REGULATORY GENETIC VARIANTS IN MENTAL ILLNESS:
FOCUS ON SEROTONIN-RELATED GENES

DISSERTATION

Presented in Partial Fulfillment of the Requirement for
the Degree Doctor of Philosophy in the Graduate School of
The Ohio State University

By

Jeong-Eun Lim, M. S.
The Ohio State University

2007

Dissertation Committee:
Professor Dr. David Saffen, Coadviser
Professor Dr. Wolfgang Sadee, Coadvisor
Professor Dr. Howard Gu
Professor Dr. Kirk Mykytyn
Professor Dr. Glenn Lin

Coadvisor
Coadvisor

Neuroscience Graduate Studies Program
ABSTRACT

My Ph.D. studies were aimed at identifying functional genetic variants that affect mRNA expression of genes implicated in psychiatric disorders. Genes encoding the serotonin transporter (*SERT*) and tryptophan hydroxylase 2 (*TPH2*) were the particular focus of my studies since disruption of brain serotonergic systems has been implicated in a variety of mental illnesses, including major depression (MD), anxiety disorders, schizophrenia, alcoholism, drug abuse, aggression and suicidal behavior. Most of the serotonin in the brain is produced by serotonergic neurons located in the raphe nuclei of the pons and brain stem. These cells express SERT, which functions to terminate the synaptic actions of serotonin and TPH2, which catalyzes the rate-limiting step in the synthesis of serotonin. Levels of expression of the *SERT* and *TPH2* genes may therefore control levels of serotonergic activity in the brain. I hypothesized that genetic variants that alter mRNA expression of these genes affect the risk of depression or suicide.

To identify functional genetic variants that influence the mRNA expression, I measured expression of mRNA in an allele-specific manner in post-mortem human pons sections containing the dorsal and median raphe nuclei. Any difference in the expression of one allele over the other indicates the presence of *cis*-acting (i.e., local) elements that differentially affect transcription and/or mRNA processing and turnover.

Using a marker SNP in the 3’ untranslated region of *SERT* mRNA, statistically significant differences in allelic mRNA levels were detected in nine out of 29 samples heterozygous for the marker SNP. While the allelic expression differences were relatively small (15–25%), they could nevertheless be physiologically relevant. In contrast to previous results, positive correlation was
not observed between SERTLPR and allelic expression ratios, implying regulation of SERT mRNA is not dependent of SERTLPR.

In TPH2, significant allelic expression imbalance (AEI) was detected ranging from 1.2 to 2.5-fold in 19 out of the 27 samples using two marker SNPs markers: rs7305115 (located in exon 7) and rs4290270 (located in exon 9), suggesting the existence of cis-acting polymorphisms that differentially affect TPH2 mRNA levels in pons. The minor A-allele yielded higher levels of TPH2 mRNA expression than the G-major allele. By genotyping twenty additional TPH2 SNPs, I identified a haplotype block comprising five tightly linked SNPs for which heterozygosity is highly correlated with AEI.

My studies have demonstrated the utility of a novel genetic approach to finding functional polymorphisms in genes of interest. Widely applied, this method should increase the ability to demonstrate the association of specific genes with mental disorders and provide a mechanistic explanation for the role of these genes in the etiology of mental illness.
ACKNOWLEDGMENTS

I would like to express my deepest gratitude to a number of people who have been essential for the completion of my Ph.D. work. I am most grateful to my advisors, Dr. David Saffen and Dr. Wolfgang Sadee, whose guidance and support has nurtured me both scientifically and personally. Without their constructive criticism and distinguished scientific insight, writing this dissertation would not have been possible.

I also thank my committee members, Dr. Howard Gu, Dr. Kirk Mykytyn and Dr. Glenn Lin for the valuable discussions and advice.

I have been lucky to be surrounded by wonderful colleagues of Dr. Sadee’s laboratory, who have provided both help with experiments and friendship. Many thanks go to Audrey Papp, Julia Pinsonneault, Danxin Wang, Gloria Smith, Zunyan Dai, Ying Zhang and Andrew Johnson for their kind cooperation and technical advice. I thank my friends for helping me get through tough times and for helping me enjoy life in Columbus. I especially want to thank Dr. Bo-Young Lee, Dr. Joung-Woo Hong, Dr. Ju-Yong Kim and Eun-Joo Lee.

I am deeply indebted to my father, mother, and my sister for their unconditional love and support. Finally, my sincere appreciation goes to my dear husband, Cheon-Jae Bahk for his love and support, and my lovely daughter, Claire. She is my joy of life and ultimate motivation.
VITA

July, 1975..........................................Born in Seoul, Korea
1994-1999.......................................... B.S. Pharmacy,

      Seoul National University, Seoul, South Korea
1999-2001......................................... M.S. Pharmaceutical science,

      Seoul National University, Seoul, South Korea
2002-present ……………………. Graduate Research Associate,

      Ohio State University

PUBLICATIONS

Research Publication


FIELDS OF STUDY

Major Field: Neuroscience
# TABLE OF CONTENTS

Abstract.................................................................................................................................................. ii
Acknowledgments........................................................................................................................................ vi
Vita............................................................................................................................................................. ix
List of Figures .............................................................................................................................................. vii
List of Tables ............................................................................................................................................... ix

Chapters :

1.  Introduction........................................................................................................................................ 1

2.  Allelic expression of serotonin transporter (*SERT*) mRNA in human pons: lack of correlation with the polymorphism *SERTLPR*........................................................................................................ 14

3.  Tryptophan hydroxylase 2 (*TPH2*) haplotypes predict levels of *TPH2* mRNA expression in human pons........................................................................................................................................... 48

4.  The effect of genetic variants on exon skipping in *TPH2* (an on-going project)............. 84

Discussion and future studies.................................................................................................................. 102

References.................................................................................................................................................. 111
LIST OF FIGURES

CHAPTER 2
2.1. A. Structure of the SERT gene and location of SERTLPR, VNTR and 3’-UTR marker SNP G2651T (rs1042173) B. Genotyping the long (l)/short (s) SERTLPR promoter polymorphism C. Genotyping the VNTR located in intron2.................................................................34
2.2. Logarithmic plots of real-time PCR amplifications.............................................35
2.3. Assay of SERT mRNA levels in pons RNA samples ...........................................36
2.4. Assay of SERT mRNA levels in RNA of non-pons regions and pons.....................37
2.5. Measurement of SERT allelic expression imbalance (AEI).....................................38
2.6. Comparison of corrected genomic DNA/RNA ratios in human pons measured with primer extension assays using the marker SNP G2651T.................................................................40
2.7. SERT mRNA expression in lymphoblasts ..........................................................41
2.8. Comparison of corrected genomic DNA/RNA rations in lymphoblast cell lines.........42

CHAPTER 3
3.1. Haplotype structure of the human TPH2 gene and locations of key SNPs...............66
3.2. Comparison of genomic DNA/RNA ratios assayed using the marker SNP rs7305115......68
3.3. Comparison of genomic DNA/RNA ratios assayed using the marker SNP rs4290270......69
3.4. Comparison of mRNA allelic expression ratios determined using the marker SNP rs7305115 and rs4290270.........................................................................................................................70
3.5. A. D’ plot for the 22 SNPs listed in Table 3-1. B. Estimated haplotypes and population
3.6. Correlation between heterozygosity of individual TPH2 SNPs and AEI of TPH2 mRNA…73
3.7. TPH2 mRNA levels in pons measured using RT-PCR………………………………………..74
3.8. Comparison of TPH2 mRNA expression levels in different tissues…………………………..75
3.9. Comparison of TPH2 mRNA stability for rs7305115 A- and G alleles………………………76

CHAPTER 4
4.1. Diagram of a segment of the TPH2 gene extending from exon 6 through exon 8…………98
4.2. Output window from the PupaSuite website…………………………………………………99
4.3. Agarose gel electrophoresis photograph of PCR products derived from mRNA containing or lacking exon7…………………………………………………………………………………..100
4.4. Capillary electrophoresis of fluorescently labeled PCR products derived from TPH2 mRNA splice variants……………………………………………………………………………..101
LIST OF TABLES

CHAPTER 2
2.1. Brain sample demographics.................................................................43
2.2. Quantification of G/T ratios depicted in Figure 2-5.................................45
2.3. Mean G/T ratios determined with primer extension assay using genomic DNA /RNA prepared from human pons.................................................................46
2.4. Estimated linkage disequilibrium and estimated individual and pairwise comparisons of these polymorphisms for three SERT polymorphisms.................................................47

CHAPTER 3
3.1. TPH2 SNPs examined in this study......................................................77
3.2. Sequences of primers used for genotyping 20 SNPs listed in Table 3-1..........78
3.3. Measurements of AEI using marker SNPs rs7305115 and rs4290270...............79
3.4. Predicted diplotypes for individuals in sample.......................................80
3.5. Correlation between heterozygosity/homozygosity and presence/absence of AEI using kappa statistics.................................................................82
3.6. Haplotype frequencies for 38 Caucasians in sample..............................83
CHAPTER 1

INTRODUCTION

Serotonin

Serotonin plays a role in a variety of processes in both central nervous system and the periphery, including mood, appetite, cognition, sex, endocrine functions, gastrointestinal functions and vascular functions (Kinney, 2005). In the central nervous system, serotoninergic neurons are primarily located in the raphe nuclei of the brain stem, and project axons to almost all regions of the brain. In addition, serotonin is synthesized by neurons in the pineal. At least 18 subtypes of the 5-HT receptor families are differentially expressed in brain and other tissues (Barnes and Sharp, 1999). Fourteen of these belong to the seven-transmembrane family of G-protein coupled receptors (GPCR) and four (5-HT3A-C,E) are 5-HT-gated ion channels.

The serotonin GPCRs are categorized into four groups according to their second messenger coupling pathways: 5-HT1 receptors, which couple to Gi proteins (5-HT1A, 5-HT1B-C and 5-HT1D-F); 5-HT2 receptors, which couple to Gq proteins (5-HT2A-C); 5-HT4, 5-HT6 and 5-HT7 receptors, which couple to Gs proteins; and the 5-HT5 receptors (5-HT5A and 5-HT5B), which resemble the previous group, but whose transduction cascade is not clear yet (Barnes and Sharp, 1999). Serotonin is also synthesized in peripheral tissues, including enterochromaffin cells in the gut, neuroepithelial bodies in the lung and the parafollicular cells in the thyroid (Russo et al., 1996; Gershon, 2004). Enterochromaffin cells store serotonin in large quantities. Approximately
95% of the body’s serotonin is located in the gut and over 90% of this is stored in enterochromaffin cells located in the enteric epithelium from the stomach through the colon. The actions of serotonin in the gut are mediated primarily through 5-HT<sub>1P</sub>, 5-HT<sub>3</sub> and 5-HT<sub>4</sub> receptors (Gershon, 2004). Agonists and antagonists to different receptor subtypes are used to treat intestinal discomfort and mobility (Tonini and Pace, 2006). It was known that serotonergic signaling and expression of SERT in the mucosa is decreased in inflammation and irritable bowel syndrome (Coates et al., 2004). Serotonin is also found in circulating platelets, which take up serotonin from the serum via SERT (Brenner et al., 2007).

**Roles of serotonin in the neural development**

Serotonergic neurons are generated early in the development of the central nervous system (CNS): from embryonic days 10 to 12 in the mouse and the sixth gestational week in humans (Whitaker-Azmitia, 2001; Gaspar et al., 2003). Serotonin was the first neurotransmitter for which a developmental role was described (Gaspar et al., 2003), and there is emerging evidence for a critical role for serotonin throughout brain development (Lipton and Kater, 1989; Lauder, 1993; Levitt et al., 1997; Azmitia, 2001; Vitalis and Parnavelas, 2003). Even transient alterations in serotonergic homeostasis in the fetal and early postnatal brain have been shown to affect emotional behaviors in adult mice (Gaspar et al., 2003). Perturbation of serotonergic functions in humans, for example by prenatal exposure to selective serotonin reuptake inhibitors (SSRIs), has been linked to preterm delivery, fetal growth restriction and delay in motor development (Chambers et al., 1996; Hendrick et al., 2003; Bairy et al., 2007).

Early pharmacological studies demonstrated that serotonin modulates a number of key events in the developing CNS, including cell division, neuronal migration, cell differentiation and synaptogenesis. In monoamine oxidase A (MAOA) knockout mice, high serotonin levels were linked to the failure to properly form “barrel fields” in the developing somatosensory cortex (Cases et al., 1995). The number of serotonergic neurons in rostral raphe cultures prepared at
E14 was increased when the cells were isolated from 5-HT_{1A} receptor (-/-) or 5-HT_{1B} receptor (-/-) mice compared to wild type mice, but was decreased by 50% in SERT (-/-) mice (Rumajogee et al., 2004). In SERT or MAOA knock out mice, there was abnormal development of the thalamocortical and retinal axons (Upton et al., 1999; Persico et al., 2001; Salichon et al., 2001).

In addition to regulatory roles for serotonin in neuronal proliferation and maturation, there are reports that disruption of the serotonergic system during early development can cause abnormal mood and behavior in adult life (Gingrich and Hen, 2001; Whitaker-Azmitia, 2001; Gross et al., 2002). Pharmacological blockade of SERT with SSRIs during the first two weeks of life resulted in abnormal emotional behaviors in adult mice (Ansorge et al., 2004). A study using genetically engineered mice that allowed conditional, tissue–specific depletion of serotonin 1A receptor, showed that expression of the serotonin 1A receptor in the forebrain in developing mice is both necessary and sufficient to ensure normal anxiety-related behavior later in life. Repression of this receptor during adulthood is ineffective, indicating the essential role of serotonin during development for the establishment of normal anxiety-modulating circuits in the brain (Gross et al., 2002).

**Serotonin and major depression**

Major depression [MD; also termed major depressive disorder (MDD) or unipolar depression (UD)] is a major psychiatric disorder, with a lifetime risk between 15% and 20% for the general population (Doris et al., 1999). A World Health Organization study has stated that MD is the leading global cause of years of life lived with disability and the fourth leading cause disability-adjusted life-years (WHO 2001). Furthermore, 15% of patients with a history of severe depression commit suicide (Greden, 2001).

Stressful life events such as childhood neglect, physical or sexual abuse, and early parental loss have been shown to be major risk factors for the development of MD (Heim and
Nemeroff, 2001; Caspi et al., 2003). Recently, evidence has emerged showing that susceptibility to MD is influenced by a combination of environmental factors and multiple genes and gene variants (Neumeister et al., 2004). This complexity makes it difficult to identify specific genes that contribute to MD. The fact that serotonin plays a key role in modulating the development of neural circuits that regulate emotion and mood, however, suggest that serotonin-related genes may contribute to development of major depression.

There is now a large amount of evidence supporting the idea that serotonin play a key role in the pathogenesis of MD and other affective (mood) disorders. Patients with MD often have altered serotonergic functions as evidenced by low plasma tryptophan levels (Coppen et al., 1972; Cowen et al., 1989), reduced cerebrospinal fluids levels of the serotonin metabolite 5-hydroxyindolacetic (5-HIAA) (Asberg et al., 1976), decreased platelet serotonin uptake (Healy and Leonard, 1987), and decreased responsiveness of brain serotonin receptors (Siever et al., 1984; Cowen and Charig, 1987; Mann et al., 1995; Meltzer and Maes, 1995). Depletion of the serotonin precursor tryptophan has been shown to induce depressive symptoms in healthy controls with high familial risk for depression (Benkelfat et al., 1994; Klaassen et al., 1999) and relapse in individual with a history of MD (Smith et al., 1997). Significantly, many drugs that are effective in treating MD, including the tricyclic antidepressants (TCAs) and SSRIs, block the reuptake of serotonin by SERT (Roman et al., 2003). Based on these observations, SERT has been extensively analyzed in psychiatric, behavioral and pharmacogenetic studies. In particular, a polymorphism in the SERT promoter (SERTLPR; also termed HTTLPR) has been extensively studied as a possible contributing factor to mental disorders, including MD (Neumeister et al., 2004). SERTLPR will be discussed in detail below.

In addition to SERT, additional serotonin-related genes have been linked to MD and other psychiatric disorders. TPH immunoreactivity, for example, has been shown to be elevated in the dorsal raphe of depressed suicide individuals compared to controls, suggesting a possible
compensation for impaired serotonin release (Boldrini et al., 2005). Zill and coworkers reported a positive association between MD and a single TPH2 SNP (the intronic SNP rs1386494) as well as a TPH2 haplotype (Zill et al., 2004b). The same authors also found a positive association between suicide and this TPH2 SNP (Zill et al., 2004a).

Positron emission tomography (PET) imaging studies have also demonstrated reduced pre- and postsynaptic 5-HT$_{1A}$ receptor binding potential in the raphe nuclei and in various cortical areas in un-medicated MD patients compared to healthy controls (Drevets et al., 1999; Sargent et al., 2000). A role for 5-HT$_{1A}$ receptors in anxiety- and depression-like behaviors has been demonstrated using mice that carry a deletion in the 5-HT$_{1A}$ gene, a result obtained with several strains of mice (Overstreet et al., 2003). A recent study using mice genetically engineered to allow tissue-specific and conditional 5-HT$_{1A}$ expression showed that expression of this receptor during the early postnatal period can prevent the anxiety-like behaviors observed in 5-HT$_{1A}$ deletion mice. By contrast, restoring 5-HT$_{1A}$ gene expression in adult knockout mice does not block anxiety-like behaviors (Gross et al., 2002). Taken together, these results imply that 5-HT$_{1A}$ receptor expression during early postnatal development regulates anxiety-like behaviors in the adult.

Recently, a putative promoter polymorphism (-C1019G; rs6295) of the human HT$_{1A}$ receptor gene (Wu and Comings, 1999) has been linked to MD and suicide (Lemonde et al., 2003). This SNP is located within a consensus repressor binding site and the G-allele has been proposed to reduce the affinity of binding of the repressor, leading to derepression of the 5-HT1A receptor gene in serotonergic raphe neurons (Lemonde et al., 2003). The resulting over-expression of inhibitory 5-HT$_{1A}$ autoreceptors is proposed to reduce the release of serotonin from the axon terminals of these neurons, thereby reducing serotonergic neurotransmission (Albert et al., 1996).
Serotonin receptors as targets for antipsychotics and antidepressants

Many 5-HT receptors, including 5-HT$_{1A}$, 5-HT$_{2A}$, 5-HT$_{2C}$, 5-HT$_{3}$, 5-HT$_{6}$, and 5-HT$_{7}$ receptors have been implicated in the action of antipsychotic drugs (Meltzer and Nash, 1991). Atypical antipsychotics, including clozapine, olanzapine, quetiapine, risperidone, sertindole, and ziprasidone, are potent 5-HT$_{2A}$ receptor antagonists and weak dopamine D2 receptor antagonists (Schotte et al., 1996). It has been proposed that the essential characteristic of atypical antipsychotics is a high ratio of 5-HT$_{2A}$ receptor to D2 receptor affinity. Numerous studies have detected decreased numbers of 5-HT2A receptors in specific cortical regions in patients with schizophrenia (Meltzer and Okayli, 1995). Several studies have shown that the nonsynonymous 5-HT$_{2A}$ receptor SNP His452Tyr (rs6314), which is present in 10–12% of the population, is associated with poor response to clozapine (Masellis et al., 1998). Some atypical antipsychotics also bind to 5-HT$_{2C}$, 5-HT$_{6}$, or 5-HT$_{7}$ receptors (Roth et al., 1994). Several atypical antipsychotics, including clozapine, ziprasidone, quetiapine, and tiospirone, are partial agonists at the 5-HT$_{1A}$ receptor (Newman-Tancredi et al., 1998).

Most antidepressants, including the SSRIs, TCAs and monoamine oxidase inhibitors (MAOIs) act by enhancing serotonergic neurotransmission. SERT inhibitors are also prescribed to the treatment of anxiety, obsessive compulsive disorders and bulimia (Charney et al., 1990; Pineyro and Blier, 1999; Coyle and Duman, 2003). Theoretically, blockade of serotonin reuptake occurs within minutes after taking SSRI, suggesting that there should be an immediate increase in serotonergic neurotransmission (Albert and Lemonde, 2004). By contrast, at least 2 to 3 weeks of daily medication are often required to obtain therapeutic effects. The reason for this delay is not yet understood. One possible explanation is that it takes several weeks of elevated serotonin release to stimulate the down-regulation of inhibitory 5-HT$_{1A}$ autoreceptors located on the cell soma of serotonergic neurons in the raphe nuclei (Albert et al., 1996; Pineyro and Blier, 1999).
Down-regulation of these receptors is thought to allow extra-cellular levels of serotonin to reach levels where therapeutic effects are obtained (Stahl, 1998; Blier and de Montigny, 1999; Hjorth et al., 2000).

**Two key serotonin-related genes: The serotonin transporter and tryptophan hydroxylase 2**

**SERT**

The human serotonin transporter gene (*SLC6A4, SERT, 5HTT*) was cloned a decade ago and SERT homologs have been identified in other species such as non-human primates, rodents, worms (*C. elegans*), fruit flies (*Drosophila*) (Androutsellis-Theotokis et al., 2003; Murphy et al., 2004). The SERT has sequence homology (~50%) with the dopamine transporter (DAT) (encoded by *SLC6A3*) and the norepinephrine transporter (NET) (encoded by *SLC6A2*), which also belong to the neurotransmitter:sodium symporter (NSS) family (Saier, 2003).

