Rare Variants Affecting Regulation of Serotonin and Dopamine Transport Contribute to the Genetic Liability of Autism

By

Nicholas George Campbell

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

In

Neuroscience

May 2014

Nashville, Tennessee

Approved

James Sutcliffe, Ph.D.

Karoly Mirnics, M.D., Ph.D.

Randy Blakely, Ph.D.

Thomas Morgan, M.D.

ACKNOWLEDGEMENTS

I would first like to thank members of my thesis committee, Dr. Karoly Mirnics M.D., Ph.D (chairman), Dr. Randy Blakely, Ph.D, Dr. Thomas Morgan M.D, and Dr. James Sutcliffe, Ph.D. I would like to especially thank my mentor, Dr. James Sutcliffe, for an opportunity to join his lab. His guidance and assistance were central to my development and the evolution of the work presented here.

I would like to acknowledge the staff support in the Vanderbilt Center of Human Genetics Research, most notably Melissa Potter. Vanderbilt's excellent DNA Resources core, have been extremely helpful for the use of equipment cell line storage, and DNA preparation. Working with the talented individuals at the Genomic Services Lab at Hudson Alpha has been a pleasure.

I am grateful to many people at Vanderbilt and I am fortunate to work in an open and friendly environment everyday. The members of the Sutcliffe lab have been a pleasure to work with and I am indebted to all the current and past members who have contributed to this work. I have been privileged to work with such a sharp and caring group of individuals.

None of this would be possible without the work of wonderful collaborators. In particular I would like to thank Peter Hamilton, Dr. Aurelio Galli, Dr. Chong Bing Zhu, Dr. Heiner Matthies, Dr. Hassane McHaourab, Dr. Chad Schafer, Dr. Kathryn Roeder, Dr. Ed Cook, Dr. Bingshan Li, The ACE group, The ARRA group, and all others, to many to name, who I've had the pleasure to have met and worked with.

TABLE OF CONTENTS

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

LIST OF ABBREVIATIONS

Abbreviation	Description
5-HT	5-hydroxytryptamine (serotonin)
A3AR; *ADORA3*	A3 subtype adenosine receptor
ADI-R	Autism diagnostic interview revised
ADOS	Autism diagnostic observation schedule
AMPH	Amphetamine
ANOVA	Analysis of variance
ASD	Autism Spectrum Disorder
BAP	Broad autism phenotype
CAST	Cohort Allelic Sums Test
CDCV	Common disease - common variant
CDRV	Common disease - rare variant
cGMP	Cyclic guanosine monophosphate
CHO	Chinese hamster ovary
CNV	Copy number variation/variant
COC	Cocaine
CV	Common variant
DA	Dopamine (5-hydroxytryptamine)
DAT; *SLC6A3*	Dopamine transporter
DNV	*De novo* variant
DSM	Diagnostic and Statistical Manual of Mental Disorders
FBAT	Family-Based Association Test
FBS	Fetal bovine serum
GPCR	G-protein coupled receptor
GWA	Genome-wide association
IB-MECA	*N*6-(3-iodobenzyl)-*N*-methyl-5'-carbamoyladenosine
MAF	Minor allele frequency

MAPK	Mitogen-activated protein kinase
MDS	Multidimensional Scaling
NGS	Next-generation sequencing
OCD	Obsessive-compulsive disorder
OR	Odds ratio
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PKC	Protein kinase C
PKG	Protein kinase G
RV	Rare variant
SERT; *SLC6A4*	Serotonin transporter
SNP	Single nucleotide polymorphism
SRS	Social Responsiveness Scale
WES	Whole exome sequencing
WPPSI	Wechsler Preschool and Primary Scale of Intelligence

CHAPTER I

Introduction

The Autism Phenotype

The year 2013 marks the 70th anniversary of Leo Kanner's initial report of a condition he termed "autistic disturbances of affective contact" [1]. His series of eleven case reports laid the foundation for past and present research, and remains the starting point for what is today considered the autism phenotype. With the publication of the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) this year, changes made to the diagnostic criteria for autism disorder were among the most controversial. At the heart of these controversies lies the latest attempt in a seventy-year ongoing effort to define the autism phenotype more accurately.

Autism spectrum disorder (ASD), or simply autism, is a neurodevelopmental disorder characterized by impairments in two domains: 1) verbal and non-verbal communication and impaired social reciprocity and 2) restricted interests and repetitive behaviors. Often noticed early in life, individuals with ASD display a spectrum of impairments ranging from mild to severe, with some individuals displaying well above average intelligence while others showing profound intellectual disability. While no age of onset is required, symptoms must be present at an early age even if they do not fully manifest until the demands exceed the subject's capacities. Individuals with ASD may show communication deficits such as inappropriate responses, misreading of nonverbal cues, and difficulties in building and maintaining relationships. Many ASD individuals

may show high sensitivity to changes in their environment, intense focus on individual items, and dependence on routines.

No diagnostically informative biological tests for ASD exists, however, there are behavioral instruments that perform with high sensitivity and specificity. Standardized interviews and direct age-adjusted observation tools have gained acceptance in research settings [2, 3], and currently a routine battery of psychometric and cognitive assessments have standardized the diagnostic routine in a clinical setting [4, 5]. Examples of these instruments include the autism diagnostic interview revised (ADI-R), a semi-structured interview given to the child's primary caregiver that consists of 93 questions, and the autism diagnostic observational schedule (ADOS), an observational session and an interview with the child. The "gold standard" for diagnosing ASD is using a combination of ADI-R and ADOS and a clinician's best judgment [6].

Diverse arrays of disorders are often co-morbid with an ASD diagnosis. These include epilepsy, sensory, motor, gastrointestinal, and sleep abnormalities [7]. Likewise, a number of co-occurring neuropsychiatric conditions are often seen including attention deficit hyperactivity disorder (ADHD), anxiety, aggression, and obsessive-compulsive behaviors [8-10]. Intellectual disability (ID), previously known as mental retardation, has historically been cited as being present in >70% of ASD subjects, however, recent reports put ID and ASD co-occurrence rate significantly lower at 40-55% [11, 12].

The public health impact of ASD is substantial, and support services are often required throughout the individual's lifetime. In 2006, it was estimated that the societal cost of caring for and treating an ASD individual is $3.2M over the course of that

Table 1.1: Diagnostic criteria for Autism Spectrum Disorder, DSM-V (APA, 2013)

Autism Spectrum Disorders 299.00 (F84.0)

1. Persistent deficits in social communication and social interaction across multiple contexts, as manifested by the following, currently or by history (examples are illustrative, not exhaustive; see text):
    1. Deficits in social-emotional reciprocity, ranging, for example, from abnormal social approach and failure of normal back-and-forth conversation; to reduced sharing of interests, emotions, or affect; to failure to initiate or respond to social interactions.
    2. Deficits in nonverbal communicative behaviors used for social interaction, ranging, for example, from poorly integrated verbal and nonverbal communication; to abnormalities in eye contact and body language or deficits in understanding and use of gestures; to a total lack of facial expressions and nonverbal communication.
    3. Deficits in developing, maintaining, and understanding relationships, ranging, for example, from difficulties adjusting behavior to suit various social contexts; to difficulties in sharing imaginative play or in making friends; to absence f interest in peers.
2. Restricted, repetitive patterns of behavior, interests, or activities, as manifested by at least two of the following, currently or by history (examples are illustrative, not exhaustive; see text):
    i. Stereotyped or repetitive motor movements, use of objects, or speech (e.g., simple motor stereotypies, lining up toys or flipping objects, echolalia, idiosyncratic phrases).
    ii. Insistence on sameness, inflexible adherence to routines, or ritualized patterns of verbal or nonverbal behavior (e.g., extreme distress at small changes, difficulties with transitions, rigid thinking patterns, greeting rituals, need to take same route or eat same food every day).
    iii. Highly restricted, fixated interests that are abnormal in intensity or focus (e.g., strong attachment to or preoccupation with unusual objects, excessively circumscribed or perseverative interests).
    iv. Hyper- or hyporeactivity to sensory input or unusual interest in sensory aspects of the environment (e.g., apparent indifference to pain/temperature, adverse response to specific sounds or textures, excessive smelling or touching of objects, visual fascination with lights or movement).
3. Symptoms must be present in the early developmental period (but may not become fully manifest until social demands exceed limited capacities, or may be masked by learned strategies in later life).
4. Symptoms cause clinically significant impairment in social, occupational, or other important areas of current functioning.
5. These disturbances are not better explained by intellectual disability (intellectual developmental disorder) or global developmental delay. Intellectual disability and autism spectrum disorder frequently co-occur; to make comorbid diagnoses of autism spectrum disorder and intellectual disability, social communication should be below that expected for general developmental level.

subject's lifetime [13]. More importantly, the impact of ASD on affected individuals and their families is universally life changing, and currently there is no curative treatment. Once considered untreatable, it wasn't until the late 1980s that Lovaas demonstrated early behavioral therapy could reduce symptoms of the disorder [14]. Today, it is widely believed that intervention services received at younger ages show the largest gains in cognitive and adaptive functioning in cases compared with those that did not receive treatment [15-18]. Unlike behavioral treatment, which attempts to target core symptoms of ASD, pharmacotherapy aims to decrease symptoms like aggression, compulsive behaviors, attention problems, and hyperactivity. Treatment with medications is quite common in ASD with estimates as high as 50% of ASD adolescents, and ASD adult individuals having been prescribed at least one psychotropic medication [19, 20]. That said, there is no consensus approach to treatment, and numerous approaches are often used tailored to specific behavioral problems and response to treatments with most ASD individuals receiving combinations of these over the course of their lifetimes [21, 22]. Risperidone, an atypical antipsychotic, has been shown to work in treating aggression and repetitive behavior [23]. Multiple studies have reported benefits from the selective serotonin reuptake inhibitors (SSRIs) clomipramine (Anafranil), fluoxetine (Prozac), and sertraline (Zoloft) in treating symptoms [24-26].

The public health impact of ASD is substantial and support services are often required throughout the individual's lifetime. In 2006, it was estimated that the societal cost of caring for, and treating an ASD individual is $3.2M over the course of that subject's lifetime [13]. More importantly, the impact of ASD on affected individuals and

their families is universally life changing, and currently there is no curative treatment. Once considered untreatable, it wasn't until the late 1980s that Lovaas demonstrated early behavioral therapy could reduce symptoms of the disorder [14]. Today, it is widely believed that intervention services at younger ages show the largest gains in cognitive and adaptive functioning compared to those that did not receive treatment [15-18]. Unlike behavioral treatment, which attempts to target core symptoms of ASD, medication treatments aim to decrease symptoms like aggression, compulsive behaviors, attention problems, and hyperactivity. Medication use is quite common in ASD with estimates as high as 50% of ASD adolescents and ASD adult individuals having used at least one psychotropic medication [19, 20]. That said, there is no consensus on treatment and numerous approaches are often used, with most ASD individuals receiving combinations of these over the course of their lifetimes [21, 22]. Risperidone, a dopamine antagonist, has been shown to work in treating aggression and repetitive behavior [23]. Likewise, three studies have reported benefits from the selective serotonin reuptake inhibitors (SSRIs) clomipramine, fluoxetine, and sertraline in treating symptoms [24-26].

Autism and the DSM

In his series of cases histories drawn from the records of more than 120 children seen in his John Hopkins psychiatry clinic, Kanner did not so much *define* autism as much as *portray* it. Kanner believed, the essential feature of autism was the subject's inability to relate [27], and eventually identified two features that he believed all his

patients had in common: profound “autistic aloneness” and “obsessive insistence on the preservation of sameness.” Interestingly, Kanner saw that the parents were almost as distinctive as the children and noted many were highly intelligent. Kanner went so far as to describe such parents as “successfully autistic.” [28, 29]. Thirty-years after his initial characterization, Kanner wrote: “there is a resemblance between their make-up and that of their children, except that their aloofness has not reached the gross proportions of a psychotic illness.” [30] Although Kanner wavered on this in his later years, he was initially adamant about the distinction of autism from schizophrenia, and he argued that the disorder was not associated with a specific medical conditions and was “inborn” [1]. With great insight for his time, Kanner conclude his discussion by saying: “We must assume that these children have come into the world with an *innate inability* to form the usual, biologically provided affective contact with people, just as other children come into the world with innate physical and intellectual handicaps. If this assumption is correct, a further study of our children may help to furnish concrete criteria regarding the still diffuse notion about the constitutional components of emotional reactivity” [1].

Kanner’s use of the term autism, which had been previously used to describe the *symptoms* of idiosyncratic and self-centered thinking often observed in schizophrenic patients was intended to suggest the notion of the child living in his or her “own world” [27], however, the use of this word immediately suggested something else: childhood schizophrenia. It was during this time when clinicians began applying an inclusive diagnosis of schizophrenia to children at earlier ages than ever before [31]. Unfortunately, interpretation of the origins were done in context of the period’s

prevailing psychoanalytic theories, which argued, in the case of autism, that the condition represented a child's response to an emotionally cold and distant mother. This argument was popularized by Bruno Bettelheim, and placed the brunt of the blame on the "refrigerator mother." So popular was this link between schizophrenia and autism that the second edition of the DSM, published in 1968, mentioned autism only in the context of childhood schizophrenia. In 1971, Kolvin released a series of clinical reports demonstrating that autism and childhood schizophrenia were distinctive in course, family history, and clinical manifestations. Near this time, psychologists began to criticize the "refrigerator mother" theory and stress a neurological and/or biological etiology [27, 32].

While Kanner claimed that autism was incompatible with ID, it would be shown that children with autism have a wide range of intellectual ability. Once intelligence and autism were no longer considered mutually exclusive, autism began to be described in children with known syndromes. Following the first descriptions of autism in a series of children with congenital rubella [33], autism was diagnosed in a wide number of neurological disorders. Current estimates have ~10% of all ASD individuals with a known syndrome, including Fragile-X syndrome (~1-2% of ASD cases), tuberous sclerosis (~1%), Rett syndrome (~0.5%), neurofibromatosis (<1%), and many others [34]. Studies would later uncover a high frequency of epilepsy in ASD adolescence [35], and establish solid evidence that autism has a strong genetic basis [36].

Autism first appeared as a separate category in 1980 with the publication of the DSM-3. Officially acknowledged as a "pervasive developmental disorder," this shift in

ethos was to emphasize its distinctiveness from schizophrenia, however, the DSM-3 did not explicitly address the multiple co-morbidities of ASD such as intellectual disability and epilepsy. The DSM-3 identified three essential features for autism, all developing in the first 30 months of life. They were (i) lack of interest in people, (ii) gross impairment in communication, and (iii) "bizarre responses" to aspects in the environment. Seven years later, the DSM-3-Revised dropped the age of requirement and offered a new category, "Pervasive Developmental Disorder, Not Otherwise Specified" (PDD-NOS), to encompass children who did note meet the full criteria for autistic disorder.

While more specific language was introduced into the diagnostic criteria of the DSM-IV, published in 1994, it paradoxically became more inclusive than that of its predecessor. Asperger syndrome, a condition similar to Kanner's description of autistic disorder, appeared for the first time allowing for children without language delays or cognitive impairment. Asperger syndrome, named after the Austrian pediatrician Hans Asperger was described unbeknownst to Kanner in 1944. Ultimately, the DSM- IV created a wide umbrella allowing in a tremendous range of individuals with various degrees of cognitive and social deficits. Not unlike the previous edition, the DSM- IV moved towards more inclusive diagnostic criteria for ASD (Table 1.2).

The DSM-V criteria in a move toward a framework that recognizes the spectrum nature of autism eliminates PDD-NOS, Asperger disorder, and the other autism subcategories in favor of a single "autism spectrum disorder." The three domains first introduced in the DSM-III are reduced to two and there is less flexibility from the diagnostic interviewer in choosing from diagnostic checklists (Table 1.2). For example,

Table 1.2: Changes in nomenclature and major shifts in the diagnostic criteria and definitions of autism

	Kanner and Eisenberg	Rutter	DSM-3	DSM-3-R	DSM-4	DSM-5
Date published	1956	1978	1980	1987	1994	2013
Larger category	-	-	PDD	PDD	PDD	-
Nomenclature	Early infantile autism	Infantile autism; autism; childhood autism	Infantile autism	Autistic disorder	Autistic disorder	Autism spectrum disorder
Age at onset of symptoms	None specified	By 30 months	By 30 months	During infancy or childhood	By 36 months	None specified
Related disorders	-	Other infantile psychoses	Atypical PDD	PDD-NOS	PDD-NOS; Asperger's syndrome; Rett's syndrome; childhood onset disintegrative disorder	-

PDD = pervasive developmental disorder
NOS – not otherwise specified

children must meet *all* three subcategories with social communication and two of the four with rigid or repetitive behavior. Children who have deficits only in social communication will be considered to have "social communication disorder." [37] In a historical context, the DSM-5 can be seen as an effort to accentuate the "autism spectrum" much as Kanner sought to accentuate the "distinctiveness of autism."

Autism epidemiology and risk factors

ASD affects more than 2 million individuals in the US and millions more worldwide. ASD is often cited as among the most common of the DSM defined neuropsychiatric diseases with a current prevalence in the US estimated at 1 in 88. ASD is demonstrated to equally affect all socioeconomic and geographic backgrounds [38, 39].

Today the likelihood of a child receiving an ASD diagnosis is nearly doubled from that just five years ago (72% increase) [38], and continues on a trend that began decades ago. Not surprisingly, this sharp increase in autism rates (or case identification) has generated great controversy, leading some groups to claim it an "epidemic." However, controversy and autism prevalence have gone hand-in-hand since 1943. In fact, Kanner wrote in a 1965 paper that "almost overnight, the country seemed to be populated by a multitude of autistic children" [40], and this was a time when the rate of autism was <1 in 10,000. Since the 1960s, many prevalence surveys of autism have been conducted, and a number of scientific reviews summarizing these surveys are published [27, 41].

Throughout the 1980s, ASD was believed to be relatively infrequent and considered more of an intriguing clinical dilemma than a major public health problem [42]. Studies published prior to 1985 estimate ASD at approximately 5 per 10,000 children and approximately 2 per 10,000 for the more narrowly defined condition (previously "autistic disorder"), but since 1985, there has been a dramatic increase in case identification and ASD is now thought to be second only to ID among the most common of the serious developmental disabilities, which include: ID, ASD, epilepsy, cerebral palsy, fetal alcohol syndrome and other disorders occurring before age 18) [11, 43]. The current prevalence estimate of 1 in 88 represents a two-fold increase from a decade ago, a ten-fold increase from the mid 90s, and a dramatic increase of more than twenty fold since 1985 (Figure 1.1).

Some have seen the rise in ASD rates as evidence of an "epidemic," identifying potential culprits like vaccines or environmental triggers [39], while others see it as a blessing with more children being identified earlier than before and more becoming eligible for special education services. While the origins of this substantial increase in prevalence estimates remain under debate, it is largely believed to reflect changes in how individuals are identified, and not necessarily a "true" increase in the number of ASD cases [39]. It is hard to contest the fact that the ever-changing diagnostic criteria have confounded the task for epidemiologists. The heightened public awareness, due in large part to efforts of parent and advocacy groups, the availability of more medical and educational resources, increased media coverage, and more training and information for physicians and service providers have complicated interpretation of the data [11].

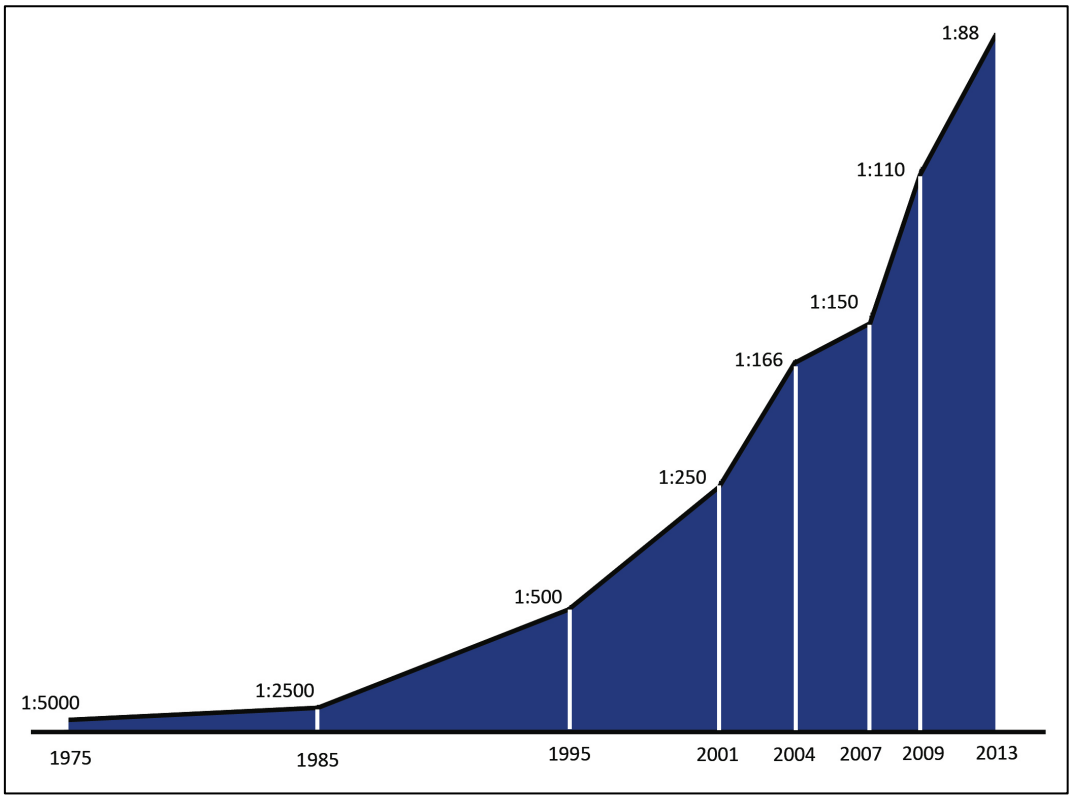


Figure 1.1: Rising rates of ASD prevalence. Figure is reproduced from estimates according to CDC's Autism and Developmental Disabilities Monitoring (ADDM) Network. Equivalent prevalence rates are reported to occur in all racial and socioeconomic groups. Many factors are responsible for the increasing rate of cases, however, it is believed in part to reflect increased awareness and a more inclusive diagnostic criteria. [38, 39]

Overall, many factors are at play, and it has proven quite challenging to dissect out legitimate factors contributing to the observed increase in rates.

The largest effect and most well known risk factor for an ASD diagnosis is a family history of a sibling or parent diagnosed with autism. Numerous studies suggest that the ASD phenotype extends beyond the DSM boundaries to a "subclinical" realm known as the broader autism phenotype (BAP); for example, studies have shown an increase in subclinical ASD-like features among the relatives of ASD cases compared to controls [44]. Twin and family studies have strongly suggested that ASD is principally a genetic condition, however, the genetics are highly complex exhibiting profound heterogeneity (discussed in Chapter 2), and opportunity for epigenetic and gene-gene interaction effects. An unexplained yet robust sex bias exists, with males being more likely affected (4.3:1 Male:Female ratio). This sex bias is modified substantially by cognitive impairment, so when ID and ASD are co-morbid the sex ratio is approximately 2:1 [45]. In addition to being male and having a family history of ASD (or BAP), other risk factors are more difficult to detect. Of them, the two most replicated are advanced paternal age and a family history of psychiatric conditions.

Less certain than (i) family history, (ii) sex, and (iii) paternal age, are possible prenatal risk factors such as maternal infection and immune dysfunction[46]. Perinatal complications and obstetric factors such as labor and delivery complications are also additional factors [47]. Environmental exposures including prescription drugs, metals, and chemicals have all been implicated in *extremely rare* ASD cases [47]. Gardener et al. perform a meta-analysis on more than 50 prenatal factors studied in 64 epidemiological

ASD studies and conclude that evidence suggests prenatal conditions *will* increase risk, yet at the time of publication claim insufficient evidence to implicate any specific prenatal factor [48]. Unfortunately, studies looking at neonatal risk factors have relatively small sample sizes and the results are often contradictory. Despite evidence for some associations with ASD, it remains unclear if these neonatal risks are causal or play a "secondary" role in determining the clinical manifestations of an individual with a genetic predisposition to ASD [49].

There was much speculation about the initial suggestion that the measles, mumps, and rubella (MMR) vaccine might be an autism risk factor [50], however, the initial report was retracted [51] and today's scientific consensus rejects any epidemiological relationship between vaccines containing thimerosal and/or mercury and ASD [52, 53]. On the other hand, prenatal exposure to valproate has been demonstrated to increase ASD risk, and it has also been suggested that prenatal psychotropic medication exposure, particularly during early gestation, *modestly* increases ASD risk [54]. Lastly, recent findings on parental education and socioeconomic status have put the once believed associations with ASD in doubt. It is now believed that the long suspected associations were a result of ascertainment bias, and no association between ASD socioeconomic status exist [55-57].

Summary

In summary, the history of autism spectrum disorder is fascinating and has provided a new appreciation for the progress made and the challenges that lie ahead.

Diagnostic criteria have become more inclusive over time reflecting the growing conviction that autism represents a wider spectrum phenotype with varying degrees of ability and cognition. This change in ideology largely explains the increasing prevalence rates. Hundreds of potential risk factors have been examined for a correlation to autism, however, the most consistent risk factors (family history and sex) are also the ones with the largest effect sizes. It is currently believed that the etiology of ASD predominately resides in the genes we are born with, and an unknown number of non-heritable factors also influencing risk in an unknown way.

CHAPTER II

Autism as a complex genetic disease

Introduction

It is now accepted that genetics contribute substantially to ASD risk, however, identifying specific genes involved has been *difficult* to detect. The fact that many early risk loci implicated by genome-wide linkage and then genome-wide association failed to replicate in later studies, it is often misunderstood that research geneticists have not identified many ASD genes. This misconception suggests that the rarity of many of these risk variations have made it difficult to appreciate the amount of progress made [58]. The fact remains that the efforts of many talented individuals have identified *hundreds* of genes and pathways contributing risk to the disorder.

It was the classical twin studies of concordance in the late 1970s that first established a strong genetic component for ASD [36, 59]. When considering the broader autism phenotype (BAP) concordance rates for dizygotic and monozygotic twins are estimated at 10% and 92%, respectively [36, 60, 61]. Furthermore, reports of sibling recurrence rates, was once estimated to be 4.5%, is now estimated to be ~25% [62-65]. Likewise, the approximate two-fold increase in concordance in full siblings compared to half siblings [66], and the observed higher rate of BAP in first-degree relatives in ASD multiplex families [44] support a genetic etiology.

Once established to be predominately genetic it was hoped that finding ASD risk genes would come easy, however, established risk loci proved extremely difficult to

characterize due to the genetic complexity of the disorder [34, 67, 68]. This genetic complexity derives from a combination of allelic and locus heterogeneity, variable expressivity, and incomplete penetrance. In spite of this complexity it is estimated that researchers can account for ~ 40% of the attributable genetic risk (Figure 2.1) [69].

Numerous "pathogenic" loci have been identified in ASD individuals with known Mendelian or genetic disorders that exhibit a phenotype inclusive of ASD. Rare CNVs detectable by (clinical or research) Chromosomal Microarray (CMA) testing estimates that ~10-15% of ASD cases are afflicted with a causal single gene, copy number, or chromosomal abnormality [34]. Some of the most frequently identified recurrent genomic factors interpreted as pathogenic include CNV duplications at 15q11-q13 (seen in ~1% of ASD cases). Fragile X (~1-2 %), tuberous sclerosis (~1%), Rett syndrome (~0.5%), Williams syndrome, Smith-Lemli-Opitz syndrome and other less common disorders have also been observed with ASD [70, 71].

Some of the earliest genetic studies used microsattelite markers spanning the genome (or at least autosomes) and tested for evidence of genetic linkage in samples of multiple ASD families. Other studies used family-based collections or case-control samples to test for common variant association at candidate genes. Virtually all of these studies assumed a common variant-common disease risk scenario. These studies identified genetic linkage and genic SNP associations to every chromosome, and in the majority of cases, these linkage or associations were not replicated. This exemplifies the Winner's Curse phenomenon, frequently observed in studying the genetics of highly complex disorders [72]. Associating risk let alone causality to any particular finding

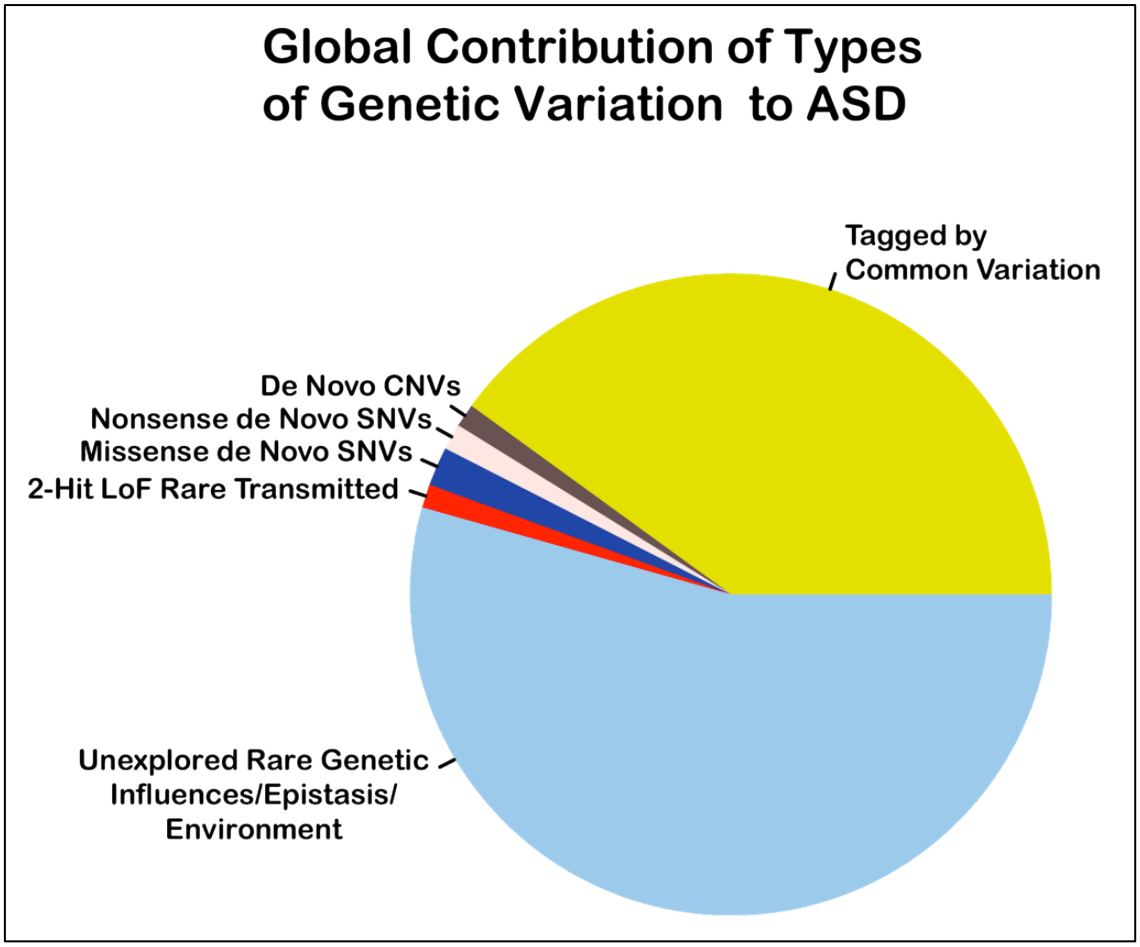


Figure 2.1: The percentage of variance explained by various forms of genetic risk factors for ASD. Common variants *acting additively* capture a large percentage of population risk for ASD, though each locus is expected to be of small effect, whereas, rare variants explain a smaller proportion of risk, each locus is expected to be of larger effect. Since syndromic forms of ASD (estimates as high as 15%) are typically excluded from most population-level studies they are not included here. Figure is reproduced from Stein et al. *Neuron* 77 (2013).

proved difficult for two reasons: few rare variant loci were common enough in the ASD population and many positive common variant associations would fail to replicate in later studies.

Newer methodological approaches such as microarray-based (SNP) genome-wide association (GWA) studies in very large samples sizes and next-generation sequencing (NGS) have changed the landscape of ASD genetics and have provided tremendous insight into the genetic architecture of ASD. In the past few years alone, well over a hundred of genes and dozens of genomic loci have been associated with ASD and/or ASD-related syndromes (Reviewed in [58]).

Contributions of common and rare variation in ASD

Considering the potential (and emerging) nature of ASD genetic susceptibility, it is important to understand in general potential genetic frameworks in which the frequency (common vs rare) of risk-conferring variants might contribute to the genetic liabilty of this as one example of complex disease [73]. As stated above, the genetic architecture of ASD is highly complex with tremendous heterogeneity, therefore two models (not mutually exclusive) have been proposed to explain genetic risk in common complex disorders like ASD: (i) *Common Disease – Common Variation* (CD-CV), and (ii) *Common Disease – Rare Variation* (CD-RV). CD-CV suggests that common variants, those found widely distributed in the population, are associated with *small* to modest increase in baseline (population) risk, while the CD-RV model suggests that a number of rare, and sometimes *de novo*, variants of *high* to major risk effect underlie the disorder. This idea

is illustrated in Figure 2.2 and *the CD-CV and CD-RV models have very different implications regarding gene-discovery strategies*. During the course of these studies considerable evidence supporting both hypothetical extremes has emerged for ASD, however, the most realistic genetic framework of ASD includes functional variants that span the territory between the two (common or rare) extremes.

The contribution of common alleles

Since ASD (and subclinical traits) is common in the general population, it was thought that common variants would be the prevailing mode of risk heritability (i.e. the CD-CV model). First tested in studies using candidate gene SNP genotyping or variable number tandem repeat (VNTR) polymorphism genotyping and subsequent association analyses, and later by large "unbiased" GWA studies. After much effort by multiple groups over a short span of approximately 4 years, there is considerable evidence that individual common variants of main effect also do not contribute substantially to ASD genetic liability [74-78]. However, we now understand that when common variation acros the genome is analyzed in the aggregate, a very different picture emerges in which many common variants that may include those conferring small main effects on ASD risk (ORs 1.05-1.20), can act additively to contribute substantially to genetic risk [79, 80].

