Candidate Gene and MicroRNA Expression Studies of Schizophrenia

by

Shannon E. Bruse

A Dissertation submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

and

The Graduate School of Biomedical Sciences

University of Medicine and Dentistry of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Microbiology and Molecular Genetics

written under the direction of

Dr. Linda Brzustowicz

and approved by

_____

_____

_____

_____

New Brunswick, New Jersey

[January, 2008]

Abstract of the Dissertation

Candidate Gene and MicroRNA Expression Studies of Schizophrenia

By Shannon Bruse

Dissertation Director:
Dr. Linda Brzustowicz

Schizophrenia is a debilitating psychiatric disorder affecting 0.5-1% of the world's population. Schizophrenia has a significant genetic component, but the early search for specific genes has been slowed by a number of factors, including a lack of inexpensive genotyping methods suitable for targeted genetic studies. This work describes the development of an inexpensive multiplexed genotyping assay ideal for candidate gene studies. Furthermore, we describe application of this technology to the study of three schizophrenia candidate genes - SNAP-25, ZDHHC8, and DGCR8. Both SNAP-25 and ZDHHC8 are involved in glutamate signaling at the NMDA receptor, and dysfunction in glutamate signaling at the NMDA receptor is one of the leading theories of the etiology of schizophrenia. DGCR8 is an RNA-binding protein necessary for the processing of microRNAs, a class of small RNAs important in the development and maintenance of the mammalian central nervous system. Though we failed to find evidence of genetic association with ZDHHC8, we did identify a genetic association between schizophrenia and variants within both the SNAP-25 and DGCR8 genes. Our results with DGCR8 provide the first genetic link between the microRNA biogenesis pathway and a major neuropsychiatric disorder. Additionally, we performed quantitative

microRNA expression profiling in post-mortem brain samples from patients with schizophrenia, bipolar disorder, and psychiatrically normal controls. We began with a pilot study comparing the expression of 158 microRNAs in the dorsolateral prefrontal cortex (BA9) of post-mortem brain tissue. Based on initial results, we assessed the expression of a number of specific microRNAs in a large, well characterized sample set from patients with schizophrenia (n=35), bipolar disorder (n=35), and psychiatrically normal controls (n=35). In the larger sample set, we observed overexpression of one microRNA, hsa-mir-372. This overexpression was observed in two separate brain regions from patients with both schizophrenia and bipolar disorder, though no differential expression of any microRNA tested was deemed statistically significant.

Dedication

This dissertation is dedicated to the following people:

First and foremost, my wife Kristy Lake-Bruse, for unfailing loyalty and love

My daughter, Amanda Bruse, who motivates me more than anyone

My mother, Susan Bruse for her love and support through the years

My sisters and brother, Cherie, Kelly, and Patrick for moral support and so much more

My grandparents, Walter and Bernice Knockle and Wil and Laura Bruse, who taught me important lessons

Mike Ennis, departed friend, whose wisdom and guidance meant so much

And John Mullen, longtime friend and fellow bibliophile

Thank you to all

Acknowledgements

Many people have contributed to this work in small and large ways, and I would like to thank Marco Azaro and Irina Tereschenko for design of the multiplexed PCR primer sets; Ray Zimmerman, Jared Hayter, and Jaime Messenger for technical assistance in the lab; Naomi McGregor and Pilar Garavito for helpful discussions; Steve Buyske for statistical analysis of microRNA expression work; Slava Saviouk for much of the PPLD analysis and help with clinical aspects of this work, including help in writing the descriptions of diagnostic categories in Chapter 5; Ron Hart and Loyal Goff for early help with microRNA expression profiling; Deborah Levy for access to the McLean samples and for help in writing a description of the quantitative phenotype in Chapter 5; my committee members Bonnie Firestein, Chris Rongo, and Steve Buyske for guidance and helpful advice along the way; and of course all of the patients and their families who made these studies possible.

Very special thanks go to Michael Moreau, fellow graduate student, friend, and close scientific collaborator on all of the work presented herein.

Finally, I am deeply grateful to my advisor, Linda Brzustowicz, for her constant support and encouragement, for allowing me to pursue my own ideas, for teaching me the intricacies of grant writing, and for helping me to grow as a scientist and as a person.

Table of Contents

Lists of tables

List of figures

Chapter 1

Introduction

Schizophrenia – general background

Schizophrenia is a debilitating psychiatric disease with a lifetime risk of 0.5-1% [1]. Studies indicate that this risk is uniform across many cultures [2] and that prevalence has remained stable since Emil Kraepelin first characterized "dementia praecox" in 1893 [3, 4]. In 1908, Eugen Bleuler formally introduced the term *schizophrenia*, and recognized that it was not a classical dementia in that some patients recovered or became less ill over time. Schizophrenia is a spectrum disorder related to schizoaffective disorder, schizotypal personality disorder, and paranoid personality disorder [5]. Symptoms of schizophrenia are generally categorized as positive, negative, or cognitive. Positive symptoms include hallucinations (generally auditory) and delusions, which frequently involve ideas of persecution or megalomania. Negative or deficit symptoms include affect flattening, anhedonia, apathy, poverty of speech, and withdrawal from social life. Cognitive symptoms include impairments in memory and problem solving. Schizophrenia occurs equally in males and females, though the typical age of onset is generally between 20-28 for males, and 26-32 for females [6]. The prognosis for those suffering from this illness is difficult to assess, given variable definitions of outcome. One study estimated that about one-third of patients make a full recovery, one-third show improvement, and one-third remain seriously ill [7]. Another study estimated a significant percentage of patients have favorable outcomes, with nearly 50% not having psychotic episodes within the previous 2 years at 15 and 25-year follow up assessments [8]. However, absence of

positive symptoms does not imply a full recovery, as negative and cognitive symptoms can be severely debilitating. In a survey study of health professionals, caregivers, and family members, active psychosis was rated the third most debilitating condition, ahead of blindness and paraplegia [9]. Schizophrenia is associated with significantly increased mortality and suicide, and a wide range of social disability and personal suffering [10].

Genetic factors are likely to play an important role in disease etiology, given that monozygotic twin concordance is 41%-65% and dizygotic concordance is 0%-17% [11], with heritability estimates of approximately 80% [12]. In addition to genetic factors, there is a significant environmental component to this disease, and varying degrees of evidence implicate poverty, infectious agents, cannabis smoking, urban living, and emotional stressors as possible environmental factors (reviewed in [13]). Additionally, stochastic developmental processes have been hypothesized to contribute to this disease [13]. Linkage studies provide an unbiased method of locating genes or variants involved in schizophrenia, and have resulted in findings of significant and suggestive linkage in many chromosomal regions. Some of these linkage regions have been replicated in further studies, though no region has been consistently replicated in a majority of studies. The most plausible explanation for this lack of consistency is genetic heterogeneity. Schizophrenia is thought to be a complex disorder, so that combinations of polymorphisms in multiple genes are required in order for the schizophrenia phenotype to emerge. Therefore, studies utilizing different populations might identify different regions of linkage. In addition, within a study there may be more than one "genetic population", so that linkage signals that are present in a subset of families can be diluted by other genetic subsets. Also potentially contributing to the lack of replicability is the fact that linkage studies lack power to detect genes of modest effect [14]. Environmental and

epigenetic factors are also likely confounders. In spite of these obstacles, two meta-analyses conclude that there are a number of linkage regions statistically likely to be genuine [15, 16], and association studies within many of these regions have revealed several strong schizophrenia candidate genes.

Disease etiology

Unlike linkage studies, association or candidate gene studies are not unbiased surveys of the genome, but are often informed by previous linkage studies and by the "biological plausibility" of genes. For schizophrenia, the biological processes most emphasized by researchers are neurodevelopment and synaptic transmission, though consistent pathophysiological findings have been difficult to identify (reviewed in [17]). Brain regions most consistently associated with schizophrenia-induced abnormalities are the hippocampus, association neocortex (prefrontal and superior temporal), and thalamus [18-20]. Reports also indicate a reduction in brain volume and weight, along with ventricular enlargement [21, 22]. This does not appear to be the result of a decreased number of neurons, at least not in the hippocampus and cerebral cortex. However, cell bodies of pyramidal neurons in the hippocampus and neocortex are smaller, and these same neuron populations have fewer dendritic spines and reduced dendritic arborizations [23-25]. Probably the most notable observation is of aberrantly clustered neurons in the lamina, entorhinal cortex, and the neocortical white matter, indicating an early developmental anomaly [26]. In addition, there is a decrease in the number and function of oligodendrocytes [27, 28]. It should be kept in mind that the pathophysiological changes observed are generally subtle, with overlapping distributions between measurements from post-mortem brain samples of schizophrenic patients and those of

non-schizophrenic patients. In other words, there is no pathophysiological measure that is diagnostic of schizophrenia, and while schizophrenia may indeed be a developmental disorder, consistent and profound anatomical and physiological malformations have been difficult to identify.