As mentioned above, SERT is a major target for SSRIs and TCAs, which block serotonin reuptake by SERT. SERT inhibitors are effective in the treatment of depression, anxiety and obsessive-compulsive disorders, and are frequently prescribed for migraine and chronic pain (Murphy et al., 2004). SERT is also a target site for drugs of abuse, including 3,4-methylenedioxy-methamphetamine (MDMA; “ecstasy”) and cocaine (Bengel et al., 1998; Sora et al., 2001). SERT deficient mice show reduced or absent locomotor responses to MDMA and have enhanced reward responses to cocaine (Bengel et al., 1998; Sora et al., 1998).

SERT is expressed in serotonergic neurons in the brain and many peripheral tissues, including enterochromaffin cells in the gut and platelets in the blood. In the brain, SERT is located in the presynaptic membranes of nerve terminals of raphe serotonergic neurons, as well as in the cell soma and dendritic arbors of these neurons. SERT functions to terminate serotonergic neurotransmission by rapidly removing serotonin from the synaptic cleft. This reuptake allows the serotonin to be recycled for subsequent release. Quaternary structure data suggest SERT molecules functions as dimers or tetramers (Kilic and Rudnick, 2000; Schmid et al., 2001).
Human SERT gene (*SLC6A4*) is located on chromosome 17q11.2 and spans 37.8 kb. The gene comprises fourteen exons and encodes a protein composed of 630 amino acids (Ramamoorthy et al., 1993; Lesch et al., 1994). Alternative splicing occurs, with skipping of exon 1B in one isoform (Bradley and Blakely, 1997). Different polyadenylation site usage was also reported (Bradley and Blakely, 1997). There are extensive studies regarding an insertion/deletion polymorphism located in the 5'-flanking region, approximately 1.4kb upstream from the transcriptional start site (Murphy et al., 2004). This polymorphism is termed the SERT linked polymorphic region (*SERTLPR*) or, alternatively, the 5HTT gene-linked polymorphic region (*5HTTLPR*). The “long” *SERTLPR* allele is thought to be a stronger promoter compared to the “short” allele, which lacks 43 bp. The *SERTLPR* is unique to humans, but a similar long/short polymorphism is present in higher primates, including the rhesus monkey, which has a 21bp insertion/deletion (*rh5HTTLPR*). The *SERTLPR* long/short polymorphism is absent in rodents (Lesch et al., 1997).

*SERTLPR* alleles generally consist of either fourteen (S-allele) or sixteen (L-allele) repeated elements (each repeat has 20 to 23 bp), although alleles with more repeat elements (“super-long”, XL and XXL, with up to twenty repeats) are occasionally observed (Heils et al., 1995; Lesch and Mossner, 1998).

The uncommon, super-long alleles are generated from duplication of an internal segment of *SERTLPR* and are predominantly observed in individuals of African descent (Delbruck et al., 1997; Delbruck et al., 2001). Genotype frequencies for *SERTLPR* alleles vary among different populations. North American/European population have allele frequencies of 57% for the L-allele and 43% for the S- allele with a genotype distribution of 32% LL, 49% LS, and 19% SS (Lesch et al., 1996). A two-fold higher occurrence of the SS genotype is found in Asians (Delbruck et al., 1997; Gelernter et al., 1997).
It has been reported that S-allele is associated with lower SERT mRNA expression in lymphocytes and weaker promoter activity in reporter gene assays (Lesch et al., 1996).

Another well-studied SERT polymorphism is a VNTR (variable number of tandem repeats) located in intron 2 (Collier et al., 1996). Variants containing nine-, ten-, or twelve-copies of a sixteen- or seventeen–base pair repeat are most common. Previous studies have associated the twelve-copy allele with higher expression for SERT mRNA in the hindbrain of transgenic embryonic mice compared to the ten-copy allele (MacKenzie and Quinn, 1999), and greater enhancer activity in embryonic stem cells (Fiskerstrand et al., 1999). Non-synonymous SNPs are uncommon in SERT gene. However, two rare SNPs, I425V and L255M located within highly conserved transmembrane regions were found to associate with complex serotonergic dysfunction-related phenotypes, including obsessive–compulsive disorder in the case of I425V (Ozaki et al., 2003) or severe depression in the case of L255M (Di Bella et al., 1996).

Additional variations have been detected in the 3’ untranslated region (UTR), including a G-to-T polymorphism in a putative adenylation site in the 3’ UTR region. However, this SNP does not affect polyadenylation site usage (Battersby et al., 1999).

**TPH2**

Tryptophan hydroxylase (TPH) is the rate-limiting enzyme in serotonin synthesis (Fitzpatrick, 1999b). It belongs to the superfamily of aromatic amino acid hydroxylase, which include tyrosine hydroxylase (TH1), phenylalanine hydroxylase (PAH), and tryptophan hydroxylase(TPH1). TPH1 was identified four decades ago and is well-characterized (Grahame-Smith, 1964). It was originally thought to be the only isoform in the body. Studies using TPH1 knockout mice, however, revealed the existence of a second isoform, THP2, which is expressed primarily in the serotonergic neurons of the raphe (Walther et al., 2003b) and in myenteric neurons in the gut (Cote et al., 2003). By contrast, TPH1 is expressed primarily in the periphery, including enterochromaffin cells in the gut. TPH1 is also expressed at high levels in the pineal
gland (Patel et al., 2004) and possibly at very low levels more widely in the brain. The *TPH1* gene is located on chromosome 11 and the *TPH2* gene on chromosome 12 (Walther et al., 2003b). The two isoforms share 71% identity in amino acid sequence, with conservation of key functional sequences (Walther and Bader, 2003). *TPH2* has been identified in other species such as mice, rats, chickens and zebrafish (Walther et al., 2003b; Teraoka et al., 2004; Coleman and Neckameyer, 2005).

Two polymorphisms *TPH1*, A218C and A779C, have been extensively studied for possible associations with neuropsychiatric disorders such as suicidal behavior, unipolar depression, bipolar depression and schizophrenia, with conflicting results (Arango et al., 2003). More recently, focus has shifted to the major brain isoform, *TPH2*. The first wave of papers have demonstrated possible associations between several *TPH2* polymorphisms and MD (Peters et al., 2004; Zill et al., 2004b; Zhang et al., 2005a; Zhou et al., 2005a), suicidal behavior (Zill et al., 2004a; Zhou et al., 2005a), attention deficit hyperactivity disorder (ADHD) (Sheehan et al., 2005; Walitza et al., 2005), autism (Coon et al., 2005), bipolar disorder (Harvey et al., 2004) and obsessive-compulsive disorder (OCD) (Mossner et al., 2006a).

Six rare, non-synonymous SNPs, L36P (rs34115267), P206S (rs17110563), A328V (rs2887147), R441H, D479E (rs7488262) in human *TPH2* and P447R in mouse *TPH2* and 3 coding synonymous SNPs, P312P (rs7305115), L327L (rs2887148), A375A (rs4290270) have been identified in human *TPH2* gene (Breidenthal et al., 2004b; Zhang et al., 2004; Zhang et al., 2005a). Significant association was observed between a TPH2 promoter region SNP (rs11178998) and unipolar depression in a northern Swedish population and protective haplotype was described (Van Den Bogaert et al., 2006). A functional magnetic resonance imaging (fMRI) study examining the effects of the promoter region SNP G(-844)T (rs4570625) on amygdala activity associated with emotion revealed a greater activity in the bilateral dorsal amygdala of T-allele carriers compared to G-homozygotes (Brown et al., 2005). The A-allele of a non-synonymous
TPH2 SNP, (G1463A; R441H), was reported to cause an approximately 80% of reduction in serotonin production in the cell culture system and to occur at relatively high frequency in geriatric patients with MD compared to control population (Zhang et al., 2005a). Several subsequent large-scale studies, however, failed to detect this SNP in the general population (Garriock et al., 2005; Van Den Bogaert et al., 2005).

Two intronic TPH2 SNPs, rs4341581 and rs11179000, were observed to have different frequencies between autistic and control samples, but failed to predict clinical phenotypes. A large scale association study reported association between TPH2 and suicide attempt, major depression, and cerebrospinal fluid 5-HIAA in 1798 cases representing four populations and controls. They identified a “yin”-haplotype that was more frequent in suicide attempters in both populations (Finnish whites and African Americans), and was linked to MD and anxiety disorders in US whites and MD in African Americans. This yin-haplotype was somewhat indicative of lower CSF 5-HIAA concentrations in controls, but not in cases (Zhou et al., 2005a). In contrast, other studies have reported a lack of association between TPH2 polymorphisms and suicidal behavior, schizophrenia, bipolar disorder or depression (De Luca et al., 2004; De Luca et al., 2005; Garriock et al., 2005). Discrepancies may be resulted from the different populations, diagnostic criteria, and the SNPs examined. In many cases the association were detected for SNPs located in introns or promoter regions where the functional properties of the SNP are difficult to assess. These results underline the importance of identifying functional polymorphisms for association studies.

**Cis acting genetic variants**

A common molecular mechanism for controlling gene expression involves the binding of transcription factors to regulatory sequences located close to their target genes. Genes encoding transcription factors, however, are usually located at sites for away from their target genes, often
on different chromosomes. Because the transcription factor genes are not physically linked to
their target gene, they are termed “trans”-acting factors. In addition, non-genetic factors such as
physiological or environmental influences are considered to be “trans”-acting factors. By
contrast, the regulatory sequences tightly linked to the gene are termed “cis”-acting factors, and
genetic variants that are located within or in close vicinity of a gene are termed “cis”-acting
variants.

The expression of each allele of an autosomal gene is influenced both by cis-acting and
trans-acting genetic variants (Singer-Sam et al., 1992; Enard et al., 2002; Whitney et al., 2003;
Pastinen and Hudson, 2004). Cis-acting variants, which include SNPs, multinucleotide
polymorphisms (MNPs), insertions and deletions (indels) and inversions can be located 5’-
(upstream) or 3’- (downstream) of the gene, as well as within the exons and introns of the gene
(Enattah et al., 2002). Cis-acting variants within promoters or enhancers influence gene
expression by changing the binding affinity of transcription factors (Day and Tuite, 1998). Cis-
acting variants within transcribed regions of the gene influence gene expression by affecting the
efficiency of mRNA splicing, transport or stability (Sheets et al., 1990; Day and Tuite, 1998;
Conne et al., 2000; Cartegni et al., 2002; Tebo et al., 2003). Cis-acting variants may also
influence the binding of regulatory non-coding RNAs to mRNAs (Novina and Sharp, 2004).

Cis- and trans-acting variants that affect mRNA expression may lead to phenotypic
differences that are expressed at the level of individuals, such as susceptibility to disease or
response to drugs (Lerer and Macciardi, 2002; Pinsonneault and Sadee, 2003). Genetic
association studies and linkage studies have identified genetic variants or sets of genetic markers
(“haplotypes”) that co-segregate with specific mental illnesses (Freedman and Leonard, 2001;
Freedman et al., 2005; 2007; Marui et al., 2007; Teltsh et al., 2007). Many of these studies,
however, have failed to undergo replication. Reasons for lack of replication include false-positive
associations due to multiple testing or population stratification and ambiguities in definitions of
phenotype (Zhao et al., 2003). An absence of knowledge concerning the properties of genetic variants used in the association study also contributes the inability to distinguish true functional variants from statistical flukes.

Prescreening genetic variants to identify those that influence gene expression should greatly facilitate the identification of genes that play a role in mental illness. In particular, the measurement of mRNA expression ratios for genetic alleles within single RNA samples is a powerful method for identifying cis-acting genetic variants that cause (or accurately predict) high- or low-levels of mRNA expression (Yan et al., 2002a; Bray et al., 2003b). Alleles identified by this method can be directly tested for association with disease, eliminating the need to test large numbers of uncharacterized genetic markers.

As described above, serotonin is a critical regulatory molecule that influences diverse physiological processes. We hypothesize that genetic variants that influence the synthesis and activities of serotonin play a role in diseases. The goal of the research described in this proposal is to identify functional genetic polymorphisms that influence the expression of two genes in the brain that encode proteins related to the synthesis and activity of serotonin: 1) SERT and 2) TPH subtype 2 (TPH2). By identifying functional genetic variants that influence the expression of these genes, I hope to lay the foundation for a better understanding of the role of serotonergic system in MD and other psychiatric disorders.
CHAPTER 2

ALLELIC EXPRESSION OF SEROTONIN TRANSPORTER (SERT) mRNA IN HUMAN PONS: LACK OF CORRELATION WITH THE POLYMORPHISM SERTLPR

INTRODUCTION

The neurotransmitter serotonin is an important modulator of physiology, behavior and psychological states (Lucki, 1998b; Ursin, 2002; Gingrich et al., 2003; Swann, 2003). Disturbances in serotonergic systems in the brain have been implicated in mental illness, including schizophrenia (Lieberman et al., 1998; Roth et al., 2004), anxiety disorders (Lesch et al., 2003), obsessive-compulsive disorders (Kaye et al., 2005), addiction (Olausson et al., 2002; Higgins and Fletcher, 2003; Oroszi and Goldman, 2004; Kreek et al., 2005), depression (Wong and Licinio, 2001; Arango et al., 2002; Nemeroff and Owens, 2002; Neumeister et al., 2004) and suicide (Arango et al., 2002; Joiner et al., 2005).

After release from axonal terminals of serotonergic neurons, synaptic effects of serotonin are terminated by reuptake into the nerve endings (Rudnick and Clark, 1993). The serotonin transporter mediating this reuptake, SERT, belongs to the Na⁺ and Cl⁻-dependent family of neurotransporters (Norregaard and Gether, 2001). SERT mRNA is expressed mainly in serotonergic neurons in the raphe nuclei of the pons and upper brainstem, which project widely to various regions of the brain (Hornung, 2003). Tricyclic antidepressants (TCAs) and serotonin selective reuptake inhibitors (SSRIs) used in the treatment of anxiety disorders and depression directly inhibit SERT (Roman et al., 2003).
The human gene encoding SERT, \textit{SLC6A4}, is located on chromosome 17q11.2 (Ramamoorthy et al., 1993). The gene is organized in 15 exons spanning \~39 kb, which encode a protein containing 630 amino acids (Ramamoorthy et al., 1993; Lesch et al., 1994). Transcriptional activity is regulated by several positive and negative regulatory elements within the \textit{SERT} promoter region (Lesch et al., 1994; Heils et al., 1996; Lesch et al., 1996; Sakai et al., 2002) and by a 17 bp variable number tandem repeat (\textit{VNTR}) element in intron 2 (MacKenzie and Quinn, 1999). Differential splicing (Bradley and Blakely, 1997; Ozsarac et al., 2002) and use of two 3’ polyadenylation sites (Battersby et al., 1999) may also contribute to the regulation of \textit{SERT} expression (Murphy et al., 2004).

Heils \textit{et al.} (Heils et al., 1995; Heils et al., 1996) first described functional variants of a repetitive sequence in the \textit{SERT} linked polymorphic region (\textit{SERTLPR}) located 1.2 kb upstream of the transcription start site. The \textit{long} (\textit{l}) variant contains 16 copies of a 20 to 23 bp GC-rich sequence and the \textit{short} (\textit{s}) variant contains 14 copies of this sequence. Transfection studies with expression vectors containing the \textit{l} or \textit{s} variants linked to a reporter gene indicated that the \textit{l} variant directs higher levels of transcription compared to the \textit{s} variant when expressed in lymphoblast cell lines (Lesch et al., 1996), a human placental choriocarcinoma cell line (JAR) (Heils et al., 1996) and in the raphe nucleus-derived cell line RN46 (Mortensen et al., 1999). In agreement with the reporter gene assays, higher \textit{SERT} mRNA levels and higher rates of serotonin uptake were measured in lymphoblasts of \textit{l/l} homozygotes compared to those containing at least one copy of the \textit{s} allele (Lesch et al., 1996). The \textit{l} variant was also associated with higher rates (\(V_{\text{max}}\)) of serotonin uptake (Greenberg et al., 1999; Nobile et al., 1999) and higher levels of serotonin binding (Stoltenberg et al., 2002) in platelets and higher serotonin levels in the blood (Hanna et al., 1998). In postmortem human midbrain tissue sections containing the dorsal and median raphe nuclei, \textit{SERT} mRNA levels were reported to be higher in subjects carrying two copies of the \textit{l} allele compared to individuals with at least one \textit{s} allele (Little et al., 1998).
Moreover, SERT ligand binding levels in the raphe region were higher in l/l individuals than in s allele carriers, with the exception of l/l alcoholics who had lower SERT ligand binding levels than l/s or s/s alcoholics (Heinz et al., 2000).

Not all studies, however, have reported a correlation between SERTLPR alleles and SERT expression. Analysis of SERT mRNA levels in 53 permanent lymphoblast cell lines with real-time PCR (using β-actin mRNA as a reference) failed to find a statistically significant correlation with SERTLPR, but did find evidence for a combined effect of SERTLPR and the intron 2 VNTR (Hranilovic et al., 2004). Also, no correlation between SERTLPR alleles and promoter activity was observed with expression vectors assayed in COS-7 (Mortensen et al., 1999; Sakai et al., 2002) or PC12 cells (Sakai et al., 2002). There were no statistically significant correlations between SERTLPR alleles and binding of the SERT ligand [123I]CIT in the thalamus-hypothalamus or mesencephalon-pons measured in SPECT scans of healthy volunteers (Willeit et al., 2001). In addition, no correlations were found between SERTLPR and binding of the SERT ligand [11C]McN5652 in healthy male subjects examined by PET (Shioe et al., 2003).

Surprisingly, most studies on SERTLPR-mediated regulation have focused on lymphocytes or heterologous cell systems, instead of the physiologically relevant cell populations, such as the serotonergic neurons in the raphe nuclei of the pons and upper brainstem. The functional effects of SERTLPR on mRNA expression in the brain therefore remain uncertain. In this study, we tested whether the SERTLPR influences mRNA expression in human pons tissue sections containing serotonergic neurons of the dorsal and median raphe nuclei, which are primary sites of SERT transcription in the CNS (McLaughlin et al., 1996).

Genetic variants that differentially affect mRNA expression from a given allele cause allelic expression imbalance (AEI), which can be quantitatively measured in individuals who are
heterozygous for a marker SNP within the mRNA (Yan et al., 2002a). The existence of AEI indicates the presence of cis-acting factors that influence intracellular levels of mRNA by modifying transcription, mRNA processing, or epigenetic effects (Yan et al., 2002a; Pastinen and Hudson, 2004; Pastinen et al., 2004; Yan and Zhou, 2004; Johnson et al., 2005). Genetic analyses suggest that cis-acting genetic variants are one of the main sources of variability in human phenotypes, including susceptibility to disease (Rockman and Wray, 2002; Bray et al., 2003b; Pastinen and Hudson, 2004; Yan and Zhou, 2004). Although a powerful approach, analysis of AEI has not been widely used, because of difficulties in obtaining reproducible measurements. We have developed a robust procedure for allele-specific measurement of mRNA expression using a fluorescence-based primer extension assay to quantify real-time PCR amplification (Pinsonneault and Sadee, 2003; Zhang et al., 2005c). We have successfully applied this approach to identifying functional cis-acting polymorphisms in MDR1 (Wang et al., 2005a) and OPRM (Zhang et al., 2005c). In the present study, we used this method to measure AEI for SERT mRNA in 48 postmortem pons tissues sections from individuals heterozygous for a marker SNP (rs1042173) in the 3’UTR of SERT mRNA. Although we have detected a limited degree of allelic variations in mRNA expression, this did not correlate with the SERTLPR promoter polymorphism. Moreover, SERTLPR also did not account for AEI of SERT mRNA in immortalized B-lymphocytes.

**MATERIALS AND METHODS**

**Materials** In total, 48 human pons sections from adult subjects and ten from fetal/neonatal were obtained from the Brain and Tissue Bank for Developmental Disorders (University of Maryland, Baltimore). The demographics of these samples are described in Table 2.1. PCR reagents were from Promega (Madison, WI). Oligonucleotide primers were designed using OLIGO 4.0
and synthesized by Integrated DNA Technologies (Coralville, IA). SNaPshot reagents were from Applied Biosystems (Foster city, CA). Other reagents for cell culture were from Sigma (St. Louis, MO) or Fisher Scientific (Pittsburgh, PA).

**Isolation of DNA and RNA from human pons** Individual sections of human pons were soaked overnight at -80°C in 10 volumes of RNA later-ICE Frozen Tissue Transition Solution (Ambion, Austin, TX). Small pieces of tissue (~500 mg) were removed for isolation of DNA and homogenized in 365 µl of nuclei lysis buffer containing 10 mM TRIS base, 400 mM NaCl, 2 mM Na$_2$EDTA and 0.7% SDS. To digest protein contaminants, 35 µl proteinase K (10 mg/ml, Invitrogen, Carlsbad, CA) were added and the mix incubated overnight at 55 ºC. The next day, digested proteins were precipitated with saturated NaCl (6M), the supernatant fractions were transferred to fresh tubes and the DNA precipitated with ethanol. The remaining tissue was homogenized in Trizol reagent (1 ml Trizol reagent per 100 mg) for isolation of RNA. RNA samples were treated with 10 µl of RNase-Free DNase (2500 Kunitz units/mg, Qiagen, Valencia, CA) for 15 min at room temperature, and total cellular RNA was purified using QIAGEN RNeasy columns. Complementary DNA (cDNA) was generated from 1 µg total RNA in a total volume of 20 µl with 1 µl (200 U) Superscript II reverse transcriptase (Invitrogen, Carlsbad, CA), 1 µl of 1 µM oligo(dT)$_{20}$ primers (Invitrogen, Carlsbad, CA), 1 µl of 10 mM dNTP mix (Invitrogen, Carlsbad, CA), 0.5 µl of 1 µl M SERT gene-specific primer (5’-TGGACACACATATT TAG-3’) to facilitate reverse transcription of the target mRNA, 4 µl of 5X first-strand buffer, 1 µl of RNaseOUT (40 units/µl), and RNase-free water to yield 20 µl. Typically, three independent cDNA preparations were made from each RNA sample. One of the 48 pons tissue samples failed to yield PCR products possibly due to degradation of genomic DNA and mRNA during the postmortem interval or during processing of the sample.