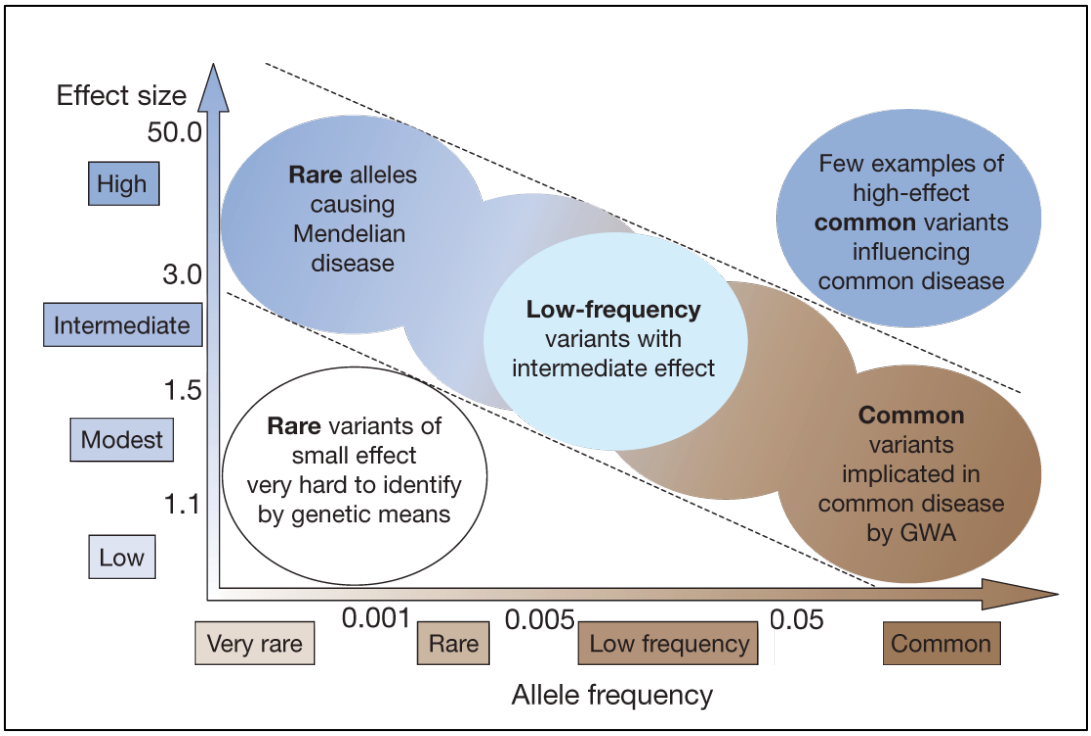


Figure 2.2: Feasibility of identifying genetic variants by risk allele frequency and strength of genetic effect (odds ratio). Most emphasis and interest lies in identifying associations with characteristics shown within diagonal dotted lines. Figure is reproduced from Manolio et al. *Nature* 461 2009.

Until recently, the impact of common variations in ASD risk was largely assessed not by GWA studies but more by selecting *candidate* genes based on their biological plausibility and what was known about ASD neurobiology. These studies assessed the contribution of one (or several) common allele(s) in or near these hypothetical risk loci. For many candidate genes considered, the majority did not detect statistically significant associations and those that did often do not replicate convincingly in independent datasets (the "Winner's Curse"). While only a minority of the common associated alleles identified in the pre-GWA era yielded solid evidence for association, even fewer exhibit functional effects, however, a handful of findings continue to warrant attention. SNPs associated with the serotonin transporter (SERT; gene symbol: *SLC6A4*) [81-83], the receptor tyrosine kinase (*MET*) [84-86], loci within 15q11-15q13 that encodes the E6-AP ubiquitin protein ligase (*UBE3A*) and three $GABA_A$ receptor subunits (*GABRB3, GABRA5,* and *GABRG3*) [87-90], the contactin-associated protein 2 (*CNTNAP2*) [91, 92], orthologous to *Drosophila* neurexin 4, integrin beta 3 (*ITGB3*) [83, 93, 94], and reelin (*RELN*) [95-97] genes are several examples of positive associations that have been identified using a candidate gene approach.

Concurrently with amassing sample sizes and improvements in microarray SNP platforms, there was a general shift from candidate gene studies to the GWA (i.e. hypothesis independent) studies. This change in approach was matched by a focus on larger sample sizes, SNP markers capturing a greater degree of linkage disequilibrium across the genome, and on the development of pipelines for GWA methodology and analysis. At present, three large, independent GWA studies have been performed, with

each looking at approximately one-half million to one million SNP markers [75-77]. In these three studies, *respectively*, GW significant association was reported at three different loci rs4307059 at 5p14.1, rs10513025 at 5p15.2, and rs4141463 at 20p12.1, replicating across all the other studies. Noticeably, all associated SNPs in candidate genes are absent at GW significant loci [75-77], however, some candidate genes such as RELN and SLC25A12 were represented in nominally significant SNPs in sub-GW-significant space. While well planned and executed, all GWA studies remain dramatically underpowered compared to those in another field, schizophrenia, which now show replicated common alleles at >100 loci suggesting many thousands of more ASD samples are needed to reach the inflection point at which CV association studies will detect replicable association. Overall, GWA studies in ASD have not yet demonstrated convincing evidence of *individual* common variants of main effect. (Table 2.1)

The contribution of rare variants

In the last decade, direct sequencing of candidate genes has implicated a number of related genes as risk factors in ASD. Early examples include the discovery of mutations in the X-linked neuroligins 3 and 4 (*NLGN3* and *NLGN4*) loci, which encode neuronal adhesion molecules present in the postsynaptic density (PSD) of glutamatergic synapses [98]. Later studies identified rare functional mutations in *NRXN1*, encoding a presynaptic binding partner for neuroligins [99, 100], and multiple SHANK genes (*SHANK2, SHANK3*) which act as postsynaptic scaffolding molecules that interact with

Table 2.1 ASD associated variants revealed by GWAS

Study	Stage	Families	Type	Findings
Wang *et al.* (2009)	1	780, 3101 participants	Multiplex	Top SNP rs4307059 ($p = 1.1x10^{-9}$) located between *CDH8* and *CDH10**
	2	1204 cases, 6491 controls	Case/control	Top SNP rs4307059 ($p = 2.2x10^{-4}$)*
	3	Combined		Top SNP rs4307059 ($p = 3.4x10^{-8}$)
Weiss *et al.* (2009)	1	1031 families, 4233 participants	Multiplex and simplex	Top SNP rs10513025 ($p = 1.7x10^{-6}$) located between *SEMA5A* and *TAS2R1**
	2	2073 Trios	Multiplex and simplex	Top SNP rs10513025 ($p = 2.1x10^{-7}$)*
Anney *et al.* (2010)	1	1369 families, 1385 cases	Multiplex and simplex	Top SNP rs4141463 ($p=2.1x10^{-8}$) in *MACROD2* (intronic)
	2	2179 families	Multiplex and simplex	Top SNP rs4141463 ($p=4.7x10^{-8}$) in *MACROD2* (intronic)

* Not genome-wide significant

PSD95 [101-103]. Candidate gene sequencing studies implicated *ADORA3*, *NRXN3*, *PTCHD1/PTCHD1AS*, *ASTN2*, *CDH8* and *CNTNAP2* as harboring ASD risk loci [101, 104-109].

Paradoxically, it appears the microarray fluorescent intensity and SNP genotype data garnered from what were also designed as GW CV association studies have arguably made the most important contributions in the study of rare variation with their ability to detect copy number variation (CNV) [110]. CNVs, often in the range of 1 kb – (>)10 Mb, are known to contribute to ASD susceptibility in ~10% of cases [71, 111]. Typically too small for karyotype detection, CNVs ≤ Mb can involve a single gene or comprise many genes. In only a few years, CNV discovery and analysis have produced, a rapid, reliable, and unbiased look at variations of that genomic size. Sebat and colleagues (including families from our collection) were the first to show that rare *de novo* variations (DNVs) were over-represented in ASD simplex families compared to controls, and compared to families in which there were multiple affected individuals [112]. Since then, five additional studies have been conducted to look at *de novo* and inherited CNVs (Table 2.2). The observed excess of *de novo* CNVs in cases compared to controls has been repeatedly replicated [101, 112-114], on the other hand, inherited CNVs are *not* observed to be increased overall in cases compared with controls [112-116]. Longstanding precedence in study of human genetic disease attributes (in general) a greater probability of rare *de novo* CNVs as being likely risk factors that confer risk of relatively large effect size. In clinical practice this can range from rare *de novo* CNVs observed in multiple unrelated ASD cases but not controls being interpreted as

pathogenic, to such *de novo* CNVs observed significantly more often in cases than controls, or rare inherited CNVs transmitted to ASD cases but not seen in control as risk factors, but not necessarily "pathogenic." They confer large effect sizes, have large ORs, but not be fully penetrant by virtue of "unaffected" parental carriers, or rare "controls" that are carriers but not affected by ASD. Rare, recurrent deletions of 16p11.2 are a prime example of this phenomenon. A significant majority of 16p11.2 deletions are *de novo,* found in cases typically presenting clinically with an ASD phenotype. However, infrequent instances of transmitted deletions have been described. When carefully examined, the carrier parent may have subclinical features of ASD or other neuropsychiatric phenotypes, but clearly not ASD [111]. As a general phenomenon, genomic variants arising *de novo* in an affected individual are likely to confer effect sizes of OR >5 [114].

Cumulatively identified in less than 2% of "idiopathic" ASD, CNVs at 16p11.2 and 15q11–q13 are two of the most frequently observed "pathogenic (i.e. causal) CNVs" in ASD cases with estimates at ~0.8% and ~1% respectively [112, 116, 117]. In the case of the 16p11.2 locus, the gestalt across many studies is that duplications confer a lower risk, and more likely to be inherited in ASD compared with deletions. Further, duplications *or* deletions at this locus demonstrate increased liability for schizophrenia and ID [112, 117, 118].

Table 2.2 Large-scale CNV studies that reveal common and rare variants associated with ASD

References	Participants	All CNVs (de novo and inherited)	Cases-controls	De novo CNVs only	Recurrent CNVs in cases	Single genes, ASD-associated genes in *de novo* CNVs
Sebat et al. (2007)	264 families (simplex, multiplex and control)	CNVs only	-	7.2% of cases; 1% of controls	De novo: 2q37.3, 3p14.2	*SLC4A10, FHIT, FLJ16237, A2BP1*
Szatmari et al. (2007)	173 families (multiplex) 196 cases, 292 unaffected sibs	624 total. No difference in frequency of CNVs in cases vs. controls	1.2 CNVs/genome cases (mean = 3.4 Mb) and 1.27 CNVs/genome controls (mean = 4.2 Mb)	5.1% of cases, 2.1% of controls	47 CNVs in cases	*NRXN1*
Marshall et al. (2008)	427 families (simplex and multiple) 427 cases and 1652 controls	2873 total. No difference in frequency of CNVs in cases vs. controls	3.1 CNVs/genome cases (mean = 603 kb) and 3.1 CNVs/genome in controls (mean = 470 kb)	27 ASD cases with DNV not found in controls	De novo: 5p15.31-p15.2, 7q31.1-q32.2, 15q11.2-q13.3, 16p11.2	*CDH18, DPYD, NLGN4, DPP6, DLGAP2, ANKRD11, SHANK3*
Pinto et al. (2010)	876 families (simplex and multiplex) 996 cases, ~3000 controls	5478. 1.19 fold increase in cases over controls	219 inherited CNVs not in controls	50 out of 876 ASD harbor de novo events.	-	*SHANK2, SYNGAP1, DLGAP2*
Sanders et al. (2011)	1124 families (SSC), 1124 cases, 872 unaffected sibs	No difference in frequency of inherited CNVs in cases vs. controls	-	5.9% of cases; 1.7% of sibs OR3.5	De novo: 1q21.1, 7q11.23, 15q13.2-q13.3, 16p11.2, 16p13.2	*CDH13, USP7, C16orf72*
Levy et al. (2011)	887 families (SSC) 858 cases and 863 unaffected sibs	No difference in frequency of inherited CNVs in cases vs. controls	-	7.9% in cases (median genes/CNV = 4), 2% in unaffected sibs (median genes/CNV = 0)	De novo: 16p11.2, 7q11.23, 16p13.2	*DDX53-PTCHd1, USP7*

Table modified and reproduced from Berg and Geschwind, *Genome Biology* 2012 13:247

Table 2.3 Whole-exome gene sequencing studies that reveal common and rare variants associated with ASD

Ref	Families	*De novo* variant characteristics	Single gene implicated	Novel findings	Trends
Neale et al. (2012)	175 families[1], simplex, multiplex and trios	Equal frequency of *de novo* mutations in cases and controls	*KATNAL2* and *CHD8* harbor *de novo* mutations in cases, none found in controls	Enriched number of protein interactions among genes with functional[3] *de novo* mutations	Greater paternal age correlates with *de novo* mutation number
O'Roak et al. (2012)	209 families[2], simplex, quads, trios	Equal frequency of *de novo* mutations in cases and controls	ASD cases harbor protein disrupting mutations in *GRIN2B, LAMC3,* and *SCN1A*.	β-catenin chromatin remodeling protein network enriched for ASD candidate genes	Greater paternal age correlates with *de novo* mutation number
Sanders et al. (2012)	238 families[2], simplex, trios, and quads	Equal frequency of *de novo* mutations in cases and controls	*SCN2A* significantly associated with ASD. *KATNAL2*, *CHD8*, and *SCN2A* associated with ASD when combined with Neale et al. and O'Roak et al.	Greater non-synonymous and nonsense *de novo* mutations in cases than unaffected sibs	Greater paternal age correlates with *de novo* mutation number
Iossifov et al. (2012)	343 families[2], simplex and quads	Equal frequency of *de novo* mutations in cases and controls	*KATNAL2, CHD8, SCN2A, DYRK1A,* and *POGZ* associated when combining all studies	Enriched functional[3] *de novo* mutations in cases than in unaffected sibs	Greater paternal age correlates with *de novo* mutation number

[1] Boston Autism Consortium
[2] Simons Simplex Collection

[3] Functional is defined as missense, nonsense, consensus splice site, and frameshift

Table modified and reproduced from Berg and Geschwind, Genome Biology 2012 13:247

Similar to observations related to 16p11.2, additional observations of identical (or overlapping) CNVs were discovered and associated with autism, ADHD, epilepsy, ID, and schizophrenia. After these discoveries, it soon became clear that variable (phenotypic) expressivity was a central feature of CNV risk, and likely other ASD risk-related variation [111]. In other words, many well-recognized "ID CNVs" may also be observed in the context of "autism," or similarly "schizophrenia." Identical highly penetrant CNVs that confer large risk effects may be seen in different individuals presenting clinically with a range of disorders such as ID, schizophrenia, and autism. These findings suggest that a number of genes (haploinsufficiency, duplication or other genic variation in which) cause a continuum of neurodevelopmental disorders, which may manifest themselves in very different ways. Notably, these highly penetrant factors occur (or re-occur) on different genetic backgrounds, and other loci in the genome may influence, even compensate for, phenotypic expression in any one individual. It is believed that specific CNVs will exhibit different penetrance and expressivity depending on the copy number sensitivity and function of the gene or genes they affect [111]. It is interesting to speculate how reciprocal changes in copy number (e.g. 1 vs. 3) can have diverse outcomes. A fascinating example, involves CNVs at 7q11.23, duplications of which are implicated in ASD, while deletions lead to Williams syndrome, and in schizophrenia (duplication/deletion) [114, 119].

Many CNVs observed to carry large effect sizes in ASD are identified as such because they arise as *de novo* events, and thus not present constitutionally in parents and not present in the general or control populations (e.g. 16p11.2 deletion; 15q11-q13

duplication; 1q21.1 deletion; 15q13.3 deletions, etc.) [120, 121]. Inherited CNVs on the other hand, are *a priori*, presumed to confer less severe effects sizes in comparison (e.g., 15q11.2 deletion, 16p11.2 duplications, etc), and may require additional risk factors in order to cross a threshold for expression of disease [99, 122]. Many presumed inherited risk-conferring CNVs may also demonstrate incomplete penetrance, and if true, this may help explain how many CNVs are observed in population controls [34].

De novo single nucleotide variants (SNVs)

Advances in high throughput or massively paralleled sequencing provide a powerful new approach to discover rare *de novo* and inherited variants [123]. In 2012, a series of reports in Nature, Neuron and PLoS Genetics presented results from large scale ASD WES studies. All together, four groups (one of which included our lab) conducted WES on approximately 1000 ASD families and over 1000 independent cases and controls [124-127]. Strikingly, across all studies (Table 2.3) a positive correlation between increased paternal age and number of "functional" DNVs was detected, which may help explain the paternal-age-dependent risk of ASD seen in epidemiological studies [128-131]. Another commonality between these studies was the finding that the rate of DNVs overall is not increased in ASD cases compared with controls, however, when studies are pooled and the *class* of variation is restricted to loss of function mutation (premature termination, consensus splice site), a significant increase in rate of DNV present in cases vs. controls. Additionally, genes involved are not random and cluster onto specific biological themes, for example chromatin remolding, or the β-catenin signaling

pathways [124, 125]. It is not how many DNVs a subject carries, *per se*, it is what type and where these variants are located in the genome that ultimately affects risk. In this respect, it should not be surprising that an enrichment of highly damaging (i.e. nonsense, frameshift, consensus splice site) DNVs are present in ASD individuals compared to chance expectation. Nonsynonymous *de novo* point mutations are extremely rare, with estimates at slightly more than one "functional" DNV per exome. Initial reports highlighted no single gene that harbored more *DNVs* than expected by chance, however, all studies demonstrated an increase in *de novo* nonsense variants. In total there are three genes (*CHD8*, *KATNAL2*, and *SCN2A*) that were shown to have recurrent nonsense DNVs.

Challenges and opportunities

Defining and understanding pathways or mechanisms of dysfunction *alongside* identification of ASD risk genes are critical steps moving forward. Efforts to explain ASD's genetic risk in the context of neurobiological pathways or relevant biology have faced many challenges due in large part to the genetic heterogeneity of the disorder. With new and suspected ASD susceptibility loci becoming more frequent with every publication, these newly discovered loci (often times disparate) may be the key to exposing relevant ASD biology. To do this, studies need to go beyond single mutations and their molecular dissections and examine the variations in a biological context. Leading this effort have been multiple studies demonstrating associated SNPs and CNVs converge onto biological pathways. For example, the first mutations identified in

idiopathic autism involved glutamatergic synaptic genes like *NLGN3* and *NLGN4X* or *SHANK3*. Later CNV and DNV analysis results support these first findings for a role of synaptic genes in ASD [132]. In the past few years, dozens of biological pathways have emerged from literature as risk factors, however, two that have garnered the most scientific support include those in chromatin modification (e.g., *CHD8*) and ion channels (e.g., *SCN2A*) [124, 126, 127, 133].

Summary

The data presented in this chapter support the view that the genetic architecture of autism spectrum disorder is highly complex involving hundreds of loci, with variation being inherited and *de novo*, common and rare, and existing in a spectrum of penetrance and expressivity. In spite of these tremendous challenges, much progress has been made with recent advances in genetic methodologies providing significant insights into the biology of ASD. Furthermore, by examining converging pathways researchers are beginning to gain a more complete picture of ASD biology.

It is worth stressing that many of these genetic findings and interpretations would not have been possible without the availability of large cohorts of ASD and control participants and shared databases. In particular, the formation of the Autism Genetic Resource Exchange (AGRE), the Simons Simplex Collection (SSC), the Autism Genome Project (AGP), and the NIH ARRA Autism Sequencing Consortium were critical. The last three of which the Sutcliffe lab has played an important role. Many important web resources cataloging ASD genetic liability are equally important in these

discoveries. They include the National Database for Autism Research (NDAR) (http://ndar.nih.gov), the SFARI Gene database (https://gene.sfari.org/autdb/), the AutDB database (http://mindspec.org/autdb.html) and the Autism Chromosome Rearrangement Database (http://projects.tcga.ca.autism).

CHAPTER III

Monoamine transporter dysfunction as a genetic and biological risk factor in psychiatric illness

Introduction

The monoamine neurotransmitters serotonin (5-HT), dopamine (DA), and norepinephrine (NE) play important roles in mood, cognition, learning, reward, sleep, and motor activities. The availability of extracellular 5-HT, DA, and NE is maintained and tightly regulated by their presynaptic transporters, SERT, DAT and NET respectively [134, 135]. Monoamine transporters serve as targets for antipsychotics, antidepressants, and drugs of abuse, emphasizing the important role these molecules play in neurotransmission and pathologies of psychiatric disorders. [134, 135] Among these drugs include fluoxetine (Prozac™), methylphenidate (Ritalin™), 3,4-methylenedioxy-*N*-methylamphetamine (MDMA; "ecstasy"), and cocaine.

The monoamine transporters, SERT, DAT and NET belong to a group of SLC6 (solute carrier family 6) Na^+ and Cl^- dependent neurotransmitter symporters [134, 136-139]. Each monoamine transporter type is responsible for reuptake of their respective neurotransmitter by sequential binding and transport of Na^+ and Cl^- ions. In the case for DAT, two Na^+ ions and one Cl^- are required for transport, while SERT and NET, only one Na^+ and one Cl^- ion are required [134]. The monoamine transporters in this family share numerous structural features, including twelve transmembrane spanning domains, a large second extracellular loop, multiple post-translational modification sites, and

cytoplasmic amino (-NH_2) and carboxy (-COOH) termini. The monoamine transporter's primary function is reuptake of synaptic neurotransmitter for further repackaging for later release.

The reuptake ability of the monoamine transporters is directly influenced by the numbers of cell surface transporters available, and by the activity of these individual transporters [134]. Regulation of monoamine transporter activity is primarily accomplished through phosphorylation and post-translational modifications. Perhaps the best understood of these is phosphorylation by the second messenger protein kinase C (PKC). PKC has been shown to downregulate SERT, DAT, and NET function through internalization of the transporter from the cell surface to cytoplasmic vesicles [134]. In addition, modulation of monoamine transporters may occur through formation of protein-protein complexes including syntaxin 1A, secretory carrier membrane protein 2 (SCAMP2), transforming growth factor beta-1-induced transcript 1 protein (Hic-5), neuronal nitric oxide synthase (nNOS), and integrin β-3. [140]

SERT in psychiatric disease

SERT and its gene *SLC6A4* are perhaps the most widely studied in psychiatry. 5-HT plays an important modulatory role in emotion, motivation and cognition, and disruption of normal 5-HT levels is implicated in conditions including affective disorders, schizophrenia, and impulse control disorders. [141] Selective serotonin reuptake inhibitors (SSRIs) are drugs that – with varying degrees of selectivity – target SERT and are used in treatment of depression, obsessive-compulsive disorder, and anxiety

disorders, thus making these antagonists the most prescribed group of compound dugs [141].

Given the significance of SERT in regulating 5-HT function and being a target for many widely used medications, the serotonin transporter gene, *SLC6A4* has been an attractive gene for study in neuropsychiatric disorders. Genetically speaking, *SLC6A4* has been associated with a number of neuropsychiatric disorders including obsessive compulsive disorder (OCD) [142, 143], anorexia [144, 145], social anxiety disorder [146], seasonal affective disorder (SAD), alcoholism [147, 148] and major depressive disorder (MDD) [149, 150] (Table 3.1).

A significant focus of *SLC6A4* genetic studies involves a common 44-base pair insertion/deletion of a repetitive sequence (a.k.a. 5-HTTLPR) in the promoter region reported to impact *SLC6A4* gene expression [146]. The long allelic version or "L" is higher expressing while the more common and short allelic version or "S" displays lower expression [151]. Quite simply, the length of the promoter has been shown to influence mRNA, protein, and 5-HT uptake levels in lymphoblastoid cell lines (LCLs), platelets, and serotonergic neurons [146, 152, 153], although one report has shown that SERT mRNA is independent of 5HTTLPR genotype in post-mortem pons tissue sections containing serotonergic neurons[154]. Most reported findings of association have pointed to the more frequent S-allele as a risk factor in anxiety-related behaviors and for increased risk of depression [155]. Homberg and Lesch (2010) argue because of its high allele frequency in the general population, any negative consequences of carrying a copy of the S-allele (e.g. anxiety related behaviors) would be explained evolutionarily if it were

Table 3.1: Association of *SLC6A4* with human disease including genetic variation and imaging studies

Gene Symbol	Aliases	Disease Association	Gene Locus	Disease Association from SNPs		Imaging Technique	Biomarker Association	Tissue Distribution
						PET		
SLC6A4	SERT	Anxiety	17q11.1-q12	I425V, F465L, L550V	OCD	SPECT	Depression	Brain
		Autism		I425V	Anorexia nervosa		Neuroticism	PNS
		Depression		L550V; I425L			Mood disorders	Placenta
		Gastrointestinal		G56A, I425L, L550V	Rigid compulsive behavior		Bipolar disorder	Epithelium cells
		Schizophrenia		G56A, I425L, F465L, L550V, K605N	Autism		Impulsivity	Platelets
		OCD		VNTR 9	Smoking		Aggression	
				VNTR 10	Migraine		Depression w/ bipolar disorder	
							Parkinson disease	
				5HTTLPR *short*	Suicide		Major depressive disorder	
					Autism		Autism	
					Antidepressant response			
					Depression			
					Alcoholism			
					PTSD			
					Drug abuse			
				5HTTLPR *long*	OCD			
					Suicide behavior			
					Antidepressant response			
					Schizophrenia			

Table modified from original production of Pramod et al, *Molecular Aspects of Medicine* 34:2013

counterbalanced or completely offset by an evolutionary gain [156]. They argue this may additionally help explain the inconsistent reports but more importantly "also lead to a more refined appreciation of allelic variation on 5-HTT function" [156].

SERT knockout mice have provided a useful tool to examine mechanisms by which SERT may regulate emotional behavior. Numerous studies have shown alterations in multiple behavioral phenotypes including anxiety, depression-like symptoms, and altered responses to drugs of abuse [157-161]. In fact, altered SERT function in mice is associated to more than 50 different phenotypic changes *in vivo* [162, 163]. Finally, SERT knockout mice have reduced platelet and brain levels of 5HT, and elevated synaptic 5-HT as measured by *in vivo* microdialysis [164].

Evidence for serotonin transporter involvement in autism spectrum disorders

It has been suggested that SERT dysfunction contributes both to the (i) genetic predisposition and (ii) neuropathologies of ASD [165-167]. Consistent with a SERT dysfunction hypothesis are clinical and scientific findings demonstrating heritable hyperserotonemia, or elevated levels of platelet 5-HT, in an estimated 25-40% of ASD families [168-172]. Hyperserotonemia, a heritable trait, is the oldest and most replicated biomarker in autism [172]. The presence of SERT on the platelet surface and its role in acquiring 5-HT from the blood provides a plausible biological mechanism for SERT involvement in hyperserotonemia. Several variables have shown correlations for platelet 5-HT levels including age, sex, and ethnicity [173].

Supporting a serotonin role in ASD are findings of reduced symptoms in SSRI treatment [24, 26, 174, 175], a worsening of symptoms in individuals depleted for tryptophan (a 5-HT precursor), and a single PET study showing altered 5-HT synthesis in some children with ASD. [24, 26, 174-179]. Moreover, studies of an animal model treated with a serotonin analog that reduces available 5-HT levels have lead to abnormal behaviors in postnatal animals reminiscent to those seen in ASD [162, 163, 180, 181]. Collectively, these data provides a compelling case for SERT's potential involvement in ASD.

Similar to other psychiatric disorders, most ASD studies have focused on the common *SLC6A4* alleles of 5-HTTLPR and VNTR, and show inconsistent results with both the S and L alleles being over-transmitted to ASD cases. Confounding many of the 5-HTTLPR analyses are two SNPs in the promoter region (rs25531 and rs25532) that can modulate activity of the L allele and reduce its expression [82, 155, 182, 183]. Despite this, significant linkage to the *SLCA4* locus (Chr17q11.2-q12) is documented and replicated [81, 184]. To ask if rare variants contribute to the linkage, Sutcliffe and colleagues resequenced exons in families with the greatest degree of linkage (allele sharing) and identified four *SLC6A4* coding variants in several of these "linked" families in which affected individuals had relatively more severe rigid-compulsive traits, as indexed from items on the ADI [81, 185]. When these variants are introduced into a human SERT expression construct, all *SLC6A4* ASD-associated variants display elevated SERT activity (i.e. gain-of-function effect). This suggests a possible mechanistic relationship between 5-HT availability and ASD [81, 186-188]. The commonalities of

altered transporter function revealed by the *SLC6A4* variants warrant special attention given the substantial number of ASD individuals who are hyperserotonemic.

Sequencing the coding region of *SLC6A4* in unrelated ASD probands from families without obligate allele sharing (linkage) at 17q11.2, another group did not detect a similar pattern of enriched coding variants [189]. It is likely that screening of multiplex probands with allele sharing in this region was an important factor in the initial discovery of novel, functional variants. Akin to Sutcliffe et al., the results from Ozaki et al. support a role of *SLC6A4* coding variants in ASD. They report an Ile425Val variant in pedigrees harboring multiple psychiatric phenotypes, including Asperger's, OCD, and anxiety disorders being the most prominent [144, 190].

Functional characterization of the Sutcliffe et al. ASD-associated SERT variants reveal each elevate 5-HT transport function via altered protein kinase G (PKG) and/or p38 mitogen activated protein kinase (MAPK) regulation [187, 188]. Characterization of one of these variants (Gly56Ala) in knock-in transgenic mice revealed elevated 5-HT clearance and a p38 MAPK-dependent transporter hyperphosphorylation *in vivo.* These effects were accompanied by deficits in behavioral domains associated with ASD [191].

Collectively, these results demonstrate that altered 5-HT availability and SERT activity (or regulation) represent an important biological endpoint in ASD risk in a subset of individuals. These results suggest that the functional impact of genetic variation at genes contributing to SERT regulation may be ASD risk factors.

DAT in psychiatric disease

The neurotransmitter DA has an important role in the central nervous system by regulating a variety of functions, including motor activity, motivation, attention and reward [192-195]. The importance of a dopaminergic role in psychiatric disease has been speculated for quite some time based on several observations. For example, many antipsychotic drugs target the dopamine system either via action at dopamine receptors or alterations in dopamine metabolism [196] and, amphetamine (a DAT substrate) abuse can mimic many of the symptoms of psychosis.

Similar to *SLC6A4*, the DAT gene, *SLC6A3*, has been the target of numerous genetic studies with the majority of studies examining the 3' untranslated region of the gene containing a common VNTR polymorphism (Table 3.2). The VNTR polymorphism is associated with substance dependence, response to medication (methylphenidate), ADHD, bipolar disorder, schizophrenia and major depressive disorder (MDD) [197, 198]. Several coding variants in *SLC6A3* were identified pre-NGS and, of these, at least two have shown *in vitro* functional consequences as a result of the mutation. Mazei-Robison and colleagues identified A559V in two ADHD subjects, and another group identified the variant in a family with bipolar disorder [199, 200]. The A559V is of special interest as ADHD is a common comorbid phenotype observed in approximately one-third of ASD [8, 201-204]. ADHD, like ASD, is heritable although analyses has revealed a complex genetic architecture [205].

In comparison to SERT there is much less convincing evidence for dopamine transporter involvement in ASD. That said, variants in DA receptor subtype genes,

Table 3.2: Association of SLC6A3 with human disease including genetic variation and imaging studies

Gene Symbol	Aliases	Disease Association	Gene Locus	Disease Association from SNPs		Imaging Technique	Biomarker Association	Tissue Distribution
SLC6A3	DAT	Parkinson disease	5p15.3	A559V; rs40184	Bipolar disorder	PET	Parkinson disease	Brain
		Tourette syndrome		A559V; E602G; Intron 8	ADHD		Dementia with ley bodies	
		ADHD		Intron 8	Cocaine addiction		ADHD	
		Addiction		A1343G; rs6347	Schizophrenia		Schizophrenia	
		Major affective disorder		VNTR 9	Smoking		Sleep behavior disorder	
				rs40184	Depression		Bipolar disorder	
				rs40184	Migraine		Autism	
				Intron 8	OCD			
				VNTR 10	Alcoholism			
				VNTR 9	Epilepsy			
				VNTR 11	Parkinson disease			
				VNTR 9	Tourette syndrome			
				VNTR 9	Methylphenidate			
				T356M	Autism			

Table modified from original production of Pramod et al, Molecular Aspects of Medicine 34:2013

including *DRD1*, *DRD3* and *DRD4* have shown association with increased risk for ASD, and correlation to the repetitive or stereotyped behaviors domain with ASD [206]. In a single positron emission tomography (PET) study in adults with ASD, DAT binding was significantly elevated in the orbitofrontal cortex [207]. It is important to note that this study has not been replicated as of this writing.

Hypothesis and thesis aims

This thesis sought to address the hypothesis that monoamine transporter gene alleles are associated with autism and to elaborate the functional impact of autism associated alleles in context of monoamine transporter dysfunction. Perhaps most broadly, the objective of this thesis was to use genetic tools to discover and describe novel ASD-associated risk factors within a biological context. Building upon the work of many before me, the foundation for this project is that *many genes* involved in ASD will *converge to biological systems* and therefore those systems, including monoamine neurotransmitter networks, are likely to provide a biological framework for ASD risk.

Hypothesis: Given that monoamine transporters are important in maintaining monoamine tone in the brain, I hypothesize that monoamine transporter dysfunction is a convergent point in the biology of ASD, and thereby is source of genetic risk in a subset of ASD individuals.

To test my hypothesis, the aims of my thesis were required to be independent yet complementary. Toward this end, the following aims were undertaken:

AIM 1: A link between ASD and SERT dysfunction is well established in some individuals with ASD (as described earlier). SERT activity is subject to acute modulation through multiple signaling pathways including activation of PKG and p38 MAPK dependent pathways. Therefore the objective of this AIM was to investigate a PKG and p38 MAPK activating receptor, the adenosine A3 receptor (A3AR), for autism associated variation. A proof of principle that variants in SERT regulatory proteins can also contribute to ASD genetic risk comes from a combination of work conducted by Carneiro and colleagues in which they found that a common variant in *ITGB3* (rs5918), confers allelic dependent effects on SERT activity and regulation [93, 208, 209], and Weiss et al previously demonstrated that this variant is statistically associated with ASD. Since my hypothesis is supported from observations of elevated SERT reuptake activity in multiple rare autism associated SERT coding variants [81, 144, 187, 188], basic understanding of any functional effects conferred by A3AR-associated variants *in vitro* will be conducted.

