ASSOCIATION STUDY OF TWO *SLC6A4* POLYMORPHISMS WITH AUTISM

By

Jacquelyn Ann Recktenwald

Thesis

Submitted to the Faculty of the

Graduate School of Vanderbilt University

In partial fulfillment of the requirements

For the degree of

MASTER OF SCIENCE

in

Interdisciplinary Studies: Neurogenetics

December, 2007

Nashville, Tennessee

Approved:

James S. Sutcliffe

Randy D. Blakely

To my parents for unending support and encouragement
David and Janet Recktenwald

To my fiance for cheering me on
Douglas Waters

ACKNOWLEDGEMENTS

The project presented here was supported by a Neuroscience Training Grant, Vanderbilt University, under the direction of Dr. Elaine Sanders-Bush.

I would like to recognize the staff in the Vanderbilt Center for Human Genetics Research, particularly Kim Gainer and Lan Jiang in the Data Analysis Core and Genea Crockett and Kelly Taylor in the Family Ascertainment Core for all their help with this project.

I would also like to acknowledge and thank the numerous families who have contributed to make this work possible. I acknowledge contributing families ascertained by our collaborators at the Tufts University/New England Medical Center under the direction of Susan Folstein, those families ascertained through Vanderbilt University, those families in the AGRE (the Autism Genetic Resource Exchange) consortium, and those contributing through our collaborations with Stanford University Medical Center and Iowa State University Medical Center.

I would like to thank the members of my thesis committee, Dr. James S. Sutcliffe and Dr. Randy Blakely, for their guidance in and outside of the classroom and support in this project. I would further like to acknowledge all the members of the Sutcliffe lab (Ryan Delahanty, Joshua Smart, and Emily Crawford) for their encouragement and friendship both in the lab and outside of the lab.

Finally, I would like to thank you, Jim (A.K.A. Dr. James Sutcliffe), for allowing me the opportunity to be a part of your lab and the research you are doing to dissect the genetics of Autism. Your enthusiasm for and dedication to genetics research is inspiring

and admirable. Your gift for teaching is obvious; I have benefited greatly from your course on Neurogenetics not just by way of your solid teaching but also by the contagiousness of your enthusiasm for genetic research. Thank you for letting me complete my thesis work under your guidance and I wish you the best in all your future research.

LIST OF ABBREVIATIONS

5-HT	5-Hydroxytrptamine/ Serotonin
5-HTT/SERT/SLC6A4	Serotonin Transporter
ADI/ADI-R	Autism Diagnostic Interview- Revised
ADOS	Autism Dianostic Observation Schedule
AGRE	Autism Genetic Resource Exchange
ASD	Autism Spectrum Disorder
BAP	Broader Autism Phenotype
bp	Base pair
cM	CentiMorgan
dNTPS	Deoxyribonuclotides
CSD	Cold shock domain
CTCF	CCTC-binding protein
DAT	Dopamine Transporter
DNA	Deoxyribonucleic Acid
DSL	Disease Susceptibility Locus
DSM-IVR	Diagnostic and Statistical Manual of Mental Disorders IV- Revised
DZ	Dizygotic twins
EMSA	Electrophoretic Mobility Shift Assay
ES	Embyronic Stem Cells
FBAT	Family Based Association Tests
HBAT	Haplotype Based Association Tests

HLOD	Heterogeneity LOD Score
HWE	Hardy-Weinberg Equilibrium
kb	Kilobase
LD	Linkage Disequilibrium
LIF	Leukemia Inhibitory Factor
LOD	Logarithm of the Odds
Mb	Megabase
MAF	Minor Allele Frequency
MAPK	Mitogen-Activated Protein Kinase
ml	Milliliter
mRNA	Messenger Ribonucleic Acid
MZ	Monozygotic Twins
NEMC	New England Medical Center
NET	Norepinephrine Transporter
ng	Nanogram
NIMH	National Institute of Mental Health
OMIM	Online Mendelian Inheritance in Man
PCR	Polymerase Chain Reaction
PDD	Pervasive Developmental Disorders
PDT	Pedigree Disequilibrium Test
PET	Positron Emission Tomography
PKC	Protein Kinase C
PKG	Protein Kinase G

PMID	PubMed-Indexed for MEDLINE
QTDT	Quantitative Trait Disequilibrium Test
QTL	Quantitative Trait Locus
RNA	Ribonucleic Acid
SNP	Single Nucleotide Polymorphism
SSRI	Selective Serotonin Reuptake Inhibitor
STin2	Serotonin Transporter Intron 2 polymorphism
T_A	Annealing Temperature
Taq	*Thermus Aquaticus Polymerase*
TD	Transmission Disequilibrium
TDT	Transmission Disequilibrium Test
TPH	Tryptophan Hydroxylase
UTR	Untranslated Region
VNTR	Variable Number Tandem Repeat
YB-1	Y box binding protein
βgal	β-Glalactosidase
μg	Microgram

TABLE OF CONTENTS

Page

Chapter

CHAPTER 1

INTRODUCTION

Autism and the Broader Phenotype

In 1943 Leo Kanner published a series of case reports that detailed his observations of eleven children presenting with what he called, "infantile autism". He distinguished these children as distinct from individuals with mental retardation on the basis of the lack of social reciprocity they displayed. Most of the eleven children he characterized were males, in fact, only three were females.[1] Kanner was first to publish features characteristic of the autism phenotype. His research laid the ground-work for further study, assessment and diagnosis of the autism phenotype.

Autism (OMIM #209850) is now recognized as a formal disorder and is characterized as part of a group of developmental disorders called Pervasive Developmental Disorders (PDDs), as classified by the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV). This disorder is distinguished from other PDDs, such as Rett Syndrome, Asperger Syndrome and PDD-Not-Otherwise-Specified (PDD-NOS), on the basis of diagnostic criteria and assessment measures as outlined by the DSM-IV, the Autism Diagnostic Interview and its revision (ADI, ADI-R)[2, 3] and the Autism Diagnostic Observation Schedule (ADOS).[4] Autism is characterized by impairments in three main areas: development and use of language, reciprocal social interaction and repetitive behaviors, restricted interests and resistance to change in the environment or in daily routines.[5] Autism represents a continuum of behavioral abnormalities and is more

accurately defined as a spectrum of disorders (Autism Spectrum Disorders, ASDs), with "classic autism" representing the more severe end of the spectrum where individuals meet diagnostic criteria defined by levels of impairment in the triad of symptom domains associated with the disorder. The presentation of impairments in individuals with autism is highly variable. Additional behavioral complications such as irritability, aggression, anxiety, hyperactivity and compulsive behavior may be present and the manifestation of impairments changes as a function of time.[6] Further, cognitive impairments exist in more than two thirds of affected individuals.[7] One study found a mental retardation rate of 66.7% in individuals with more narrowly defined autism.[8] In addition, gastrointestinal disturbances such as diarrhea, abdominal pain and/or constipation are frequently present in individuals with autism.[9] Seizures are a frequent co-morbid condition in autism; various studies have found the frequency of epilepsy in autism to be anywhere from 5% - 38.3% (reviewed in reference 10).[10]

Prevalence and Heritability of Autism

The current estimated prevalence of classic autism is 1/500, which increases to 1/150 (PMID: 17287715) when all PDDs (not Rett Syndrome) are included.[7, 8] Males are affected four times as frequently as females.[8] Twin studies lend strong support for genetic determinants in autism: monozygotic twins have a concordance rate of 60% for classic autism and up to 92% when broader criteria are used, as compared to dizygotic twin concordance rates of 0% and 10% for classic autism and more broadly defined impairments, respectively.[11, 12] In addition, sibling recurrence rates are estimated to be ~6-8% when "stoppage rules" (when the parents of a child with autism decide not to have

any more children) are taken into consideration.[13, 14] This rate is much greater than would be expected based strictly on the population prevalence of autism, which is 1/500 for narrowly defined autism. It has been estimated that there are likely 15 or more genetic loci that contribute to risk for autism.[15] Thus, it seems that the mode of inheritance is likely oligogenic with different families possessing different sets of risk alleles,[5] thus contributing to the range of phenotypes seen within autism.

Serotonin in Autism

Serotonergic dysregulation has long been suspected to play a role in autism. In 1961 Schain and Freedman discovered about one-third of individuals with autism have elevated blood platelet serotonin levels (hyperserotonemia)[16] and this finding has been replicated by other groups.[17] In addition, selective serotonin reuptake inhibitors (SSRIs) are frequently effective in the treating of ritualized, repetitive behaviors, stereotypies, anger, and anxiety associated with autism.[18-20] Further, when the essential precursor of serotonin (tryptophan) is depleted, autistic symptoms in affected individuals worsen.[21] In addition, a positron emission tomography (PET) study performed on a cohort of boys with autism revealed a decrease in serotonin synthesis in the left thalamus and left frontal cortex in five of the seven autistic boys studied and in the right thalamus and right frontal cortex of the other two boys (the one female with autism did not have these asymmetries).[22] PET studies also reveal a developmental difference in the capacity for serotonin synthesis between affected and unaffected individuals.[6, 23] In a normal individual, serotonin synthesis in the brain is equal to or greater than 200% that of adult synthesis up until approximately five years of age, and then levels begin to decline

gradually until reaching adult levels.[6, 23] In individuals with autism, serotonin synthesis gradually rises from age two until approximately age fifteen at which point it levels off at 150% that of normal adult values. [21] Thus some individuals with autism from the age of two until five (and likely under two) have less capacity for central nervous system serotonin synthesis than children without autism.