It is possible that perturbations in synaptic transmission and/or neurodevelopment involve changes not observable at gross levels of anatomy and physiology, and only by looking at the molecular level are defects observed. Indeed, post-mortem gene expression studies have indicated alterations in a variety of signaling pathways, including dopamine, glutamate, gamma-aminobutyric acid (GABA), and opioid signaling. The question of whether these changes are primary (causing schizophrenia) or secondary (caused by schizophrenia) has not been answered by these studies, nor have the findings from post-mortem gene expression studies been robust. Genetic studies increasingly implicate glutamate, and to a lesser extent dopamine, in the etiology of schizophrenia. Historically, dopamine has received the most attention. Briefly, the dopamine hypothesis states that excess dopamine signaling in distinct anatomical brain regions (mainly the prefrontal cortex) is responsible for schizophrenia. Two key pharmacological observations support this hypothesis – drugs that cause release of dopamine (amphetamines) can trigger psychosis, and most drugs used to treat schizophrenia block dopamine (D2) receptors. However, these D2 antagonists seem to alleviate the positive but not the negative symptoms of schizophrenia. More recently, dysfunction in glutamate transmission has been gaining attention. Generally, the glutamate hypothesis posits hypofunction in glutamate signaling through the N-methyl-D-aspartic acid receptor (NMDAR) as a causative factor in schizophrenia. Noncompetitive NMDAR antagonists, such as phencyclidine (PCP) and ketamine, mimic a fuller range of symptoms than do

amphetamines, and significantly, can induce both positive and negative symptoms [29]. These NMDAR antagonists also produce an increase in dopamine in the prefrontal cortex, arguing that the primary defect is in glutamate signaling and that defects in dopamine signaling are secondary [30]. Normal patients administered ketamine exhibit both positive and negative symptoms **and** perform normally on the Mini-Mental Status Exam, mimicking results observed in schizophrenic patients [31]. Additionally, adjunctive treatment with NMDAR agonists such as glycine or D-serine can improve the symptoms of schizophrenia [32]. It is important to consider that blocking the NMDAR results in an increased release of glutamate in the prefrontal cortex. This glutamate is potentially available to be used at α-amino-3-hydroxy-5-methylisoxazole-4- propionic acid (AMPA), kainate, and metabotropic receptors, so that considering only the NMDAR in disease etiology may be an oversimplification [33]. At this time most of the candidate genes (discussed below) identified by genetic studies are involved in some way with signaling at the NMDAR, making genes involved in NMDAR signaling prime candidates for further genetic investigation.

The NMDAR complex contains over 160 known proteins that can be grouped into five classes: neurotransmitter receptors, scaffolding proteins, adaptors, signaling enzymes and cytoskeletal proteins. In addition, there are many proteins in the presynaptic cells and in the supporting oligodendrocytes that influence NMDAR signaling. These include transporters, enzymes, and endogenous ligands of the glycine modulatory site. The NMDAR is comprised of a combination of the NR1 and NR2 (A-D) subunits [34]. Postsynaptic density-95 (PSD-95) is one of a family of four (others are PSD-93, SAP 102, and SAP 107) PDZ-containing membrane-associated guanylate kinase (MAGUK) proteins that bind to the NMDA receptor. PDZ domains are involved in protein-protein

interactions, and are often found in scaffold proteins that help to form large molecular complexes. At the post-synaptic density the PSD-95 family helps to anchor and cluster NMDA receptors, and is likely involved in protein trafficking and second-messenger signaling through binding of proteins to the PDZ (as well as SH3 and GK) domains. More than one study has found altered expression of PSD-95 and NMDAR subunit proteins in patients with schizophrenia [34]. Interestingly, PCP applied to primary neuronal/glial cell cultures results in increased expression of NR2B and decreased expression of NR2A and PSD-95, which is the pattern seen in at least one post-mortem study of patients with schizophrenia [35].

Candidate gene studies

In addition to the pharmacological and expression data outlined above, candidate gene studies also point to a role of glutamate signaling at the NMDAR in the etiology of schizophrenia. Fine mapping of a strong linkage region (LOD = 6.50) has identified nitric oxide synthase 1 adaptor protein (NOS1AP; formerly known as CAPON) as a schizophrenia candidate gene [36]. Neuronal nitric acid synthase (nNOS) binds to a PSD-95 PDZ domain, and in turn has its own PDZ domain that binds to NOS1AP. NOS1AP is thus involved in NOS signaling through the NMDAR/PSD-95 complex. Like NOS1AP, other candidate genes are involved in glutamate signaling. Genetic studies and a transgenic mouse model suggest that the zinc finger DHHC domain containing 8 gene (ZDHHC8), which palmitoylates PSD-95 and perhaps PSD-93, is associated with schizophrenia [37]. It is known that synaptic strength at the NMDAR is regulated by palmitate cycling on PSD-95, and there is evidence that ZDHHC8 palmitoylates PSD-95

[37, 38]. Multiple genetic studies have associated Neuregulin1 (NRG1) with schizophrenia, and a recent report shows that Neuregulin1 regulates transcription of PSD-95 [39, 40]. Dysbindin1, part of the dystrophin glycoprotein complex, is another strong candidate gene that acts at the glutamatergic synapse. Dysbindin1 has been associated with schizophrenia in 11 separate genetic studies using a diversity of ethnic groups [41]. D-amino-acid oxidase (DAAO) and D-amino acid oxidase activator (DAOA) together degrade D-serine, a potent agonist of the glycine modulatory site at the NMDA receptor. Both genes have been associated with schizophrenia [42]. PSZA11q14, a proposed anti-sense regulator of PSD-93, has also been associated with the disease [43]. The large number of identified candidate genes involved in glutamate signaling strengthens the glutamate hypothesis of schizophrenia. However, it should be noted that association studies of glutamate/NMDAR genes, like linkage analyses, have been difficult to replicate. Despite the findings of positive association outlined above, there is in each case at least one subsequent study (when multiple studies are available for a particular candidate gene) which report finding no variant within the gene that is associated with disease. Also, separate studies have reported opposite alleles as the disease causing variants [17]. In fact, despite intense discovery efforts, no allele in any gene has been established as an unequivocal disease-causing variant. While evidence for the glutamate hypothesis is accumulating, it is by no means fully established as a causative factor in schizophrenia.

A portion of the work described herein involves the study of two genes involved in glutamate signaling. Genes were selected from a list compiled using various sources, including current literature and protein interaction databases, to identify proteins that function at the NMDA receptor. We focused on the post-synaptic density, as well as

presynaptic and other more peripheral proteins [44]. Below (Table 1.1) is a list of NMDAR-associated genes located in linkage regions as defined by the meta-analyses or by Brzustowicz et al. al. [45].

Table 1.1. Identified and novel schizophrenia candidate genes located within regions of established genetic linkage

Meta-analysis regions	Identified candidate genes	Novel candidate genes
2p12-q22.1		Sema4C (2q11.2)
		Rab3A (2q21.3)
5q23.2-q34		Alpha-CamKII (5q32)
3p25.3-p22.1	Synapsin2 (3p25)	ATP2B2 (3p25.3)
11q22.3-q24.1		
6pter-p22.3	Dysbindin1 (6p22.3)	SynGap1 (6p21.32)
2q22.1-q23.3		
1p13.3-q23.3*	NOS1AP (1q23.3)	
22pter-q12.3	ZDDHC8 (22q11.21)	Stargazin (22q12.3)
	COMT (22q11.21)	
8p22-p21.1*	Neuregulin1 (8p21-p12)	Pyk2 (8p21.1)
	PPP3CC (8p21.3)	FAK2 (8p21.1
6p22.3-p21.11		
20p12.3-p11		SNAP-25 (20p12-p11.2)
14pter-q13.1		
Brzustowicz et. al. regions		
13q22-q32		
1p21.1-p13		

Table 1.1 continued		
2p25.1		
11q13-q23	**PSD-93 (11q14.1)**	**Neurexin2 (11q13.1)**
		SHANK1 and 2 (11q13.4)
		Calpain (11q13)
		MLK-3 (11q13.1-q13.3)
3q13.12		
Table 1.1 continued		
17p12-p11		**Dexras1 (17p11.12)**
12q22-24	**DAAO (12q24)**	**Citron (12q24)**
*** = also Brzustowicz et al**		

In light of the glutamate hypothesis of schizophrenia, two genes were selected (in red, **Table 1.1**) for inclusion in genetic studies based on their involvement in NMDAR signaling. First, we performed a candidate gene study of SNAP-25, a well-known gene encoding for a protein broadly involved in presynaptic release of neurotransmitter, including glutamate. A second study involved investigation of the gene encoding for the zinc finger DHHC domain-containing protein 8 (ZDHHC8), a palmitoyl transferase involved in palmitoylation of PSD-95. ZDHHC8, which had previously been associated with schizophrenia through a family-based genetic study, is contained in a schizophrenia susceptibility region on chromosome 22. This susceptibility region has long been hypothesized to contain one or more genetic loci contributing to schizophrenia. We

performed a candidate gene study of a third gene, DiGeorge critical region-8 (DGCR8), which is located in close proximity to ZDHHC8 and is discussed in more detail below.

TagSNP selection

When performing candidate gene studies, the question of which genotyping method to utilize and which single nucleotide polymorphisms (SNPs) to genotype is not trivial. Choosing the wrong SNPs can result in failure to find genetic association even when that association is present, and existing methods of genotyping can be costly and time consuming. This work will describe development of an in-house multiplexed genotyping method that is accurate, rapid, and cost effective. Additionally, a tagSNP approach was utilized in all candidate gene studies to reduce the genotyping burden. Randomly choosing SNPs at a reasonably high density of 1 SNP per 5 KB excludes at least 80% of available SNPs, and *a priori* reduces the power of a study. However, intelligent utilization of linkage disequilibrium (LD) information can greatly offset this loss of power. It is well known that LD decays with distance; therefore, SNPs in relatively close proximity to each other may share significant LD. The task, then, is to choose a region that is sufficiently small, or a density of SNPs that is sufficiently large, so that the genotyped SNPs are likely to be in LD with the unknown, disease-causing variant. However, for a typical candidate gene this can still mean genotyping prohibitively large numbers of SNPs with no assurance that the SNPs chosen are in LD with the disease causing variant. Researchers have therefore developed SNP-tagging methods that utilize the information from experimentally determined LD patterns to help select a minimal set of SNPs that provide a maximum amount of information. During the course of our studies, LD databases, such as those offered by HapMap, Perlegen, and Celera, have been

continually refined and improved. Therefore, our tagSNPs were selected from whichever database was deemed most complete at the time a study was undertaken, and our strategies are outlined in the Methods sections of the appropriate chapters.