Sub aim A: To analyze the A3AR locus (*ADORA3*) for ASD related variation. A3AR was evaluated for (a) evidence of *common* (≥5% MAF) allele association and (b) discovery of novel variants in ASD subjects and gene based association testing of rare variants.

Sub aim B: **To investigate the functional impact of putative A3AR ASD-associated variation on protein function and SERT activity**. Functional analysis of variants will be prioritized on the basis of evidence for genetic risk to ASD, likelihood for damaging effect on the protein in question, and knowledge of SERT regulatory pathways and ASD risk loci.

AIM 2: DAT is strongly implicated in ADHD, a common co-morbidity in ASD, and DA dysfunction is linked to several other psychiatric diseases. Little is known regarding DAT dysfunction in autism, therefore, the objective of this aim was to functionally, structurally, and behaviorally characterize a *de novo* variant identified in a single ASD proband.

To test my hypothesis that dysfunction of monoamine transporters is a biological risk factor in ASD, I set out to identify ASD-associated variants in a SERT-regulatory gene (*ADORA3*), then to test identified variants for functional effects on SERT (Chapter 4). Next, I set out to characterize the functional consequences of a *de novo* variant in DAT identified in an ASD proband (Chapter 5). Lastly, novel methodological techniques were conducted that validated the use of lymphoblastoid cell line (LCL) derived DNA, a common source of DNA for discovery of novel variation. Through these next chapters I set out to demonstrate that genetically derived monoamine transporter dysfunction contributes to ASD liability.

Chapter IV

Rare coding variants of the adenosine A3 receptor are increased in autism: on the trail of the serotonin transporter regulome

Introduction

The genetic and functional analysis of ASD-associated adenosine A3 receptor (A3AR) variants is the focus of this chapter and much of this work can be found in Campbell et al. [104]. The publication of this manuscript encompasses AIM1 of my thesis and is summarized in this chapter.

Modulation of synaptic 5-HT tone is a dynamic and tightly controlled process subject to influence through multiple signaling pathways and interacting proteins that act on SERT [140]. SERT, a monoamine transporter, is the main pharmacological target of numerous psychiatric medications, and dysregulation of SERT is implicated in a wide array of psychiatric disorders (discussed in Chapter 3). Additionally, the gene encoding the SERT protein, *SLC6A4*, is the target of numerous genetic studies across multiple psychiatric conditions.

It has been shown that SERT activity can be enhanced via PKG and p38 MAPK signaling pathways acting through (i) trafficking-dependent and (ii) trafficking-independent mechanisms [140, 210, 211]. A trigger for both of these uptake-enhancing pathways is activation of the A3 adenosine receptor (A3AR, gene symbol: *ADORA3*). A3ARs are G-protein-coupled receptors (GPCR), which exert their influence on SERT activity via a Gq-linked stimulation of guanyl cyclase (GC)-mediated cGMP synthesis

[212]. A3ARs have been shown to be expressed and physically interact with SERT at synaptic terminals in 5-HT synthesizing neurons *in vivo* [210, 213]. Moreover, A3AR stimulation of SERT activity is lost in A3AR knockout mice [214].

In light of the role of A3ARs in SERT regulation, a reasonable approach to address if SERT regulating proteins harbor functional ASD-associated variants, was to leverage *ADORA3* as a candidate locus and ask if *rare or common* functional variants at this locus are correlated with ASD risk. To test for common allele effects on ASD risk, SNPs indexing common haplotypes at *ADORA3* were genotyped and assessed using family-based association methods. The alternative rare variant hypothesis was tested by direct Sanger sequencing of the *ADORA3* transcript in a sample of ASD probands and an ethnically comparison group. Initial rare variant association analysis was followed-up in a replication analysis using WES data generated from a large cohort of independent ASD cases and matched controls. Moreover, two nonsynonymous variants identified as being increased in cases from these studies were evaluated for changes in basal and agonist-activated modulation of SERT activity by heterologous expression of wildtype and variant A3ARs *in vitro*.

Methods

Sample for allelic association analysis

The sample for analysis of common alleles consisted of 958 combined simplex and multiplex ASD families (1649 probands; 4150 samples) recruited across multiple centers or obtained from the NIMH Genetics Repository (http://nimhgenetics.org). All ASD

probands were assessed with the Autism Diagnostic Interview (ADI) or its revision (ADI-R) [2] and most with the Autism Diagnostic Observation Schedule (ADOS) [3]. Affection status was assigned using a classification scheme employed by the Autism Genome Project [77]. Other "unaffected" family members were designated as "unknown" for purposes of genetic analyses described in this chapter. Demographic information for this sample is shown in Table 4.1. Studies were approved by the Vanderbilt Institutional Review Board and with the informed consent of participating families.

Genotyping and analysis of common variants

Tag-SNPs were selected to capture common alleles across the *ADORA3* loci at an $r^2 \geq 0.8$ and minor allele frequency (MAF) ≥5%. TaqMan™ allelic discrimination assays for four SNPs that span 14.4 kb (Figure 4.1) were obtained from Applied Biosystems (ABI, Foster City, CA) as Assays-on-Demand or designed as Assays-by-Design.

PCR amplification was conducted in a 5 µl volume, and went as follows: Cycling conditions included an initial denaturation at 95°C for 7 min, followed by 50 cycles of 92°C for 15 s and 60°C for 1 min. Allelic discrimination was performed on the ABI 7900HT and genotypes called using the ABI Sequence Detection System software. Quality control procedures included genotype completeness (≥98%), inter- and intra-plate replicates, and removal of families containing Mendelian inconsistencies identified by PEDCHECK [215].

Single marker analysis was conducted using the family based association test (FBAT) [216] under the additive model, and significance determined using the empirical

variance (-e) option - as this provides a more conservative estimate of association - given the presence of multiplex families in the dataset. Power calculations were conducted using the Genetic Power Calculator [217] under the following conditions: Prevalence rate of 1 in a 100 and a sample size of 1000 trios, assuming minor allele frequencies (MAFs) of 5%, 10%, and 30%, respectively and a D'=0.8. Based on these calculations, we would have ≥80% power to detect odds ratios (ORs) of ≥1.59, ≥1.44 and ≥1.36 for risk allele frequencies of ≥5%, 10%, and 30%, respectively.

Determination of sample ancestry

To permit association analysis within the major subset of the family-based sample and for matching of cases and controls (see below), ancestry was determined using multiple platforms and two methods: STRUCTURE [218] and multidimensional scaling (MDS) in PLINK [219]. Many multiplex families were genotyped by the Autism Genome Project (AGP) using the Affymetrix 10k SNP platform and analyzed using STRUCTURE [99]. Other families had genotypes from Illumina 550k and/or 1M SNP arrays and analyzed using MDS [75]. Coincidently, numerous families were genotyped and analyzed using both applications, allowing for confirmation that both STRUCTURE and MDS yielded consistent and robust assignments that also agreed with self-report information. A small number of families from the overall association sample did not have genome-wide genotypes, in which case self report information determined classification, or had neither genome-wide genotype data nor self-report information, in which case they were classified as being of “unknown” ancestry.

Table 4.1: Demographics of the sample cohort of common allele analysis at *ADORA3*

	All	Caucasian	African	Asian	Hispanic	Other[b]
Families	958	763	29	26	39	101
Individuals	4150	3311	116	134	171	418
ASD diagnosis[a]	1649	1316	48	52	69	164
Female	316	248	8	8	21	31
Male	1333	1068	40	44	48	133
Number of families with:						
1 affected	317	252	9	2	13	41
2 or more affected	641	511	20	24	26	60

[a] Indicates number of individuals who met criteria for (i) "autism" on the ADI-R "ASD" on both the ADI-R and ADOS or "autism" on the ADOS alone

[b] Represents families with two ancestry calls or families with at least one founder of undetermined or unknown ancestry

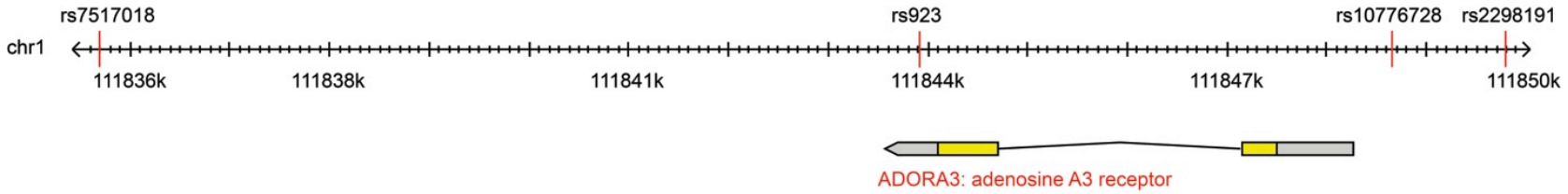


Figure 4.1: Chromosomal location of the four *ADORA3* genotyped SNPs. Genotyped SNPs are shown above the basepair position ruler and a vertical red harsh highlights positions. Tag SNPs were chosen to index haplotypes ≥5% across the locus using an r^2 cutoff of 0.8.

Discovery and analysis of rare variants

Initial screening for sequence variants at *ADORA3* utilized whole blood derived DNA samples from 185 unrelated cases. Non-clinical comparison samples were drawn from reference collections and consisted of lymphoblastoid cell line (LCL) derived DNA from 288 samples of European ancestry taken from the Coriell Human Genome Diversity panel and from the Human Random Control (HRC) collection. Twenty-four additional non-clinical comparison samples were obtained from the neurologically normal NINDS/Coriell African-American panel to allow for appropriate associations between case and control. While limited in power to detect rare risk variants, my *a priori* expectation was that functional *ADORA3* risk variants would confer large genetic effects, thereby allowing for detection in a small data set. Admittedly, this study would be significantly underpowered if hypothetical rare variants conferred ORs <2.

PCR amplifying primers (Table 4.2) were designed using Primer3 and subjected to a BLAST-Like Alignment Tool search to ensure no matches existed elsewhere in the genome. Amplification was optimized using a gradient of annealing temperatures, typically ranging from 56-64°C, followed by agarose gel electrophoresis to determine ideal annealing temperature. PCR was performed on the case (N=185) and control (N=305) screening samples containing 7.1 nm of amplifying primers, 10 µl of 2X Mastermix, and 12 ng of DNA in a final volume of 20 µl. Amplification parameters were 7 min at 95°C, 40 cycles of 15 s at 95°C, 12 s at the annealing temperature, and 60 s at 72°C, with a final extension at 72°C for 7 min.

Table 4.2: *ADORA3* PCR amplifying primers

Primer Set	Forward	Reverse	Amplicon Length	Annealing Temperature
1	CGCCATTGTTGTTACTGCTG	CAGCAAAGATCCTTGGTCAAA	585	60.2°C
2	GGCTAAGCAGGTGTGATGCT	TTCAGGGGTGTTTCAGGAAG	469	60.2°C
3	TGCTCTTTCCATCTTTTTGCT	GCCAGCAAGATCCGTCTGTA	425	58.3°C
4	AGACTGTCACTGCACATGGA	GCCCTCTTTCAACATCAAGG	589	60.2°C
5	GCAGAAGATTGGAGAATTAAGAGA	AATCAGAGGGATGGCAGACC	700	60.2°C
6	CCAACTCCATGATGAACCCTA	GGCTCCAAGTAGCAAGCAAG	749	62.3°C

Following PCR, each product was examined by gel-electrophoresis, and once verified for expected size and specificity, excess primers and nucleotides were removed using Millipore 96 well filter plates. Samples were quantitated, subjected to Sanger sequencing by the DNA Sequencing Core at Vanderbilt University, and data analyzed using Sequencher v4.9 (Gene Codes Corporation, Ann Arbor, MI). Specific variants were validated by independent PCR, and segregation analysis performed by sequencing available family members, affected or unaffected.

Variants were assessed using five diverse approaches including (i) documentation in dbSNP or the 1000 Genomes project, (ii) *in silico* algorithms such as PolyPhen2, SIFT and SNAP to predict amino acid substitution as "damaging" or benign [220-222] (iii) cross-species conservation of the amino acid and surrounding residues, (iv) available literature regarding structural features of the A3AR protein, and (v) a global burden test of rare (<1%) coding variants. CAST, the cohort allelic sums test, is a gene burden test in which the number of individuals with one or more variants in a gene is compared between affected and unaffected individuals [223, 224]. Rare variant counts in matched cases and controls in a 2 x 2 contingency table were tallied and a Fisher's Exact test performed.

An independent replication sample corresponded to ASD cases and controls from the NIMH Repository. Cases and controls were pair-matched for ancestry, and samples used for WES analysis were described in Neale et al. [125], with the exception that a number of case-control pairs were excluded if control subjects had any history of obsessive compulsive behavior, anxiety disorders, and substance abuse - three disorders

associated to SERT dysregulation. WES variants at *ADORA3* in 339 case-control pairs were examined for "functional" (i.e. missense, nonsense, consensus splice site and read-through) variants. Details regarding quality control procedures and read depth for case and control samples are detailed in the Supplemental Material of Neale et al.

Functional Studies

Constructs: A full-length human SERT cDNA in the mammalian expression vector pcDNA3.1, and a full-length human A3AR (myc-A3AR/pCMV) cDNA were used in this study and are described in [213, 214]. *In vitro* mutagenesis of the A3AR cDNA clones to introduce Val90 and Ile171 variants were performed using the QuikChange mutagenesis kit (Stratagene, La Jolla, CA), and variants were confirmed by Sanger sequencing at the DNA Sequencing Core at Vanderbilt University.

Cell culture and transfection*:* Chinese Hamster Ovary (CHO) cells were maintained at 37°C in Dulbecco's modified Eagle's medium containing 10% fetal bovine serum, 1% L-glutamine, 100 IU/mL penicillin, and 100 µg/mL streptomycin. Transient transfections were performed using the *Trans*IT–CHO Transfection reagent (Mirus, Madison, WI). In brief, 50 ng/well SERT and 20 ng/well of A3AR cDNAs were pre-incubated with Mirus reagent for 30 min at room temperature prior to adding to plated cells, seeded at 20,000 cells/well. Transfected cells were cultured for an additional 24-48 hr after transfection and prior to functional analyses.

cGMP activity assays: The concentration of the second messenger cyclic guanosine monophosphate (cGMP) was measured using the CatchPoint cGMP fluorescent assay kit

(Molecular Devices, Sunnyvale, CA). In brief, CHO cells were co-transfected with wildtype A3AR or A3AR coding variants and SERT cDNAs were plated at 50,000 cells/well. Cells were initially washed in a Krebs-Ringer bicarbonate pre-stimulation buffer containing 0.8 mM 3-isobutyl-1-methylxanthine (IBMX), an inhibitor of cGMP-phosphodiesterases. Following a 10 min incubation period, cGMP accumulation was measured in response to a 1 μM IB-MECA or vehicle treatment at various time points (0 min, 10 min, 20 min, 30 min, 40 min, 60 min) according to the manufacture's protocol. The 1 μM IB-MECA concentration was selected on the basis of previous studies that found optimal stimulation of A3ARs under those conditions [210, 213, 214]. All values were normalized to basal wildtype cGMP levels (i.e. WT at time = 0 min). Statistical analyses comparing the treatment effects of the adenosine A3 selective agonist and adenosine analog IB-MECA on variant and wildtype A3ARs cells were performed with Prism (GraphPad, La Jolla, CA) using two-way analysis of variance (ANOVA). A t-test to compare overall differences over the full time course of IB-MECA treatment was calculated with Prism using total cGMP production. Statistical means and standard errors of the mean were calculated and P-values < 0.05 were considered significant.

5-HT transport assays: Assays measuring transport of [^{3}H]5-HT were conducted as described previously [210, 213, 214]. Briefly, media from CHO cells were removed and cells washed with Krebs-Ringer-Hepes (KRH) buffer containing 130 nM NaCl, 1.3 nM KCl, 2.2 mM $CaCl_2$, 1.2 mM $MgSO_4$, 1.2 mM KH_2P0_4, 1.8 g/L glucose, 10 mM HEPES, pH 7.4, 100 mM pargyline, and 100 mM ascorbic acid. Cells were incubated in triplicate at 37°C with and without IB-MECA treatment. Following treatment with IB-MECA, cells were

incubated for 10 min with [^{3}H]5-HT (20 nM) at 37°C. Cells were washed with ice-cold KRH buffer, solubilized with 0.5 mL Microscint-20, and [^{3}H]5-HT accumulation was quantified using a TopCount plate scintillation counter (PerkinElmer Life and Analytical Sciences, Waltham, MA). Specific 5-HT uptake was determined by subtracting the amount of [^{3}H]5-HT accumulated in the presence of 10 μM paroxetine, a selective SERT inhibitor. A minimum of three independent replicates was performed for each experiment. Analyses comparing basal and agonist-induced uptake in wildtype and variant A3AR/SERT co-transfections at individual time points was calculated and plotted using the software package Prism.

Molecular modeling of A3AR coding variants

The A3AR residue positions of Leu90 and Val171 were determined for the human adenosine A2a receptor (A2aAR) by amino acid sequence alignment and correspond to residues Val84 and Val171, respectively. The relative positions and structural conformation of the corresponding A2aAR residues were determined using structural data obtained using an adenosine-bound thermostabilized A2aAR (PDB ID: Y2DO) as a model. Images were prepared using PyMOL (DeLano Scientific Ltd).

Results

Family based association analysis

To test if common alleles at *ADORA3* contribute significant genetic risk in a sample of 958 ASD families, four SNPs representing common haplotypes spanning the locus

were genotyped. FBAT analyses of genotype data showed no evidence to support common variant association at *ADORA3* at the SNPs tested (Table 4.3). Analysis of haplotypes (>5%) across the locus that were captured by these SNPs also showed no significant association, consistent with single marker analysis (data not shown).

Sequence discovery of functional variation at ADORA3

To identify potential functional risk variants in *ADORA3*, all exons and the promoter region in 185 unrelated ASD cases and 305 non-clinical comparison samples by direct Sanger sequencing were screened. Numerous synonymous and nonsynonymous variants were identified and are documented in Table 4.4. Three novel variants, two nonsynonymous and one synonymous, were detected in cases (Leu90Val, Val171Ile, Cys194). A novel Ala195Thr substitution was identified in a single control sample.

Parental transmission was analyzed and no *de novo* variants were detected. Bioinformatic algorithms applied to rare (<1%) ns-variants predicted Val171Ile to be "damaging", however, Leu90Val along with Ala195Thr were predicted to be benign substitutions. Following discovery of Leu90Val and Ile171Val variants, but during follow-up functional and genetic experiments, both ASD identified variants emerged from 1k Genomes sequence data and were thus deposited into dbSNP, as reflected in Table 4.4.

Table 4.3: Association analysis of common variants at the *ADORA3* loci

SNP	Major/ minor allele	MAF[a]	Inf fams[b]	Major T_{OBS}[c]	Major T_{EXP}[d]	Minor T_{OBS}	Minor T_{EXP}	Z	p-value[e]
rs7517018	T/C	0.145	266	671	668.2	283	285.8	0.247	0.809
rs923	A/G	0.180	287	723	710.5	303	315.5	1.025	0.338
rs10776728	A/T	0.323	403	883	861.2	535	556.8	1.453	0.162
rs2298191	T/C	0.336	395	801	815.7	589	574.3	0.983	0.343

[a]MAF indicates minor allele frequency
[b]Inf Fams indicates number of informative families
[c]T_{OBS} indicates transmissions observed; equivalent to the "S" statistic in FBAT
[d]T_{EXP} indicates transmissions expected; equivalent to the "E(S)" statistic in FBAT
[e]p-value corresponds to empirical variance "-e" option in FBAT

Leu90Val and Val171Ile were detected only in ASD cases and not controls (3/370 case chromosomes for each variant vs. 0/562 control chromosomes). A case-control comparison of these *individual* rare alleles does not reach significance given the small number of observations (Fisher's Exact p=0.064 for Leu90Val and Val171Ile each). Due to the limitation in power to compare rates of single *rare* alleles, a gene-wide collapsing method suggests a better approach is to model all "functional" (i.e. nonsynonymous, consensus splice site and read-through) rare variants simultaneously for case-control comparisons. CAST testing on total rare "functional" variants obtained from sequence discovery revealed a significant burden effect (6/185 ASD individuals vs. 1/310 controls; p=0.013; OR=10.19, CI = 1.20-81.56) (Table 4.5A).

Consistent with published data on accepted risk CNVs (e.g. 16p11.2, 1q21.1, 22q11.2, 22q13.3, etc.) [111, 118, 121] novel variants Leu90Val and Val171Ile did not always segregate to only - or all - affected individuals in a family, consistent with incomplete penetrance (Figure 4.2).

In-silico analysis reveals functional effects of ASD associated variants

The availability of crystal structure for a human *ligand-bound* adenosine A2a receptor (A2aAR) [225] provides an important source of structure-function information that can inform predictions of a functional impact of the Leu90Val and Ile171Leu variants. Figure 4.3 depicts the structure of the A2aAR, modeling the residues equivalent to A3AR-Leu90 and -Val171, in relation to bound adenosine. This model demonstrates that the A3AR-Leu90 and -Val171 residues flank the binding site,

Table 4.4 : Sanger sequencing discovery of variation at *ADORA3*

				Unrelated probands		Caucasian controls		African American controls	
SNP ID	HGVS[a]	mRNA position[b] (NM_000677.3)	mRNA Location[c]	Chromosomes[d]	MAF[e]	Chromosomes[d]	MAF[e]	Chromosomes[d]	MAF[e]
rs10776728	c.108-13736A>T		-376	105/370	0.284	178/562	0.317	17/48	0.354
rs140137165	c.108-13682G>C		-322	0/370	0.000	0/562	0.000	2/48	0.042
	c.108-13643G>T		-283	0/370	0.000	1/562	0.002	0/48	0.000
rs10776727	c.108-13642G>T		-282	164/370	0.443	251/562	0.447	13/48	0.271
rs114241928	c.108-13599C>G		-239	0/370	0.000	0/562	0.000	2/48	0.042
	c.108-13532C>T		-172	1/370	0.003	0/562	0.000	0/48	0.000
	c.108-13371C>T		-11	0/370	0.000	2/562	0.004	0/48	0.000
	c.-724G>A	44	5'UTR	1/370	0.003	1/562	0.002	0/48	0.000
rs1544223	c.-581G>A	187	5'UTR	71/370	0.192	119/562	0.212	11/48	0.229
rs1544224	c.-564T>C	204	5'UTR	73/370	0.197	130/562	0.231	11/48	0.229
	c.-563G>A	205	5'UTR	0/370	0.000	1/562	0.002	0/48	0.000
	c.-479A>G	289	5'UTR	0/370	0.000	1/562	0.002	0/48	0.000
rs41282522	c.-221G>C	547	5'UTR	52/370	0.141	81/562	0.144	6/48	0.125
rs41282520	c.-105A>C	663	5'UTR	1/370	0.003	4/562	0.007	0/48	0.000
	c.-85G>A	683	5'UTR	0/370	0.000	1/562	0.002	0/48	0.000
rs35789323	c.265C>T	1,032	Leu89Leu	1/370	0.003	0/562	0.000	0/48	0.000
rs77883500	c.268C>G	1,035	Leu90Val	3/370	0.008	0/562	0.000	0/48	0.000
rs76934313	c.345C>T	1,112	Thr115Thr	2/370	0.006	3/562	0.005	0/48	0.000
rs2789537	c.390C>T	1,157	Ala130Ala	6/370	0.016	5/562	0.009	1/48	0.021
rs139935750	c.511G>A	1,278	Val171Ile	3/370	0.008	0/562	0.000	0/48	0.000
	c.582C>T	1,349	Cys194Cys	1/370	0.003	1/562	0.002	0/48	0.000
rs143962803	c.583G>A	1,350	Ala195Thr	0/370	0.000	1/562	0.002	0/48	0.000
rs35511654	c.742A>C	1,509	Ile248Leu	54/370	0.146	84/562	0.158	6/48	0.125
rs2800889	c.797T>A	1,564	Met266Lys	6/370	0.016	5/562	0.009	1/48	0.021
rs2229155	c.897T>C	1,664	Ala299Ala	52/370	0.154	100/562	0.178	9/48	0.188
	c.*97C>A	1,821	3'UTR	0/370	0.000	1/562	0.002	0/48	0.000
rs923	c.*189C>T	1,913	3'UTR	51/370	0.149	100/562	0.178	9/48	0.188
rs1415793	c.*336C>T	2,060	3'UTR	51/370	0.149	100/562	0.178	9/48	0.188
rs75048140	c.*350A>G	2,074	3'UTR	2/370	0.005	0/562	0.000	0/48	0.000
rs1415792	c.*377A>G	2,101	3'UTR	50/370	0.146	100/562	0.178	9/48	0.188
	c.*409C>T	2,133	3'UTR	0/370	0.000	1/562	0.002	0/48	0.000
rs3393	c.*423G>A	2,147	3'UTR	123/370	0.435	258/562	0.459	9/48	0.188
rs3394	c.*441A>G	2,165	3'UTR	51/370	0.149	99/562	0.176	9/48	0.188

[a] Human Genome Variation Society (HGVS) recommended nomenclature

[b] Indicates base-pair position within mRNA structure. RefSeq Accession number is given in parenthesis.

[c] Indicates location within mRNA structure (Exon). Negative (-) numbers indicate base pair position 5' to mRNA. Amino acid substitutions are shown relative to protein position.

[d] Indicates the number of minor allele counts over the total number of chromosomes sequenced.

[e] MAF indicates minor allele frequency

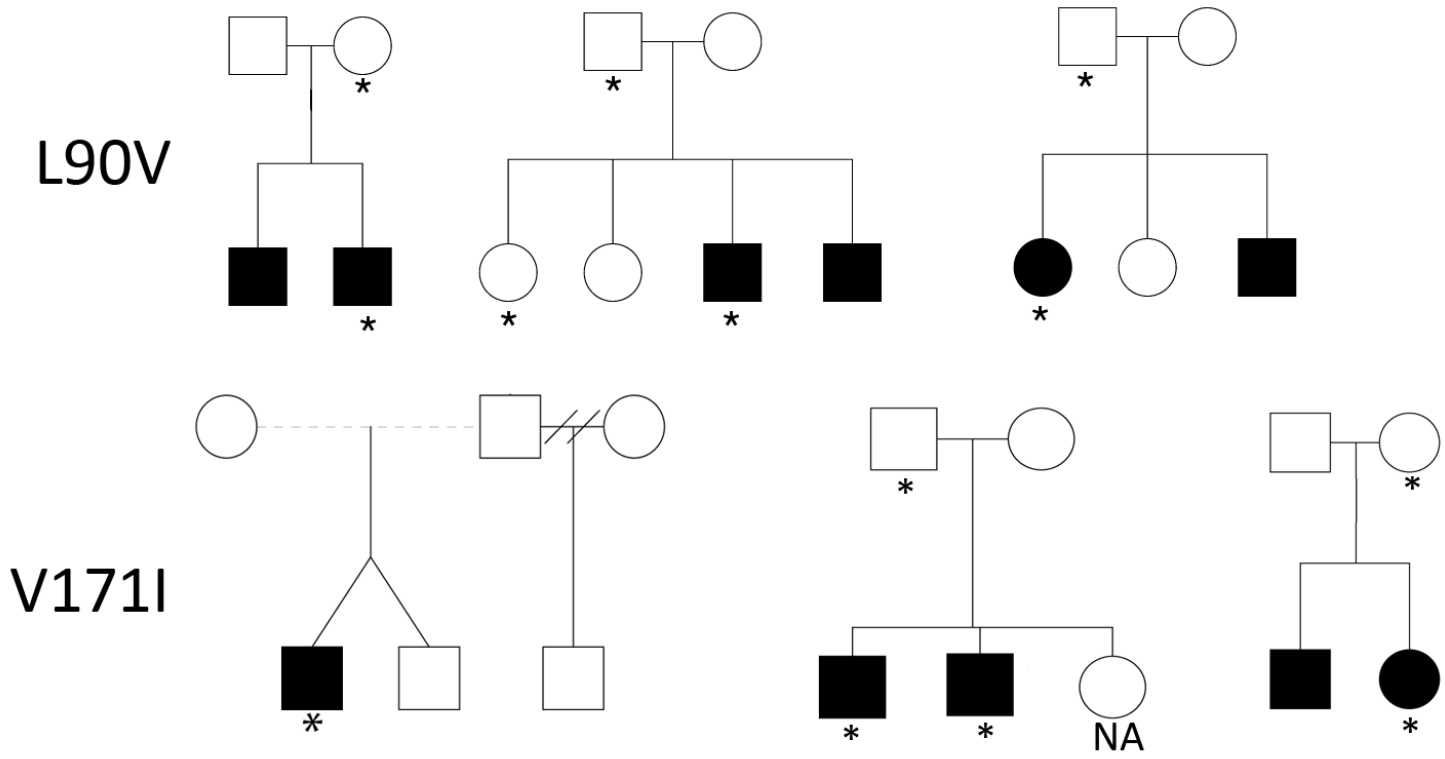


Figure 4.2. Pedigree structure of ASD families harboring the Leu90Val or Val171Ile variant. An asterisk indicates subjects harboring either variant. Individuals for whom DNA was not available are labeled NA. Individuals with ASD re indicated by filled circles or squares, while unfilled elements reflect individuals without an ASD diagnosis.

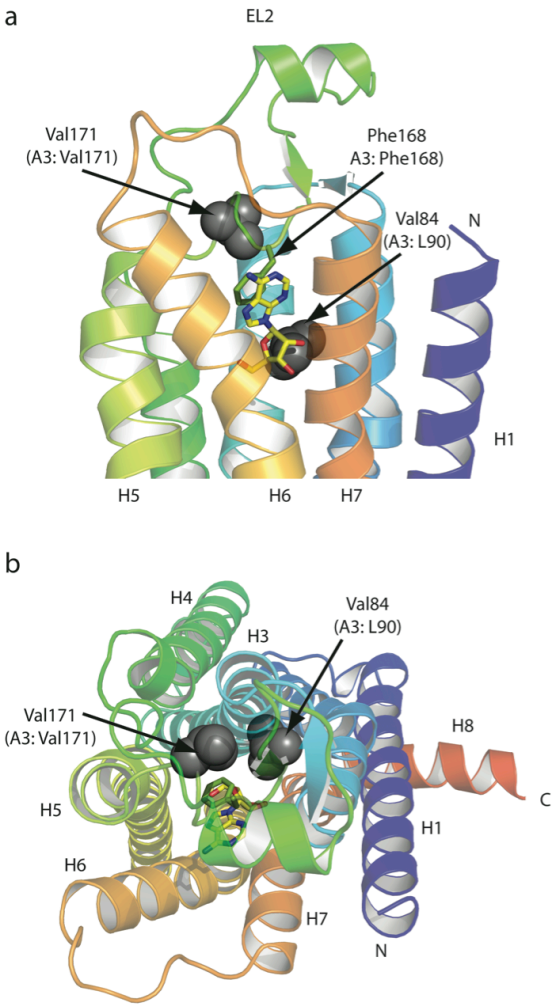


Figure 4.3. Structure of the A2a adenosine receptor identifying corresponding positions of Leu90 and Val171 residues in the A3 receptor. The structure of the thermostabilized A2a receptor, A2aAR-GL31 (PDB code 2YDO), is depicted as a cartoon in rainbow coloration (N-terminus in blue, C-terminus in red), viewed either **(A)** parallel to the membrane plane or **(B)** from the extracellular surface perpendicular to the membrane plane. The endogenous ligand adenosine is shown as a stick model (C, yellow, N, blue, O, red) as is the side chain for Phe168 (C, green). The transmembrane helices are labeled H1-H7, the extracellular loop 2 (EL2) is indicated and H8 is the C-terminal amphipathic helix lying parallel to the membrane plane. A3AR residues Leu90 and Val171 correspond to Val84 and Val171 in A2aAR, and they are both shown as space-filling models (grey)

predicting a functional effect. Figure 4.4 demonstrates consistent cross-species conservation in mammals for Leu90 and Val171, while Ala195 is less conserved. Functional studies therefore focused solely on the two variants identified in cases.

Functional analysis of Leu90Val- and Val171Ile-A3AR effects on cGMP Production and SERT activity

To test for potential functional effects produced by the Leu90Val and Val171Ile substitutions, we engineered human A3AR cDNA expression constructs to harbor either Leu90Val or Val171Ile variants in a heterologous transfection system. Western blot analysis of cell lysates confirmed equivalent expression of myc-tagged wildtype A3AR compared to both variant A3AR constructs. cGMP levels for A3AR/SERT variants, before and after A3AR activation with IB-MECA, demonstrate cells transfected with the Leu90Val variant display elevated basal cGMP levels compared to wildtype transfected cells (L90V: 207.85% ± 45.71 vs. WT: 100.0% ± 0.02; P=0.015, n=3) (Figure 4.5A) Stimulation of the Leu90Val A3AR by IB-MECA (1 μM) revealed that the elevated cGMP production seen in the basal state persisted over the full time course and production paralleled wildtype A3AR with both returning to their respective basal states after 40 minutes (Figure 4.5A).