These studies indicate a role for serotonin in autism, but the argument could be made that serotonergic dysregulation is secondary to autism rather than etiological. Therefore studies looking at serotonin levels in unaffected first degree relatives are advantageous. One study found a positive correlation between serotonin levels in subjects with autism and their unaffected first-degree relatives.[24] Further, affected individuals with an affected sibling have greater serotonin levels than those without an affected sibling.[25] In addition, rat pups injected with a serotonin analog during development exhibit behavioral and neurochemical features reminiscent of autism.[26] The investigator of this study hypothesizes that elevated blood serotonin levels will lead to a decrease in the brain serotonin neurocircuitry by activating negative serotonergic feedback loops prematurely.[26] This would lead to the decreased serotonin neurocircuitry found in the brains of individuals with autism. Further, in consideration of findings of hyperserotonemia in ~25% of individuals with autism, it may be noted that higher levels of serotonin in blood platelets may, conceptually, be considered as equal to lower amounts of serotonin signaling in the brain since more is being retained within the neuron. These studies are supportive of the idea that serotonin dysregulation is etiological rather than secondary to autism. Therefore any genes whose protein products

play a role in serotonergic signaling are functional candidate genes for autism susceptibility.

Genetic Studies of Autism

To study genetic determinants of autism risk, several groups have employed genome-wide linkage screens using multiplex families to detect regions of the genome that tend to be inherited in the affected offspring, and/or utilized allelic association tests to determine if common alleles at a given locus confer increased risk for autism. Genome-wide linkage screens based on categorical phenotypes within autism have revealed significant linkage peaks at chromosomal regions 7q, 2q and 17q.[15, 27-42] Follow-up studies for the 17q region with independent samples demonstrated significant linkage at or close to the serotonin transporter (SERT) gene (*SLC6A4*), located at 17q11.2.[38, 43]

SLC6A4 as a Functional Candidate Gene

The *SLC6A4* gene is a functional candidate gene in autism as its protein product, SERT, is a key regulator of serotonin levels in the presynaptic neuron terminals and post-synaptic synapses of serotonergic neurons which project from the raphe nuclei of the brainstem to cortical areas (reviewed by Purselle and Nemeroff, 2003).[44] In addition, SERT is expressed on thalamocortical neurons for a short period during development.[45] Further, SERT is expressed in blood platelets where serotonin is taken up and stored in dense granules for later release in hemostasis (reviewed by Horiuchi, 2006).[46] *SLC6A4* is a member of the *SLC6A* gene family of Na+ and Cl- coupled transporters and is most similar to the norepinephrine transporter (NET) and dopamine transporter (DAT).[47]

SLC6A4 is expressed from a single gene that undergoes no alternative splicing of its coding exons thus the translated 630 amino acid SERT protein expressed in the brain is identical to that expressed in the periphery in blood platelets and lymphocytes.[47-49] More than twenty polymorphisms in the coding region have been identified in *SLC6A4*. SERT has twelve transmembrane-spanning domains[48] with intracellular NH_2 and COOH termini.[50-52] Further, SERT is rapidly modulated by protein kinase G (PKG) and p38 mitogen-activated protein kinase (MAPK) pathways in addition to various G-protein coupled receptors.[47, 53-57] Internalization of SERT can be mediated by activation of protein kinase C (PKC) pathways, and this is dependent on extracellular serotonin levels.[54]

Increased Linkage at 17q11.2 Based on Sex and Phenotype

Importantly, the strength of genetic linkage to 17q11.2 is dependant on sex and phenotype.[39, 58] Our lab did a study of 137 multiplex families that revealed a multipoint heterogeneity logarithm of the odds (HLOD) score of 2.74 which increased to 3.62 in a subset of families that display comparatively more severe rigid-compulsive behaviors.[39] A further study of 341 multiplex families revealed that when families with affected males only (MO; n=202) were analyzed independently of families with at least one affected female (female-containing, FC), the peak recessive HLOD score at chromosome 17q11.2 went from 5.8 to 8.0, with little contribution from the female-containing families.[58] An HLOD score of 8.0 for a complex psychiatric disorder is striking. In Cystic Fibrosis, there is some allelic heterogeneity, but there is a single, more prominent mutation, ΔF508, that only has an HLOD score of 6.0.[59]

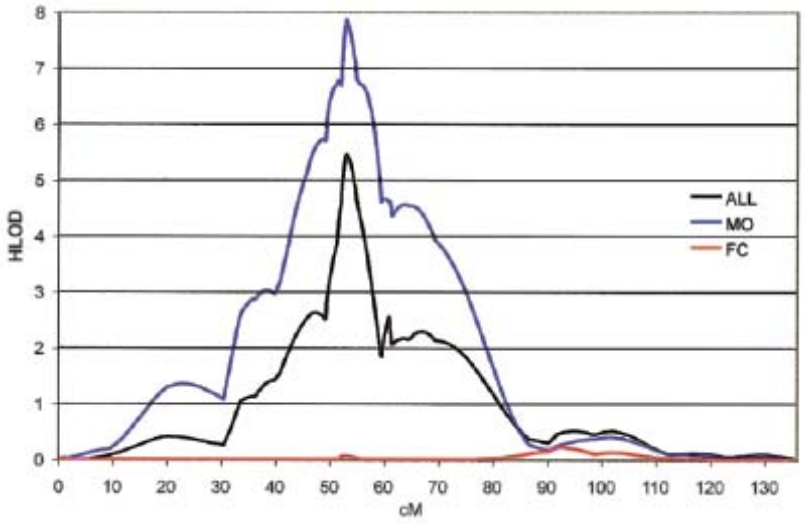


Figure 1. Linkage of autism at 17q11.2 is predominantly driven by male-only families. Recessive multipoint HLOD scores are plotted along chromosome 17 as a position of the marker in centimorgans (cM) for the pooled autism dataset (All; n=341; black line), families in which only males are affected with autism (MO; n=202; blue line), and families in which at least one female is affected with autism (FC; n=138; red line). Peak HLOD score for the pooled dataset is 5.8 and increases to 8.0 in MO families.[58] Figure taken from Sutcliffe et al. (2005).[58]

Rare Variants in *SLC6A4*

Our lab went on to probe the possibility that rare or novel variants of *SLC6A4* in contribute to autism susceptibility. We used allelic discrimination assays to test for the presence of known, rare variants in autism families. We also selected 120 unrelated probands who contributed most to the linkage peak at 17q11.2 and screened their *SLC6A4* promoter and exons for novel variants.[58] The Gly56Ala variant was identified in multiple affected families, particularly within the "linked" subgroup, and included three homozygous individuals. Thus in the "linked" families, Gly56Ala was present with a minor allele frequency (MAF) of ~2.3%, whereas a previous study reported a MAF of ~0.44% in a nonclinical sample.[58, 60] In addition, three novel variants were identified in

the screen (Ile425Leu, Leu550Val, and Phe465Leu), all of which are coding substitutions that occur in transmembrane domains at highly conserved residues.[58]

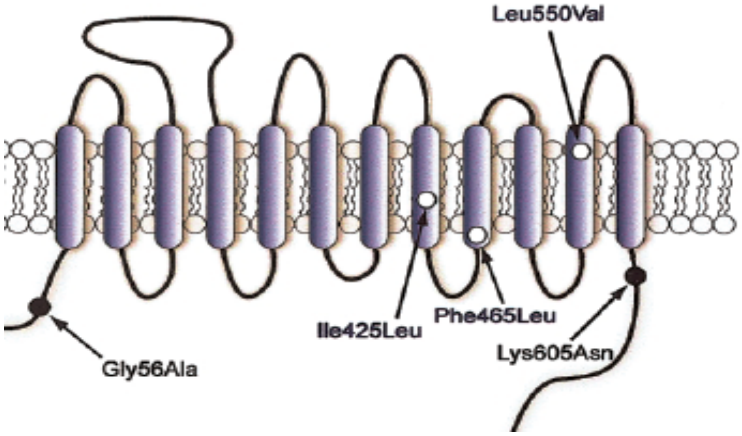


Figure 2. Schematic representation of rare variants discovered in SLC6A4 of affected families. Three of the variants are located within transmembrane domains (white dots; Ile425Leu, Phe465Leu, Leu550Val) and two are located in the intracellular amino and carboxyl terminal regions (black dots; Gly56Ala and Lys605Asn, respectively). Gly56Ala has previously been reported to exhibit a dose-dependent increase in basal 5-HT transport activity and to be resistant to PKG- or MAPK-stimulated 5-HT uptake. Ile425Leu occurs at a nucleotide that has also been affected by a previously reported polymorphism (Ile425Val) where the minor allele tracks with individuals presenting with a complex psychiatric phenotype including OCD and Asperger syndrome. These novel variants on SERT exhibit a gain-of-function phenotype. Figure taken from Sutcliffe et al. (2005).[58]

Gly56Ala SERT exhibits a dose-dependant increase in basal 5-HT transport activity in EBV-transformed lymphocytes and transfected cells with Ala56 homozygous cell lines having an ~75% increase compared to Gly56 homozygous cell lines.[58] The Ala56 SERT is also refractory to serotonin uptake stimulation induced by acute application of PKG or p38 MAPK activators.[47, 58] Ile425Leu occurs at a nucleotide that was previously identified as a variant and leading to an Ile425Val substitution in two families segregating a complex neuropsychiatric phenotype including obsessive-

compulsive disorder (OCD) and Asperger syndrome.[61] Cells expressing Val425 SERT exhibit significantly increased surface expression of this protein resulting in elevated basal serotonin uptake.[47] Val425 encoded SERT does not display altered regulation by kinase activation as Gly56 encoded SERT does, but both protein variants do exhibit elevated basal serotonin uptake.[47, 58] The three newly identified variants in addition to Gly56Ala are associated with more severe rigid-compulsive behaviors.[58]

SLC6A4 Repeat Polymorphisms: HTTPLR and VNTR

The most commonly studied variants of *SLC6A4* are two repeat polymorphisms located in noncoding regions of *SLC6A4* that affect gene expression.[62, 63] The first is a 44 base pair (bp) insertion/deletion polymorphism (HTTLPR) located 1.2 kb upstream of the first exon (1a) that results in a long (*l*) and a short (*s*) allele.