MicroRNAs

A gene encoding for the protein DGCR8 is located in very close proximity to ZDHHC8, but has not previously been considered a gene of interest and thus has not been investigated as a schizophrenia susceptibility gene. DGCR8, as a gene, is relatively recently characterized, and its function even more recently established. DGCR8 is involved in processing of microRNAs (miRNAs), a novel class of small RNAs with a widespread role in mammalian gene regulation. Given the lack of an established model of the pathogenesis of schizophrenia, in spite of decades of intense research, we were eager to explore aspects of human biology or classes of candidate genes not previously studied. Evidence that small RNAs might play an unexpected role in the regulation of gene expression began to emerge during the latter half of the 1990's. In 1998, Craig Mello and Andrew Fire's seminal paper showed that double stranded RNA (dsRNA) was a potent and specific regulator of gene expression [46]. Surprisingly, exogenous dsRNA fed to *C. elegans* was shown to degrade homologous, endogenous mRNA, resulting in post-transcriptional gene silencing. This phenomenon was termed RNA interference (RNAi). It soon became apparent that quelling in *Neurospora crassa* and post-transcriptional gene silencing in plants (PTGS) were RNAi mediated. RNAi has since been demonstrated in a wide diversity of organisms, from simple yeast to mammalian cells [47-49]. While RNAi does not occur naturally in mammals, these simple but powerful experiments spurred an

explosion of interest in RNA as a functional molecule, and have changed our understanding of the landscape of the genome. The ENCODE project, which seeks to annotate the functional elements of the genome, paints a complex picture of massive, overlapping transcription occurring on both strands and in both directions, showing little respect for conventional boundaries that define protein coding genes [32]. The FANTOM consortium, whose aim is to provide a comprehensive picture of the mammalian transcriptome, has estimated that 63% of the mouse genome is transcribed, yet only 1.5% of the genome is for protein coding genes [50]. It is unclear how many of these non-protein coding transcripts are functionally important, but certain classes of non-coding RNAs are undoubtedly functional. Emerging experimental evidence indicates that non-coding RNAs are likely to be particularly important in the central nervous system (CNS) [51]. The classes of non-coding RNAs known to be functionally relevant in the CNS include miRNAs, small nucleolar RNAs, longer non-coding RNAs, transfer RNAs, ribosomal RNAs, and RNA trinucleotide expansions [51]. These RNAs regulate gene expression by functioning in a diversity of processes, including chromatin modification, translation, transcription, alternative splicing, imprinting, and RNA editing.

MiRNAs are an important class of non-coding RNAs that function in post-transcriptional gene regulation by targeting mRNAs for translational repression or degradation. While miRNAs were discovered by Victor Ambros in 1993 [52] - prior to the RNAi experiments of Mello and Fire - the importance of miRNAs in animal gene regulation didn't emerge until 2001 [53, 54]. There are currently 533 human miRNAs in the Sanger database, and estimates of the number of human miRNAs are as high as 800 [55]. They tend to be located in clusters throughout the genome, and are transcribed by RNA polymerase II [56]. A global survey of their genomic locations reveals that more

than half are located within introns of known protein coding genes [57]. Initially, it was thought that miRNA targeting in animals involves imperfect base-pairing between miRNA and mRNA targets, resulting in repression of protein expression and no alteration in mRNA expression level. This stood in contrast to plants, where perfect Watson-Crick complentarity between miRNA and mRNA targets results in endonucleytic cleavage and degradation of the mRNA. However, while it is true that the vast majority of animal miRNA/mRNA target pairs involve imperfect base-pairing, recent evidence suggests that even imperfect base-pairing impacts mRNA expression levels through a deadenylation process [58, 59]. The specificity of miRNAs for their target mRNAs is largely determined by the first 1 to 8 base pairs at the 5' end of the miRNA (the "seed" region) [60]. Given that an individual miRNA may have hundreds of targets, it is likely that a significant portion of protein coding genes are under miRNA regulation.

Processing of miRNAs is similar to that of dsRNA in the RNAi pathway. The transcripts giving rise to miRNAs are called pri-miRNAs, which in some instances can double as mRNAs [61]. An RNAse III enzyme, Drosha, cleaves the pri-miRNAs to form a 65 nucleotide hairpin structure known as pre-miRNA [62]. It has recently been discovered that DiGeorge Syndrome Critical Region Gene 8 (DGCR8) is a necessary co-factor for Drosha processing [63]. The pre-miRNA then associates with Exp5 and Ran-GTP and is exported from the nucleus to the cytoplasm [64]. At this point the processing of miRNAs and dsRNAs (which give rise to siRNAs, the counterpart to miRNAs) is thought to be identical. Another RNAse III enzyme, Dicer, cleaves the hairpin to give a 18-23 base pair double-stranded RNA [65]. The miRNA is quickly unwound to a single-stranded RNA which associates with an miRNA-induced silencing complex (miRISC), which contains proteins from the Argonaute complex [66].

Little is known about specific functions of miRNAs in mammals, though some examples exist. A pancreatic islet-specific miRNA, mir-375, regulates insulin secretion [67]. It likely does so through control of myotrophin expression, which is involved in glucose-induced insulin exocytosis. Another miRNA, mir-143, has been shown to influence hematopoesis in human cell cultures [68]. Many miRNAs have striking tissue specific expression patterns, including a number of clusters that are expressed only in the brain and only in human embryonic stem cells [69]. In fact, early on it was recognized that miRNAs are abundantly expressed in developing and adult mammalian brains, often with spatiotemporal specificity, strongly implying functionality [70, 71]. Though the function of the vast majority of brain-expressed miRNAs is unknown, studies are emerging which show specific functions in the brain. MiRNAs have emerged as important regulators of animal neurodevelopment [72]. An *in vitro* model of mouse embryonic stem cell derived neurogenesis identified specific miRNAs important in progenitor differentiation to neurons or astrocytes [71], and *in vitro* results correlated with expression patterns of *in vivo* mouse embryonal neurogenesis [70, 73]. A recent paper demonstrated that mir-134 regulates dendritic spine size in the adult rat brain through targeting of the LIMK1 gene [74]. Mir-430 has been shown in zebrafish to be important for proper neural tube formation, neural stem cell maintenance, and axonal pathfinding [75]. Target prediction algorithms report enrichment for genes important in the CNS. One study took note of the fact that there was enrichment in targets at the NMDAR complex, including all NMDAR subunits, all PSD-95 family mRNAs (DLG1-4), and certain PSD-95 associated mRNAs [76]. This represents intriguing (though speculative) evidence for the importance of miRNAs at excitatory synapses. This is

notable given the evidence, presented earlier, for involvement of this signaling complex in schizophrenia.

In addition to the above identified functions in the human brain, there have also been tentative links between miRNAs and neuropsychiatric disease. A mutation in SLITRK1, which alters the hsa-mir-189 binding site, has been associated with Tourrette's Syndrome [77]. Fragile X mental retardation protein (FMRP) has been shown, through biochemical and genetic experiments, to associate with the miRNA RISC complex, hinting at potential involvement of miRNAs in Fragile X syndrome [78]. The FMR1 gene (FMRP is the protein) causes over 99% of fragile X syndrome cases. A working model of FMRP function is that it is involved in mRNA transport and local protein translation at the synapse. Similar to human patients with schizophrenia, FMR1 knockout mice have altered dendritic spine morphology, though post-mortem samples from patients with schizophrenia indicate fewer dendritic spines than average while FMR1 knockout mice have more [79]. Interestingly, the same study which took note of increased targeting of the NMDA receptor complex also predicted a significant enrichment in the RNA cargoes of FMRP [80]. This study predicted that miRNAs, which comprise 1% of all human genes, regulate 10% or more of all protein coding genes (though later studies have revised this estimate upwards to 30-100%). There are approximately 400 mRNA cargoes of FMRP, and about 300 of them were strongly predicted targets of miRNAs. Given that FMRP is associated with the RISC complex, this suggests a model where FMRP is an adaptor whose cargo RNAs are regulated through interaction with the miRNA pathway, providing an as yet unverified model that allows for rapid translational control at the synapse.

As outlined earlier, leading theories of schizophrenia propose defects in the processes of synaptic transmission and/or neurodevelopment. Specific miRNAs have been shown to play an important role in both of these complex and interrelated biological processes, and many more miRNAs whose functions have not been determined are likely to play a role as well. Many primate-specific miRNAs have been identified [81] whose function has not been determined, making this class of genes attractive candidates for involvement in disorders of brain function, including schizophrenia. Our work on miRNAs can be divided into two parts: (1) a candidate gene study of the miRNA-processing gene DGCR8 and (2) expression profiling of mature miRNAs in post-mortem brain samples. For both the candidate gene and expression studies, the sample sets available allowed us to investigate both schizophrenia and bipolar disorder. This is significant given that multiple lines of evidence from linkage, association, functional imaging, and gene expression studies point to some shared etiology between schizophrenia and bipolar disorder (reviewed in [82-87]), with current thinking favoring the hypothesis that these disorders share genetic susceptibility loci.

In summary, we will describe the development of a genotyping assay that allowed us to perform rapid and inexpensive candidate gene studies. Furthermore, we describe application of this technology to the study of three schizophrenia candidate genes - SNAP-25, ZDHHC8, and DGCR8. Both SNAP-25 and ZDHHC8 are involved in glutamate signaling at the NMDA receptor, and dysfunction in glutamate signaling at the NMDA receptor is one of the leading theories of the etiology of schizophrenia. DGCR8 is an RNA-binding protein necessary for the processing of miRNAs, and represents a novel area of neuropsychiatric research. In addition to the genetic study of DGCR8, we

performed miRNA expression profiling of post-mortem brain samples from patients with schizophrenia, bipolar disorder, and psychiatrically normal controls.

