Exploration of Functional Genetic Variants in Candidate Genes for Psychiatric Disorders

DISSERTATION

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

By

Robert A. Moyer

Graduate Program in Integrated Biomedical Science Program

The Ohio State University

2010

Dissertation Committee:
Professor Wolfgang Sadée, Advisor
Professor Rene Anand
Professor R. Thomas Boyd
Professor Howard Gu
Abstract

Drug addiction is a chronic disorder characterized by compulsive drug seeking and drug taking despite serious negative consequences. While twin studies indicate a significant contribution of genetic factors to an individual’s susceptibility to addiction, identifying and replicating significant genetic associations for addiction has proven difficult. Association studies have identified a number of candidate genes harboring variants associated with susceptibility to addiction, but the causative variants frequently remain unknown. Instead, the association is found between the disease and a polymorphism that is linked to the causal variant. As allele frequencies and linkage patterns may vary between populations, associations found are frequently not replicated in follow-up studies, in part due to a lack of consideration of the molecular mechanisms underlying the effects of the polymorphisms.

A solution to this problem is the thorough characterization of functional genetic polymorphisms and their effects on regulation of candidate genes for addiction and other psychiatric disorders. Historically, many genetic studies have focused on nonsynonymous single nucleotide polymorphisms (SNPs) that alter amino acid sequence, even though regulatory polymorphisms in non-protein coding regions are likely to be more prevalent. Regulatory polymorphisms can result in altered transcription, alternative splicing, and mRNA folding and stability, to name but a few mechanisms. The first step
in our approach is to measure allelic mRNA expression and/or differential expression of functionally distinct splice variants for candidate genes in relevant brain regions from human autopsies. We use these measurements as phenotypic traits to identify causative genetic variants affecting the amount and/or nature of the mRNA. Finally, we test the putative functional variant for association with human phenotype to establish its clinical relevance.

I focused on a number of candidate genes for drug addiction, and made some novel and significant findings: 1) The effects of two intronic SNPs in the human dopamine receptor D2 (DRD2) gene, previously shown to alter DRD2 alternative splicing, were confirmed in human brain autopsy tissues (putamen and prefrontal cortex) obtained from a population of cocaine abusers and controls. The same SNPs were significantly associated with cocaine abuse with an odds ratio of ~2. Furthermore, I detected a significant interaction between these SNPs and a functional variant in the dopamine transporter (DAT) gene in the same population. 2) A synonymous SNP in the human βarrestin2 (ARRB2) gene was associated with a small but significant increase in ARRB2 mRNA expression in prefrontal cortex obtained from Alzheimer’s patients. However, there was no genotype effect on βarrestin2 protein levels in the same samples. 3) I identified a number of samples with significant allelic expression imbalance (AEI) of DRD3 mRNA in nucleus accumbens obtained from a cohort of normal controls and patients diagnosed with schizophrenia, bipolar disorder, or depression, indicating the presence of a cis-acting regulatory variant affecting mRNA expression. However, none
of the 11 SNPs genotyped across the DRD3 gene locus could account for the observed AEI.

Identification of functional genetic variants and characterization of the underlying molecular genetic mechanisms are critical for interpretation of clinical association studies and selection of valid biomarkers affecting disease susceptibility and treatment response. The striking association of the DRD2 SNPs with cocaine abuse supports the notion that the modulation of DRD2 alternative splicing is an important regulator of dopaminergic activity with phenotypic consequences, although the data do not prove a causative link between alternative splicing and cocaine abuse. The results of this study contribute to our understanding of the functional genetic variation in three important candidate genes and have the potential to improve the prevention and treatment of drug addiction and other psychiatric disorders.
Acknowledgments

I thank my adviser, Dr. Wolfgang Sadée, for teaching me to think and ask questions like a scientist, and for teaching me a great deal about the many other facets to becoming a successful scientist in a relatively short amount of time. I also thank our lab manager, Audrey Papp, for countless helpful discussions about biology, problem solving, and life in general. I thank them both for providing an excellent work environment that allowed me ask and pursue my own questions while keeping me on track. I am also grateful to my committee members, Drs. Rene Anand, R. Thomas Boyd, Howard Gu, and Lane Wallace, for their time and for their guidance along the way.

I am thankful to the rest of the Sadée laboratory members, both past and present, for their advice, challenging questions, and friendly and often entertaining dicussions: Ryan Smith, Dr. Danxin Wang, Dr. Julia Pinsonneault, Mandy Zaret, Gloria Smith, Dr. Ying Zhang, Dave Straka, Diane Delobel, Beth Barrie, Jonathon Sanford, Kumiko Aizawa, Danielle Sullivan, and Sarah Fritz.

I thank Dr. Laura Bohn, my advisor before she moved to another institution, for getting me started in biomedical research and involving me in collaboration with the Sadée laboratory before she left. I am also grateful to her for her support and understanding before, during, and after my deployment to Iraq. I also thank the members of her laboratory for all the coaching, collaborating, and some pretty good times both in
and out of the lab: Dr. Kirsten Raehal, Chad Groer, Lori Hudson, Cullen Schmid, and
Alex Jaeger.

I am thankful for the support of IBGP program staff, without whom there would
be no Integrated Biomedical Science Graduate Program and who were always there to
help me navigate the system successfully: Christine Kerr, Dr. Allan Yates, Dr. Virginia
Sanders, Dr. R. Thomas Boyd, Angie Thomas, Darlene Johns, Amy Lahmers, and Kelly
Dillon. I am also grateful to the Pharmacology Department staff (Sherry Ring, Gina
Pace) for all their patience and help with important paperwork and procedures.

I thank my collaborators outside the laboratory whose work is also presented here:
Dr. Deborah Mash, Linda Duque, Dr. Alessandro Bertolino, and Dr. Maree Webster.

I thank the anonymous tissue donors and their families without whom this work
would not have been possible.

I am forever grateful to my parents, Tom and Rose Moyer, who brought me up
curious and were always supportive. I would not be the person I am today were it not for
their example and the values they instilled in me from very early on.

I thank my wife, Katie, who has been a wonderful mother to our daughter in
addition to supporting me every step of the way during this dissertation research. Her
patience, kindness, and love throughout all the ups and downs of graduate school
(including my deployment to Iraq right in the middle of it) have been invaluable for me to
reach this point.

Finally, I thank my baby daughter Emmalyn, for really opening my eyes and
changing my perspective on just about everything.

vi
Vita

1992.................................................................Chanel High School

1997.................................................................B.S. in Education, Kent State University

2004 to present ..............................................Ph.D. in Integrated Biomedical Science, The Ohio State University

Publications


Fields of Study

Major Field: Integrated Biomedical Science Program
Table of Contents

Abstract .................................................................................................................................................. ii

Acknowledgments .................................................................................................................................... v

Vita ........................................................................................................................................................ vii

List of Tables .......................................................................................................................................... x

List of Figures ......................................................................................................................................... xi

Chapter 1: Introduction ........................................................................................................................ 1

1.1 Drug addiction and genetics ........................................................................................................ 1

1.2 The role of cis-acting regulatory polymorphisms in human genetic variation ........... 3

1.3 Epigenetic Factors ...................................................................................................................... 5

1.4 Cis-Acting Genetic Regulation of Alternative Splicing ....................................................... 7

1.5 Review of DRD2 Genotype-Phenotype Associations ........................................................ 7

1.6 Review of ARRB2 Genotype-Phenotype Associations ...................................................... 9

1.7 Review of DRD3 Genotype-Phenotype Associations ........................................................ 9

1.8 Summary of the Main Studies ................................................................................................. 10

Chapter 2: Intronic Polymorphisms Affecting Alternative Splicing of Human D2

Receptor are Associated with Cocaine Abuse ................................................................................. 13
List of Tables

Table 2.1. Results of sequencing DNA from five samples that displayed different D2L/D2S allelic ratios..........................................................22
Table 2.2. Demographic Characteristics....................................................31
Table 2.3. Pairwise Linkage Disequilibrium Values.................................32
Table 2.4. Minor allele frequencies in cocaine abusers and controls, plus results of statistical group comparisons..............................................33
Table 2.5. Haplotype Analysis in Cocaine Abusers and Controls..............35
Table 2.6. DRD2-DAT Gene Interaction....................................................36
Table 2.7. Clinical Associations of SNPs in LD (R2>0.6) with rs2283265.......45
Table 2.8. Sequences of Oligonucleotides (Primers) Used.......................46-47
Table 3.1. Demographic Characteristics..................................................71
Table 3.2. Sequences of Oligonucleotides (Primers) Used.......................73
Table 4.1. Sequences of Oligonucleotides (Primers) Used.......................86
List of Figures

Figure 1.1. Quantitative allele-specific expression analysis……………………………..5
Figure 2.1. *DRD2* gene map, with blown up view of the region that was sequenced…..21
Figure 2.2. Alternative splicing from *DRD2* minigenes in HEK 293 cells………………24
Figure 2.3. *DRD2* mRNA expression…………………………………………………...25
Figure 2.4. *DRD2* mRNA expression grouped by race and cocaine abuse status……....26
Figure 2.5. Expression of D2S relative to D2L…………………………………………28
Figure 2.6. Expression of D2S relative to D2L grouped by race………………………29
Figure 2.7. LD plot of surrogate markers for the intron 5 SNP rs2283265 (SNP1) as the seed………………………………………………………………………………………43
Figure 3.1. *ARRB2* gene map……………………………………………………………52
Figure 3.2. *ARRB2* mRNA expression………………………………………………..59-60
Figure 3.3. Allelic *ARRB2* mRNA expression ratios……………………………...62-63
Figure 3.4. Western Blot………………………………………………………………...65
Figure 3.5. Quantitative detection of *ARRB2* mRNA splice variants…………………67
Figure 3.6. LD plot of the *ARRB2* genomic locus…………………………………..72
Figure 4.1. *DRD3* gene map…………………………………………………………….77
Figure 4.2. Allelic *DRD3* mRNA expression ratios…………………………………….81
Figure 4.3. Quantitative detection of *DRD3* mRNA splice variants…………………..83
Chapter 1: Introduction

1.1 Drug addiction and genetics

Drug addiction is a chronic disorder that is defined by compulsive drug seeking and drug use, even in the face of dire consequences including the loss of job, reputation, home, and family. In addition to their serious negative effects on individuals and families, including failure in work and school, domestic abuse, and other crimes, drug abuse and addiction are a major economic burden to society: Recent estimates for the annual overall costs of substance abuse including lost productivity as well as health- and crime-related costs were $181 billion for illicit drugs (Office of National Drug Control Policy, 2004), $168 billion for tobacco (Centers for Disease Control and Prevention, 2005), and $185 billion for alcohol (Harwood, 2000). By comparison, cancer and diabetes were estimated to carry a cost of $172 billion (American Cancer Society, 2003) and $132 billion (Petersen, 2003), respectively, underscoring the magnitude of the economic impact of drug addiction on society.

It is well documented that the influence on behavior of both addictive drugs and natural rewards is due to their effect of causing an increase in synaptic dopamine levels in the nucleus accumbens (NAc) in the ventral striatum (Wise and Bozarth, 1987; Koob and
Bloom, 1988; Di Chiara, 1998; Wise, 1998). The NAc is a part of the mesolimbic dopamine pathway, which originates in the ventral tegmental area (VTA) with projections to the amygdala, hippocampus, and prefrontal cortex (PFC) as well as NAc. Although the mechanisms differ between drug classes (Johnson and North, 1992; Jones et al., 1998; Tapper et al., 2004; Waldhoer et al. 2004; Justinova et al., 2005), all addictive drugs activate this pathway, which has long been postulated to be critical for their reinforcing properties. Dopamine signaling in the mesolimbic pathway has been intensely studied for its role in addiction, as have many of the genes encoding receptors, enzymes, and other proteins thought to be important for dopaminergic signaling and regulation. Despite these efforts, the genetic mechanisms underlying susceptibility have remained elusive.

Genetic factors contribute to the susceptibility to drug addiction, with multiple candidate genes (Hiroi and Agatsuma, 2005; Howard et al., 2002; Kreek et al., 2004; Tyndale, 2003). Epidemiological and twin studies have established that 40-60% of an individual’s risk for addiction, whether it is to alcohol, opiates or cocaine, is genetic (Kendler et al., 1994; Kender et al., 2000; Prescott and Kendler, 1999; Tsuang et al., 1998). Genetic association studies have identified a number of plausible candidate genes harboring variants associated with susceptibility for addiction, although such associations are often not replicated in follow-up studies (Kreek et al., 2004; Kreek et al., 2005; Le Foll et al., 2009). A limitation of many phenotype-genotype association studies is a lack of consideration of the mechanisms underlying the effects of polymorphisms,
underscoring the importance of identifying functional polymorphisms in candidate genes and characterizing their molecular genetic effects.

1.2 The role of \textit{cis}-acting regulatory polymorphisms in human genetic variation

Many association studies have focused on polymorphisms in protein coding regions that alter the amino acid sequence, though recent studies indicate that regulatory polymorphisms in non-coding regions occur more frequently and have yet to be discovered (Bray et al., 2003; Johnson et al., 2005; Rockman and Wray, 2002; Wray, 2007; Yan et al., 2002). Since functional variants altering gene function and directly contributing to susceptibility are generally not known, the reported associations between phenotypes and a polymorphism typically rely on a marker SNP in linkage disequilibrium (LD) with the causal variant. As allele frequencies and LD patterns may vary widely between populations, results may vary between association studies, underscoring the importance of identifying functional genetic variants in candidate genes.

Regulatory polymorphisms that do not affect amino acid sequence may influence gene function through altered genetic regulatory regions and stability and processing of mRNA. The existence of regulatory variants affecting transcription (especially in the promoter region) has long been recognized (Callahan and Balbinde, 1970; Roberts, 1969). As numerous proteins interact with heteronuclear RNA and mRNA at multiple sites, \textit{cis}-regulatory polymorphisms can enhance or inhibit transcription through the modification of regulatory sites or structure of regulatory elements. Furthermore, the folding of single stranded mRNA is readily altered by SNP substitutions, with subsequent
effects on mRNA stability and processing (Johnson et al., 2008b, Johnson et al., 2005). Alternative splicing represents another important source of protein diversity and regulation that can be affected by genetic polymorphisms and is discussed in more detail below. Transcription and mRNA processing are also affected by noncoding RNAs, the mRNA docking sites for which may be altered by genetic polymorphisms (Saunders et al., 2007). Although largely undiscovered, regulatory polymorphisms are relatively common and estimated to be present in nearly all genes (Rockman and Wray, 2002).

Regardless of the mechanism, cis-acting regulatory variants share the common feature of affecting the amount and nature of mRNA generated from one allele versus the other. One powerful approach to identifying cis-acting polymorphisms relies on allelic expression imbalance (AEI) as the phenotypic measure (Johnson et al., 2008b, Johnson et al., 2005): Using a heterozygous marker SNP in the transcribed region of the gene of interest, the mRNA specifically generated from each allele is measured (Fig. 1.1). Since both alleles are subject to the same trans-acting factors, any imbalance in allelic expression must be due to cis-acting factors. Upon identifying a gene that displays significant AEI, the gene locus is scanned for candidate SNPs associated with AEI. This powerful approach has already been successful in identifying regulatory cis-acting variants in a number of genes (Johnson et al., 2009; Lim et al., 2007; Wang et al., 2008a; Wang et al., 2005; Zhang et al., 2007; Zhang et al., 2005), including some with clinical associations (Johnson et al., 2009; Wang et al., 2008a; Zhang et al., 2007).
Figure 1.1. **Quantitative allele-specific expression analysis.** A segment of DNA or RNA (cDNA) surrounding a heterozygous marker SNP is amplified by PCR. The PCR products are subjected to a single base extension reaction (SNaPshot), which adds a single fluorophore-labeled (blue or green, in this case) nucleotide complementary to the marker SNP. Each peak represents a distinct allele, and the peak heights are proportional to the amount of PCR product specifically generated for each allele.