**Cell culture** Epstein-Barr virus transformed lymphoblastoid cell lines were obtained from the
Coriell Cell Repositories (Coriell Institute for Medical Research, Camden, NJ). Cells were cultured at 37 ºC in a humidified incubator at 5 % CO₂ in RPMI medium containing 2mM L-glutamine plus 15 % fetal bovine serum, 100 U/ml penicillin and 100 µg/ml streptomycin. When the growth medium turned yellow, the cells were supplied with fresh medium, doubling the volume. Fifty ml of cell culture typically yielded 50 to 100 million cells. Medium containing cells (20 ml) was centrifuged at 1000 x g, lysed in Sucrose-Triton solution prior to treatment with proteinase K (10 mg/ml), and incubated for 4 to 5 hours at 55ºC. 6M NaCl was added (125 µl) to precipitate the digested proteins. The supernatant fractions were transferred to new tubes and the DNA precipitated by addition of 1 ml of ethanol. RNA was isolated in Trizol reagent, followed by RNA purification with QIAGEN RNeasy mini prep kits. DNA in RNA samples was eliminated by treating with a RNase-Free DNase Set (2500K units/mg, Qiagen, Valencia, CA).

**Genotyping** A marker SNP (G2651T; rs1042173) located in the 3’-untranslated region (UTR) of human *SERT* (Figure. 2-1.A) was genotyped for 48 pons and 12 lymphoblast genomic DNA samples using the SNaPshot assay. The l/s SERTLPR promoter polymorphism was genotyped by PCR amplification of the region using flanking oligonucleotide primers: 5’-ATGCCAGCAC-CTAACCCCTAATGT-3’ (forward primer) and 5’-GGACCGCAAGGTGGGCGGGA-3’ (reverse primer). PCR amplification reactions were carried out using the GC-rich PCR System (Roche Applied Science, Indianapolis, IN) with the following cycles: [1x (30s at 95 ºC); 35x (30s at 62 ºC, 45s at 72 ºC)]. For genotyping of the VNTR, we PCR-amplified the genomic DNA segment containing the VNTR using flanking oligonucleotide DNA primers. A modified forward primer containing fluorescent dye (6-FAM) covalently attached to its 5’-end was used in this reaction. The fluorescent PCR products were resolved by electrophoresis on an ABI3730 DNA sequence analyzer to determine the number of repeats in the VNTR. Amplification reaction for the 17-bp VNTR was performed in a total volume of 25 µl containing 25ng DNA, 200 µM dNTP mix,
200nM of each of the primers, 10X buffer and Taq DNA polymerase. Applied primer sequences were 5’-GTCAGTATCACAGGCTGCGA 3’ (forward) and 5’-TGTTCTAGTCTTACGCCA GTG-3’ (reverse). The PCR program consisted of an initial denaturation at 94 °C for 2 min, 35 cycles of 20s at 95 °C, 30s at 62 °C, 1 min at 72 °C, and a final extension for 10 min at 72 °C. For genotyping of polyA SNP, we performed allele-specific PCR amplification using an ABI 7000 DNA sequence detection system (Applied Biosystems, Foster City, CA). Briefly, two reverse primers were designed, with the 3’ base of each primer matching only one of the biallelic SNP bases to be evaluated. To increase the specificity of the amplification from each allele, different mismatches were introduced into the third nucleotide from the 3’ end of each reverse primer (Okimoto and Dodgson, 1996). A common forward primer was designed to anneal upstream of the polymorphic site. Independent PCR amplification reactions were carried out with each allele-specific primer. The sequences of the primers were: 5’-CAAATATATGAATTCCCCAAATTTTTTC-3’ (forward primer), 5’-CAC-AATTGAGTTGGTAGAATTTGTGAA-3’ (allele-specific reverse primer 1) and 5’-ACACAA-TTGAGTTGGTAGAATTTGTGAAC-3’ (allele-specific reverse primer 2). 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 sec and primer annealing and extension for 1 min at 60°C.

**Allelic expression imbalance (AEI) assays**  A detailed protocol for this method has been previously described.(Pinsonneault et al., 2004a) Briefly, a segment of genomic DNA or cDNA (205 bp) spanning the marker SNP was amplified by PCR [1x (30s at 95 °C); 35x (30s at 56 °C, 45s at 72 °C)] using the following oligonucleotide primers: 5’-TATCTGTTTGCTTA- AAGGTTTC-3’ (forward primer); 5’-TGAGCAGACTATTTTCTATTAG-3’ (reverse primer). Unincorporated dNTPs and excess primers were inactivated with 2 units of exonuclease I (New England Biolabs, Beverly, MA) and 5 units of antarctic alkaline phosphatase (New England
Biolabs, Beverly, MA). The PCR products were then used as templates in SNaPshot (Applied Biosystems, Foster City, CA) primer extension assays. Primer extension reactions were carried out independently in both directions (for verification) using primers that terminate immediately adjacent to the marker polymorphism (“forward” primer: 5’-GCCATATATT-TCTGAGTAGCATATA-3’; “reverse” primer: 5’-GGTTCTAGTAGATCCAGCAATAAAATT-3’). Following incorporation of a single fluorophore-labeled dideoxynucleoside triphosphate (ddNTP) complementary to the nucleoside at the polymorphic site, the resulting primer extension products were resolved by capillary electrophoresis (ABI 3730 DNA Sequencer, Applied Biosystems) and analyzed using Gene Mapper 3.0 software (Applied Biosystems). To measure the expression ratios of the two alleles, we used DNA and cDNA from individuals who are heterozygous for the marker SNP. Incorporation of different fluorescently labeled dideoxynucleotides into the primers produced oligonucleotides with similar electrophoretic mobilities but distinct fluorescence spectra. Standard curves using peak areas and different ratios of allelic DNA were linear. Because 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 using a correction factor based on the mean of the genomic DNA ratios, and the cDNA ratios of heterozygous samples adjusted accordingly. SNaPshot assays were performed 3x with genomic DNA and 3x with three independent cDNA preparations per sample.

Statistical analysis Data are expressed as mean ± standard deviation (SD). Statistical analyses were performed using Excel (Microsoft, Inc.) for ANOVA, t-tests (two-tailed) and linear regressions. Prism (GraphPad, San Diego, CA) was used for evaluating differences in relative SERT mRNA expression levels for l/l, l/s and s/s genotypes. We assessed allelic expression imbalance (AEI) in individual samples by comparing the mean allelic ratio for RNA with the
mean allelic ratio for DNA. The mean allelic ratio for RNA was considered to significantly differ from that of DNA if 1) the means differed by more than 10%, and 2) there was no overlap between 2-times the standard error for each mean. Calculations of pairwise linkage disequilibrium (D’) for three SERT polymorphisms and the contributions of these polymorphisms to SERT mRNA allelic expression imbalance (AEI) were carried out using programs Linkage Disequilibrium View and Two Genetic Loci Plot programs in HelixTree® (GoldenHelix Inc., Bozeman, MT).

RESULTS

Genotyping

The locations of the four SERT polymorphisms examined in this study are shown in Figure. 2-1.A. We genotyped the marker SNP in the 3’ UTR (G2651T) in 47 human pons samples using the SNaPshot primer extension assay. Fourteen individuals were T/T homozygotes, 4 were G/G homozygotes, and 29 were G/T heterozygotes. The l/s SERTLPR polymorphism was genotyped in the 29 G/T heterozygous individuals using agarose gel electrophoresis to distinguish the PCR products derived from the l (419 bp) and s variants (375 bp) (Figure. 2-1.B). Twelve individuals were l/l, 15 were l/s and 2 were s/s. We genotyped the intron 2 VNTR in the 29 G/T heterozygotes using capillary gel electrophoresis to distinguish the PCR products derived from the 10-repeat (265 bp) and 12-repeat variants (299 bp) (Figure. 2-1.C). Three individuals were homozygous for 10 repeats (10/10), 9 were homozygous for 12 repeats (12/12), and 17 were heterozygous for 10 and 12 repeats (10/12). We also genotyped the 3’ UTR SNP (A2858C; rs3813034) located 207 bp downstream from the marker SNP within the second SERT mRNA polyadenylation site using the SNaPshot primer extension assay.
Twenty-five individuals heterozygous for the marker SNP were also heterozygous (A/C) for rs3813034, 3 were homozygous C/C and 1 was homozygous A/A.

**Analysis of mRNA expression ratios in human pons**

To determine if there were sufficient numbers of SERT mRNA molecules to accurately measure allelic expression ratios, we used quantitative real-time PCR to determine the number of SERT mRNA molecules in 29 independent preparations of pons RNA. As shown in Figure 2-2, genomic DNA was used to standardize our quantitative real-time PCR measurements. Based on this standard curve, we determined that our human pons RNA samples contained approximately 1000 to 64,000 molecules of SERT mRNA (Figure 2-3), numbers sufficiently large for reliable measurements of allelic expression ratios. We also used quantitative real-time PCR to show that SERT mRNA levels are significantly higher in pons compared to the cerebellum or cortex (regions that contain fewer serotonergic neurons), suggesting that the SERT mRNA in the pons derives from serotonergic neurons in the raphe nuclei (Figure 2-4).

We next carried out a small-scale experiment to determine if allelic expression ratios could be accurately measured using human pons RNA. The four fluorescent (dR6G-, dTAMRA-, dR110- and dROX-) dideoxynucleotide derivatives used in the primer extension assays have different fluorescence yields, incorporation efficiencies, and electrophoretic mobilities. To determine if different combinations of fluorophores differentially affect the measured expression ratios, we carried out independent allelic expression ratio determinations with RNA or DNA isolated from six individuals heterozygous for the mRNA marker SNP using extension primers that anneal to the anti-sense or sense DNA strand, respectively. The “forward” extension primer reaction produced G/T ratios derived from the fluorescence peak areas of dR110 and dROX (Figure 2-5.A, top; 2-5.B, left). The “reverse” primer extension reaction produced C/A ratios derived from dTAMRA and dR6G fluorescence peaks (Figure 2-5.B, bottom; 2-5.B, right).
Allelic ratios of PCR products derived from genomic DNA using the forward and reverse extension primers yielded an average ratio of 1.15 ± 0.08 SD; reverse primer yielded an average ratio of (0.432 ± 0.021 SD). No sample yielded genomic allele ratios that significantly deviated from the mean, suggesting an absence of chromosomal instability-related gene dosage effects. The small standard deviations (SDs) of these measurements indicate that allelic ratios can be determined with good precision. Since each allele is present in equal amounts in genomic DNA, deviations from the expected allelic ratio of 1.0 likely arise from the measurement process itself, e.g. from differences in the efficiency of incorporation between the fluorescently-tagged dideoxynucleotide triphosphates used in the primer extension assay. To normalize the DNA-based allelic ratios, a correction factor was calculated as the inverse of the mean of all DNA allelic ratio measurements \([n = 18; F = 1/(mean \text{ allelic ratio for DNA})]\). Correction factors were determined independently for the forward and backward primers. Each measured DNA allelic ratio was multiplied by this factor, and the mean of these values (± standard deviation) tabulated. (Table 2-2.A and B, left columns) \textit{SERT} mRNA expression ratios were also measured using RNA samples isolated from the pons of six individuals heterozygous for the 3’UTR marker SNP using both forward and reverse extension primers. The measured ratios were multiplied by the correction factors determined from the genomic DNA, and the average of these values (± standard deviation) tabulated (Table 2-2.A and B, right columns). As indicated, allelic mRNA ratios determined using the forward and reverse primers were not significantly different. Moreover, none of the measured allelic mRNA ratios deviated significantly from unity, indicating an absence of AEI in these samples, even though the samples included \(l/s\) heterozygotes where one would have expected AEI.

The above experiment demonstrates that we can reproducibly measure allelic ratios for \textit{SERT} mRNA using total RNA isolated from sections of human pons. To look for evidence of AEI of \textit{SERT} mRNA in a larger set of samples, we used RNA isolated from 29 individuals
heterozygous for the 3'UTR marker SNP, including the 6 individuals analyzed above (Figure 2-6 and Table 2-3). After correction of the ratios as explained above, we found small differences in allelic mRNA expression in 9 of the 29 samples [#1054 (l/l), 1103 (l/l), 1112 (l/l), 1209 (l/l), 1279 (l/l), 914 (l/s), 1027 (l/s), 1429 (l/s), 1500 (l/s); range of AEI = 1.12 to 1.27]. The samples that showed AEI for SERT mRNA, however, were not all heterozygous at the SERTLPR locus (i.e., l/s genotype), as would be expected if the l and s forms of the promoter were the dominant regulatory elements for this gene. Also, there were no significant differences in mRNA expression between G and T alleles in 11 out of 15 samples of individuals who are also heterozygous for the l/s promoter polymorphism. Finally, an ANOVA revealed no significant differences in SERT AEI among the l/l, l/s and s/s genotypes (Table 2-3; p = 0.381). Taken together, these results indicate that the SERTLPR genotype by itself does not predict measured SERT mRNA levels in these pons tissue sections.

AEI in 4 samples heterozygous for l/s suggests that there may be functional cis-acting polymorphisms in partial linkage disequilibrium with SERTLPR. Similarly, 5 out of 12 samples showed AEI for individuals homozygous for the l form, suggesting that the l allele associates with a functional polymorphism with a frequency of 20-25%. Since only two homozygous s carriers were among the study population, with no imbalance observed, we cannot ascertain whether this polymorphism is also associated with some s alleles.

**Allele-specific mRNA expression in lymphoblasts**

Our data using brain RNA differs from the results of a previous study of SERT expression in Epstein Barr virus-transformed B-lymphoblasts, where the l promoter variant was found to produce 1.7-fold higher levels of endogenous SERT mRNA compared to the s-promoter variant (Lesch et al., 1996). To determine whether there are differences between SERT expression in pons and lymphocytes, we measured allelic expression ratios for SERT mRNA using RNA isolated from Epstein Barr virus-transformed B-lymphoblasts prepared from 9 individuals heterozygous...
for the 3’-UTR marker SNP. Genomic DNA from these cells was used to genotype SERTLPR. As for pons RNA samples, quantitative real-time PCR analysis demonstrated that there were sufficient numbers of SERT mRNA in the lymphoblast RNA samples to accurately determine allelic expression ratios (Figure 2-7). Consistent with results from brain samples, we found small, statistically significant differences in allelic mRNA expression in 5 out of 9 samples (Figure 2-8). Unlike the study of Lesch and coworkers (Lesch et al., 1996), however, these differences did not strictly correlate with l/l, l/s or s/s genotypes. Interestingly, one s/s homozygous sample (6991) showed the largest AEI observed in this study, indicating that a cis-acting polymorphism that regulates SERT expression could be in partial linkage disequilibrium with the s allele.

**Estimated linkage between SERT polymorphisms and contributions of these polymorphisms to SERT AEI**

Because the SERT polymorphisms are distributed within a 39 kb segment of the chromosome, it is experimentally difficult to determine which alleles are physically linked in any given individual. To determine the probability of physical linkage between alleles we calculated estimated linkage disequilibrium for pairs of polymorphisms examined in this study. These results are summarized in Table 2-4.A. Consistent with previous estimates (Conroy et al., 2004), the data show that there is relatively low linkage disequilibrium between the SERTLPR and the three other markers (the VNTR, the marker SNP and polyA SNP). By contrast, the VNTR and 3’ UTR SNPs show a high degree of linkage disequilibrium. These results suggest that the SERTLPR and the three other polymorphisms are located on different haplotype blocks, a conclusion consistent with previous results (Conroy et al., 2004) and with the haplotype structure predicted for SERT from analysis of HapMAP(2003) SNPs (http://www.hapmap.org/index.en).
As mentioned above, a recent study by Hranilovic et al (Hranilovic et al., 2004) reported a correlation between SERT mRNA expression in lymphocytes with the combined SERTLPR and VNTR markers. To determine if the VNTR contributes to the small AEI observed in pons sections, we examined the association of SERT mRNA AEI with VNTR alone and in combination with SERTLPR or the marker SNP. As shown in Table 2-4.B, no individual marker showed a statistically significant association with SERT AEI. By contrast, SERTLPR and the VNTR in combination showed a weak, but statistically significant (in the absence of corrections for multiple testing) association with SERT AEI. Even in this case, however, the extent of SERT AEI was at most 1.27-fold (range 1.12 to 1.27), considerably less than the 1.7-fold difference in SERT mRNA expression previously reported for the l and s alleles in lymphocytes.

**DISCUSSION**

Previous studies have investigated the influence of l and s SERTLPR alleles on SERT promoter activity in heterologous expression systems (Heils et al., 1996; Lesch et al., 1996; Mortensen et al., 1999; Sakai et al., 2002), SERT mRNA levels in lymphoblast cell lines (Lesch et al., 1996; Hranilovic et al., 2004) and brain (Little et al., 1998), SERT ligand binding (Stoltenberg et al., 2002) and transport activity in platelets (Greenberg et al., 1999), and SERT ligand binding levels in postmortem brain sections (Little et al., 1998) and living brain (Naylor et al., 1998; Heinz et al., 2000; Willeit et al., 2001; Shioe et al., 2003). Clinical association studies have further examined the relationship between SERTLPR alleles and psychiatric disorders and anxiety-related personality traits (Anguelova et al., 2003a; Lotrich and Pollock, 2004; Schinka et al., 2004; Sen et al., 2004; Cho et al., 2005; Fan and Sklar, 2005; Lasky-Su et al., 2005; Munafo et al., 2005), suicide (Anguelova et al., 2003b; Lin and Tsai, 2004), alcoholism (Heinz et al., 2003; Feinn et al., 2005), response to antidepressants (Peters et al., 2004; Lesch and Gutknecht,
and brain function as revealed by MRI scans (Hariri et al., 2002; Bertolino et al., 2005; Hariri et al., 2005; Pezawas et al., 2005). Although many studies have reported correlations between SERTLPR alleles and SERT expression levels or clinical phenotype, almost an equal number of studies have not. Where correlations have been replicated or survived meta-analysis, the estimated contributions of the SERTLPR polymorphism are small.

Most studies on SERTLPR to date (more than 350 currently cited in PubMed) were motivated by early work showing that the l form of the SERT promoter is more active than the s form. Curiously, only a few studies have examined the relationship between SERTLPR genotype and SERT mRNA or protein expression in the relevant population of cells: serotonergic neurons in the pons and brainstem. The goal of the present study was to examine the relationship between SERTLPR genotype and SERT mRNA expression in tissue sections from the rostral pons that contain dorsal and median raphe nuclei.

Because it is difficult to make quantitative comparisons of mRNA levels between autopsy tissue sections, we decided to compare allele-specific expression of SERT mRNA using total cellular RNA isolated from the same tissue section. In this experimental design, the mRNA transcribed from one allele acts as a control for the other, since both are presumably exposed to the same trans-acting gene regulators, time to tissue harvest (post mortem interval), and medication and medical histories. Allele-specific differences in mRNA expression in these samples are taken to reflect differences in cis-acting factors, i.e., the presence of genetic variants located on the same chromosome, likely in the vicinity of the gene, that affect mRNA levels through modulation of mRNA transcription, splicing and/or degradation.

Recent studies have demonstrated that the expression of a large percentage of human genes are regulated by cis-acting genetic elements (Yan et al., 2002a; Pinsonneault et al., 2004a). Given the paucity of functional polymorphisms in the coding region of SERT affecting protein structure and function (Hahn and Blakely, 2002; Ozaki et al., 2003; Sutcliffe et al., 2005), we
suspect that functional differences in SERT expression among individuals likely arise from differences in regulation of the gene by cis-acting genetic elements located outside of the coding region.

To study allele-specific regulation of the SERT mRNA expression, we have developed an assay in which a common SNP (G2651T, rs1042173) in the 3’-UTR is used to quantify relative levels of mRNA transcribed from each allele (Figure. 2-1.A). This SNP represents the entire SERT gene region for each of the two alleles in a given individual, regardless of linkage disequilibrium, which is determined from a population average. As shown in Figure. 2-5 and Table. 2-2, this assay is highly reproducible, allowing us to reliably detect differences of 15-20% in expression between alleles. Using this assay we demonstrated a statistically significant AEI in 9 out of 29 pons samples heterozygous for the marker SNP, indicating the presence of a cis-acting factor in ~30% of the population. While the extent of the imbalance was relatively small (15-25%), this could be sufficient for biological effects. For example, an AEI of maximally 30% was previously suggested to be sufficient to account for an association between catecholamine O-methyl transferase (COMT) and schizophrenia (Bray et al., 2003c).