In contrast with Leu90Val, cells co-transfected with Val171Ile-A3AR/SERT displayed similar basal cGMP levels to those observed in wildtype A3AR/SERT transfected cells (V171I: 116.1% ±16.3 vs. WT: 100.0% ± 9.2, P=n.s., n=3). IB-MECA treatment on Val171Ile/SERT cells failed to induce significant cGMP production compared to that seen

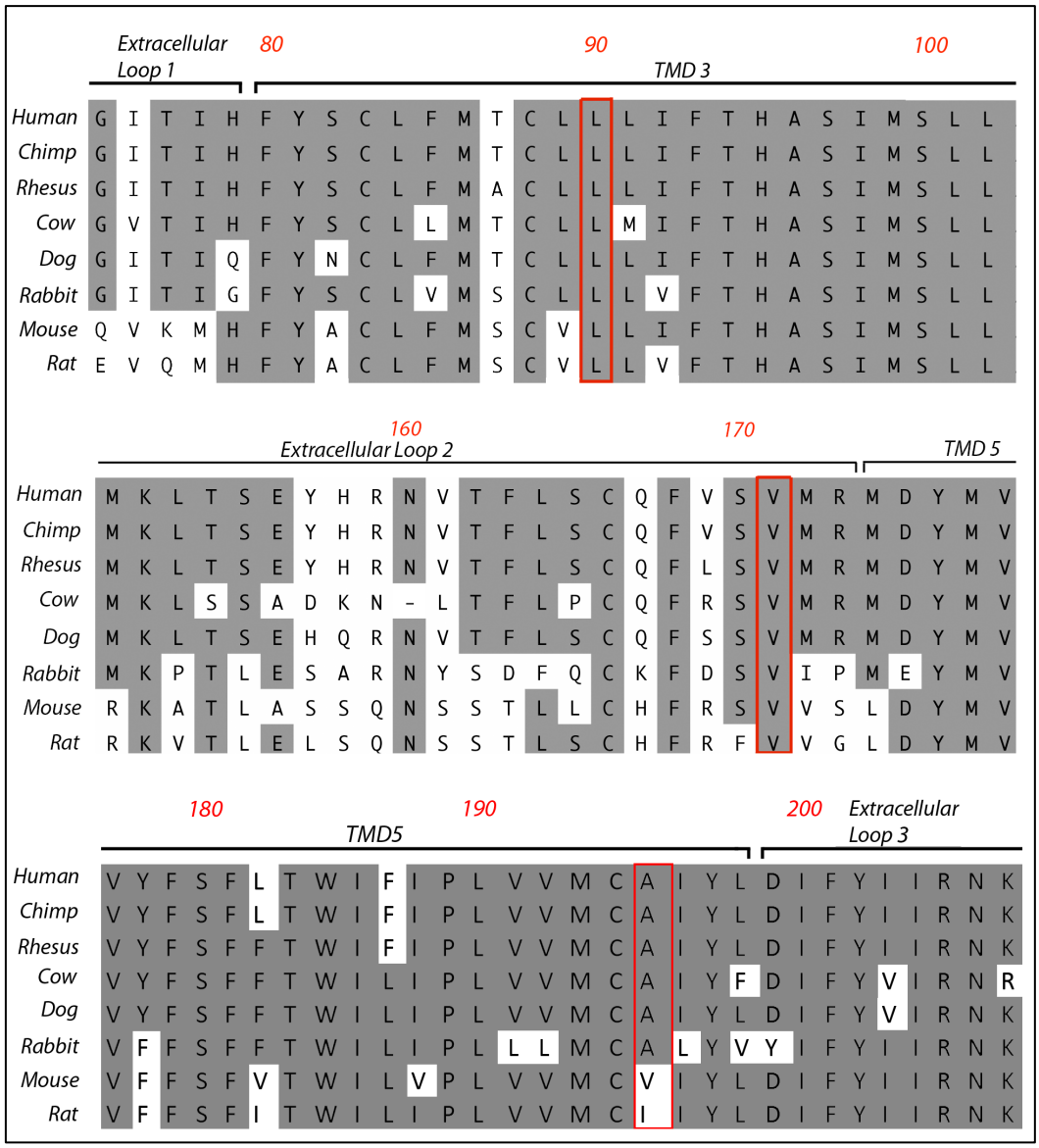


Figure 4.4. Cross-species conservation at *ADORA3* variant sites detected by Sanger sequencing. Amino acid sequences of the A3AR protein encoded by *ADORA3* are aligned for the three variants and their flanking residues: Leu90Val, Val171Ile (cases), and Ala195Thr (controls).

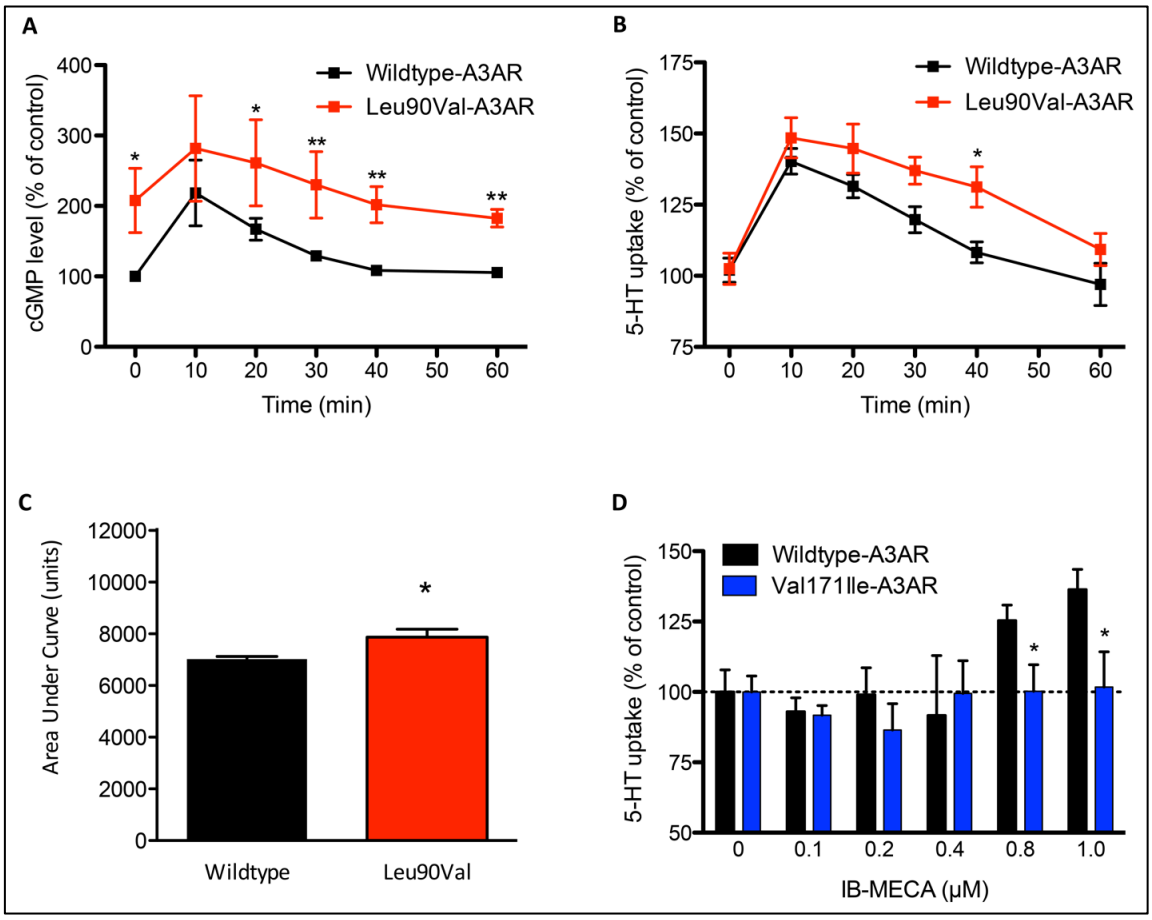

A
cGMP level (% of control)
400
300
200
100
0
Wildtype-A3AR
Leu90Val-A3AR
*
*
**
**
**
0
10
20
30
40
50
60
Time (min)
B
5-HT uptake (% of control)
175
150
125
100
75
Wildtype-A3AR
Leu90Val-A3AR
*
0
10
20
30
40
50
60
Time (min)
C
Area Under Curve (units)
12000
10000
8000
6000
4000
2000
0
*
Wildtype
Leu90Val
D
5-HT uptake (% of control)
150
125
100
75
50
Wildtype-A3AR
Val171Ile-A3AR
*
*
0
0.1
0.2
0.4
0.8
1.0
IB-MECA (μM)

Figure 4.5: Functional consequences of the A3AR specific agonist IB-MECA in WT-A3AR and Leu90Val-A3AR or Val171Ile-A3AR expressing cells. CHO cells co-transfected with either WT-A3AR or Leu90Val-A3AR or Val171Ile-A3AR and SERT and were stimulated using IB-MECA (1 μM). **(A)** Levels of cGMP production from wildtype-A3AR and Leu90Val-A3AR cells were measured prior to and after IB-MECA stimulation. Leu90Val-A3AR cells display elevated basal cGMP production revealing enhanced overall production of cGMP. **(B)** 60 min time-course of IB-MECA induced 5-HT uptake as measured as percent of control (wildtype; time = 0 min). CHO cells were co-transfected with either wildtype-A3AR and SERT or Leu90Val-A3AR and SERT, and 5-HT uptake was stimulated using IB-MECA (1 μM). Cells were incubated with [^{3}H]5-HT (20 nM) for the indicated period. Over the time course (0-60 min) in Leu90Val-A3AR/SERT expressing cells, 5-HT uptake is elevated compared with the wildtype A3AR counterpart. **(C)** At the time when the control cells had already returned to baseline levels, greater 5-HT uptake was seen in Leu90Val A3AR cells compared to wildtype A3AR/SERT. **(D)** [^{3}H]5-HT accumulation into A3AR/SERT cells was and indicate a failure of IB-MECA treatment of Val171Ile-A3AR to induce SERT dependent 5-HT re-uptake across all tested concentrations. Compared to the peak concentrations (0.8 μM and 1.0 μM). $P<0.05$ findings are indicated by an asterisk (*).

with wildtype A3AR/SERT. This inability of the Val171Ile receptor to be stimulated by IB-MECA resulted in a non-significant reduction in cGMP synthesis compared with wildtype (V171I: 90.14% ± 12.22 vs. WT: 100.0% ± 16.18; 1-tailed t-test P=n.s.)

Maximum 5-HT uptake (i.e. at 10 min) in Leu90Val A3AR/SERT cells was equivalent to that of wildtype A3AR/SERT (L90V: 148.5% ± 7.1 vs. WT: 140.3% ± 4.5; n=4), however, cells expressing Leu90Val A3AR/SERT displayed a persistent increase in IB-MECA-induced 5-HT uptake (Figure 4.5B). At the time which wildtype A3AR-stimulated SERT activity had already returned to baseline levels, Leu90Val A3AR/SERT cells were elevated in comparison (L90V: 131.3% ± 7.1 vs. WT: 108.3% ± 3.7; two way ANOVA p=0.001, n=4). Overall, greater 5-HT uptake was seen in Leu90Val A3AR/SERT cells compared to wildtype A3AR/SERT cells (Figure 4.5C; L90V: 112.2% ± 8.80 vs. WT: 100.0% ± 3.22; 1-tailed t-test p=0.020, n=4)

In striking contrast to the enhanced SERT activity observed with Val90 co-transfected cells, IB-MECA treatment of Ile171 failed to stimulate SERT activity (Figure 4.5D) across all doses tested.

Replication test of an increased rate of rare functional variants in ASD cases vs. controls

Examining whole exome sequence data from 339 cases and pair-matched controls of European ancestry, generated by the NIH ARRA Autism Sequencing Consortium, revealed numerous functional variants at *ADORA3*. Nine case subjects were found to harbor nonsynonymous variants compared to three in controls. Although trending, this

increase in variants seen in cases was not statistically significant (Fisher's Exact 1-tailed p=0.071; OR = 3.06, CI: 0.82-10.99; Table 4.5B).

In more detail, Leu90Val was identified in one additional ASD case and in zero controls, and Val171Ile was found in two ASD cases and one control. Of the remainder, four additional missense variants (Ile22Thr, Phe48Ser, Ala69Ser, and Leu294Thr) were found in one case each and no controls, and a single read-through mutation (*319Gln) predicted to add an additional 38 residues (>10% of the native protein) was found in two cases and zero controls. Two more missense variants were detected in controls (Phe180Leu and Ala273Thr). All variant calls passed quality control thresholds and have a high positive predictive value for being valid calls, although they have not been experimentally validated.

Combining both discovery and replication findings strengthens the evidence for an increase of "functional" variants in ASD. In total, 15 out of 524 cases compared to 4 out of 644 controls carried rare *ADORA3* nonsynonymous alleles (Fisher's Exact 1-tailed, Bonferroni-corrected p=0.0025; OR = 4.72, CI: 1.56-14.30). Likewise, the enrichment of the Val90 allele in ASD cases vs. controls becomes statistically significant when taking into account replication data, (Fisher's Exact 1-tailed p=0.040). Although present in greater numbers in ASD cases, enrichment for Leu171Ile by itself did not reach significance (Fisher's Exact 1-tailed p=0.068).

Examination of available phenotype data showed no consistent pattern of elevated or diminished ADI-R domain scores in carriers vs. non-carriers. A similar lack of

Table 4.5A The cohort allelic sums test on rare *ADORA3* variants in the discovery sample

rs#	mRNA	Cases	Controls
rs77883500	Leu90Val	3	0
rs139935750	Val171Ile	3	0
rs143962803	Ala195Thr	0	1
	Carriers	6	1
	Non-carriers	179	304

P=0.013 (OR=10.19, CI: 1.22-85.37)

Table 4.5B The cohort allelic sums test on rare *ADORA3* variants in the replication sample

rs#	mRNA	Cases	Controls
rs112045912	Ile22Thr	1	0
	Phe48Ser	1	0
	Ala69Ser	1	0
rs77883500	Leu90Val	1	0
rs139935750	Val171Ile	2	1
	Phe180Leu	0	1
	Ala273Thr	0	1
	Leu294Phe	1	0
rs112042574	*319Gln	2	0
	Carriers	9	3
	Non-carriers	330	336

P=0.071 (OR=3.06, CI: 0.82-11.39)

correlation existed upon examination of subsets of the ADI-derived principal components analysis-derived scores (Appendix A) [226]. Of the 15 ASD subjects with rare *ADORA3* variants CNV data were available on six cases. A single *de novo* duplication of approximately 211 kb (hg18: ChrX:153263157-153474401) was found in the male proband (NIMH ID: 217-14216-3470) harboring a Ala69Ser substitution in *ADORA3*, however, most CNVs also corresponded to regions detected in controls.

Discussion

Given that (i) SERT is an essential regulator of 5-HT signaling, (ii) SERT and 5-HT have a longstanding connection to ASD, and (iii) A3AR can regulate SERT via PKG and p38 MAPK signaling pathways, I postulated that rare *ADORA3* variants would be enriched in an ASD sample, and these A3AR rare ASD-associated variants would confer functional consequences on SERT activity. These experiments were largely premised on the findings of elevated SERT-dependent 5-HT uptake activity seen with multiple rare, ASD-associated SERT coding variants [81, 144, 187, 188].

Consistent with results from large-scale GWA scans for common allele risk loci I found no evidence to support a main effect on ASD attributable to common variants at *ADORA3*. Since 1000 families provides limited power to detect alleles of small effect size (e.g. OR < 1.3), the possibility that common *ADORA3* variants confer small effect sizes (or interact with alleles at other genes) to confer risk is very real.

Mechanistically speaking, the most expected scenario was for *ADORA3* to harbor coding variants, likely ≥5% minor allele frequency, that would impact A3 function

directly and SERT function indirectly. Numerous studies have documented that rare variation affecting a large number of genes is collectively a substantial source of genetic liability in ASD. As noted in the Methods, there is low power to detect rare variants of only modest risk in the sequencing sample. Nevertheless, a nominal increase in rare "functional" variants was discovered in the discovery case vs. control sample (p=0.013). These results are consistent with *ADORA3* contributing genetic liability in the form of rare variation in a small number of ASD individuals. Subsequent availability and analysis of exome sequence data from cases and clinically screened controls showed a greater number of "functional" variants at *ADORA3* in cases compared with controls. While the replication stage merely suggested an increase in rare "functional" variants (p=0.07), combining Sanger discovery and WES data strengthened evidence for association, even after Bonferroni correction for multiple comparisons (Fisher's exact 1-tailed p=0.0025; OR = 4.72, CI: 1.56-14.30).

Encouraged by the initial discovery of Leu90Val and Val171Ile, functional studies focused on these two nonsynonymous variants. Analysis of Leu90Val and Val171Ile entailed examining the effect of receptor stimulation to induce cGMP synthesis or on downstream SERT-dependent 5-HT uptake. IB-MECA stimulation of the Val90-encoded A3AR showed enhanced cGMP synthesis under basal conditions, and enhanced cGMP levels extended over the full time course ending with a return to the *elevated* baseline. Furthermore, IB-MECA treatment prompted an overall increase in SERT-dependent 5-HT uptake activity demonstrating a delayed return to baseline in cells transfected with Leu90Val A3AR/SERT. Increases in both cGMP and 5-HT uptake activity parallel the

recent report by Zhu and colleagues of enhanced Leu90Val-A3AR/SERT *complex* formation [213]. These data imply enhanced basal A3AR receptor-G protein coupling, and/or reduced receptor desensitization, and/or possible differences in the binding kinetics of IB-MECA to Val90-A3AR as a consequence of its location proximal to the A3AR ligand-binding site. Taken together, the Leu90Val-A3AR increases on SERT activity are consistent with prior data on increased 5-HT uptake caused by rare SERT ASD-associated mutations [81, 144, 187, 188].

In contrast to the functional consequences of Leu90Val, the Val171Ile variant rendered A3AR *insensitive* to induce cGMP synthesis and a downstream increase in SERT-dependent 5-HT uptake, as a response to the IB-MECA agonist. The molecular mechanism underlying this effect is not yet clear, however, similar to Leu90Val, the proximity of the Val171 residue to the ligand binding pocket may hinder the adenosine analog IB-MECA to bind A3AR. This effect may result in a more rapid dissociation and/or less efficient (or absence of) receptor-G protein coupling.

Additional experiments are required to fully elucidate the molecular mechanisms of these two coding variants on receptor function. Although the functional impact of the two ASD associated variants on SERT are in opposite directions, it is possible that both *elevated and diminished* capacity for regulation of SERT will influence ASD risk. Therefore it is conceivable that A3AR-related alterations in SERT activty may disrupt 5-HT's ability to coordinate brain development and/or adult 5-HT signaling [227-230].

I recognize that the statistical association of rare *ADORA3* variants (especially Leu90Val) requires additional validation in a larger ASD cohort; however, this confound

is characteristic of rare variant analysis in complex disease. The predominant challenge moving forward is in detecting susceptibility variants from those of normal human variation. This challenge is complicated by the recent appreciation that all ASD implicated loci will be littered with rare variation prompting many investigators to emphasize the *network* as a better model to elucidate the underlying mechanisms (and genes) implicated in autism spectrum disorders.

A single male case harboring an Ala69Ser variant (of unknown functional effect) at *ADORA3* possessed a *de novo* duplication of an X-linked interval including *RPL10*, a gene that has been previously associated with ASD and ID [231]. While the duplication effects of *RPL10* are unknown, this CNV is likely to confer risk.

While compelling, the precise magnitude of risk conferred by the *ADORA3* variants is uncertain and there are some caveats regarding the studies presented. First, while the discovery sequence sample was ethnically matched, a subset of individuals were not matched based on genome-wide genotype data but instead on self-report. It is possible that subtle population stratification effects could lead to inflation of the observed increase in numbers of rare nonsynonymous variants in cases vs. controls, and evidence for gene based association in the discovery ($p=0.0143$) and combined ($p=0.0025$) sample sets should be interpreted with caution. Secondly, while multiple comparisons were conducted in functional experiments, the parallel increases in cGMP production and 5-HT uptake support my conclusion regarding the ability of Leu90Val A3AR to augment SERT-dependent 5-HT uptake over time. Finally, all functional experiments were conducted *in vitro* in CHO cells, and may not accurately reflect the function *in vivo*.

While this is certainly possible, existing data on multiple variants in *SLC6A4* and in proteins that influence SERT present a consistent picture from *in vitro* based experiments to effects on SERT function in mouse models harboring these variants [186, 209, 232].

Summary

The results presented in this chapter support the hypothesis that SERT regulatory proteins harbor an excess of rare, functional variants in an ASD sample that impact SERT activity and regulation. Taken together with previous data on the functional implications of an ASD-associated SNP in *ITGB3* (also a SERT regulator), and several in *SLC6A4* itself, this data is consistent with my hypothesis that monoamine dysregulation represents one subset of biological risk in ASD individuals. These results encourage further investigation of the SERT-regulatory network, and more broadly dysfunction of monoamine transporters, as potential genetic and biologically relevant ASD risk factors. The studies presented herein add to the growing body of data implicating specific gene/protein networks in contributing to ASD liability. It is hoped that in the future, additional studies within the SERT regulatory network may provide new leads to ASD therapeutics.

CHAPTER V

De novo mutation in the dopamine transporter gene associated dopamine dysfunction with autism spectrum disorder

Introduction

The functional characterization of a de novo DAT variant identified in a simplex ASD family is the focus of this chapter. The majority of this work can be found in Hamilton and Campbell et al. 2013 [233]. The publication of this manuscript encompasses AIM2 of my thesis and is summarized here.

As discussed in Chapter 2, genetic factors have been implicated as important components in the etiology of autism spectrum disorder (ASD). It is now accepted that rare genetic variation affecting single nucleotides of protein coding DNA as well as rare genomic copy number variants are significant ASD risk factors [34, 112, 114, 126]. In particular, increasing evidence suggests that *de novo* genetic variation, both SNV and CNV, comprise a considerable class of risk factors in ASD and other neuropsychiatric diseases [34, 101, 111, 113, 114]. Several groups have conducted whole-exome sequencing (WES) on ASD families, and collectively, these studies indicate that discrete *de novo* mutation (single nucleotide variation or small insertions and deletions) contribute to the overall genetic risk of ASD [124-127]. Among these variations is the first ASD-associated, *de novo* mutation found in the human dopamine (DA) transporter (DAT) gene (*SLC6A3*) [125]. This mutation results in a Thr to Met substitution at position 356 (DAT T356M). The consequences of this *de novo* mutation and its impact on DA

neurotransmission have yet to be elucidated.

The neurotransmitter DA has an important role in the central nervous system by regulating a variety of functions, including motor activity, motivation, attention and reward [192-195]. Disrupted DA function is implicated in a number of neuropsychiatric disorders, including bipolar disorder, schizophrenia, attention deficit hyperactivity disorder (ADHD) [234-236] and, more recently, ASD. (See Chapter 3) The dopamine transporter (DAT) is a presynaptic membrane protein that regulates DA neurotransmission via the high-affinity reuptake of synaptically released DA [237]. It is the major molecular target of cocaine (COC), amphetamine (AMPH; Adderall) and methylphenidate (Ritalin) [238-242]. Owing to DAT's role in DA neurotransmission, *SLC6A3* variants have been a focus of genetic association studies linking the etiology of brain disorders to dysregulated DA neurotransmission [243, 244]. Recent studies have identified a rare, inherited, functional missense *SLC6A3* variant, DAT A559V (rs28364997), and that has been associated with ADHD, a phenotype commonly comorbid in ASD subjects [8, 201-204, 245, 246]. These studies point to a contribution of DAT genetic variants in complex brain disorders.

Although an important role for DA in ADHD has been established [247]. DA's role in ASD is poorly understood. [248] Many individuals with ASD exhibit co-occurrence of ADHD symptoms (~20–45%). The comorbid nature of ADHD with ASD points to dysregulation of common signaling pathways (for example, DA) as a mechanism underlying these neuropsychiatric disorders. [249]

In this chapter, I present the characterization of the first ASD-identified, *de novo*

mutation in DAT by structural, functional and behavioral analysis. These results implicate altered DA homeostasis as a potential liability in ASD risk.

Methods

Subjects and clinical assessment

Subjects from this family include the proband, both parents and an unaffected sibling who were recruited by the Boston Autism Consortium as described previously [125, 250]. Clinical assessment followed standard research criteria for ASD diagnosis. The proband was classified as having a comparatively narrow "autism" diagnosis (as opposed to a "broader ASD") on the basis of diagnostic algorithms from the Autism Diagnostic Interview Revised (ADI-R) [2] using criteria described by Risi et. al [251], and classification resulting from the diagnostic algorithm of the Autism Diagnostic Observational Schedule (ADOS) [3]. Proband IQ was assessed at age 5 years, 9 months of age using the Wechsler Preschool and Primary Scale of Intelligence (WPPSI; Wechsler, D. (1967)). The Social Responsiveness Scale (SRS; Western Psychological Services) was performed on both parents to index the presence and severity of broader autism phenotype traits. The biological mother provided medical and family history.

SLC6A3 T356M *de novo* discovery

Methodological details and results for the discovery and validation of the T356M *de novo* mutation were published by Neale et al [125]. In brief, DNA derived from whole blood on both parents and the proband were exome captured, sequenced, putative de

novo variation called, and subsequently validated. Exons were captured in this trio using the Agilent 38Mb SureSelect v2 and sequenced using the IlluminaHiSeq2000 at the Broad Institute (Boston, MA). Read processing was performed using Picard (http://picard.sourceforge.net), mapped to RefSeq19 using BWA, and variable sites called using GATK [252]. The T356M variant, identified as a heterozygote in the proband and absent in both parents, was experimentally validated and confirmed to be a *de novo* mutation that does not appear in the unaffected sibling.

Cell culture and transfection

The GFP-synDAT-pCIHygro expression vectors containing GFP-tagged DAT WT and DAT T356M sequence were engineered and amino acid substitution was confirmed by direct sequencing. Expression vectors were transiently transfected into Chinese hamster ovary (CHO) cells using the FuGENE-6 (Promega) system in a 6:1 transfection reagent:DNA ratio. Cells were stored in a 5% CO_2 incubator at 37°C and maintained in Ham's F-12 medium supplemented with 10% fetal bovine serum (FBS), 1 mM l-glutamine, 100 U/mL penicillin, and 100 µg/mL streptomycin. Assays were conducted 24-48 hr post-transfection.

Amperometry and patch clamp electrophysiology

Wildtype DAT and T356M DAT expressing cells were plated at a density of ~30,000 per 35-mm culture dish and experiments conducted as described in Bowton *et al*. [199, 200]. Patch pipettes for loading of DA were generated using a programmable

puller, model P-2000 (Sutter Instruments), to confer a resistance of 5 MΩ. The pipettes were filled with a 2 mM DA containing internal solution (120 mM KCl, 10 mM HEPES, 0.1 mM $CaCl_2$, 2 mM $MgCl_2$, 1.1 mM EGTA, 30 mM Dextrose, pH 7.35, and 275 mOsm). Cells were washed with the external bath solution (130 mM NaCl, 10 mM HEPES, 34 mM D-glucose, 1.5 mM $CaCl_2$, 0.5 mM $MgSO_4$, 1.3 mM KH_2PO_4, adjusted pH to 7.35, and 300 mOsm). Following internal access to the transfected cells, the DA containing solution was allowed to diffuse into the cell for 10 minutes. A carbon fiber electrode (ProCFE; fiber diameter of 5 μm) was located neighboring the plasma membrane and held at +700 mV (a potential greater than the oxidation potential of DA). Amperometric currents were recorded using an Axopatch 200B amplifier (Molecular Devices) in response to an addition of 10 μM AMPH or 10 μM COC. Amperometric traces were digitally filtered offline at 1 Hz using Clampex9 software (Molecular Devices). DA efflux was quantified as the peak value (in picoamperes) of the amperometric current.

[^{3}H]Dopamine uptake

Wildtype and T356M DAT expressing cells were seeded into polyornithine coated, 24-well plates 48 hr before assaying performed as previously described [253, 254]. Prior to measurement, cells were washed with KRH buffer (130 mM NaCl, 1.3 mM KCl, 2.2 mM $CaCl_2$, 10 mM HEPES, and 1.2 mM KH_2PO_4 supplemented with 10 mM D-glucose, 0.1 mM ascorbic acid, 1 mM tropolone, and 0.1 mM pargyline, pH 7.4). DAT activity was assayed with 50 nM [^{3}H]DA (dihydroxyphenylethylamine 3,4-[ring-2,5,6-^{3}H]) for 15 min at 37°C . Nonspecific activity was determined by 5 min pretreatment with 10

μM COC. [^{3}H]DA uptake was quantified with the TopCount NXT™ Microplate Scintillation Counter (PerkinElmer). All data are presented as non-specific subtracted.

Amphetamine uptake

Plated wildtype and T356M DAT expressing were washed three times with KRH buffer and incubated for 5 min at 37°C in the presence of 10 nM AMPH. Cells were washed three times with ice-cold KRH. After solubilization in extraction buffer, AMPH was quantified using HPLC.

Cell surface biotinylation and western blo*t*

For cell surface biotinylation assays and western blots, CHO cells were cultures in 6-well plates and transiently transfected with wildtype DAT or T356M DAT expression constructs. For cell surface biotinylation assays, cells were labeled with sulfo-NHS-SS-biotin (1.0 mg/ml) and analysis conducted as in Mazei et al [200]. DAT protein was detected using a rat monoclonal primary antibody to the N-terminus of hDAT (1:1000) and a goat-anti-rat-HRP-conjugated secondary antibody (1:5000)

Double Electron Electron Resonance

Cysteine residues were introduced using site directed mutagenesis into LeuT and LeuT A289M constructs and experiments conducted as described in Claxton et al. [255]. In Figure 5.4, Apo refers to ion Na^+ and leucine-free transporter while the +NaL state was obtained by addition of 200 mM NaCl and 4-fold molar excess of Leu relative to

LeuT. Double electron electron resonance [256] was performed at 83K on a Bruker 580 pulsed EPR spectrometer operating at Q-band frequency [257] using a standard 4-pulse sequence as previously described. Double electron electron resonance echo decays were analyzed to obtain distance distributions [258].

Drosophila genetics

Drosophila homozygotes for the DAT null allele DAT^{fmn} (dDAT KO) [259] and flies harboring TH-Gal4 [260] were outcrossed to a control line (Bloomington Indiana (BI) 6326) and selected by PCR or by eye color. TH-GAL4 (Bl 8848) and M[vas-int.Dm]ZH-2A, M[3xP3-RFP.attP'] ZH-22A (Bl 24481) were obtained from the BI stock center and outcrossed to dDAT KO flies carrying the *white* (w^{1118}) mutation (BI stock number 6236) for 5–10 generations. Transgenes (hDAT or hDAT T356M) were cloned into pBI-UASC [261] and constructs were injected into embryos from M[vas-int.Dm]ZH-2A, M[3xP3-RFP.attP']ZH-22A (Bl 24481). Initial potential transformants were isolated, selected, and maintained by standard methods.

Behavioral analysis

Three-day post eclosion, male *Drosophila* were collected and placed into tubes with food for 72 hr. Locomotion was recorded by beam breaks and analyzed using equipment/software from Trikinetics (www.trikinetics.com). For the AMPH-induced locomotion, males were starved for 6 hr and then fed sucrose (5 mM) containing either

AMPH (10 mM) or vehicle. Data were analyzed by one-way analysis of variance followed by a Newman-Keuls Multiple Comparison Post-test.

Results

T356M *de novo* DAT variant has impaired function

A recent study assessed the role of de novo variation in ASD by using whole-exome sequencing in 175 ASD parent–child trios [125]. In the data presented in this chapter, a *de novo SLC6A3* variant was identified, resulting in a Thr to Met substitution at site 356. Given the rarity of non-synonymous de novo events, it is not surprising that this mutation was absent in the ~1000 unrelated ASD cases and controls and has not been deposited in 1000 genomes [262], single-nucleotide polymorphism database (build 137) [263], or the NHLBI (National Heart, Lung, and Blood Institute) Exome Sequencing Project. The subject carried no other coding *de novo* mutations. The T356 residue is completely conserved across several species (Figure 5.1A). Importantly, residue 356 is located in the seventh transmembrane domain of DAT and resides in a highly conserved region implicated in ion binding [264]. *In silico* molecular modeling of the T356 mutant protein is shown in Figure 5.1B.

The subject harboring this *de novo* mutation is the elder male child of healthy non-consanguineous Caucasian parents (proband has a healthy younger sibling). There is no immediate family history of ASD or related psychiatric conditions. The subject has a normal IQ (94) and has no history of seizures or other co-morbidities. By the age of 6 years, the proband was diagnosed with ASD (see Appendix B for full clinical reports).