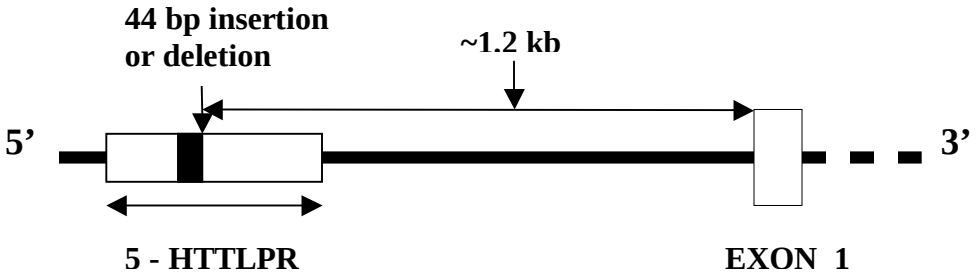


Figure 3. Schematic representation of HTTLPR. The 44 bp insertion/deletion polymorphism is located approximately 1.2 kb upstream of the first exon of *SLC6A4*. The presence of the 44 bp insertion results in a long allele (*l*). Deletion of the 44 bp region results in a short allele (*s*). The *l* allele is associated with increased expression of SERT. (kb) kilobases, (bp) base pair.

In vitro lymphoblast studies demonstrate the *s* allele of HTTLPR decreases *SLC6A4* promoter transcriptional efficiency resulting in a lower concentration of SERT

and a reduction in the uptake of serotonin.[62] In addition, *in vivo* imaging studies have found a significant increase in raphe SERT levels for *l/l* homozygous individuals as compared to levels in carriers of the *s* allele.[64]

A number of studies have been done looking at association of the *s* allele with various personality traits or preponderance to psychiatric symptoms. For instance, one study found the *s* allele to be associated with neuroticism.[65] Another study found that carriers of the *s* allele have an exacerbated response toward developing depressive symptoms as a result of tryptophan depletion regardless of family history of depression; although, in heterozygous individuals a family history of depression resulted in an increased tendency toward depressive symptoms comparable in degree to homozygous carriers of the *s* allele.[66]

Association Studies of HTTLPR Polymorphisms with Autism

Studies looking at association of either HTTLPR allele with autism have been largely inconclusive. Some groups have found positive association of the *l* allele with autism[67, 68], some with the *s* allele with autism,[39, 69-71] and a few studies found no association of either allele with autism.[72-75] Reasons for this discrepancy may be the low degree of statistical power resulting from small sample sizes in these studies, the difficulty of genotyping this region due to the nature of it being a very GC rich segment, and/or differing populations being genotyped and thus differing risk alleles. Of importance, the differing risk alleles would likely be associated with different aspects of the autism phenotype, different developmental patterns and/or different traits represented

differentially in various autism samples. In addition, diagnostic criteria and ascertainment measures may differ amongst the various groups performing the studies.

VNTR

The second polymorphism of *SLC6A4* is a variable nucleotide tandem repeat (VNTR) of a 16-17 bp element located in intron 2 of *SLC6A4* (STin2, Serotonin Transporter intron 2). There are typically 9, 10 or 12 repeats of this element commonly referred to as STin2.9, STin2.10, STin12, respectively. VNTR variants demonstrate differential expression both *in vitro* and *in vivo*.[63, 76, 77] Fiskerstrand and colleagues performed a study in which they transfected luciferase constructs containing Stin2.10 and Stin2.12 (they could not produce a stable construct of STin2.9) into embryonic stem (ES) cells and looked at the relative expression differences between the two alleles after removing LIF (Leukemia Inhibitory Factor), which is known to keep the cell in an undifferentiated state. STin2.10 increased expression of the luciferase construct by a 5.77 fold increase, demonstrating an ability of STin2.10 to enhance gene expression of a reporter gene. STin2.12, however, enhanced expression of the reporter gene by a 167.85 fold increase.[76] However, these results do appear to be cell specific as STin2 expression constructs do not have observable effects in HeLa cells.[76]

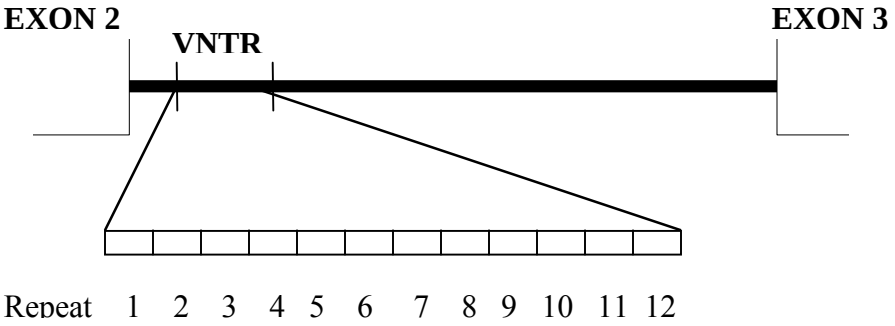


Figure 3. Schematic representation of the VNTR in *SLC6A4*. This polymorphic region is located between the second and third exon of *SLC6A4*. Depicted are repeats one through twelve of the 16-17 bp element and thus a representation of STin2.12. Removal of repeats nine and ten would represent Stin2.10 while removal of repeats six, seven, and eight would result in a representation of STin2.9. STin2.12 is associated with increased expression of SERT. This figure is adapted from Klenova et al. 2004.[78]

To examine STin2 allelic effects *in vivo*, the same group performed pronuclear injections of expression constructs of Stin2.10 and Stin2.12 variants placed upstream of a heterologous promoter driving LacZ into mouse embryos and looked at β-galactosidase (βgal) expression in the developing murine brain. At embryonic day 10.5 (E10.5) βgal expression was observed in two regions on either side of the midbrain, neural tube floor plate and floor plate of the hindbrain, particularly in rhombomeres four and five for both constructs. However, differential βgal expression patterns were seen in the rostral portion of the hindbrain in the regions of rhombomeres one and two; Stin2.10 animals had low levels of βgal expression in these rhombomeres whereas those with Stin2.12 had high levels of βgal expression in these rhombomeres, comparable to levels of βgal in rhombomeres four and five of both variants.[63] This study suggests that not only do the VNTR polymorphic variants STin2.10 and STin2.12 exhibit differential enhancement

capabilities of reporter gene expression within a given cell type, but may also exhibit differential effects on spatial patterns of reporter gene expression.

Association Studies of the Intron 2 VNTR with Autism

A number of studies testing association of STin2 polymorphisms with autism have been undertaken. Two studies found no preferential transmission of any STin2 polymorphism with autism.[72, 73] A German study also did not find preferential transmission of STin2 alleles with autism, but did find evidence for association of a haplotype containing the *l* allele of HTTLPR and STin2.12.[67] One early study by the Cook laboratory found no preferential transmission of any STin2 polymorphism with autism, but did find evidence of significant association of a haplotype containing the *s* allele of HTTLPR and STin2.12.[59] Subsequently, this group tested 81 new trios and obtained the same results of preferential transmission for a haplotype containing the *s* allele of HTTLPR and STin2.12.[74]

One study of 125 Dutch patients with DSM-IV-TR, ADI-R or ADOS characterized PDD did not find association between STin2 alleles and PDD, however, taking specific autism phentoypes into account by using multivariate analysis to test the continuum of behavioral subtypes revealed an association between individuals with the homozygous STin2.12 genotype and more severe rigid-compulsive tendencies. Similarly, quantitative transmission disequilibrium test (QTDT) analysis revealed an association between the STin2.12 allele and more severe rigid-compulsive tendencies.[79] The 2006 study mentioned above by Cook and colleagues tested for association between specific autism behavioral subcategories (from ADI-R or ADOS) and STin2 genotype, but found

no significant associations.[80] Finally, one Irish study of 84 families found excess transmission to autistic probands for multiple haplotypes of *SLC6A4*, the most significant being a haplotype defined by three polymorphisms including the STin12 allele.[70]

Describing the Autism Phenotype

Lord and colleagues (1994) describe the ADI-R as an effective diagnostic measure that is used to determine if behaviors commonly associated with autism or PDDs are found in the individual being examined and to evaluate how severe each behavior might be.[3] This tool is used widely and has been studied extensively to assess and ensure its accuracy and relevancy. It differs from the ADOS in that the primary caregiver of the individual who is presenting with these behaviors is interviewed by a trained investigator. The assessment is not minor, there are over 100 items on the partially-structured diagnostic and a certain threshold must be reached in four key areas for a diagnosis of autism to be made: communication, social interaction, repetitive behaviors, and age at onset of some symptoms.[3, 81] Tadevosyan-Leyfer and colleagues performed a principle components analysis of items pertaining to the ADI-R providing a useful set of key components characteristic of the autism phenotype.[81] These six factors can be used to mathematically approximate the nature of an affected individual's phenotype.[81] While the autism phenotype is described as being comprised of three key features (development and use of language, reciprocal social interaction and repetitive behaviors, restricted interests and resistance to change in the environment or in daily routines.[5]) that range in severity from person to person, these three categories likely do not accurately represent genetically relevant categories. Thus, the value of a principle components analysis based

on a thorough evaluation of the autism phenotype, such as the ADI-R, is the ability to parse out a group of components that adequately describe the autism phenotype and delineate its features into categories that are likely to be genetically relevant.[81] These components are: spoken language, social intent, compulsions, developmental milestones, savant skills and sensory aversions.[81] The first three components, spoken language, social intent and compulsions are more closely related to the three categorical features used in diagnosing autism.[81] Spoken language relates to how well, if at all, an individual can produce verbal output, and does not include aspects of language related to individual expression of a feeling, idea or desire.[81] The second component, social intent, does include these aspects of language (including gesturing and greeting) in addition to nonverbal communication and ability to relate socially.[81] The third component, compulsions, reflects the repetitive behaviors, restricted interests and resistance to change aspects used to diagnose autism, but does not include motor mannerisms that are complex as these characteristics seem to be more related to features of social interaction.[81] In fact, in a factor analysis study done by Cuccaro and colleagues 'repetitive sensory motor actions' was identified as a category distinct from 'resistance to change'.[82] The first three components in the principle components analysis are necessary for delineating diagnostic criteria for autism, but they are not sufficient for a holistic description of the autism phenotype. The second three components, developmental milestones, savant skills and sensory aversions, vary from individual to individual to a much greater extent than the first three components but are important aspects of the autism phenotype at large and thus important to genetic studies of autism.[81]

Hypothesis

Autism is a highly heritable disorder. Given the multiple lines of evidence supporting serotonergic dysregulation in autism etiology, along with linkage studies showing evidence for linkage at the *SLC6A4* chromosomal region, I aim to dissect the role two repeat polymorphisms of SERT may play in contributing to susceptibility for autism. The goal of my study is to determine if alleles at STin2 and/or HTTLPR individually or as haplotypes are associated with autism susceptibility in our sample. I hypothesize that preferential transmission will exist for HTTLPR and VNTR alleles, as well as a haplotype of these markers to individuals with autism in our sample.