On the basis of previous findings that the l promoter variant generates higher levels of mRNA expression compared to the s variant (Heils et al., 1996; Lesch et al., 1996), we had predicted that l/s heterozygous samples would show differences in allelic mRNA expression. Contrary to this prediction, however, we observed no consistent differences in levels of expression of individual SERT alleles in the pons of l/s individuals. It is possible that a cis-acting functional polymorphism is associated with the l form in a portion of alleles; however, we were unable to ascertain the phasing between the marker SNP and the l/s forms due to the low linkage disequilibrium between these markers (Table 2-4). This prevented unambiguous assignment of the l form to the more highly expressed allele. Our observation that 12/16 l/s heterozygous
samples showed no allelic imbalance appears to rule out SERTLPR as the determining factor for AEI of SERT mRNA expression in the rostral pons.

Because the SNaPshot analysis of SERT mRNA in the pons produced unexpected results, we performed a series of control experiments to validate the assays. A critical variable is the quality of the mRNA extracted from postmortem brain tissues. This problem is in part alleviated by the use of a ratio method, which as mentioned above, bypasses issues such as differences in mRNA decay between tissue samples. (We cannot completely rule out, however, the possibility of differential rates of decay between alleles during the postmortem interval or tissue processing). As an extra precaution, we used gene-specific primers for the synthesis of cDNA (see Materials and Methods), with the oligonucleotide targeting a sequence close to the marker SNP. This ensures a high yield of the region of interest, but possibly masks differences in splicing at other sites of the transcribed gene. The most likely sources of variability in the assay are in mRNA extraction, cDNA synthesis and/or PCR amplification. To account for this variability, we carried out three independent mRNA/cDNA preparations, and performed PCR in triplicate for each assay. Duplicate analysis of the PCR products using forward and backward primers in the SNaPshot reaction revealed that the detection step is precise and does not significantly contribute to overall measurement errors. Furthermore, we ascertained that the amount of mRNA available from pons tissue sections is in a range where we can expect reproducible results after PCR amplification. Allelic DNA ratios can be determined with an average standard deviation of 4% (range: 0-11% for individual samples) across all samples, while the average standard deviation for mRNA/cDNA measurements was 5% (range: 0-10% for individual samples).

To provide a direct comparison with previously published studies, we tested allelic mRNA expression of SERT in Epstein Barr virus-transformed B-lymphocytes. Although we observed significant AEI in 5 out of 9 samples (15 - 40%; Figure.2-8), this did not correlate with l/s heterozygosity in SERTLPR, consistent with our result with pons tissue sections. Absolute
levels of *SERT* mRNA were also measured in B-lymphoblast cell lines (14 l/l, 12 l/s and 8 s/s) by real-time PCR, a more advanced technique than the competitive PCR method used previously (Lesch et al., 1996). Large variations in *SERT* mRNA levels were observed, but again there was no significant correlation with the l or s promoter alleles (Figure.2-7).

Using quantitative real-time PCR to quantify levels of *SERT* mRNA, Hranilovic et al (Hranilovic et al., 2004) reported a trend towards higher expression in EB virus-transformed lymphoblasts from l/l individuals compared to lymphoblasts from l/s and s/s individuals, but these differences were not statistically significant. These authors did observe a statistically significant correlation, however, between the combined *SERTLPR* and *VNTR* genotypes and *SERT* expression. By contrast, we observed no effect of *VNTR* by itself on AEI of *SERT* mRNA, and only a weak association for *VNTR* in combination with *SERTPR*. (Statistical significance was not observed following corrections for multiple testing.) Taken together, these studies suggest that cis-acting polymorphisms regulate small allele-specific differences in *SERT* mRNA expression, but argue against *SERTLPR* or the *VNTR per se* as playing a significant role in regulating the *SERT* gene.

To our knowledge, only three previous studies have examined the relationship between *SERTPR* genotype and *SERT* mRNA or protein expression in serotonergic neurons of the dorsal and median raphe nuclei. Little et al (Little et al., 1998), observed higher levels of binding of the SERT ligand [123I]CIT in the dorsal and median raphe nuclei of l/l individuals (n = 16) compared to l/s individuals (n = 10). [123I]CIT binding levels in s/s individuals (n = 4), however, were close to those of l/l individuals, an effect attributed to the small size of the s/s sample. Moreover, [123I]CIT binding levels were lower in l/l chronic alcohol users compared to l/s or s/s chronic alcohol users, which was taken to suggest that *SERT* expression is influenced by environmental factors. Analysis of *SERT* mRNA expression by in situ hybridization, showed higher levels of *SERT* expression in l/l individuals (n =13) compared to l/s (n = 8) and s/s individuals (n = 3).
Although these results are consistent with higher SERT expression in l/l compared to l/s and s/s individuals, the small number of individuals examined warrants caution in building upon these results.

Heinz et al., (Heinz et al., 2000) used [123I]CIT binding measured in vivo by SPECT to examine the relationship between SERTLPR genotype and SERT protein expression in the raphe area of the dorsal brainstem of alcohol users and abstinent volunteers. These studies revealed that l/l control subjects (n =3) had 2-fold higher levels of [1123]CIT binding in the brainstem compared to l/s and s/s subjects (n = 5). Consistent with the in vitro binding studies of Little et al.,(Little et al., 1998) [123I]CIT binding levels in l/l alcoholic individuals were lower than those in l/s and s/s alcoholic individuals. These results again emphasize that SERT expression may be modulated by environmental factors. Heinz and coworkers suggested that a possible explanation for the low [123I]CIT binding levels in l/l alcoholic individuals is that serotonergic neurons in l/l individuals are more susceptible to the neurotoxic effects of chronic alcohol use. In contrast to the results of Heinz and colleagues, a similar study by Willeit et al., (Willeit et al., 2001) failed to detect a correlation between SERTLPR genotype and [123I]CIT binding levels in the thalamus-hypothalamus or mesencephlon-pons measured by SPECT scans of healthy subjects [l/l (n = 3); l/s (n = 9); s/s (n = 4)]. Clearly more work will be required to resolve these conflicting reports and sort out the relative importance of SERTLPR genotype and environmental factors in determining SERT expression in the raphe nuclei.

While a direct association between SERTLPR and psychiatric traits remains elusive, other experiments suggest a gene-environmental interaction as a determining factor. In mice with disrupted SERT, homozygous and heterozygous strains displayed more fearful behavior and greater increases in the stress hormone adrenocorticotropin in response to stress compared to homozygous (SERT +/+ ) controls, but in the absence of stress, no differences related to genotype were observed (Murphy et al., 2001). In rhesus macaques, with an SERTLPR analogous to that of
humans, the short allele is associated with decreased serotonergic function among monkeys reared in stressful conditions, but not among normally reared monkeys (Bennett et al., 2002). In humans, individuals with one or two copies of short allele of the SERT promoter polymorphism exhibited more depressive symptoms, diagnosable depression, and suicidality in relation to stressful life events than individuals homozygous for the long allele; however, no direct association between the SERT gene and depression was observed (Caspi et al., 2003). Taken together, these studies show that there may not be a direct relationship between SERTLPR and a variety of biochemical and clinical phenotypes; yet, intriguing results do imply a role for SERTLPR in behavioral disorders when environmental factors are taken into account.

While our results appear to contradict previous assertions that the SERTLPR regulates gene expression, the following possibilities need to be considered: 1) The l and s forms harbor functional polymorphisms of lower frequency (~20%) that affect mRNA levels. Our study was not designed to identify these putative functional polymorphism, which will require a larger sample size, 2) The pons tissue sections used in this study were from anonymous donors. It is possible that in afflicted individuals (e.g., with depression disorder), the frequency of another functional polymorphisms in linkage disequilibrium with SERTLPR is significantly higher, giving rise to clinical associations indirectly between SERTLPR and phenotypic trait, and 3) AEI of SERT occurs only following exposure to stress, when different transcriptional regulation could involve the SERTLPR.

This is the first study to measure mRNA expression of SERT using an allele-specific assay to evaluate the potential role of cis-acting polymorphisms in the human pons. Although we observed statistically significant differences in mRNA expression between alleles, our results argue against a role for SERTLPR in determining SERT mRNA expression. Our approach lends itself to finding functional polymorphisms in SERT and assessing their relative role in inter-individual variability.
Figure 2.1. A. Structure of the SERT gene and the location of SERTLPR, VNTR and the 3’UTR marker SNP G2651T (rs1042173). B. Genotyping the long (l) /short (s) SERTLPR promoter polymorphism. Synthetic oligonucleotide primers flanking the l/s polymorphic region of the SERT promoter were used for PCR amplification of genomic DNA from three individuals (a, b and c). PCR products were resolved by electrophoresis in a 2% agarose gel and visualized by staining with ethidium bromide. The deduced genotypes for individuals a, b and c are l/l, l/s and s/s, respectively. C. Genotyping the VNTR located in intron 2. A modified forward primer containing a fluorescent dye (6-FAM) attached to its 5’ end was used for PCR amplification of genomic DNA from three individuals (a’, b’, and c’; different from those in A). PCR products were separated by capillary electrophoresis on an ABI 3730 DNA analyzer. The genotypes for individuals a’, b’, and c’ are 10/12, 12/12 and 10/10 repeats, respectively.
Figure. 2.2. Logarithmic plots of real-time PCR amplifications with increasing amounts of template DNA performed on the ABI Prism® 7000 Sequence Detection System using SYBR Green I fluorescence to detect the PCR products. C_T (cycle threshold) is the fractional cycle number at which the total fluorescence passes a fixed threshold (T). A. PCR amplification of 1, 5, 10, 20, 30 and 50 ng template DNA, corresponding to approximately 300, 1500, 3000, 6000, 9000 and 15000 molecules, respectively. These traces were used to determine the C_T for each amount of template DNA. B. The log (number of DNA molecules) was plotted versus C_T as determined in A. The data points were fitted to the equation: C_T = (slope) \times \log_{10}(\text{number of DNA molecules}) + (Y\text{-intercept}); R^2 = 0.985; (\epsilon = 10^{-1/\text{slope}} - 1).
Figure 2.3. Assay of SERT mRNA levels in pons RNA samples using quantitative real-time PCR. Number of samples examined: (l/l) = 12, (l/s) = 15 and (s/s) = 2. \( N_0 \) = estimated initial number of SERT mRNA molecules in each sample.
Figure 2.4. Assay of *SERT* mRNA levels in RNA of non-pons brain regions (n = 5) and pons (n = 26) using quantitative real-time PCR. cDNA was prepared from 1 µg of total RNA isolated from pons, cerebellum or the occipital, frontal, parietal and temporal lobes of the cortex and 1/20 of this cDNA was used as a template for quantitative real-time PCR. *SERT* mRNA levels were significantly higher in pons compared to other brain regions (One-way ANOVA, p = 1.36E-17). $N_0 =$ initial number of *SERT* mRNA molecules in each sample.
Figure 2.5. Measurement of SERT allelic expression imbalance (AEI). Segments of genomic DNA or cDNA containing the SERT marker SNP G2651T were amplified by PCR using flanking oligonucleotide primers. The PCR products were then used as templates in SNaPshot primer extension assays. (A) Diagram of extended primers annealed to SERT PCR products. Upper box: extended forward primers annealed to noncoding DNA strands derived from the ‘G’ (top) or ‘T’ (bottom) SERT allele. The extended primers are labeled with dR110 or dROX, respectively. Lower box: extended reverse primers annealed to coding strand of the ‘G’ (top) or ‘T’ (bottom) allele. The extended primers are labeled with dTAMRA or dR6G, respectively. (B) Quantification of extended primers by capillary electrophoresis and fluorescence detection using an ABI 3730 DNA sequence analyzer. Left box: the forward extension primers produce ‘G’ and ‘T’ peaks, labeled with dR110-(dark gray) or dROX-(medium gray) terminal dideoxynucleotides, respectively. Right box: the reverse extension primers produce ‘A’ and ‘C’ peaks, labeled with dR6G-(light gray) and dTAMRA-(black) terminal dideoxynucleotides, respectively. The peaks in the top half of each box are derived from genomic DNA and those in the bottom half from cDNA. Relative amounts of ‘G’ and ‘T’ extension products measured using the forward primer were determined by dividing the area of the dR110 peak by the area of the dROX peak. Relative amounts of ‘C’ and ‘A’ extension products measured using the reverse primer were calculated by dividing the area of the dTAMRA peak by the area of the dR6G peak. The inverse of the mean of the G/T or C/A ratios produced from genomic DNA was used as a correction factor for individual DNA and cDNA-derived ratios obtained with the forward and reverse primers, respectively.
Figure. 2.6. Comparison of corrected genomic DNA ratios and corrected RNA ratios in human pons samples measured with SNaPshot primer extension assays using the marker SNP G2651T (rs1042173). Twenty-nine of the 48 individuals in our collection were heterozygous (G/T) for this SNP. Data are expressed as G/T ratios. The DNA ratios are the average of three measurements, each from independent preparations of genomic DNA, and the RNA ratios are the average of three measurements, each from an independent reverse-transcription reaction using RNA from a single preparation. (*) samples where the means of allelic ratios determined from DNA and RNA differed by more than 10% and where there was no overlap between 2 X s.e.m.
Figure. 2.7. SERT mRNA expression in Epstein-Barr virus transformed lymphoblasts was determined by quantitative real-time PCR. One µg of total cellular RNA was used for cDNA synthesis and 1/20 of this cDNA was used as a template for quantitative real-time PCR. Number of samples examined: (I/I) = 11, (I/S) = 9 and (S/S) = 7. There was no statistical difference in cycle threshold (Cₜₛ) among the three SERT genotypes (One-way ANOVA, p=0.980). N₀ = initial number of SERT mRNA molecules in each sample.
Figure. 2.8. Comparison of corrected genomic DNA and mRNA ratios for Epstein–Barr virus transformed lymphoblast cell lines. Data are expressed as G/T ratios for the markers SNP G2651T. The ratios are the average of three measurements from one preparation of genomic DNA and three measurements from three independent reverse transcription reactions. (*) samples where the means of allelic ratios determined from DNA and RNA differed by more than 10% and where there was no overlap between the standard error of each mean.
<table>
<thead>
<tr>
<th>UMS#</th>
<th>Tissue</th>
<th>Sex</th>
<th>PMI</th>
<th>Cause of death (health history/medications)</th>
<th>Disorder</th>
<th>Age</th>
<th>Sex</th>
<th>Race</th>
</tr>
</thead>
<tbody>
<tr>
<td>602</td>
<td>midbrain/pons</td>
<td>3</td>
<td>15</td>
<td>accident; multiple injuries</td>
<td>control</td>
<td>27</td>
<td>M</td>
<td>Caucasian</td>
</tr>
<tr>
<td>479</td>
<td>midbrain/pons</td>
<td>3</td>
<td>15</td>
<td>multiple injuries</td>
<td>control</td>
<td>21</td>
<td>M</td>
<td>Caucasian</td>
</tr>
<tr>
<td>914</td>
<td>midbrain/pons</td>
<td>3</td>
<td>15</td>
<td>accident</td>
<td>control</td>
<td>20</td>
<td>M</td>
<td>Caucasian</td>
</tr>
<tr>
<td>1029</td>
<td>midbrain/pons</td>
<td>3</td>
<td>15</td>
<td>multiple injuries</td>
<td>control</td>
<td>20</td>
<td>M</td>
<td>Hispanic</td>
</tr>
<tr>
<td>1076</td>
<td>midbrain/pons</td>
<td>3</td>
<td>15</td>
<td>accident; multiple injuries</td>
<td>control</td>
<td>17</td>
<td>F</td>
<td>Caucasian</td>
</tr>
<tr>
<td>1106</td>
<td>midbrain/pons</td>
<td>3</td>
<td>15</td>
<td>accident; multiple injuries</td>
<td>control</td>
<td>16</td>
<td>M</td>
<td>Caucasian</td>
</tr>
<tr>
<td>1112</td>
<td>midbrain/pons</td>
<td>3</td>
<td>15</td>
<td>suicide; diphenhydramine intoxication (history of depression)</td>
<td>depression</td>
<td>16</td>
<td>F</td>
<td>Caucasian</td>
</tr>
<tr>
<td>1115</td>
<td>midbrain/pons</td>
<td>3</td>
<td>16</td>
<td>acute drug intoxication (history of polycystic kidney disease and depression; medications: venlafaxine, divalproex, alprazolam and clonazepam)</td>
<td>depression</td>
<td>42</td>
<td>F</td>
<td>Caucasian</td>
</tr>
<tr>
<td>1230</td>
<td>midbrain/pons</td>
<td>3</td>
<td>16</td>
<td>multiple injuries</td>
<td>control</td>
<td>15</td>
<td>M</td>
<td>AA</td>
</tr>
<tr>
<td>1297</td>
<td>midbrain/pons</td>
<td>3</td>
<td>16</td>
<td>multiple injuries</td>
<td>control</td>
<td>22</td>
<td>F</td>
<td>Caucasian</td>
</tr>
<tr>
<td>1406</td>
<td>midbrain/pons</td>
<td>3</td>
<td>16</td>
<td>multiple injuries</td>
<td>control</td>
<td>26</td>
<td>F</td>
<td>Asian</td>
</tr>
<tr>
<td>1499</td>
<td>midbrain/pons</td>
<td>3</td>
<td>16</td>
<td>drowning (history of depression)</td>
<td>depression</td>
<td>6</td>
<td>M</td>
<td>Caucasian</td>
</tr>
<tr>
<td>1506</td>
<td>midbrain/pons</td>
<td>3</td>
<td>16</td>
<td>multiple injuries</td>
<td>control</td>
<td>28</td>
<td>M</td>
<td>Caucasian</td>
</tr>
<tr>
<td>1546</td>
<td>midbrain/pons</td>
<td>3</td>
<td>16</td>
<td>multiple injuries</td>
<td>control</td>
<td>17</td>
<td>F</td>
<td>Caucasian</td>
</tr>
<tr>
<td>1607</td>
<td>midbrain/pons</td>
<td>3</td>
<td>16</td>
<td>multiple drug intoxication</td>
<td>control</td>
<td>41</td>
<td>F</td>
<td>Caucasian</td>
</tr>
<tr>
<td>1135</td>
<td>midbrain/pons</td>
<td>3</td>
<td>16</td>
<td>multiple injuries (end stage renal disease and depression; medications: lorazepam, oxycodone)</td>
<td>depression</td>
<td>33</td>
<td>M</td>
<td>AA</td>
</tr>
</tbody>
</table>

Table 2.1. Brain sample demographics
<table>
<thead>
<tr>
<th>ID</th>
<th>gender</th>
<th>injury/illness</th>
<th>control</th>
<th>sex</th>
<th>race</th>
</tr>
</thead>
<tbody>
<tr>
<td>1365</td>
<td></td>
<td>multiple injuries (treated for heret problem in her late teens and also treated for drug abuse during the same time)</td>
<td>control</td>
<td>26</td>
<td>M</td>
</tr>
<tr>
<td>1436</td>
<td></td>
<td>asthma</td>
<td>control</td>
<td>36</td>
<td>M</td>
</tr>
<tr>
<td>1539</td>
<td></td>
<td>multiple injuries</td>
<td>control</td>
<td>33</td>
<td>F</td>
</tr>
<tr>
<td>1551</td>
<td></td>
<td>multiple injuries (history of depression, medication: serotonin, somnus)</td>
<td>depression</td>
<td>43</td>
<td>F</td>
</tr>
<tr>
<td>1609</td>
<td></td>
<td>coronary artery thrombosis</td>
<td>control</td>
<td>33</td>
<td>F</td>
</tr>
<tr>
<td>1672</td>
<td></td>
<td>asphyxiation (history of suicide attempts due to problems with her girlfriend)</td>
<td>suicide</td>
<td>19</td>
<td>M</td>
</tr>
<tr>
<td>1678</td>
<td></td>
<td>cardiac arrhythmias</td>
<td>control</td>
<td>52</td>
<td>M</td>
</tr>
<tr>
<td>1744</td>
<td></td>
<td>arteriosclerotic cardiovascular disease</td>
<td>control</td>
<td>48</td>
<td>F</td>
</tr>
<tr>
<td>1745</td>
<td></td>
<td>arteriosclerotic cardiovascular disease</td>
<td>control</td>
<td>46</td>
<td>F</td>
</tr>
<tr>
<td>1638</td>
<td></td>
<td>multiple injuries</td>
<td>control</td>
<td>21</td>
<td>M</td>
</tr>
<tr>
<td>133</td>
<td></td>
<td>accident, multiple injuries (history of depression and migraines headaches)</td>
<td>control</td>
<td>36</td>
<td>F</td>
</tr>
<tr>
<td>917</td>
<td></td>
<td>accident, multiple injuries</td>
<td>control</td>
<td>14</td>
<td>F</td>
</tr>
<tr>
<td>1027</td>
<td></td>
<td>accident, multiple injuries</td>
<td>control</td>
<td>22</td>
<td>M</td>
</tr>
<tr>
<td>1065</td>
<td></td>
<td>suicide, gunshot wound</td>
<td>suicide</td>
<td>23</td>
<td>F</td>
</tr>
<tr>
<td>1101</td>
<td></td>
<td>accident, multiple injuries</td>
<td>control</td>
<td>19</td>
<td>F</td>
</tr>
<tr>
<td>1102</td>
<td></td>
<td>accident, multiple injuries</td>
<td>control</td>
<td>21</td>
<td>M</td>
</tr>
<tr>
<td>1209</td>
<td></td>
<td>chest and abdominal injuries</td>
<td>control</td>
<td>39</td>
<td>M</td>
</tr>
<tr>
<td>1257</td>
<td></td>
<td>multiple injuries (head on collision, hematoxia for maxilla, parietal bone fracture, white matter)</td>
<td>control</td>
<td>16</td>
<td>M</td>
</tr>
</tbody>
</table>
Table 2.2. Samples from six individuals heterozygous for the 3’ UTR marker SNP were assayed using forward and reverse extension primers as shown in Figure 2.5.A. Ratios were calculated by dividing the area of two peaks and corrected as describe in Materials and methods. (A) Quantification of G/T ratios determined in SNaPshot primer extension assays depicted in Figure 2.5.A (top) and B (left), and (B) C/A ratios were determined from peak areas obtained with SNaPshot primer extension assays depicted in Figure 2.5.A (bottom) and B (right).
Table 2.3. Mean G/T ratios (7s.d.) determined with SNaPshot primer extension assays using genomic DNA or RNA prepared from human pons. Individual ratios were calculated by dividing the ‘G’ peak areas by ‘T’ peak areas. The mean ratios are the average of three measurements using independent preparations of genomic DNA or RNA. The ratios of RNAs were normalized as described in Materials and methods. The samples are grouped based upon the SERTLPR genotype: s = short promoter allele; l = long promoter allele. Genotypes for the intron 2 VNTR (10 or 12 repeats) and the 3’UTR SNP (A2858C; rs3813034) located within the putative SERT mRNA polyadenylation site (Poly-A SNP) are listed for each sample. There were no statistically significant differences in AEI among the l/l, l/s and s/s SERT genotypes (P = 0.381; one-way ANOVA).
Table 2.4 (A) Estimated linkage disequilibrium (D') for three SERT polymorphisms\textsuperscript{a} and (B) estimated individual and pairwise contributions of these polymorphisms to SERT mRNA expression in human pons\textsuperscript{b}.