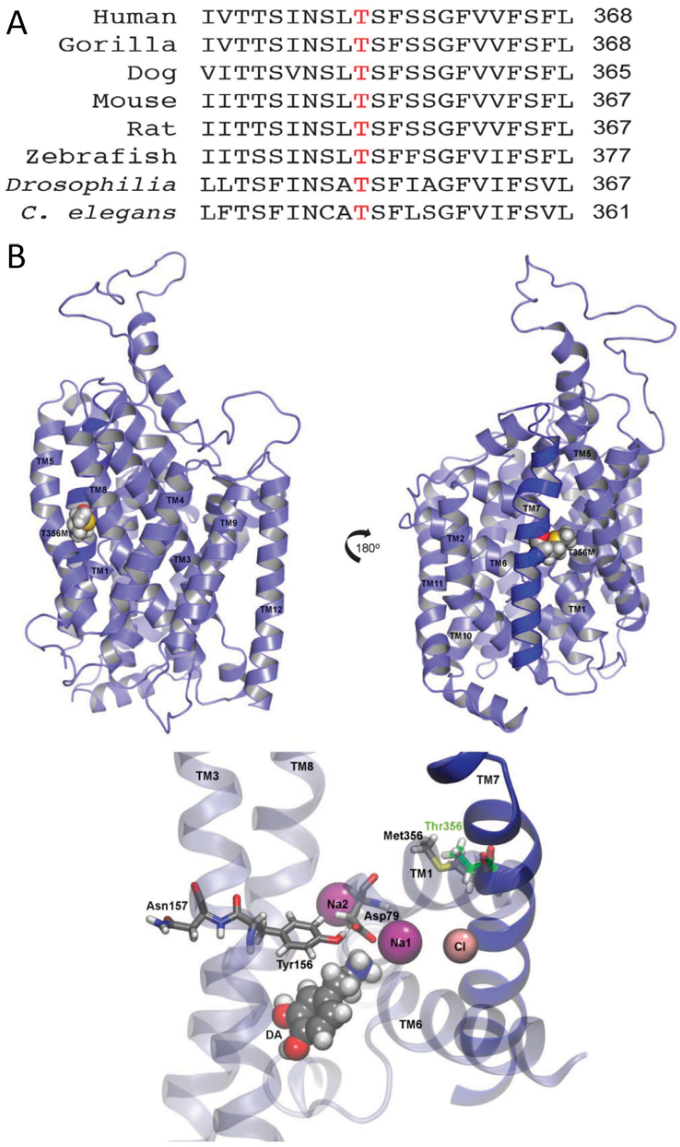
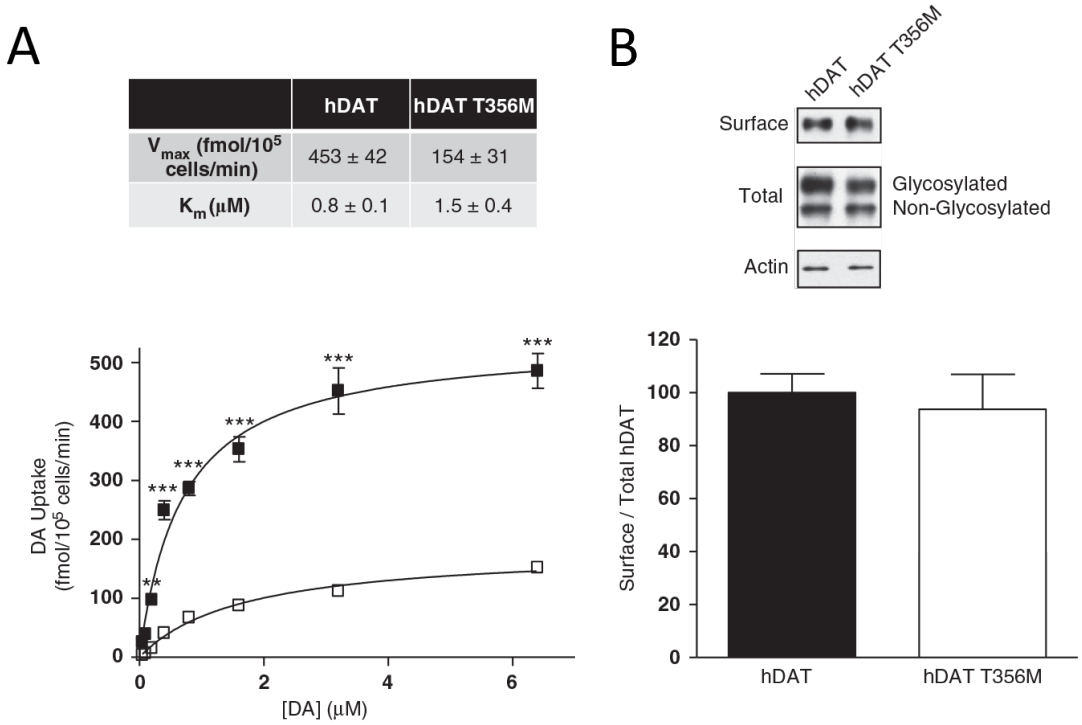


Figure 5.1: Cross-species conservation and *in silico* mutagenesis of T356. (**A**) Alignment of the dopamine transporter (DAT), amino-acid sequence across multiple species. Threonine 356 is represented in red. (B) In an equilibrated three-dimensional homology model of hDAT, the T356M mutation is located on transmembrane domain 7 (TM7). Top: schematic views representing a 180° rotation show T356M with respect to the TM helices. TM7 is shown in dark blue. Bottom: critical resides that interact with dopamine (DA) are shown, as well as the bound Na^+ and Cl^- ions. The methionine is rendered together at position 356 with the wildtype threonine (green).

To evaluate the T356M variant, the activities of wildtype and T356M DAT were compared in a heterologous expression system and radioactive [^{3}H]DA uptake and affinity were calculated. In DAT T356M cells, the maximal velocity of DA influx (V_{max}) was significantly reduced, whereas the apparent DA affinity (K_m) of hDAT T356M was not significantly different from that of hDAT (Figure 5.2A, top). A representative plot of DA uptake kinetics for DAT and DAT T356M is shown in Figure 5.2A, bottom. The reduced [^{3}H]DA transport capacity was not associated with a reduction in either total or DAT surface expression (Figure 5.2B, top), as assessed by measuring changes in DAT proteins in the total and biotinylated fraction, respectively. The total fraction for DAT and DAT T356M contained both glycosylated and non-glycosylated forms of the DAT. Surface fractions were quantitated, normalized to total DAT (glycosylated) and expressed as a percent of DAT (Figure 5.2B; bottom). Furthermore, normalizing the total DAT fraction (glycosylated) to actin-loading control yielded no significant differences between DAT and DAT T356M total expression. Data are expressed as a percentage of wiltdtype DAT (wildtype DAT 100±17.6% *vs* T356M DAT 96.4±13.7%; $P \geq 0.87$ by Student's t-test; n = 8–11).

DAT T356M displays anomalous DA efflux

Although DAT T356M displays similar surface expression to that of DAT, it demonstrates reduced ability to accumulate intracellular DA. One possibility is that constitutive DA efflux, here referred to as anomalous DA efflux (ADE), contributes to this reduced DA uptake in the DAT T356M cells. This efflux would impede the intracellular



	hDAT	hDAT T356M
V_{max} (fmol/10^5 cells/min)	453 ± 42	154 ± 31
K_m (μM)	0.8 ± 0.1	1.5 ± 0.4

Figure 5.2: Human dopamine transporter (DAT) T356M has impaired function. (**A**) Top: kinetic parameters (V_{max} and K_m) for DAT and DAT T356M (V_{max}: P≥0.005; by Student's t-test; n=3, in triplicate; K_m: P=0.20; by Student's t-test; n=3, in triplicate). Bottom: representative plot of [^{3}H]DA uptake kinetics in DAT (filled squares) or DAT T356M (empty squares) cells (**=P≥0.01, ***=P≥0.001; by two-way analysis of variance (ANOVA) followed by Bonferroni post-test; n=3, in triplicate). (**B**) Representative immunoblots for biotinylated (surface) and total protein fractions from DAT and DAT T356M cells. Surface fractions were quantitated, normalized to total DAT (glycosylated) and expressed as a percent of DAT (P=0.05; by Student's t-test; n=8–11).

accumulation of DA. To determine whether DAT T356M exhibits ADE, cells were whole-cell patch clamped and perfused for 10 min with an internal solution containing 2 mM DA [199]. The electrode, in current clamp configuration, allows the cell to control its membrane voltage. In addition, this technique ensures that cells expressing either DAT or DAT T356M were equally loaded with DA. DA efflux was quantified via amperometry [199]. It has been previously shown that, in the presence of ADE, COC decreases the amperometric signal through blockade of DAT [200]. In DAT cells, amperometric currents were unaffected by the application of COC (Figure 5.3A, top right) indicating no ADE. In contrast, amperometric signals from DAT T356M cells were significantly reduced by the application of COC, revealing ADE (Figure 5.3A, top left). DAT T356M-expressing cells displayed a significant increase in ADE compared with DAT-transfected cells (Figure 5.3A, bottom).

Cell membrane depolarization has been shown to support DA efflux [265]. Figure 5.3B reveals that there is not a significant difference in resting membrane potential (measured in current clamp) between cells expressing DAT or DAT T356M. This indicates that differences in the function of DAT and DAT T356M are not a result of resting membrane potential.

To examine potential changes in AMPH to induce DA efflux, cells were patch clamped and DA dopamine delivered into either wildtype DAT or DAT T356M expressing cells while recording DA efflux with amperometry. AMPH dose-response assays (measuring the peak of the amperometric current at different AMPH concentrations) revealed that DAT T356M and DAT cells have comparable AMPH EC_{50} (EC_{50}; DAT:

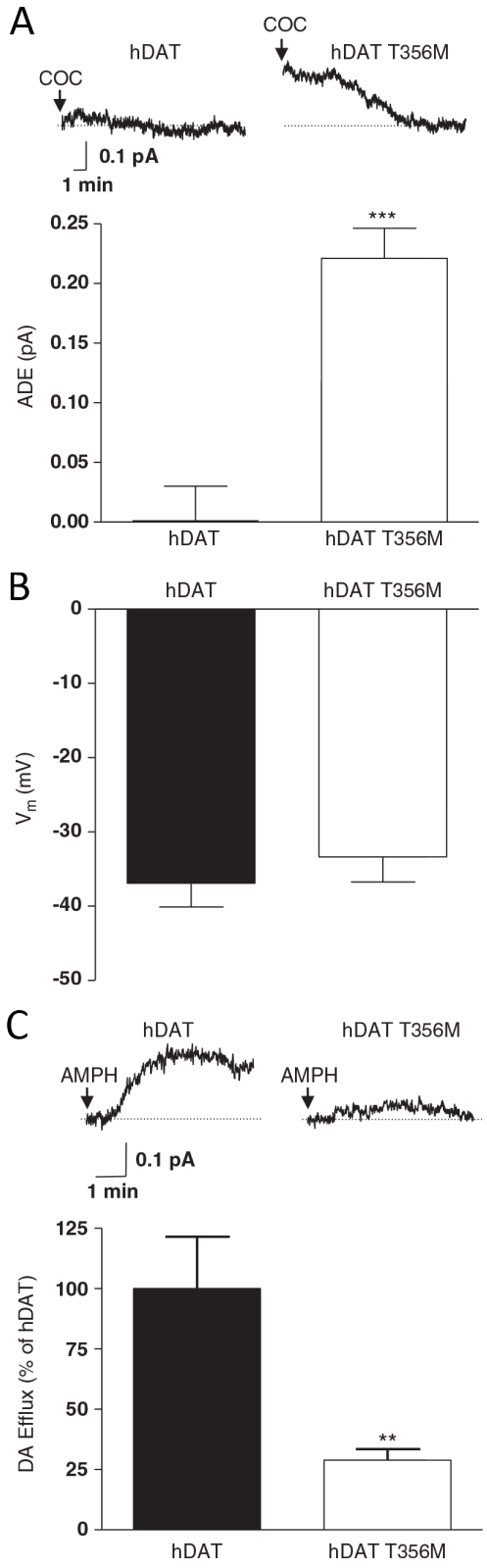

A
hDAT
COC
hDAT T356M
COC
0.1 pA
1 min
***
0.25
0.20
0.15
0.10
0.05
0.00
ADE (pA)
hDAT
hDAT T356M
B
hDAT
hDAT T356M
0
-10
-20
-30
-40
-50
V_m (mV)
C
hDAT
hDAT T356M
AMPH
AMPH
0.1 pA
1 min
125
100
75
50
25
0
DA Efflux (% of hDAT)
**
hDAT
hDAT T356M

Figure 5.3: Human dopamine transporter (DAT) T356M exhibits robust ADE. (**A**) Top: representative amperometric currents recorded from DAT and DAT T356M cells. Arrows indicate application of 10 µM cocaine (COC). Bottom: quantitation of the COC-induced reduction in the amperometric current (ADE). Data are reported as maximal current (***=P≤0.001 by Student's t-test; n=4–5). **(B)** DAT T356M cells do not display altered resting membrane potential with respect to DAT cells (P≥0.05 by Student's t-test, n=9–25). (**C**) Representative amphetamine (AMPH)-induced amperometric currents recorded from hDAT and DAT T356M cells. Arrows indicate application of 10 µM AMPH. Bottom: quantitation of AMPH-induced dopamine (DA) efflux. Data are represented as maximal current expressed as percent of the current recorded in hDAT cells (**=P≤0.01 by Student's t-test; n=5–7).

0.15±0.05 µM; DAT T356M: 0.16±0.07 µM; n=4; p≥0.95 by Student's t-test). Then, using a saturating AMPH concentration (10 µM), DA efflux was calculated in either DAT or DAT T356M cells (Figure 5.3C, top). AMPH-induced DA efflux was significantly reduced in DAT T356M cells compared with wildtype DAT cells (Figure 5.3C, bottom). These results strongly suggest that ADE does not share common mechanisms with AMPH-induced DA efflux.

Substitution of LeuT Ala289Met promotes an outward facing conformation

To investigate the structural consequences of the T356M *de novo* mutation in a DAT homolog with a known crystal structure, we analyzed changes in the conformational cycle of the leucine transporter (LeuT). We substituted A289 (the homologous amino acid to T356) with a Met (LeuT A289M). We measured distances between pairs of spin labels (r (Å)) and the distance distribution (P(r), the probability of a given distance between the two labels) monitoring the intra- and extracellular gates by double electron electron resonance [266]. First, we examined the pair 309/480 (Figure 5.4A, left) that monitors the relative movement of the extracellular loop 4 in LeuT. This loop obstructs the permeation pathway in the Apo conformation [266], as indicated by the close proximity of the pair 309/480 (Figure 5.4A, middle Apo black line). Upon Na^+ binding, the distance between extracellular loop 4 and TM12 increases [266], indicating opening of the extracellular vestibule and enabling substrate access. Subsequent Leu binding resets the closed extracellular loop 4 conformation in the occluded conformation of the transporter (Na^+ and Leu bound in the vestibule) (Figure

5.4A, right; +Na/L black line). The extracellular Apo site (Figure 5.4A, middle; Apo; compare red and black lines) as well as Na^+-bound state is similar in LeuT and LeuT A289M. Yet, LeuT A289M with Na^+ and Leu bound in the vestibule has a destabilized bound structure with fluctuations on the extracellular side (Figure 5.4A, right; +Na/L; compare red and black lines). The probability distribution in the Na/Leu-bound state contains distinct populations of conformations that indicate fluctuations of LeuT A289M to a permeation pathway that has increased probability to be open to the outside, relative to LeuT (Figure 5.4A, right; red line, arrows).

We then examined the pair 7/86 (Figure 5.4B, left) to determine the distance between the N-terminus and intracellular loop 1 (IL1) and to monitor fluctuation dynamics on the intracellular side. This is necessary to describe changes in the population of transporters with an inward-facing conformation [266]. In the LeuT background, distance distributions between spin [267] or fluorescent probes [268] in the Apo state are bimodal, reflecting the equilibrium of this intracellular gate between closed and open conformations (Figure 5.4B, middle; Apo; black lines). Introduction of the A289M leads to a shift in the equilibrium to favor the closed conformation side (Figure 5.4B, middle; Apo; compare red and black lines, arrow). Na+ binding does not alter the equilibrium between the two conformations (data not shown), whereas Na. and Leu binding resets this shift to LeuT-like conformations (Figure 5.4B, middle; +Na/L; compare red and black lines).

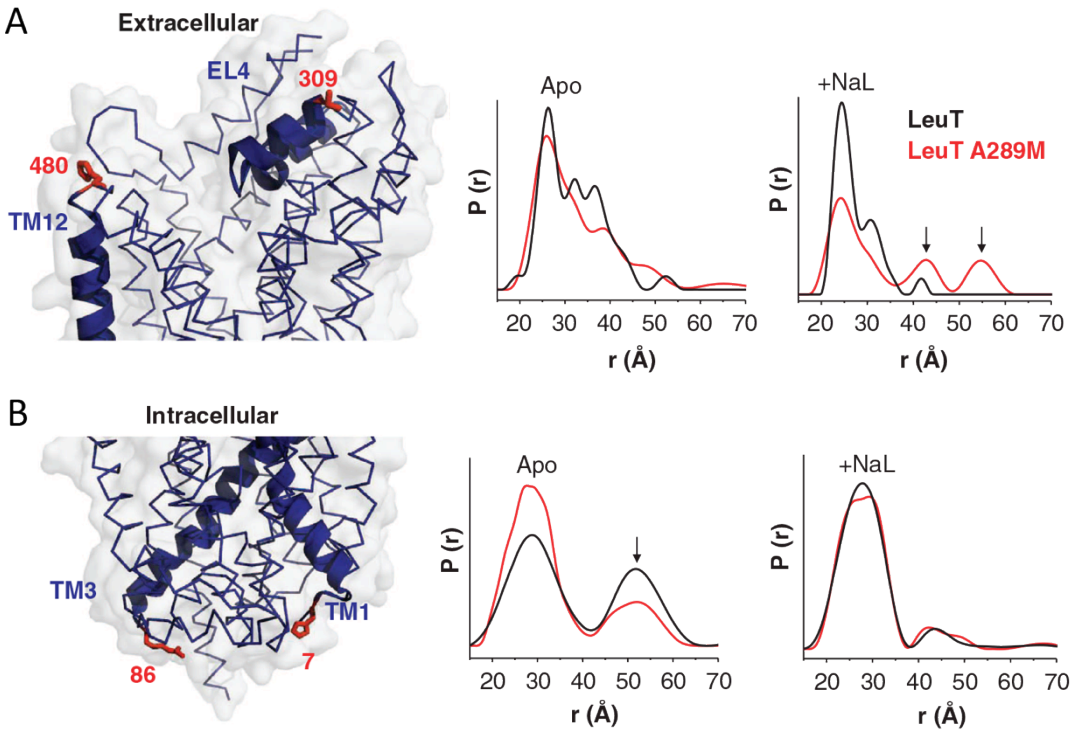


Figure 5.4: In leucine transporter (LeuT), substitution of Ala289 with a Met supports an outward-open facing conformation. Distance distributions of extracellular and intracellular spin labeled Cys pairs in LeuT reveal changes in the conformational equilibrium caused by mutating Ala289 to a Met. (**A**) Left: extracellular reporter pairs (309–480) tagged on three-dimensional structure of LeuT. Right: distance of the extracellular reporter pair for LeuT (black) and A289M (red), in the Apo conformation (Apo) and in the presence of Na^+ and Leu (+NaL). (**B**) Left: intracellular reporter pairs (7–86) tagged on the three-dimensional structure of LeuT. Right: distance of the intracellular reporter pair for LeuT (black) and A289M (red), in the Apo conformation (Apo) and in the presence of Na^+ and Leu (+NaL). The LeuT structure was obtained from PDB 2A65. The structures were generated using PyMOL.

Our results demonstrate that, because of the A289M, the presence of substrate and Na^+ fails to completely close the extracellular pathway as in LeuT, inducing fluctuations on the extracellular side. These fluctuations to an open-to-the-outside permeation pathway persist, possibly enabling substrate release. This is in contrast to LeuT, where substrate binding closes the extracellular permeation pathway.

Drosophila expressing hDAT T356M in DA neurons are hyperactive

Locomotion is a fundamental behavior regulated across species by DA [269-271] therefore, locomotion in flies offers a powerful model for exposing a behavioral impact of ADE associated with the DAT T356M. To use locomotion as a behavioral output, flies expressing either hDAT or hDAT T356M in DA neurons were generated using a Gal4/UAS system in a dDAT KO mutant background. To generate these transgenic flies, the phiC31-based integration was used, which promotes equivalent expression of mRNA levels for the relevant transgenes. Locomotion was calculated by beam crossings over a >24 hr period and data binned in 15 min intervals during both the light and dark cycles. Consistent with previous reports [259], dDAT KO flies in this study are hyperactive as compared to flies expressing hDAT in DA neurons, validating the approach used (Figure 5.5A, compare hDAT with dDAT KO). Flies expressing hDAT T356M show a significant increase in locomotion (Figure 5.5A), predicted to be a consequence of increase extracellular DA as a result of persistent ADE. Total activity of hDAT T356M and dDAT KO flies are significantly higher than that of wildtype hDAT flies. This is illustrated in Figure 5.5B, which shows total locomotor activity over a 24 hr period across the different fly

lines.

Since it was shown in Figure 5.3 that hDAT T356M cells display compromised AMPH-induced DA efflux, it was reasoned that a reduced increase in locomotion in response to AMPH would be seen in flies expression hDAT T356M. To test this, changes in locomotion were determined upon AMPH or vehicle exposure (15 min) and calculated as beam crosses. No significant increase in locomotion was seen in hDAT T356M flies when exposed to AMPH as compared with vehicle control (hDAT T356M (vehicle) 9.7 ± 0.7 beam breaks vs. hDAT T356M (AMPH) 12.7 ± 1.6 beam breaks; $n = 24$, $P \geq 0.05$). This is in contrast to flies expressing hDAT, where AMPH induced a significant increase in locomotion (hDAT (vehicle) 6.2 ± 0.9 beam breaks vs hDAT (AMPH) 18.2 ± 1.0 beam breaks; $n = 24$; $P \leq 0.001$). In the dDAT KO flies, AMPH failed to provoke a significant increase in locomotion, a result similar to that of hDAT T356M.

Discussion

Alterations in DA tone underlie multiple neuropsychiatric disorders, including bipolar disorder, schizophrenia and ADHD [234-236]. With respect to ADHD, altered DA signaling, including changes in DAT function, may contribute to the cognitive and hyperactive traits of the disorder [199, 200] ASD, like ADHD, is phenotypically and etiologically complex, with evidence residing, at least in part, to dopaminergic factors. Genes harboring *de novo* events are highly meaningful to better our understanding the etiology of ASD [124-127]. Therefore, the biological networks identified from these *de novo* events, and the broader pathways they function within, are candidate risk factors

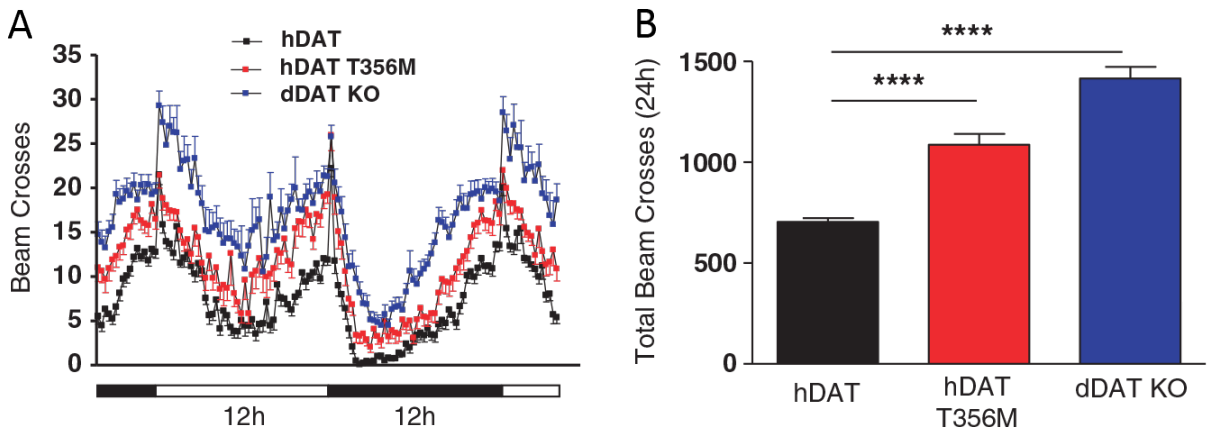


Figure 5.5: Expression of human dopamine transporter (hDAT) T356M in *Drosophila* leads to hyperactivity. hDAT or hDAT T356M was expressed in DA neurons of dDAT KO flies. (**a**) Locomotor activity was assayed over 32 h during the light (horizontal white bars) or dark (horizontal black bars) cycle. Flies expressing hDAT T356M (red squares), as well as dDAT KO flies (blue squares), were hyperactive throughout the 32 h period with respect to flies expressing wild-type hDAT (black squares) (n=32; beam breaks binned in 15 min intervals). (**B**) Quantitation of total beam crosses over 24 h for hDAT, hDAT T356M and dDAT KO flies (****=P≤0.0001; n=32)

for ASD. The motivation for the study presented here was to characterize the mutant transporter using a diverse array of approaches.

The amperometric recordings demonstrate that the *de novo* DAT T356M mutation confers COC sensitive ADE. It is intriguing to speculate that anomalous transporter mediated neurotransmitter efflux may be an unappreciated source of risk for mental illness, especially in disorders associated with altered DA signaling. It is possible that ADE, driven by DAT variants or variants in other genes in the DAT regulatory network, could impact risk for ASD. This point has been made in the context of ADHD for the DAT A559V variant [199, 200].

To address how the T356M *de novo* point mutation perturbs transporter structure to trigger ADE, this study leveraged the DAT bacterial homolog, LeuT. As the crystal structure of hDAT is unavailable, changes we analyzed in the conformational cycle of LeuT. In LeuT A289M, we measured the distances between pairs of spin labels monitoring the intra- and extracellular gates by double electron electron resonance. In contrast to LeuT, the spin labels monitoring the extracellular gate to LeuT A289M show that the presence of Na^+ and leucine promotes a permeation pathway unoccluded to the outside. By no means definitive, this study strongly suggests that the mechanism of ADE for DAT T356M is distinct from that of DAT A559V, which is a result of a tonic activation of DRD2 and the downstream kinase CaMKII [126]. One intriguing possibility is that there may be multiple mechanistic routes to promote DAT-mediated ADE, and ADE might support a common convergent point in the shared etiology of ASD and ADHD.

In vivo hDAT T356M may alter extracellular DA levels and, as a consequence, increase locomotion [259]. dDAT KO flies were engineered to selectively expressed hDAT T356M specifically in DA neurons. *Drosophila* expressing hDAT T356M exhibited prominent hyperactivity compared to Drosophila expressing hDAT. Lastly, AMPH has an diminished ability to increase locomotion in hDAT T356M (and dDAT KO) flies, hypothetically from a decreased ability of AMPH to induce DA efflux in hDAT T356M expressing neurons.

Summary

The results presented in this chapter support the hypothesis that dysfunction of DAT, a monoamine transporter, may influence ASD susceptibility. These results encourage further genetic investigation of DAT, and proteins that exert influence on DAT, as candidate ASD risk genes. By reporting the functional, structural and behavioral consequences of the first identified *de novo* DAT variant associated with ASD the work presented herein adds to the growing body of data implicating dopamine transporter dysfunction, and more broadly dopamine dysfunction, in psychiatric illness.

CHAPTER VI

Whole exome sequencing reveals minimal differences between cell line and whole blood derived DNA

Introduction

The data presented in this chapter reflect a significant amount of method validation and the results influenced much of the conclusions drawn from the manuscript by Schafer et al. [272].

Next-generation sequencing (NGS) has become an affordable tool to probe the human genome for rare variants affecting risk for disease. Association analyses are being applied to NGS data to discover loci with rare variants that affect the risk of a wide spectrum of diseases. Nevertheless, it is becoming increasingly apparent that very large sample sizes and widespread availability will be required to attain significant results in many of these studies. To attain large samples, scientists would make use of large collections of DNA from subjects with immortalized lymphoblastoid cell lines (LCLs) in repositories such as the NIMH Center for Collaborative Genomic Disorders on Mental Disorders (hereafter the "NIMH Repository") at the Rutgers Cell and DNA Repository (RUCDR). However, due to reports of large numbers (20 per exome) of non-germline mutations in LCLs [273], concerns now exist about the use of LCL-derived DNA for sequencing studies. The concern is well founded, because immortalization involves transformation of lymphocytes with cytomegalovirus (CMV) and this, combined with extensive serial passaging of cells, could lead to DNA sequence changes. Ultimately,

mutation underlies all heritable genetic variation but it has recently become apparent that mutations that have arising *de novo* can be highly informative in identifying genetic risk factors for disorders such as autism and schizophrenia [124-127, 274, 275]. Non-germline mutations (and potential artifacts) are particularly troubling in this setting, thus motivating a careful investigation of the quality of LCLs as a basis for such studies.

The work presented in this chapter, sought to empirically test the degree to which mutations that arise in LCLs might impact results from WES studies. Therefore, WES on paired DNA samples derived from whole blood (WB) and LCL of low passage number for 16 subjects were compared. Cell lines were obtained from the NIMH Repository at RUCDR or from the Icahn School of Medicine at Mount Sinai. Specifically, LCLs had been cultured through a maximum of four passages, in which a relatively small number of cells seed a new culture to greatly expand the cell populations. These cell populations are used for DNA isolation or cryopreservation of multiple new culture aliquots. This chapter presents 1) initial detection of genotype calls discordant between paired WB and LCL samples; 2) a novel filtering and prioritization algorithm that effectively assigns putative differences into categories of "confidence" as demonstrated by subsequent validation experiments; 3) that the vast majority of called WB/LCL differences are false positives and that those that do validate reflect DNA mutations that arose in the cell line; 4) in nearly all cases true positives appear as low level mosaic non-reference alleles; and finally 5) only 3-4 valid sequence differences per subject exist.

Methods

Validation of selected candidates

Validation of the selected variant calls was carried out using standard Sanger dye-terminator sequencing of amplimers for regions containing putative WB/LCL sequence differences. Primers were designed using Primer 3 software and subjected to a BLAST-like alignment tool to ensure amplification specificity. PCR products were amplified from ~20 ng DNA using AmpliTaq Gold PCR MasterMix (Life Technologies; Carlsbad, CA) with individually optimized PCR conditions in a total volume of 20 µl. Genomic and LCL DNA from the same subject were amplified simultaneously. Amplified DNA fragments were then purified using Multiscreen PCR filtration plates (Millipore; Billerica, MA) and yields were determined using a NanoDrop spectrophotomer ND-1000. Sequencing reactions were carried out using the BigDye Terminator v3.1 Cycle Sequencing Kit, and reactions were subjected to analysis on an ABI 3730xl DNA analyzer. Sequence electropherogram data were viewed to examine sites of putative difference using the computer software Sequencer v5.0.1 (GeneCodes; Ann Arbor, MI).

All variants in Groups 0, 1, and 2 were sequenced. In addition validation of a subset of Group 3 were conducted. Instead of validating a random subset of Group 3, results from the first three groups were exploited to determine what, if any, features of the candidates increased their chances of being real mismatches. It was determined that among the confirmed mismatches, almost all shared two characteristics: first, the proportion of the WB DNA calls that corresponded to the reference allele was either very close to zero or very close to one. Second, the cell line reads for both alleles were

well represented on both the forward and reverse strands. Hence, a filter was applied that labeled a candidate as "strong" if (1) the homozygous call the proportion of reads for the reference allele was either larger than 0.97 or smaller than 0.03 and (2) the proportion of reads in a single direction did not exceed 0.95 for either allele present in the cell line. The remaining candidates were labeled "weak." Group 3 validation was therefore pursued chiefly for "strong" candidates.

Results

Segregation of the candidate disagreements between the 16 WB and LCL pairs

In total 15,099 call pairs were candidates for disagreement between the WB and LCL pairs. Filtering of the 15,099 candidates resulted in 864 candidate differences. These 864 candidates were segregated into four groups of "confidence" based on inspection of read quality, evidence for strand bias, and the relative balance of reference/alternative calls, and also by whether the heterozygote fell in blood or cell line. The candidates in each of Group 0, 1, and 2 are shown in Appendix C. Overall 44% of the entries in Table 6.1 declare blood heterozygous and cell homozygous, it is notable that only 19% of the confident differences (levels 0-2) differ in this direction.

Validation of selected candidates

For Group 0, all ten of the candidates for blood/cell line differences were confirmed to be mismatches by follow-up Sanger sequencing (see Table 6.1), which revealed small amplitude non-reference peaks on both forward and reverse strands. For

Table 6.1 Summary of validation results for candidate mismatches sorted by decreasing level of confidence

Group	Total count	Valid attempt	Valid success	Confirmed mismatch	Confirmed rate	Blood heterozygous	Ti/Tv ratio
0	10	10	10	10	100%	0	4.00
1	28	28	26	18	695	2	2.11
2	107	107	104	15	14%	26	0.91
3 "Strong"	34	24	13	3	1.1%[a]	352	0.47
"Weak"	685	9	9	0			

[a] Percent is estimated based on a weighted average

Groups 1 and 2, however, the percentage validated was 69% and 14% respectively. There are a handful of instances for which the validation was unsuccessful due to sequencing problems; these are excluded from the count in "Validations Successful" column.

The results from the validation of Group 3 candidates are also shown in Table 6.1, divided into "strong" and "weak" candidates, as described in the methods of Schafer et al. [272]. Taking a weighed average of the two subgroups, the estimated rate of mismatch in Group 3 is 1.1%. Assessing the Ti/Tv ratio of all candidates in each of the groups, there is a clear decreasing trend as quality decreases. Among all confirmed mismatches, the Ti/Tv ratio is 3.18.

For each candidate validated we determined that WB was actually homozygous, or LCL was heterozygous leading to concordant calls. This supports our conjecture that in the process of creating LCL from WB, errors are much more likely to result in homozygous loci becoming heterozygous.

Inferences

A notable feature of the validation results is that in those cases for which a blood/cell line disagreement was confirmed, the locus was homozygous for the allele on blood, but then, with only one possible exception, displayed the characteristics of mosaicism within the cell line. The mosaic nature of these loci is evident not only in the results of the targeted validation via Sanger sequencing, but also in the original WES results. For example, consider one of the candidates in Group 2 (Individual 4, on

chromosome 10 at position 103772671). In a pileup diagram showing the individual NGS reads as horizontal lines (Figure 6.1), all but one of the reads in the blood sample were for T, the reference allele; yet the cell line sample, a majority (62 of 73) of the reads were for T, but a quarter returned C. This is far from the 50/50 split one would expect if the site were homozygous. Sanger sequencing results were consistent in that the electropherogram (Figure 6.2) reveals a peak for both T and C for the cell line sample. This example of relative over-representation of the reference allele is consistent with virtually all of the other instances of validated blood/cell line disagreements. The exception is a single candidate for what the Sanger trace is balanced in one direction and unbalanced in the other direction.

In addition there were no confirmed mismatches among the candidates for which blood was called heterozygous and cell line called homozygous. Hence, the overall *mismatch* rate per person is also our estimate of the overall rate of *mosaics* per person. This estimate can be obtained by taking a weighted average of the rate within each group with weight calculated using the proportion of the candidates within that group (Table 6.1). Hence, the probability of a randomly chosen candidate being a mosaic is approximately 6.1%.