Chapter 2

Development of an Inexpensive Bead-Based Oligonucleotide Ligation Assay for SNP Genotyping

Introduction

Single nucleotide polymorphisms (SNPs) represent a major source of genetic variation in human beings; the NCBI database dbSNP contains nearly 12 million unique SNPs, which correlates to an average of 1 SNP every 250 base pairs. Due to their frequency, SNPs are an important type of marker used in association studies of human disease, and can themselves induce functional changes that contribute to disease and drug response variability [88, 89]. In a relatively short time, SNP genotyping technologies have evolved from rudimentary and simplexed gel-based assays to highly multiplexed arrays designed to capture most of the genetic variation within an individual sample. Ultra-high throughput array-based genotyping technologies, such as those offered by Illumina and Affymetrix, are utilized in genome-wide studies requiring genotyping of hundreds of thousands of SNPs in hundreds (or low thousands) of samples. However, fine-mapping, candidate gene, and other targeted genetic studies often require genotyping of dozens or hundreds of SNPs in hundreds or thousands of individuals, and ultra-high throughput methods are not ideal for these applications. Though too numerous to list in their entirety, popular medium-throughput commercial methods include simplexed technologies such as Biotage's pyrosequencing, Third Waves's Invader assays, and ABI's TaqMan SNP assays; multiplexed technologies include Beckman-Coulter's SNPstream and ABI's SNPplex platforms. Drawbacks of commercial genotyping platforms include equipment

costs and the need to purchase proprietary reagents. The genotyping method presented here utilizes the Luminex flow cytometer, which is within the budget of many academic labs, and is a relatively open source platform, readily amenable to "home-brew" assays not requiring expensive and proprietary reagents. Additionally, unlike many SNP genotyping platforms, the Luminex flow cytometer is one of the few instruments capable of quantitating both proteins and nucleic acids (DNA, mRNA, and miRNA) [90-93], providing it with uses beyond the SNP genotyping assay presented here.

Our method uses the robust chemistry of the oligonucleotide ligation assay (OLA) in conjunction with the Luminex flow cytometry platform (**Figure 2.1**) [94-99]. The assay is performed in 96-well plate format and has an upper multiplex capability of 50 SNPs per well, making it suitable for a moderate number of SNPs in a large number of samples. Two key modifications make this method less expensive and easier to perform than similar published methods. First, we use fewer beads than are traditionally used in quantitative Luminex assays [100-103]. The most significant expense of typical Luminex-based assays is the polystyrene beads used. Typical recommendations for SNP typing assays are to use 2500 input beads while counting 100 beads, but our experience indicates that these numbers can be substantially reduced, with a significant savings in cost per genotype.

Second, a universal biotinylated oligonucleotide was employed for signal quantification. For a biallelic SNP the OLA assay requires two allele specific oligonucleotides and a common oligonucleotide. Traditionally, the common oligonucleotide is labeled (during synthesis) with a detector molecule (i.e. biotin), which allows for quantification of the OLA product. In contrast, we synthesize the common oligonucleotide to have an attached sequence which is complementary to a universal

oligonucleotide double-labeled with biotin, thus saving on the cost of synthesizing biotin-labeled common oligonucleotides for each SNP assay. In addition to being less expensive, the method presented here does not employ typical wash or centrifugation steps, making it simpler to perform.

Materials and Methods

Oligonucleotide primer and probe sets

The multiplex PCR primer selection program used in this study is a significant improvement over the program that was used to create 1000+ plex PCRs for parallel genotyping [104]. The algorithm uses a caching strategy to avoid redundant computations, enabling the interaction algorithm to run in O(n) time (i.e., the approximate run-time is linearly proportional to the number of input sequences) versus O(n^2) time (i.e., the approximate run-time is proportional to the square of the number of input sequences) for the older program. The newer program can assemble a set of 22,000 multiplex primers in approximately 16 hrs on a PC with a 1 GHz Pentium III processor and 1 GB of RAM. Amplicon size in multiplexed PCRs is generally below 400 base pairs, though we have successfully genotyped from amplicons as large as 1200 base pairs. PCR primers were ordered in 96-well plate format (IDT, Coralville, Iowa) and purified using standard desalting.

The OLA requires two allele-specific and one common probe for each SNP being assayed. Allele-specific probe pairs consist of a 5' tag sequence and a 3' locus specific portion which differs only at the terminal position of the probe. Each allele specific probe contains a unique 24-base FlexMAP™ tag (Luminex® Corporation, Austin, TX) at the 5' end to allow hybridization to a reverse complement anti-

tag coupled to a unique FlexMAP™ microsphere. Common probes contain a locus specific portion at the 5' end, a universal capture sequence at the 3' end, and are 5' phosphorylated by the manufacturer (Integrated DNA Technologies, Coralville, IA). The universal capture sequence is the reverse complement of a doubly biotinylated "universal oligonucleotide" which is included in the bead hybridization step. For the universal capture sequence we employed the "tag" of FlexMAP™ bead 100 and the "universal oligonucleotide" sequence is the anti-tag of bead 100; for this reason, FlexMAP™ bead 100 is not utilized in our assays. All OLA probe sets are designed so that the locus-specific portion of both the common and allele-specific probes have a melting temperature of approximately 64 C. Module 1 of the HyTher server was used to determine melting temperature (http://ozone3.chem.wayne.edu/). OLA probe sets were ordered in 96-well plate format and purified using standard desalting. All primer and probe sets are contained in Supplementary Material 1 – Primer and Probe Sets.

Preparation of DNA and Multiplex PCR Reaction

Genomic DNA was isolated from whole blood using the QIAamp Blood Kit (Qiagen, Santa Clara, CA). Amplifications were performed in a solution (30 μl) containing 1.5 U AmpliTaq Gold DNA polymerase (Applied Biosystems); 1:10 dilution of 10X Buffer II (supplied with AmpliTaq Gold); 200 μM of each dNTP; 166 nM of each primer; 2.5 mM MgCl; and 40 ng genomic DNA. Reactions were initially heated at 94 ˚C for 10 minutes, followed by 40 cycles of 94 ˚C for 40 seconds, 60 ˚C for 30 seconds, ramping from 60 ˚C to 65 ˚C using fifty 0.1 ˚C two second cycles, and 72 ˚C for 2 minutes. Reactions were completed at 72 ˚C for 10 minutes and 20 ˚C for 5 minutes. All reactions were performed in 96-well plates.

Oligonucleotide Ligation Reaction

OLAs were performed in a solution (15 μl) containing 20 mM Tris/HCl buffer pH 7.6; 25 mM KOAc; 10 mM MgOAc; 1 mM NAD^+; 10 mM DTT; 0.1% Triton X-100; 10 nM (200 fmol) of each OLA probe; 2 μl of multiplexed PCR product; and 3 U of Taq DNA ligase (New England Biolabs, Beverly, MA). Reactions were initially heated for 1 minute at 95 ˚C, followed by 32 thermal cycles of 95 ˚C for 15 seconds (denaturation) and 58 ˚C for 2 minutes (annealing/ligation). Reactions were then cooled to 4 ˚C, and were used immediately in the hybridization step or stored at -20 ˚C for up to 1 week before proceeding with the hybridization step. All reactions were performed in 96-well plates.

Hybridization of OLA Products to Microspheres

50 μl of TMAC hybridization solution (3 M tetramethylammonium chloride, 50 mM Tris-HCl, pH 8.0, 3 mM EDTA, pH 8.0, 0.10% SDS) containing 200 beads (unless otherwise noted) from each FlexMAP™ microsphere set was added directly to a well containing a completed 15 μl OLA reaction. Hybridization reactions were then denatured at 95 ˚C for 90 seconds and hybridized at 37 ˚C for 20 minutes.

Fluorescent Labeling and Flow Cytometric Analysis

Following hybridization, 6 μl of TMAC hybridization solution containing 180 ng streptavidin-R-phycoerythrin (Molecular Probes – Invitrogen Corporation, Carlsbad, CA) was added directly to the well containing the hybridization reaction. Labeling reactions were incubated at 37 ˚C for 40 minutes. Care must be taken to avoid exposing SA-PE to

excessive amounts of light. Unless otherwise noted, 20 of each FlexMAP™ bead included in the multiplexed assay were sorted and quantitated using a Luminex 100 or Luminex 200 flow cytometer.

Direct sequencing and pyrosequencing

Semiautomated fluorescent direct sequencing was performed on the Beckman CEQ 8000 instrument. Pyrosequencing assays were performed as simplexed reactions on the automated PSQ HS96A platform [105, 106]. PCR primers were designed using Primer3 (Whitehead Institute) and the sequencing primer used for the Pyrosequencing assay was designed using the Pyrosequencing SNP Primer Design Software v1.0. PCR reactions contained 40 ng of template DNA, 0.5 U Taq polymerase, 0.03 µM of each primer, 0.1 mM of dNTP, 3.0 mM of MgCl2, 50mM of KCl, 10mM of Tris—HCl (pH 9.0), and 0.1% Triton X-100 in a 20 µl volume. A touchdown amplification program was used. After 8 minutes at 95°C, 15 cycles were done at 94°C for 30 s, annealing for 20 seconds, and at 72°C for 20 seconds, with the annealing temperature starting at 60°C and decreasing by 0.5°C for each cycle. This was followed by 35 cycles at 94°C for 20 seconds, at 52°C for 20 seconds, at 72°C for 30 seconds, and then a final extension step at 72°C for 15 minutes.

Genotype calling for OLA/Luminex assay

Manual inspection of scatterplots was used to determine genotype cutoffs for each SNP assay. For assessment of concordance with other SNP genotyping methods, all Luminex OLA calls were made blinded to the results obtained from those other methods. Scatter

plots show the median fluorescence intensity (MFI) ratio of one allele on the Y axis versus sample number on the X axis. The allelic ratio is calculated as follows:

$$\frac{MFI(a_1)}{MFI(a_1 + a_2)}$$

Thus, a biallelic SNP can form three clusters representing the three possible genotypes. Assays not forming distinct clusters, and assays where a significant percentage of samples are below signal threshold, are considered failed assays. In general, the signal threshold for a "callable" genotype is an MFI, for either allele, of greater than 1000 (using the high PMT setting when calibrating the Luminex flow cytometer). Quality of clustering is a subjective determination, though a formal measure of cluster quality can be used; in this report Silhouette scores were used for assessing quality of clusters [107]. Silhouette scores measure the distances between and within manually assigned clusters, and Lovmar et. al. determined that "genotypes can be unequivocally assigned without manual inspection when the Silhouette score for a SNP assay is > 0.65". This threshold is likely to be similar across different genotyping platforms.