CHAPTER II

MATERIALS AND METHODS

DNA Sample

The sample for this study consisted of 694 combined multiplex and parent-child trio families (see Table 1). The families were ascertained from multiple sources: 151 families from the Tufts-Vanderbilt Consortium, 327 from the AGRE Consortium, 131 from Stanford University and 85 from the University of Iowa. Within each family is at least one proband who demonstrates full ADI criteria for autism and at least one other sibling who also met full ADI criteria for autism or presented on the broader spectrum.[58] The ADI or ADI-R were used to assess each affected individual. In addition, data was analyzed based on individualized scores for the six principle components of the autism phenotype as based on analysis by Tadevosyan-Leyfer and colleagues.[81]

In our ascertainment, subjects that had aberrant karyotypes, a genetic disorder of known causation, dysmorphic features or having a diagnosis of Fragile X syndrome were not included. Lymphoblastoid cells, buccal cells or peripheral blood was used to isolate DNA from each individual, following the protocol recommended by the Pure Gene manufacturer (Gentra Systems, Minneapolis, MN).

Table 1. Genotyped families with autism organized by center. Labels are as follows: (AGR) Autism Genetics Resource Exchange Consortium, (IOW) Iowa State University Medical Center, (STA) Stanford University Medical Center, (TUF) Tufts University Medical Center, (VAN) Vanderbilt University Medical Center, (ADI) Autism Diagnostic Interview, (ADI-R) Autism Diagnostic Interview-Revised. Stanford University Medical Center used an alternate form of the ADI for which not all questions were available when characterizing ascertained subjects.

	ALL	AGR	IOW	STA	TUF	VAN
Families	694	327	85	131	98	53
Individuals	2823	1420	361	522	353	167
Number with autism	1256	623	162	259	153	59
Female	268	147	26	52	34	9
Male	988	476	136	207	119	50
Number without diagnosis	1567	797	199	263	200	108
Female	814	416	103	139	102	54
Male	752	381	96	124	98	53
Number of Families with:						
0 affected	26	15	1	0	9	1
1 affected	143	36	19	16	27	45
2 affected	471	244	55	105	60	7
3 affected	45	29	7	7	2	0
4 affected	9	3	3	3	0	0
ADI version		ADI / ADI-R	ADI	ADI*	ADI	ADI / ADI-R

Determination of Genotype

Genotypes for HTTLPR were determined using PCR followed by gel-based size discrimination of alleles. Gels were 3% NuSieve (3:1) agarose (FMC Bioproducts; Rockland, ME), which allowed for ample discrimination between the longer allele (525 bp) versus the shorter allele (484 bp). The following amplifying PCR primers were used: 5'-CTGAATGCCAGCACCTAACCCCTAATGT-3' and

5'-GGGGAATACTGGTAGGGTGCAAGGAGAA-3.' "

PCR reaction volumes for HTTLPR were 20 µl, with 40 ng of DNA template and 0.8 µl of 10 picomolar (pM) primer solution. The Eppendorf Mastermix 2.5X kit

(Eppendorf, North America) supplied the buffer, dNTPS and Taq. Cycling conditions consisted of an initial denaturing step at 95 ºC for three minutes followed by 50 cycles of 95 ºC for 30 seconds, 61 ºC for 30 seconds, and 72 ºC for one minute. A final extension step was 72 ºC for seven minutes.

Genotypes for VNTR were also determined using PCR followed by gel-based size discrimination of alleles. Gels were 3.5% NuSieve (3:1) agarose, which allowed for clear discrimination between the three STin2 alleles: STin2.12 (390 bp), STin2.10 (360 bp) and STin2.9 (345 bp). The following amplifying PCR primers were used:

Forward: TGGATTTCCTTCTCTCAGTGATTGG and

Reverse: TCATGTTCCTAGTCTTACGCCAGTG

PCR reaction volumes were 20 µl, with 20 ng of DNA template and 0.8 µl of 10pM primer solution. The Eppendorf Mastermix 2.5X kit (Eppendorf, North America) supplied the buffer, dNTPS and Taq. Cycling conditions consisted of an initial denaturing step at 95 ºC for three minutes followed by 50 cycles of 95 ºC for 30 seconds, 64 ºC for 30 seconds, and 72 ºC for one minute. A final extension step was 72 ºC for seven minutes.

Statistical Analysis

The first step of analyzing the HTTLPR and VNTR data involved quality control checks to verify internal controls and to be sure our data was consistent with Mendelian inheritance. In addition, we verified the genotyping efficiency was robust.

As reviewed by Laird and Lange (2006), the most common and simplest family based association test is the transmission disequilibrium test (TDT).[83] This test uses trio

or multiplex families to test whether there is an increased transmission of a given allele or alleles to the affected offspring. In rejecting the null hypothesis that no preferential transmission of alleles exists in affected individuals, the marker can be assumed to be in linkage disequilibrium (LD) with a disease susceptibility locus (DSL).[83] The TDT is limited in its use however, particularly with cases of missing parents, complex diseases, or cases where the exact allele representation on homologous chromosomes is not specified (missing phase).[83] An extension of the TDT which allows for testing of association despite these factors is the non-parametric family based association tests, or FBATs.[83, 84] The FBAT is refractory to problems arising due to non-specified complex disease models, missing parents, general pedigrees and/or in cases where the distribution of a disease in a population is unknown.[83, 84] In addition, the FBAT is unaffected by population admixture or stratification, and is useful in cases where multiple comparisons must be made.[83, 84] This test is ideal for studying the association of a given gene in a complex disorder, such as the serotonin transporter gene in autism.

Thus, we used the FBAT[84], including the Haplotype Based Association Test (HBAT) to determine if any HTTLPR or VNTR alleles or haplotypes of these alleles are associated with autism in our sample. An increased number of permutations results in greater analytical precision, thus we performed 1,000 permutations on our sample. Given the gender bias of autism spectrum disorders, we separated our dataset into families containing affected males only and families containing affected females only.

We performed our analysis based on the evidence for linkage driving our dataset based on previous findings in our lab. In addition, we performed a quantitative

transmission disequilibrium test based on the six components key to the autism phenotype, as based on the work done by Tadevosyan-Leyfer and colleagues (2003).[81]

CHAPTER III

RESULTS AND CONCLUSIONS

HTTLPR Results and Discussion

There is no significant association of either HTTLPR allele with autism in our sample. Further, there was no significant association of either HTTLPR allele with autism in families with affected males only or in families with affected females only.

Table 2. FBAT analysis of association of HTTLPR polymorphisms with autism. Labels are as follows: (Freq.) allele frequency, (S) seen, (E(S)) Expected to be seen, Var (S) variance. There is no significant association of either HTTLPR allele with autism in the total sample, within families with affected males only or within families with affected females only.

All Famililes			**P**	**P (-e)**	**P (-p)**
			0.19	0.22	0.19
Allele	afreq	S	E(S)	Var(S)	P
484	0.458	521	500	281.4	0.22
528	0.542	517	538	281.4	0.22
Affected Males Only			**P**	**P (-e)**	**P (-p)**
			0.23	0.26	0.22
Allele	afreq	S	E(S)	Var(S)	P
484	0.458	424	407	221.7	0.26
528	0.542	440	457	221.7	0.26
Affected Females Only			**P**	**P (-e)**	**P (-p)**
			0.59	0.59	0.64
Allele	afreq	S	E(S)	Var(S)	P
484	0.458	131	127	55.6	0.59
528	0.542	147	151	55.6	0.59

These results are consistent with a previous finding of no association between HTTLPR genotype and autism,[72-75, 85] but disparate with a number of other studies that have found association of either the *l* or *s* HTTLPR allele with autism.[39, 67-71]

As mentioned above, the inconsistency amongst studies of HTTLPR could be due to a number of things such as small sample size, difficulty genotyping this region, differing family collections possessing somewhat different properties vis-à-vis etiologic heterogeneity and thus differing representation of risk alleles. One very important possible reason for the discrepancy among HTTLPR association studies in autism may be a failure to account for the effect of a SNP located within the *l* allele of HTTLPR.