\textsuperscript{a}For this analysis, genotypes for three polymorphisms in 47 samples were imported into the HelixTree\textsuperscript{R} Linkage Disequilibrium program. Values of D' range from 0 (no linkage disequilibrium) to 1.0 (complete association). D0 = (p_{AB} - p_A p_B) / D_{max}, where D_{max} is the maximum possible value of |p_{AB} - p_A p_B| for allele frequencies p_A and p_B. p_A = frequency of allele 1 (locus A); p_B = frequency of allele 1 (locus B); p_{AB} = frequency of allele 1(A)/allele 1 (B) haplotype. P-values are listed in \textit{italics}.

\textsuperscript{b} The HelixTree\textsuperscript{R} Two-Loci genetic plot is used to calculate the statistical significance for associations between pairs of genetic markers and a response variable. In the present analysis, 29 samples heterozygous for G2651T were divided into two groups: one group consisting of samples with significant AEI and a second group with nonsignificant AEI. Probabilities for associations between pairs of markers and AEI were then calculated from the genotypes. The values listed were not corrected for multiple testing.

<table>
<thead>
<tr>
<th></th>
<th>SERTLPR</th>
<th>VNTR</th>
<th>G2651T</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SERTLPR</td>
<td>0.152</td>
<td>0.0319</td>
<td></td>
</tr>
<tr>
<td>VNTR</td>
<td>0.323</td>
<td>0.0069</td>
<td></td>
</tr>
<tr>
<td>G2651T</td>
<td>0.331</td>
<td>0.725</td>
<td></td>
</tr>
<tr>
<td>(B)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SERTLPR</td>
<td>0.316</td>
<td>0.648</td>
<td></td>
</tr>
<tr>
<td>VNTR</td>
<td>0.170</td>
<td>0.170</td>
<td></td>
</tr>
<tr>
<td>G2651T</td>
<td></td>
<td>0.998</td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER 3

TRYPTOPHAN HYDROXYLASE 2 (TPH2) HAPLOTYPES PREDICT LEVELS OF TPH2 mRNA EXPRESSION IN HUMAN PONS

INTRODUCTION

Tryptophan hydroxylase (TPH) catalyzes the rate-limiting step in the synthesis of serotonin (5-hydroxytryptamine; 5-HT) (Fitzpatrick, 1999a), a neurotransmitter that plays an important role in the regulation of mood (Lucki, 1998a). Dysregulation of serotonergic activity has been associated with major depression, anxiety disorders and suicidal behavior (Ressler and Nemeroff, 2000). 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 (Walther et al., 2003a; Bach-Mizrachi et al., 2005; Zill et al., 2005). The dorsal and media raphe nuclei are the major source of serotonin in the forebrain (Walther et al., 2003a), including areas implicated in mood and anxiety disorders.

Because 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 (Breidenthal et al., 2004a; Zhou et al., 2005b; Zhang et al., 2006). 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 (Yan et al., 2002b; Bray et al., 2003a). In this method, relative levels of mRNA expressed from each of two alleles are measured using RNA isolated from individuals who 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%. Because 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) (Pinsonneault et al., 2004b), p-glycoprotein (MDR1) (Wang et al., 2005b), the µ–opiate receptor (OPRM) (Zhang et al., 2005c), and the serotonin transporter (SERT) (Lim et al., 2006).

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
Isolation of DNA and RNA from human pons  Isolation of DNA and RNA from the tissue samples in our collection has been described previously (Lim et al., 2006). 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 (Lim et al., 2006). 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 3-1. The locations of these
SNPs within the TPH2 gene are shown in Figure. 3-1. Sequences of the PCR amplification and primer extension primers and reaction conditions for each primer set used for genotyping are shown in Table 3-2.

**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/haplovew/](http://www.broad.mit.edu/mpg/haplovew/)) (Barrett et al., 2005). Predicted diplotypes for each individual in our collection were calculated from the genotyping data using HelixTree<sup>RT</sup> (GoldenHelix, Inc., Bozeman, MT, USA).

**AEI measurements**  Measurements of allele-specific mRNA expression were carried out as described previously (Lim et al., 2006). 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 1µM oligo(dT)20 primer (Invitrogen), 1 µl of 10mM dNTP mix (Invitrogen), 0.5 µl of 1 µM TPH2 gene-specific primer (5’-TTAATTCTCCAATGGAGGAAAGGA-3’), 4 µl of 5 X first-strand buffer (Invitrogen), 1 µl of RNaseOUT (40 U/ml), 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’-ACGAGACTT-TCTGCGAGGACTG-3’, and the reverse primer 5’-TTAATTCTCCAATGGAGGAAAGGA-3’ with the following cycles: (1 X (5 min at 95 °C); 35 X (30 s at 95 °C , 30 s at 60 °C , 1 min at 72 °C ) ; 1 X (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’-GATC-CCCTCTACACCACCC-3’ for rs7305115 and 5’-AAAGGA-
GTCCTGCTCCATA-3’ for rs4290270 with the following cycles: (25 X (10 s at 96 °C, 5 s at 50 °C, 30 s at 72 °C)). Unincorporated fluorescent dNTP analogs were removed by incubation with 1.0U 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 X with genomic DNA and 3 X 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’-TTA-ATTCTCAATGGAGAAAGGA-3’ (TPH2) or 5’-GTGTGGGACTGAGTGTG-3’ (GAPDH). Segments of TPH2 or GAPDH cDNAs were amplified using TPH2- or GAPDH-specific primer sets and heatactivated 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’-ACGAGACTTTTCTGGCAGGACTG-3’ (forward) and 5’-TTAATTCTCCAATGGAGGAAGGA-3’ (reverse) and the GAPDH amplification primers were: 5’-CAGCAAGAGCACAAGAGGAAGAGA-3’ (forward) and 5’-GTGTGGTG-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 (CT) and normalized by subtracting CT 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 $TPH2$ 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 3-1 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 3-2 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 3.3). 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 3-3 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 3-3). 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 3-4 shows that there is an excellent correlation between AEI measurements made using the two marker SNPs.

The results in Figures 3-2-3-4 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 3-1 and Figure 3-1 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 3-5.A shows 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 3-1) and US and Finnish populations (Zhou et al., 2005a). The frequencies of haplotypes within each block are listed in Figure 3-5.B, and the predicted diplotypes for each individual in our collection are listed in Table 3.4..

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 3-5. The strength of each correlation was assessed using the Kappa-statistic (Saffen et al., 1999). As shown in Figure 3-6, 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\textsuperscript{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 3-4). 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 3-6 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), 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 Zhou et al (Zhou et al., 2005a).

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 CT) were normalized by subtracting CT 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 3-7 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). CT 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 3-8, 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).

**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 3-6). 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 3-1), 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 (Zuker, 2003) 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 3-9). 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., 2004)), 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) (Cartegni et al., 2002). Exons containing a nonfunctional or partially functional ESE are often skipped during RNA splicing (Cartegni et al., 2002), 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 (Cartegni et al., 2002). 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 3.2–3.4) and lower levels of TPH2 mRNA expression in G/A heterozygotes and G/G homozygotes compared to A/A homozygotes (Figure 3-7).
To determine if aberrant \textit{TPH2} mRNAs lacking exon 7 are expressed in pons, we carried out RT-PCR amplification of \textit{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-containing 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 \textit{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 \textit{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 (Wang et al., 2006b). The existence of positive selection would indicate that \textit{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 \textit{TPH2} mRNA expression are identified, knowledge of marker SNPs and haplotypes that strongly predict high or low levels of \textit{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% (Ebmeier et al., 2006).

Since the discovery of the TPH2 gene in 2003 (Walther et al., 2003a), 14 published studies have examined possible associations between TPH2 SNPs and various mental disorders including major depression (Zill et al., 2004b; Zhang et al., 2005a; Zhou et al., 2005a), bipolar disorder (Harvey et al., 2004), anxiety disorders (Mossner et al., 2006b; Mossner et al., 2006a), attention-deficit/hyperactivity disorder (ADHD)(Sheehan et al., 2005; Walitza et al., 2005), autism (Coon et al., 2005) or suicidal behavior (De Luca et al., 2004; Zill et al., 2004a; De Luca et al., 2005; Zhou et al., 2005a; De Luca et al., 2006; Mergen et al., 2006). 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 (Harvey et al., 2004; Zill et al., 2004b; Zill et al., 2004a; Coon et al., 2005; Sheehan et al., 2005; Walitza et al., 2005; Zhang et al., 2005a; Zhou et al., 2005a; Mossner et al., 2006a), and five showing no significant associations (De Luca et al., 2004; De Luca et al., 2005; De Luca et al., 2006; Mergen et al., 2006; Mossner et al., 2006b). 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 (Zhang et al., 2005a), but this was not replicated in five subsequent studies (Garriock et al., 2005; Glatt et al.,...
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. (Zill et al., 2004b) 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 3-1 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 (Zill et al., 2004a), 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. (Zhou et al., 2005a) 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.(Harvey et al., 2004) 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.(Mossner et al., 2006a) 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.(Walitza et al., 2005) described a weak association between regulatory region SNPs (rs4570625 and rs11178997) and attention-deficit hyperactivity disorder (ADHD). Sheehan et al.(Sheehan et al., 2005) 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., (Coon et al., 2005) reported statistically significant associations between autism and TPH2 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 amygdale activity (Brown et al., 2005; Canli et al., 2005) or ‘emotional processing’(Herrmann et al., 2007). Brown et al.(Brown et al., 2005) 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., (Canli et al., 2005) observed increased responses in both the right and left amygdala of T-allele carriers viewing fearful, happy or sad faces compared to faces with neutral expression. Hermann et al.,(Herrmann et al., 2007) 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 (Heils et al., 1995; Lesch et al., 1996) 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 3.6 and Table 3.5). 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 3-5) 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 (Whitaker-Azmitia, 2001; Gaspar et al., 2003), it is possible that differential expression of TPH2 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 \textit{TPH2} and mental disorders should consider the following points: (1) The contribution of \textit{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 \textit{SERT} promoter polymorphisms (Caspi et al., 2003), 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 \textit{SERT} promoter polymorphisms to correlate more highly with ‘neuroticism,’ a personality trait highly associated with depression, than with depression itself (Munafo et al., 2006), (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, $> 2$, and (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 \textit{TPH2}. AEI measurements, as described here for \textit{TPH2}, have already revealed the presence of frequent functional polymorphisms in other genes previously implicated in mental disorders, including the m-opiate (OPRM1) (Zhang et al., 2005b), monoamine oxidase A (MAOA) (Pinsonneault et al., 2006) 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.
**Figure 3.1.** 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).
Figure 3.2. 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.
Figure. 3.3. 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 (*).
Figure. 3.4. 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; r² = 0.86).
Figure 3.5.A. D’ plot for the 22 SNPs listed in Table 3-1 (main text) based upon genotyping data from 36 Caucasian individuals in our collection. The plot was generated using Haplovview (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.
Figure 3.5.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.”
Figure. 3.6. Correlations between heterozygosity of individual $TPH2$ SNPs and AEI of $TPH2$ mRNA. Y-axis: Kappa-coefficients were calculated from the data in Table 3-5 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 ((**): $P < 0.001$; (*): $P = 0.003$).
Figure. 3.7. *TPH2* mRNA levels in pons measured using real-time PCR. The Y-axis plots the difference between CT determined for *GAPDH* and *TPH2* mRNAs. Individuals were 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 3.8. Comparison of TPH2 mRNA expression levels in different tissues. The Y-axis plots the difference between CT for GAPDH and TPH2 mRNAs. Results obtained from 27 pons samples, five non pons brain regions (cerebellum and occipital, frontal, parietal and temporal cortices) 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).
Figure. 3-9. 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/l actinomycin D (added 24 h after transfection). As indicated, 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.
<table>
<thead>
<tr>
<th>SNP no.</th>
<th>DbSNP no.</th>
<th>Location on chromosome 12</th>
<th>Location within TPH2 gene</th>
<th>Allele frequencies</th>
<th>Heterozygosity</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>rs4570625</td>
<td>70618190</td>
<td>Upstream</td>
<td>G/T = 0.72/0.28</td>
<td>0.403</td>
</tr>
<tr>
<td>02</td>
<td>rs10748185</td>
<td>70622122</td>
<td>Intron 2</td>
<td>A/G = 0.51/0.49</td>
<td>0.500</td>
</tr>
<tr>
<td>03</td>
<td>rs2129575</td>
<td>70626340</td>
<td>Intron 4</td>
<td>G/T = 0.74/0.26</td>
<td>0.385</td>
</tr>
<tr>
<td>04</td>
<td>rs1386488</td>
<td>70630885</td>
<td>Intron 5</td>
<td>A/C = 0.85/0.15</td>
<td>0.255</td>
</tr>
<tr>
<td>05</td>
<td>rs1843809</td>
<td>70634965</td>
<td>Intron 5</td>
<td>T/G = 0.83/0.17</td>
<td>0.282</td>
</tr>
<tr>
<td>06</td>
<td>rs1386495</td>
<td>70638589</td>
<td>Intron 5</td>
<td>T/C = 0.83/0.17</td>
<td>0.282</td>
</tr>
<tr>
<td>07</td>
<td>rs1386494</td>
<td>70638710</td>
<td>Intron 5</td>
<td>G/A = 0.88/0.12</td>
<td>0.211</td>
</tr>
<tr>
<td>08</td>
<td>rs6582072</td>
<td>70640744</td>
<td>Intron 5</td>
<td>G/A = 0.83/0.17</td>
<td>0.282</td>
</tr>
<tr>
<td>09</td>
<td>rs2171363</td>
<td>70646531</td>
<td>Intron 5</td>
<td>C/T = 0.53/0.47</td>
<td>0.498</td>
</tr>
<tr>
<td>10</td>
<td>rs4760815</td>
<td>70652496</td>
<td>Intron 6</td>
<td>T/A = 0.65/0.35</td>
<td>0.455</td>
</tr>
<tr>
<td>11</td>
<td>rs7305115</td>
<td>70659129</td>
<td>Exon 7</td>
<td>G/A = 0.65/0.35</td>
<td>0.455</td>
</tr>
<tr>
<td>12</td>
<td>rs6582078</td>
<td>70661158</td>
<td>Intron 7</td>
<td>T/G = 0.60/0.40</td>
<td>0.480</td>
</tr>
<tr>
<td>13</td>
<td>rs1023900</td>
<td>70668514</td>
<td>Intron 7</td>
<td>T/C = 0.79/0.21</td>
<td>0.332</td>
</tr>
<tr>
<td>14</td>
<td>rs1007023</td>
<td>70674641</td>
<td>Intron 8</td>
<td>T/G = 0.89/0.11</td>
<td>0.196</td>
</tr>
<tr>
<td>15</td>
<td>rs1352251</td>
<td>70684161</td>
<td>Intron 8</td>
<td>T/C = 0.59/0.41</td>
<td>0.484</td>
</tr>
<tr>
<td>16</td>
<td>rs1473473</td>
<td>70690645</td>
<td>Intron 8</td>
<td>G/A = 0.88/0.12</td>
<td>0.211</td>
</tr>
<tr>
<td>17</td>
<td>rs9325202</td>
<td>70693744</td>
<td>Intron 8</td>
<td>G/A = 0.65/0.35</td>
<td>0.455</td>
</tr>
<tr>
<td>18</td>
<td>rs1487275</td>
<td>70696559</td>
<td>Intron 8</td>
<td>T/G = 0.78/0.22</td>
<td>0.343</td>
</tr>
<tr>
<td>19</td>
<td>rs1386486</td>
<td>70699487</td>
<td>Intron 8</td>
<td>C/T = 0.61/0.39</td>
<td>0.476</td>
</tr>
<tr>
<td>20</td>
<td>rs4290270</td>
<td>70702502</td>
<td>Exon 9</td>
<td>A/T = 0.63/0.37</td>
<td>0.466</td>
</tr>
<tr>
<td>21</td>
<td>rs1872824</td>
<td>70716581</td>
<td>Intron 9</td>
<td>C/T = 0.64/0.36</td>
<td>0.461</td>
</tr>
<tr>
<td>22</td>
<td>rs1352252</td>
<td>70733308</td>
<td>Downstream</td>
<td>A/G = 0.56/0.44</td>
<td>0.493</td>
</tr>
</tbody>
</table>

Table 3.1. TPH2 SNPs examined in this study
Table 3.2. Primers used for genotyping 20 SNPs in TPH2

<table>
<thead>
<tr>
<th>snp</th>
<th>GC primers 1</th>
<th>GC primers 2</th>
<th>Common primers 1</th>
<th>Common primers 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>n4578425</td>
<td>5'--CACCTTCCGACACAACACTG-3'</td>
<td>5'--GCGCCGCGGGTCTGCTAATAAG-3'</td>
<td>5'--ACACTACCTTTCCTTACACATG-3'</td>
<td></td>
</tr>
<tr>
<td>n15748181</td>
<td>5'--CTGACGACAGCGGCTTCTG-3'</td>
<td>5'--GCGCCGCGGGTCTGCTAATAAG-3'</td>
<td>5'--TGGTTGTTTCTGGTTCTCTTATT-3'</td>
<td></td>
</tr>
<tr>
<td>n21129575</td>
<td>5'--GACGCTGACGACAGCGGCTTCTG-3'</td>
<td>5'--GCGCCGCGGGTCTGCTAATAAG-3'</td>
<td>5'--GGTTGTTTCTGGTTCTCTTATT-3'</td>
<td></td>
</tr>
<tr>
<td>n1286488</td>
<td>5'--TCTATCGGTCGCAACAGCGGCTTCTG-3'</td>
<td>5'--GCGCCGCGGGTCTGCTAATAAG-3'</td>
<td>5'--GGTTGTTTCTGGTTCTCTTATT-3'</td>
<td></td>
</tr>
<tr>
<td>n1843809</td>
<td>5'--GCACGCTGACGACAGCGGCTTCTG-3'</td>
<td>5'--GCGCCGCGGGTCTGCTAATAAG-3'</td>
<td>5'--GGTTGTTTCTGGTTCTCTTATT-3'</td>
<td></td>
</tr>
<tr>
<td>n1388485</td>
<td>5'--GACGCTGACGACAGCGGCTTCTG-3'</td>
<td>5'--GCGCCGCGGGTCTGCTAATAAG-3'</td>
<td>5'--GGTTGTTTCTGGTTCTCTTATT-3'</td>
<td></td>
</tr>
<tr>
<td>n1286484</td>
<td>5'--CACCTTCCGACACAACACTG-3'</td>
<td>5'--GCGCCGCGGGTCTGCTAATAAG-3'</td>
<td>5'--ACACTACCTTTCCTTACACATG-3'</td>
<td></td>
</tr>
<tr>
<td>n8582275</td>
<td>5'--GGTTGTTTCTGGTTCTCTTATT-3'</td>
<td>5'--GCGCCGCGGGTCTGCTAATAAG-3'</td>
<td>5'--GGTTGTTTCTGGTTCTCTTATT-3'</td>
<td></td>
</tr>
<tr>
<td>n2117183</td>
<td>5'--GACGCTGACGACAGCGGCTTCTG-3'</td>
<td>5'--GCGCCGCGGGTCTGCTAATAAG-3'</td>
<td>5'--GGTTGTTTCTGGTTCTCTTATT-3'</td>
<td></td>
</tr>
<tr>
<td>n4708815</td>
<td>5'--ACACGCTGACGACAGCGGCTTCTG-3'</td>
<td>5'--GCGCCGCGGGTCTGCTAATAAG-3'</td>
<td>5'--GGTTGTTTCTGGTTCTCTTATT-3'</td>
<td></td>
</tr>
<tr>
<td>n4708878</td>
<td>5'--GCACGCTGACGACAGCGGCTTCTG-3'</td>
<td>5'--GCGCCGCGGGTCTGCTAATAAG-3'</td>
<td>5'--GGTTGTTTCTGGTTCTCTTATT-3'</td>
<td></td>
</tr>
<tr>
<td>n1828896</td>
<td>5'--GACGCTGACGACAGCGGCTTCTG-3'</td>
<td>5'--GCGCCGCGGGTCTGCTAATAAG-3'</td>
<td>5'--GGTTGTTTCTGGTTCTCTTATT-3'</td>
<td></td>
</tr>
<tr>
<td>n1870722</td>
<td>5'--GACGCTGACGACAGCGGCTTCTG-3'</td>
<td>5'--GCGCCGCGGGTCTGCTAATAAG-3'</td>
<td>5'--GGTTGTTTCTGGTTCTCTTATT-3'</td>
<td></td>
</tr>
<tr>
<td>n1352551</td>
<td>5'--ACACGCTGACGACAGCGGCTTCTG-3'</td>
<td>5'--GCGCCGCGGGTCTGCTAATAAG-3'</td>
<td>5'--GGTTGTTTCTGGTTCTCTTATT-3'</td>
<td></td>
</tr>
<tr>
<td>n1472477</td>
<td>5'--ACACGCTGACGACAGCGGCTTCTG-3'</td>
<td>5'--GCGCCGCGGGTCTGCTAATAAG-3'</td>
<td>5'--GGTTGTTTCTGGTTCTCTTATT-3'</td>
<td></td>
</tr>
<tr>
<td>n9253502</td>
<td>5'--ACACGCTGACGACAGCGGCTTCTG-3'</td>
<td>5'--GCGCCGCGGGTCTGCTAATAAG-3'</td>
<td>5'--GGTTGTTTCTGGTTCTCTTATT-3'</td>
<td></td>
</tr>
<tr>
<td>n1482775</td>
<td>5'--ACACGCTGACGACAGCGGCTTCTG-3'</td>
<td>5'--GCGCCGCGGGTCTGCTAATAAG-3'</td>
<td>5'--GGTTGTTTCTGGTTCTCTTATT-3'</td>
<td></td>
</tr>
<tr>
<td>n1388486</td>
<td>5'--GACGCTGACGACAGCGGCTTCTG-3'</td>
<td>5'--GCGCCGCGGGTCTGCTAATAAG-3'</td>
<td>5'--GGTTGTTTCTGGTTCTCTTATT-3'</td>
<td></td>
</tr>
<tr>
<td>n1870724</td>
<td>5'--ACACGCTGACGACAGCGGCTTCTG-3'</td>
<td>5'--GCGCCGCGGGTCTGCTAATAAG-3'</td>
<td>5'--GGTTGTTTCTGGTTCTCTTATT-3'</td>
<td></td>
</tr>
<tr>
<td>n1352552</td>
<td>5'--ACACGCTGACGACAGCGGCTTCTG-3'</td>
<td>5'--GCGCCGCGGGTCTGCTAATAAG-3'</td>
<td>5'--GGTTGTTTCTGGTTCTCTTATT-3'</td>
<td></td>
</tr>
</tbody>
</table>
Table 3-3. Measurements of AEI using the marker SNPs rs7305115 and rs4290270. * = statistically significant. Bold characters indicate samples showing statistically significant AEI for the marker SNPs rs7305115 and/or rs4290270.
Table 3.4. Predicted diplotypes for individuals in sample. 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 AE1 (Kappa coefficient > 0.66) are listed in bold type.