Results

As shown in **Figure 2.1**, this SNP genotyping assay can be divided into four steps: 1) multiplexed PCR, 2) multiplexed OLA, 3) hybridization/labeling 4) and detection using the Luminex flow cytometer.

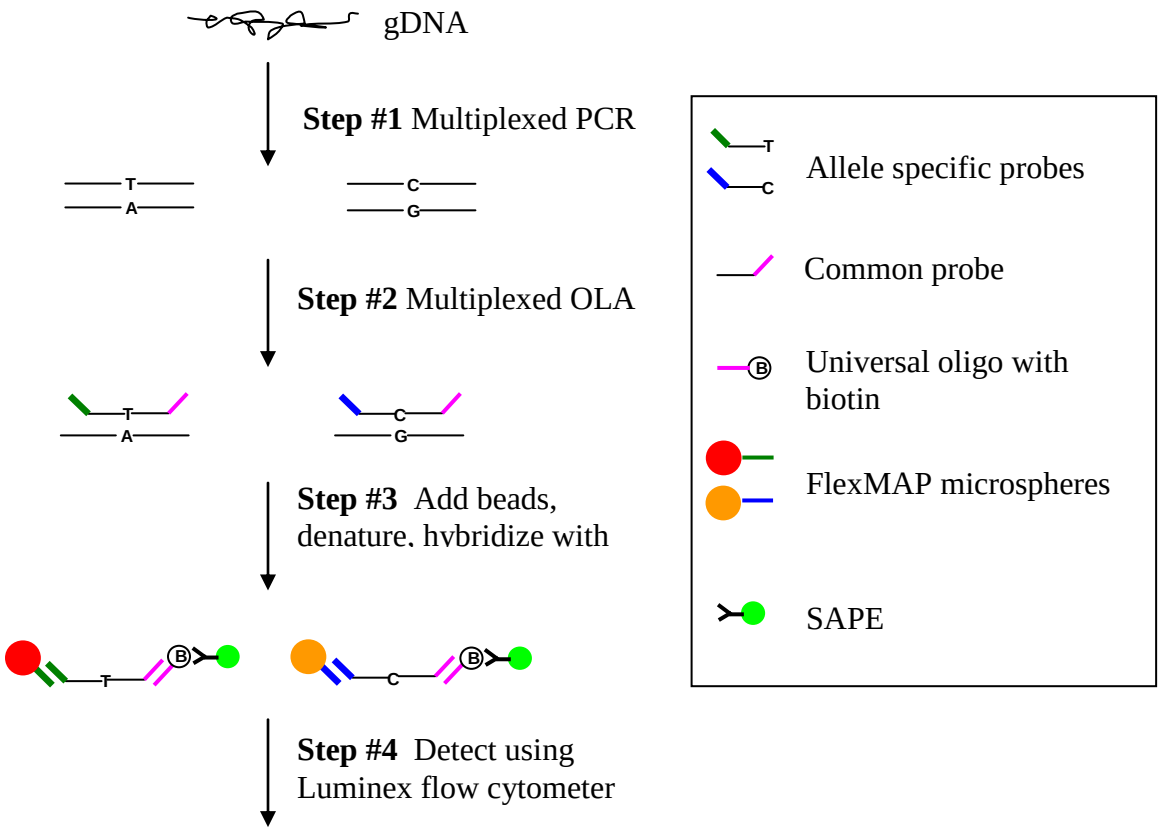

Figure 2.1. Overview of Luminex OLA assay: The assay consists of four steps; 1) Multiplexed PCR 2) multiplexed OLA 3) hybridization with beads and SAPE 4) Detection using the Luminex flow cytometer. Steps 2-4 take place in the same 96-well plate.

As shown in **Figure 2.1**, for an individual SNP the OLA involves a pair of allele specific primers and a 5' phosphorylated common primer, with tag and capture sequences utilized in downstream hybridization and labeling reactions. If an allele is present, Taq DNA ligase facilitates ligation between the allele specific and common oligonucleotides, using PCR product as a template. The OLA product is then hybridized to FlexMAP™ beads and the universal oligonucleotide, followed by labeling with streptavidin-phycoerythrin (SA-PE). The Luminex flow cytometer is then used to sort the FlexMAP™ beads and quantitate the signal from each unique bead,

thereby allowing determination of the absence or presence of each allele. Of note, different signal to noise ratios are achieved with different SA-PE conjugates. Five SA-PE conjugates were tested, one from Invitrogen/Molecular Probes and four from ProZyme (PJ31S, PJ70S, PJ35S, PJ37S) with the best signal to noise ratio achieved with the Invitrogen/Molecular Probes SA-PE (data not shown).

A control experiment was performed to demonstrate that counting as few as 20 beads allows one to distinguish a positive and negative population of beads. In order to mimic an OLA reaction, we used a synthesized OLA product as a positive control and unligated OLA probes as a negative control (see Supplementary Material 1 – Primer and Probe Sets). Sixteen replicate wells were assessed for each of four conditions: low beads/positive control, low beads/negative control, high beads/positive control, and high beads/negative control. Each of the four conditions was performed in a separate quadrant of the same 96-well plate, and positive and negative wells were alternated within each quadrant. **Table 2.1** shows that inter-well CVs and range of MFIs are very similar when counting 20 beads versus 300, and that positive and negative MFIs differ by approximately an order of magnitude, providing a signal to noise ratio more than sufficient for accurate genotype calling. Importantly, inter-well CV and MFI were based on the median values obtained from each number of beads counted (20 or 300), and not the mean. Inter-well and intra-well CVs calculated using the mean were higher when counting low versus high number of beads (data not shown), likely due to a small percentage of bead carryover that occurs with the Luminex flow cytometer.

Table 2.1. Inter-well CV and MFI range using low and high number of beads in control experiment

200 input/20 count	**Positive control**	**Negative control**
Avg. inter-well CV	2%	26%
Avg. MFI	21,968	192
MFI range	21,003-22,842	62-299
5000 input/300 count	**Positive control**	**Negative control**
Avg. inter-well CV	2%	20%
Avg. MFI	22,619	162
MFI range	21,496-23,022	115-205

Note: Median fluorescence intensity (MFI) was assessed using the high PMT setting of the Luminex 100 flow cytometer. Inter-well CVs were calculated manually and are based on the median.

The Luminex-based OLA assay, using 200 input beads and counting 20, was compared to two other genotyping methods to assess validity. First, eight SNPs were compared to direct sequencing in fourteen human DNA samples representing both homozygotes and heterozygotes, and concordance with direct sequencing was 100%. Second, a comparison was made between the multiplexed Luminex OLA and simplexed pyrosequencing. We compared 1471 genotypes from seven SNPs (**Table2.2**), and concordance between Luminex OLA and pyrosequencing was excellent, averaging 99.52%, with the lowest concordance at 98.80 %.

Table 2.2. Pyrosequencing concordance

SNP	concordant calls	% concordant
ID345549	231/232	99.60%
rs1123005	287/289	99.30%
rs11806859	168/170	98.80%
rs12122048	58/58	100.00%
rs4657179	290/290	100.00%
rs4657187	167/168	99.40%
rs905720	263/264	99.62%
Total	**1464/1471**	**99.52%**

To assess robustness when counting 20 beads, replicate genotyping of an 8-plex assay was performed in 84 human genomic DNA samples (see Supplementary Material 1 – Primer and Probe Sets). Two separate PCR amplifications were performed, and an OLA was performed in duplicate on each PCR product, for a total of four OLA assays on the same 8 SNPs in the same 84 individuals. 2685 of 2688 genotypes were "callable", with a concordance of 100% between the 99.9% of genotypes that were callable (**Table 2.3**).

Table 2.3. Call rates of technical replicates

Assay	call rate	% call rate
OLA 1 – PCR 1	670/672	99.7%
OLA 2 – PCR 1	672/672	100%
OLA 3 – PCR 2	671/672	99.9%
OLA 4 – PCR 2	672/672	100%
Total	**2685/2688**	**99.9%**

To further assess robustness, a 19-plex assay is presented in order to demonstrate typical results obtained from a "first-pass" genotyping effort (see Supplementary Material 1 – Primer and Probe Sets – Chr. 22 panel). 750 samples were genotyped from a large collection of pedigrees routinely used in the Brzustowicz lab for genetic studies of psychiatric disorders. Fifteen of nineteen assays were converted, with four failing due to the majority of MFI values falling below 1000. The approximately 80% conversion rate observed in this 19-plex assay is typical of the results observed across dozens of similar multiplexed panels in use in the Brzustowicz lab. **Figure 2.2** shows genotype clusters for the 15 working assays within the 19-plex assay, while **Table 2.4** gives the complete set of call rates and silhouette scores for each SNP. 11,207 of 11,250 genotypes were callable (first-pass call rate 99.6 %), and the average silhouette score is >0.90.