Ten SNPs have been identified within HTTLPR.[86] A prevalent A → G polymorphism located in one of the two additional repeat units of the l allele is associated with altered transcriptional efficiency of HTTLPR.[87] This makes HTTLPR a tri-allelic locus comprised of the *s* allele and the A and G variants of the *l* allele, represented as l_G and l_A, respectively. The l_A allele results a 2.8-fold increase in reporter gene expression as compared to s allele within transfected RN46A cells. However, the effect of the l_G allele on reporter gene expression is equivalent to that of the s allele.[87] Further, quantitation of SERT mRNA levels in lymphoblastoid cell lines indicates the l_A / l_A genotype results in the highest levels of SERT mRNA while the s/s genotype results in the lowest levels of SERT mRNA.[87]

The l_G allele results in lower transcriptional activity as compared to the l_A allele due to the corresponding presence of an AP-2 binding site, where AP-2 can suppress gene transcription.[87] The AP-2 family of transcription factors are important in neural

development and AP-2 binding sites are present in regulatory sites in a number of genes that are part of monoamine neurotransmitter systems.[88] AP-2 acts a suppressor of HTTLPR l_G and if an oligonucleotide containing the sequence for the l_G allele AP2 binding-site is added into the cell, reporter gene expression between l_G and l_A will be equivalent, indicating that if AP2 fails to bind to its site on the l_G allele, transcriptional efficiency between both *l* allele is the same.[87]

Xian-Zhang Hu and colleagues studied the frequency of each allele in three populations and found the s:l_G:l_A ratio to be 2.5:5:2.5 in African Americans, 4:5:1 in Caucasians, and 2:1:0 in American Indians.[87]

Frequently HTTLPR association studies consider this locus to be bi-allelic, not taking into account this A→G polymorphism that results in gene expression levels equivalent to those of the s allele. This would presumably result in failure to detect possible associations accurately since grouping both *l* alleles into one group results in a dampening effect.[87] One large (2,998 total individuals) genotyping study done by Hu and colleagues found an s allele frequency of .25 in subjects who were of African descent, .35-.40 in Caucasians, and .64-.66 in American Indians, revealing a disparate HTTLPR allele frequency among differing populations.[87]

In addition, the inconsistency between these studies may partly be explained by failure to consider clinical heterogeneity, with which genetic heterogeneity has been correlated. For instance, Cook and colleagues tested the possibility of a genotype-phenotype interaction by comparing ADI-R and ADOS scores of individuals within the different genotype categories of HTTLPR. Individuals with one or two copies of the s allele were found to have higher (more severe) scores on the "failure to use nonverbal

communication to regulate social interaction" subdomain of the ADI-R as compared to their *l/l* counterparts. The *l/l* individuals tended to have more severe scores on the "stereotyped and repetitive motor mannerisms" subdomain of the ADI-R and also had higher ADOS scores on "directed facial expressions" and "unusual sensory interests", reflecting the important concept that genotypic heterogeneity often underlies phenotypic heterogeneity in complex disorders such as autism.[80] This idea is further supported by the findings of our lab, which previously observed an increase in the evidence for linkage at 17q11.2 at *SLC6A4*, with the multipoint HLOD score increasing from 2.74 to 3.62 in families with more severe rigid-compulsive behaviors.[39] In addition, our lab also found rare variants in *SLC6A4* to be associated with more severe rigid-compulsive behaviors.[58] It is possible that our analysis of HTTLPR did not yield any significant associations because we did not test for the l_G and l_A polymorphisms and because we did not take phenotypic heterogeneity into account in this study. Further analysis to this end is an important next step in the study of these data.

VNTR Results

There is no significant association of any of the VNTR alleles with autism in the total sample, within families with affected males only or within families with affected females only. There were only enough individuals in the families with affected females only category with the 9 repeat VNTR allele to warrant analysis of this allele.

Table 3. FBAT analysis of association of VNTR polymorphisms with autism. Labels are as follows: (Freq.) allele frequency, (S) seen, (E(S)) Expected to be seen, Var (S) variance. There is no significant association of any of the VNTR alleles with autism in the total sample, within families with affected males only or within families with affected females only.

All Families			**P**	**P (-e)**	**P (-p)**
			0.25	0.28	0.16
Allele	afreq	S	E(S)	Var(S)	P
10	0.358	380	406	284.0	0.12
12	0.629	692	665	280.9	0.11
Affected Male Only			**P**	**P (-e)**	**P (-p)**
			0.53	0.55	0.42
Allele	afreq	S	E(S)	Var(S)	P
10	0.358	300	315	214.3	0.30
12	0.629	545	529	214.2	0.28
Affected Female Only			**P**	**P (-e)**	**P (-p)**
			0.38	0.4	0.27
Allele	afreq	S	E(S)	Var(S)	P
9	0.014	7	7	2.8	0.97
10	0.358	99	110	61.6	0.17
12	0.629	182	171	63.4	0.18

Two previous studies reported no evidence of association at STin2 with autism,[72, 73] two studies reported an association of a haplotype containing the *s* variant of HTTLPR and STin2.12 with autism, [59, 74] and one study found evidence for association of a haplotype containing the *l* variant of HTTLPR and STin2.12 with autism.[67] Further, one group found excess transmission to autistic probands for multiple haplotypes of *SLC6A4*, the most significant being a haplotype defined by three polymorphisms including the STin2.12 allele.[70] In addition, two independent groups tested the possibility of

association of VNTR polymorphisms with specific autism phenotypes with disparate results: Cook and colleagues found no significant association of any STin2 genotype with specific behavioral subcategories from ADI-R or ADOS, while Conroy and colleagues found an association between individuals with the homozygous STin2.12 genotype and more severe rigid-compulsive tendencies. Similarly, QTDT analysis revealed an association between the STin2.12 allele and more severe rigid-compulsive tendencies in this study.[79]

Discussion of VNTR

Transcription Factors of VNTR Affecting SERT Expression

The individual repeat units of STin2 are not all identical in sequence or length.[76] In fact, repeats four, seven, nine, and eleven are only 16 nucleotides in length and there are seven unique repeat sequences among the 12 units.[76] Only residues 6, 10, 11, 12, 14 and 15 may be polymorphic, and changes are conservative as only C/T and A/G substitutions are observed.[76] Unlike trinucleotide repeats which would likely not contain enough primary sequence variation to result in multiple transcription factor consensus binding motifs, STin2 repeats are relatively large and variable in primary sequence information. These factors increase the probability of transcription factor binding motifs that may be specific to a given allele of STin2.

Y box binding protein (YB-1), a transcription factor, is known to interact with STin2.[78] In fact, there are 12 potential binding sites within the 12 repeat allele of STin2 and significantly, these DNA consensus sites span the repeat units of STin2.

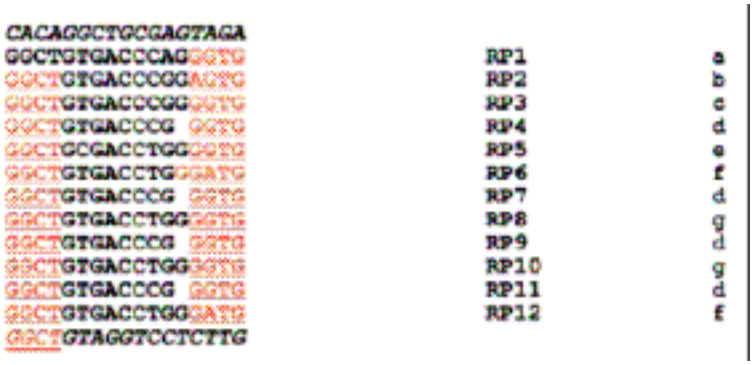


Figure 5. Primary sequence of each repeat and location of YB-1 consensus motifs in VNTR. Shown are the sequences of each repeat unit of the VNTR in *SLC6A4*. The sequence is written 5′ to 3′ of the sense strand. The seven unique sequences among the 12 repeats are labeled a-g. Repeats four, seven, nine and eleven are only 16 nucleotides in length; all the rest are 17 nucleotides in length. Highlighted in red are the 12 putative binding consensus motifs for YB-1 comprised of the sequences GATG/CATC or GGTG/CACC. These motifs encompass the region of possible sequence variation and possible variation in the number of nucleotides per repeat, 16 vs. 17. Figure taken from Klenova et al. 2004.[78]

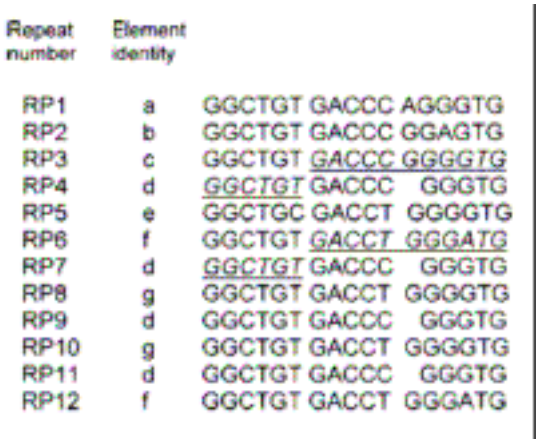
The GATG/CATC and GGTG/CACC motifs are consensus motifs recognized and bound by YB-1.[78, 89] The location of these motifs is important because these sites encompass the region of possible sequence variation and possible variation in the number of nucleotides per repeat (16 vs. 17 bp).[78] Therefore, it is possible that YB-1 may affect STin2 expression in a differential pattern based on the primary sequence of the repeat and the size of the repeat. Cotransfection assays in COS7 cells using STin2 variants upstream of an SV40 promoter attached to the Luciferase gene (termed pStin2.9Luc, Stin2.10Luc and Stin2.12Luc) in addition to a vector containing YB-1, reveal YB-1 differentially activates each polymorphism of STin2 in this specific cell line. PStin2.9Luc expression was strongly affected by the presence of YB-1 and an increase in YB-1 concentration resulted in increasingly stronger activation of this variant. On the other hand, the pStin2.10Luc expression was not affected by the presence of YB-1 and pStin2.12Luc expression was slightly repressed by the presence of YB-1. This result may partially be

cell-specific as pStin2.10Luc was activated by the presence of a relatively low concentration of YB-1 in HEK293T cells in addition to pStin2.9Luc being activated by YB-1; this lends evidence to the idea that modulation of YB-1's effect on STin2 expression is affected by factors that may be present in some types of cells and not other types.[78]

One possible explanation for differential modulation of STin2 variants by YB-1 could be the frequency of G/A polymorphisms located in the consensus binding motifs in STin2.9. Most of the consensus binding sites located in STin2.12 (the longest STin2 variant) have the sequence, GGTGGGCT, but two consensus binding sequences, located at the junction between repeat six and seven, and at the end of repeat 12 have the sequence GATGGGCT. The use of the GGTGGGCT consensus sequence between repeat three and four (rep 3/4) along with the GATGGGCT consensus sequence (rep 6/7) in EMSA competition experiments with a specific inhibitor oligonucleotide and YB-1, demonstrates differential binding of YB-1 to each type of motif. Specifically, rep 3/4 (GGTGGGCT) formed four complexes with YB-1 and only two of these complexes could be outcompeted by the presence of the inhibitor oligonucleotide, however, rep 6/7 (GATGGGCT) formed two complexes with YB-1 (with a faint third complex occasionally being observed) and all complexes could be outcompeted by the inhibitor oligonucleotide, although it did take more oligonucleotide to produce this effect than it did for the two complexes that dissociated in rep 3/4.[78] This suggests differential binding of YB-1 to GATTGGGCT motifs vs. GGTGGGCT motifs; the oligonucleotide with the GGTTGGGCT motif (rep 3/4) formed more (four total) complexes with YB-1 proteins and only two of them could be outcompeted by the inhibitor oligonucleotide.