<table>
<thead>
<tr>
<th>Sample #</th>
<th>Haplotype 1</th>
<th>Haplotype 2</th>
<th>EM-p</th>
</tr>
</thead>
<tbody>
<tr>
<td>602</td>
<td>TGGCTTTGACATTAGTTGTTGGAGGTTCG</td>
<td>GGGCGCAAATAGTGCAGGTTTGG</td>
<td>1</td>
</tr>
<tr>
<td>010</td>
<td>AGAATTCTGCTTTAATTCAACAGAATTCTGTTG</td>
<td>TGGATTGGCTTACAGCTTTACATG</td>
<td>0.99</td>
</tr>
<tr>
<td>679</td>
<td>AGAATTCTGCTTTAATTCAACAGAATTCTGTTG</td>
<td>GGGCGCAAATAGTGCAGGTTTGG</td>
<td>0.95</td>
</tr>
<tr>
<td>914</td>
<td>AGAATTCTGCTTTAATTCAACAGAATTCTGTTG</td>
<td>GGGCGCAAATAGTGCAGGTTTGG</td>
<td>0.99</td>
</tr>
<tr>
<td>247</td>
<td>AGAATTCTGCTTTAATTCAACAGAATTCTGTTG</td>
<td>TGGATTGGCTTACAGCTTTACATG</td>
<td>1</td>
</tr>
<tr>
<td>1028</td>
<td>AGAATTCTGCTTTAATTCAACAGAATTCTGTTG</td>
<td>GGGCGCAAATAGTGCAGGTTTGG</td>
<td>1</td>
</tr>
<tr>
<td>1027</td>
<td>AGAATTCTGCTTTAATTCAACAGAATTCTGTTG</td>
<td>TGGATTGGCTTACAGCTTTACATG</td>
<td>1</td>
</tr>
<tr>
<td>1029</td>
<td>AGAATTCTGCTTTAATTCAACAGAATTCTGTTG</td>
<td>GGGCGCAAATAGTGCAGGTTTGG</td>
<td>1</td>
</tr>
<tr>
<td>1054</td>
<td>AGAATTCTGCTTTAATTCAACAGAATTCTGTTG</td>
<td>GGGCGCAAATAGTGCAGGTTTGG</td>
<td>1</td>
</tr>
<tr>
<td>1083</td>
<td>AGAATTCTGCTTTAATTCAACAGAATTCTGTTG</td>
<td>GGGCGCAAATAGTGCAGGTTTGG</td>
<td>0.99</td>
</tr>
<tr>
<td>1070</td>
<td>AGAATTCTGCTTTAATTCAACAGAATTCTGTTG</td>
<td>GGGCGCAAATAGTGCAGGTTTGG</td>
<td>0.99</td>
</tr>
<tr>
<td>1076</td>
<td>AGAATTCTGCTTTAATTCAACAGAATTCTGTTG</td>
<td>GGGCGCAAATAGTGCAGGTTTGG</td>
<td>0.99</td>
</tr>
<tr>
<td>1101</td>
<td>AGAATTCTGCTTTAATTCAACAGAATTCTGTTG</td>
<td>GGGCGCAAATAGTGCAGGTTTGG</td>
<td>0.99</td>
</tr>
<tr>
<td>1109</td>
<td>AGAATTCTGCTTTAATTCAACAGAATTCTGTTG</td>
<td>GGGCGCAAATAGTGCAGGTTTGG</td>
<td>0.99</td>
</tr>
<tr>
<td>1104</td>
<td>AGAATTCTGCTTTAATTCAACAGAATTCTGTTG</td>
<td>GGGCGCAAATAGTGCAGGTTTGG</td>
<td>0.99</td>
</tr>
<tr>
<td>1103</td>
<td>AGAATTCTGCTTTAATTCAACAGAATTCTGTTG</td>
<td>GGGCGCAAATAGTGCAGGTTTGG</td>
<td>1</td>
</tr>
<tr>
<td>1112</td>
<td>AGAATTCTGCTTTAATTCAACAGAATTCTGTTG</td>
<td>GGGCGCAAATAGTGCAGGTTTGG</td>
<td>0.99</td>
</tr>
<tr>
<td>1115</td>
<td>AGAATTCTGCTTTAATTCAACAGAATTCTGTTG</td>
<td>GGGCGCAAATAGTGCAGGTTTGG</td>
<td>1</td>
</tr>
<tr>
<td>1125</td>
<td>AGAATTCTGCTTTAATTCAACAGAATTCTGTTG</td>
<td>GGGCGCAAATAGTGCAGGTTTGG</td>
<td>1</td>
</tr>
<tr>
<td>1169</td>
<td>AGAATTCTGCTTTAATTCAACAGAATTCTGTTG</td>
<td>GGGCGCAAATAGTGCAGGTTTGG</td>
<td>0.95</td>
</tr>
<tr>
<td>1209</td>
<td>AGAATTCTGCTTTAATTCAACAGAATTCTGTTG</td>
<td>GGGCGCAAATAGTGCAGGTTTGG</td>
<td>1</td>
</tr>
<tr>
<td>1220</td>
<td>AGAATTCTGCTTTAATTCAACAGAATTCTGTTG</td>
<td>GGGCGCAAATAGTGCAGGTTTGG</td>
<td>1</td>
</tr>
<tr>
<td>1257</td>
<td>AGAATTCTGCTTTAATTCAACAGAATTCTGTTG</td>
<td>GGGCGCAAATAGTGCAGGTTTGG</td>
<td>1</td>
</tr>
<tr>
<td>1269</td>
<td>AGAATTCTGCTTTAATTCAACAGAATTCTGTTG</td>
<td>GGGCGCAAATAGTGCAGGTTTGG</td>
<td>1</td>
</tr>
</tbody>
</table>

continued
Table 3.4. continued

| 1279 | G_A_G_A_T_T_G_C_T_G_T_T_T_A_T_A_G_C_A_C_A | T_G_A_T_T_T_G_T_A_R_G_C_T_C_A_R_G_T_T_T_G | 0.99 |
| 1287 | G_G_A_T_T_G_G_C_T_G_T_T_T_A_T_A_G_T_T_T_C_A | T_A_T_T_G_G_T_G_T_T_T_A_G_T_T_T_G | 0.99 |
| 1307 | G_A_C_A_T_T_G_C_T_T_T_T_A_T_A_G_C_A_C_A | G_O_A_T_T_G_G_T_G_T_T_T_A_G_T_T_C_A | 1.0 |
| 1305 | T_G_A_C_A_T_T_G_C_T_T_T_T_A_T_A_G_C_A_C_A | G_O_A_T_T_G_G_T_G_T_T_T_A_G_T_T_C_A | 1.0 |
| 1407 | T_G_A_C_A_T_T_G_C_T_T_T_T_A_T_A_G_C_A_C_A | T_G_A_C_A_T_T_G_C_T_T_T_T_A_G_T_T_C_A | 1.0 |
| 1409 | G_A_C_A_T_T_G_C_T_T_T_T_A_T_A_G_C_A_C_A | G_A_T_T_G_G_T_G_T_T_T_A_G_T_T_C_A | 1.0 |
| 1419 | G_A_C_A_T_T_G_C_T_T_T_T_A_T_A_G_C_A_C_A | G_A_T_T_G_G_T_G_T_T_T_A_G_T_T_C_A | 1.0 |
| 1420 | G_A_C_A_T_T_G_C_T_T_T_T_A_T_A_G_C_A_C_A | G_A_T_T_G_G_T_G_T_T_T_A_G_T_T_C_A | 1.0 |
| 1442 | G_A_C_A_T_T_G_C_T_T_T_T_A_T_A_G_C_A_C_A | G_G_A_T_T_G_G_C_T_G_T_T_T_A_G_T_T_C_A | 1.0 |
| 1466 | T_G_A_C_A_T_T_G_C_T_T_T_T_A_T_A_G_C_A_C_A | T_G_A_T_T_G_A_T_A_T_T_T_A_G_T_T_C_A | 1.0 |
| 1449 | T_G_A_C_A_T_T_G_C_T_T_T_T_A_T_A_G_C_A_C_A | T_G_A_C_A_T_T_G_C_T_T_T_T_A_G_T_T_C_A | 1.0 |
| 1500 | G_A_C_A_T_T_G_C_T_T_T_T_A_T_A_G_C_A_C_A | C_C_A_T_T_C_C_T_T_T_T_A_G_T_T_C_A | 1.0 |
| 1539 | G_A_C_A_T_T_G_C_T_T_T_T_A_T_A_G_C_A_C_A | G_A_T_T_G_G_T_G_T_T_T_A_G_T_T_C_A | 1.0 |
| 1540 | G_A_C_A_T_T_G_C_T_T_T_T_A_T_A_G_C_A_C_A | G_G_C_C_A_T_T_A_R_G_C_G_T_C_A | 1.0 |
| 1546 | T_G_A_C_A_T_T_G_C_T_T_T_T_A_T_A_G_C_A_C_A | T_G_A_C_A_T_T_G_C_T_T_T_T_A_G_T_T_C_A | 1.0 |
| 1546 | T_G_A_C_A_T_T_G_C_T_T_T_T_A_T_A_G_C_A_C_A | T_G_A_C_A_T_T_G_C_T_T_T_T_A_G_T_T_C_A | 1.0 |
| 1546 | T_G_A_C_A_T_T_G_C_T_T_T_T_A_T_A_G_C_A_C_A | T_G_A_C_A_T_T_G_C_T_T_T_T_A_G_T_T_C_A | 1.0 |
| 1546 | T_G_A_C_A_T_T_G_C_T_T_T_T_A_T_A_G_C_A_C_A | T_G_A_C_A_T_T_G_C_T_T_T_T_A_G_T_T_C_A | 1.0 |
| 1546 | T_G_A_C_A_T_T_G_C_T_T_T_T_A_T_A_G_C_A_C_A | T_G_A_C_A_T_T_G_C_T_T_T_T_A_G_T_T_C_A | 1.0 |
| 1639 | G_A_C_A_T_T_G_C_T_T_T_T_A_T_A_G_C_A_C_A | G_A_T_T_G_G_T_G_T_T_T_A_G_T_T_C_A | 1.0 |
| 1607 | C_A_C_A_T_T_C_C_T_T_T_T_A_T_A_G_C_A_C_A | C_C_C_C_A_T_A_R_G_C_T_C_C_C_G_T_T_C_A | 0.98 |
| 1609 | G_A_C_A_T_T_G_C_T_T_T_T_A_T_A_G_C_A_C_A | T_G_A_T_T_G_A_T_A_T_T_T_A_G_T_T_C_A | 1.0 |
| 1639 | G_A_C_A_T_T_G_C_T_T_T_T_A_T_A_G_C_A_C_A | G_A_T_T_G_G_T_G_T_T_T_A_G_T_T_C_A | 1.0 |
| 1639 | G_A_C_A_T_T_G_C_T_T_T_T_A_T_A_G_C_A_C_A | G_A_T_T_G_G_T_G_T_T_T_A_G_T_T_C_A | 1.0 |
| 1639 | G_A_C_A_T_T_G_C_T_T_T_T_A_T_A_G_C_A_C_A | G_A_T_T_G_G_T_G_T_T_T_A_G_T_T_C_A | 1.0 |
| 1639 | G_A_C_A_T_T_G_C_T_T_T_T_A_T_A_G_C_A_C_A | G_A_T_T_G_G_T_G_T_T_T_A_G_T_T_C_A | 1.0 |
| 1639 | G_A_C_A_T_T_G_C_T_T_T_T_A_T_A_G_C_A_C_A | G_A_T_T_G_G_T_G_T_T_T_A_G_T_T_C_A | 1.0 |
| 1639 | G_A_C_A_T_T_G_C_T_T_T_T_A_T_A_G_C_A_C_A | G_A_T_T_G_G_T_G_T_T_T_A_G_T_T_C_A | 1.0 |
| 1639 | G_A_C_A_T_T_G_C_T_T_T_T_A_T_A_G_C_A_C_A | G_A_T_T_G_G_T_G_T_T_T_A_G_T_T_C_A | 1.0 |
| 1639 | G_A_C_A_T_T_G_C_T_T_T_T_A_T_A_G_C_A_C_A | G_A_T_T_G_G_T_G_T_T_T_A_G_T_T_C_A | 1.0 |
| 1639 | G_A_C_A_T_T_G_C_T_T_T_T_A_T_A_G_C_A_C_A | G_A_T_T_G_G_T_G_T_T_T_A_G_T_T_C_A | 1.0 |
| 1639 | G_A_C_A_T_T_G_C_T_T_T_T_A_T_A_G_C_A_C_A | G_A_T_T_G_G_T_G_T_T_T_A_G_T_T_C_A | 1.0 |
Table 3.5. ω = (ad-bc)/(p1q2+p2q1); 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(-); p1 = proportion of samples that are heterozygous for a give SNP (see Table S3); q1 = proportion of samples that are homozygous for a given SNP (see Table S3); p2 = proportion of samples that are AEI(+) = 0.667 (18/27); q2 = 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. ω > 0.75: excellent agreement; 0.4 to 0.75 = fair to good agreement; < 0.4 = poor agreement.

<table>
<thead>
<tr>
<th>#</th>
<th>dbSNP</th>
<th>a = Hetero &amp; AEI (+)</th>
<th>b = Hetero &amp; AEI (-)</th>
<th>c = Homo &amp; AEI (+)</th>
<th>d = Homo &amp; AEI (-)</th>
<th>ω</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>rs4570625</td>
<td>0.370 (10/27)</td>
<td>0.037 (1/27)</td>
<td>0.333 (9/27)</td>
<td>0.259 (7/27)</td>
<td>0.311</td>
<td>0.053</td>
</tr>
<tr>
<td>02</td>
<td>rs10478185</td>
<td>0.566 (15/27)</td>
<td>0.111 (3/27)</td>
<td>0.148 (4/27)</td>
<td>0.185 (5/27)</td>
<td>0.400</td>
<td>0.037</td>
</tr>
<tr>
<td>03</td>
<td>rs2129575</td>
<td>0.370 (10/27)</td>
<td>0.037 (1/27)</td>
<td>0.333 (9/27)</td>
<td>0.259 (7/27)</td>
<td>0.311</td>
<td>0.053</td>
</tr>
<tr>
<td>04</td>
<td>rs1386488</td>
<td>0.222 (6/27)</td>
<td>0.074 (2/27)</td>
<td>0.481 (13/27)</td>
<td>0.222 (6/27)</td>
<td>0.047</td>
<td>0.732</td>
</tr>
<tr>
<td>05</td>
<td>rs1643809</td>
<td>0.148 (4/27)</td>
<td>0.111 (3/27)</td>
<td>0.519 (14/27)</td>
<td>0.222 (6/27)</td>
<td>-0.115</td>
<td>0.373</td>
</tr>
<tr>
<td>06</td>
<td>rs1386495</td>
<td>0.148 (4/27)</td>
<td>0.111 (3/27)</td>
<td>0.556 (15/27)</td>
<td>0.185 (5/27)</td>
<td>-0.115</td>
<td>0.373</td>
</tr>
<tr>
<td>07</td>
<td>rs1386494</td>
<td>0.148 (4/27)</td>
<td>0.074 (2/27)</td>
<td>0.556 (15/27)</td>
<td>0.222 (6/27)</td>
<td>-0.027</td>
<td>0.822</td>
</tr>
<tr>
<td>08</td>
<td>rs6582072</td>
<td>0.222 (6/27)</td>
<td>0.074 (2/27)</td>
<td>0.481 (13/27)</td>
<td>0.222 (6/27)</td>
<td>0.047</td>
<td>0.732</td>
</tr>
<tr>
<td>09</td>
<td>rs2171363</td>
<td>0.630 (17/27)</td>
<td>0.037 (1/27)</td>
<td>0.074 (2/27)</td>
<td>0.259 (7/27)</td>
<td>0.743</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>10</td>
<td>rs4760815</td>
<td>0.630 (17/27)</td>
<td>0.037 (1/27)</td>
<td>0.074 (2/27)</td>
<td>0.259 (7/27)</td>
<td>0.743</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>11</td>
<td>rs7305115</td>
<td>0.630 (17/27)</td>
<td>0.037 (1/27)</td>
<td>0.074 (2/27)</td>
<td>0.259 (7/27)</td>
<td>0.743</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>12</td>
<td>rs6582078</td>
<td>0.593 (16/27)</td>
<td>0.037 (1/27)</td>
<td>0.111 (3/27)</td>
<td>0.259 (7/27)</td>
<td>0.669</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>13</td>
<td>rs1023990</td>
<td>0.481 (13/27)</td>
<td>0.111 (3/27)</td>
<td>0.222 (6/27)</td>
<td>0.185 (5/27)</td>
<td>0.279</td>
<td>0.135</td>
</tr>
<tr>
<td>14</td>
<td>rs1007023</td>
<td>0.148 (4/27)</td>
<td>0.074 (2/27)</td>
<td>0.556 (15/27)</td>
<td>0.222 (6/27)</td>
<td>-0.027</td>
<td>0.822</td>
</tr>
<tr>
<td>15</td>
<td>rs1332251</td>
<td>0.519 (14/27)</td>
<td>0.037 (1/27)</td>
<td>0.185 (5/27)</td>
<td>0.259 (7/27)</td>
<td>0.534</td>
<td>0.003</td>
</tr>
<tr>
<td>16</td>
<td>rs1473473</td>
<td>0.148 (4/27)</td>
<td>0.111 (3/27)</td>
<td>0.556 (15/27)</td>
<td>0.185 (5/27)</td>
<td>-0.115</td>
<td>0.373</td>
</tr>
<tr>
<td>17</td>
<td>rs9325202</td>
<td>0.593 (16/27)</td>
<td>0.037 (1/27)</td>
<td>0.111 (3/27)</td>
<td>0.259 (7/27)</td>
<td>0.669</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>18</td>
<td>rs1487275</td>
<td>0.407 (11/27)</td>
<td>0.111 (3/27)</td>
<td>0.296 (8/27)</td>
<td>0.185 (5/27)</td>
<td>0.173</td>
<td>0.333</td>
</tr>
<tr>
<td>19</td>
<td>rs1386496</td>
<td>0.519 (14/27)</td>
<td>0.259 (7/27)</td>
<td>0.185 (5/27)</td>
<td>0.037 (1/27)</td>
<td>-0.149</td>
<td>0.430</td>
</tr>
<tr>
<td>20</td>
<td>rs4290270</td>
<td>0.556 (15/27)</td>
<td>0.259 (7/27)</td>
<td>0.148 (4/27)</td>
<td>0.037 (1/27)</td>
<td>-0.096</td>
<td>0.601</td>
</tr>
<tr>
<td>21</td>
<td>rs1872824</td>
<td>0.444 (12/27)</td>
<td>0.148 (4/27)</td>
<td>0.259 (7/27)</td>
<td>0.148 (4/27)</td>
<td>0.119</td>
<td>0.525</td>
</tr>
<tr>
<td>22</td>
<td>rs1352252</td>
<td>0.407 (11/27)</td>
<td>0.111 (3/27)</td>
<td>0.296 (8/27)</td>
<td>0.185 (5/27)</td>
<td>0.173</td>
<td>0.333</td>
</tr>
</tbody>
</table>
Table 3-6. Haplotype frequencies for 38 Caucasians in sample (76 chromosomes)
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.