1.3 Epigenetic Factors

Changes in gene function that are not caused by variation in DNA sequence are referred to as epigenetic. Well characterized examples of epigenetic mechanisms include methylation of DNA in regions with high CG content (CpG islands) and chromatin remodeling. CpG methylation attracts binding proteins which in turn recruit additional
specialized proteins that can alter chromatin structure through methylation, acetylation, and phosphorylation (Chen et al., 2003; Laird, 2005). Chromatin structure plays an important role in the regulation of gene expression by restricting or permitting access of various enzymes to the DNA, either obstructing or enabling transcription, respectively (Wolffe and Matzke, 1999; Workman and Kingston, 1998). Genetic and epigenetic factors are not independent of each other and their effects on gene regulation must be considered together. For example, promoter polymorphisms may affect CpG methylation and lead to allele-specific changes in methylation, as in the case of a G/A promoter SNP in SLC12A6, which introduces an additional methylation site on the opposite strand when the rare G allele is present (Moser et al., 2009). Furthermore, the authors showed the G allele is associated with decreased gene expression (Moser et al., 2009). Another example is MAOA, which has a variable number tandem repeat (VNTR) repeat polymorphism in the promoter region that has been associated with a number of psychiatric disorders (Herman et al., 2005; Huang et al., 2004; Manuck et al., 2000). In this case the two different alleles are 3- and 4-repeats, with large allelic differences in CpG methylation that correlate with AEI ratios in female samples, independent of X-inactivation (Pinsonneault et al., 2006). Moreover, CpG methylation can affect the rate of transcription which could in turn affect mRNA processing (including splicing) since the proteins involved may be present in limiting supply.
1.4 Cis-Acting Genetic Regulation of Alternative Splicing

Alternative mRNA splicing is thought to be a major source of proteomic diversity in eukaryotes (Nilsen and Graveley, 2010). In humans, over 90% of genes are alternatively spliced and ~85% have a minor isoform frequency of 15% or more (Wang et al., 2008b). Furthermore, mRNA splicing is an especially important source of diversity and regulation for neuronal genes (Lee et al., 2003). There are abundant examples of cis-acting genetic variants affecting splicing with subsequent effects on human phenotypes: A novel acceptor splice site mutation in the \( PAX3 \) gene has been shown to cause aberrant splicing resulting in Waardenburg syndrome, an autosomal dominant disorder characterized by developmental defects including hearing loss (Attaie et al., 1997). In \( DRD2 \), two intronic SNPs that shift the balance of two functionally distinct isoforms modulate memory processing and dopamine pathway activity in healthy humans (Zhang et al., 2007) and are associated with cocaine abuse in Caucasians (Moyer, manuscript in preparation). Another gene where a SNP is shown to exert a robust and clinically relevant effect on alternative splicing is \( MUC1 \), the protein product of which is ligand for \textit{Helicobacter pylori}, where it plays a role in gastric carcinogenesis. An exonic SNP in \( MUC1 \) controls alternative splicing of the same gene (Ng et al., 2008) and has been associated with an increased risk of stomach cancer (Jia et al., 2010). Taken together, these results indicate an important regulatory role for genetic polymorphisms in alternative mRNA splicing.

1.5 Review of \( DRD2 \) Genotype-Phenotype Associations

\( DRD2 \) is an intensely studied gene for its potential role in drug addiction and other psychiatric disorders thought to involve dopamine dysregulation. The most intensely
studied variant is Taq1A (rs1800497), a SNP located ~ 10 kb upstream of \textit{DRD2} that was recently shown to be in the coding region of another gene, namely \textit{ANKK1} (Neville et al., 2004). Approximately 100 association studies have assessed the linkage of Taq1A to clinical phenotypes including schizophrenia, addiction, and antipsychotic treatment response, with results that are not always replicable (for reviews see Le Foll et al., 2009; Munafo et al., 2004; Munafo et al., 2007; Thelma et al., 2008; Zhang et al., 2010). Of special interest are studies reporting associations of Taq1A with addiction to various drugs, including cocaine (Noble et al., 1993), alcohol (Smith et al., 2008), heroin (Lawford et al., 2000), and nicotine (Huang et al., 2009). Another frequently studied \textit{DRD2} SNP is the -141C Ins/Del promoter polymorphism (rs1799732), which has been associated with schizophrenia (Arinami et al., 1997) and alcoholism (Ishiguro et al., 1998). Another promoter SNP, rs12364283, was shown to be predictive of avoidance learning in healthy humans (Frank and Hutchison, 2009). The synonymous SNP C957T (rs6277), located in exon 7, was reported to be associated with schizophrenia (Lawford et al., 2005). In addition to the often more intensely studied promoter and exonic SNPs, a number of intronic SNPs in \textit{DRD2} have recently been associated with clinical phenotypes, most notably rs2283265 (intron 5) and rs1076560 (intron 6), associated with cocaine abuse (Moyer et al., manuscript in preparation), altered emotional control (Blasi et al., 2009), schizophrenia (Glatt et al., 2009) and schizophrenia-related phenotypes, (Bertolino et al., 2009b), alcoholism (Sasabe et al., 2007), opiate addiction (Doehring et al., 2009), and relative avoidance to reward-seeking behavior (Frank and Hutchison, 2009). Both SNPs have been shown to modulate DRD2 alternative splicing (Zhang et al.,
2007; Moyer et al., manuscript in preparation) and are part of a large haplotype block that includes a number of other SNPs with clinical associations (Table 2.7).

1.6 Review of *ARRB2* Genotype-Phenotype Associations

Genetic variants in *ARRB2* have been associated with psychiatric disorders, including drug addiction (Ikeda et al., 2007; Sun et al., 2008), attention deficit hyperactivity disorder (Brookes et al., 2006) and tardive dyskinesia (Liou et al., 2007). Furthermore, *ARRB2* has also been implicated in altered response to drug therapy: Cancer patients who did not respond to morphine for pain management and/or who experienced intolerable adverse effects were found to be less likely to carry the minor allele of the synonymous *ARRB2* SNP rs1045280 than patients who responded to morphine (Ross et al., 2005).

1.7 Review of *DRD3* Genotype-Phenotype Associations

Like *DRD2*, *DRD3* has been intensely studied for its potential role in disorders thought to involve dopamine dysregulation. One *DRD3* SNP in particular, the nonsynonymous SNP rs6280 (also called Ser9Gly or BallI), has received the most attention. The rs6280 variant has been associated with nicotine dependence (Huang et al., 2008), tardive dyskinesia (Lerer et al., 2002; Steen et al., 1997), schizophrenia (Jonsson et al., 2003), familial essential tremor (Jeanneteau et al., 2006), and antipsychotic treatment response (Lane et al., 2005; Scharfetter et al., 1999; Szekeres et al., 2004). The Gly-9 variant of rs6280 has an increased affinity for dopamine in vitro (Jeanneteau et al., 2006), evidence that the amino acid substitution causes a functional change to the protein that could account for
some of the clinical associations, although other genetic factors could contribute as well. Clinical associations for other DRD3 SNPs have been reported, including rs167771 associated with autism spectrum disorder (de Krom et al., 2009) and risperidone-induced extrapyramidal symptoms (Gasso et al., 2009). A study that assessed the contribution of a number of SNPs across the DRD3 gene locus in two independent populations identified a common haplotype spanning most of the gene that was associated with schizophrenia (Talkowski et al., 2006).

1.8 Summary of the Main Studies

The dopamine receptor D2 (encoded by DRD2) is implicated in susceptibility to mental disorders and cocaine abuse, but mechanisms responsible for this relationship remain uncertain. DRD2 mRNA exists in two main splice isoforms with distinct functions: D2 long (D2L) and D2 short (D2S, lacking exon 6), expressed mainly post-synaptically and pre-synaptically, respectively. Two intronic single nucleotide polymorphisms (SNPs rs2283265 (intron 5) and rs1076560 (intron 6)) in high linkage disequilibrium with each other have been reported to alter D2S/D2L splicing and several behavioral traits in human subjects, such as memory processing. To assess the role of DRD2 variants in cocaine abuse, I measured levels of D2S and D2L mRNA in human brain autopsy tissues (prefrontal cortex and putamen) obtained from cocaine abusers (n=119) and controls (n=95), and genotyped a panel of DRD2 SNPs in these tissues. Robust effects of rs2283265 and rs1076560 on D2S/L splicing were confirmed, with additive effects for each minor allele in reducing D2S. The minor alleles of rs2283265/rs1076560 were considerably more frequent in Caucasians (18%) compared to African Americans (7%).
Also, in Caucasians, rs2283265/rs1076560 minor alleles were significantly overrepresented in cocaine abusers compared to controls (rs2283265: 25% to 9%, respectively; p=0.001; OR=3.4 (1.7-7.1)), with possible additive effects for the minor allele (minor allele homozygous: p=0.052; OR=8.3 (0.98-70)). My results confirm the role of rs2283265/rs1076560 in D2 alternative splicing and support a strong role in susceptibility to cocaine abuse.

βArrestin2 (encoded by *ARRB2*) is an important regulator of most G-protein coupled receptors that also initiates its own distinct signaling pathways. Genetic variants of *ARRB2* have been implicated in psychiatric disorders, but the underlying mechanisms remain unknown. One particular synonymous SNP located in exon 11, rs1045280, has a number of clinical associations and is an ideal candidate marker SNP for analysis of allelic expression. To assess the effects of rs1045280 on *ARRB2* expression and splicing, I measured allelic and total mRNA expression and the levels of the two major *ARRB2* splice isoforms in human brain autopsy PFC from Alzheimer’s patients (n=26) and normal controls (n=20) as well as in cultured human lymphocytes. In the Alzheimer’s samples, I also assessed βarrestin2 protein levels by Western blotting. The presence of significant AEI in a number of Alzheimer’s and lymphocyte samples suggest that rs1045280 is in LD with a regulatory variant or variants causing increased transcription, and rs1045280 was associated with increased expression of overall *ARRB2* mRNA in the Alzheimer’s cohort (p<0.01), although there was no correlation between rs1045280 genotype and βarrestin2 protein levels in the Alzheimer’s samples. Alternative splicing did not differ across tissues or between samples, with the major full-length transcript
representing the overwhelming majority of \textit{ARRB2} mRNA in all three cohorts. My results indicate that rs1045280 is in LD with a regulatory variant exerting a small but significant effect on \textit{ARRB2} mRNA expression that is not maintained at the protein level.

The dopamine receptor D3 (encoded by \textit{DRD3}) is implicated in susceptibility to mental disorders including schizophrenia and addiction, but the mechanisms responsible for remain uncertain. \textit{DRD3} mRNA exists in two main splice isoforms: D3 long (D3L), encoding the functional receptor, and the truncated D3nf, which may have a dominant negative effect. \textit{DRD3} has high homology with other dopamine receptor genes and is expressed at lower levels, presenting challenges for the study of its RNA species and necessitating that assessment of DRD3 mRNA expression and alternative splicing be performed in tissues where it is most highly expressed. Using AEI as the phenotypic measure, I searched for regulatory variants in DRD3 affecting mRNA expression in human brain autopsy nucleus accumbens (NAc) from a cohort of normal controls and patients diagnosed with schizophrenia, bipolar disorder, or depression (n=15 in each group). Additionally, I assessed alternative splicing of \textit{DRD3} in the same samples. The presence of significant AEI in 4 out of 29 samples tested suggests the presence of a \textit{cis}-acting regulatory variant or variants with a modest effect on mRNA expression. However, none of the 11 SNPs genotyped were associated with the observed AEI. Furthermore, D3nf mRNA was not detectable in any of the samples. My results indicate that there is at least one \textit{cis}-acting regulatory variant with a small but significant effect on \textit{DRD3} mRNA expression that is yet to be identified. The absence of measureable D3nf mRNA suggests that it is not expressed in the NAc samples tested.
Chapter 2: Intronic Polymorphisms Affecting Alternative Splicing of Human D2 Receptor are Associated with Cocaine Abuse

2.1 Introduction

The dopamine receptor D2 (encoded by *DRD2*) is a member of the D2-like family of dopamine receptor subtypes, which are coupled to \( G_i \) inhibitory G-proteins and generally reduce formation of intracellular cAMP when activated (Civelli et al., 1993; Strange, 1993). D2 receptors are robustly expressed in the striatum and prefrontal cortex (PFC), areas involved in the primary reinforcing effects of drugs of abuse (Hyman et al., 2006) and cognitive processes (Seamans and Yang, 2004), respectively. The involvement of D2 receptors with addiction and cognition is well documented. In rodent models, response to drugs of abuse, including cocaine, ethanol, nicotine, and morphine, is disrupted by administration of D2 agonists or antagonists, and in transgenic mice lacking D2 (reviewed by Le Foll et al., 2009). In human cocaine abusers, D2 receptor availability and activity in frontal brain regions is reduced, even after months of detoxification (Volkow et al., 1993).
*DRD2* has been linked to drug addiction, although associations with specific genetic variants are not always replicable (for recent review see Le Foll et al., 2009). Association studies have typically focused on markers in protein coding regions used, in part, for historical reasons, whereas recent studies indicate that regulatory polymorphisms in non-coding regions – most as yet undiscovered – occur more frequently (Johnson et al., 2005; Rockman and Wray, 2002; Wray, 2007). Since functional variants contributing to disease are generally unknown for multigenic non-Mendelian disorders, such as drug addiction, reported associations between the disease and a polymorphism typically rely on marker SNPs in linkage disequilibrium (LD) with the causal variant. As allele frequencies and LD patterns between surrogate marker SNPs and the responsible functional variants may differ across ethnic populations, results obtained using surrogate markers may vary between association studies, underscoring the importance of identifying functional genetic variants in candidate genes.

One important mechanism for modulating D2 signaling is alternative splicing of exon 6, resulting in long (D2L) and short (D2S) isoforms expressed mainly postsynaptically and presynaptically, respectively (Khan et al., 1998). Studies in mice lacking D2L but not D2S have shown that D2S functions primarily as an autoreceptor inhibiting D1 dopamine receptor-mediated functions, whereas post-synaptic D2L receptors can act synergistically with D1 receptors (Usiello et al., 2000). Further studies in mice lacking D2L have identified altered dopamine-dependent modulation of GABA and glutamate transmission in the striatum (Centonze et al., 2003; Centonze et al., 2004) and altered behavioral responses to morphine (Smith et al., 2002) and psychotic and
antipsychotic agents (Xu et al., 2002). In postmortem human striatum and PFC, we have previously characterized two \( DRD2 \) intronic single nucleotide polymorphisms (SNPs rs1076560 and rs2283265, referred to hereafter as SNP1 and SNP2, existing in high LD with each other (\( D' = 1.0 \)), which were associated with decreased relative expression of D2S mRNA, and a promoter SNP (rs12364283) associated with enhanced \( DRD2 \) mRNA expression (Zhang et al., 2007). In the same study, SNP1/SNP2 also reduced cognitive performance in healthy humans on working memory and attentional control tasks (Zhang et al., 2007). Both SNPs modulate putative splice factor binding sites, and both were found to independently affect splicing \( \textit{in vitro} \) using a minigene construct (Zhang et al., 2007). Interestingly, the extensively studied rs1800497 (also known as Taq1A) did not appear to contribute directly to \( DRD2 \) mRNA expression or splicing, but the Taq1A1 allele is in LD with the minor alleles of SNP1/SNP2 (\( D' = 0.855 \), Zhang et al., 2007), providing a potential mechanistic basis for the clinical associations observed with Taq1A. Other recent studies have linked SNP1 and SNP2 to additional human phenotypes. Individually, SNP1 genotype has been shown to interact with dopamine transporter (\( DAT \)) \( 3' \)-variable number tandem repeat (VNTR) genotype to modulate response of the working memory network, including the striatum and PFC (Bertolino et al., 2009a). SNP1 has also been associated with altered emotional control (Blasi et al., 2009), schizophrenia-related phenotypes, (Bertolino et al., 2009b), and reduced binding of DAT and D2 receptor radiotracers in human striatum (Bertolino et al., 2010). Furthermore, SNP1 has been associated with alcoholism (Sasabe et al., 2007), opiate addiction (Doehring et al., 2009), and apomorphine-induced growth hormone response in alcohol
dependent patients (Lucht et al., 2010), and SNP2 has been associated with schizophrenia (Glatt et al., 2009). Both SNPs have been linked to relative avoidance to reward-seeking behavior (Frank and Hutchison, 2009). Taken together, these results support the notion that SNP1/SNP2 modulate the D2S/L ratio in striatum and PFC, which in turn affects dopamine signaling and related phenotypes in humans. The functional effect of SNP1/SNP2 could account for the phenotypic associations of other DRD2 SNPs in high LD with SNP1/SNP2.