The estimated number of mosaics per individual is 0.0609 x 864/16 = 3.29. This estimate is predicated on the assumption that loci that were filtered in the initial steps of the processing consist entirely of mismatches that can be attributed to variant caller errors and poor quality reads. (Recall that the list of 864 candidates was culled form the initial set of 15,099 blood/cell line differences). It is natural to assume of the 14,235 loci

that were excluded from the focused read-level analysis as forming a fifth group, “Group 4”, whose

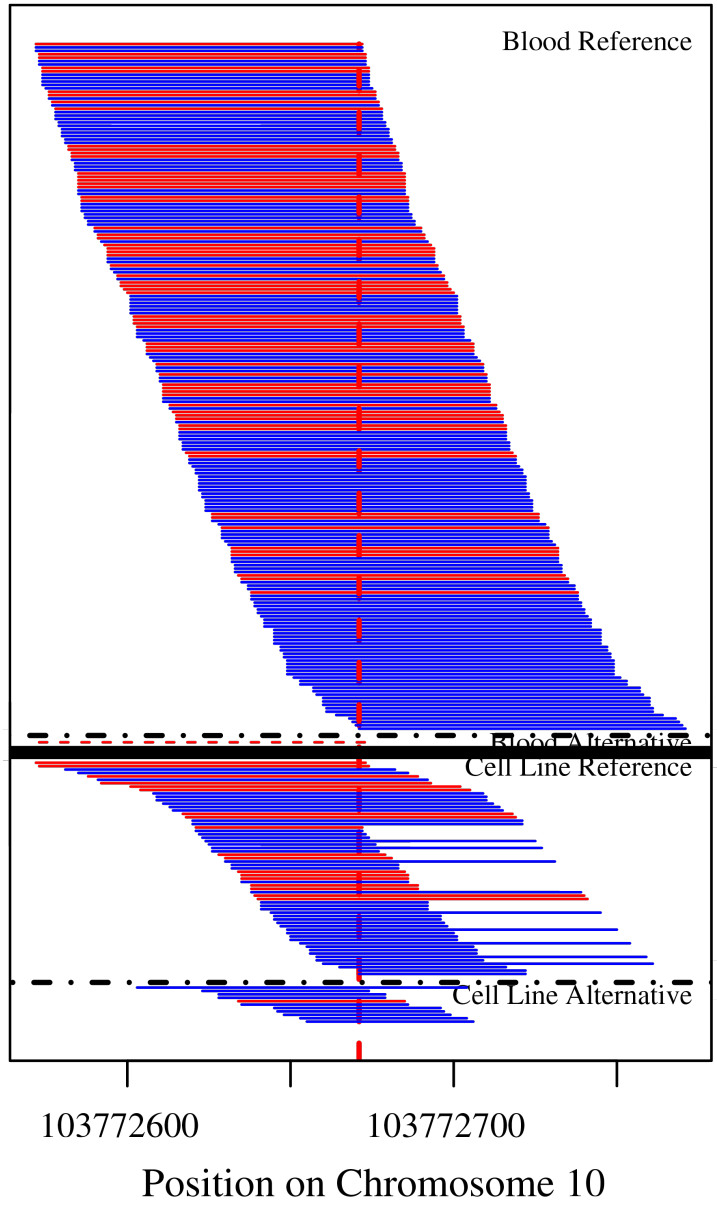


Figure 6.1: Pileup diagram from the sequencing results of Individual 4, on chromosome 10 at position 103772671, a confirmed mosaic. Note that although there is a large number of reads for the alternative allele on the cell line, the proportion of such reads is much less than one half of all the reads. The dashed line under “Blood

Alternative" represents a read that did not match either the reference or alternative allele for this site. The dashed line under "Cell Line Alternative" represents reads that did not match either the reference for this site.

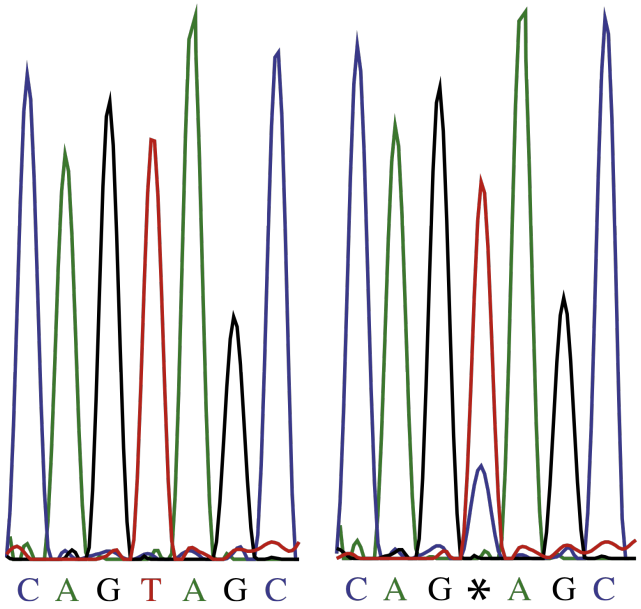


Figure 6.2: Results from Sanger sequencing of a confirmed mosaic. Sequence trace is from Individual 4, on chromosome 10 at position 103772671, as were the results from Figure 6.1. Note the additional peak in the chromatogram on the right. This result confirmed that this site was a mosaic as suggested in Figure 6.1.

quality metrics are even worse than the weak candidates in Group 3, and that among this group, a very small percentage would be true blood/cell line mismatches.

Discussion

To identify mutations in LCLs, results from WES of DNA from both whole blood and LCLs for 16 subjects were analyzed. All LCLs were at low passage (<5) and, therefore, the samples are much more likely to be representative of those in widely accessed public repositories. After filtering, 864 discordant genotype calls between blood and LCL samples (approximately 50/subject) were selected. These candidates were further prioritized based on read depth, base-call, mapping quality and read directions. This permitted segregation of candidate WB/LCL differences into four categories of decreasing "confidence." Sanger sequencing to evaluate each candidate variant in the top 3 confidence levels (N=145 variants) as well as some of the lower level candidates was preformed. The results of this validation study yielded few observations of confirmed WB/LCL differences, and in virtually every instance the difference manifested as low-level mosaicism in the LCL. Aside from these apparently mosaic loci, there was only one instance of a sequencing variant arising in the cell line and showing approximately equal allelic representation. The evidence for mosaicism was present not only in the Sanger sequencing, but also in the WES results. Hence, most of the 864 candidate differences were false positives and the remainder had characteristics that would cause them to be filtered by standard variant calling procedures.

There are two possible sources for the mosaicism. First there is a possibility that mutations accumulate in the cell lines during transformation and subsequent serial passaging. Second, lymphocytes represent a mixture of different cells that have been undergoing genomic changes from the clonal progeny they are derived. Any of these clones may have a point mutation that could be amplified during subsequent culturing. In either case, low passage of the LCLs and the careful analysis of sequencing traces for mosaicism would help identify such instances.

Although we have not tested higher passage LCLs, our results do point towards an elevated risk of using such samples. We identified 46 mosaic LCL mutations in the exomes of 16 subjects and predict fewer than ten remain in the unvalidated sample suggesting a rate of 3.3 per person per exome. Prior studies predict a rate of 1 germline mutation per exome per person. Our estimate of 3.3:1 non-germline to germline *de novo* mutation compares well with a prior estimate of 1:1 and differs sharply from the estimate 20:1 reported from HapMap Samples [273]. As noted, a difference in mutation rate could be caused by number of passages of the cell line as well as other factors, including external factors affecting mutation rate. The reduced and tractable number of mutations we observed likely reflects the focus on lower passage cells.

In earlier research, the suitability of LCL for use in genetic studies was evaluated in the context of SNP arrays [276]. To test for genotypic errors potentially induced by the Epstein-Barr virus transformation process, this study compared SNP genotype calls in WB vs. LCLC from the same individual using cells that were not passaged after immortalization. Genotypic discrepancies found in the matched WB and LCL pairs were

not notable relative to differences observed among control pairs, suggesting that most genotypic discrepancies were due to technical artifacts rather than the transformation process. Prior studies also supported the conclusion that LCL have a minor effect on genomic structural variation [277-279]. Even in the HapMap samples, putative LCL-specific genomic errors accounted for less that 0.5% of observed deletions [277].

The results of this work could be used to identify mosaic loci in studies that make use of LCL DNA. In particular, the evidence for mosaicism is most clearly defined by significant imbalance in the Sanger sequence trace showing the reference and variant allele, but it is also frequently apparent in WES statistics. Even without further filtering, the number of mosaics is sufficiently modest such that LCL-derived DNA is appropriate for gene discovery using case-control analyses of rare variants. If the mosaic loci were called as variant alleles, there would be minimal loss of power in gene discovery studies in complex genetics, because many samples are used and multiple independent hits are required for statistical evidence of association. For analyses of *de novo* variants in families, the statistics are more sensitive to errors. Nevertheless, multiple hits in the same gene are required for implicating a novel gene as one affecting risk [126], providing some measure of protection against false discoveries. Moreover, it is standard practice to validate *de novo* calls and this would typically identify mosaics. Important findings can be further validated in blood samples, when available, as well as independent samples. If validation in blood is not feasible, observations derived from individuals with numerous *de novo* mutations could be down weighted in analysis.

Summary

In summary, the results in this chapter strongly support the use of LCL-derived DNA for high-throughput sequencing in research, provided the cell lines have not been subject to extended serial passages. This work demonstrates DNA derived from low passage number LCL samples are appropriate for use in WES studies, and in turn validates many of the NGS studies where DNA derived from low passage number LCLs were used. This work also highlights the utility for rigorous filtering algorithms when handling and discriminating variants most likely to later validate from NGS raw reads.

CHAPTER VII

Conclusion

Interdisciplinary by design, this thesis sought to use genetics as a tool to gain insight into the biological complexity of autism spectrum disorder (ASD), through the lens of monoamine transporter dysfunction. In order to achieve my thesis goals, two independent aims were addressed in this thesis. In conclusion, the results are highly supportive that both a genetic and functional liability exists within the broad context of monoamine dysfunction and ASD. The results of this work add to the growing literature of monoamine dysfunction in psychiatric disease, specifically ASD.

APPENDIX

A. Autism phenotypic profile of rare variants in the *ADORA3* gene

	Leu90Val				Val171Ile				
NIMH Sample ID[1]	165-3551-0001	2103_256200*	60-1072-008	218-20036-1233001	165-3583-001	2220_1*	60-1001-001	72-0899-03	72-0907-302
Discovery vs. Replication	D	D	D	R	D	D	D	R	R
Gender	M	M	M	M	M	M	M	M	M
Ancestry[2]	AA	EUR	AA	EUR	EUR	EUR	EUR	EUR	EUR
ASD diagnostic Classification[3]	Strict	ASD	Strict	Strict	Strict	Strict	ASD	Strict	Strict
ADI-R									
Age ADI-R (mo)	45	36	139	65	66	48	113	118	98
Social domain	18	24	29	25	13	11	9	20	14
Communication domain verbal/non verbal	16 (v)	10	16 (v)	10	15 (v)	13 (v)	16 (v)	19 (v)	9 (v)
Restricted, repetitive behavior domain	5	6	4	6	10	5	10	3	5
Abnormality of Development	5	2	4	5	3	5	4	3	3

[1]Samples labeled with an asterisk (*) are not in the NIMH Repository and Autism Genome Project (AGP) IDs are provided
[2]Ancestry based on eigenvector analyses determined to be European (EUR) or African-American (AA)
[3]Autism corresponds to an AGP classification of “strict” or “narrow” autism; ASD corresponds to an AGP “spectrum” classification

A. continued

	Ile22Thr	Phe48Ser	Ala69Ser	Leu294Phe	*319Gln	*319Gln
NIMH Sample ID[1]	217-14276-3990	211-5202-3	217-14216-3470	215-13218-2403	74-0580-03	215-13003-43
Discovery vs. Replication	R	R	R	R	R	R
Gender	M	F	M	M	M	M
Ancestry[2]	EUR	EUR	EUR	EUR	EUR	EUR
ASD diagnostic Classification[3]	Strict	Strict	Strict	Strict	Strict	Strict
ADI-R						
Age ADI-R (mo)	62	70	67	157	185	71
Social domain	25	24	18	26	29	20
Communication domain verbal/non verbal	10 (nv)	20 (v)	8 (nv)	21 (v)	16 (v)	18 (v)
Restricted, repetitive behavior domain	3	6	3	8	11	8
Abnormality of Development	3	2	3	3	3	3

B. Clinical history on the nuclear family of T356M SLC6A3 proband

Clinical Information: Proband (male) is the eldest of two children and heterozygous for the *SLC6A3 de novo* Thr356Met mutation described in Chapter 5.

Patient ID: AC04-0029-01 (proband)

At the time of testing, patient was a 66-69-month-old non-Hispanic male Caucasian diagnosed with autism (on both ADOS and ADIR). Maternal interview on pregnancy provided the following details: Mother experienced anemia during pregnancy (gestational age = 39 weeks). Labor was induced (use of Pitocin) because of failure to progress (Note: no C-section was done). Child was born 8 lbs 6 oz (134 ounces), with no other delivery issues.

Subset and composite scores on the Wechsler Preschool and Primary Scale of Intelligence (WPPSI) indicate normal intelligence. Full Scale IQ was Average (94), and the subsets of Verbal-IQ and Performance-IQ was scored as Average (101) and Low Average (86) respectively, indicating normal intelligence.

ADIR revealed deficits across all four subsets: (1) Reciprocal Social Interaction (score=13; cutoff=10); (2) Abnormalities in Communication (score=13; cutoff=8); (3) Restricted, Repetitive and Stereotyped Patterns of Behavior (score=4; cutoff=3); (4) Development Evident at or before 36 Months (score=1; cutoff=1).

Patient experienced a delay in speech requiring therapy. No history of seizures, gastrointestinal conditions, sleep deficits, and no diet restrictions. Currently taking multivitamins, with no use of other medications (besides amoxicillin, Tylenol, and Benadryl).

No family history for psychiatric illness requiring hospitalization. Known history of Asperger's Disorder (Mother's cousin's son), and depression (Father's paternal uncle).

Patient ID: AC04-0029-02 (father)

Father is an adult non-Hispanic Caucasian male. Age at conception of proband is 36. Slightly above normative range of intelligence (IQ=118; Wechsler Adult Intelligence Scale). No presence of broader autism phenotype, and no psychiatric medication use current or past. No co-morbid diagnoses. He holds a postgraduate degree and reports an annual household income of $81-101k.

Patient ID: AC04-0029-03 (mother)

Mother is an adult non-Hispanic Caucasian female. Age at conception of proband is 32. Above normative range of intelligence (IQ=131; Wechsler Adult Intelligence Scale); no presence of broader autism phenotype, and no psychiatric medication use current or past. No co-morbid diagnoses. She holds a postgraduate degree and reports an annual household income of $81-101k. No medication use reported for mother before, during, or after pregnancy except for epidural during labor.

Patient ID: AC04-0029-04 (sibling)

Sibling is a non-Hispanic Caucasian 3 year old (37 months) of unspecified sex. Normative intelligence (IQ=118; WPPSI). No behavioral problems reported. No medication use endorsed for current or past. No comorbid diagnoses.

C. Detailed validation results on putative differences in whole blood vs. cell line derived DNA

Group 0: Validation results

Indiv.	Chr.	Position	Blood Call	Validation	Cell Call	Validation	Comment
3	X	90691093	AAC[C/C]GCC	TRUE	AAC[C/T]GCC	TRUE	MOSAIC
6	X	123538932	CCT[C/C]TGC	TRUE	CCT[C/T]TGC	TRUE	MOSAIC
7	4	123533877	TTC[A/A]CTT	TRUE	TTC[A/G]CTT	TRUE	MOSAIC
11	2	228882979	ACC[G/G]TTG	TRUE	ACC[G/A]TTG	TRUE	MOSAIC
11	8	42256366	AAA[T/T]CTC	TRUE	AAA[T/A]CTC	TRUE	MOSAIC
12	6	131188570	GCT[C/C]GAA	TRUE	GCT[C/T]GAA	TRUE	MOSAIC
12	20	13846164	CTG[T/T]ACA	TRUE	CTG[T/G]ACA	TRUE	MOSAIC
16	1	167403251	AAC[A/A]CTC	TRUE	AAC[A/G]CTC	TRUE	MOSAIC
16	12	109684116	CTG[C/C]GGG	TRUE	CTG[C/T]GGG	TRUE	MOSAIC
16	18	58038892	TGC[C/C]GGG	TRUE	TGC[C/T]GGG	TRUE	MOSAIC

Group 1: Validation results

Indiv.	Chr.	Position	Blood Call	Validation	Cell Call	Validation	Comment
1	4	69180008	GAC[G/G]AAG	TRUE	GAC[G/A]AAG	TRUE	MOSAIC
4	1	10706375	AGG[C/C]AAG	TRUE	AGG[C/A]AAG	TRUE	MOSAIC
4	6	32629194	AAC[G/G]CCA		AAC[G/A]CCA		Impossible primer design.
4	6	32629199	CAC[T/T]CAG		CAC[T/C]CAG		Impossible primer design.
4	10	98078170	CTT[C/C]AAG	TRUE	CTT[C/T]AAG	FALSE	All “C” reads on forward strand.
4	11	47644317	CAT[C/C]CAC	TRUE	CAT[C/A]CAC	TRUE	MOSAIC
4	11	56143156	AGA[C/C]CAT	TRUE	AGA[C/T]CAT	FALSE	Possibly an alignment problem? Note next candidate is adjacent.
4	11	56143158	ACC[A/A]TCA	TRUE	ACC[A/G]TCA	FALSE	See previous.
4	12	57662804	CTG[C/C]TGG	TRUE	CTG[C/G]TGG	FALSE	All “G” reads on the forward strand.
4	12	77424135	TAA[C/T]GAA	FALSE	TAA[C/C]GAA	TRUE	Almost all reads on the reverse strand.
4	22	50553671	GAC[G/G]GAA	TRUE	GAC[G/T]GAA	FALSE	All “T” reads on the reverse strand.
4	X	55249176	GAA[G/G]CAG	TRUE	GAA[G/A]CAG	TRUE	MOSAIC
5	11	67378012	CCC[G/T]CCT	FALSE	CCC[G/G]CCT	TRUE	All “T” reads on reverse strand.
5	16	89246670	GGA[T/T]GAG	TRUE	GGA[T/C]GAG	FALSE	All but one “C” read is highly skewed.
6	X	99661793	TCG[G/G]CCA	TRUE	TCG[G/A]CCA	TRUE	MOSAIC
6	X	118767373	GGA[C/C]GAA	TRUE	GGA[C/T]GAA	TRUE	MOSAIC

7	12	54803111	GTG[C/C]GCC	TRUE	GTG[C/T]GCC	TRUE	MOSAIC
7	20	210435	ACT[C/C]CAA	TRUE	ACT[C/T]CAA	TRUE	MOSAIC
8	5	137801559	GAG[G/G]AGA	TRUE	GAG[G/C]AGA	TRUE	MOSAIC
8	9	119106793	CTC[C/C]TTC	TRUE	CTC[C/T]TTC	TRUE	MOSAIC
9	2	198299652	TTA[G/G]GCT	TRUE	TTA[G/C]GCT	TRUE	MOSAIC, but noisy background
9	11	134241679	AAT[C/C]GTG	TRUE	AAT[C/T]GTG	TRUE	MOSAIC
11	2	116599816	ATC[T/T]GAG	TRUE	ATC[T/G]GAG	TRUE	MOSAIC
11	7	72397315	ACA[G/G]CTA	TRUE	ACA[G/A]CTA	TRUE	MOSAIC (very small additional peak)
11	19	56300185	TTA[A/A]GTA	TRUE	TTA[A/G]GTA	TRUE	MOSAIC
12	11	27389932	TTT[C/C]GGG	TRUE	TTT[C/T]GGG	TRUE	MOSAIC
15	11	116797981	AGC[T/T]TTT	TRUE	AGC[T/C]TTT	TRUE	MOSAIC
16	11	18313224	GTC[G/G]TTC	TRUE	GTC[G/T]TTC	TRUE	MOSAIC

Group 2: Validation results

Indiv.	Chr.	Position	Blood Call	Validation	Cell Call	Validation	Comment
1	1	65313240	GGG[T/T]CTA	TRUE	GGG[T/G]CTA	FALSE	All “G” reads on the forward strand, and these reads are skewed.
1	3	122354873	ACA[C/C]GAT	TRUE	ACA[C/T]GAT	TRUE	MOSAIC
1	4	46264063	GTC[C/C]ATG	TRUE	GTC[C/A]ATG	FALSE	Large total depth, and very few “A” reads on cell line.
1	9	135982045	CTT[C/C]CCA	TRUE	CTT[C/T]CCA	FALSE	All “T” reads on the forward strand.
1	11	62378536	CCG[C/A]CAC	FALSE	CCG[C/C]CAC	TRUE	Blood “A” reads are skewed, and few.
1	12	132623810	GAG[C/A]CCC	FALSE	GAG[C/C]CCC	TRUE	Blood “A” reads are skiewed, and few.
1	17	78931437	GGT[G/G]GTG	TRUE	GGT[G/T]GTG	FALSE	The forward “T” reads are skewed to the left, the reverse “T” reads are skewed to the right.
1	19	15278133	TGG[T/T]GCC	TRUE	TGG[T/G]GCC	FALSE	All “G” reads are on the forward strand.
1	19	51958638	AGG[T/T]GTG	TRUE	AGG[T/G]GTG	FALSE	All “G” reads are on the forward strand.
1	20	3928845	GGG[A/A]AAA	TRUE	GGG[A/G]AAA	FALSE	All but one “G” read is on the forward strand, and are highly skewed.
1	20	61459302	CCC[G/G]GAG	TRUE	CCC[G/T]GAG	FALSE	The forward “T” reads are skewed to the left, the reverse “T” reads are skewed to the right.
1	Y	14958873	TCC[C/C]CAA	TRUE	TCC[C/T]CAA	TRUE	MOSAIC
2	7	44613258	CGA[G/G]CCA	TRUE	CGA[G/T]CCA	FALSE	All “T” reads are skewed, and there are few.
2	9	8484293	TAT[G/G]ATT	TRUE	TAT[G/T]ATT	FALSE	Low proportion of “T” reads on cell line, and almost all are skewed.

2	9	101498842	TGG[C/A]ACT	FALSE	TGG[C/C]ACT	TRUE	The forward “A” reads are skewed to the left, the reverse “A” reads are skewed to the right.
2	9	101498847	CTG[C/A]TCT	FALSE	CTG[C/C]TCT	TRUE	The forward “A” reads are skewed to the left, the reverse “A” reads are skewed to the right. Also, note proximity to previous.
2	18	63511013	AAC[C/C]GGT	TRUE	AAC[C/T]GGT	TRUE	MOSAIC
2	19	54725995	GCT[A/A]TAG	FALSE	GCT[G/A]TAG	TRUE	This appears to be a mosaic on both blood and cell line. Large imbalance in “A” versus “G” calls, but consistent between blood and cell line.
2	21	35169715	TGG[A/A]GCA	TRUE	TGG[A/G]GCA	FALSE	All “G” reads are on the forward strand.
2	X	10107414	AAT[C/C]AAT	TRUE	AAT[C/A]AAT	FALSE	All “A” reads are on the forward strand.
3	1	2441539	TGA[G/G]GGA	TRUE	TGA[G/A]GGA	FALSE	
3	1	27995021	GTC[C/A]GAG	FALSE	GTC[C/C]GAG	TRUE	
3	15	75970085	GCA[G/T]CCT	FALSE	GCA[G/G]CCT	TRUE	
3	17	33459447	CTT[C/C]GGG	TRUE	CTT[C/T]GGG	FALSE	
3	19	55148043	GTC[T/T]GGG	TRUE	GTC[T/C]GGG	FALSE	
3	21	33077817	TTA[C/C]TTA	TRUE	TTA[C/T]TTA	FALSE	
4	1	16748468	CTC[T/T]GGA	TRUE	CTC[T/G]GGA	FALSE	
4	1	35870652	TGG[A/A]GCC	TRUE	TGG[A/G]GCC	FALSE	
4	1	46499441	TCC[A/A]GCT	TRUE	TCC[A/C]GCT	FALSE	
4	1	57398975	CTT[C/C]CAC	TRUE	CTT[C/T]CAC	FALSE	
4	1	94461686	TCC[A/A]GCA	TRUE	TCC[A/C]GCA	FALSE	
4	2	11780500	CAG[T/T]TCC	TRUE	CAG[T/G]TCC	FALSE	
4	2	242204025	TGG[A/A]CCA	TRUE	TGG[A/C]CCA	FALSE	
4	5	43173577	GTT[C/C]TTA	TRUE	GTT[C/T]TTA	FALSE	
4	6	11735807	CTG[A/A]TGT	TRUE	CTG[A/G]TGT	FALSE	
4	6	32632795	CTC[C/C]GTC	TRUE	CTC[C/T]GTC	FALSE	
4	6	32632850	GCG[G/G]GGA	TRUE	GCG[G/A]GGA	FALSE	
4	6	32634318	TCG[C/C]CAG		TCG[C/A]CAG		Failure
4	7	6461451	CGT[T/T]GCA	TRUE	CGT[T/G]GCA	FALSE	
4	7	44151241	CTG[T/T]GGG	TRUE	CTG[T/G]GGG	FALSE	
4	8	2020398	GCC[C/C]TGA	TRUE	GCC[C/A]TGA	FALSE	
4	9	100315518	CTG[T/T]TCC	TRUE	CTG[T/G]TCC	FALSE	

4	10	94822655	TGG[A/A]CGA	TRUE	TGG[A/G]CGA	FALSE	
4	10	97397146	TCC[A/A]CTG	TRUE	TCC[A/C]CTG	FALSE	
4	10	103772671	CAG[T/T]AGC	TRUE	CAG[T/C]AGC	TRUE	MOSAIC
4	11	68131385	AGG[T/T]GCT	TRUE	AGG[T/G]GCT	FALSE	
4	11	116706900	TCC[A/A]CCT	TRUE	TCC[A/C]CCT	FALSE	
4	13	21063460	GGG[A/A]GAC	TRUE	GGG[A/G]GAC	FALSE	
4	14	23564357	CTT[C/C]GGG	TRUE	CTT[C/T]GGG	FALSE	
4	14	23867949	CAG[T/T]TGC	TRUE	CAG[T/G]TGC	FALSE	
4	14	54416937	TGG[T/T]CAG	TRUE	TGG[T/C]CAG	FALSE	
4	14	57085352	CCG[T/T]CCT	TRUE	CCG[T/G]CCT	FALSE	
4	16	2498902	CTC[A/A]CGG	TRUE	CTC[A/C]CGG	FALSE	
4	16	4445265	CCC[A/A]CCC	TRUE	CCC[A/C]CCC	FALSE	
4	17	7707597	AGG[A/A]GCT	TRUE	AGG[A/G]GCT	FALSE	
4	17	21319369	ACC[G/G]AGG	TRUE	ACC[G/A]AGG	FALSE	
4	19	313490	CCC[A/A]CCT	TRUE	CCC[A/C]CCT	FALSE	
4	19	44252108	CCC[T/T]CCT	TRUE	CCC[T/C]CCT	FALSE	
4	19	50322605	GTG[A/A]ATG	TRUE	GTG[A/C]ATG	FALSE	
4	20	47592567	CTC[A/A]CGG	TRUE	CTC[A/C]CGG	FALSE	
4	21	46640865	AGC[T/T]GGG	TRUE	AGC[T/G]GGG	FALSE	
5	3	179527471	AAA[G/A]AAG	TRUE	AAA[G/G]AAG	FALSE	
5	5	140079996	TCC[A/C]AGC	FALSE	TCC[A/A]AGC	TRUE	
5	5	148712424	CAA[G/G]CCC	TRUE	CAA[G/T]CCC	FALSE	
5	6	7405175	TAT[C/C]ATT	TRUE	TAT[C/A]ATT	FALSE	
5	7	66648157	GAA[C/T]GTG	TRUE	GAA[T/T]GTG	FALSE	
5	11	117789342	GCC[T/C]GGG	TRUE	GCC[T/T]GGG	FALSE	
5	11	117789345	TGG[G/C]CTG	TRUE	TGG[G/G]CTG	FALSE	
5	15	43712957	GGA[G/A]AAG	FALSE	GGA[G/G]AAG	TRUE	
5	20	62658494	CGG[G/A]AAG	FALSE	CGG[G/G]AAG	TRUE	
6	6	32632598	CGG[A/T]ACG		CGG[A/A]ACG		Failure
6	16	135354	GAG[G/T]ACA	FALSE	GAG[G/G]ACA	TRUE	
6	22	24583727	GTG[G/C]CAG	FALSE	GTG[G/G]CAG	TRUE	
7	9	4663267	CAC[C/C]ATC	TRUE	CAC[C/A]ATC	TRUE	MOSAIC
7	9	72006567	CTC[G/G]CGG	TRUE	CTC[G/A]CGG	TRUE	MOSAIC
7	9	87425462	TTT[G/G]TTT	TRUE	TTT[G/A]TTT	TRUE	MOSAIC

7	10	15138591	CGA[C/C]TTT	TRUE	CGA[C/T]TTT	TRUE	MOSAIC
7	22	42912029	CTT[G/G]GGA	TRUE	CTT[G/T]GGA	FALSE	
8	1	54360125	AGC[G/G]ACT	TRUE	AGC[G/A]ACT	TRUE	MOSAIC
8	11	63672427	GTG[A/C]CCC	TRUE	GTG[A/A]CCC	FALSE	
8	20	10036141	TTT[C/T]TTT	TRUE	TTT[C/C]TTT	FALSE	
9	2	130912789	CAA[A/A]CAA	TRUE	CAA[A/C]CAA	FALSE	
9	5	840785	GAG[C/A]CAC	TRUE	GAG[C/C]CAC	FALSE	
9	6	75823294	CAT[G/G]TGC	TRUE	CAT[G/A]TGC	TRUE	MOSAIC
9	11	2416725	AGG[C/C]TGT	TRUE	AGG[C/A]TGT	FALSE	
9	15	48459567	ATC[C/C]TTA	TRUE	ATC[C/T]TTA	TRUE	MOSAIC
10	1	68942568	CTT[C/T]ATT	TRUE	CTT[C/C]ATT	FALSE	
10	5	60241142	GAA[A/A]GAG	TRUE	GAA[G/A]GAG	FALSE	
10	7	128483043	TGC[A/C]CCT	TRUE	TGC[A/A]CCT	FALSE	
10	10	3174669	CTC[A/A]CCG		CTC[A/C]CCG		Failure
10	19	46375892	GCC[A/A]GAA	TRUE	GCC[A/G]GAA	FALSE	
11	1	12921127	ATT[A/G]GAA	TRUE	ATT[A/A]GAA	FALSE	All G calls are in a single direction
11	2	139318349	AAC[A/A]TTT	TRUE	AAC[A/T]TTT	FALSE	All T calls are in a single direction
11	3	48464209	AAG[C/C]GAT	TRUE	AAG[C/T]GAT	TRUE	MOSAIC
12	3	100091564	GCG[G/T]TAC	FALSE	GCG[G/G]TAC	TRUE	All T calls are in a single direction
12	4	155506864	CAC[G/G]GGA	TRUE	CAC[G/A]GGA	TRUE	MOSAIC
12	10	117825111	GCC[C/C]GAC	TRUE	GCC[C/T]GAC	TRUE	All T calls are in a single direction
12	19	33467325	TTC[A/A]TTG	TRUE	TTC[A/G]TTG	TRUE	MOSAIC
13	17	21318770	CGC[G/A]GCG	FALSE	CGC[G/G]GCG	TRUE	All A calls are in a single direction
13	17	21318773	GGC[G/A]CAG	FALSE	GGC[G/G]CAG	TRUE	All A calls are in a single direction
14	7	26251843	GTT[C/T]TTT	FALSE	GTT[C/C]TTT	TRUE	Very slight bump for T on blood
14	15	20588693	CAA[C/C]ATC	FALSE	CAA[C/T]ATC	TRUE	Appears to be mosaic on both
15	14	96180232	AAG[C/C]TCC	TRUE	AAG[C/T]TCC	FALSE	All T calls are in a single direction
16	3	48505302	GGC[A/A]TCT	TRUE	GGC[A/G]TCT	FALSE	No clear explanation for mistake
16	6	151334926	TGG[T/T]GTT	TRUE	TGG[T/G]GTT	FALSE	All G calls are in a single direction
16	11	19251484	TGG[T/T]TTG	TRUE	TGG[T/G]TTG	TRUE	MOSAIC
16	14	81223208	ACA[A/A]TAA	TRUE	ACA[A/C]TAA	FALSE	All C calls are in a single direction

Group 3: Validation results

Indiv.	Chr.	Position	Blood Call	Validation	Cell Call	Validation	Comment
1	7	44099733	TGC[C/C]CTG	TRUE	TGC[C/A]CTG	FALSE	Strong Candidate
1	20	1895794	CGT[G/G]TTG		CGT[G/A]TTG		Strong Candidate; Failure
1	20	1895797	GTT[G/G]GTT		GTT[G/A]GTT		Strong Candidate; Failure
2	1	185137463	TAT[T/T]TAT		TAT[T/A]TAT		Strong Canddiate; Failure
2	3	16264288	AGA[G/G]CCT	TRUE	AGA[G/T]CCT	FALSE	Strong Candidate
2	7	912157	CAC[G/G]TAT	TRUE	CAC[G/T]TAT	FALSE	Strong Candidate
2	8	73480055	TCC[G/G]GAG	TRUE	TCC[G/A]GAG	FALSE	Strong Candidate
3	1	240371031	AGC[A/A]GGA		AGC[A/G]GGA		Strong Candidate; Failure
4	5	115230799	ACA[T/T]TAG	TRUE	ACA[T/A]TAG	FALSE	Strong Candidate
4	5	115230800	CAT[T/T]AGA	TRUE	CAT[T/C]AGA	FALSE	Strong Candidate
4	6	32712996	AGT[A/A]CAC	TRUE	AGT[A/T]CAC	FALSE	Strong Candidate
4	X	154721357	AAA[C/C]AGA		AAA[G/C]AGA		Strong Candidate; Failure
5	6	148861617	CTC[G/G]TTG	TRUE	CTC[G/A]TTG	TRUE	Strong Candidate; MOSAIC
6	19	51919894	CAC[A/A]CCC	TRUE	CAC[A/G]CCC	FALSE	Strong Candidate
7	19	54942009	CAG[C/C]GAG		CAG[C/A]GAG		Strong Candidate; Failure
8	12	51740409	CCA[T/T]AAA		CCA[T/G]AAA		Strong Candidate; Failure
8	12	51740410	CAT[A/A]AAG		CAT[A/G]AAG		Strong Candidate; Failure
9	6	32609368	CTT[A/A]AGT		CTT[A/G]AGT		Strong Candidate; Failure
9	8	28651299	CAA[A/A]AAA	TRUE	CAA[A/C]AAA	FALSE	Strong Candidate
11	2	27840357	TGG[A/G]AAA	FALSE	TGG[A/A]AAA	TRUE	
11	3	10114944	GAA[A/C]CAA	FALSE	GAA[A/A]CAA	TRUE	
11	5	180335598	AGG[T/T]GTT	TRUE	AGG[T/G]GTT	FALSE	
12	4	72319317	CAT[T/T]TAT	TRUE	CAT[T/G]TAT	TRUE	Strong Candidate; MOSAIC
12	16	88926388	GCC[C/C]TAC	TRUE	GCC[C/T]TAC	FALSE	
16	1	15808767	TCC[G/G]GGA		TCC[G/A]GGA		Strong Candidate; Failure
16	1	202743892	ATC[T/T]CAT	TRUE	ATC[T/C]CAT	FALSE	
16	5	120022221	GAT[T/T]CCT	TRUE	GAT[T/C]CCT	TRUE	Strong Candidate; MOSAIC
16	7	99361466	GTA[T/T]CAT	TRUE	GTA[C/T]CAT	FALSE	
16	8	7718187	AAT[T/G]AGA	TRUE	AAT[G/G]AGA	FALSE	
16	8	87229948	CGA[A/A]GAT		CGA[T/A]GAT		Strong Candidate; Failure
16	9	135153668	CAA[T/T]AGT	TRUE	CAA[C/T]AGT	FALSE	

16	19	52249211	AAC[G/G]CCA	TRUE	AAC[T/G]CCA	FALSE	Strong Candidate
16	21	15516948	TTT[T/T]TGG	TRUE	TTT[C/T]TGG	FALSE	

REFERENCES

1. Kanner L: **Autistic disturbances of affective contact.** *Acta Paedopsychiatr* 1968, **35:**100-136.
2. 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.** *J Autism Dev Disord* 1994, **24:**659-685.
3. Lord C, Risi S, Lambrecht L, Cook EH, Jr., Leventhal BL, DiLavore PC, Pickles A, Rutter M: **The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism.** *J Autism Dev Disord* 2000, **30:**205-223.
4. Wall DP, Kosmicki J, Deluca TF, Harstad E, Fusaro VA: **Use of machine learning to shorten observation-based screening and diagnosis of autism.** *Transl Psychiatry* 2012, **2:**e100.
5. Wall DP, Dally R, Luyster R, Jung JY, Deluca TF: **Use of artificial intelligence to shorten the behavioral diagnosis of autism.** *PLoS One* 2012, **7:**e43855.
6. Reaven JA, Hepburn SL, Ross RG: **Use of the ADOS and ADI-R in children with psychosis: importance of clinical judgment.** *Clin Child Psychol Psychiatry* 2008, **13:**81-94.
7. Geschwind DH: **Advances in autism.** *Annu Rev Med* 2009, **60:**367-380.
8. Leyfer OT, Folstein SE, Bacalman S, Davis NO, Dinh E, Morgan J, Tager-Flusberg H, Lainhart JE: **Comorbid psychiatric disorders in children with autism: interview development and rates of disorders.** *J Autism Dev Disord* 2006, **36:**849-861.
9. Hartley SL, Sikora DM, McCoy R: **Prevalence and risk factors of maladaptive behaviour in young children with Autistic Disorder.** *J Intellect Disabil Res* 2008, **52:**819-829.
10. Simonoff E, Pickles A, Charman T, Chandler S, Loucas T, Baird G: **Psychiatric disorders in children with autism spectrum disorders: prevalence, comorbidity, and associated factors in a population-derived sample.** *Journal of the American Academy of Child and Adolescent Psychiatry* 2008, **47:**921-929.