<table>
<thead>
<tr>
<th>rs7305115 G-allele haplotypes</th>
<th>rs7305115 A-allele haplotypes</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>C T G T G</td>
<td>T A A G A</td>
<td>0.553 (42/76)</td>
</tr>
<tr>
<td>T A A G A</td>
<td>0.316 (24/76)</td>
<td></td>
</tr>
<tr>
<td>T A A G G</td>
<td>0.053 (4/76)</td>
<td></td>
</tr>
<tr>
<td>T T A G G</td>
<td>0.026 (2/76)</td>
<td></td>
</tr>
<tr>
<td>T T G T A</td>
<td>0.026 (2/76)</td>
<td></td>
</tr>
<tr>
<td>C T G T A</td>
<td>0.013 (1/76)</td>
<td></td>
</tr>
<tr>
<td>C T G G G</td>
<td>0.013 (1/76)</td>
<td></td>
</tr>
<tr>
<td>Total G-allele haplotypes</td>
<td>0.605 (46/76)</td>
<td></td>
</tr>
<tr>
<td>Total A-allele haplotypes</td>
<td>0.395 (30/76)</td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER 4

THE EFFECT OF GENETIC VARIANTS ON EXON SKIPPING IN TPH2
(AN ON-GOING PROJECT)

INTRODUCTION

RNA splicing

Most human genes are transcribed as precursor RNAs that are typically tens of kilobases in length. To produce mature mRNAs, the precursor RNAs undergo “splicing” the process by which large introns are removed and flanking exons precisely ligated together. RNA splicing is catalyzed by an assembly of short nuclear (sn)RNPs and proteins that together constitute the “spliceosome” (Cartegni et al., 2002; Fairbrother et al., 2002). mRNA splicing requires the assembly of the basal splicing machinery in splicesome complexes on consensus sequences present at boundaries between introns and exons. Alternative splicing is a mechanism that allows multiple mRNA isoforms to be generated from a single precursor mRNA. It is important for generating proteins with distinct properties and functions from individual genes, and may also play a role in the timing and location of gene expression (Ladd and Cooper, 2002; Maniatis and Tasic, 2002). Recent research suggests that more than 60% of genes in the human genome undergo alternative splicing (Maniatis and Tasic, 2002; Modrek and Lee, 2002).

Splicing regulatory elements

The specificity of splicing is defined in part by splice-site and branch-site sequences located near the 5’- and 3’-ends of introns (Black, 1995). There are additional elements that
differentially regulate splicing during development and in different type of cells (Cartegni et al., 2002). These elements are conserved among species and similarly regulated genes (Yeo et al., 2004). They are located within exons or introns. Intronic elements are located upstream, downstream, or on both sides of the regulated exon and act by stimulating or repressing splice-site functions. Depending on their location and their effect on the recognition of alternative splice sites, these regulatory elements are referred to as exonic splicing enhancers (ESEs), silencers (ESSs), intronic splicing enhancers (ISEs), or silencers (ISSs). Among these, many ESEs have been characterized (Liu et al., 1998; Schaal and Maniatis, 1999). ESEs serve as binding sites for specific serine/arginine-rich (SR) proteins that participate in splicing. SR proteins that are bound to ESEs can promote exon definition by recruiting proteins of the splicing machinery through their RS (Arg/Ser dipeptides) domain and/or by antagonizing the action of nearby silencer elements (Cartegni et al., 2002).

**Genetic variants that influence splicing**

It has been reported that 60% of human genes undergo alternative splicing (Maniatis and Tasic, 2002; Modrek and Lee, 2002), and that up to 50% of all point mutations that cause genetic diseases do so by causing aberrant splicing (Cartegni et al., 2002). Mutations can disrupt splicing by: i) directly inactivating or creating a splice site, ii) activating a cryptic splice site or iii) interfering with splicing regulatory elements (Krawczak et al., 1992). Mutations that disrupt splicing often result in the elimination of an exon from the mature RNA, a process referred to as exon skipping. mRNAs lacking key exons often produce truncated or defective proteins with dramatically altered sequences. mRNAs that encode truncated proteins are often degraded by a recently discovered pathway termed nonsense-mediated mRNA decay (Cartegni et al., 2002). Exon skipping may be one of the mechanisms that generate allelic expression imbalance, thus contributing to genetic variability between individuals (Lee and Irizarry, 2003). We can test for
the presence of cis-acting polymorphisms affecting splicing by measuring relative amounts of each splice (i.e. mRNA +/- skipped exon).

Identification of cis-regulatory elements

Large data sets including genomic DNA, mRNA and ESTs have been used to identify regulatory elements functioning in alternative splicing. During the last few years, several useful bioinformatics tools have been developed to predict the effects of genetic variants on RNA splicing, including PupaSNP (http://pupasuite.bioinfo.cipf.es/), RESCUE-ESE (http://genes.mit.edu/burgelab/rescue-ese/) and ESEfinder (http://rulai.cshl.edu/tools/ESE/). Computational approaches to identify putative elements have been validated for many genes including, ACF (Dance et al., 2002), BRCA1 (Liu et al., 2001), BRCA2 (Fackenthal et al., 2002), FBN1 (Caputi et al., 2002), IGF1 (Smith et al., 2002), SMN1 (Cartegni and Krainer, 2002), and CFTR (Pagani et al., 2003). Expression of mini-gene pre-mRNAs by transient transfection provides a rapid assay for loss-of-function and gain-of-function analyses for cis-elements and trans-acting factors that affect splicing regulation and also can determine whether an allelic variant has an effect on splicing efficiency.

As described in Chapter 3, we observed significant allelic expression imbalance of TPH2 mRNA, suggesting the presence of functional cis-acting polymorphisms within or near the TPH2 gene. From analysis of possible effects of TPH2 SNPs using PupaSNP, the G-allele of SNP rs7305115 appeared to disrupt the binding of ESEs, suggesting that this allele might favor skipping of exon 7. This prediction was consistent with the observed lower expression of mRNA from G-allele. To investigate underlying mechanisms, we constructed mini-gene expression vectors, introduced these into cultured cell lines, and examined the functional effects of the A-and G-alleles on exon 7 skipping.
MATERIALS AND METHODS

Materials

As described above, frozen pons sections from 48 individuals were obtained from the Brain and Tissue Bank for Developmental Disorders (University of Maryland, Baltimore). The demographics of these samples are described in table 2-1. PCR reagents were purchased from Promega (Madison, WI, USA). Oligonucleotide primers were designed using OLIGO 4.0 (National Biosciences Inc., Plymouth, MN, USA) and synthesized by Sigma-Genosys (St. Louis, MO, USA). Other reagents for cell culture were obtained from Invitrogen (Carlsbad, California, USA) or Fisher Scientific (Pittsburgh, PA, USA). RN46A cells were kindly provided by Dr. Scott Whittemore (University of Louisville).

Expression of exon 7–excluded mRNA in the human brain

Total RNA was isolated from frozen human pons sections as described in Chapter 2. Complementary (c)DNA was reverse-transcribed from 1µg of total RNA using SuperScript™ II Reverse Transcriptase (Invitrogen) with Oligo(dT) primers in a total reaction volume of 20 µl. PCR amplification from this cDNA template was carried out using a forward-primer (5’-TAAATACTGTGGCTACAGAGGACA-3’) derived from exon 6 and a reverse-primer (5’-TTCTTGTGAAAACTGAGCAAGTGA-3’) derived from exon 8. After denaturing at 95°C for 3 min, the PCR reactions were run for 35-cycles under the following conditions: 95°C for 30 s, 60°C for 30 s and 72 °C for 45 s. PCR products were resolved in a 2% agarose gel and visualized by ethidium bromide staining. The PCR product containing exon 7 was predicted to contain 274 bp and the PCR lacking exon 7 was predicted to contain 138 bp.
TPH2 mini-genes

As depicted in Figure 4-1, mini-gene expression vectors comprised a segment of the TPH2 gene containing: i) exons 6, 7 and 8; ii) partially deleted introns 6 and 7 and iii) approximately 500 bp of 5’ and 3’-flanking introns 5 and 8 sequences inserted in the multi-cloning site (MCS) of pcDNA 3.1 (Invitrogen, Carlsbad, California, USA). Approximately 6 kb and 14 kb within introns 6 and 7, respectively, were deleted to reduce the size of the mini-gene. Construction of this plasmid was carried out sequentially as follows: 1) A DNA segment (F1) comprising the entire exon 6 and 410 bp of intron 6 was produced by PCR amplification of human genomic DNA using sequence specific oligonucleotide primers. Sequences corresponding to Kpn I and EcoR I restriction enzyme sites were added to the 5’-end of the forward- and reverse-primers, respectively, to facilitate plasmid construction (forward primer: 5’-GGGGTACCCTG-GTCAGCCCATTCCCAGG-3’, reverse primer: 5’-GGAATTCCTC-TAGCCAGC-ACAGCAG-3’). The PCR reaction was performed under following conditions: i) 1x 95°C for 3 min; ii) 5x [95°C for 30 s, 56°C for 30 s, 72°C for 45 s], iii) 25x [95°C for 30s, 63°C for 30 s, 72°C for 45 s], iiiii) 1x 72°C for 10 min using GoTaq DNA polymerase (Promega, Madison, WI, USA). Following digestion of with Kpn I and EcoR I, the F1 fragment was cloned between the Kpn I and EcoR I sites of pcDNA3.1 (Invitrogen) to obtain the plasmid pcDNA3.1-F1.

A second DNA fragment (F2) comprising approximately 600 bp of intron 6, the entire exon 7 and 507 bp of intron 7 was PCR amplified from genomic DNA using forward and reverse primers containing EcoR I and Xho I restriction enzyme sites, respectively (forward-primer 5’-GGAATTCCTCCTTACCATAAGGCAATAGAG-CAAGC-3’, reverse-primer 5’-GCCCTCGAGGGCACTTGCTGTTCCTCCCTGACTTG-3’). The PCR reaction was performed under following conditions: i) 1x 95°C for 3 min; ii) 5x [95°C for 30s, 56°C for 30 s, 72°C for 45 s], iii) 25x [95°C for 30 s, 63°C for 30 s, 72°C for 45 s], iiiii) 1x 72°C for 10 min. Following
digestion with *Eco*RI and *Xho*I, the F2 fragment was cloned between the *Eco*RI and *Xho*I sites of pcDNA3.1-F1 to obtain the plasmid pcDNA3.1-F1F2.

A third DNA fragment (F3) comprising the entire exon 8 and 501 bp of intron 7 was produced by PCR amplification of genomic DNA using forward and reverse primers containing *Xho*I and *Xba*I restriction enzyme sites, respectively (forward-primer: 5’-CCCTCGAGGGCCTGGTTGGTTGCCC-TATCCCTGATCTAC-3’, reverse-primer: 5’-GCTCTAGAAGGGTGGCCTAG-TTTCGGAACATCTTCATCT-3’). The PCR reaction was performed under the following conditions: i) 1 x 95°C for 3 min; ii) 5 x [95°C for 30 s, 56°C for 30 s, 72°C for 45 s], iii) 25 x [95°C for 30 s, 63°C for 30 s, 72°C for 45 s], iii) 1x 72°C for 10 min. Following digestion with *Xho*I and *Xba*I, the F3 fragment was cloned between the *Xho*I and *Xba*I sites of pcDNA3F1F2 to yield the final minigene expression vector pTPH2mg-A.

This expression vector contains the *A*-allele of rs7305115. To obtain a mini-gene expression vector containing the *G*-allele (pTPH2mg-*G*), the plasmid was subjected to site-specific mutagenesis using the QuikChange® II Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA) with antisense-primer: 5’-CATGTGTCTGGTTCCGGGGTGTAGAGGG-3’ and the sense-primer: 5’- CCCTCTACACCCCGGAACCAGACACATG-3’). The sequences of the *A*- and *G*-allele minigene constructs were confirmed by DNA sequencing.

**Cell culture and transfection**

RN46A cells were grown in DMEM/F12 I (Invitrogen, Carlsbad, California, USA) supplemented with 10 units/ml penicillin, 45ng/ml streptomycin and 10% FBS (HyClone) at 33°C as described in (White et al., 1994). The RN46A cell line was developed directly from embryonic day 13 rat medullary raphe nucleus and they express enzymes necessary for serotonin synthesis, TPH and AAAD. These cells were immortalized by infection with a retrovirus encoding a temperature-sensitive mutant of SV40 large T-antigen. At 33°C, RN46A cells express nuclear T-
antigen and divide with a doubling time of 9 hr. G-418 (250 µg/ml; Invitrogen) was added to the culture medium one day after recovering the cells from frozen cultures. The cells were re-plated 24 hr prior to transfection to obtain cultures approximately 90% confluent on the day of transfection. Cells were transfected with the indicated amounts of the TPH2 mini-gene expression vectors using FuGENE® HD Transfection Reagent (Roche, Indianapolis, IN).

Total RNA was isolated 24 hr after transfection using Trizol reagent (Invitrogen). cDNA was reverse-transcribed from 1 µg of total RNA using the SuperScript II Reaction Kit (Invitrogen) with 0.5 µg of oligo (dT) primer in total reaction volume of 20 µl.

**Quantitative analysis of splice variants**

To measure the relative amounts of splice variants with different sizes, we performed PCR amplifications of cDNA using fluorescently labeled primers. A pair of PCR primers flanking the splice site was designed using OLIGO 4.0 (National Biosciences Inc., Plymouth, MN, USA), with one primer labeled with fluorescent dye FAM (Sigma, St Louis, MO, USA). After initial denaturing at 95°C for 3 min, the PCR reactions were run for 25-cycles under the following conditions: 95°C for 30 s, 60°C for 30 s and 72°C for 45 s with the forward primer, 5’-TAAATACTGTGGCTACAGAGGACA-3’, and the reverse primer, 5’-TTCTTGTG-AAAACCTGAGCAAAGTTAG-3’, using GoTaq DNA polymerase (Promega). The PCR amplification products were resolved on an ABI 3730 Analyzer (Applied Biosystems, Foster City, California, USA) and relative amounts of each PCR product quantified using Gene Mapper 3.0 software. Splice variants with different molecular weights yielded peaks with different retention times. The peak area for each splice variant is proportional to the amount of cDNA amplified. The amount of each splice variant was expressed as the percentage of total transcripts from each splice locus.
RESULTS

SNP rs7305115 is predicted to influence splicing of TPH2 mRNA

Based on our previous studies showing allelic expression imbalance in TPH2 mRNA, we decided to investigate the functional effect of rs7305115. In particular we wondered whether rs7305115 is involved in alternative splicing, since this SNP is located at the end of the exon. To investigate this possibility, we used the Pupa Suite program (http://pupasuite.bioinfo.cipf.es/) to determine if rs7305115 is located within a putative Exonic Splicing Enhancer (ESE) or Exonic Splicing Silencer (ESS). According to this program, the G-allele of rs7305115 is predicted to reduce the binding of the ESE proteins sc35, sf2 and srp40, possibly causing an increase in skipping of exon 7 (Figure 4-2.)

Preferential skipping of exon 7 is a plausible mechanism for generating AEI for TPH2 mRNA, since mRNA lacking this exon would not be measured in the primer extension assay. Thus, even if the TPH2 A- and G- alleles are transcribed with equal efficiency, the G-allele would produce less measurable mRNAs in our assay. In addition to this mechanism, preferential skipping of exon 7 could result in lower levels of the G-allele mRNA due to non-sense mediated RNA decay (Cartegni et al., 2002), thus producing AEI when measured with a marker SNP outside of exon 7 (e.g. rs4290270 in exon 9). TPH2 mRNA lacking exon 7 is predicted to produce a truncated form of the protein, and thus possibly activate non-sense mediated mRNA decay. However, our current assay was not able to measure G-allele mRNA generated because certain proportions of mRNAs are under non-sense mediated mRNA decay. Further studies are required to measure accurate levels of G-allele mRNA with blocking non-sense mediated mRNA decay.
Identification of a novel TPH2 splice variant in human brain

To test our hypothesis that SNP rs7305115 is involved in skipping of exon 7, we first tested the presence of exon7- skipped mRNA in human brain. SNP rs7305115 is located at end of exon 7. We performed PCR amplification using a forward-primer from exon 6 and a reverse-primer from exon 8 with cDNA templates produced by reverse-transcription of human mRNA. PCR products derived from the two splice variants were resolved by electrophoresis in 2% agarose gel and visualized by ethidium bromide staining (Figure 4-3). The 274 bp PCR product was amplified from mRNA containing exon 7 and the 138 bp of PCR product from mRNA lacking exon 7. Both PCR products were confirmed by sequencing.

Quantitative analysis of splice variants

TPH2 minigenes, pTPH2mg-A and pTPH2mg-G, were constructed as described in the Materials and Methods. These constructs contain sequences from exon 6 through exon 8 plus approximately 500 bp of flanking intronic sequences (Figure. 4-1). pTPH2mg-A contains the A-allele of rs7305115 and pTPH2mg-G contains the G-allele of this SNP. Each minigene construct was transfected into neuronal RN46A cells separately, and RNA was isolated at 24 hr after transfection.

We performed PCR amplification using cDNA produced from reverse-transcribed mRNA isolated from transfected cells and forward primers covalently linked to the fluorescent dye (6-FAM). The fluorescently labeled PCR products were resolved on an ABI 3730 analyzer and peak areas from each splice variant quantified using the GeneMapperR (ABI) program. Figure. 4-4 depicts the signals produced by the fluorescent PCR products derived from the two TPH2 splice variants (i.e., plus/minus exon 7). The GeneMapper program calculated the areas under each peak. The relative amount of each splice variant was expressed as the percentage of total transcripts. Consistent with data from human brain, exon skipping was observed in RN46A cells transfected
with TPH2 minigene. However, the amount of exon7-skipped variant was relatively small, accounting for only 3 to 15% of total transcripts. In the cells transfected with pTPH2mg-A, the relative amount of the major splice variant (exon 7–included) was 85% and minor splice variant was 15% of total transcripts. In the cells expressing pTPH2mg-G, major and minor splice variants accounted for 97% and 3% of total transcripts, respectively. This result does not agree with our prediction that a higher percentage of exon 7-skipped mRNA would be observed in cells expressing the TPH2 G-allele. Possible explanations for this result are discussed below.

DISCUSSION

The goal of this study is to investigate the molecular mechanisms underlying allelic expression imbalance of TPH2 mRNA in human pons. As described above, a web-based program, PupaSNP, predicted that an A- to G-transversion at position of rs7305115 changed the consensus ESE sequence that serves as a binding site for the SR proteins sc35, sf2 and srp40 (Figure 4-2). We hypothesized that the G-allele of rs7305115 interferes with the binding of these proteins, resulting in an increase in exon 7 skipping. Skipping exon 7 results in a shift in the TPH2 reading frame, giving rise to premature termination codon. This change may lead to preferential degradation of this mRNA via nonsense-mediated mRNA decay.

Our study is the first to show that the exon 7-skipped TPH2 splice variant is expressed in human pons (Fig4-3). However, the relative amount of this splice variant was too small to allow variations between individuals to be accurately measured. To study exon-skipping mechanisms, we constructed minigenes containing exons 6-8 and shortened introns 7 and 8. Transcripts from these minigenes expressed in eRN46A cells underwent splicing. In this study, we exploited quantitative methods to measure the relative expression of mRNA splice variants, as previously described for the L-type voltage-dependent calcium channel (CACNA1C) gene (Wang et al.,
Contrary to our predictions, we observed higher amounts of exon 7-skipped mRNA in cells expressing the TPH2 minigene carrying A-allele compared to the minigene carrying the G-allele, which did not agree with our prediction (Figure 4-4). These are preliminary observations, however, and important control experiments must be carried out before ruling out exon skipping as a mechanism to generate allelic expression imbalance.

The most important control experiment will be to determine whether nonsense mediated decay (NMD) of TPH2 mRNA prevents accurate measurement the exon 7-skipped splice variant. To investigate this possibility, we will quantify levels of exon 7 +/- TPH2 splice variants in RN46A cells treated with the protein synthesis inhibitor cycloheximide, a treatment previously shown to inhibit NMD (Busi and Cresteil, 2005). Specifically, we will compare cellular levels of the exon 7-skipped splice variant with or without cycloheximide pretreatment in RA46A cells transfected with minigene expression vectors carrying the A- or G-allele. We hypothesize that we will detect higher levels of the exon 7-skipped splice variant in cells expressing the TPH2 G-allele compared with cells expressing the A-allele when the NMD pathway is blocked.