Conversely, the GATGGGCT oligonucleotide motif only formed two to three complexes with YB-1 and all complexes could be outcompeted (although with more inhibitor nucleotide than it took to out-compete the two GGTTGGGCT motif complexes).

One study done by Lovejoy and colleagues further supports the hypothesis that STin2 repeat domains differ in their ability to regulate enhancement of reporter gene expression.[77] This group used ES cells to study the enhancement capabilities of each individual repeat, and also that of two larger oligonucleotides that span two separate repeat units.[77]



Repeat number	Element identity	
RP1	a	GGCTGT GACCC AGGGTG
RP2	b	GGCTGT GACCC GGAGTG
RP3	c	GGCTGT *GACCC GGGGTG*
RP4	d	*GGCTGT* GACCC GGGTG
RP5	e	GGCTGC GACCT GGGGTG
RP6	f	GGCTGT *GACCT GGGATG*
RP7	d	*GGCTGT* GACCC GGGTG
RP8	g	GGCTGT GACCT GGGGTG
RP9	d	GGCTGT GACCC GGGTG
RP10	g	GGCTGT GACCT GGGGTG
RP11	d	GGCTGT GACCC GGGTG
RP12	f	GGCTGT GACCT GGGATG

Figure 6. Sequence of the "spanning" oligonucleotides of VNTR tested for reporter gene expression enhancement in ES cells. The sequence of each repeat is written 5' to 3' along the sense strand. The seven unique sequences among the 12 repeats are labeled a-g. Underlined are the two "spanning" oligonucleotides tested for enhancing ability in an expression assay. The "spanning" c/d oligonucleotide has the sequence GACCCGGGGTGGGCTGT while the "spanning" f/d oligonucleotide has sequence, GACCTGGGATGGGCTGT. (RP) repeat. This figure was taken from Lovejoy et. al. 2003.[77]

Figure 5 shows the seven unique repeat sequences tested in addition to two "spanning" oligonucleotides (termed c/d and f/d) that were also tested for enhancer activity. The "spanning" c/d oligonucleotide has the sequence GACCCGGGGTGGGCTGT while the "spanning" f/d oligonucleotide has sequence,

GACCTGGGATGGGCTGT. Repeat units c and g showed a small-fold increase in enhancer activity upon removal of LIF, whereas repeat unit e did not show a significant increase in enhancer activity.[77] Surprisingly, the f/d oligonucleotide construct supported a 4.1 fold increase and the c/d construct supported a 13.5 fold increase in enhancer activity even though these two constructs differ by just two nucleotide residues.[77] Thus, it appears that sequence variation, not just repeat number, of the STin2 repeat polymorphism, is an important determinant in enhancer ability, at least in murine ES cells.[77] Interestingly, when these oligonucleotides are tested in the human placental cell line, JAR (often used to test *SLC6A4* promoter function) the results are quite different with there not being a significant difference in the enhancing ability of the c repeat vs. the c/d repeat.[77] The researchers in this study conclude that transcription factors that may bind to the differing transcription factor motifs within the VNTR domain are activated by tissue specific stimuli, not ubiquitous non-specific stimuli like that which would bind to a constitutively active regulatory domain.[77]

CTCF: Interaction with YB-1 for Regulation of STin2

The implications for the differential binding properties of the two types of YB-1 consensus motifs in STin2 are further amplified when factors modulating YB-1's interaction with STin2 are taken into account. One example of a factor that affects the effect of YB-1 on STin2 polymorphisms is CCTC-binding protein (CTCF).[78] CTCF has an 11-zinc finger DNA-binding domain and can bind to very divergent DNA sequences of approx. 50 bp in length within the promoter regions of human c-myc genes (reviewed in reference 72).[90] Baculoviral-produced CTCF is able to cause dissociation of YB-1 -

STin2.12 complexes as well as YB-1–rep 3/4 and YB-1–rep 6/7 complexes.[78] In addition, co-transfection studies reveal that CTCF results in a strong inhibition of STin2.9 expression in COS7 and HEK 293T cells, but only when YB-1 is present.[78] In fact, in the absence of YB-1, CTCF has no effect on the expression levels of any of the STin2 polymorphisms in COS7 or HEK 293T cells.[78] PStin2.10Luc expression is increased by the presence of YB-1 in HEK 293T cells but when CTCF is added, the increase in expression is not as substantial.[78] These experiments indicate that CTCF negatively regulates the enhancement effects of YB-1 on STin2 expression. The investigators suggest that this regulation could occur through either of two distinct mechanisms: first, CTCF could bind to the YB-1 consensus motif on STin2, resulting in stereotaxic inhibition of this region for activation by YB-1. Alternatively, CTCF might bind to YB-1, resulting in the loss of YB-1's ability to bind to its recognition motif on STin2. The latter postulation is likely based on studies revealing that CTCF binds to YB-1 through interaction of CTCF's zinc-finger region with YB-1.[90] Truncation interaction studies reveal that YB-1 binds to the zinc-finger region of CTCF through YB-1's DNA-binding cold-shock domain (CSD) located on the N-terminal.[78] Finally, if CTCF was exerting its inhibitory effect through stereotaxic inhibition of YB-1's recognition motif, one would expect a CTCF-STin2 complex present in the EMSA studies mentioned above and this was not observed.[78]

Possible Reasons for Differential Expression

To summarize, YB-1 differentially regulates STin2 polymorphisms *in vitro* and this is dependent upon cell type. In addition, at least one factor that we know of, CTCF,

can diminish the enhancement effects of YB-1 on STin2 expression for those variants of STin2 that are responsive to YB-1. The question that arises, is, why is there such a drastic difference in how each STin2 polymorphism is affected by YB-1?

Although observed STin2 alleles are highly similar, differences exist that may have far-reaching implications for effects on expression. The effects of YB-1 may depend on the secondary structure of the DNA to which it is binding.[78, 89] Thus, we must consider the possibility that differences in both repeat number and complement (i.e. the number of 16 bp vs. 17 bp repeats and the presence of C/T or G/A variants) for STin2 may affect the binding of YB-1 (and possibly many other transcription factors) through secondary structure effects on the DNA of this region. Repeats six, seven, and eight are absent from the STin2.9 allele, and repeats nine and ten are absent in STin2.10.[78] This means that one of the two GATGGGCT motifs (rep 6/7) is not present in STin2.9. Based on the strength of STin2.9 activation by YB-1 *in vitro*, it is possible that when the rep 6/7 motif is bound by YB-1 in STin2.10 or STin2.12, it may diminish enhancement of STin2 expression by YB-1, perhaps through secondary DNA structure effects.[78, 89] In addition, the presence of a single nucleotide polymorphism (within STin2.10 and STin2.12; G/A) at two locations, can affect how strongly YB-1 may bind, but also likely produces new transcription factor binding motifs recognized by entirely different transcription factors and indeed, this is the case.[78] In fact, the presence of the G/A polymorphism in repeat six results in the creation of a number of predicted transcription factor binding sites for such transcription factors as TCF11/KCR-F1/Nrf1 homodimers, activator protein-1 and retinoic acid receptor-related orphan receptor alpha 1.[78] STin2.9 would only carry one GATGGCT motif and thus would only contain one binding site for these other

transcription factors unlike STin2.10 and STin2.12 which would each have two binding sites for these factors. However data are not yet available showing that any of the predicted factors functionally act on these sequences. Also unaddressed in this study was a discussion of why observable differences exist between the enhancing capabilities of STin2.10 and STin2.12, not just differences between STin2.9 and both STin2.10 and STin2.12. Unlike STin2.9, differential frequency of G/A polymorphisms would not likely be a possible explanation for differential regulatory abilities of STin2.10 vs. STin2.12, since both alleles have the two G/A polymorphisms present. One possible explanation for the difference in enhancing capacities is the size difference in the alleles, and thus the likely alteration in secondary DNA structure. At this point, however, no further assumptions can be made without additional expression and transcription factor binding studies.

STin2.12 is known to enhance reporter gene expression in ES cells by a 167 fold increase as compared to STin2.10 which enhanced reporter gene expression by only 5-6 fold.[76] Thus one possible effect of this polymorphism is enhanced expression of *SLC6A4*. We know that depending on the cell type, the various repeats of STin2 and the various polymorphisms seem to have independent regulatory effects on reporter gene transcription. Thus, while there is evidence for differential enhancer properties of the individual variants even within a given cell type, we must also consider the enhancer profile of each variant depending upon cell type. Different complements of transcription factors exist in differing cell types and the various STin2 polymorphisms may be affected differently. Recall that STin2 variants produced differential expression patterns in the brain of developing mouse embryos: Stin2.10 embyros had low levels of reporter-gene

expression in the region of rhombomeres one and two whereas Stin2.12 embryos had high levels of reporter-gene expression in these rhombomeres, comparable to levels of reporter-gene in rhombomeres four and five of both variants.[63] This rostral portion of the hindbrain expresses *SLC6A4* mRNA and is involved in the formation of serotonergic cell clusters.[63] Thus, one could speculate that having one or more copies of the STin2.12 allele could result in higher expression of *SLC6A4* anywhere *SLC6A4* is expressed, but could also, and perhaps more importantly, result in a significantly different spatial profile of SERT in the brain of a given individual. Given the differences between the male and female brain anatomy, this differential spatial profile may be an important factor to consider in light of a possible STin2.12 gender bias.

While Stin2, or VNTR, plays a role in the expression of *SLC6A4*, it may not play enough of a role to be associated with a complex neuropsychiatric disorder such as autism.