11. Yeargin-Allsopp M, Rice C, Karapurkar T, Doernberg N, Boyle C, Murphy C: **Prevalence of autism in a US metropolitan area.** *JAMA* 2003, **289:**49-55.
12. Chakrabarti S, Fombonne E: **Pervasive developmental disorders in preschool children.** *JAMA* 2001, **285:**3093-3099.
13. Ganz ML: **The lifetime distribution of the incremental societal costs of autism.** *Arch Pediatr Adolesc Med* 2007, **161:**343-349.
14. Lovaas OI: **Behavioral treatment and normal educational and intellectual functioning in young autistic children.** *J Consult Clin Psychol* 1987, **55:**3-9.
15. Dawson G, Rogers S, Munson J, Smith M, Winter J, Greenson J, Donaldson A, Varley J: **Randomized, controlled trial of an intervention for toddlers with autism: the Early Start Denver Model.** *Pediatrics* 2010, **125:**e17-23.
16. Smith T, Groen AD, Wynn JW: **Randomized trial of intensive early intervention for children with pervasive developmental disorder.** *Am J Ment Retard* 2000, **105:**269-285.
17. Cohen H, Amerine-Dickens M, Smith T: **Early intensive behavioral treatment: replication of the UCLA model in a community setting.** *J Dev Behav Pediatr* 2006, **27:**S145-155.
18. Remington B, Hastings RP, Kovshoff H, degli Espinosa F, Jahr E, Brown T, Alsford P, Lemaic M, Ward N: **Early intensive behavioral intervention: outcomes for children with autism and their parents after two years.** *Am J Ment Retard* 2007, **112:**418-438.
19. Mandell DS, Morales KH, Marcus SC, Stahmer AC, Doshi J, Polsky DE: **Psychotropic medication use among Medicaid-enrolled children with autism spectrum disorders.** *Pediatrics* 2008, **121:**e441-448.
20. Rosenberg RE, Mandell DS, Farmer JE, Law JK, Marvin AR, Law PA: **Psychotropic medication use among children with autism spectrum disorders enrolled in a national registry, 2007-2008.** *J Autism Dev Disord* 2010, **40:**342-351.

21. Dove D, Warren Z, McPheeters ML, Taylor JL, Sathe NA, Veenstra-VanderWeele J: **Medications for adolescents and young adults with autism spectrum disorders: a systematic review.** *Pediatrics* 2012, **130:**717-726.

22. Taylor JL, McPheeters ML, Sathe NA, Dove D, Veenstra-Vanderweele J, Warren Z: **A systematic review of vocational interventions for young adults with autism spectrum disorders.** *Pediatrics* 2012, **130:**531-538.

23. McDougle CJ, Holmes JP, Carlson DC, Pelton GH, Cohen DJ, Price LH: **A double-blind, placebo-controlled study of risperidone in adults with autistic disorder and other pervasive developmental disorders.** *Arch Gen Psychiatry* 1998, **55:**633-641.

24. McDougle CJ, Brodkin ES, Naylor ST, Carlson DC, Cohen DJ, Price LH: **Sertraline in adults with pervasive developmental disorders: a prospective open-label investigation.** *J Clin Psychopharmacol* 1998, **18:**62-66.

25. Brodkin ES, McDougle CJ, Naylor ST, Cohen DJ, Price LH: **Clomipramine in adults with pervasive developmental disorders: a prospective open-label investigation.** *J Child Adolesc Psychopharmacol* 1997, **7:**109-121.

26. Cook EH, Jr., Rowlett R, Jaselskis C, Leventhal BL: **Fluoxetine treatment of children and adults with autistic disorder and mental retardation.** *Journal of the American Academy of Child and Adolescent Psychiatry* 1992, **31:**739-745.

27. Volkmar FR: *Autism and pervasive developmental disorders.* Cambridge ; New York, NY: Cambridge University Press; 1998.

28. Kanner L, Eisenberg L: **Review of psychiatric progress 1954; child psychiatry and mental deficiency.** *Am J Psychiatry* 1955, **111:**520-523.

29. Eisenberg L, Kanner L: **Childhood schizophrenia; symposium, 1955. VI. Early infantile autism, 1943-55.** *Am J Orthopsychiatry* 1956, **26:**556-566.

30. Kanner L: **The birth of early infantile autism.** *J Autism Child Schizophr* 1973, **3:**93-95.

31. Bender L: **Childhood schizophrenia; clinical study on one hundred schizophrenic children.** *Am J Orthopsychiatry* 1947, **17:**40-56.

32. Kolvin I: **Studies in the childhood psychoses. I. Diagnostic criteria and classification.** *Br J Psychiatry* 1971, **118:**381-384.
33. Chess S: **Autism in children with congenital rubella.** *J Autism Child Schizophr* 1971, **1:**33-47.
34. Devlin B, Scherer SW: **Genetic architecture in autism spectrum disorder.** *Curr Opin Genet Dev* 2012, **22:**229-237.
35. Rutter M, Bartak L: **Causes of infantile autism: some considerations from recent research.** *J Autism Child Schizophr* 1971, **1:**20-32.
36. Folstein S, Rutter M: **Genetic influences and infantile autism.** *Nature* 1977, **265:**726-728.
37. Ozonoff S: **Editorial perspective: Autism spectrum disorders in DSM-5--an historical perspective and the need for change.** *J Child Psychol Psychiatry* 2012, **53:**1092-1094.
38. Autism, Developmental Disabilities Monitoring Network Surveillance Year Principal I, Centers for Disease C, Prevention: **Prevalence of autism spectrum disorders--Autism and Developmental Disabilities Monitoring Network, 14 sites, United States, 2008.** *MMWR Surveill Summ* 2012, **61:**1-19.
39. Newschaffer CJ, Croen LA, Daniels J, Giarelli E, Grether JK, Levy SE, Mandell DS, Miller LA, Pinto-Martin J, Reaven J, et al: **The epidemiology of autism spectrum disorders.** *Annu Rev Public Health* 2007, **28:**235-258.
40. Kanner L: **Infantile autism and the schizophrenias.** *Behav Sci* 1965, **10:**412-420.
41. Wing L, Potter D: **The epidemiology of autistic spectrum disorders: is the prevalence rising?** *Ment Retard Dev Disabil Res Rev* 2002, **8:**151-161.
42. Gillberg C, Steffenburg S, Schaumann H: **Is autism more common now than ten years ago?** *Br J Psychiatry* 1991, **158:**403-409.
43. Bhasin TK, Brocksen S, Avchen RN, Van Naarden Braun K: **Prevalence of four developmental disabilities among children aged 8 years--Metropolitan Atlanta Developmental Disabilities Surveillance Program, 1996 and 2000.** *MMWR Surveill Summ* 2006, **55:**1-9.

44. Piven J, Palmer P, Jacobi D, Childress D, Arndt S: **Broader autism phenotype: evidence from a family history study of multiple-incidence autism families.** *Am J Psychiatry* 1997, **154:**185-190.

45. Werling DM, Geschwind DH: **Sex differences in autism spectrum disorders.** *Curr Opin Neurol* 2013, **26:**146-153.

46. Patterson PH: **Maternal infection and autism.** *Brain, behavior, and immunity* 2012, **26:**393.

47. Chaste P, Leboyer M: **Autism risk factors: genes, environment, and gene-environment interactions.** *Dialogues Clin Neurosci* 2012, **14:**281-292.

48. Gardener H, Spiegelman D, Buka SL: **Prenatal risk factors for autism: comprehensive meta-analysis.** *Br J Psychiatry* 2009, **195:**7-14.

49. Guinchat V, Thorsen P, Laurent C, Cans C, Bodeau N, Cohen D: **Pre-, peri- and neonatal risk factors for autism.** *Acta Obstet Gynecol Scand* 2012, **91:**287-300.

50. Wakefield AJ, Murch SH, Anthony A, Linnell J, Casson DM, Malik M, Berelowitz M, Dhillon AP, Thomson MA, Harvey P, et al: **Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children.** *Lancet* 1998, **351:**637-641.

51. Eggertson L: **Lancet retracts 12-year-old article linking autism to MMR vaccines.** *CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne* 2010, **182:**E199-200.

52. Parker SK, Schwartz B, Todd J, Pickering LK: **Thimerosal-containing vaccines and autistic spectrum disorder: a critical review of published original data.** *Pediatrics* 2004, **114:**793-804.

53. Hviid A, Melbye M, Pasternak B: **Use of selective serotonin reuptake inhibitors during pregnancy and risk of autism.** *The New England journal of medicine* 2013, **369:**2406-2415.

54. Croen LA, Grether JK, Yoshida CK, Odouli R, Hendrick V: **Antidepressant use during pregnancy and childhood autism spectrum disorders.** *Arch Gen Psychiatry* 2011, **68:**1104-1112.

55. Wing L: **Childhood autism and social class: a question of selection?** *Br J Psychiatry* 1980, **137:**410-417.

56. Bhasin TK, Schendel D: **Sociodemographic risk factors for autism in a US metropolitan area.** *J Autism Dev Disord* 2007, **37:**667-677.

57. Larsson HJ, Eaton WW, Madsen KM, Vestergaard M, Olesen AV, Agerbo E, Schendel D, Thorsen P, Mortensen PB: **Risk factors for autism: perinatal factors, parental psychiatric history, and socioeconomic status.** *Am J Epidemiol* 2005, **161:**916-925; discussion 926-918.

58. Betancur C: **Etiological heterogeneity in autism spectrum disorders: more than 100 genetic and genomic disorders and still counting.** *Brain Res* 2011, **1380:**42-77.

59. Folstein S, Rutter M: **Infantile autism: a genetic study of 21 twin pairs.** *J Child Psychol Psychiatry* 1977, **18:**297-321.

60. Bailey A, Le Couteur A, Gottesman I, Bolton P, Simonoff E, Yuzda E, Rutter M: **Autism as a strongly genetic disorder: evidence from a British twin study.** *Psychol Med* 1995, **25:**63-77.

61. Ronald A, Hoekstra RA: **Autism spectrum disorders and autistic traits: a decade of new twin studies.** *Am J Med Genet B Neuropsychiatr Genet* 2011, **156B:**255-274.

62. Jorde LB, Hasstedt SJ, Ritvo ER, Mason-Brothers A, Freeman BJ, Pingree C, McMahon WM, Petersen B, Jenson WR, Mo A: **Complex segregation analysis of autism.** *Am J Hum Genet* 1991, **49:**932-938.

63. Ozonoff S, Young GS, Carter A, Messinger D, Yirmiya N, Zwaigenbaum L, Bryson S, Carver LJ, Constantino JN, Dobkins K, et al: **Recurrence risk for autism spectrum disorders: a Baby Siblings Research Consortium study.** *Pediatrics* 2011, **128:**e488-495.

64. Hallmayer J, Cleveland S, Torres A, Phillips J, Cohen B, Torigoe T, Miller J, Fedele A, Collins J, Smith K, et al: **Genetic heritability and shared environmental factors among twin pairs with autism.** *Arch Gen Psychiatry* 2011, **68:**1095-1102.

65. Rosenberg RE, Law JK, Yenokyan G, McGready J, Kaufmann WE, Law PA: **Characteristics and concordance of autism spectrum disorders among 277 twin pairs.** *Arch Pediatr Adolesc Med* 2009, **163:**907-914.

66. Constantino JN, Todorov A, Hilton C, Law P, Zhang Y, Molloy E, Fitzgerald R, Geschwind D: **Autism recurrence in half siblings: strong support for genetic mechanisms of transmission in ASD.** *Molecular psychiatry* 2013, **18:**137-138.

67. Folstein SE, Rosen-Sheidley B: **Genetics of autism: complex aetiology for a heterogeneous disorder.** *Nat Rev Genet* 2001, **2:**943-955.

68. O'Roak BJ, State MW: **Autism genetics: strategies, challenges, and opportunities.** *Autism Res* 2008, **1:**4-17.

69. Stein JL, Parikshak NN, Geschwind DH: **Rare inherited variation in autism: beginning to see the forest and a few trees.** *Neuron* 2013, **77:**209-211.

70. Abrahams BS, Geschwind DH: **Connecting genes to brain in the autism spectrum disorders.** *Arch Neurol* 2010, **67:**395-399.

71. Berg JM, Geschwind DH: **Autism genetics: searching for specificity and convergence.** *Genome Biol* 2012, **13:**247.

72. Kraft P: **Curses--winner's and otherwise--in genetic epidemiology.** *Epidemiology* 2008, **19:**649-651; discussion 657-648.

73. Schork NJ, Murray SS, Frazer KA, Topol EJ: **Common vs. rare allele hypotheses for complex diseases.** *Curr Opin Genet Dev* 2009, **19:**212-219.

74. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, et al: **Finding the missing heritability of complex diseases.** *Nature* 2009, **461:**747-753.

75. Wang K, Zhang H, Ma D, Bucan M, Glessner JT, Abrahams BS, Salyakina D, Imielinski M, Bradfield JP, Sleiman PM, et al: **Common genetic variants on 5p14.1 associate with autism spectrum disorders.** *Nature* 2009, **459:**528-533.

76. Weiss LA, Arking DE, Gene Discovery Project of Johns H, the Autism C, Daly MJ, Chakravarti A: **A genome-wide linkage and association scan reveals novel loci for autism.** *Nature* 2009, **461:**802-808.

77. Anney R, Klei L, Pinto D, Regan R, Conroy J, Magalhaes TR, Correia C, Abrahams BS, Sykes N, Pagnamenta AT, et al: **A genome-wide scan for common alleles affecting risk for autism.** *Hum Mol Genet* 2010, **19:**4072-4082.

78. Devlin B, Melhem N, Roeder K: **Do common variants play a role in risk for autism? Evidence and theoretical musings.** *Brain Res* 2011, **1380:**78-84.

79. Klei L, Sanders SJ, Murtha MT, Hus V, Lowe JK, Willsey AJ, Moreno-De-Luca D, Yu TW, Fombonne E, Geschwind D, et al: **Common genetic variants, acting additively, are a major source of risk for autism.** *Mol Autism* 2012, **3:**9.

80. Devlin B, Melhem N, Roeder K: **Do common variants play a role in risk for autism? Evidence and theoretical musings.** *Brain Res* 2010.

81. Sutcliffe JS, Delahanty RJ, Prasad HC, McCauley JL, Han Q, Jiang L, Li C, Folstein SE, Blakely RD: **Allelic heterogeneity at the serotonin transporter locus (SLC6A4) confers susceptibility to autism and rigid-compulsive behaviors.** *Am J Hum Genet* 2005, **77:**265-279.

82. McCauley JL, Olson LM, Dowd M, Amin T, Steele A, Blakely RD, Folstein SE, Haines JL, Sutcliffe JS: **Linkage and association analysis at the serotonin transporter (SLC6A4) locus in a rigid-compulsive subset of autism.** *Am J Med Genet B Neuropsychiatr Genet* 2004, **127B:**104-112.

83. Coutinho AM, Sousa I, Martins M, Correia C, Morgadinho T, Bento C, Marques C, Ataide A, Miguel TS, Moore JH, et al: **Evidence for epistasis between SLC6A4 and ITGB3 in autism etiology and in the determination of platelet serotonin levels.** *Hum Genet* 2007, **121:**243-256.

84. Campbell DB, Sutcliffe JS, Ebert PJ, Militerni R, Bravaccio C, Trillo S, Elia M, Schneider C, Melmed R, Sacco R, et al: **A genetic variant that disrupts MET transcription is associated with autism.** *Proceedings of the National Academy of Sciences of the United States of America* 2006, **103:**16834-16839.

85. Campbell DB, Li C, Sutcliffe JS, Persico AM, Levitt P: **Genetic evidence implicating multiple genes in the MET receptor tyrosine kinase pathway in autism spectrum disorder.** *Autism Res* 2008, **1:**159-168.

86. Campbell DB, Warren D, Sutcliffe JS, Lee EB, Levitt P: **Association of MET with social and communication phenotypes in individuals with autism spectrum disorder.** *Am J Med Genet B Neuropsychiatr Genet* 2010, **153B:**438-446.

87. McCauley JL, Olson LM, Delahanty R, Amin T, Nurmi EL, Organ EL, Jacobs MM, Folstein SE, Haines JL, Sutcliffe JS: **A linkage disequilibrium map of the 1-Mb 15q12 GABA(A) receptor subunit cluster and association to autism.** *Am J Med Genet B Neuropsychiatr Genet* 2004, **131B:**51-59.

88. Delahanty RJ, Kang JQ, Brune CW, Kistner EO, Courchesne E, Cox NJ, Cook EH, Jr., Macdonald RL, Sutcliffe JS: **Maternal transmission of a rare GABRB3 signal peptide variant is associated with autism.** *Molecular psychiatry* 2011, **16:**86-96.

89. Buxbaum JD, Silverman JM, Smith CJ, Greenberg DA, Kilifarski M, Reichert J, Cook EH, Jr., Fang Y, Song CY, Vitale R: **Association between a GABRB3 polymorphism and autism.** *Molecular psychiatry* 2002, **7:**311-316.

90. Cook EH, Jr., Courchesne RY, Cox NJ, Lord C, Gonen D, Guter SJ, Lincoln A, Nix K, Haas R, Leventhal BL, Courchesne E: **Linkage-disequilibrium mapping of autistic disorder, with 15q11-13 markers.** *Am J Hum Genet* 1998, **62:**1077-1083.

91. Arking DE, Cutler DJ, Brune CW, Teslovich TM, West K, Ikeda M, Rea A, Guy M, Lin S, Cook EH, Chakravarti A: **A common genetic variant in the neurexin superfamily member CNTNAP2 increases familial risk of autism.** *Am J Hum Genet* 2008, **82:**160-164.

92. Anney R, Klei L, Pinto D, Almeida J, Bacchelli E, Baird G, Bolshakova N, Bolte S, Bolton PF, Bourgeron T, et al: **Individual common variants exert weak effects on the risk for autism spectrum disorderspi.** *Hum Mol Genet* 2012, **21:**4781-4792.

93. Weiss LA, Kosova G, Delahanty RJ, Jiang L, Cook EH, Ober C, Sutcliffe JS: **Variation in ITGB3 is associated with whole-blood serotonin level and autism susceptibility.** *Eur J Hum Genet* 2006, **14:**923-931.

94. Napolioni V, Lombardi F, Sacco R, Curatolo P, Manzi B, Alessandrelli R, Militerni R, Bravaccio C, Lenti C, Saccani M, et al: **Family-based association study of ITGB3**

in autism spectrum disorder and its endophenotypes. *Eur J Hum Genet* 2011, **19:**353-359.

95. Persico AM, D'Agruma L, Maiorano N, Totaro A, Militerni R, Bravaccio C, Wassink TH, Schneider C, Melmed R, Trillo S, et al: **Reelin gene alleles and haplotypes as a factor predisposing to autistic disorder.** *Molecular psychiatry* 2001, **6:**150-159.
96. Holt R, Barnby G, Maestrini E, Bacchelli E, Brocklebank D, Sousa I, Mulder EJ, Kantojarvi K, Jarvela I, Klauck SM, et al: **Linkage and candidate gene studies of autism spectrum disorders in European populations.** *Eur J Hum Genet* 2010, **18:**1013-1019.
97. Skaar DA, Shao Y, Haines JL, Stenger JE, Jaworski J, Martin ER, DeLong GR, Moore JH, McCauley JL, Sutcliffe JS, et al: **Analysis of the RELN gene as a genetic risk factor for autism.** *Molecular psychiatry* 2005, **10:**563-571.
98. Jamain S, Quach H, Betancur C, Rastam M, Colineaux C, Gillberg IC, Soderstrom H, Giros B, Leboyer M, Gillberg C, et al: **Mutations of the X-linked genes encoding neuroligins NLGN3 and NLGN4 are associated with autism.** *Nat Genet* 2003, **34:**27-29.
99. Autism Genome Project C, Szatmari P, Paterson AD, Zwaigenbaum L, Roberts W, Brian J, Liu XQ, Vincent JB, Skaug JL, Thompson AP, et al: **Mapping autism risk loci using genetic linkage and chromosomal rearrangements.** *Nat Genet* 2007, **39:**319-328.
100. Kim HG, Kishikawa S, Higgins AW, Seong IS, Donovan DJ, Shen Y, Lally E, Weiss LA, Najm J, Kutsche K, et al: **Disruption of neurexin 1 associated with autism spectrum disorder.** *Am J Hum Genet* 2008, **82:**199-207.
101. Pinto D, Pagnamenta AT, Klei L, Anney R, Merico D, Regan R, Conroy J, Magalhaes TR, Correia C, Abrahams BS, et al: **Functional impact of global rare copy number variation in autism spectrum disorders.** *Nature* 2010, **466:**368-372.
102. Durand CM, Betancur C, Boeckers TM, Bockmann J, Chaste P, Fauchereau F, Nygren G, Rastam M, Gillberg IC, Anckarsater H, et al: **Mutations in the gene**

encoding the synaptic scaffolding protein SHANK3 are associated with autism spectrum disorders. *Nat Genet* 2007, **39:**25-27.

103. Moessner R, Marshall CR, Sutcliffe JS, Skaug J, Pinto D, Vincent J, Zwaigenbaum L, Fernandez B, Roberts W, Szatmari P, Scherer SW: **Contribution of SHANK3 mutations to autism spectrum disorder.** *Am J Hum Genet* 2007, **81:**1289-1297.

104. Campbell NG, Zhu CB, Lindler KM, Yaspan BL, Kistner-Griffin E, Consortium NA, Hewlett WA, Tate CG, Blakely RD, Sutcliffe JS: **Rare coding variants of the adenosine A3 receptor are increased in autism: on the trail of the serotonin transporter regulome.** *Mol Autism* 2013, **4:**28.

105. Vaags AK, Lionel AC, Sato D, Goodenberger M, Stein QP, Curran S, Ogilvie C, Ahn JW, Drmic I, Senman L, et al: **Rare deletions at the neurexin 3 locus in autism spectrum disorder.** *Am J Hum Genet* 2012, **90:**133-141.

106. Noor A, Whibley A, Marshall CR, Gianakopoulos PJ, Piton A, Carson AR, Orlic-Milacic M, Lionel AC, Sato D, Pinto D, et al: **Disruption at the PTCHD1 Locus on Xp22.11 in Autism spectrum disorder and intellectual disability.** *Sci Transl Med* 2010, **2:**49ra68.

107. Glessner JT, Wang K, Cai G, Korvatska O, Kim CE, Wood S, Zhang H, Estes A, Brune CW, Bradfield JP, et al: **Autism genome-wide copy number variation reveals ubiquitin and neuronal genes.** *Nature* 2009, **459:**569-573.

108. Pagnamenta AT, Khan H, Walker S, Gerrelli D, Wing K, Bonaglia MC, Giorda R, Berney T, Mani E, Molteni M, et al: **Rare familial 16q21 microdeletions under a linkage peak implicate cadherin 8 (CDH8) in susceptibility to autism and learning disability.** *J Med Genet* 2011, **48:**48-54.

109. Bakkaloglu B, O'Roak BJ, Louvi A, Gupta AR, Abelson JF, Morgan TM, Chawarska K, Klin A, Ercan-Sencicek AG, Stillman AA, et al: **Molecular cytogenetic analysis and resequencing of contactin associated protein-like 2 in autism spectrum disorders.** *Am J Hum Genet* 2008, **82:**165-173.

110. State MW, Levitt P: **The conundrums of understanding genetic risks for autism spectrum disorders.** *Nat Neurosci* 2011, **14:**1499-1506.

111. Cook EH, Jr., Scherer SW: **Copy-number variations associated with neuropsychiatric conditions.** *Nature* 2008, **455:**919-923.

112. Sebat J, Lakshmi B, Malhotra D, Troge J, Lese-Martin C, Walsh T, Yamrom B, Yoon S, Krasnitz A, Kendall J, et al: **Strong association of de novo copy number mutations with autism.** *Science* 2007, **316:**445-449.

113. Levy D, Ronemus M, Yamrom B, Lee YH, Leotta A, Kendall J, Marks S, Lakshmi B, Pai D, Ye K, et al: **Rare de novo and transmitted copy-number variation in autistic spectrum disorders.** *Neuron* 2011, **70:**886-897.

114. Sanders SJ, Ercan-Sencicek AG, Hus V, Luo R, Murtha MT, Moreno-De-Luca D, Chu SH, Moreau MP, Gupta AR, Thomson SA, et al: **Multiple recurrent de novo CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism.** *Neuron* 2011, **70:**863-885.

115. Szatmari P, Paterson AD, Zwaigenbaum L, Roberts W, Brian J, Liu XQ, Vincent JB, Skaug JL, Thompson AP, Senman L, et al: **Mapping autism risk loci using genetic linkage and chromosomal rearrangements.** *Nature genetics* 2007, **39:**319-328.

116. Marshall CR, Noor A, Vincent JB, Lionel AC, Feuk L, Skaug J, Shago M, Moessner R, Pinto D, Ren Y, et al: **Structural variation of chromosomes in autism spectrum disorder.** *American journal of human genetics* 2008, **82:**477-488.

117. Weiss LA, Shen Y, Korn JM, Arking DE, Miller DT, Fossdal R, Saemundsen E, Stefansson H, Ferreira MA, Green T, et al: **Association between microdeletion and microduplication at 16p11.2 and autism.** *The New England journal of medicine* 2008, **358:**667-675.

118. Kumar RA, KaraMohamed S, Sudi J, Conrad DF, Brune C, Badner JA, Gilliam TC, Nowak NJ, Cook EH, Jr., Dobyns WB, Christian SL: **Recurrent 16p11.2 microdeletions in autism.** *Hum Mol Genet* 2008, **17:**628-638.

119. Pober BR: **Williams-Beuren syndrome.** *The New England journal of medicine* 2010, **362:**239-252.

120. Bolton PF, Veltman MW, Weisblatt E, Holmes JR, Thomas NS, Youings SA, Thompson RJ, Roberts SE, Dennis NR, Browne CE, et al: **Chromosome 15q11-13**

abnormalities and other medical conditions in individuals with autism spectrum disorders. *Psychiatr Genet* 2004, **14:**131-137.

121. Christian SL, Brune CW, Sudi J, Kumar RA, Liu S, Karamohamed S, Badner JA, Matsui S, Conroy J, McQuaid D, et al: **Novel submicroscopic chromosomal abnormalities detected in autism spectrum disorder.** *Biol Psychiatry* 2008, **63:**1111-1117.

122. Ahn K, Gotay N, Andersen TM, Anvari AA, Gochman P, Lee Y, Sanders S, Guha S, Darvasi A, Glessner JT, et al: **High rate of disease-related copy number variations in childhood onset schizophrenia.** *Molecular psychiatry* 2013.

123. Cirulli ET, Goldstein DB: **Uncovering the roles of rare variants in common disease through whole-genome sequencing.** *Nat Rev Genet* 2010, **11:**415-425.

124. O'Roak BJ, Vives L, Girirajan S, Karakoc E, Krumm N, Coe BP, Levy R, Ko A, Lee C, Smith JD, et al: **Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations.** *Nature* 2012, **485:**246-250.

125. Neale BM, Kou Y, Liu L, Ma'ayan A, Samocha KE, Sabo A, Lin CF, Stevens C, Wang LS, Makarov V, et al: **Patterns and rates of exonic de novo mutations in autism spectrum disorders.** *Nature* 2012, **485:**242-245.

126. Sanders SJ, Murtha MT, Gupta AR, Murdoch JD, Raubeson MJ, Willsey AJ, Ercan-Sencicek AG, DiLullo NM, Parikshak NN, Stein JL, et al: **De novo mutations revealed by whole-exome sequencing are strongly associated with autism.** *Nature* 2012, **485:**237-241.

127. Iossifov I, Ronemus M, Levy D, Wang Z, Hakker I, Rosenbaum J, Yamrom B, Lee YH, Narzisi G, Leotta A, et al: **De novo gene disruptions in children on the autistic spectrum.** *Neuron* 2012, **74:**285-299.

128. Lauritsen MB, Pedersen CB, Mortensen PB: **Effects of familial risk factors and place of birth on the risk of autism: a nationwide register-based study.** *J Child Psychol Psychiatry* 2005, **46:**963-971.

129. Reichenberg A, Gross R, Weiser M, Bresnahan M, Silverman J, Harlap S, Rabinowitz J, Shulman C, Malaspina D, Lubin G, et al: **Advancing paternal age and autism.** *Arch Gen Psychiatry* 2006, **63:**1026-1032.

130. Gabis L, Raz R, Kesner-Baruch Y: **Paternal age in autism spectrum disorders and ADHD.** *Pediatr Neurol* 2010, **43:**300-302.

131. Lundstrom S, Haworth CM, Carlstrom E, Gillberg C, Mill J, Rastam M, Hultman CM, Ronald A, Anckarsater H, Plomin R, et al: **Trajectories leading to autism spectrum disorders are affected by paternal age: findings from two nationally representative twin studies.** *J Child Psychol Psychiatry* 2010, **51:**850-856.

132. Zoghbi HY, Bear MF: **Synaptic dysfunction in neurodevelopmental disorders associated with autism and intellectual disabilities.** *Cold Spring Harb Perspect Biol* 2012, **4**.

133. Weiss LA, Escayg A, Kearney JA, Trudeau M, MacDonald BT, Mori M, Reichert J, Buxbaum JD, Meisler MH: **Sodium channels SCN1A, SCN2A and SCN3A in familial autism.** *Molecular psychiatry* 2003, **8:**186-194.

134. Torres GE, Gainetdinov RR, Caron MG: **Plasma membrane monoamine transporters: structure, regulation and function.** *Nat Rev Neurosci* 2003, **4:**13-25.

135. Hahn MK, Blakely RD: **The functional impact of SLC6 transporter genetic variation.** *Annu Rev Pharmacol Toxicol* 2007, **47:**401-441.

136. Kilty JE, Lorang D, Amara SG: **Cloning and expression of a cocaine-sensitive rat dopamine transporter.** *Science* 1991, **254:**578-579.

137. Shimada S, Kitayama S, Lin CL, Patel A, Nanthakumar E, Gregor P, Kuhar M, Uhl G: **Cloning and expression of a cocaine-sensitive dopamine transporter complementary DNA.** *Science* 1991, **254:**576-578.

138. Blakely RD, Berson HE, Fremeau RT, Jr., Caron MG, Peek MM, Prince HK, Bradley CC: **Cloning and expression of a functional serotonin transporter from rat brain.** *Nature* 1991, **354:**66-70.

139. Hoffman BJ, Mezey E, Brownstein MJ: **Cloning of a serotonin transporter affected by antidepressants.** *Science* 1991, **254:**579-580.

140. Steiner JA, Carneiro AM, Blakely RD: **Going with the flow: trafficking-dependent and -independent regulation of serotonin transport.** *Traffic* 2008, **9:**1393-1402.