Alternative splicing is regulated by combinational actions through both cis-acting and trans-acting factors. Cis-acting elements refer to the DNA sequences near to and within the gene that influence gene expression. Intragenic sites that function in gene expression are often located within 5’- or 3’-splice sites, intronic branch sites, or splice factor binding sites located in exons or introns (Hastings and Krainer, 2001; Nissim-Rafinia and Kerem, 2002). Exonic splicing enhancers (ESEs) are cis-acting elements that promote splicing. They have been identified within exons that are expressed in a tissue-specific or developmentally regulated manner (Cartegni et al., 2002). These exons typically have intrinsically “weak” splice sites and require the ESE for exon inclusion (Fairbrother et al., 2002). In some cases, ESEs are specifically recognized by one or more SR proteins, which are expressed at different levels in different tissues (Zahler et al., 1993).
The major *trans*-acting factors that regulate alternative splicing are a family of highly conserved serine/arginine-rich RNA-binding proteins (SR proteins), which bind to ESEs and recruit additional splicing factors to the splice sites of adjacent introns (Wu and Maniatis, 1993). Nine human SR proteins are presently known: SF2/ASF, SC35, SRp20, SRp40, SRp75, SRp55, 9G8, SRp30c (Liu et al., 1998). Mutations in the binding site of any of these factors can cause exon skipping or alternative splicing (Cartegni and Krainer, 2002). Our experiments were designed to examine whether the *G*-allele of rs7305115 facilitates exon skipping by disrupting a putative ESE in exon 7. In addition to rs7305115, however, the intronic SNPs rs4760815 and rs6582078 are also highly correlated with allelic expression imbalance of TPH2 mRNA. Since these SNPs are in close linkage disequilibrium with SNP rs7305115, we cannot *a priori* exclude the possibility that they play a role in exon 7 skipping. Further experiments will be required to investigate the possible regulatory roles of these SNPs.

Several bioinformatic webtools are currently available for analyzing the potential effects of SNPs on ESEs, and these have been successfully used to identify functional polymorphisms. For example, experiments using a minigene expression system demonstrated that a mutation in the breast cancer susceptibility gene *BRCA1* predicted to disrupt the binding of SF2/ASF to a ESE increases exon skipping *in vivo* (Liu et al., 2001). Spinal muscular atrophy is caused by mutations in the survival of motor neuron 1 (*SMN1*) gene, even though there is a nearly identical *SMN2* gene that could in theory compensate for the loss of *SMN1* function (Lorson et al., 1999). The reason that *SMN2* cannot fill-in for the missing *SMN1* has recently been shown to be due to disruption of a SF2/ASF-dependent ESE by a SNP in *SMN2*, which results in inefficient inclusion of exon 7 (Singh et al., 2006). *SMN2* is thus unable to produce sufficient full-length SMN protein to compensate for loss of *SMN1* function.

As described above, the PupsSNP webtool predicted that the *A*- to *G*-transition at position of SNP rs7305115 lowers the ESE motif scores, suggesting inefficient recognition by SR proteins.
We used minigenes to determine whether an allelic variant has an effect on splicing efficiency. We made internal deletions of intron 6 and 7 to facilitate construction of the minigene expression vector. The decision to incorporate deleted introns in the minigene expression vector was based on current knowledge that the primary elements regulating alternative splicing are typically within 200-300 nucleotides upstream and/or downstream of the regulated exon (Cooper, 2005). We therefore retained approximately 500 nucleotides of flanking intronic sequences upstream and downstream from the exons in our expression plasmids. Our minigene transcripts were spliced in vitro assay. However, contrary to expectations, we detected higher amount of mRNA lacking exon 7 in cells transfected with the A-allele of rs7305115, which was predicted to have a better ESE consensus sequence.

In interpreting these results, we must consider the possibility that the presence of a high-score motif in a sequence does not necessarily identify a functional ESE, i.e., there may not be a strict correlation between numerical scores and ESE activity. Several important variables, such as the local sequence context, the splice-site strengths, the position of the ESE along the exon and the presence of silencer elements are likely to play a significant role in ESE activity. It is also possible that occurrence of alternative splicing in TPH2 is a normal phenomenon and mutation at rs7305115 simply interferes with a normal splicing event.

To examine these possibilities, it will be necessary to test several minigenes to define a minimal responsive genomic segment to facilitate identification of the relevant cis-element, and we may need to include larger genomic segments to demonstrate the regulation. Because there are examples of genes where cell-specific regulatory elements are located more than one kilobase from the splice site, future experiments may need to be carried out using minigene expression vectors that contain intact introns 6 and 7.

In addition, since splicing patterns may differ between cell lines, it may also be necessary to examine levels of exon 7-skipping in additional cell lines. To investigate the interaction
between SR proteins and ESE motif containing SNP rs7305115, we can introduce another substitution within the low-score motif to recreate a high-score motif and see whether exon-7 inclusion is restored.

Clearly more work will be required to elucidate the molecular mechanisms underlying allelic expression imbalance of \textit{TPH2} mRNA in human brain. Understanding these mechanisms is important, however, to gain insights into the molecular mechanisms that predispose individuals to disorders associated with low serotonin levels, including major depression, anxiety disorders and suicidal behavior.
Figure 4.1. Diagram of a segment of the TPH2 gene extending from exon 6 (E6) through exon 8 (E8) (I: intron). The marker SNP (rs7305115) is located at the end of exon 7 (E7). The locations of PCR primers used to generate the PCR products F1, F2 and F3 are indicated. Those fragments were sequentially inserted into pcDNA3.1 vector to produce the minigene expression vector pTPH2mg-Δ. (See Methods section for details.)
Figure 4.2. Output window from PupaSuite website (http://pupasuite.bioinfo.cipf.es/), listing putative exonic splicing enhancers within TPH2 exons. SNP rs7305115 is predicted to reside within binding sites for the ESE proteins, sc35, sf2 and srp40. The $A$-allele contributes to the canonical binding sequences, while the $G$-allele partially disrupts these sequences. These observations suggest that the $G$-allele may increase skipping of exon 7. The first score is related to the likelihood that the site is a real ESE. A new score is calculated for alleles within these putative binding sites. A large decrease in the score suggests that the allele may disrupt the binding of ESE proteins.
Figure 4.3. Agarose gel electrophoresis photograph of PCR products derived from mRNA containing or lacking exon 7. PCR amplifications were performed with primers spanning exon 6 to exon 8 using mRNA isolated from human pons. PCR products were resolved by electrophoresis in a 2% agarose gel and visualized by staining with ethidium bromide. The 274 bp PCR product contains exon 7 sequences, while the 136 bp PCR product does not.
Figure 4.4. Capillary electrophoresis of fluorescently labeled PCR products derived from TPH2 mRNA splice variants. TPH2 minigene constructs, pTPH2mg-\(A\) and pTPH2mg-\(G\) were transfected into RN46A cells and total cellular RNA was isolated from cells 24 hr after transfection. Complementary DNA fragments spanning exon 6 to exon 8 were amplified with PCR primers, one of which was labeled with HEX fluorescent dye. The PCR products were resolved on a ABI 3730 DNA analyzer. Panel A depicts the PCR products derived from mRNA produced from pTPH2mg-\(A\) and B panel from pTPH2mg-\(G\). The 274 bp PCR product shown on the right represents the splice variant containing exon 7 and 138 bp PCR product on the left represents the exon-7 skipped splice variant.
The goal of the studies described in this dissertation was to identify functional genetic variants that affect the expression of genes implicated in psychiatric disorders. Evidence from twin studies and family-based studies have demonstrated substantial heritability for many mental disorders, suggesting that genes directly contribute to susceptibility to mental illness. Environmental factors, however, cannot be ignored, so that gene-environmental interactions are likely to be required for a complete accounting of susceptibility to mental disorders. Various mechanisms including cis, trans-acting polymorphisms or epigenetic modifications can affect gene expression, and thus produce phenotypic differences among individuals. There is now abundant evidence that functional cis-regulatory variation is widespread in the human genome and contributes to variation in gene expression in a substantial portion of genes surveyed (Rockman and Wray, 2002). The majority of cis-regulatory variants, however, remain to be discovered.

Mental disorders, including major depression, bipolar, schizophrenia, ADHD, and suicidal behavior, are complex diseases that are the product of multiple genes and environmental factors. My current work is based on the hypothesis that common cis-regulatory variants that influence gene expression by affecting transcription, RNA stability and/or mRNA processing contribute to susceptibility to mental disorders. Genome-wide association studies and linkage analyses have been widely used to identify genes or loci that contribute to mental illnesses, but, with a few notable exceptions, the yield of validated candidate genes has been disappointing: follow-up studies have often failed to replicate the reported associations. One of the major goals
of my research is to identify validated functional genetic variants. The use of these genetic variants in association studies should greatly increase the power to detect and replicate associations, and immediately provide molecular mechanisms to explain how the variants contribute to diseases.

In my studies, I examined the expression of mRNA encoding the serotonin transporter (SERT) and tryptophan hydroxylase 2 (TPH2), since changes in the expression of these genes has the potential to directly influence levels of serotonin, a neurotransmitter strongly implicated in major depression (MD), anxiety disorders and suicidal behavior. To identify cis-acting regulatory genetic variants that influence the expression of SERT and TPH2 mRNAs, I utilized a powerful and efficient method for measuring relative expression of mRNA from individual alleles. In the case of TPH2, I identified “yin” and “yang” haplotypes (comprising alleles of 5 TPH2 SNPs) that predict low- or high-expression of TPH2, respectively. Work to determine the functional genetic variant(s) that directly controls levels of TPH2 mRNA in brain is still in progress. In the case of SERT, I demonstrated that, contrary to expectations, the widely studied SERTLPR is not the primary regulator of SERT mRNA expression in the adult brains in our collection. Further studies will be need to determine if SERTLPR differentially regulates SERT mRNA expression during brain development or during times of stress.

I. Serotonin transporter

The SERTLPR has been extensively studied as a possible factor that contributes to mental disorders. Although many positive associations have been reported, especially for the “short-SERTLPR allele”, many additional studies have failed to replicate these findings. My study also did not find a correlation between SERTLPR and SERT allelic expression imbalance (AEI). Recent studies however, suggest that the SERTLPR cannot simply be characterized as being “short” or “long.” By contrast, the SERTLPR region has been determined to be highly
polymorphic: at least 14 variants have been identified, each occurring at different frequencies in different populations (Nakamura et al., 2000). Among these, SNP rs25531 (A/G), which is located within the SERTLPR, has been taken into account in many studies. Hu et al (Hu et al., 2005), recently reported that the A-allele of the “long-” SERTLPR variant is a stronger promoter than the G-allele. In fact, the LG variant has the promoter strength close to that of the “short-" SERTLPR, suggesting that the LG and S SERTLPR variants should be grouped together in association studies. The G-allele is predicted to alter the binding of the transcription factor activator protein 2. In addition, the G-allele showed the association with antidepressant response (Kraft et al., 2005). By contrast, the high expression LA allele was associated with response to alcohol (Hu et al., 2005). These results suggest that I should re-analyze my results, by including rs25531 genotype information. It should be noted, however, that taking into account the genotype of rs25531 will still not explain the lack of significant SERT AEI in adult pons. The possibility that the SERTLPR differentially regulates SERT mRNA expression during exposure to stress or during development of the brain will be discussed in below.

II. Tryptophan hydroxylase 2 (TPH2)

To identify the functional genetic variants that underlie TPH2 AEI, we plan to investigate the possible contribution of the follow molecular mechanisms.

1) Exon 7 skipping

We detected skipping of exon 7 in RNA samples from adult pons. Based upon the predicted effects of the G- and A-alleles of rs7305115 on the binding of splicing enhancer proteins, we expected that exon 7 would be preferentially skipped during processing of pre-TPH2 mRNA containing G-allele. A quantitative comparison of the relative amounts of splice variants in individuals harboring the A-allele or G-allele of rs7305115 was not possible, however, due to the extremely low levels of exon 7-skipped mRNA in human brain. In fact, exon 7-skipped
mRNA could only be reliably detected in a PCR-based assay (using forward and reverse primers flanking exon 7) when PCR products derived from the full-length transcript were eliminated by digestion with a restriction enzyme that cut within exon 7. As an alternative, the use of PCR primers sets where one of the primers anneal to the exon 6-exon 8 junction may provide a more sensitive method for detecting exon 7-skipped mRNA, and thus allow a comparison of skipping for mRNAs containing the rs7305115 A- or G-alleles.

Quantifying the true extent of exon 7 skipping in human brain, however, is confounded by the possible rapid degradation of exon 7-skipped mRNA by nonsense-mediated decay (NMD), a mechanism that has evolved to eliminate mRNAs encoding prematurely terminated proteins. Indirect evidence for (or against) extensive NMD of exon 7-skipped mRNA can be obtained by comparing the extent of TPH2 AEI measured using the exon 7 SNP rs7305115 versus AEI measured using the exon 9 SNP rs4290270. If exon 7 skipping significantly contributes to TPH2 AEI, but NMD does not occur, AEI ratios measured using rs4290270 would be expected to be significantly smaller than those obtained using rs7305115. By contrast, AEI ratios measured using rs7305115 and rs4290270 would be expected to be same, if the exon 7-deleted mRNA was extensively degraded. The AEI experiments described in Chapter 3 did, in fact, show a close correlation between measurements using each SNP (Fig 3.4). In these experiment, however, both sets of measurements were made using cDNA prepared from PCR primer sets containing a forward primer located within exon 7. To use AEI measurements to look for evidence of NMD, it will be necessary to repeat these experiments using PCR primers that flank exon 7.

To examine exon-skipping under conditions that permit experimental manipulation, we will express TPH2 minigenes containing the A- or G-allele in cells in tissue culture. An important caveat of these studies is that it is not known whether the in vitro system will faithfully replicates the pattern of exon skipping that takes place in brain. The expression of splicing proteins, including exonic splicing enhancers (ESEs) or ribonucleic proteins (RNPs) could be different in
different tissues, cell types or developmental stages. An advantage of this experimental system, however, is that will allows levels of mRNA splice variants to be examined under conditions where NMD is inhibited (by the addition of protein synthesis inhibitors, such as cycloheximide).

2) The effect of TPH2 SNPs on protein translation

Our TPH2 AEI data described in Chapter 3 suggested the presence of cis-acting polymorphisms that regulate mRNA expression. Another possibility that needs to be examined is the possible influence of specific SNPs on TPH2 protein expression. For this purpose, I will transfect cDNA constructs carrying different SNP alleles into cell culture system and quantify levels of TPH2 protein.

SNPs may influence translation by altering mRNA secondary structure. For example, Nackely and Shabalina (Nackley et al., 2006) showed that synonymous changes in the COMT coding region produced large reductions in COMT enzymatic activity by reducing protein translation. The SNPs in question (rs4633 and rs4840) reduced protein translation by altering critical stem-loop secondary structures in the COMT mRNA. Likewise, a synonymous SNP in DRD2 was also predicted to change mRNA folding, and thus cause a decrease in mRNA stability and protein translation (Duan et al., 2003). Other mechanisms by which SNPs can affect translation include interfering with the binding of ribosomal proteins or translational initiation factors or by switching codon usage from a plentiful to rare subtype of tRNA (Kimchi-Sarfaty et al., 2007).

3) Identification of functional intronic SNPs

If it turns out that the exon 7 SNP rs7305115 does not differentially affect TPH2 mRNA expression or protein translation, I will systematically examine the four intronic SNPs (rs2171363, rs4760815, rs6582078 and rs9325202) that were highly correlated with TPH2 AEI for effects on mRNA expression. Two of these SNPs (rs4760815 and rs6582078) are located very close to
exon-intron boundaries and may therefore affect mRNA splicing. I will examine the influence of each of these SNPs using the minigene expression system described above.

SNPs located within enhancer or silencer binding sites can also influence mRNA expression. I will use web-based tools to identify potential enhancer and silencer binding site in the vicinity of the TPH2 gene and look for overlap with identified SNPs. SNPs predicted to lie within these sites will be tested for their affect on mRNA expression in a heterologous expression system. The effects of SNPs on the binding of proteins to DNA will be tested using electrophoretic mobility shift assays. Intronic SNP can also differentially affect levels of pre-mRNA (hnRNA). A recent study showed that AEI measurements could be carried out on CYP3A4 hnRNA using intronic SNPs (Hirota et al., 2004). Although it may be difficult to isolate hnRNA from postmortem tissues, it may still be useful to attempt to make these measurements.

4) Clinical association study with TPH2 SNPs

Since the discovery of TPH2 gene in 2003, many studies have examined the potential association of TPH2 genetic variants with mental disorders. Association of specific SNPs in human TPH2 have been reported with major depression (Zill et al., 2004b; Zhang et al., 2005a; Van Den Bogaert et al., 2006), affective disorders (Harvey et al., 2004; Lopez et al., 2007), suicidality (Zill et al., 2004a; Ke et al., 2006; Van Den Bogaert et al., 2006; Jollant et al., 2007; Lopez de Lara et al., 2007; Lopez et al., 2007), autism (Coon et al., 2005), early onset obsessive-compulsive disorder (Mossner et al., 2006a), attention deficit hyperactivity disorder (Sheehan et al., 2005; Walitza et al., 2005) and panic disorder (Maron et al., 2007).

Polymorphisms in the TPH2 regulatory region have also been reported. A common promoter polymorphism -703G/T (rs4570625) was reported to be associated with ADHD (Walitza et al., 2005), amygdala responsiveness (Brown et al., 2005; Canli et al., 2005), emotional processing (Herrmann et al., 2007), personality traits and disorders related to emotional dysregulation (Gutknecht et al., 2007; Reuter et al., 2007), while another common polymorphism
-473T/A (rs11178997) in the 5’ regulatory region was associated with ADHD (Walitza et al., 2005), and unipolar and bipolar disorder (Van Den Bogaert et al., 2006). The SNP rs7305115, which is associated with \( TPH2 \) mRNA expression in my study, has shown association with suicide and major depression in several clinical studies.

Ke L et al reported the association of SNP rs7305115 with suicide behavior in major depression patients. The \( A \)-allele, associated with higher \( TPH2 \) mRNA expression in my study, was significantly less frequent in attempters that in non-attempters (\( p=0.0067 \)) suggesting protective effect of \( A \)-allele against suicide attempt in major depression patients. An association study with haplotype tagging SNPs identified a TPH2 haplotype block that correlated with protection against MD (Van Den Bogaert et al., 2006). By extrapolation, this haplotype block was determined to contain the \( A \)-allele of SNP rs7305115

Taken together, many studies have demonstrated a possible association of \( TPH2 \) SNPs with mental disorders. However, the statistical power of most of these studies was low, In addition, different studies identified different sets of associated SNPs and haplotypes. A possible explanation for the weak associations reported in many studies is that functional polymorphisms were not included: i.e., the genetic markers used were only in partial linkage with functional genetic variants. The \( TPH2 \) SNPs we identified in our study are good candidates for association study because they are highly predictive of levels of \( TPH2 \) mRNA expression. In particular, we have identified specific alleles that correlate with low \( TPH2 \) mRNA expression, and these are likely to be present in disorders associated with low serotonin levels. Perhaps the most important next-step in this research will be to test this hypothesis in association studies.

**III. Developmental role of serotonin**

Serotonin has been reported to regulate brain development as a trophic factor before it begins to function as a neurotransmitter. In the human brain, serotonergic neurons appear by the
5th week of gestation and continue to differentiate through the 10th week of gestation. Serotonergic neurons are found almost exclusively within the raphe nuclei of the pons and brainstem. Clusters of these neurons can be divided into two groups based on their distribution, developmental origin, function and major projection. The rostral serotonergic neurons located in the midbrain and upper pons send their axons to the cerebral cortex, thalamus, hypothalamus, basal ganglia, hippocampus and amygdala. The caudal cell bodies in the medulla and caudal pons project to other brain stem sites, cerebellum and spinal cord. This group modulates respiratory chemosensitivity, cardiovascular function, thermoregulation, upper airway reflexes, motor activity, pain and arousal.

Multiple studies have suggested a developmental role for serotonin. Impairment of serotonergic transmission during development may predispose individuals to increased risk of developing mental disorders. For example, transient inhibition of SERT with the SSRI fluoxetine during early development gave rise to abnormal emotional behaviors in adult mice (Ansorge et al., 2004). Expression of serotonin 1A receptor during critical time of development was sufficient to rescue the behavioral phenotype of the knockout mice (Gross et al., 2002).

As discussed above, we did not observe a correlation between SERTLPR and SERT AEI in adult human pons. One possibility that remains to be examined, however, is that the SERTLPR plays a role in controlling mRNA expression during development of the brain. Individuals with short-allele may have excess of serotonin during development, which can possibly influence the formation of brain circuits related to emotion and anxiety. Effects of this abnormal brain development might predispose individuals to mental disorder later in life.

To examine the possibility that SERT mRNA is differentially regulated by the SERTLPR in the developing brain, we measured SERT AEI in RNA samples isolated from the rostral pons of ten fetal or neonatal brains. Consistent with the results from adult brains, only minor AEI was detected and this did not correlate with SERTLPR alleles. Although more samples must be
examined before reaching a conclusion, these preliminary results suggest that the SERTLPR may not differentially regulate SERT mRNA expression during development. Another possibility that must be examined is whether the SERTLPR regulates SERT mRNA expression during times of stress. To examine this possibility, we plan to examine SERT AEI in the pons of suicide victims, who chose violent means of death. These individuals were presumably highly stressed at the time of death and may therefore provide a unique window into the regulation of this important gene.
LIST OF REFERENCES


- 132 -