Epistatic interaction between genetic loci is thought to be an important factor in assessing the genetic contribution to disease risk, and is likely to be especially important in complex multigenic disorders like drug addiction. The risk associated with a given genetic locus may be influenced by the genotype at another locus, confounding attempts to detect phenotypic associations in a single locus analysis (Carlson et al., 2004). However, the sheer number of possible interactions between genetic variants presents a challenge to identifying real epistatic interactions with biological consequences. One approach to this problem is to limit the analysis to putative functional variants in genes in related pathways (the dopamine pathway in this case). Encoding the molecular target of cocaine, the DAT gene represents another important candidate gene for addiction. Importantly, DAT physically interacts with D2 receptors in vivo and a statistical interaction between SNP1 genotype and DAT 3’ variable number tandem repeat (VNTR) genotype was detected and shown to modulate the response of the working memory network in humans (Bertolino et al., 2009a). In light of the evidence for interactions between DRD2 and DAT, we asked whether there is a significant interaction of
SNP1/SNP2 with the DAT 3’-VNTR or other DAT variants with evidence for being functional, including the intron 8 VNTR (Guindalini et al., 2006) and rs6347 and rs27072 (Pinsonneault, manuscript in preparation).

Because of the importance of D2 alternative splicing, one goal of the current study was to replicate the relationship between SNP1/SNP2 and D2 splicing in a different set of brain tissues. Moreover, we searched for additional genetic variants that could affect D2 splicing, because most, but not all of the observed splicing variability for D2 could be accounted for by SNP1/SNP2 (see Zhang et al. 2007, Fig. 3a). In addition, we assessed the relationship between D2S/L splicing, SNP1/SNP2 genotype, and cocaine abuse status in human brain autopsy PFC and putamen tissues from chronic cocaine abusers who died from cocaine overdose and age-matched drug-free controls. We also tested for interaction of SNP2 with four putative functional DAT SNPs and identified a significant interaction with the DAT intron 8 VNTR. While we did not identify any additional polymorphisms between exon 5 and exon 7 affecting D2S/L splicing, we did confirm the role of SNP1/SNP2 in D2S/L splicing and identified both SNPs as potential risk factors for cocaine abuse.

2.2 Materials and Methods

Postmortem Human Brain Tissues.

Miami Brain Bank. PFC (Brodmann’s area 46) and putamen tissue samples from 119 cocaine abusers who died from cocaine intoxication and 95 age-matched drug-free controls were provided by the Miami Brain Endowment Bank (University of Miami, Miami, FL). Supplemental brain and blood toxicology and neuropathologic evaluations
were done in every case, and most of the drug-exposure cases obtained at autopsy came from individuals who had evidence of a number of surrogate measures of chronic cocaine abuse (drug-related pathology, arrest records, and hospital and treatment admissions). Drug-free age-matched control subjects were selected from accidental or cardiac sudden deaths with negative urine screens for all common drugs and there was no history of psychiatric disorders or licit or illicit drug use prior to death.

Stanley Array Collection. DNA and RNA extracted from postmortem brain tissues were donated by The Stanley Medical Research Institute's (Chevy Chase, MD) brain collection. We obtained genomic DNA and total RNA extracted from 105 individuals previously diagnosed with schizophrenia or bipolar disorder (35 of each) and 35 controls.

DNA and RNA Extraction. DNA and RNA were extracted as described in Zhang et al. 2005. cDNA was synthesized with SuperScript™ III (Life Technologies, Foster City, CA) using both gene-specific primers and oligo(dT), except for the minigene experiment (see below).

DNA Sequencing. DNA sequencing of PCR product amplified from genomic DNA was performed with a 3730 DNA analyzer using the BigDye® Terminator Cycle Sequencing kit (Life Technologies, Foster City, CA).

Cell Culture and Minigene Splicing. Human embryonic kidney (HEK) 293 cells were cultured in DMEM/F12 media containing 10% FBS and 1% penicillin-streptomycin under 5% CO₂. We used the Stratagene QuikChange Lightning site-directed mutagenesis kit (Agilent Technologies, Santa Clara, CA) to introduce rs35608204 into a DRD2 minigene consisting of exons 5-7 and introns 5 and 6 in pcDNA3 (Zhang et al. 2007).
The constructs were sequenced to confirm the intended genotypes and transfected into HEK 293 cells using Lipofectamine 2000 (Life Technologies, Foster City, CA). RNA was extracted 24 hr later and cDNA synthesized using only a plasmid-specific primer, SP6, to avoid synthesis of endogenous DRD2 cDNA. Splice variants were assayed as described below.

**Quantitative mRNA Analysis by Real-Time PCR.** Real-time PCR was performed on cDNA samples on an ABI 7000 sequence detection system with Power SYBR® Green PCR Master Mix (Life Technologies, Foster City, CA). DRD2 expression levels, in arbitrary units, were calculated by subtracting the β-actin cycle threshold (Ct) from the DRD2 Ct to get a ΔCt as described previously (Pinsonneault et al., 2006). Arbitrary units for each sample = 100(2^-ΔCt).

**Quantitative Detection of Splice Isoforms.** D2S and D2L were measured following PCR amplification of cDNA with a fluorescent Fam-labeled exon 5 primer and an exon 7 reverse primer as described in Wang et al., 2006. Standard curves were prepared with varying mixtures of cloned D2S and D2L cDNA (Zhang et al., 2007).

**Genotyping Methods.** SNPs were genotyped by at least one of four methods: SNPlex (Life Technologies, Foster City, CA), SNaPshot (Life Technologies, Foster City, CA) (Zhang et al. 2005), and allele-specific primers as described in Papp et al., 2003. The fourth method was a modified PCR-restriction fragment length polymorphism (RFLP) method (Haliassos et al., 1989a; Haliassos et al., 1989b). For this method, one of the PCR primers was Fam-labeled, allowing for resolution of the restriction fragments by
capillary electrophoresis with an Applied Biosystems 3730 DNA analyzer (Life Technologies, Foster City, CA).

**Statistical Analysis.** HelixTree 6.4.3 (Golden Helix, Bozeman, MT) was used to test for Hardy-Weinberg equilibrium, calculate LD and $R^2$ values, haplotype estimation, and a basic allele chi-squared test for association with cocaine status. One-way ANOVA and student’s t tests were performed with GraphPad Prism software (GraphPad Software, La Jolla, CA). The associations between genotypes or haplotypes and disease risk as well as the contribution of each minor allele to disease risk were analyzed using logistic regression model performed using SAS 9.1 software (SAS Institute Inc., Cary, NC). The suitability of model fitting was judged by deviance goodness of fit statistics $p$ value and score test $p$ value, both of which should be equal to or below 0.05.

### 2.3 Results

**Search for Additional Polymorphisms Associated with Alternative Splicing.** Based on results of earlier experiments measuring allelic mRNA ratios of D2S and D2L in human postmortem tissues (Zhang et al., 2007), we selected five samples homozygous for the major alleles of SNP1/SNP2 for further analysis. Three of the samples had displayed substantial variation in the allelic mRNA ratio of D2S to D2L, indicative of other variants potentially affecting splicing, while two samples lacking allelic differences served as controls. An approximately 3 kb region around exon 6 was sequenced (Fig. 2.1), and the results analyzed for additional polymorphisms associated with differential alternative splicing. One control sample was heterozygous for rs12363125, which was excluded as potentially affecting splicing. One sample with an allelic D2L/D2S ratio $>$1...
was a heterozygous carrier for the minor allele of rs35608204 (C/T genotype), located in intron 5 (Table 2.1), suggesting a possible effect of rs35608204 on D2 splicing. We did not identify any other sequence differences in the samples.

![Figure 2.1. DRD2 gene map, with blown up view of the region that was sequenced.](image)
The exon that is included in D2L but absent in D2S (exon 6) is in gray. The locations of SNPs genotyped in this study are marked.
Table 2.1. Results of sequencing DNA from five samples that displayed different D2L/D2S allelic ratios

<table>
<thead>
<tr>
<th>Sample ID (with AEI (L) / AEI (S) ratio)</th>
<th>Low 1 (0.74)</th>
<th>Low 2 (0.74)</th>
<th>High (1.67)</th>
<th>Control 1 (0.99)</th>
<th>Control 2 (0.86)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs1110977</td>
<td>AA</td>
<td>AA</td>
<td>AA</td>
<td>AA</td>
<td>AA</td>
</tr>
<tr>
<td>rs71653614</td>
<td>AA</td>
<td>AA</td>
<td>AA</td>
<td>AA</td>
<td>AA</td>
</tr>
<tr>
<td>rs6277</td>
<td>CT</td>
<td>CT</td>
<td>CT</td>
<td>CT</td>
<td>CT</td>
</tr>
<tr>
<td>rs6275</td>
<td>CT</td>
<td>CT</td>
<td>CT</td>
<td>CT</td>
<td>CT</td>
</tr>
<tr>
<td>rs1801028</td>
<td>CC</td>
<td>CC</td>
<td>CC</td>
<td>CC</td>
<td>CC</td>
</tr>
<tr>
<td>rs1800496</td>
<td>CC</td>
<td>CC</td>
<td>CC</td>
<td>CC</td>
<td>CC</td>
</tr>
<tr>
<td>rs4986921</td>
<td>CC</td>
<td>CC</td>
<td>CC</td>
<td>CC</td>
<td>CC</td>
</tr>
<tr>
<td>rs1076561</td>
<td>AA</td>
<td>AA</td>
<td>AA</td>
<td>AA</td>
<td>AA</td>
</tr>
<tr>
<td>rs1076560</td>
<td>GG</td>
<td>GG</td>
<td>GG</td>
<td>GG</td>
<td>GG</td>
</tr>
<tr>
<td>rs4986920</td>
<td>CC</td>
<td>CC</td>
<td>CC</td>
<td>CC</td>
<td>CC</td>
</tr>
<tr>
<td>rs34684460</td>
<td>no deletion</td>
<td>no deletion</td>
<td>no deletion</td>
<td>no deletion</td>
<td>no deletion</td>
</tr>
<tr>
<td>rs1079721</td>
<td>GG</td>
<td>GG</td>
<td>GG</td>
<td>GG</td>
<td>GG</td>
</tr>
<tr>
<td>rs1110976</td>
<td>AA</td>
<td>AA</td>
<td>AA</td>
<td>AA</td>
<td>AA</td>
</tr>
<tr>
<td>rs1110975</td>
<td>CC</td>
<td>CC</td>
<td>CC</td>
<td>CC</td>
<td>CC</td>
</tr>
<tr>
<td>rs2362873</td>
<td>CC</td>
<td>CC</td>
<td>CC</td>
<td>CC</td>
<td>CC</td>
</tr>
<tr>
<td>rs55900980</td>
<td>no deletion</td>
<td>no deletion</td>
<td>no deletion</td>
<td>no deletion</td>
<td>no deletion</td>
</tr>
<tr>
<td>rs35319840</td>
<td>no insertion</td>
<td>no insertion</td>
<td>no insertion</td>
<td>no insertion</td>
<td>no insertion</td>
</tr>
<tr>
<td>rs72153544</td>
<td>no insertion</td>
<td>no insertion</td>
<td>no insertion</td>
<td>no insertion</td>
<td>no insertion</td>
</tr>
<tr>
<td>rs34525623</td>
<td>no insertion</td>
<td>no insertion</td>
<td>no insertion</td>
<td>no insertion</td>
<td>no insertion</td>
</tr>
<tr>
<td>rs2511521</td>
<td>TT</td>
<td>TT</td>
<td>TT</td>
<td>TT</td>
<td>TT</td>
</tr>
<tr>
<td>rs2283265</td>
<td>GG</td>
<td>GG</td>
<td>GG</td>
<td>GG</td>
<td>GG</td>
</tr>
<tr>
<td>rs57837178</td>
<td>CC</td>
<td>CC</td>
<td>CC</td>
<td>CC</td>
<td>CC</td>
</tr>
<tr>
<td>rs56014166</td>
<td>CC</td>
<td>CC</td>
<td>CC</td>
<td>CC</td>
<td>CC</td>
</tr>
<tr>
<td>rs35608204</td>
<td>TT</td>
<td>TT</td>
<td>CT</td>
<td>TT</td>
<td>TT</td>
</tr>
<tr>
<td>rs35579712</td>
<td>no deletion</td>
<td>no deletion</td>
<td>no deletion</td>
<td>no deletion</td>
<td>no deletion</td>
</tr>
<tr>
<td>rs12363125</td>
<td>AG</td>
<td>AG</td>
<td>AG</td>
<td>AA</td>
<td>AG</td>
</tr>
<tr>
<td>rs34674946</td>
<td>no insertion</td>
<td>no insertion</td>
<td>no insertion</td>
<td>no insertion</td>
<td>no insertion</td>
</tr>
<tr>
<td>5’rs71653613</td>
<td>CC</td>
<td>CC</td>
<td>CC</td>
<td>CC</td>
<td>CC</td>
</tr>
</tbody>
</table>

All SNPs (according to the dbSNP database) within the region sequenced are listed along with the observed genotype in each sample. The only two positions where the samples differed from each other are indicated by the shaded boxes. The rs6277 and rs6275 SNPs were used as marker SNPs for the allelic expression assay and therefore are heterozygous in these samples because only heterozygous samples were selected.
**Alternative Splicing of a *DRD2* Minigene Carrying rs35608204.** We introduced the minor ‘C’ allele SNP for rs35608204 into a partial gene construct containing exon 5 to exon 7 and carrying the major alleles of SNP1 and SNP2 in an expression vector (Zhang et al. 2007). This construct along with two others, one carrying the minor alleles of both SNP1 and SNP2 with the major allele of rs35608204 and a reference construct carrying the major alleles of all three SNPs, were transfected separately into HEK 293 cells. We extracted RNA from the cells, made cDNA using a plasmid-specific primer, and measured relative expression of D2S and D2L mRNA (cDNA), using PCR with fluorescently labeled primers. As shown previously by Zhang et al. 2007, the minor alleles of SNP1/SNP2 generated significantly less D2S relative to D2L (Fig. 2.2). However, rs35608204 did not affect D2S/L splicing (Fig. 2.2). These results confirm that the minor T alleles of SNP1/SNP2 reduce the formation of D2S relative to D2L but do not support a role for rs35608204 in D2S/L splicing regulation.
Figure 2.2. Alternative splicing from \textit{DRD2} minigenes in HEK 293 cells. Minigenes carrying the indicated genotypes were transfected into HEK 293 cells and D2S relative to D2L expression levels measured. Minor alleles are indicated by bold type. Data are mean ± SEM (n = 3 separate transfections). *, P<0.05, one-way ANOVA with Bonferroni post-test, compared to GGT and GGC.

\textit{DRD2} mRNA Expression in Human PFC and Putamen from Cocaine Abusers and Controls. We measured overall \textit{DRD2} mRNA expression with real time PCR in postmortem human PFC and putamen from cocaine abusers and age-matched controls. We did not observe a statistically significant difference in overall \textit{DRD2} mRNA
expression between cocaine abusers and controls in either brain region in the complete cohort (Fig. 2.3) or in Caucasians or African Americans when analyzed separately (Fig. 2.4). No SNP tested correlated with overall DRD2 mRNA levels (data not shown); any potential small effects likely masked by large interindividual variability in total mRNA levels in autopsy brain tissues.

Figure 2.3. DRD2 mRNA expression. Real-time PCR was applied to quantify mRNA, with β-actin as an internal control. DRD2 mRNA levels are presented in arbitrary units (see Materials and Methods). Data are mean ± SEM, single measurements in PFC and duplicate measurements in putamen.
Figure 2.4. \textit{DRD2} mRNA expression grouped by race and cocaine abuse status. Real-time PCR was applied to quantify mRNA, with $\beta$-actin as an internal control. \textit{DRD2} mRNA levels are presented in arbitrary units (see Materials and Methods). Data are mean ± SEM, single measurements in PFC and duplicate measurements in putamen.

**Alternative Splicing of \textit{DRD2} mRNA in Human PFC and Putamen from Cocaine Abusers and Controls.** To assess D2S/L splicing in the context of cocaine abuse, we measured relative expression of D2S and D2L mRNA in postmortem human PFC and putamen from cocaine abusers and age-matched controls. Consistent with previous reports (Khan et al., 1998; Zhang et al., 2007), relative D2S expression was higher in PFC (27 ± 5%, n = 73) than in putamen (17 ± 4%, n = 70) ($p<0.0001$, Student’s $t$ test). Importantly, subjects carrying the minor allele T of SNP1/SNP2 had significantly reduced relative expression of D2S mRNA in both brain regions (Fig. 2.5a). This result is consistent with an earlier report (Zhang et al., 2007) and confirms a strong effect of SNP1/SNP2 on D2S/L splicing in both tissues, in a different cohort. The D2S/L ratio
was reduced to a greater extent in individuals homozygous for the minor allele than in heterozygotes, indicating that the effects of SNP1/SNP2 on D2S/L splicing are additive (Fig. 2.5a). Consistent with the results of the in vitro minigene experiments, rs35608204 was not associated with the D2S/L ratio (data not shown). No difference was observed in relative expression of D2S between cocaine abusers and controls in either brain region tested (Fig. 2.5b). Since allele frequencies and linkage patterns may differ between populations, we analyzed the relationship of SNP1/SNP2 with D2S/L splicing separately in Caucasians and African Americans and obtained similar results across ethnicity (Fig. 2.6), mirroring the combined analysis shown in fig. 2.5a, and further supporting the role of SNP1/SNP2 as functional variants affecting splicing.
Figure 2.5. Expression of D2S relative to D2L. (a) Relative D2S mRNA expression grouped by SNP1/SNP2 genotype (T is minor allele). (b) Relative D2S mRNA expression grouped by cocaine abuse status. Data are mean ± SEM, at least three measurements per sample, *, P<0.05, ***, P<0.001, one-way ANOVA with Bonferroni post-test.
Figure 2.6. Expression of D2S relative to D2L grouped by race. Relative D2S mRNA expression grouped by race and SNP1/SNP2 genotype (T is minor allele). Data are mean ± SEM, at least three measurements per sample, *, P<0.05, ***, P<0.001, one-way ANOVA with Bonferroni post-test.