141. Murphy DL, Lerner A, Rudnick G, Lesch KP: **Serotonin transporter: gene, genetic disorders, and pharmacogenetics.** *Mol Interv* 2004, **4:**109-123.

142. Hanna GL, Himle JA, Curtis GC, Koram DQ, Veenstra-VanderWeele J, Leventhal BL, Cook EH, Jr.: **Serotonin transporter and seasonal variation in blood serotonin in families with obsessive-compulsive disorder.** *Neuropsychopharmacology* 1998, **18:**102-111.

143. Hu XZ, Lipsky RH, Zhu G, Akhtar LA, Taubman J, Greenberg BD, Xu K, Arnold PD, Richter MA, Kennedy JL, et al: **Serotonin transporter promoter gain-of-function genotypes are linked to obsessive-compulsive disorder.** *Am J Hum Genet* 2006, **78:**815-826.

144. Ozaki N, Goldman D, Kaye WH, Plotnicov K, Greenberg BD, Lappalainen J, Rudnick G, Murphy DL: **Serotonin transporter missense mutation associated with a complex neuropsychiatric phenotype.** *Molecular psychiatry* 2003, **8:**933-936.

145. Akkermann K, Paaver M, Nordquist N, Oreland L, Harro J: **Association of 5-HTT gene polymorphism, platelet MAO activity, and drive for thinness in a population-based sample of adolescent girls.** *Int J Eat Disord* 2008, **41:**399-404.

146. Lesch KP, Bengel D, Heils A, Sabol SZ, Greenberg BD, Petri S, Benjamin J, Muller CR, Hamer DH, Murphy DL: **Association of anxiety-related traits with a polymorphism in the serotonin transporter gene regulatory region.** *Science* 1996, **274:**1527-1531.

147. Pinto E, Reggers J, Gorwood P, Boni C, Scantamburlo G, Pitchot W, Ansseau M: **The short allele of the serotonin transporter promoter polymorphism influences relapse in alcohol dependence.** *Alcohol Alcohol* 2008, **43:**398-400.

148. Florez G, Saiz P, Garcia-Portilla P, Alvarez S, Nogueiras L, Morales B, Alvarez V, Coto E, Bobes J: **Association between the Stin2 VNTR polymorphism of the serotonin transporter gene and treatment outcome in alcohol-dependent patients.** *Alcohol Alcohol* 2008, **43:**516-522.

149. Caspi A, Sugden K, Moffitt TE, Taylor A, Craig IW, Harrington H, McClay J, Mill J, Martin J, Braithwaite A, Poulton R: **Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene.** *Science* 2003, **301:**386-389.

150. Gillespie NA, Whitfield JB, Williams B, Heath AC, Martin NG: **The relationship between stressful life events, the serotonin transporter (5-HTTLPR) genotype and major depression.** *Psychol Med* 2005, **35:**101-111.

151. Mortensen OV, Thomassen M, Larsen MB, Whittemore SR, Wiborg O: **Functional analysis of a novel human serotonin transporter gene promoter in immortalized raphe cells.** *Brain Res Mol Brain Res* 1999, **68:**141-148.

152. Greenberg BD, Tolliver TJ, Huang SJ, Li Q, Bengel D, Murphy DL: **Genetic variation in the serotonin transporter promoter region affects serotonin uptake in human blood platelets.** *Am J Med Genet* 1999, **88:**83-87.

153. Little KY, McLaughlin DP, Zhang L, Livermore CS, Dalack GW, McFinton PR, DelProposto ZS, Hill E, Cassin BJ, Watson SJ, Cook EH: **Cocaine, ethanol, and genotype effects on human midbrain serotonin transporter binding sites and mRNA levels.** *Am J Psychiatry* 1998, **155:**207-213.

154. Lim JE, Papp A, Pinsonneault J, Sadee W, Saffen D: **Allelic expression of serotonin transporter (SERT) mRNA in human pons: lack of correlation with the polymorphism SERTLPR.** *Molecular psychiatry* 2006, **11:**649-662.

155. Devlin B, Cook EH, Jr., Coon H, Dawson G, Grigorenko EL, McMahon W, Minshew N, Pauls D, Smith M, Spence MA, et al: **Autism and the serotonin transporter: the long and short of it.** *Molecular psychiatry* 2005, **10:**1110-1116.

156. Homberg JR, Lesch KP: **Looking on the bright side of serotonin transporter gene variation.** *Biol Psychiatry* 2011, **69:**513-519.

157. Adamec R, Burton P, Blundell J, Murphy DL, Holmes A: **Vulnerability to mild predator stress in serotonin transporter knockout mice.** *Behav Brain Res* 2006, **170:**126-140.

158. Holmes A, Yang RJ, Murphy DL, Crawley JN: **Evaluation of antidepressant-related behavioral responses in mice lacking the serotonin transporter.** *Neuropsychopharmacology* 2002, **27:**914-923.

159. Lira A, Zhou M, Castanon N, Ansorge MS, Gordon JA, Francis JH, Bradley-Moore M, Lira J, Underwood MD, Arango V, et al: **Altered depression-related behaviors and functional changes in the dorsal raphe nucleus of serotonin transporter-deficient mice.** *Biol Psychiatry* 2003, **54:**960-971.

160. Bengel D, Greenberg BD, Cora-Locatelli G, Altemus M, Heils A, Li Q, Murphy DL: **Association of the serotonin transporter promoter regulatory region polymorphism and obsessive-compulsive disorder.** *Molecular psychiatry* 1999, **4:**463-466.

161. Sora I, Wichems C, Takahashi N, Li XF, Zeng Z, Revay R, Lesch KP, Murphy DL, Uhl GR: **Cocaine reward models: conditioned place preference can be established in dopamine- and in serotonin-transporter knockout mice.** *Proceedings of the National Academy of Sciences of the United States of America* 1998, **95:**7699-7704.

162. Murphy DL, Lesch KP: **Targeting the murine serotonin transporter: insights into human neurobiology.** *Nature reviews Neuroscience* 2008, **9:**85-96.

163. Murphy DL, Fox MA, Timpano KR, Moya PR, Ren-Patterson R, Andrews AM, Holmes A, Lesch KP, Wendland JR: **How the serotonin story is being rewritten by new gene-based discoveries principally related to SLC6A4, the serotonin transporter gene, which functions to influence all cellular serotonin systems.** *Neuropharmacology* 2008, **55:**932-960.

164. Kim DK, Tolliver TJ, Huang SJ, Martin BJ, Andrews AM, Wichems C, Holmes A, Lesch KP, Murphy DL: **Altered serotonin synthesis, turnover and dynamic**

regulation in multiple brain regions of mice lacking the serotonin transporter. *Neuropharmacology* 2005, **49:**798-810.

165. Cook EH, Leventhal BL: **The serotonin system in autism.** *Curr Opin Pediatr* 1996, **8:**348-354.

166. Bartlett CW, Gharani N, Millonig JH, Brzustowicz LM: **Three autism candidate genes: a synthesis of human genetic analysis with other disciplines.** *Int J Dev Neurosci* 2005, **23:**221-234.

167. Pardo CA, Eberhart CG: **The neurobiology of autism.** *Brain Pathol* 2007, **17:**434-447.

168. Lam KS, Aman MG, Arnold LE: **Neurochemical correlates of autistic disorder: a review of the literature.** *Res Dev Disabil* 2006, **27:**254-289.

169. Abramson RK, Wright HH, Carpenter R, Brennan W, Lumpuy O, Cole E, Young SR: **Elevated blood serotonin in autistic probands and their first-degree relatives.** *J Autism Dev Disord* 1989, **19:**397-407.

170. Cook EH, Jr., Leventhal BL, Heller W, Metz J, Wainwright M, Freedman DX: **Autistic children and their first-degree relatives: relationships between serotonin and norepinephrine levels and intelligence.** *J Neuropsychiatry Clin Neurosci* 1990, **2:**268-274.

171. Piven J, Tsai GC, Nehme E, Coyle JT, Chase GA, Folstein SE: **Platelet serotonin, a possible marker for familial autism.** *J Autism Dev Disord* 1991, **21:**51-59.

172. Schain RJ, Freedman DX: **Studies on 5-hydroxyindole metabolism in autistic and other mentally retarded children.** *J Pediatr* 1961, **58:**315-320.

173. McBride PA, Anderson GM, Hertzig ME, Snow ME, Thompson SM, Khait VD, Shapiro T, Cohen DJ: **Effects of diagnosis, race, and puberty on platelet serotonin levels in autism and mental retardation.** *Journal of the American Academy of Child and Adolescent Psychiatry* 1998, **37:**767-776.

174. McDougle CJ, Naylor ST, Cohen DJ, Volkmar FR, Heninger GR, Price LH: **A double-blind, placebo-controlled study of fluvoxamine in adults with autistic disorder.** *Arch Gen Psychiatry* 1996, **53:**1001-1008.

175. Namerow LB, Thomas P, Bostic JQ, Prince J, Monuteaux MC: **Use of citalopram in pervasive developmental disorders.** *J Dev Behav Pediatr* 2003, **24:**104-108.

176. Chugani DC, Muzik O, Behen M, Rothermel R, Janisse JJ, Lee J, Chugani HT: **Developmental changes in brain serotonin synthesis capacity in autistic and nonautistic children.** *Ann Neurol* 1999, **45:**287-295.

177. Chugani DC: **Role of altered brain serotonin mechanisms in autism.** *Molecular psychiatry* 2002, **7 Suppl 2:**S16-17.

178. McDougle CJ, Naylor ST, Goodman WK, Volkmar FR, Cohen DJ, Price LH: **Acute tryptophan depletion in autistic disorder: a controlled case study.** *Biol Psychiatry* 1993, **33:**547-550.

179. McDougle CJ, Naylor ST, Cohen DJ, Aghajanian GK, Heninger GR, Price LH: **Effects of tryptophan depletion in drug-free adults with autistic disorder.** *Arch Gen Psychiatry* 1996, **53:**993-1000.

180. Whitaker-Azmitia PM: **Behavioral and cellular consequences of increasing serotonergic activity during brain development: a role in autism?** *Int J Dev Neurosci* 2005, **23:**75-83.

181. McNamara IM, Borella AW, Bialowas LA, Whitaker-Azmitia PM: **Further studies in the developmental hyperserotonemia model (DHS) of autism: social, behavioral and peptide changes.** *Brain Res* 2008, **1189:**203-214.

182. Kim SJ, Cox N, Courchesne R, Lord C, Corsello C, Akshoomoff N, Guter S, Leventhal BL, Courchesne E, Cook EH, Jr.: **Transmission disequilibrium mapping at the serotonin transporter gene (SLC6A4) region in autistic disorder.** *Molecular psychiatry* 2002, **7:**278-288.

183. Huang CH, Santangelo SL: **Autism and serotonin transporter gene polymorphisms: a systematic review and meta-analysis.** *Am J Med Genet B Neuropsychiatr Genet* 2008, **147B:**903-913.

184. Cook EH, Jr., Courchesne R, Lord C, Cox NJ, Yan S, Lincoln A, Haas R, Courchesne E, Leventhal BL: **Evidence of linkage between the serotonin transporter and autistic disorder.** *Molecular psychiatry* 1997, **2:**247-250.

185. Tadevosyan-Leyfer O, Dowd M, Mankoski R, Winklosky B, Putnam S, McGrath L, Tager-Flusberg H, Folstein SE: **A principal components analysis of the Autism Diagnostic Interview-Revised.** *Journal of the American Academy of Child and Adolescent Psychiatry* 2003, **42:**864-872.

186. Veenstra-Vanderweele J, Jessen TN, Thompson BJ, Carter M, Prasad HC, Steiner JA, Sutcliffe JS, Blakely RD: **Modeling rare gene variation to gain insight into the oldest biomarker in autism: construction of the serotonin transporter Gly56Ala knock-in mouse.** *J Neurodev Disord* 2009, **1:**158-171.

187. Prasad HC, Steiner JA, Sutcliffe JS, Blakely RD: **Enhanced activity of human serotonin transporter variants associated with autism.** *Philos Trans R Soc Lond B Biol Sci* 2009, **364:**163-173.

188. Prasad HC, Zhu CB, McCauley JL, Samuvel DJ, Ramamoorthy S, Shelton RC, Hewlett WA, Sutcliffe JS, Blakely RD: **Human serotonin transporter variants display altered sensitivity to protein kinase G and p38 mitogen-activated protein kinase.** *Proceedings of the National Academy of Sciences of the United States of America* 2005, **102:**11545-11550.

189. Sakurai T, Reichert J, Hoffman EJ, Cai G, Jones HB, Faham M, Buxbaum JD: **A large-scale screen for coding variants in SERT/SLC6A4 in autism spectrum disorders.** *Autism Res* 2008, **1:**251-257.

190. Delorme R, Betancur C, Wagner M, Krebs MO, Gorwood P, Pearl P, Nygren G, Durand CM, Buhtz F, Pickering P, et al: **Support for the association between the rare functional variant I425V of the serotonin transporter gene and susceptibility to obsessive compulsive disorder.** *Molecular psychiatry* 2005, **10:**1059-1061.

191. Veenstra-VanderWeele J, Muller CL, Iwamoto H, Sauer JE, Owens WA, Shah CR, Cohen J, Mannangatti P, Jessen T, Thompson BJ, et al: **Autism gene variant causes hyperserotonemia, serotonin receptor hypersensitivity, social impairment and repetitive behavior.** *Proceedings of the National Academy of Sciences* 2012.

192. Bjorklund A, Dunnett SB: **Fifty years of dopamine research.** *Trends Neurosci* 2007, **30:**185-187.

193. Giros B, Caron MG: **Molecular characterization of the dopamine transporter.** *Trends Pharmacol Sci* 1993, **14:**43-49.

194. Carlsson A: **Perspectives on the discovery of central monoaminergic neurotransmission.** *Annu Rev Neurosci* 1987, **10:**19-40.

195. Palmiter RD: **Dopamine signaling in the dorsal striatum is essential for motivated behaviors: lessons from dopamine-deficient mice.** *Ann N Y Acad Sci* 2008, **1129:**35-46.

196. Dunlop BW, Nemeroff CB: **The role of dopamine in the pathophysiology of depression.** *Arch Gen Psychiatry* 2007, **64:**327-337.

197. Cordeiro Q, Siqueira-Roberto J, Vallada H: **Association between the SLC6A3 A1343G polymorphism and schizophrenia.** *Arq Neuropsiquiatr* 2010, **68:**716-719.

198. Pattarachotanant N, Sritharathikhun T, Suttirat S, Tencomnao T: **Association of C/T polymorphism in intron 14 of the dopamine transporter gene (rs40184) with major depression in a northeastern Thai population.** *Genet Mol Res* 2010, **9:**565-572.

199. Bowton E, Saunders C, Erreger K, Sakrikar D, Matthies HJ, Sen N, Jessen T, Colbran RJ, Caron MG, Javitch JA, et al: **Dysregulation of dopamine transporters via dopamine D2 autoreceptors triggers anomalous dopamine efflux associated with attention-deficit hyperactivity disorder.** *J Neurosci* 2010, **30:**6048-6057.

200. Mazei-Robison MS, Bowton E, Holy M, Schmudermaier M, Freissmuth M, Sitte HH, Galli A, Blakely RD: **Anomalous dopamine release associated with a human dopamine transporter coding variant.** *J Neurosci* 2008, **28:**7040-7046.

201. Ronald A, Simonoff E, Kuntsi J, Asherson P, Plomin R: **Evidence for overlapping genetic influences on autistic and ADHD behaviours in a community twin sample.** *J Child Psychol Psychiatry* 2008, **49:**535-542.

202. Gadow KD, DeVincent CJ, Pomeroy J: **ADHD symptom subtypes in children with pervasive developmental disorder.** *J Autism Dev Disord* 2006, **36:**271-283.

203. Goldstein S, Schwebach AJ: **The comorbidity of Pervasive Developmental Disorder and Attention Deficit Hyperactivity Disorder: results of a retrospective chart review.** *J Autism Dev Disord* 2004, **34:**329-339.

204. Roman T, Rohde LA, Hutz MH: **Polymorphisms of the dopamine transporter gene: influence on response to methylphenidate in attention deficit-hyperactivity disorder.** *Am J Pharmacogenomics* 2004, **4:**83-92.

205. Cook EH, Jr., Stein MA, Krasowski MD, Cox NJ, Olkon DM, Kieffer JE, Leventhal BL: **Association of attention-deficit disorder and the dopamine transporter gene.** *Am J Hum Genet* 1995, **56:**993-998.

206. Hettinger JA, Liu X, Schwartz CE, Michaelis RC, Holden JJ: **A DRD1 haplotype is associated with risk for autism spectrum disorders in male-only affected sib-pair families.** *Am J Med Genet B Neuropsychiatr Genet* 2008, **147B:**628-636.

207. Nakamura K, Sekine Y, Ouchi Y, Tsujii M, Yoshikawa E, Futatsubashi M, Tsuchiya KJ, Sugihara G, Iwata Y, Suzuki K, et al: **Brain serotonin and dopamine transporter bindings in adults with high-functioning autism.** *Arch Gen Psychiatry* 2010, **67:**59-68.

208. Weiss LA, Veenstra-Vanderweele J, Newman DL, Kim SJ, Dytch H, McPeek MS, Cheng S, Ober C, Cook EH, Jr., Abney M: **Genome-wide association study identifies ITGB3 as a QTL for whole blood serotonin.** *European journal of human genetics : EJHG* 2004, **12:**949-954.

209. Carneiro AM, Cook EH, Murphy DL, Blakely RD: **Interactions between integrin alphaIIbbeta3 and the serotonin transporter regulate serotonin transport and platelet aggregation in mice and humans.** *J Clin Invest* 2008, **118:**1544-1552.

210. Zhu CB, Hewlett WA, Feoktistov I, Biaggioni I, Blakely RD: **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* 2004, **65:**1462-1474.

211. Zhu CB, Carneiro AM, Dostmann WR, Hewlett WA, Blakely RD: **p38 MAPK activation elevates serotonin transport activity via a trafficking-independent, protein phosphatase 2A-dependent process.** *J Biol Chem* 2005, **280:**15649-15658.

212. Miller KJ, Hoffman BJ: **Adenosine A3 receptors regulate serotonin transport via nitric oxide and cGMP.** *J Biol Chem* 1994, **269:**27351-27356.

213. Zhu CB, Lindler KM, Campbell NG, Sutcliffe JS, Hewlett WA, Blakely RD: **Colocalization and regulated physical association of presynaptic serotonin transporters with A(3) adenosine receptors.** *Mol Pharmacol* 2011, **80:**458-465.

214. Zhu CB, Steiner JA, Munn JL, Daws LC, Hewlett WA, Blakely RD: **Rapid stimulation of presynaptic serotonin transport by A(3) adenosine receptors.** *J Pharmacol Exp Ther* 2007, **322:**332-340.

215. O'Connell JR, Weeks DE: **PedCheck: a program for identification of genotype incompatibilities in linkage analysis.** *Am J Hum Genet* 1998, **63:**259-266.

216. Horvath S, Xu X, Laird NM: **The family based association test method: strategies for studying general genotype--phenotype associations.** *Eur J Hum Genet* 2001, **9:**301-306.

217. Purcell S, Cherny SS, Sham PC: **Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits.** *Bioinformatics* 2003, **19:**149-150.

218. Pritchard JK, Stephens M, Donnelly P: **Inference of population structure using multilocus genotype data.** *Genetics* 2000, **155:**945-959.

219. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC: **PLINK: a tool set for whole-genome association and population-based linkage analyses.** *Am J Hum Genet* 2007, **81:**559-575.

220. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR: **A method and server for predicting damaging missense mutations.** *Nat Methods* 2010, **7:**248-249.

221. Adzhubei I, Jordan DM, Sunyaev SR: **Predicting functional effect of human missense mutations using PolyPhen-2.** *Curr Protoc Hum Genet* 2013, **Chapter 7:**Unit7 20.

222. Bromberg Y, Rost B: **SNAP: predict effect of non-synonymous polymorphisms on function.** *Nucleic Acids Res* 2007, **35:**3823-3835.

223. Morgenthaler S, Thilly WG: **A strategy to discover genes that carry multi-allelic or mono-allelic risk for common diseases: a cohort allelic sums test (CAST).** *Mutat Res* 2007, **615:**28-56.

224. Li B, Liu DJ, Leal SM: **Identifying rare variants associated with complex traits via sequencing.** *Curr Protoc Hum Genet* 2013, **Chapter 1:**Unit 1 26.

225. Lebon G, Warne T, Edwards PC, Bennett K, Langmead CJ, Leslie AG, Tate CG: **Agonist-bound adenosine A2A receptor structures reveal common features of GPCR activation.** *Nature* 2011, **474:**521-525.

226. Bucan M, Abrahams BS, Wang K, Glessner JT, Herman EI, Sonnenblick LI, Alvarez Retuerto AI, Imielinski M, Hadley D, Bradfield JP, et al: **Genome-wide analyses of exonic copy number variants in a family-based study point to novel autism susceptibility genes.** *PLoS Genet* 2009, **5:**e1000536.

227. Bonnin A, Levitt P: **Fetal, maternal, and placental sources of serotonin and new implications for developmental programming of the brain.** *Neuroscience* 2011, **197:**1-7.

228. Gaspar P, Cases O, Maroteaux L: **The developmental role of serotonin: news from mouse molecular genetics.** *Nat Rev Neurosci* 2003, **4:**1002-1012.

229. Deneris ES, Wyler SC: **Serotonergic transcriptional networks and potential importance to mental health.** *Nat Neurosci* 2012, **15:**519-527.

230. Murphy DL, Lesch KP: **Targeting the murine serotonin transporter: insights into human neurobiology.** *Nat Rev Neurosci* 2008, **9:**85-96.

231. Klauck SM, Felder B, Kolb-Kokocinski A, Schuster C, Chiocchetti A, Schupp I, Wellenreuther R, Schmotzer G, Poustka F, Breitenbach-Koller L, Poustka A:

Mutations in the ribosomal protein gene RPL10 suggest a novel modulating disease mechanism for autism. *Mol Psychiatry* 2006, **11:**1073-1084.

232. Veenstra-VanderWeele J, Muller CL, Iwamoto H, Sauer JE, Owens WA, Shah CR, Cohen J, Mannangatti P, Jessen T, Thompson BJ, et al: **Autism gene variant causes hyperserotonemia, serotonin receptor hypersensitivity, social impairment and repetitive behavior.** *Proceedings of the National Academy of Sciences of the United States of America* 2012, **109:**5469-5474.

233. Hamilton PJ, Campbell NG, Sharma S, Erreger K, Herborg Hansen F, Saunders C, Belovich AN, Consortium NAAS, Sahai MA, Cook EH, et al: **De novo mutation in the dopamine transporter gene associates dopamine dysfunction with autism spectrum disorder.** *Molecular psychiatry* 2013.

234. Seeman P, Niznik HB: **Dopamine receptors and transporters in Parkinson's disease and schizophrenia.** *FASEB J* 1990, **4:**2737-2744.

235. Volkow ND, Wang GJ, Newcorn J, Telang F, Solanto MV, Fowler JS, Logan J, Ma Y, Schulz K, Pradhan K, et al: **Depressed dopamine activity in caudate and preliminary evidence of limbic involvement in adults with attention-deficit/hyperactivity disorder.** *Arch Gen Psychiatry* 2007, **64:**932-940.

236. Cousins DA, Butts K, Young AH: **The role of dopamine in bipolar disorder.** *Bipolar Disord* 2009, **11:**787-806.

237. Kristensen AS, Andersen J, Jorgensen TN, Sorensen L, Eriksen J, Loland CJ, Stromgaard K, Gether U: **SLC6 neurotransmitter transporters: structure, function, and regulation.** *Pharmacol Rev* 2011, **63:**585-640.

238. Kahlig KM, Binda F, Khoshbouei H, Blakely RD, McMahon DG, Javitch JA, Galli A: **Amphetamine induces dopamine efflux through a dopamine transporter channel.** *Proceedings of the National Academy of Sciences of the United States of America* 2005, **102:**3495-3500.

239. Khoshbouei H, Sen N, Guptaroy B, Johnson L, Lund D, Gnegy ME, Galli A, Javitch JA: **N-terminal phosphorylation of the dopamine transporter is required for amphetamine-induced efflux.** *PLoS Biol* 2004, **2:**E78.

240. Giros B, Jaber M, Jones SR, Wightman RM, Caron MG: **Hyperlocomotion and indifference to cocaine and amphetamine in mice lacking the dopamine transporter.** *Nature* 1996, **379:**606-612.

241. Jones SR, Gainetdinov RR, Wightman RM, Caron MG: **Mechanisms of amphetamine action revealed in mice lacking the dopamine transporter.** *J Neurosci* 1998, **18:**1979-1986.

242. Sulzer D, Sonders MS, Poulsen NW, Galli A: **Mechanisms of neurotransmitter release by amphetamines: a review.** *Prog Neurobiol* 2005, **75:**406-433.

243. Kurian MA, Zhen J, Cheng SY, Li Y, Mordekar SR, Jardine P, Morgan NV, Meyer E, Tee L, Pasha S, et al: **Homozygous loss-of-function mutations in the gene encoding the dopamine transporter are associated with infantile parkinsonism-dystonia.** *J Clin Invest* 2009, **119:**1595-1603.

244. Kurian MA, Li Y, Zhen J, Meyer E, Hai N, Christen HJ, Hoffmann GF, Jardine P, von Moers A, Mordekar SR, et al: **Clinical and molecular characterisation of hereditary dopamine transporter deficiency syndrome: an observational cohort and experimental study.** *Lancet Neurol* 2011, **10:**54-62.

245. Grunhage F, Schulze TG, Muller DJ, Lanczik M, Franzek E, Albus M, Borrmann-Hassenbach M, Knapp M, Cichon S, Maier W, et al: **Systematic screening for DNA sequence variation in the coding region of the human dopamine transporter gene (DAT1).** *Molecular psychiatry* 2000, **5:**275-282.

246. Mazei-Robison MS, Couch RS, Shelton RC, Stein MA, Blakely RD: **Sequence variation in the human dopamine transporter gene in children with attention deficit hyperactivity disorder.** *Neuropharmacology* 2005, **49:**724-736.

247. Nemoda Z, Szekely A, Sasvari-Szekely M: **Psychopathological aspects of dopaminergic gene polymorphisms in adolescence and young adulthood.** *Neurosci Biobehav Rev* 2011, **35:**1665-1686.

248. Dichter GS, Damiano CA, Allen JA: **Reward circuitry dysfunction in psychiatric and neurodevelopmental disorders and genetic syndromes: animal models and clinical findings.** *J Neurodev Disord* 2012, **4:**19.

249. Di Martino A, Zuo XN, Kelly C, Grzadzinski R, Mennes M, Schvarcz A, Rodman J, Lord C, Castellanos FX, Milham MP: **Shared and distinct intrinsic functional network centrality in autism and attention-deficit/hyperactivity disorder.** *Biol Psychiatry* 2013, **74:**623-632.

250. Wolfson W: **Boston Autism Consortium searches for genetic clues to autism's puzzle.** *Chem Biol* 2007, **14:**117-118.

251. Risi S, Lord C, Gotham K, Corsello C, Chrysler C, Szatmari P, Cook EH, Jr., Leventhal BL, Pickles A: **Combining information from multiple sources in the diagnosis of autism spectrum disorders.** *Journal of the American Academy of Child and Adolescent Psychiatry* 2006, **45:**1094-1103.

252. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, et al: **A framework for variation discovery and genotyping using next-generation DNA sequencing data.** *Nat Genet* 2011, **43:**491-498.

253. Rickhag M, Hansen FH, Sorensen G, Strandfelt KN, Andresen B, Gotfryd K, Madsen KL, Vestergaard-Klewe I, Ammendrup-Johnsen I, Eriksen J, et al: **A C-terminal PDZ domain-binding sequence is required for striatal distribution of the dopamine transporter.** *Nat Commun* 2013, **4:**1580.

254. Rasmussen TN, Plenge P, Bay T, Egebjerg J, Gether U: **A single nucleotide polymorphism in the human serotonin transporter introduces a new site for N-linked glycosylation.** *Neuropharmacology* 2009, **57:**287-294.

255. Claxton DP, Quick M, Shi L, de Carvalho FD, Weinstein H, Javitch JA, McHaourab HS: **Ion/substrate-dependent conformational dynamics of a bacterial homolog of neurotransmitter:sodium symporters.** *Nat Struct Mol Biol* 2010, **17:**822-829.

256. Jeschke G, Polyhach Y: **Distance measurements on spin-labelled biomacromolecules by pulsed electron paramagnetic resonance.** *Phys Chem Chem Phys* 2007, **9:**1895-1910.

257. Zou P, McHaourab HS: **Increased sensitivity and extended range of distance measurements in spin-labeled membrane proteins: Q-band double electron-electron resonance and nanoscale bilayers.** *Biophys J* 2010, **98:**L18-20.

258. Jeschke G, Koch A, Jonas U, Godt A: **Direct conversion of EPR dipolar time evolution data to distance distributions.** *J Magn Reson* 2002, **155:**72-82.

259. Kume K, Kume S, Park SK, Hirsh J, Jackson FR: **Dopamine is a regulator of arousal in the fruit fly.** *J Neurosci* 2005, **25:**7377-7384.

260. Friggi-Grelin F, Coulom H, Meller M, Gomez D, Hirsh J, Birman S: **Targeted gene expression in Drosophila dopaminergic cells using regulatory sequences from tyrosine hydroxylase.** *J Neurobiol* 2003, **54:**618-627.

261. Wang JW, Beck ES, McCabe BD: **A modular toolset for recombination transgenesis and neurogenetic analysis of Drosophila.** *PLoS One* 2012, **7:**e42102.

262. Genomes Project C, Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, Kang HM, Marth GT, McVean GA: **An integrated map of genetic variation from 1,092 human genomes.** *Nature* 2012, **491:**56-65.

263. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K: **dbSNP: the NCBI database of genetic variation.** *Nucleic Acids Res* 2001, **29:**308-311.

264. Yamashita A, Singh SK, Kawate T, Jin Y, Gouaux E: **Crystal structure of a bacterial homologue of Na+/Cl--dependent neurotransmitter transporters.** *Nature* 2005, **437:**215-223.

265. Khoshbouei H, Wang H, Lechleiter JD, Javitch JA, Galli A: **Amphetamine-induced dopamine efflux. A voltage-sensitive and intracellular Na+-dependent mechanism.** *J Biol Chem* 2003, **278:**12070-12077.

266. McHaourab HS, Steed PR, Kazmier K: **Toward the fourth dimension of membrane protein structure: insight into dynamics from spin-labeling EPR spectroscopy.** *Structure* 2011, **19:**1549-1561.

267. McHaourab HS, Lin YL, Spiller BW: **Crystal structure of an activated variant of small heat shock protein Hsp16.5.** *Biochemistry* 2012, **51:**5105-5112.

268. Zhao Y, Terry DS, Shi L, Quick M, Weinstein H, Blanchard SC, Javitch JA: **Substrate-modulated gating dynamics in a Na+-coupled neurotransmitter transporter homologue.** *Nature* 2011, **474:**109-113.

269. Wicker-Thomas C, Hamann M: **Interaction of dopamine, female pheromones, locomotion and sex behavior in Drosophila melanogaster.** *J Insect Physiol* 2008, **54:**1423-1431.

270. Pendleton RG, Rasheed A, Sardina T, Tully T, Hillman R: **Effects of tyrosine hydroxylase mutants on locomotor activity in Drosophila: a study in functional genomics.** *Behav Genet* 2002, **32:**89-94.

271. Pizzo AB, Karam CS, Zhang Y, Yano H, Freyberg RJ, Karam DS, Freyberg Z, Yamamoto A, McCabe BD, Javitch JA: **The membrane raft protein Flotillin-1 is essential in dopamine neurons for amphetamine-induced behavior in Drosophila.** *Molecular psychiatry* 2013, **18:**824-833.

272. Schafer CM, Campbell NG, Cai G, Yu F, Makarov V, Yoon S, Daly MJ, Gibbs RA, Schellenberg GD, Devlin B, et al: **Whole exome sequencing reveals minimal differences between cell line and whole blood derived DNA.** *Genomics* 2013, **102:**270-277.

273. Conrad DF, Keebler JE, DePristo MA, Lindsay SJ, Zhang Y, Casals F, Idaghdour Y, Hartl CL, Torroja C, Garimella KV, et al: **Variation in genome-wide mutation rates within and between human families.** *Nat Genet* 2011, **43:**712-714.

274. Awadalla P, Gauthier J, Myers RA, Casals F, Hamdan FF, Griffing AR, Cote M, Henrion E, Spiegelman D, Tarabeux J, et al: **Direct measure of the de novo mutation rate in autism and schizophrenia cohorts.** *Am J Hum Genet* 2010, **87:**316-324.

275. Gauthier J, Champagne N, Lafreniere RG, Xiong L, Spiegelman D, Brustein E, Lapointe M, Peng H, Cote M, Noreau A, et al: **De novo mutations in the gene encoding the synaptic scaffolding protein SHANK3 in patients ascertained for**

schizophrenia. *Proceedings of the National Academy of Sciences of the United States of America* 2010, **107:**7863-7868.

276. Herbeck JT, Gottlieb GS, Wong K, Detels R, Phair JP, Rinaldo CR, Jacobson LP, Margolick JB, Mullins JI: **Fidelity of SNP array genotyping using Epstein Barr virus-transformed B-lymphocyte cell lines: implications for genome-wide association studies.** *PLoS One* 2009, **4:**e6915.

277. Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W, et al: **Global variation in copy number in the human genome.** *Nature* 2006, **444:**444-454.

278. Jeon JP, Shim SM, Nam HY, Baik SY, Kim JW, Han BG: **Copy number increase of 1p36.33 and mitochondrial genome amplification in Epstein-Barr virus-transformed lymphoblastoid cell lines.** *Cancer Genet Cytogenet* 2007, **173:**122-130.

279. Simon-Sanchez J, Scholz S, Fung HC, Matarin M, Hernandez D, Gibbs JR, Britton A, de Vrieze FW, Peckham E, Gwinn-Hardy K, et al: **Genome-wide SNP assay reveals structural genomic variation, extended homozygosity and cell-line induced alterations in normal individuals.** *Hum Mol Genet* 2007, **16:**1-14.