A C A	1
1500	G A G A T T G G **C T G T** T T T A **G** T C A C A	G G G A T T G G **C T G T** T T T A **G** T C A C A	1
1539	G A G A T T G G **C T G T** T T T A **G** T C A C A	G A G A T T G G **C T G T** T T T A **G** T C A C A	1
1540	G A G A T T G G **C T G T** T T T A **G** T C A C A	G G G C G C A A **T A A G** T G C G **A** T C A C A	1
1546	T G G A G C G G **T T G T** C T C A **A** G C A T G	T G G A G C G G **T T G T** T T C A **A** T T T T G	0.50
1546	T G G A G C G G **T T G T** T T C A **A** G C A T G	T G G A G C G G **T T G T** C T C A **A** T T T T G	0.50
1551	G A G A T T G G **C T G T** T T T A **G** G T T T G	G G G C G C A A **T A A G** T G C G **A** G T T T G	1
1584	G A G A T T G G **C T G T** T T T A **G** T C A C A	G A G A T T G G **C T G T** T T T A **G** T C A C A	1
1607	G A G A T T G G **C T G T** T T T A **G** T C A C A	G G G C G C A A **T A A G** T G C G **A** G T T T G	0.98
1609	G A G A T T G G **C T G T** C T C A **G** T C A C A	T G T A T T G G **T A A G** T T C A **A** T C A C A	1
1613	G A G A T T G G **T A A G** C T C A **A** T C A C A	G G G C G C A A **T A A G** T G C G **A** G T T T G	0.73
1613	G G G C G C A A **T A A G** T G C G **A** T C A C A	G A G A T T G G **T A A G** C T C A **A** G T T T G	0.27
1614	T G T A T T G G **T A A G** T T C A **A** T C A C A	T G T C T T G A **T A A G** C T C A **G** T T T C A	1
1672	G G G A T T G G **T A A G** C T C A **A** T C A C A	G G G C G C A A **T A A G** T G T A **G** T C A C A	1
1675	G A G A T T G G **C T G T** T T T A **G** T C A C A	G A G A T T G G **C T G T** T T T A **G** T C A C A	1
1744	G A G A T T G G **C T G T** T T T A **G** T C A C A	G G G A T T G G **C T G T** T T T A **G** T C A C A	1
1745	G A G C G C A A **T A G G** T G C G **A** G T T T G	G G G C G C A A **T A G G** T G C G **A** G T T T G	1

Diplotypes were predicted from genotyping data for the 48 individuals in our sample for the 22 SNPs listed in Table S2 using HelixTree. Only one predicted diplotype is shown for cases where the estimation-maximum probability (EM-p) was 0.98 or greater. Accurate predictions could not be made for three SNPs (#19, 20, 21) in sample 1486: X = C/T; Y = A/T; Z = C/T. Alleles of SNPs for which heterozygosity is highly correlated with TPH2 AEI (Kappa coefficient > 0.66) are listed in **bold** type.

Table S2.

#	dbSNP	a = Hetero & AEI (+)	b = Hetero & AEI (-)	c = Homo & AEI (+)	d = Homo & AEI (-)	κ	p-value
01	rs4570625	0.370 (10/27)	0.037 (1/27)	0.333 (9/27)	0.259 (7/27)	0.311	0.053
02	rs10748185	0.556 (15/27)	0.111 (3/27)	0.148 (4/27)	0.185 (5/27)	0.400	0.037
03	rs2129575	0.370 (10/27)	0.037 (1/27)	0.333 (9/27)	0.259 (7/27)	0.311	0.053
04	rs1386488	0.222 (6/27)	0.074 (2/27)	0.481 (13/27)	0.222 (6/27)	0.047	0.732
05	rs1843809	0.148 (4/27)	0.111 (3/27)	0.519 (14/27)	0.222 (6/27)	-0.115	0.373
06	rs1386495	0.148 (4/27)	0.111 (3/27)	0.556 (15/27)	0.185 (5/27)	-0.115	0.373
07	rs1386494	0.148 (4/27)	0.074 (2/27)	0.556 (15/27)	0.222 (6/27)	-0.027	0.822
08	rs6582072	0.222 (6/27)	0.074 (2/27)	0.481 (13/27)	0.222 (6/27)	0.047	0.732
09	**rs2171363**	0.630 (17/27)	0.037 (1/27)	0.074 (2/27)	0.259 (7/27)	**0.743**	**< 0.001**
10	**rs4760815**	0.630 (17/27)	0.037 (1/27)	0.074 (2/27)	0.259 (7/27)	**0.743**	**< 0.001**
11	**rs7305115**	0.630 (17/27)	0.037 (1/27)	0.074 (2/27)	0.259 (7/27)	**0.743**	**< 0.001**
12	**rs6582078**	0.593 (16/27)	0.037 (1/27)	0.111 (3/27)	0.259 (7/27)	**0.669**	**< 0.001**
13	rs1023990	0.481 (13/27)	0.111 (3/27)	0.222 (6/27)	0.185 (5/27)	0.279	0.135
14	rs1007023	0.148 (4/27)	0.074 (2/27)	0.556 (15/27)	0.222 (6/27)	-0.027	0.822
15	**rs1352251**	0.519 (14/27)	0.037 (1/27)	0.185 (5/27)	0.259 (7/27)	**0.534**	**0.003**
16	rs1473473	0.148 (4/27)	0.111 (3/27)	0.556 (15/27)	0.185 (5/27)	-0.115	0.373
17	**rs9325202**	0.593 (16/27)	0.037 (1/27)	0.111 (3/27)	0.259 (7/27)	**0.669**	**< 0.001**
18	rs1487275	0.407 (11/27)	0.111 (3/27)	0.296 (8/27)	0.185 (5/27)	0.173	0.333
19	rs1386486	0.519 (14/27)	0.259 (7/27)	0.185 (5/27)	0.037 (1/27)	-0.149	0.430
20	rs4290270	0.556 (15/27)	0.259 (7/27)	0.148 (4/27)	0.037 (1/27)	-0.096	0.601
21	rs1872824	0.444 (12/27)	0.148 (4/27)	0.259 (7/27)	0.148 (4/27)	0.119	0.525
22	rs1352252	0.407 (11/27)	0.111 (3/27)	0.296 (8/27)	0.185 (5/27)	0.173	0.333

$\kappa = 2(ad-bc)/(p_1q_2+p_2q_1)$; where a = proportion of samples heterozygous & AEI(+); b = proportion of samples heterozygous & AEI (-); c = proportion of samples homozygous & AEI(+); d = proportion of samples homozygous & AEI(-); p_1 = proportion of samples that are heterozygous for a give SNP (see Table S3); q_1 = proportion of samples that are homozygous for a given SNP (see Table S3); p_2 = proportion of samples that are AEI(+) = 0.667 (18/27); q_2 = proportion of samples that are AEI(-) = 0.333 (9/27). Sample size = number of samples where AEI measurements were possible = number of samples heterozygous for marker SNP rs7305115 or rs4290270 = 27. k > 0.75: excellent agreement; 0.4 to 0.75 = fair to good agreement; < 0.4 = poor agreement.