Cocaine Status and DRD2 Genotype. Given the well-established role of dopamine signaling in drug addiction and the mounting evidence that SNP1/SNP2 affect the D2S/D2L balance in relevant brain regions, we tested both SNPs for association with cocaine abuse in a population of cocaine abusers (n=119) and age-matched controls (n=95) obtained from an archived biorepository housed in the Miami Brain Endowment Bank (population characteristics are provided in Table 2.2). We also genotyped the SNP we initially suspected might have an effect on D2S/L splicing, rs35608204, as well as the promoter SNP rs12364283, previously shown to alter allelic DRD2 mRNA expression (Zhang et al., 2007) and predictive of avoidance learning in healthy humans (Frank and Hutchison, 2009). In addition, we selected three other DRD2 SNPs for genotyping based
on their associations with clinical phenotypes and/or utility as haplotype tagging SNPs: rs1125394 (intron 1, Hwang et al., 2005; Hwang et al., 2006), rs4648318 (intron 1, Huang et al., 2009), and rs11214608 (5’ upstream). All SNPs tested were in Hardy-Weinberg equilibrium (p>0.05) except SNP1, SNP2, and a third SNP, rs1125394, in high LD with SNP1/SNP2, (D’=1.0, Table 2.3). This violation of Hardy-Weinberg equilibrium appears to be driven by the cocaine abuse cohort, as all three SNPs were in Hardy-Weinberg equilibrium in the controls. Genotyping errors are unlikely to be the cause, as SNP1 and SNP2 genotypes were verified with two independent methods.

We found that SNP1, SNP2, and rs1125394 were associated with cocaine abuse, with minor allele frequencies approximately twice as high in cocaine abusers as in controls (SNP2: p=0.01, OR=2.3, 95% CI 1.2 - 4.2, Table 2.4). When the data were analyzed separately in Caucasians and African Americans, association of SNP1/SNP2 with cocaine abuse was even stronger in Caucasians (SNP2: p=0.001, OR=3.4, 95% CI 1.7 – 7.1), whereas there was no significant association in African Americans, who have a considerably lower minor allele frequency (7% versus 18% in Caucasians; Table 2.4). As the splicing data suggest an additive effect of each minor allele of SNP1/SNP2, we further analyzed the number of minor alleles of SNP2 (i.e. GG=0, GT=1, TT=2) vs. cocaine status using a logistic regression model. The odds ratio (OR) of each minor allele in the complete cohort was 2.0 (95% CI 1.1 - 3.7, p=0.02; in Caucasians OR=3.0, 95% CI 1.5 - 6.0, p=0.003), suggesting the OR for minor allele homozygotes would be at least four. When we analyzed the genotype effect (i.e. SNP2 GG, GT and TT) vs. cocaine status, the association did not reach statistical significance in the complete cohort
(TT vs. GG, p=0.06, OR=7.3, 95% CI 0.90 - 60) or in the Caucasian subgroup (TT vs. GG, p=0.052, OR=8.3, 95% CI 0.98 – 70) despite the large odds ratios, possibly due to the relatively small sample size.

To assess the effect of cocaine abuse on D2S/L splicing independent of the effects of SNP1/SNP2, we analyzed the splicing data from both brain regions broken down by SNP1/SNP2 genotype and further analyzed across controls and cocaine abusers within each group. No significant differences in D2S/L splicing were detectable between controls and cocaine abusers (data not shown).

Table 2.2. Demographic Characteristics

<table>
<thead>
<tr>
<th></th>
<th>Controls</th>
<th>Cocaine Abusers</th>
</tr>
</thead>
<tbody>
<tr>
<td>n=</td>
<td>93</td>
<td>118</td>
</tr>
<tr>
<td>Age (mean ± SD)</td>
<td>35 ± 11</td>
<td>37 ± 10</td>
</tr>
<tr>
<td>% Male</td>
<td>87%</td>
<td>85%</td>
</tr>
<tr>
<td>Caucasian</td>
<td>63</td>
<td>74</td>
</tr>
<tr>
<td>African American</td>
<td>30</td>
<td>44</td>
</tr>
</tbody>
</table>

Three individuals were excluded from all analyses and are not included in this table. One was the only subject of Asian ancestry, and we did not obtain quality DNA/reliable genotyping results from the other two.
Table 2.3. Pairwise Linkage Disequilibrium Values

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th>R²</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>rs1076560</td>
<td>rs2283265</td>
<td>rs35608204</td>
<td>rs1125394</td>
<td>rs4648318</td>
<td>rs11214608</td>
<td>rs12364283</td>
</tr>
<tr>
<td>rs1076560</td>
<td>0.94</td>
<td>0</td>
<td>0.75</td>
<td>0.05</td>
<td>0.12</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>rs2283265</td>
<td>0.97</td>
<td>0.96</td>
<td>0.81</td>
<td>0.05</td>
<td>0.13</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>rs35608204</td>
<td>0.97</td>
<td>1</td>
<td>0.01</td>
<td>0.01</td>
<td>0.03</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td>rs1125394</td>
<td>0.95</td>
<td>0.99</td>
<td>0.05</td>
<td>0.18</td>
<td>0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs4648318</td>
<td>0.89</td>
<td>0.99</td>
<td>1</td>
<td>0.77</td>
<td>0.32</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>rs11214608</td>
<td>0.92</td>
<td>0.99</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs12364283</td>
<td>0.12</td>
<td>0.13</td>
<td>0.56</td>
<td>0.18</td>
<td>0.40</td>
<td>0.20</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Russians</th>
<th>R²</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>rs1076560</td>
<td>rs2283265</td>
<td>rs35608204</td>
<td>rs1125394</td>
<td>rs4648318</td>
<td>rs11214608</td>
<td>rs12364283</td>
</tr>
<tr>
<td>rs1076560</td>
<td>0.94</td>
<td>0.01</td>
<td>0.90</td>
<td>0.06</td>
<td>0.29</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>rs2283265</td>
<td>0.97</td>
<td>0.01</td>
<td>0.97</td>
<td>0.06</td>
<td>0.31</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>rs35608204</td>
<td>0.98</td>
<td>0.99</td>
<td>0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td>rs1125394</td>
<td>0.96</td>
<td>0.99</td>
<td>0.06</td>
<td>0.31</td>
<td>0.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs4648318</td>
<td>1</td>
<td>0.99</td>
<td>1</td>
<td>0.39</td>
<td>0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs11214608</td>
<td>0.94</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs12364283</td>
<td>0.09</td>
<td>0.61</td>
<td>0.22</td>
<td>0.62</td>
<td>0.26</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>D'</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Caucasians</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs1076560</td>
<td>0.93</td>
<td>0.90</td>
<td>0.06</td>
<td>0.29</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs2283265</td>
<td>0.98</td>
<td>0.97</td>
<td>0.01</td>
<td>0.02</td>
<td>0.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs35608204</td>
<td>0.96</td>
<td>0.99</td>
<td>0.06</td>
<td>0.31</td>
<td>0.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs1125394</td>
<td>1</td>
<td>1</td>
<td>0.39</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs4648318</td>
<td>0.94</td>
<td>0.99</td>
<td>1</td>
<td>1</td>
<td>0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs11214608</td>
<td>0.09</td>
<td>0.61</td>
<td>0.22</td>
<td>0.62</td>
<td>0.26</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>D'</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>African</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs1076560</td>
<td>0.93</td>
<td>0.90</td>
<td>0.06</td>
<td>0.29</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs2283265</td>
<td>0.98</td>
<td>0.97</td>
<td>0.01</td>
<td>0.02</td>
<td>0.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs35608204</td>
<td>0.96</td>
<td>0.99</td>
<td>0.06</td>
<td>0.31</td>
<td>0.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs1125394</td>
<td>1</td>
<td>1</td>
<td>0.39</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs4648318</td>
<td>0.94</td>
<td>0.99</td>
<td>0.01</td>
<td>0.04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs11214608</td>
<td>0.09</td>
<td>0.61</td>
<td>0.02</td>
<td>0.96</td>
<td>0.99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs12364283</td>
<td>1</td>
<td>1</td>
<td>0.02</td>
<td>0.06</td>
<td>0.26</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2.4. Minor allele frequencies in cocaine abusers and controls, plus results of statistical group comparisons

<table>
<thead>
<tr>
<th>SNP</th>
<th>All (n=211)</th>
<th>Controls (n=93)</th>
<th>Cocaine (n=118)</th>
<th>$\chi^2$, P Value</th>
<th>Odds Ratio (95% C.I.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs1076560 (SNP1)</td>
<td>0.14</td>
<td>0.10</td>
<td>0.17</td>
<td>0.03</td>
<td>1.94 (1.07 - 3.50)</td>
</tr>
<tr>
<td>rs2283265 (SNP2)</td>
<td>0.14</td>
<td>0.09</td>
<td>0.18</td>
<td>0.01</td>
<td>2.27 (1.23 - 4.19)</td>
</tr>
<tr>
<td>rs35608204</td>
<td>0.02</td>
<td>0.03</td>
<td>0.02</td>
<td>0.30</td>
<td>0.52 (0.14 - 1.86)</td>
</tr>
<tr>
<td>rs1125394*</td>
<td>0.17</td>
<td>0.12</td>
<td>0.22</td>
<td>0.03</td>
<td>2.05 (1.07 - 3.94)</td>
</tr>
<tr>
<td>rs4648318*</td>
<td>0.28</td>
<td>0.27</td>
<td>0.29</td>
<td>0.64</td>
<td>1.13 (0.67 - 1.89)</td>
</tr>
<tr>
<td>rs11214608*</td>
<td>0.45</td>
<td>0.51</td>
<td>0.41</td>
<td>0.09</td>
<td>0.67 (0.42 - 1.06)</td>
</tr>
<tr>
<td>rs12364283</td>
<td>0.06</td>
<td>0.05</td>
<td>0.07</td>
<td>0.53</td>
<td>1.31 (0.56 - 3.07)</td>
</tr>
</tbody>
</table>

Caucasians (n=137)

<table>
<thead>
<tr>
<th>SNP</th>
<th>All (n=137)</th>
<th>Controls (n=63)</th>
<th>Cocaine (n=74)</th>
<th>$\chi^2$, P Value</th>
<th>Odds Ratio (95% C.I.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs1076560 (SNP1)</td>
<td>0.18</td>
<td>0.10</td>
<td>0.24</td>
<td>0.003</td>
<td>2.74 (1.38 - 5.45)</td>
</tr>
<tr>
<td>rs2283265 (SNP2)</td>
<td>0.18</td>
<td>0.09</td>
<td>0.25</td>
<td>0.001</td>
<td>3.42 (1.66 - 7.05)</td>
</tr>
<tr>
<td>rs35608204</td>
<td>0.03</td>
<td>0.05</td>
<td>0.02</td>
<td>0.21</td>
<td>0.42 (0.10 - 1.71)</td>
</tr>
<tr>
<td>rs1125394*</td>
<td>0.17</td>
<td>0.09</td>
<td>0.25</td>
<td>0.004</td>
<td>3.33 (1.42 - 7.80)</td>
</tr>
<tr>
<td>rs4648318*</td>
<td>0.21</td>
<td>0.23</td>
<td>0.19</td>
<td>0.46</td>
<td>0.78 (0.39 - 1.53)</td>
</tr>
<tr>
<td>rs11214608*</td>
<td>0.41</td>
<td>0.36</td>
<td>0.45</td>
<td>0.19</td>
<td>1.46 (0.83 - 2.59)</td>
</tr>
<tr>
<td>rs12364283</td>
<td>0.08</td>
<td>0.07</td>
<td>0.09</td>
<td>0.63</td>
<td>1.25 (0.51 - 3.02)</td>
</tr>
</tbody>
</table>

African Americans (n=74)

<table>
<thead>
<tr>
<th>SNP</th>
<th>All (n=74)</th>
<th>Controls (n=30)</th>
<th>Cocaine (n=44)</th>
<th>$\chi^2$, P Value</th>
<th>Odds Ratio (95% C.I.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs1076560 (SNP1)</td>
<td>0.07</td>
<td>0.08</td>
<td>0.06</td>
<td>0.53</td>
<td>0.66 (0.18 - 2.40)</td>
</tr>
<tr>
<td>rs2283265 (SNP2)</td>
<td>0.07</td>
<td>0.08</td>
<td>0.06</td>
<td>0.53</td>
<td>0.66 (0.18 - 2.40)</td>
</tr>
<tr>
<td>rs35608204</td>
<td>0.01</td>
<td>0.00</td>
<td>0.01</td>
<td>0.42</td>
<td>N.D.</td>
</tr>
<tr>
<td>rs1125394*</td>
<td>0.18</td>
<td>0.19</td>
<td>0.17</td>
<td>0.73</td>
<td>0.83 (0.28 - 2.41)</td>
</tr>
<tr>
<td>rs4648318*</td>
<td>0.43</td>
<td>0.35</td>
<td>0.47</td>
<td>0.28</td>
<td>1.61 (0.68 - 3.82)</td>
</tr>
<tr>
<td>rs11214608*</td>
<td>0.15</td>
<td>0.17</td>
<td>0.15</td>
<td>0.78</td>
<td>0.85 (0.28 - 2.62)</td>
</tr>
<tr>
<td>rs12364283</td>
<td>0.01</td>
<td>0.00</td>
<td>0.02</td>
<td>0.25</td>
<td>N.D.</td>
</tr>
</tbody>
</table>

Minor allele frequencies were determined for all samples (n=211), controls (n=93), and cocaine abusers (n=118) separately within each group. Approximately 70% of samples were successfully genotyped with SNPlex for SNPs annotated with a *. The rest of the SNPs were successfully genotyped in >95% of samples with a number of techniques (see Materials and Methods).
**Haplotype Analysis.** We conducted a haplotype analysis for all seven SNPs and identified four major haplotypes that occurred at >2% frequency in the complete cohort (Table 2.5). Of the two haplotypes containing the minor alleles of SNP1/SNP2, the more frequent haplotype H3 was present at almost twice the frequency in cocaine abusers vs. controls (14% vs. 8%), although the association with cocaine abuse did not reach statistical significance (p=0.2). One relatively rare haplotype (H6) was also more prevalent in cocaine abusers, (4% vs. <1% in controls, Table 2.5). Interestingly, the only difference between this haplotype and the second most frequent haplotype (H2) is rs12364283, the promoter SNP previously associated with increased *DRD2* mRNA expression (Zhang et al., 2007). However, none of the haplotypes was significantly associated with cocaine abuse in the complete cohort (data not shown), requiring a larger sample size.

The haplotype distribution (Table 2.5) and LD pattern (Table 2.3) differed between Caucasians and African Americans in our cohort. The frequent haplotype carrying the minor alleles of SNP1/SNP2, H3, was significantly associated with cocaine abuse in Caucasians (p=0.049, OR=3.0, 95% CI 1.0 – 9.1). However, in African Americans it was much less frequent and tends to be overrepresented in controls (Table 2.5), although this did not approach statistical significance (p=0.6). No other haplotypes were associated with cocaine abuse in the complete cohort or in the Caucasian or African American subgroups, although the rare haplotype carrying the functional promoter SNP rs12364283 (H6) tends to be more prevalent in cocaine abusers in both Caucasians and African Americans (Table 2.5).
Table 2.5. Haplotype Analysis in Cocaine Abusers and Controls

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Genotype</th>
<th>Haplotype Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>All</td>
</tr>
<tr>
<td>H1</td>
<td>G G A A T A</td>
<td>42%</td>
</tr>
<tr>
<td>H2</td>
<td>G G A A G C</td>
<td>26%</td>
</tr>
<tr>
<td>H3</td>
<td>T T A G A C</td>
<td>11%</td>
</tr>
<tr>
<td>H4</td>
<td>G G A A A C</td>
<td>10%</td>
</tr>
<tr>
<td>H5</td>
<td>G G G A C A</td>
<td>2%</td>
</tr>
<tr>
<td>H6</td>
<td>G G A A G C</td>
<td>2%</td>
</tr>
<tr>
<td>H7</td>
<td>T T A G A C</td>
<td>1%</td>
</tr>
</tbody>
</table>

Haplotypes occurring at frequencies of at least 2% in at least two groups were included in the table.