Haplotype Analysis Results

There was no significant association of a haplotype containing specific alleles of HTTLPR or VNTR with autism in our sample. A haplotype containing the *s* allele of HTTLPR and STin2.12 has been previously associated with autism in two separate studies,[59, 74] and one study found association of the *l* allele of HTTLPR and STin2.12 in autism.[67]

Table 4. Haplotype Based Association Test for association of an HTTLP/VNTR haplotype with autism. There was no significant association of a haplotype containing specific alleles of HTTLPR or VNTR with autism in our sample

All Families	**P**	**P (-e)**	**P (-p)**
	0.11	0.17	0.07
Affected Male Only	**P**	**P (-e)**	**P (-p)**
	0.16	0.25	0.12
Affected Female Only	**P**	**P (-e)**	**P (-p)**
	0.41	0.44	0.35

The results of our analysis make sense since there was no significant association of any of the HTTLPR or VNTR alleles with autism in our sample.

Quantitative Transmission Disequilibrium Test Results

Finally, we performed a quantitative transmission disequilibrium test based on the six components key to the autism phenotype.[81] There were no significant associations of HTTLPR or VNTR alleles with any of the six components in the combined dataset. Maternal inheritance of HTTLPR or VNTR alleles was not significantly associated with any of the six components. However, paternal inheritance of VNTR alleles was associated with two components, the social intent component and the milestones component ($p = .02$ and $p = .05$, respectively).

Table 5. QTDT analysis for parental transmission of HTTLPR and VNTR alleles. There were no significant associations of HTTLPR or VNTR alleles with any of the six components in the combined dataset. Maternal inheritance of HTTLPR or VNTR alleles was not significantly associated with any of the six components. However, paternal inheritance of VNTR alleles was associated with two components, the social intent component and the milestones component (p = .02 and p = .05, respectively).

OVERALL	Language_t	Social	Milestones_t	Savant_t	Rigid_t	Sensory_t
VNTR	0.90	0.40	0.70	1.00	0.90	0.50
HTTLPR	0.90	0.30	0.50	0.60	0.10	0.90
Maternal Only	Language_t	Social	Milestones_t	Savant_t	Rigid_t	Sensory_t
VNTR	0.600	0.500	0.100	0.400	0.100	0.700
HTTLPR	0.500	0.600	0.300	0.200	0.400	0.700
Paternal Only	Language_t	Social	Milestones_t	Savant_t	Rigid_t	Sensory_t
VNTR	0.700	0.020	0.050	0.900	0.400	0.600
HTTLPR	0.800	0.400	0.700	0.300	0.400	0.300

The social intent component is related to individual expression of a idea, desire or feeling and includes gesturing and greeting in addition to nonverbal communication and ability to relate socially.[81] The milestones component relates to the attainment of milestones in physical and mental development.[81]

Given the lack of association for any HTTLPR or VNTR alleles with autism, we have to consider the possibility that these two association results could be due to chance. This would seem most likely for the milestones component as it is nominally significant. Further study needs to be done to determine if paternal inheritance of VNTR alleles contributes to susceptibility for autism.

REFERENCES

1. Kanner, L. Autistic Disturbance of Affective Contact. *Nervous Child* **2**, 217-250 (1943).
2. Lord, C. et al. Diagnosing autism: analyses of data from the Autism Diagnostic Interview. *Journal of autism and developmental disorders* **27**, 501-517 (1997).
3. Lord, C., Rutter, M. & Le Couteur, A. Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of autism and developmental disorders* **24**, 659-685 (1994).
4. Lord, C. et al. Autism diagnostic observation schedule: a standardized observation of communicative and social behavior. *Journal of autism and developmental disorders* **19**, 185-212 (1989).
5. Folstein, S.E. & Rosen-Sheidley, B. Genetics of autism: complex aetiology for a heterogeneous disorder. *Nat Rev Genet* **2**, 943-955 (2001).
6. Bethea, T.C. & Sikich, L. Early pharmacological treatment of autism: a rationale for developmental treatment. *Biological psychiatry* **61**, 521-537 (2007).
7. Fombonne, E. The prevalence of autism. *Jama* **289**, 87-89 (2003).
8. Chakrabarti, S. & Fombonne, E. Pervasive developmental disorders in preschool children: confirmation of high prevalence. *The American journal of psychiatry* **162**, 1133-1141 (2005).
9. Levy, S.E. et al. Relationship of dietary intake to gastrointestinal symptoms in children with autistic spectrum disorders. *Biological psychiatry* **61**, 492-497 (2007).
10. Tuchman, R. & Rapin, I. Epilepsy in autism. *Lancet neurology* **1**, 352-358 (2002).
11. Folstein, S. & Rutter, M. Infantile autism: a genetic study of 21 twin pairs. *J Child Psychol Psychiatry* **18**, 297-321 (1977).
12. Rutter, M. et al. Genetic factors in child psychiatric disorders--II. Empirical findings. *J Child Psychol Psychiatry* **31**, 39-83 (1990).
13. Jones, M.B. & Szatmari, P. Stoppage rules and genetic studies of autism. *J Autism Dev Disord* **18**, 31-40 (1988).
14. Ritvo, E.R. et al. The UCLA-University of Utah epidemiologic survey of autism: prevalence. *Am J Psychiatry* **146**, 194-199 (1989).
15. Risch, N. et al. A genomic screen of autism: evidence for a multilocus etiology. *American journal of human genetics* **65**, 493-507 (1999).
16. Schain R.J., F.D.X. Studies on 5-hydroxyindole metabolism in autistic and other mentally retarded children. . *J. Pediatr* **58**, 315-320 (1961).
17. Cook, E.H. & Leventhal, B.L. The serotonin system in autism. *Curr Opin Pediatr* **8**, 348-354 (1996).
18. Gordon, C.T., State, R.C., Nelson, J.E., Hamburger, S.D. & Rapoport, J.L. A double-blind comparison of clomipramine, desipramine, and placebo in the treatment of autistic disorder. *Arch Gen Psychiatry* **50**, 441-447 (1993).
19. McDougle, C.J. et al. A double-blind, placebo-controlled study of fluvoxamine in adults with autistic disorder. *Arch Gen Psychiatry* **53**, 1001-1008 (1996).

20. Kolevzon, A., Mathewson, K.A. & Hollander, E. Selective serotonin reuptake inhibitors in autism: a review of efficacy and tolerability. *J Clin Psychiatry* **67**, 407-414 (2006).
21. McDougle, C.J. et al. Acute tryptophan depletion in autistic disorder: a controlled case study. *Biological psychiatry* **33**, 547-550 (1993).
22. Chugani, D.C. et al. Altered serotonin synthesis in the dentatothalamocortical pathway in autistic boys. *Annals of neurology* **42**, 666-669 (1997).
23. Chugani, D.C. et al. Developmental changes in brain serotonin synthesis capacity in autistic and nonautistic children. *Annals of neurology* **45**, 287-295 (1999).
24. Cook, E.H., Jr. et al. Autistic children and their first-degree relatives: relationships between serotonin and norepinephrine levels and intelligence. *J Neuropsychiatry Clin Neurosci* **2**, 268-274 (1990).
25. Piven, J. et al. Platelet serotonin, a possible marker for familial autism. *J Autism Dev Disord* **21**, 51-59 (1991).
26. Whitaker-Azmitia, P.M. Behavioral and cellular consequences of increasing serotonergic activity during brain development: a role in autism? *Int J Dev Neurosci* **23**, 75-83 (2005).
27. Bennett, J.S. Platelet-fibrinogen interactions. *Ann N Y Acad Sci* **936**, 340-354 (2001).
28. Bennett, J.S. & Mousa, S. Platelet function inhibitors in the Year 2000. *Thromb Haemost* **85**, 395-400 (2001).
29. Ashley-Koch, A. et al. Genetic studies of autistic disorder and chromosome 7. *Genomics* **61**, 227-236 (1999).
30. An autosomal genomic screen for autism. *Am J Med Genet* **105**, 609-615 (2001).
31. Philippe, A. et al. Genome-wide scan for autism susceptibility genes. Paris Autism Research International Sibpair Study. *Hum Mol Genet* **8**, 805-812 (1999).
32. Auranen, M. et al. Analysis of autism susceptibility gene loci on chromosomes 1p, 4p, 6q, 7q, 13q, 15q, 16p, 17q, 19q and 22q in Finnish multiplex families. *Mol Psychiatry* **5**, 320-322 (2000).
33. Buxbaum, J.D. et al. Linkage analysis for autism in a subset families with obsessive-compulsive behaviors: evidence for an autism susceptibility gene on chromosome 1 and further support for susceptibility genes on chromosome 6 and 19. *Mol Psychiatry* **9**, 144-150 (2004).
34. Buxbaum, J.D. et al. Evidence for a susceptibility gene for autism on chromosome 2 and for genetic heterogeneity. *Am J Hum Genet* **68**, 1514-1520 (2001).
35. Liu, J. et al. A genomewide screen for autism susceptibility loci. *Am J Hum Genet* **69**, 327-340 (2001).
36. Shao, Y. et al. Phenotypic homogeneity provides increased support for linkage on chromosome 2 in autistic disorder. *Am J Hum Genet* **70**, 1058-1061 (2002).
37. Shao, Y. et al. Genomic screen and follow-up analysis for autistic disorder. *Am J Med Genet* **114**, 99-105 (2002).
38. Yonan, A.L. et al. A genomewide screen of 345 families for autism-susceptibility loci. *Am J Hum Genet* **73**, 886-897 (2003).
39. McCauley, J.L. et al. Linkage and association analysis at the serotonin transporter (SLC6A4) locus in a rigid-compulsive subset of autism. *Am J Med Genet B Neuropsychiatr Genet* **127**, 104-112 (2004).