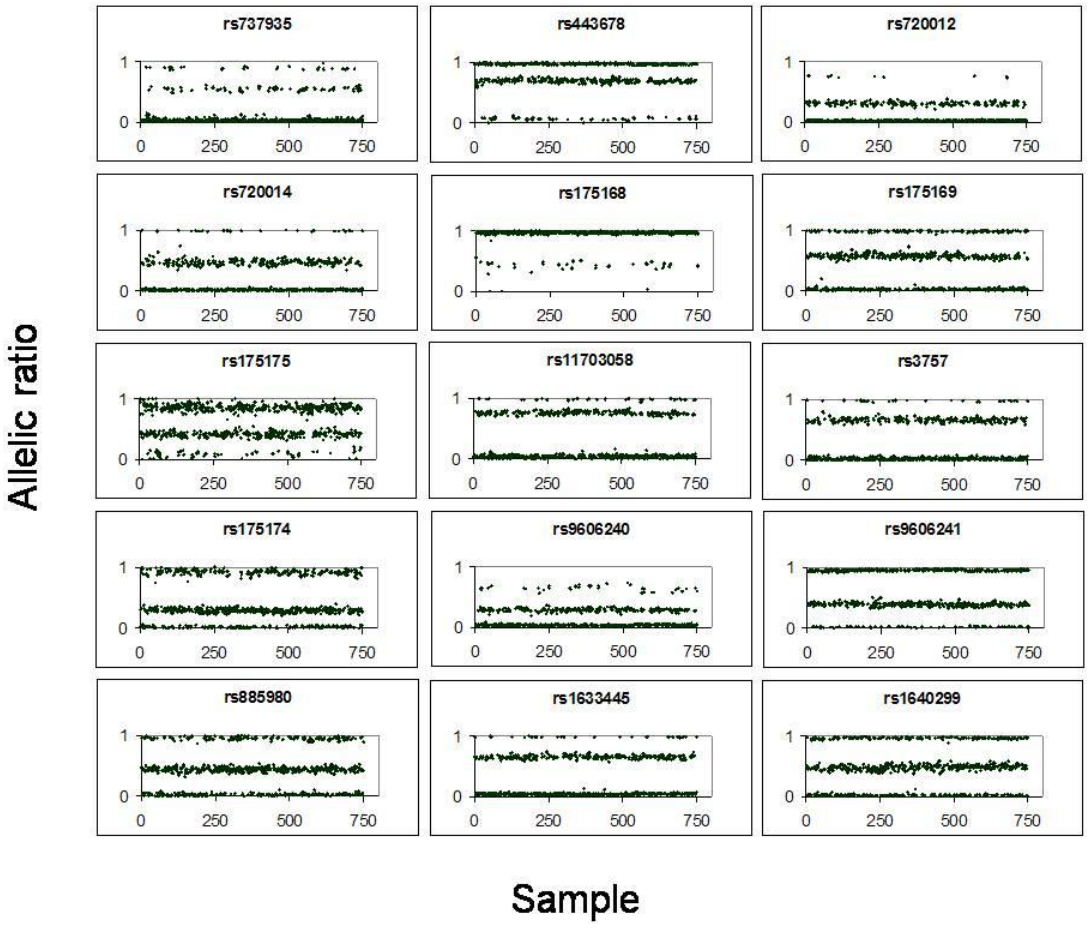

Figure 2.2. Genotype clusters for 15-plex assay Individual figures represent the allelic ratios (X axis) produced from a first pass genotyping of 750 human genomic DNA samples (Y axis). Variability in allelic ratios is observed for different SNP assays, generally with skewing of heterozyote clusters, though homozygote clusters can also be skewed (i.e. rs9606240). Nonetheless, for each SNP genotype clusters are robust (see 2.4 for silhouette scores).

Table 2.4. Call rates and silhouette scores for 15-plex assay

	Calls made	Percent callable	Silhouette score
rs737935	747/750	99.6%	0.923
rs443678	748/750	99.7%	0.905
rs720012	749/750	99.9%	0.906
rs720014	748/750	99.7%	0.94
rs175168	748/750	99.7%	0.914
rs175169	746/750	99.5%	0.925
rs175175	741/750	98.8%	0.808
rs11703058	746/750	99.5%	0.905
rs3757	746/750	99.5%	0.909
rs175174	745/750	99.3%	0.842
rs9606240	750/750	100.0%	0.816
rs9606241	748/750	99.7%	0.945
rs885980	746/750	99.5%	0.906
rs1633445	750/750	100.0%	0.933
rs1640299	749/750	99.9%	0.937
Total or average	**11207/11250**	**99.6%**	**0.901**

Discussion

We have presented a method for multiplexed SNP genotyping on the Luminex platform that is accurate, robust, and relatively inexpensive. Differences between this method and previously published methods include the use of significantly fewer beads, use of a universal biotinylated oligonucleotide for signal quantification, and no wash or centrifugation steps. We assessed concordance with both direct sequencing and pyrosequencing to assess validity, and observed 100% and >99% concordance respectively. Additional experiments with separate multiplexed assays demonstrated complete concordance across technical replicates and high call rates (>99%). The cost is approximately $0.085 per genotype when performing both the PCR and OLA reactions in a 15-plex format (**Table 2.5**).

Table 2.5. Cost per genotype using 2500 versus 200 FlexMAP™ beads

Reagent	2500 input beads	200 input beads
FlexMAP™ microspheres	0.7000	0.0560
PCR + OLA oligos	0.0200	0.0200
Taq DNA ligase	0.0030	0.0030
SA-PE	0.0020	0.0020
96-well plate	0.0030	0.0030
Hybridization buffer	0.0002	0.0002
Sheath fluid	0.0007	0.0007
Total	**$0.7289**	**$0.0849**

Note: Calculations show the cost of each reagent per genotype generated. Cost of each reagent assumes that 15 SNPs are multiplexed at both the PCR and OLA steps; reagent

costs decrease as multiplexing increases. Cost of PCR and OLA oligonucleotides assumes a sample size of 2000; oligonucleotide costs decrease as sample size increases, given that only a small portion of a typical OLA oligonucleotide order is utilized to genotype 2000 samples.

No reports directly assess the "bead count" issue when performing SNP genotyping assays on the Luminex platform. One previous report has assessed the variability in measurements when counting different numbers of beads using limiting amounts of SA-PE or analyte, and concluded that under the appropriate conditions counting as few as ten beads can distinguish two populations [108]. The appropriate conditions included a high amount of SA-PE (or analyte) relative to bead number. The conditions of our assay utilize massively saturating amounts of SA-PE (equivalent to 1.5 x 10^8 molecules per bead) and analyte, so it is not surprising that we can clearly and robustly distinguish two populations when counting as few as 20 beads (and perhaps less). Of note, run times on the Luminex flow cytometer are impacted by the number of input beads and the proportion of those input bead one attempts to count. A typical 96-well plate assay using 5000 input beads (per well) and counting 100 beads has an approximately 30 minute run time, while using 200 input beads and counting 20 takes approximately 45 minutes; the degree of multiplexing does not dramatically affect run times. Issues of machine time-out begin to appear when trying to use less than 200 beads while still counting 20, or when trying to count more than 20 beads with only 200 input beads. Also, their may be a small percentage of well-to-well bead carryover using the Luminex flow cytometer. Using the median, rather than the mean, ensures that this carryover will not significantly affect the MFI and subsequent genotype calls when counting 20 beads. It is important to point out that we have only assessed using low bead counts for a qualitative, OLA SNP genotyping assay. In quantitative assays, where

minimizing intra and inter-well variance is paramount, it may not be appropriate to use bead counts as low as 20, and we have not addressed this question in our current study.

Chapter 3

Genetic Association of SNAP-25 with Schizophrenia

Introduction

Schizophrenia is a severely debilitating psychiatric disorder with 0.5-1% prevalence in the general population worldwide. It is a disorder characterized by disturbed thought, perception, emotion, and behavior. The precise pathophysiological causes of schizophrenia have not yet been elucidated, but family, twin, and adoption studies support a disease etiology based on predisposing genetic factors. The two major neurobiological models of schizophrenia focus on defects in neurodevelopment and neurotransmission. Neuroanatomical studies have identified structural and functional brain abnormalities in schizophrenic individuals [109], including enlarged ventricles and reduced tissue volume in adjacent areas [110], as well as dendritic spine defects. The other major theory of disease pathology posits that deregulation of neurotransmission is a principal causal factor. Synaptosomal-associated protein of 25 kDa (SNAP-25) is a neuron-specific protein localized to the cytoplasmic surface of the plasma membrane, and plays a critical role in the vesicle-mediated exocytosis of neurotransmitters [111]. Although nervous system development remains normal *in utero*, genetic ablation of SNAP-25 causes complete loss of action potential-dependent neuronal signaling in SNAP-25 null mutant mice, suggesting that SNAP-25 is essential for neural signaling but not required in neurodevelopmental pathways [112]. The well-characterized molecular function of SNAP-25 as a mediator of vesicle fusion, along with the gene's cytological

location within a schizophrenia susceptibility region cited by multiple linkage studies, and also alteration of SNAP-25 protein expression in various brain regions of schizophrenic patients, cumulatively suggest that SNAP-25 is a prime candidate gene for association analysis.

Several genome-wide linkage scans for schizophrenia susceptibility loci have been performed to date, and results have provided significant or suggestive evidence for linkage in many chromosomal regions [15, 45, 113-127]. While certain regions have been investigated in multiple studies, positive linkage findings have not been replicated in a majority of studies for any chromosomal region [15]. This lack of consistency is likely due to genetic heterogeneity and the innate complexity of the schizophrenia disorder, which may require a combination of mutations at various loci to result in phenotypic manifestation. Given these confounding variables, it is feasible that studies utilizing distinct populations could discover different regions of linkage. Despite these obstacles, rank-based genome scan meta-analysis was applied to data from 20 schizophrenia genome-wide scans [15], revealing a degree of consistency among various studies not previously recognized. The cytological location of the SNAP-25 gene on 20p12-p11.2 falls within one of the regions implicated in the meta-analysis. The chromosomal location of SNAP-25 suggests that fine mapping of a schizophrenia disease locus in or around the SNAP-25 gene is a worthwhile undertaking.