Interaction of SNP2 with a Functional Variant in the Dopamine Transporter Gene (DAT). To test for a statistical interaction between SNP2 and each of the four DAT variants, logistic regression was performed in SNP2 major allele homozygotes and minor
allele carriers separately (allele frequencies separated by cocaine status and SNP2 genotype are shown in Table 2.6). We found a significant interaction of SNP2 with the intron 8 VNTR (OR for association with cocaine abuse in SNP2 minor allele carriers = 0.28, 95% CI 0.072 - 1.03, p=0.055). In SNP2 major allele homozygotes, OR = 2.1, 95% CI 1.1 ~ 3.9, p=0.028, indicating that the intron 8 VNTR minor allele five repeat is protective in individuals carrying the SNP2 minor allele, while increasing risk in individuals homozygous for the major allele of SNP2. No other significant interactions were identified, and none of the DAT variants were associated with cocaine abuse when tested separately.

Table 2.6. DRD2-DAT Gene Interaction

<table>
<thead>
<tr>
<th></th>
<th>Controls DRD2 SNP1 Minor Allele Carriers (n=15) MAF in This Group</th>
<th>Cocaine Abusers DRD2 SNP2 Minor Allele Carriers (n=35) MAF in This Group</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>DAT intron 8 VNTR</em></td>
<td>0.42</td>
<td>0.26</td>
</tr>
<tr>
<td><em>DAT rs6347</em></td>
<td>0.22</td>
<td>0.22</td>
</tr>
<tr>
<td><em>DAT rs727072</em></td>
<td>0.20</td>
<td>0.20</td>
</tr>
<tr>
<td><em>DAT 3 VNTR</em></td>
<td>0.20</td>
<td>0.20</td>
</tr>
<tr>
<td><em>DAT rs727072</em></td>
<td>0.20</td>
<td>0.20</td>
</tr>
<tr>
<td><em>DAT 3 VNTR</em></td>
<td>0.20</td>
<td>0.20</td>
</tr>
</tbody>
</table>

Minor allele frequencies were determined for controls (n=93), and cocaine abusers (n=118) separately within each group.

2.4 Discussion

The present study confirms a robust effect of SNP1/SNP2 on D2S/L splicing in human PFC and putamen autopsy tissues in a cohort of cocaine abusers with a documented history of heavy abuse, and age-matched controls. Sequencing the region of DRD2
between exon 5 and exon 7 suggested an additional candidate SNP, rs35608204, but a role in splicing was not corroborated by analysis of D2S/D2L mRNA in human PFC and putamen, or by DRD2 minigene experiments *in vitro*. A genetic association analysis revealed that SNP1, SNP2, and a third SNP also in high LD, rs1125394, but lacking evidence for any biological function, are associated with cocaine abuse in Caucasians but not African Americans.

The association of SNP1/SNP2 with decreased expression of D2S relative to D2L in human PFC and putamen was consistent with our earlier report (Zhang et al., 2007), replicating the finding in a separate population with a different pathophysiology. While this study failed to reveal any additional polymorphisms affecting splicing, the validation of SNP1/SNP2 as DRD2 variants affecting D2S/D2L splicing is important because only a firm understanding of the molecular genetic effect can guide the interpretation of clinical association studies. A number of SNPs are in high LD with SNP1/SNP2. Fig. 2.7 demonstrates the large LD block and the many SNPs that could have served as surrogate markers for SNP1/SNP2 (http://www.broadinstitute.org/mpg/snap/ldplot.php; Johnson et al., 2008a). All of these can thus be considered markers for a true genetic effect with demonstrable impact on human behavioral phenotypes (Table 2.7). It is further noted that the LD block extends in to the adjacent gene ANNKJ (Fig. 2.7), including the Taq1A variant used as a marker in numerous clinical association studies (Table 2.7). While ANNKJ could well represent a risk gene on its own, the strong LD to SNP1/SNP2 in the DRD2 locus must be considered. Lastly, the haplotype block is significantly shorter in
subjects of African descent when measured using the Yoruban HapMap population (not shown), indicating that strong ethnic differences in association studies can be expected.

As the importance of dopamine in the pathophysiology of addiction is well established, perturbations in dopamine signaling could impact an individual’s response to drugs of abuse and/or risk for addiction. Alternative splicing of \textit{DRD2} to form D2S and D2L represents a regulatory event that could affect dopamine signaling: There is ample evidence that D2S and D2L are functionally distinct (Smith et al., 2002; Usiello et al., 2000; Wang et al., 2000; Xu et al., 2002), and D2S levels could influence dopamine transport as well as release. DAT is regulated by D2S receptors through a protein-protein interaction which promotes localization of intracellular DAT to the cell membrane, resulting in enhanced clearance of extracellular dopamine (Bolan et al., 2007; Lee et al., 2007). In light of the compelling evidence for the effects of SNP1 and SNP2 on D2S/L splicing, we tested for and found a positive association of SNP1 and SNP2 with cocaine abuse in our cohort (Table 2.4). When analyzing the Caucasian and African American subgroups separately, an even stronger association emerged with cocaine abuse in Caucasians, but no association in African Americans (Table 2.4). Allele frequencies were substantially higher in the Caucasian cocaine group (SNP1 24% and SNP2 25%), compared to the controls (SNP1 10% and SNP2 9% (p=0.003 and 0.001, respectively, Table 2.4). The positive association in one racial group but not another could be interpreted as evidence against SNP1/SNP2 being the functional variants driving the association. However, the African American subgroup of our cohort is smaller (n=74 vs. 137 Caucasians) and both SNPs are less frequent in African Americans (MAF=7% vs.
18% in Caucasians) (Table 2.4), resulting in a loss of power to detect statistically significant associations in the African American subgroup. In fact, there were only nine minor allele carriers of SNP1/SNP2 in the African American subgroup, including only one individual homozygous for the minor allele, who was a cocaine abuser. Furthermore, SNP1/SNP2 exist in a different genomic context in Caucasians vs. African Americans, given the different LD pattern (Table 2.3) and haplotype distribution (Table 2.5) in the two groups, differential contribution of other genetic factors in each group must be considered. However, the biological effect of SNP1/SNP2 on D2S/L splicing was recapitulated in both groups (Fig. 2.6).

Another SNP tested, rs1125394, was also associated with cocaine abuse in the complete cohort and in Caucasians (p=0.03 and p=0.004, respectively, Table 2.4). This third SNP is in tight LD with SNP1/SNP2 (D'=1.0, Table 2.3) and has also been shown to be predictive of clozapine response in schizophrenic patients (Hwang et al., 2005; Hwang et al., 2006). In light of the data presented here and by Zhang et al. 2007 supporting SNP1 and SNP2 being functional variants modulating D2S/L splicing, it is likely that the clinical associations of rs1125394 are due to its LD with SNP1/SNP2.

The apparent additive effect of the minor alleles of SNP1/SNP2 on D2S/L splicing (Fig. 2.5a) led us to ask whether they had an additive effect on susceptibility to cocaine abuse. An analysis of SNP2 revealed the OR of each minor allele to be 2.0 (95% CI 1.1 - 3.7, p=0.02), suggesting that the OR for homozygotes would be ~4. However, when we analyzed the genotype effect vs. cocaine status, comparison of minor allele homozygotes vs. major allele homozygotes only displayed a trend towards statistical
significance (p=0.052 in Caucasians, with a large odds ratio), indicating that susceptibility of carriers homozygous for the SNP1/SNP2 minor alleles to cocaine abuse must be studied in a larger cohort.

The D2S/L ratio was not measurably affected by cocaine abuse, nor was it correlated with cocaine abuse in our cohort, even though SNP1 and SNP2 were associated with reduced relative expression of D2S and with cocaine abuse. The contribution of external trans-acting factors to variability in transcription and alternative splicing must be considered, introducing sufficient noise so that any correlations between splicing and cocaine abuse are no longer apparent. Previous evidence supports varied trans effects on splicing. Cocaine abuse has been shown to act as a trans-factor modulating overall gene expression in humans and animals (reviewed in Lull et al., 2008) and altering differential expression of specific isoforms of brain-derived neurotrophic factor in human brain (Jiang et al., 2009). Furthermore, D2S and D2L expression in mouse striatum and ventral tegmental area is differentially affected following treatment with amphetamine, another psychostimulant (Giordano et al., 2006). A host of other trans-acting factors, including postmortem decay, and the variability in extracting RNA from postmortem human tissues could have contributed to increased variability in the measurements of D2S/L splicing, confounding the three-way relationship between genotype, splicing, and cocaine abuse. Another possibility is that D2S/L splicing in the brain regions tested is not important in the etiology of cocaine abuse and the association of SNP1/SNP2 with cocaine abuse is due to altered D2S/L splicing in other brain regions (nucleus accumbens, for instance) or to another mechanism yet to be identified. Further
studies are needed to better characterize the interaction between SNP1/SNP2 genotype, cocaine abuse, and D2S/L splicing in specific brain regions.

A haplotype analysis was conducted to test for an interaction between SNP1/SNP2 and the functional promoter SNP rs12364283 and to assess the DRD2 haplotype distribution in Caucasians and African Americans separately. The haplotype distribution differed considerably between the two groups (Table 2.5), requiring haplotype associations with cocaine abuse to be performed for each group separately. A haplotype carrying the minor alleles of SNP1/SNP2 (H3) was significantly associated with cocaine abuse in Caucasians, while no other haplotype showed significant associations in the complete cohort or in the Caucasian or African American subgroups (Table 2.5). The only two haplotypes carrying the minor allele of rs12364283 (H6 and H7) tended to be overrepresented in cocaine abusers, but they are too rare (<3%) to permit a meaningful statistical analysis in our cohort (Table 2.5).

Given the evidence for interactions between DRD2 and DAT (Bertolino et al., 2009a), we assessed the interaction of SNP2 with four DAT variants thought to be functional, and identified a significant interaction between SNP2 and the DAT intron 8 VNTR, previously associated with altered DAT expression in vitro (Guindalini et al., 2006). The data indicate that the minor allele five repeat is protective in SNP2 minor allele carriers, while increasing susceptibility to cocaine abuse in individuals homozygous for the major allele of SNP2. It is interesting to speculate about the potential mechanism for this relationship: The D2S receptor interacts with DAT and “holds” it in the plasma membrane (Bolan et al., 2007; Lee et al., 2007), so D2S levels
influence the level of functional DAT that is localized to the cell membrane. *In vitro*, the intron 8 VNTR five repeat mediates higher *DAT* expression than the six repeat, and the effect of genotype on *DAT* expression is influenced by the cellular environment: the five repeat mediates even higher expression when the cells are treated with cocaine (Guindalini et al., 2006). However, the 5 repeat allele is associated with lower *DAT* mRNA expression in human substantia nigra autopsy samples (Pinsonneault et al., manuscript in preparation). Further study is needed to determine the functional effects of the *DAT* intron 8 VNTR in relevant brain regions, particularly in the context of cocaine abuse and other environmental factors. In light of the evidence for the effects of SNP1/SNP2 on D2S/L splicing and for the effects of the intron 8 VNTR on *DAT* expression, the interaction between the two variants in the context of cocaine abuse reported here is striking. As drug addiction is thought to be a complex, multigenic disorder involving a large number of genes with low penetrance, the finding of a significant interaction between two genetic variants in plausible candidate genes is an important finding requiring replication in a larger cohort.

In the present study we examined the relationship between *DRD2* genotype, D2S/L splicing, and cocaine abuse. We replicated our earlier finding that SNP1 and SNP2 decrease the relative expression of D2S in human brain in a different population and found both SNPs to be associated with cocaine abuse. The latter finding is supported by numerous earlier clinical association results in various disorders with these same SNPs, or SNPs in high LD with them. Furthermore, we identified a significant interaction between SNP2 and the DAT intron 8 VNTR. Taking all previous studies and
the present results together, SNP1/SNP2 are regulatory variants that modulate the ratio of D2S to D2L in human brain, affecting cognitive processes and psychiatric disorders including drug addiction. The large odds ratios observed for SNP1/SNP2 in the present study for risk of cocaine abuse supports a large effect size in this population of subjects with documented heavy abuse, leading eventually to cocaine overdose and death.

Figure 2.7. LD plot of surrogate markers for the intron 5 SNP rs2283265 (SNP1) as the seed. (http://www.broadinstitute.org/mpg/snap/ldplot.php) (Johnson et al., 2008a). The sequence window spans 25 kb on either side of the seed SNP, overlapping with both the DRD2 and ANNK1 genes (green arrows, reflecting gene orientation), and the threshold was set at $r^2=0.8$. The plot also shows the recombination rate, which is low
throughout this entire region but increases strongly at the borders (not shown). The LD results are from the CEU population (Utah residents with ancestry from northern and western Europe; HapMap release 22, Frazer et al., 2007). In Yoruban subjects in Ibadan, Nigeria (YRI), a significantly shortened LD block was observed, not extending beyond the \textit{DRD2} locus (not shown). SNPs with clinical associations (Table 2.7) are labeled.
Table 2.7. Clinical Associations of SNPs in LD ($R^2>0.6$) with rs2283265

<table>
<thead>
<tr>
<th>$R^2$ (with rs2283265)&lt;sup&gt;1&lt;/sup&gt;</th>
<th>SNP</th>
<th>Clinical Association</th>
<th>Population</th>
<th>Reference&lt;sup&gt;2&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>rs2283265 (SNP)&lt;sup&gt;2&lt;/sup&gt;</td>
<td>Schizophrenia, reduced cognitive performance, relative avoidance to reward-seeking behavior</td>
<td>Han Chinese, Caucasian</td>
<td>Glatt et al. 2009, Zhang et al. 2007, Frank and Hutchison 2009</td>
</tr>
<tr>
<td>1</td>
<td>rs1079594</td>
<td>Major allele assoc. with increased alcohol intake and cigarette smoking</td>
<td>Caucasian</td>
<td>Preuss et al. 2007</td>
</tr>
<tr>
<td>1</td>
<td>rs2075654</td>
<td>Commission errors on a continuous performance task test</td>
<td>79% Caucasian, 16% African American</td>
<td>Kollins et al. 2008</td>
</tr>
<tr>
<td>1</td>
<td>rs2075654</td>
<td>Affects working memory performance in the context of a COMT SNP</td>
<td>Participants in the Netherlands Twin Register</td>
<td>Gosso et al. 2008</td>
</tr>
<tr>
<td>1</td>
<td>rs1079727</td>
<td>Schizophrenia</td>
<td>Han Chinese</td>
<td>Glatt et al. 2009</td>
</tr>
<tr>
<td>0.94</td>
<td>rs1079596</td>
<td>Commission errors on a continuous performance task test</td>
<td>79% Caucasian, 16% African American</td>
<td>Kollins et al. 2008</td>
</tr>
<tr>
<td>0.94</td>
<td>rs1079597</td>
<td>Gilles de la Tourette Syndrome</td>
<td>South American</td>
<td>Herzberg et al. 2010</td>
</tr>
<tr>
<td>0.94</td>
<td>rs1125394</td>
<td>Clozapine treatment response</td>
<td>76% Caucasian, 24% African American</td>
<td>Hwang et al. 2005, Hwang et al. 2006</td>
</tr>
<tr>
<td>0.94</td>
<td>rs1075660 (SNP)&lt;sup&gt;1&lt;/sup&gt;</td>
<td>Alcoholism, altered response of the working memory network (interaction with DAT 3 VNTR), apomorphine induced growth hormone response in alcohol dependent patients, opiate addiction, reduced cognitive performance, reduced binding of dopamine transporter, and D2 receptor radiotracers in human striatum, relative avoidance to reward-seeking behavior, altered emotional control, schizophrenia-related phenotypes</td>
<td>Japanese, Caucasian, Caucasiun, Caucasian, Caucasian, Caucasian, Caucasian, Caucasian, Caucasian, Caucasiun, Caucasian, Caucasian</td>
<td>Sasabe et al. 2007, Bertolino et al. 2009a, Lu et al. 2010, Deehan et al. 2009, Zhang et al. 2007, Bertolino et al. 2010, Frank and Hutchison 2009, Blasi et al. 2009, Bertolino et al. 2009b</td>
</tr>
</tbody>
</table>