40. McCauley, J.L. et al. Genome-wide and Ordered-Subset linkage analyses provide support for autism loci on 17q and 19p with evidence of phenotypic and interlocus genetic correlates. *BMC Med Genet* **6**, 1 (2005).
41. Stone, J.L. et al. Evidence for sex-specific risk alleles in autism spectrum disorder. *Am J Hum Genet* **75**, 1117-1123 (2004).
42. Cantor, R.M. et al. Replication of autism linkage: fine-mapping peak at 17q21. *Am J Hum Genet* **76**, 1050-1056 (2005).
43. A genomewide screen for autism: strong evidence for linkage to chromosomes 2q, 7q, and 16p. *Am J Hum Genet* **69**, 570-581 (2001).
44. Purselle, D.C. & Nemeroff, C.B. Serotonin transporter: a potential substrate in the biology of suicide. *Neuropsychopharmacology* **28**, 613-619 (2003).
45. Lebrand, C. et al. Transient uptake and storage of serotonin in developing thalamic neurons. *Neuron* **17**, 823-835 (1996).
46. Horiuchi, H. Recent advance in antiplatelet therapy: the mechanisms, evidence and approach to the problems. *Annals of medicine* **38**, 162-172 (2006).
47. Prasad, H.C. et al. Human serotonin transporter variants display altered sensitivity to protein kinase G and p38 mitogen-activated protein kinase. *Proc Natl Acad Sci U S A* **102**, 11545-11550 (2005).
48. Lesch, K.P., Wolozin, B.L., Murphy, D.L. & Reiderer, P. Primary structure of the human platelet serotonin uptake site: identity with the brain serotonin transporter. *Journal of neurochemistry* **60**, 2319-2322 (1993).
49. Ramamoorthy, S. et al. Antidepressant- and cocaine-sensitive human serotonin transporter: molecular cloning, expression, and chromosomal localization. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 2542-2546 (1993).
50. Hoffman, B.J., Mezey, E. & Brownstein, M.J. Cloning of a serotonin transporter affected by antidepressants. *Science* **254**, 579-580 (1991).
51. Chen, J.G., Liu-Chen, S. & Rudnick, G. Determination of external loop topology in the serotonin transporter by site-directed chemical labeling. *J Biol Chem* **273**, 12675-12681 (1998).
52. Miner, L.H., Schroeter, S., Blakely, R.D. & Sesack, S.R. Ultrastructural localization of the serotonin transporter in superficial and deep layers of the rat prelimbic prefrontal cortex and its spatial relationship to dopamine terminals. *J Comp Neurol* **427**, 220-234 (2000).
53. Ramamoorthy, S., Giovanetti, E., Qian, Y. & Blakely, R.D. Phosphorylation and regulation of antidepressant-sensitive serotonin transporters. *J Biol Chem* **273**, 2458-2466 (1998).
54. Ramamoorthy, S. & Blakely, R.D. Phosphorylation and sequestration of serotonin transporters differentially modulated by psychostimulants. *Science* **285**, 763-766 (1999).
55. Samuvel, D.J., Jayanthi, L.D., Bhat, N.R. & Ramamoorthy, S. A role for p38 mitogen-activated protein kinase in the regulation of the serotonin transporter: evidence for distinct cellular mechanisms involved in transporter surface expression. *J Neurosci* **25**, 29-41 (2005).
56. Zhu, C.B., Hewlett, W.A., Feoktistov, I., Biaggioni, I. & Blakely, R.D. Adenosine receptor, protein kinase G, and p38 mitogen-activated protein kinase-dependent

up-regulation of serotonin transporters involves both transporter trafficking and activation. *Mol Pharmacol* **65**, 1462-1474 (2004).
57. Zhu, C.B., Carneiro, A.M., Dostmann, W.R., Hewlett, W.A. & Blakely, R.D. p38 MAPK activation elevates serotonin transport activity via a trafficking-independent, protein phosphatase 2A-dependent process. *J Biol Chem* **280**, 15649-15658 (2005).
58. Sutcliffe, J.S. et al. Allelic heterogeneity at the serotonin transporter locus (SLC6A4) confers susceptibility to autism and rigid-compulsive behaviors. *American journal of human genetics* **77**, 265-279 (2005).
59. Kerem, B. et al. Identification of the cystic fibrosis gene: genetic analysis. *Science* **245**, 1073-1080 (1989).
60. Glatt, C.E. et al. Screening a large reference sample to identify very low frequency sequence variants: comparisons between two genes. *Nat Genet* **27**, 435-438 (2001).
61. Ozaki, N. et al. Serotonin transporter missense mutation associated with a complex neuropsychiatric phenotype. *Mol Psychiatry* **8**, 933-936 (2003).
62. Lesch, K.P. et al. Association of anxiety-related traits with a polymorphism in the serotonin transporter gene regulatory region. *Science* **274**, 1527-1531 (1996).
63. MacKenzie, A. & Quinn, J. A serotonin transporter gene intron 2 polymorphic region, correlated with affective disorders, has allele-dependent differential enhancer-like properties in the mouse embryo. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 15251-15255 (1999).
64. Heinz, A. et al. A relationship between serotonin transporter genotype and in vivo protein expression and alcohol neurotoxicity. *Biological psychiatry* **47**, 643-649 (2000).
65. Greenberg, B.D. et al. Association between the serotonin transporter promoter polymorphism and personality traits in a primarily female population sample. *American journal of medical genetics* **96**, 202-216 (2000).
66. Neumeister, A. et al. Association between serotonin transporter gene promoter polymorphism (5HTTLPR) and behavioral responses to tryptophan depletion in healthy women with and without family history of depression. *Archives of general psychiatry* **59**, 613-620 (2002).
67. Klauck, S.M., Poustka, F., Benner, A., Lesch, K.P. & Poustka, A. Serotonin transporter (5-HTT) gene variants associated with autism? *Human molecular genetics* **6**, 2233-2238 (1997).
68. Yirmiya, N. et al. Evidence for an association with the serotonin transporter promoter region polymorphism and autism. *American journal of medical genetics* **105**, 381-386 (2001).
69. Cook, E.H., Jr. et al. Evidence of linkage between the serotonin transporter and autistic disorder. *Mol Psychiatry* **2**, 247-250 (1997).
70. Conroy, J. et al. Serotonin transporter gene and autism: a haplotype analysis in an Irish autistic population. *Mol Psychiatry* **9**, 587-593 (2004).
71. Devlin, B. et al. Autism and the serotonin transporter: the long and short of it. *Molecular psychiatry* **10**, 1110-1116 (2005).
72. Maestrini, E. et al. Serotonin transporter (5-HTT) and gamma-aminobutyric acid receptor subunit beta3 (GABRB3) gene polymorphisms are not associated with

autism in the IMGSA families. The International Molecular Genetic Study of Autism Consortium. *American journal of medical genetics* **88**, 492-496 (1999).
73. Betancur, C. et al. Serotonin transporter gene polymorphisms and hyperserotonemia in autistic disorder. *Molecular psychiatry* **7**, 67-71 (2002).
74. Kim, S.J. et al. Transmission disequilibrium mapping at the serotonin transporter gene (SLC6A4) region in autistic disorder. *Molecular psychiatry* **7**, 278-288 (2002).
75. Persico, A.M. et al. Lack of association between serotonin transporter gene promoter variants and autistic disorder in two ethnically distinct samples. *American journal of medical genetics* **96**, 123-127 (2000).
76. Fiskerstrand, C.E., Lovejoy, E.A. & Quinn, J.P. An intronic polymorphic domain often associated with susceptibility to affective disorders has allele dependent differential enhancer activity in embryonic stem cells. *FEBS letters* **458**, 171-174 (1999).
77. Lovejoy, E.A., Scott, A.C., Fiskerstrand, C.E., Bubb, V.J. & Quinn, J.P. The serotonin transporter intronic VNTR enhancer correlated with a predisposition to affective disorders has distinct regulatory elements within the domain based on the primary DNA sequence of the repeat unit. *The European journal of neuroscience* **17**, 417-420 (2003).
78. Klenova, E. et al. YB-1 and CTCF differentially regulate the 5-HTT polymorphic intron 2 enhancer which predisposes to a variety of neurological disorders. *J Neurosci* **24**, 5966-5973 (2004).
79. Mulder, E.J. et al. Serotonin transporter intron 2 polymorphism associated with rigid-compulsive behaviors in Dutch individuals with pervasive developmental disorder. *Am J Med Genet B Neuropsychiatr Genet* **133**, 93-96 (2005).
80. Brune, C.W. et al. 5-HTTLPR Genotype-Specific Phenotype in Children and Adolescents With Autism. *The American journal of psychiatry* **163**, 2148-2156 (2006).
81. Tadevosyan-Leyfer, O. et al. A principal components analysis of the Autism Diagnostic Interview-Revised. *Journal of the American Academy of Child and Adolescent Psychiatry* **42**, 864-872 (2003).
82. Cuccaro, M.L. et al. Factor analysis of restricted and repetitive behaviors in autism using the Autism Diagnostic Interview-R. *Child psychiatry and human development* **34**, 3-17 (2003).
83. Laird, N.M. & Lange, C. Family-based designs in the age of large-scale gene-association studies. *Nature reviews* **7**, 385-394 (2006).
84. Horvath, S., Xu, X. & Laird, N.M. The family based association test method: strategies for studying general genotype--phenotype associations. *Eur J Hum Genet* **9**, 301-306 (2001).
85. Tordjman, S. et al. Role of the serotonin transporter gene in the behavioral expression of autism. *Molecular psychiatry* **6**, 434-439 (2001).
86. Nakamura, M., Ueno, S., Sano, A. & Tanabe, H. The human serotonin transporter gene linked polymorphism (5-HTTLPR) shows ten novel allelic variants. *Molecular psychiatry* **5**, 32-38 (2000).

87. Hu, X.Z. et al. Serotonin transporter promoter gain-of-function genotypes are linked to obsessive-compulsive disorder. *American journal of human genetics* **78**, 815-826 (2006).
88. Damberg, M. Transcription factor AP-2 and monoaminergic functions in the central nervous system. *J Neural Transm* **112**, 1281-1296 (2005).
89. Zasedateleva, O.A. et al. Specificity of mammalian Y-box binding protein p50 in interaction with ss and ds DNA analyzed with generic oligonucleotide microchip. *Journal of molecular biology* **324**, 73-87 (2002).
90. Chernukhin, I.V. et al. Physical and functional interaction between two pluripotent proteins, the Y-box DNA/RNA-binding factor, YB-1, and the multivalent zinc finger factor, CTCF. *The Journal of biological chemistry* **275**, 29915-29921 (2000).