The well-characterized molecular function of SNAP-25 further justifies its selection as a candidate gene for association analysis. The arrival of an action potential at the axon terminal triggers calcium ion influx through voltage-gated ion channels, which then stimulates the subsequent exocytotic discharge of neurotransmitters into the synapse. Locally synthesized neurotransmitters are often contained within synaptic vesicles, and

vesicle trafficking within the axon terminal is a complex, multi-step process with numerous regulatory proteins at each step [128, 129]. A group of proteins called N-ethylmaleimide-sensitive factor attachment protein receptors (SNAREs) are especially important in regulating vesicle docking at the cell plasma membrane as well as lipid membrane fusion [130]. While calcium-dependent proteins including synaptotagmin and the complexins may modulate the speed of exocytosis in the vesicle priming step [131, 132], the actual fusion of the lipid membranes involves a core complex composed of three distinct SNARE proteins. These include two essential components on the target cell membrane (t-SNAREs), and one critical protein on the vesicle membrane (v-SNARE) [133]. The v-SNARE synaptobrevin (i.e. VAMP, vesicle-associated membrane protein) interacts with the t-SNAREs syntaxin and SNAP-25 to form a structural motif composed of four interwoven helices, and the "zippering" of the helices helps to facilitate inter-membrane contact and unification [134]. To reiterate, SNAP-25 is an essential member of a core complex of SNAREs that are required in the vesicle trafficking steps of docking and fusion, and therefore SNAP-25 appears to be essential for action potential-dependent neuronal signaling.

One genetic study failed to find association of a tandem repeat polymorphism in the 5' upstream region of SNAP-25 with schizophrenia [135], but there have been no reported attempts to select and analyze several polymorphic markers throughout the SNAP-25 gene. We attempted to demonstrate genetic association in a family-based study by genotyping 13 tagSNP markers in two medium-sized sample groups. One sample consisted of 24 extended families from a Canadian Celtic population (n = 392), and the other consisted of 39 North American nuclear families of European descent (n = 214).

Results were statistically analyzed using the program PSEUDOMARKER, and evidence of linkage disequilibrium with a putative schizophrenia disease locus was discovered for several of the chosen SNPs.

Materials and Methods

Samples

One sample group consisted of 392 Canadian subjects of mostly Celtic background from 25 extended pedigrees. These pedigrees had been investigated previously by the Brzustowicz lab, and a genome-wide linkage scan for schizophrenia susceptibility loci was performed with this sample group [45]. Additionally, multiple sample sets were provided through the NIMH Human Genetics Initiative, and a sample of 214 subjects and controls from 39 North American Caucasian pedigrees was used in this study. Details of clinical diagnosis methods have been described in previous studies [136, 137]. In addition to affection status, the diagnostic classifications "narrow" and "broad" were used. Individuals were considered affected under the narrow diagnostic classification if they were diagnosed with schizophrenia or chronic schizoaffective disorder. Individuals were considered affected under the broad diagnostic classification if they had been diagnosed with one of those disorders or with a nonaffective psychotic disorder, schizotypal personality disorder, or paranoid personality disorder [45].

Selecting tagSNPs

A recent report presented a SNP tagging method that utilizes linkage disequilibrium units (LDU), and showed that it outperforms block-based methods while allowing researchers to use significantly fewer tagging SNPs [138]. SNPbrower^{TM} (Celera), a freely available

software that utilizes the method of SNP tagging presented in that report, was used to select minimum sets of tagging SNPs based on linkage disequilibrium maps. SNPbrower™ is a flexible program that contains linkage disequilibrium maps derived from four ethnic populations, allowing optimal SNP tagging that can be tailored to the population under study. For the present study, the SNP Wizard feature of SNPbrower™ software 3.0 was set to select tagSNPs at a density of 1 SNP/0.5 LDU for a Caucasian population, resulting in selection of 13 tagSNPs for genotyping. The positions of these 13 SNPs within the SNAP-25 gene are shown in Figure 3.1.

Multiplex PCR

Fragments containing all 13 of the selected SNPs were amplified from genomic DNA simultaneously using standard multiplex PCR protocols. The 30 μL reaction volume had the following components: 12.9 μL sterile water, 3 μL 10x PCR buffer, 3 μL $MgCl_2$ (25 mM), 4 μL dNTPs (1 mM), 5 μL primer mix (each 1 mM), 0.1 μL AmpliTaq Gold taq polymerase (1U), 2 μL genomic DNA (20 ng/μL). Reactions were initially heated at 94 ˚C for 10 minutes, followed by 40 cycles of 94 ˚C for 40 seconds, 60 ˚C for 30 seconds, ramping from 60 ˚C to 65 ˚C using fifty 0.1 ˚C two second cycles, and 72 ˚C for 2 minutes. Reactions were completed at 72 ˚C for 10 minutes and 20 ˚C for 5 minutes. All reactions were performed in 96-well plates. PCR products were verified using ethidium bromide stained polyacrylamide gel electrophoresis.

Genotyping

A novel genotyping strategy was utilized for SNP-typing of all subjects and controls at the 13 selected markers. The first procedural step is the oligo ligation assay (OLA). For

each SNP, there are three specifically designed oligonucleotides added to the reaction mix. There are two allele-specific oligonucleotides (one for each of the polymorphic variants) and each allele specific oligo has a unique 5' extension that does not base pair to the PCR product. The third oligo is a common oligo, which is the same for both of the polymorphisms and contains a universal 3' extension. The OLA reaction is carried out in a PTC-200 thermocycler with alternating denaturing and annealing steps, and Taq DNA ligase present in the reaction mixture joins the oligos together by creating a new phosphodiester linkage. The next step is the hybridization step, where differently colored microsphere beads are added to hybridize to the unique 5' extension tags on the two different allele specific oligos. A biotinylated universal oligo is also added to hybridize to the universal 3' extension. After this hybridization step, a fluorescent steptavidin-phycoerethryn conjugate (SAPE for short) is added and the reaction mixture is incubated at 37ºC for an hour. Finally, two-color flow cytometric analysis is performed on the Luminex 100 platform, where the fluorescent tag and the allele-specific color beads are detected to call the genotypes. The 15 μL reaction volume for the ligation detection reaction had the following components: 11.3 μL sterile water, 1.5 μL 10x Taq DNA ligase buffer, 0.15 μL OLA oligo mix (each 1 mM), 0.05 μL Taq DNA ligase, 2 μL PCR product. Reactions were initially heated for 1 minute at 95 ˚C, followed by 32 thermal cycles of 95 ˚C for 15 seconds (denaturation) and 58 ˚C for 2 minutes (annealing/ligation). Reactions were then cooled to 4 ˚C, and were used immediately in the hybridization step or stored at -20 ˚C for up to 1 week before proceeding with the hybridization step. All reactions were performed in 96-well plates. 50 μl of TMAC hybridization solution (3 M tetramethylammonium chloride, 50 mM Tris-HCl, pH 8.0, 3 mM EDTA, pH 8.0, 0.10% SDS) containing 200 beads (unless otherwise

noted) from each FlexMAP™ microsphere set was added directly to a well containing a completed 15 μl OLA reaction. Hybridization reactions were then denatured at 95 ˚C for 90 seconds and hybridized at 37 ˚C for 20 minutes. Fluorescent labeling was performed by adding the following mixture to each reaction well: 0.12 μL streptavidin-R-phycoerythrin (SAPE, Molecular Probes), 5.88 μL hybridization buffer. The total mixture was once again incubated at 37ºC for a minimum of 40 minutes. Two color flow cytometric analysis was then performed using the Luminex 100 flow cytometer.

Statistical analysis

PedCheck was used to detect Mendelian errors. Mendelian errors (~0.15% of all genotypes) were re-genotyped, and all samples with Mendelian errors gave consistent genotyping results and were therefore removed from further analysis. Tests of Hardy-Weinberg equilibrium and tests for genetic association were performed using PSEUDOMARKER v0.9.7 [139]. Pseudomarker is likelihood-based test of linkage and/or linkage disequilibrium which tests under both dominant and recessive genetic models, as well as under a model-based condition. Under all three models, five p-values are calculated under five different assumptions: LD | linkage, LD | no linkage, linkage | LD, linkage | no LD, and a joint test of LD and linkage.

Results

Two distinct sample groups totaling 606 subjects and controls were genotyped at thirteen polymorphic markers, all of which were single nucleotide polymorphisms (SNPs). The genotype distributions at each of the selected SNPs did not deviate from Hardy-Weinberg

equilibrium (data not shown). Genotyping results were analyzed using the statistical program PSEUDOMARKER. Nominally significant p-values indicating evidence of linkage disequilibrium between the marker and a putative schizophrenia disease locus under the narrow phenotype definition are displayed in **Table 3.1** below. Three out of the thirteen analyzed markers showed significant evidence of LD in the NIMH sample group, under a recessive genetic model.

Table 3.1. Nominally significant p-values supporting evidence of LD, narrow phenotype

Marker	Canadian	NIMH	Combined
rs363039	p = **0.025516** (LD/Linkage, -d) p = **0.0385827** (LD/Linkage, -m)	NS	p = **0.0109904** (LD/Linkage, -d) p = **0.0268784** (LD/Linkage, -r) p = **0.014655** (LD/Linkage, -m)
rs8119844	p = **0.0439283** (LD/Linkage, -r)	NS	NS
rs362602	NS	p = **0.035959** (LD/Linkage, -r) p = **0.0429014** (LD/No Linkage, -r)	NS
rs6039807	NS	p = **0.0209833** (LD/Linkage, -r) p = **0.02089634** (LD/No Linkage, -r)	NS
rs363022	NS	p = **0.0232683** (LD/Linkage, -r) p = **0.0232683** (LD/No Linkage, -r)	NS

Table 3.2, below, displays nominally significant p-values indicating evidence of LD under the broad phenotype definition. Five of the thirteen SNPs showed significant evidence of LD in the NIMH sample group, under a recessive genetic model.

Table 3.2. Nominally significant p-values supporting evidence of LD, broad phenotype

Marker	Canadian	NIMH	Combined
rs4813023	NS	p = **0.0477315** (LD/No Linkage , -m)	NS
rs362993	NS	p = **0.00588443** (LD/Linkage, -r) p = **0.00893194** (LD/No Linkage , -r)	NS
rs362548	NS	p = **0.00747155** (LD/Linkage, -r) p = **0.00755522**(LD/No Linkage , -r)	NS
rs8636	NS	p = **0.0439283**(LD/Linkage, -r) p = **0.0477315** (LD/No Linkage , -r)	NS
rs6039807	NS	p = **0.0411645** (LD/No Linkage , -r)	NS
rs363022	NS	p = **0.0203908** (LD/Linkage, -r) p = **0.0190382** (LD/No Linkage , -r)	NS

The positions of selected SNPs within the SNAP-25 gene are displayed in **Figure 3.1**, and the markers showing nominally significant evidence of LD are indicated.