<sup>1</sup>In HapMap CEU population
<sup>2</sup>Clinical associations for Taq1A are too many to list individually; references given are reviews and meta-analyses
<table>
<thead>
<tr>
<th>Experiment</th>
<th>Oligos</th>
<th>Comments</th>
</tr>
</thead>
</table>
| **Real Time PCR to Assess DRD2 mRNA Expression** | Forward: 5'CCAGCTGACTCTCCCGAC  
Reverse: 5'GCATGCCCATTCTTCTCTGG | |
| **Splice Variant Detection** | Forward: 5'ACATTGTCTCCCGGACAGCAG  
Reverse: 5'GCATGCCCATTCTTCTCTGG | Forward primer was Fam labeled |
| **Minigene Mutagenesis** | Forward: 5'GCCCGGCGCGCCCTGTCCTCAGTT  
Reverse: 5'TGAATTACATCGGATGCTCTGC | For introduction of rs35608204 into minigene |
| **Minigene cDNA Synthesis** | SP6: 5'CATTAGGTGACACTATAG | Used for cDNA synthesis from RNA derived from plasmid |
| rs12364283 GC Clamp Genotyping | SNP Forward: 5'GCCCGGCGCGCCCTGTCCTCAGTT  
WT Forward: 5'CTGTCCTCAGTTGGCAGGA  
Reverse: 5'CAGCACCTGTCTTAAGCCTCTAGT | |
| rs12364283 SNaPshot Genotyping | Forward: 5'GCCCGGCGCGCCCTGTCCTCAGTT  
Reverse: 5'CAGCACCTGTCTTAAGCCTCTAGT  
PEP: 5'GTGACTTCTGATATGACACAGAA | |
| rs35608204 GC Clamp Genotyping | Forward: 5'TCTTGTCTCTCTGTGCCCAG  
WT Reverse: 5'TGAATTACATCGGATGCTCTGC  
SNP Forward: 5'GCCCGGCGCGCCCTGTCCTCAGTT  
Reverse: 5'CGACCTGTCTTAAGCCTCTAGT  
GACTCTGAC | |
| rs35608204 RFLP Genotyping | Forward: 5'TCTTGTCTCTCTGTGCCCAG  
Reverse: 5'TGAATTACATCGGATGCTCTGC  | Forward primer was Fam labeled, restriction enzyme was |
| rs35608204 SNaPshot Genotyping | Forward: 5'CTCTTTTCTCGGTCTCTGCTCTC  
5'TTTGATTCCCATTTAACAGATGAGG  
PEP: 5'GGAACAGGCTCATAGAAGTGATGC | |
| rs2283265 GC Clamp Genotyping | Forward: 5'CTCTTTTCTCGGTCTCTGCTCTC  
5'TTTGATTCCCATTTAACAGATGAGG  
PEP: 5'GGAACAGGCTCATAGAAGTGATGC | |
<table>
<thead>
<tr>
<th>rs2283265 SNaPshot Genotyping</th>
<th>Forward: 5’CTCTTTTCTGGTTTCTCTGTCTCACT</th>
<th>5’TTTGATTCCCATTTAACAGATGAGG</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PEP: 5’TTGAGGAACAGGCTCATAGAAGGTAAG</td>
<td></td>
</tr>
<tr>
<td>rs1076560 GC Clamp Genotyping</td>
<td>Forward: 5’CTGCACCAGAGGCAGAGG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WT Reverse: 5’TTGCAGGAGTCTTCAGAGTGCG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SNP Reverse:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5’CGGGCCGGCCCGCCGAGTTGCAGAGGAGTTCGAG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TCAGAGCGT</td>
<td></td>
</tr>
<tr>
<td>rs1076560 SNaPshot Genotyping</td>
<td>Forward: 5’CACCCATCTCAGCTGGCCC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reverse: 5’GTCGGTGTGGTTGGCAGGAGTCTTCAG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PEP: 5’GTGGTGCAGGAGTCTTCAGAGGG</td>
<td></td>
</tr>
</tbody>
</table>
3.1 Introduction

β-Arrestin2 is an important mediator in the desensitization and internalization of G-protein coupled receptors (GPCRs) (Ferguson, 2001; Ferguson et al., 1996; Luttrell and Lefkowitz, 2002; Perry and Lefkowitz, 2002). More recently, it has also been shown to activate signaling cascades independent of G-protein activation (DeWire et al., 2007; Lin et al., 1998; Luttrell et al., 2001). It is ubiquitously expressed and interacts with most GPCRs, including dopamine receptors (Beaulieu et al., 2005; Klewe et al., 2008; Lan et al., 2009). Mice lacking β-arrestin2 display a number of altered phenotypes relevant to psychiatric disorders, including enhanced morphine antinociception (Bohn et al., 1999) and reward (Bohn et al., 2003) as well as decreased locomotor activity and disrupted dopaminergic signaling and dopamine-related behaviors (Beaulieu et al., 2005). Furthermore, it has been suggested that identification of “biased ligands” that selectively activate G-protein or β-arrestin2-mediated signaling could represent a novel approach to drug discovery (Violin and Lefkowitz, 2007). In fact, it has been recently shown that lithium, which is used to treat psychiatric disorders including schizophrenia, bipolar
disorder, and depression, disrupts a signaling complex that includes β-arrestin2 and mice lacking β-arrestin2 are resistant to the behavioral changes induced by lithium (Beaulieu et al., 2008), underscoring the potential impact of β-arrestin2 on drug therapy.

Variants of the gene encoding β-arrestin2 (ARRB2) have been associated with psychiatric disorders and with altered response to drug therapy in human populations. One single nucleotide polymorphism (SNP) located in exon 11 of the ARRB2 gene, rs1045280, comes up repeatedly in the literature: It has been linked to tardive dyskinesia in Chinese patients with schizophrenia (Liou et al., 2008) and was part of a haplotype associated with nicotine dependence in a European American population (Sun et al., 2008). In a Japanese population, rs1045280 was associated with methamphetamine use disorder but not schizophrenia (Ikeda et al., 2007). Another recent study showed that cancer patients who did not respond to morphine for pain management and/or who experienced intolerable adverse effects were less likely to carry the minor allele of rs1045280 than patients who responded to morphine (Ross et al., 2005).

Association studies have historically focused on SNPs in protein coding regions, since variants that change the amino acid sequence or introduce a premature stop codon are the most obvious candidates for genetic variants affecting the amount or function of the gene product. However, recent studies indicate that regulatory polymorphisms in non-coding regions – most as yet undiscovered – occur more frequently (Johnson et al., 2005; Rockman and Wray, 2002; Wray, 2007). As rs1045280 is a synonymous SNP that does not change the amino acid sequence, its functional consequences (if any) must be due to its regulatory effects at the RNA level (e.g. transcription, mRNA processing,
mRNA stability, and translation) rather than structural alterations in the protein. One limitation of most association studies is a lack of consideration of the molecular genetic mechanisms underlying the effects of polymorphisms, resulting in the association of phenotype with genetic variants that may be functional or, more likely, are simply in linkage disequilibrium (LD) with a variant that is functional. Since allele frequencies and linkage patterns may differ between populations, results may vary between association studies, underscoring the importance of identifying functional genetic variants in candidate genes.

The *ARRB2* gene occupies ~11 kb on chromosome 17. It consists of 15 exons (Fig. 3.1) and undergoes alternative splicing to form at least 22 distinct splice variants (http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/; Thierry-Mieg and Thierry-Mieg, 2006). At least 90% of human genes undergo alternative splicing (Wang et al., 2008b), a process that has emerged as an especially important source of diversity and regulation for genes expressed in the brain (Lee et al., 2003). NCBI’s RefSeq database (Pruitt et al., 2007) annotates two major *ARRB2* transcripts, the full-length isoform and a short isoform lacking exon 4 (Fig. 3.1), although little is known about the functional significance of the different *ARRB2* splice variants. Located in exon 11, the synonymous SNP rs1045280 is one of five SNPs in the coding region of *ARRB2* and the only one of the five with >5% minor allele frequency. *ARRB2* is highly expressed in the central nervous system as well as in other tissues all over the body. In summary, *ARRB2* is biologically important, implicated in psychiatric disorders and drug response, highly expressed, and contains a frequent marker SNP in the coding region (rs1045280) amenable to analysis of allelic
expression. Given all of these factors, I asked whether rs1045280, previously associated with a number of human phenotypes, affects transcription and/or alternative splicing in human brain.

I assessed overall and allelic expression of ARRB2 mRNA in postmortem human prefrontal cortex tissue samples from Alzheimer’s patients (n=26) and normal controls (n=20) as well as in cultured human lymphocytes. In addition, I evaluated the levels of the two major ARRB2 mRNA transcripts in all three cohorts and the βarrestin2 protein levels in the Alzheimer’s cohort. The major full-length transcript represented the overwhelming majority of ARRB2 mRNA in all three cohorts. In the Alzheimer’s and lymphocyte samples, allelic expression imbalance (AEI), in a number of samples suggest that rs1045280 is in LD with a regulatory variant or variants causing increased transcription, and rs1045280 was associated with increased expression of overall ARRB2 mRNA in the Alzheimer’s cohort. However, I tested for and did not identify any relationship between rs1045280 genotype and βarrestin2 protein levels in the Alzheimer’s cohort.
Figure 3.1. *ARRB2* gene map. The exon that is included in the full length isoform of *ARRB2* mRNA but absent in the short isoform tested for in this study (exon 4), is in gray. The location of the synonymous SNP rs1045280 in exon 11 is indicated.

3.2 Materials and Methods

**Human Tissue Selection and Sources**

Post-mortem prefrontal cortex (PFC) tissue samples were obtained from different sources as follows: 26 from patients with histologically confirmed Alzheimer’s disease from the Ohio State University Neurodegenerative Disease Brain Tissue Repository (Buckeye Brain Bank), courtesy of Dr. Maria Kataki and 20 from normal controls from the Miami Brain Endowment Bank (University of Miami, Miami, FL), courtesy of Dr. Deborah Mash. Ninety Epstein-Barr virus-transformed B-lymphoblast cell lines, consisting of 30 Caucasian family trios, were obtained from the Coriell Cell Repository (Coriell Institute, Camden, NJ).

**DNA and RNA Preparation**

Genomic DNA and total RNA were prepared from frozen brain samples or cultured lymphocytes. Samples were treated with sucrose-Triton (with SDS for brain tissues only).
and digested overnight with proteinase K, followed by sodium chloride precipitation to separate proteins and ethanol precipitation of DNA. Total RNA was extracted with TRIzol reagent (Life Technologies, Foster City, CA) and purified using RNeasy (QIAGEN, Germantown, MD) columns according to the manufacturer’s instructions, including an on-column DNase treatment to remove genomic DNA. Complementary DNA (cDNA) was generated from 500 ng of total RNA with Superscript III reverse transcriptase (Life Technologies, Foster City, CA) using oligo(dT) and, for the Alzheimer’s and Coriell cultured lymphocyte samples, ARRB2-specific primers.

**Quantitative mRNA Analysis by Real-Time PCR**

Real-time PCR was performed on cDNA samples on an ABI 7000 sequence detection system with Power SYBR® Green PCR Master Mix (Life Technologies, Foster City, CA). PCR reactions were performed in 20 µL reaction volumes in standard 96-well plates. Cycle thresholds (Ct), the point at which a reaction reaches a fluorescent intensity above background, were determined with ABI 7000 SDS software. Replicate cycle thresholds were averaged and ARRB2 expression levels in arbitrary units were calculated by subtracting the β-actin Ct from the ARRB2 Ct to get a ΔCt as described previously (Pinsonneault et al., 2006). Arbitrary units for each sample = 100(2^-ΔCt).

**Allelic Expression Imbalance**

Allele-specific mRNA expression in samples heterozygous for selected marker SNPs was measured with SNaPshot (Life Technologies, Foster City, CA). The SNaPshot assay is a single base extension method that allows quantitative comparison of two alleles. The first step was PCR amplification of the target region of DNA containing the SNP of
interest. PCR conditions were as follows: Initial denaturation of 5 min at 95°C, followed by 30 cycles of 15 sec denaturation at 95°C, 45 sec primer annealing at 60°C, 1 min of extension at 72°C, and one final extension step at 72°C for 10 min. Following amplification, the reaction mixture was treated with exonuclease I (New England Biolabs, Ipswich, MA) to degrade unused primers and bacterial Antarctic alkaline phosphatase (New England Biolabs, Ipswich, MA) to remove the 5’ phosphate groups from unused deoxynucleotides. The single base extension reaction followed, where a single primer with its 3’ end one base from the target SNP was extended by one base complementary to the SNP. I used the SNaPshot kit (Life Technologies, Foster City, CA) that includes fluorescently labeled dideoxynucleotides that allow for extension by a single base and detection of the individual nucleotides in the product. SNaPshot reaction conditions were as follows: 25 cycles of 95°C denaturation for 10 sec, 55°C for 5 sec, and 60°C for 30 sec. Each reaction mixture was then treated with calf intestinal phosphatase (New England Biolabs, Ipswich, MA) to remove the 5’ phosphate groups from unused dideoxynucleotides before being analyzed on an Applied Biosystems 3730 DNA Analyzer and the results compiled and quantitated with GeneMapper software (Life Technologies, Foster City, CA). Each of the four possible dideoxynucleotides was labeled with a different color fluorophore, allowing for quantitative detection of individual alleles in the amplified product. In heterozygous genomic DNA samples, the amounts of each allele are assumed to be equal. However, since different fluorophores may influence nucleotide incorporation rates and migration of the product through the capillaries in the ABI 3730, peak height ratios for DNA as well as cDNA were
normalized to the mean of the DNA ratios. As an additional control for variation introduced by different fluorophores, two primers (one upstream and one downstream) were utilized independently for the SNaPshot reaction, allowing for the comparison of the results utilizing one pair of fluorescently labeled dideoxynucleotides to the other pair.

**Genotyping Methods**

Genotyping of rs1045280 was accomplished by at least one of two methods: SNaPshot, as described above, or GC clamp. The GC clamp genotyping assay is described in detail in (Papp et al., 2003). Briefly, I designed two different allele specific primers along with one common primer in the opposite direction. A GC clamp of 10-15 bases was introduced to the 5’ end of one of the allele specific primers, allowing the two different allelic PCR products to be differentiated by their melting temperatures by an ABI 7000 Sequence Detection System real time PCR instrument.

**Western Blotting**

Frozen tissues from human autopsy samples were homogenized in RIPA buffer (20 mM Tris-HCl, pH 7.5, 150 mM NaCl, 2 mM EDTA, 0.1% SDS, 1% NP-40, 0.25% deoxycholate, 1 mM PMSF, 1 mM NaF, with Complete Mini, EDTA-free protease inhibitor cocktail tablet (Roche Diagnostics, IN). As a positive control and to enable identification of βarrestin2 vs. βarrestin1, each blot included cell lysates from untransfected HEK 293 cells as well as cells transiently transfected with human *ARRB2* (3 µg or 6 µg cDNA). Protein levels were determined with the use of the detergent-compatible protein assay system (Bio-Rad, Hercules, CA), and equal amounts of protein per sample (50 µg from cells or 75 µg from tissue) were resolved by 1-D gel
electrophoresis on 10% NUPAGE Bis-Tris gels (Life Technologies, Foster City, CA). Proteins were transferred to polyvinylidene fluoride (PVDF) membranes (Immobilon P; Millipore, Billerica, MA) and immunoblotted for βarrestin2 (βarrestin2 (H-9 sc-13140); Santa Cruz Biotechnology, Inc., Santa Cruz, CA). Blots were stripped and blotted for actin levels (MAB1501 antibody; Millipore, Billerica MA), which were used to normalize the overall βarrestin2 levels between samples. Chemiluminescence was detected and quantified using the Kodak 2000R imaging system (Eastman Kodak Company, NY). Data are mean ± SEM, n=3 measurements per sample.

**Cell Culture and Transient Transfection**

HEK 293 cells were cultured in minimum essential medium (MEM; Life Technologies, Foster City, CA) containing 10% heat inactivated fetal bovine serum (FBS; Life Technologies, Foster City, CA) and 1% penicillin-streptomycin (Life Technologies, Foster City, CA) at 37°C under 5% CO₂. Transient transfections were performed by electroporation using the Gene Pulser II system (Bio-Rad, Hercules, CA). Cells were resuspended in MEM + 10% FBS + 5 mM BES at a concentration of approximately 10⁶ cells in 0.5 ml per 0.4-cm cuvette with 3 μg or 6 μg of human βarrestin2 cDNA in a pcDNA 3.1 (-) vector. A single pulse at 260 V, 1000 μF, was used. Additional complete media without BES was added immediately to the cells, which were then returned to the incubator. The media was changed 4-6 hours later and the cells harvested 24-48 hours later in cold RIPA buffer as described above.
Quantitative Detection of Splice Isoforms. The *ARRB2* splice variants were measured following PCR amplification of cDNA with a fluorescent Fam-labeled exon 3 primer and an exon 5 reverse primer as described in Wang et al., 2006.

Statistical Analysis

Genomic DNA and cDNA peak height ratios were normalized as described above in “Allelic Expression Imbalance”. The presence of AEI was determined with normalized cDNA ratios compared to the normalized genomic DNA ratio. The standard deviation (S.D.) of the normalized genomic DNA ratio was determined and samples that had significant allelic ratio differences between cDNA and DNA (> 2 S.D.) were considered to show significant AEI. One-way ANOVA was performed with Prism software (GraphPad Software, La Jolla, CA). Haplotype estimation and generation of the LD plot were performed with Haploview 4.2 (www.broadinstitute.org/haploview/haploview; Barrett et al., 2005)

3.3 Results

Total *ARRB2* mRNA Expression

*PFC from Alzheimer’s Patients and Normal Controls*

I measured overall *ARRB2* mRNA expression with real time PCR in postmortem human PFC from Alzheimer’s patients (n=26) and normal controls (n=20). The *ARRB2* mRNA levels measured were approximately 10-fold higher in the Alzheimer’s group than in the controls, possibly due to the lack of *ARRB2*-specific primers in the cDNA synthesis for the latter. In the Alzheimer’s group, the minor allele C of rs1045280 was significantly associated with increased *ARRB2* mRNA expression (TT vs. CC, P<0.01, Fig. 3.2a).
the control group, there was no correlation between rs1045280 genotype and ARRB2 mRNA expression (Fig. 3.2b).

*Coriell Cultured Lymphocytes*

Since ARRB2 is ubiquitously expressed, I was able to measure overall ARRB2 mRNA expression with real time PCR in cultured lymphocytes as well (n=82). ARRB2 mRNA levels in the lymphocytes were lower than in the Alzheimer’s samples but comparable to the levels observed in the control PFC samples, not surprising considering that ARRB2 is more highly expressed in brain than in lymphocytes (Unigene EST profile; Wheeler et al., 2003) but the cDNA synthesis for the RNA from the control brains was the only one that did not utilize ARRB2-specific primers. There was no significant correlation between rs1045280 genotype and ARRB2 mRNA expression in the lymphocytes (Fig. 3.2c).
Figure 3.2. **ARRB2 mRNA expression.** Real-time PCR was applied to quantify mRNA, with β-actin as an internal control. *ARRB2* mRNA levels are presented in arbitrary units (see Materials and Methods). Data are mean ± SEM, two cDNA syntheses and triplicate measurements for each cDNA synthesis per sample for the Alzheimer’s brains (a), and one cDNA synthesis with duplicate measurements in the control brains (b), and lymphocytes (c). ***, P<0.01, one-way ANOVA with Bonferroni post-test.
Figure 3.2. *ARRB2* mRNA expression
Allelic *ARRB2* mRNA Expression

*PFC from Alzheimer’s Patients and Normal Controls*

Measurement of the mRNA generated from each allele requires the use of frequent marker SNPs located in the transcribed region. The *ARRB2* transcript contains only one SNP with >5% heterozygosity, rs1045280, which was used as the marker SNP for all of the AEI measurements. To control for fluorescent dye bias in the SNaPshot reaction, extension primers were included for both orientations in separate reactions and analyzed separately. The two different extension primers yielded similar results, which are averaged together for each sample (Fig. 3.3). I measured allelic mRNA expression in samples heterozygous for rs1045280 from the Alzheimer’s patients (n=12) and normal controls (n=8). In the Alzheimer’s cohort, I found significant AEI in three samples (Fig. 3.3a), all below unity, indicating that the minor allele C yielded more mRNA than the major allele T in these samples. The fact that the AEI ratios were all in the same direction suggest that the marker SNP is in LD with a regulatory variant causing the AEI. None of the control samples displayed significant AEI (Fig. 3.3b).

*Coriell Cultured Lymphocytes*

I measured allelic mRNA expression in 34 samples heterozygous for rs1045280 and found significant AEI in four of them (Fig. 3.3c), all below unity to a similar extent as in the Alzheimer’s samples (Fig. 3.3a). Similar to the results obtained from the Alzheimer’s cohort, these results indicate that the minor allele C yielded more mRNA than the major allele T and that the functional SNP is linked to the marker SNP rs1045280.
Figure 3.3. Allelic ARRB2 mRNA expression ratios. Ratios were normalized to genomic DNA and are mean ± SEM, two cDNA syntheses and at least triplicate measurements for each cDNA synthesis per sample in the Alzheimer’s brains (a), one cDNA synthesis with triplicate measurements in the control brains (b), and one cDNA synthesis with either one or triplicate measurements (samples with error bars representing the latter) in lymphocytes (c). Samples with significant AEI are denoted with a * and were determined as described in Materials and Methods.
Figure 3.3. Allelic ARRB2 mRNA expression ratios.
\textbf{βarrestin2 Protein Expression in Human PFC from Alzheimer’s Patients Assessed by Western Blotting}

In light of the observation that the minor allele of rs1045280 was associated with increased \textit{ARRB2} mRNA expression in the Alzheimer’s cohort (Fig. 3.2a), I assessed βarrestin2 protein in the same samples (n=26) by Western blotting to determine whether rs1045280 was associated with increased protein levels as well. Either 50 or 75 µg of total protein was loaded per lane and resolved by 1-D gel electrophoresis on 10% bis-tris gels. Since the βarrestin2 antibody also reacts to βarrestin1, which has a slightly higher but similar molecular weight, cell lysates from HEK 293 cells transiently transfected with human \textit{ARRB2} cDNA were included on each gel as a positive control and to enable identification of the βarrestin2 band (see representative blot; Fig. 3.4a). βArrestin1 has been shown to be the predominant arrestin in several human brain regions (Bychkov et al., 2008), and it predominates in our human PFC samples as well (Fig. 3.4a). Both arrestins are expressed in HEK 293 cells. However, βarrestin1 is expressed at lower levels in HEK 293 cells than βarrestin2 (BioGPS database; Wu et al., 2009), which accounts for the lack of a visible band corresponding to βarrestin1 in the HEK 293 cell lysates (Fig. 3.4a). The inclusion of the transfected HEK 293 cell lysates allowed for unambiguous identification of the βarrestin2 band on each blot. I detected no correlation between rs1045280 genotype and βarrestin2 protein levels (Fig. 3.4b).
Figure 3.4. Western Blot. (a) Representative Western blot and (b) densitometric analysis of relative β-arrestin2 protein levels in extract prepared from human PFC from Alzheimer’s patients (75 µg total protein per lane) or HEK 293 cells (50 µg total protein per lane) transfected with the indicated amounts of human *ARRB2* cDNA. β-Arrestin2 levels were normalized to actin. Data are mean ± SEM; n=3 measurements per sample.
**Alternative Splicing of ARRB2 mRNA**

To test for genetic effects on splicing, I assessed the levels of the two major splice variants of ARRB2, the full-length long isoform and a short isoform lacking exon 4. The experiments revealed that the full-length isoform represented the overwhelming majority of ARRB2 mRNA in the Alzheimer’s (n=26) and normal brains (n=20) as well as the cultured lymphocytes (n=14, Fig. 3.5). Interestingly, I did not detect the expected short isoform but did detect additional fluorescent peaks representing PCR products that did not correspond in size with any reported ARRB2 splice variant, representing either novel mRNA splice variants or PCR artifacts. The peaks in question appeared consistently in nearly all samples from the lymphocytes and both brain regions at approximately the same intensity (Fig. 3.5), suggesting that even if they do represent real ARRB2 mRNA splice variants, there is very little variability in alternative splicing in this region between samples or even between tissues. Given the lack of evidence that the ARRB2 splice variants are functionally distinct and the lack of variability of the splice variants measured in our samples, I did not pursue the characterization of ARRB2 splice variants any further.
3.4 Discussion

The present study indicates that the synonymous SNP rs1045280 is associated with enhanced \textit{ARRB2} mRNA expression in human PFC autopsy tissues in a cohort of Alzheimer’s patients but not in PFC from normal controls or in cultured human lymphocytes. An assessment of allelic \textit{ARRB2} mRNA expression in the same samples suggested that rs1045280 is in LD with a functional variant or variants with a small but significant effect (increase) on expression, an effect detected in the lymphocytes and the Alzheimer’s cohort but not in the controls. In the Alzheimer’s cohort, measurement of
β-arrestin2 protein by Western blotting did not reveal any association between rs1045280 genotype and β-arrestin2 protein levels.

Since genetic regulatory mechanisms may differ between tissues, it is necessary to assess the genetic regulation of a candidate gene in tissues where it is functionally expressed. *ARRB2* is ubiquitously expressed but I focused on brain since rs1045280 has been linked to psychiatric disorders including tardive dyskinesia (Liou et al., 2008) and drug addiction (Ikeda et al., 2007; Sun et al., 2008). Furthermore, I evaluated the genetic regulation of *ARRB2* in lymphocytes in order to determine whether *ARRB2* is differentially regulated in the two different tissues. It is well documented that gene expression is altered in neurodegenerative disorders including Alzheimer’s (Liang et al., 2008; Loring et al., 2001). Moreover, *ARRB2* mRNA expression is elevated in striatum in patients with Parkinson’s disease with dementia (Bychkov et al., 2008). In the current study, *ARRB2* mRNA levels were approximately 10-fold higher in the Alzheimer’s patients than in the normal controls (Fig. 3.2a,b). This difference could be due to the lack of *ARRB2*-specific primers in the cDNA synthesis for the normal controls and/or the effect of Alzheimer’s disease on gene expression. Other differences between the two cohorts are also potentially contributing factors: The average age of the controls was much lower than that of the Alzheimer’s patients, and the controls were all female (Table 3.1).

The association of rs1045280 with increased overall *ARRB2* mRNA expression in the Alzheimer’s cohort is consistent with the only significant AEI ratios being <1 in the same samples, since the AEI ratio is calculated as major allele expression/minor (SNP)
allele expression. The observation that not every sample heterozygous for rs1045280 displayed significant AEI (Fig. 3.3) is an indication that it is not a functional SNP affecting transcription. However, all of the significant ratios observed were in the same direction (Fig. 3.3), suggesting that rs1045280 is in LD with a functional variant. Given the haplotype structure of ARRB2, this is not a surprising finding: ARRB2 is a relatively small gene (~11 kb) that sits in a large haplotype block (Fig. 3.6), so any cis-acting regulatory SNP in the gene locus is likely to be in LD with rs1045280. Given the association of rs1045280 with enhanced ARRB2 mRNA expression in the Alzheimer’s cohort, I assessed the relationship between rs1045280 genotype and βarrestin2 protein levels in the same tissues and found no correlation (Fig. 3.4b), most likely due to a compensatory translational mechanism or experimental variability inherent in extracting RNA and protein from autopsy samples.

Alternative mRNA splicing is an important source of diversity and regulation for genes expressed in the brain (Lee et al., 2003) and is thought to be a major source of proteomic diversity in all eukaryotes (Nilsen and Graveley, 2010). At least 22 distinct ARRB2 splice variants have been identified, (http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/; Thierry-Mieg and Thierry-Mieg, 2006), with two major isoforms (http://www.ncbi.nlm.nih.gov/refseq/; Pruitt et al., 2007). We have previously characterized two SNPs that modulate alternative splicing of dopamine D2 receptor mRNA, resulting in a shift in the balance of two functionally distinct isoforms with effects on human phenotypes including cognitive performance (Zhang et al., 2007) and risk for cocaine addiction (Moyer et al., 2010 (manuscript in
preparation)). Given the importance of alternative splicing, the evidence that ARRBB2 mRNA is alternatively spliced, and our previous findings that SNPs can modulate alternative splicing with consequences for human phenotypes, I assessed the levels of the two major ARRBB2 splice variants in each cohort. The striking similarity between cohorts and tissue types (Fig. 3.5) and between samples within each cohort (data not shown) is an indication that there is little variability in alternative splicing in the region tested (exon 3-5) in our samples, evidence that alternative splicing of exon 4 is not important for β-arrestin2 regulation in human PFC or lymphocytes.

In the present study I examined the relationship between rs1045280, previously associated with human phenotypes, and ARRBB2 mRNA and expression and assessed alternative splicing of ARRBB2 mRNA in human PFC autopsy tissues and in cultured human lymphocytes. Alternative splicing did not differ between tissue types or between samples. The minor allele of rs1045280 was associated with increased overall ARRBB2 mRNA expression in PFC from Alzheimer’s patients but not in PFC from normal controls or in cultured lymphocytes. Analysis of allelic mRNA expression in the same samples indicated that rs1045280 likely does not directly influence transcription but is in LD with a functional variant that does, and examination of the haplotype structure of the gene locus revealed that ARRBB2 lies in a large haplotype block, so there are likely numerous SNPs in LD with the functional variant(s) that could serve as surrogate markers. There was no correlation between rs1045280 genotype and β-arrestin2 protein levels, indicating that the must be a compensatory translational mechanism or the difference in protein level is too small to be seen by Western blotting. Taken together,
the results suggest that rs1045280 is in LD with a regulatory variant exerting a small but significant effect on *ARRB2* mRNA expression that is not maintained at the protein level.

### Table 3.1. Demographic Characteristics

<table>
<thead>
<tr>
<th></th>
<th>Controls</th>
<th>Alzheimer’s Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>n=</td>
<td>20</td>
<td>26</td>
</tr>
<tr>
<td>Age (mean ± SD)</td>
<td>59 ± 7</td>
<td>80 ± 9</td>
</tr>
<tr>
<td>% Female</td>
<td>100%</td>
<td>73%</td>
</tr>
<tr>
<td>Caucasian</td>
<td>19</td>
<td>26</td>
</tr>
<tr>
<td>African American</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
Figure 3.6. LD plot of the ARRBB2 genomic locus. (www.broadinstitute.org/haploview/haploview; Barrett et al., 2005). The sequence window spans 5 kb on either side of ARRBB2 and linkage disequilibrium values (D’) are indicated as percent in each box, with darker boxes indicating higher linkage between alleles (D’=1 for the darkest boxes). The results are from the CEU population (Utah residents with ancestry from northern and western Europe; Frazer et al., 2007).
### Table 3.2. Sequences of Oligonucleotides (Primers) Used

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Oligos</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNaPshot (AEI measurement)</td>
<td>Forward: 5'CCAGGTATCTCCAGCTCCA</td>
<td>Also used for rs1045280 genotyping</td>
</tr>
<tr>
<td></td>
<td>Reverse: 5'CAAAGGGTAGGTCCCGC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PEF: 5'ACCATAACCCACCTGCTAG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PER: 5'CGCTTCTCCGGGTTGTC</td>
<td></td>
</tr>
<tr>
<td>Splice Variant Detection</td>
<td>Forward: 5'GCCAGGACTTCGTAGATCACC</td>
<td>Forward primer was Fam labeled</td>
</tr>
<tr>
<td></td>
<td>Reverse: 5'CCAGCACAATCCAGGTCTTCA</td>
<td></td>
</tr>
<tr>
<td>Real Time PCR to Assess ARRB2 mRNA Expression</td>
<td>Forward: 5'CCAGGTATCTCCAGCTCCA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reverse: 5'CAAAGGGTAGGTCCCGC</td>
<td></td>
</tr>
</tbody>
</table>
Chapter 4: Allelic Imbalance in DRD3 mRNA Expression

4.1 Introduction

The dopamine receptor D3 (encoded by DRD3) is a member of the D2-like family of dopamine receptor subtypes, which are coupled to G_i inhibitory G-proteins and generally reduce formation of intracellular cAMP when activated (Civelli et al., 1993; Strange, 1993). DRD3 is expressed predominantly in the brain’s limbic system and nucleus accumbens (NAc), brain regions thought to be involved in the pathophysiology of drug addiction and schizophrenia (Suzuki et al., 1998). The involvement of D3 receptors with addiction is well documented in preclinical studies: Pharmacological blockade of D3 receptors or genetic ablation of DRD3 in rodents both result in reduced cue-induced reinstatement of drug-seeking behavior for a number of drugs of abuse including ethanol, nicotine, opiates, and cocaine (reviewed by Le Foll et al., 2009). In humans, increased levels of DRD3 mRNA have been reported in the NAc of cocaine fatalities (Segal et al., 1997; Staley and Mash, 1996), and decreased DRD3 mRNA expression was observed in cortical brain regions in patients with schizophrenia (Schmauss et al., 1993). Changes in D3 receptor density and/or function could be clinically relevant, since the D3 receptor is thought to mediate the actions of antipsychotic drugs (Schwartz et al., 2000; Sokoloff et
al., 1992) and it is a potential drug target for the treatment of addiction (Pilla et al., 1999; Xi and Gardner, 2007).

Genetic variants of \( \textit{DRD}3 \) have been scrutinized in a number of association studies, often with inconsistent results (for recent review see Le Foll et al., 2009). Association studies are often biased toward polymorphisms in the protein coding in the gene(s) of interest, and association studies of \( \textit{DRD}3 \) are no exception; most have evaluated the nonsynonymous SNP rs6280, also called Ser9Gly or Ball. The Gly-9 variant has an increased affinity for dopamine \textit{in vitro} (Jeanneteau et al., 2006), evidence for a gain of function. A number of clinical associations with rs6280 have been found, although they are often not replicable (Le Foll et al., 2009). It is also noted that the Ser allele frequency varies widely across ethnic populations (Crocq et al., 1996; HapMap release 22, Frazer et al., 2007; Lerer et al., 2002). Taken together, the inconsistent results of clinical association studies and varied allelic distribution of rs6280 suggest that there may be other functional variants that contribute to the phenotypes tested. Given their potential impact and predicted prevalence, regulatory polymorphisms in non-coding regions must be considered (Johnson et al., 2005; Rockman and Wray, 2002; Wray, 2007).

The \( \textit{DRD}3 \) gene occupies ~50 kb on chromosome 3. The full-length isoform (D3L) consists of 7 exons (Fig. 4.1) and \( \textit{DRD}3 \) mRNA is alternatively spliced to form at least 5 distinct splice variants (http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/; Thierry-Mieg and Thierry-Mieg, 2006). Alternative splicing is an important source of proteomic diversity and regulation (Nilsen and Graveley, 2010) that occurs in nearly all
human genes (Wang et al., 2008b) and is thought to be an especially important source of
diversity and regulation for neuronal genes (Lee et al., 2003). NCBI's RefSeq database
(Pruitt et al., 2007) annotates two major DRD3 transcripts, the full-length isoform (D3L)
and a truncated isoform (D3nf) with a frame shift mutation and a premature stop codon,
resulting in a shortened transcript with an altered sequence beginning in exon 6 (Fig. 4.1).
D3nf mRNA is translated into a protein lacking D3-typical transmembrane-spanning
domains 6 and 7 that has been detected in human brain (Liu et al., 1994). In vitro co-
immunoprecipitation studies have shown that D3nf protein physically interacts with the
functional D3 receptor and causes it to accumulate in an intracellular compartment rather
than inserting into the plasma membrane, indicating a dominant negative effect for D3nf
(Elmhurst et al., 2000; Karpa et al., 2000).

DRD3 is expressed at low levels in the brain compared to other dopamine
receptor genes, so it is critical to examine a physiologically relevant brain region where
DRD3 is functionally expressed at levels high enough to measure reliably. To this end, I
measured allelic mRNA expression and alternative splicing in human brain autopsy NAc
obtained from a cohort of normal controls as well as patients diagnosed with
schizophrenia, bipolar disorder, or depression (n=15 in each group) and searched for
SNPs associated with altered transcript levels. Using rs6280 as the marker SNP, I
detected significant AEI in 4 out of 29 samples tested, although none of the 11 SNPs
genotyped were associated with the observed AEI. Additionally, I assessed D3L and
D3nf mRNA levels in the same samples and found only D3L.
4.2 Materials and Methods

**Human Tissue Selection and Sources**

Postmortem brain tissues were donated by The Stanley Medical Research Institute's (Chevy Chase, MD) Neuropathology Consortium Brain Collection, courtesy of Dr. Maree Webster. We obtained frozen sections from 60 individuals previously diagnosed with schizophrenia, bipolar disorder, or depression (15 of each) and 15 controls.

**DNA and RNA Preparation**

Genomic DNA and total RNA were prepared from frozen brain samples. Samples were treated with sucrose-Triton with SDS and digested overnight with proteinase K, followed by sodium chloride precipitation to separate proteins and ethanol precipitation of DNA. Total RNA was extracted with TRIzol reagent (Life Technologies, Foster City, CA) and...
purified using RNase (QIAGEN, Germantown, MD) columns according to the manufacturer’s instructions, including an on-column DNase treatment to remove genomic DNA. Complementary DNA (cDNA) was generated from 500 ng of total RNA with Superscript III reverse transcriptase (Life Technologies, Foster City, CA) using oligo(dT) and \textit{DRD3}-specific primers.

\textbf{Allelic Expression Imbalance}

Allele-specific mRNA expression in samples heterozygous for selected marker SNPs was measured with SNaPshot (Life Technologies, Foster City, CA). The SNaPshot assay is a single base extension method that allows quantitative comparison of two alleles. The first step was PCR amplification of the target region of DNA containing the SNP of interest. PCR conditions were as follows: Initial denaturation of 5 min at 95°C, followed by 30 cycles of 15 sec denaturation at 95°C, 45 sec primer annealing at 60°C, 1 min of extension at 72°C, and one final extension step at 72°C for 10 min. Following amplification, the reaction mixture was treated with exonuclease I (New England Biolabs, Ipswich, MA) to degrade unused primers and bacterial Antarctic alkaline phosphatase (New England Biolabs, Ipswich, MA) to remove the 5’ phosphate groups from unused deoxynucleotides. The single base extension reaction followed, where a single primer with its 3’ end one base from the target SNP was extended by one base complementary to the SNP. I used the SNaPshot kit (Life Technologies, Foster City, CA) that includes fluorescently labeled dideoxynucleotides that allow for extension by a single base and detection of the individual nucleotides in the product. SNaPshot reaction conditions were as follows: 25 cycles of 95°C denaturation for 10 sec, 55°C for 5 sec,
and 60°C for 30 sec. Each reaction mixture was then treated with calf intestinal phosphatase (New England Biolabs, Ipswich, MA) to remove the 5’ phosphate groups from unused dideoxynucleotides before being analyzed on an Applied Biosystems 3730 DNA Analyzer (Life Technologies, Foster City, CA) and the results compiled and quantitated with GeneMapper software (Life Technologies, Foster City, CA). Each of the four possible dideoxynucleotides was labeled with a different color fluorophore, allowing for quantitative detection of individual alleles in the amplified product. In heterozygous genomic DNA samples, the amounts of each allele are assumed to be equal. However, since different fluorophores may influence nucleotide incorporation rates and migration of the product through the capillaries in the ABI 3730, peak height ratios for DNA as well as cDNA were normalized to the mean of the DNA ratios. As an additional control for variation introduced by different fluorophores, two primers (one upstream and one downstream) were utilized independently for the SNaPshot reaction, allowing for the comparison of the results utilizing one pair of fluorescently labeled dideoxynucleotides to the other pair.

**Genotyping Methods**

Genotyping was accomplished by at least one of three methods: SNPlex (Life Technologies, Foster City, CA), SNaPshot, as described above, or GC clamp. The GC clamp genotyping assay is described in detail in (Papp et al., 2003). Briefly, I designed two different allele specific primers along with one common primer in the opposite direction. A GC clamp of 10-15 bases was introduced to the 5’ end of one of the allele specific primers, allowing the two different allelic PCR products to be differentiated by
their melting temperatures by an ABI 7000 Sequence Detection System real time PCR instrument.

**Quantitative Detection of Splice Isoforms.** The *DRD3* splice variants were measured following PCR amplification of cDNA with a fluorescent Fam-labeled exon 6 primer and an exon 8 reverse primer as described in Wang et al., 2006.

**Statistical Analysis**

Genomic DNA and cDNA peak height ratios were normalized as described above in “Allelic Expression Imbalance”. The presence of AEI was determined with normalized cDNA ratios compared to the normalized genomic DNA ratio. The standard deviation (S.D.) of the normalized genomic DNA ratio was determined and samples that had significant allelic ratio differences between cDNA and DNA (> 2 S.D.) were considered to show significant AEI. Haplotype tagging SNPs were selected with Haplovview 4.2 (www.broadinstitute.org/haplovview/haplovview; Barrett et al., 2005).

4.3 Results

**Allelic *DRD3* mRNA Expression**

Measurement of the mRNA generated from each allele requires the use of frequent marker SNPs located in the transcribed region. The *DRD3* transcript contains only one SNP with >10% heterozygosity, rs6280, which was used as the marker SNP for all of the AEI measurements. I measured allelic mRNA expression in samples heterozygous for rs6280 (n=29), and identified four samples that displayed significant AEI (Fig. 4.2). The fact that the AEI ratios go in both directions (i.e. <1 and >1) suggest that the marker SNP
is not in LD with a regulatory variant causing the AEI or that there may be more than one variant influencing mRNA expression.

Figure 4.2. Allelic DRD3 mRNA expression ratios. Ratios were normalized to genomic DNA and are mean ± SEM, one cDNA syntheses and at least triplicate measurements per sample. Samples with significant AEI are denoted with a * and were determined as described in Materials and Methods.

SNP Scanning

To scan for regulatory polymorphisms affecting mRNA expression, I genotyped a series of 11 SNPs in the DRD3 gene locus (Fig. 4.1) in all samples. Except for rs6280 (the marker SNP), SNPs were chosen based on their utility as haplotype tagging SNPs in the genomic region from 20 kb upstream to 20 kb downstream of DRD3: The 11 SNPs captured 30 of 44 (68%) alleles at $r^2 > 0.8$, as assessed with Haploview 4.2.
(www.broadinstitute.org/haploview/haploview; Barrett et al., 2005). No SNP tested was significantly associated with the observed AEI.

**Alternative Splicing of** **DRD3** **mRNA**

To test for genetic effects on splicing, I assessed the levels of the two major splice variants of **DRD3**; D3L and D3nf and detected only D3L (Fig. 4.3) in all samples. The absence of PCR product corresponding to D3nf was not due to complete failure of the PCR reaction, since virtually all samples displayed robust amplification of a single product corresponding to D3L (Fig. 4.3).
Figure 4.3. Quantitative detection of *DRD3* mRNA splice variants. The peaks correspond to PCR products of the size indicated on the X axis that have incorporated the fam-labeled primer. Peak heights are proportional to the amount of PCR product. Representative results from three NAc samples are shown.
4.4 Discussion

The present study indicates that there is at least one regulatory *cis*-acting variant with a modest effect on *DRD3* mRNA expression, although none of the 11 SNPs tested could account for the observed AEI. Measurements of *DRD3* mRNA splice variants indicate that D3nf is not expressed at detectable levels in the samples tested.

Since gene regulation is tissue specific, it is critical to examine the regulation of a candidate gene in tissues where it is functionally expressed. This is an especially important consideration for *DRD3*, due to its relatively low expression and high homology with other dopamine receptor genes. NAc was an ideal brain region for this study, since *DRD3* is most highly expressed in NAc and other limbic areas of the brain (Sokoloff et al., 1990; Bouthenet et al., 1991).

Significant AEI was observed in only 4 out of 29 samples tested, an indication that the marker SNP (rs6280) does not influence mRNA expression. That there was AEI in some samples suggests the existence of at least one *cis*-acting regulatory variant affecting *DRD3* mRNA expression. Moreover, the allelic ratios of the samples with significant AEI were not all in the same direction, indicating that rs6280 is not in LD with the causal variant and/or that there may be more than one causal variant.

Given the potential dominant negative effect of D3nf (Elmhurst et al., 2000; Karpa et al., 2000), alternative splicing of *DRD3* to form D3L or D3nf represents an important regulatory event that could have important biological consequences. Furthermore, it is well-documented that *cis*-acting genetic polymorphisms can modulate alternative splicing (Attaie et al., 1997; Ng et al., 2008; Zhang et al., 2007; Moyer,
manuscript in preparation). Given the functional significance of *DRD3* mRNA splicing and the evidence that cis-regulatory variants can affect alternative mRNA splicing, I assessed D3L and D3nf mRNA (cDNA) levels in postmortem human NAc samples that I had already genotyped a number of SNPs across the *DRD3* gene locus. I did not detect D3nf, a surprising finding since it has previously been detected at levels comparable to D3L in 12 different regions in postmortem human brain, including NAc (Schmauss et al., 1993). These disparate results could be due to a number of factors, including different experimental conditions (primers, PCR conditions) or differences in the populations tested including age, sex, race, or postmortem interval and condition of the samples.

In the present study I assessed allelic mRNA expression and alternative splicing in human brain autopsy NAc obtained from a cohort of normal controls as well as patients diagnosed with schizophrenia, bipolar disorder, or depression (n=15 in each group). Significant AEI was detected in 4 out of 29 samples tested, indicating the existence of at least one *cis*-acting regulatory variant affecting mRNA expression. However, none of the 11 SNPs tested could account for the observed AEI. Evaluation of D3L and D3nf mRNA levels revealed detectable levels of D3L only, an indication that D3nf is not expressed in the samples tested. These results suggest the existence of a *cis*-acting regulatory variant affecting *DRD3* mRNA expression that is not in LD with any SNP tested and call into question whether D3nf mRNA is universally expressed in human NAc.
<table>
<thead>
<tr>
<th>Experiment</th>
<th>Oligos</th>
<th>Comments</th>
</tr>
</thead>
</table>
| SNaPshot (AEI measurement)     | Forward: 5'GAAGCCCCCTTGGGATCAC  
Reverse: 5'GCCCAACAGGATGTAGTTCAG  
PEP: 5'TCTCTGAGCCAGCTGAGT | Also used for rs6280 genotyping                                         |
| Splice Variant Detection       | Forward: 5'AAAGAGAGGAGAGAAGACTCGGAATTC  
Reverse: 5'TGGGTTATGGAGACATGGGTCA | Forward primer was Fam labeled                                           |
| rs9825563 GC Clamp Genotyping  | Forward: 5'ACTCTTTGACCCCCGTGG  
WT Reverse: 5'GAAGTCCAGAGAAAGGATCGCT  
SNP Reverse: 5'CGCGCCGGCGGCGGAAGTGCCAGAGAAAGGATCGCT | SNaPshot (AEI measurement)                                          |
| rs9824856 GC Clamp Genotyping  | WT Forward: 5'TAGAGCTCAGGCACATGCTCA  
SNP Forward: 5'GGCAGCTTCATGCTGACC  
SNP Reverse: 5'ACCACTACTAGAGGAGTTCAGAGGATCGCT | SnaPshot (AEI measurement)                                          |
| rs324036 SNaPshot Genotyping   | Forward: 5'TGCTAGGATTATAGGTTGAGGCA  
Reverse: 5'AATACCTGGAATGGGCTCGCA  
PEP: 5'GGCACTGTGCTGACCCTT | Also used for rs6280 genotyping                                         |
| rs167771 SNaPshot Genotyping   | Forward: 5'TCATCTCTTCCAGGGTGC  
Reverse: 5'GAATCAGACTGTTGGAGGATCGTCA  
PEP: 5'GAGCTATTTCATGCTGACCCTT | Also used for rs6280 genotyping                                         |
| rs12491239 SNaPshot Genotyping | Forward: 5'TGATATGAGCAATTGATGGAATGAG  
Reverse: 5'CCAAGCATTTATGTTGCTCTTCA  
PEP: 5'GCAGTTTTTCATGCTGACCCTT | Also used for rs6280 genotyping                                         |
| rs3732783 SNaPshot Genotyping  | Forward: 5'GCATCTCTGAGCCAGCTGAG  
Reverse: 5'GGCTGGCCACTGTTGGAGT  
PEP: 5'CTGAACACTACCTGTCGGGC | Also used for rs6280 genotyping                                         |
| rs2630351 SNaPshot Genotyping  | Forward: 5'GGGGCTGAGCCAGGATTAG  
Reverse: 5'AGAAGGCAGGATGTTGGA  
PEP: 5'CTATTGGACTCAGTTAGTAATCTTAA | Also used for rs6280 genotyping                                         |


Associated With Clinical Outcomes in Hypertension. *Clinical Pharmacology & Therapeutics* 85(1): 36-44.