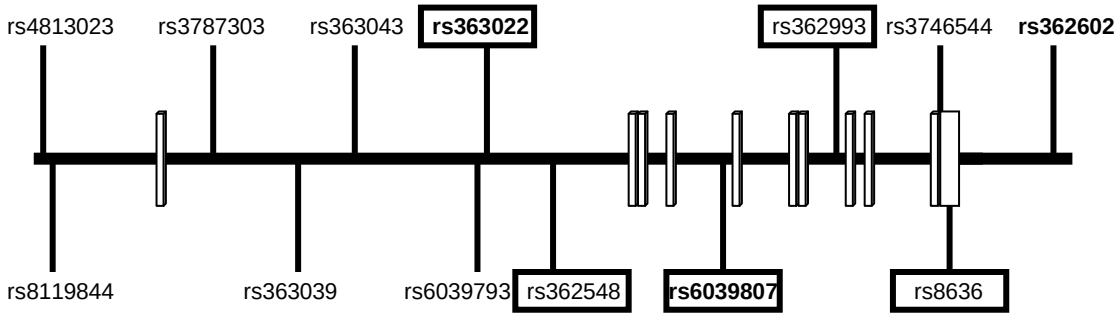

Figure 3.1. Selected SNPs interspersed throughout the SNAP-25 gene, SNPs showing positive LD: This figure displays the positions of 11 exons within the SNAP-25 gene; the first and last exons also mark the boundaries of the gene. The positions of 13 selected single nucleotide polymorphisms are displayed along with dbSNP identification numbers. SNPs in bold type indicate nominally significant findings of linkage disequilibrium under the narrow phenotype definition, boxed SNPs show nominally significant evidence of LD under the broad phenotype definition.

Discussion

In the present study, we report suggestive evidence of association between genetic variation of SNAP-25 and schizophrenia in a Caucasian, North American sample group. In testing for evidence of linkage disequilibrium between selected markers and a putative causal mutation for schizophrenia, we sought a method of analysis suitable for both the extended pedigree structure of one sample group and the predominantly nuclear pedigree structure of the other sample group. We elected to use PSEUDOMARKER v0.9.7 to conduct analyses of LD [139], which has demonstrated power in detecting LD in both

large and small family pedigrees [36]. PSEUDOMARKER, unlike many popular TDT-based programs, explicitly separates the evidence for linkage and LD in families with multiple affected individuals. For each marker, PSEUDOMARKER provides separate evaluations for evidence of linkage, of LD given linkage, of LD given no linkage, of linkage given LD, and for a joint test of linkage and LD. Unlike other programs that only examine allelic transmission within pedigrees to detect association, PSEUDOMARKER examines genotypes under various models, including a dominant model, a recessive model, and a model-based condition.

The only association study present in the literature reports a failure to find positive association of a single tandem repeat polymorphism in the 5' upstream region of SNAP-25 with schizophrenia [135]. Several of the thirteen examined polymorphic markers in our study displayed nominally significant evidence of linkage disequilibrium with a schizophrenia susceptibility locus. It would be of interest to correlate genetic variation with alteration of SNAP-25 function or protein expression, though none of the selected markers are believed to be functional mutations. Little is known about intronic cis-regulatory sequences within the SNAP-25 gene, and there has been no effort to screen for such mutations considering a lack of prior association findings. Figure 3.1 displays the positions of SNPs showing nominally significant evidence of LD, and all of these markers cluster within the 3' half of the gene.

Altered expression levels of SNAP-25 protein have been discovered in various brain regions of schizophrenic subjects relative to controls. Decreased SNAP-25 protein expression in schizophrenic subjects was reported in the inferior temporal cortex and areas of the prefrontal association cortex, while elevated protein levels were reported in another region of the prefrontal cortex [140]. Decreased SNAP-25 immunoreactivity in

the hippocampus [141, 142], as well as elevated SNAP-25 protein in the cerebrospinal fluid of schizophrenic individuals [143] have also been observed. Overexpression of dystrobrevin-binding protein 1 (dysbindin), the product of another strongly implicated schizophrenia candidate gene, was found to up-regulate expression of SNAP-25 and synapsin I in cortical neuronal cell culture [144]. This same study reported increased levels of extracellular glutamate, suggesting that dysbindin and SNAP-25 may participate in a biochemical pathway involved in glutamate neurotransmission. Despite evidence of SNAP-25 protein expression abnormalities in the brains of schizophrenic individuals, there are no known mutations mapped within the SNAP-25 gene to account for these disparities.

The SNAP-25 protein is known to exist in two isoforms resulting from alternative splicing of divergent variations of exon 5 [145]. Nine amino acid residues differ between the SNAP-25a and SNAP-25b isoforms, which results in altered relative positioning of two sites of posttranslational fatty acylation. This disparity may result in differential membrane binding characteristics of the SNAP-25 protein, which would therefore alter its functional contribution to the core complex of SNARE proteins that are essential for membrane fusion and exocytotic discharge of neurotransmitters. The marker rs6039807 is roughly 1000 nucleotides upstream of the splice acceptor site for exon 5, and this marker would be expected to be in tightest LD with a splice site mutation within the exon. This marker did show nominally significant evidence of LD under the narrow ($p = 0.0209833$) and broad ($p = 0.0411645$) phenotype definitions in the NIMH sample group under a recessive model. However, the most significant p-value ($p = 0.00588443$, NIMH sample group, recessive model, broad phenotype) was observed for the marker rs362993, which is the closest marker downstream of exon 5 and the possible site of alternative

splicing. Of note, all p-values presented here are nominal p-values. Methods which correct for multiple testing run the risk of being overly conservative when correcting for the testing of multiple SNPs in candidate gene studies. Additionally, there is no consensus among researchers as to which method is most appropriate, and it has been suggested that presenting nominally significant p-values would aid in the standardization of results presented in the literature. However, under a strict Bonferroni correction none of the individual SNP p-values presented here would be statistically significant.

This research demonstrated support for genetic association between SNAP-25 and schizophrenia. Evidence of LD with several SNPs suggests that a causal mutation is present in the 3' half of the SNAP-25 gene and this mutation are segregating in pedigrees according to a recessive genetic model. The presence of several nominally significant markers near exon 5 suggest that alternative splicing of this exon and the resultant SNAP-25 isoforms may be involved in schizophrenia etiology, possibly due to the altered capacity of SNAP-25 to regulate synaptic vesicle fusion.

Chapter 4

Genetic Association of the MicroRNA Processing Gene DGCR8 with Schizophrenia and Schizoaffective Disorder

Introduction

Schizophrenia is a debilitating psychiatric disease affecting 0.5-1% of the world's population [2]. Current studies estimate monozygotic and dizygotic twin concordance at 41%-65% and 0%-28% respectively, and give heritability estimates of 80-85% [11]. In addition to a strong genetic component, environmental, epigenetic, and stochastic processes have been implicated in or hypothesized to contribute to disease etiology. Genetic linkage and association studies have been inconsistent, though meta-analyses have identified linkage regions very likely to contain susceptibility variants [15, 16], and promising candidate genes have emerged [146, 147]. Chromosome 22q11 is one such region that has shown significant or suggestive linkage in independent linkage studies and also contains putative candidate genes [15, 16, 113, 148-150].

22q11 deletion syndrome (22q11DS), known alternatively as velo-cardiofacial or DiGeorge syndrome, involves a common (1 in 4000 live births) hemizygous deletion that results in a range of clinical phenotypes. The deletion is typically 3.1 Mb and encompasses ~40-50 genes, though in a minority of cases (~15%) deletions of smaller size are observed [151]. Notably, 22q11DS leads to an increased risk of psychiatric disorders, including schizophrenia and bipolar illness. It has been estimated that 22q11DS increases the risk for schizophrenia ~25-fold, and that the prevalence of the deletion is 40-fold higher in patients with schizophrenia (~1%) than in the general

population (~0.025%). Thus, 22q11DS is a strong genetic risk factor for schizophrenia. While the 22q11 deletion is observed in only a small percentage of patients with schizophrenia, common variants of genes within this region may contribute to risk in non-deleted patients with schizophrenia. Genetic polymorphisms from a number of genes within the deletion region - including COMT, TBX1, PICK4CA, PRODH, ZDHHC8, CLDN5, DGCR14, and DGCR2 - have been associated with schizophrenia in candidate gene studies (reviewed in [152]). A fine mapping study of common variants in a so-called 1.5 Mb "critical region" identified two subregions that were positively associated with schizophrenia [153]; one of these subregions contained a disease-associated haplotype block that encompasses four closely positioned genes examined in this study (DGCR8, HTF9C, RANBP, and ZDHHC8). A later study identified a functional variant within the ZDHHC8 gene that was positively associated with schizophrenia, and a ZDHHC8 knockout mouse model resulted in compelling schizophrenia-like behavioral phenotypes [154]. However, a number of subsequent studies have failed to replicate the genetic association of the purported ZDHHC8 functional variant and schizophrenia [155-160].

DGCR8 presents as a plausible schizophrenia candidate gene based on chromosomal location and biological function. DGCR8 is a ~31.5 Kb gene located at 22q11.21 that encodes an RNA-binding protein necessary for proper maturation of miRNAs (**Figure 4.1**). miRNAs are small RNAs (~22 nucleotides) that regulate gene expression in a wide diversity of organisms by targeting mRNAs for translational repression or degradation [161]. It is believed that miRNAs begin as part of long primary transcripts (pri-miRNA) typically derived from RNA polymerase II transcription [162]. Pri-miRNAs are then subjected to cleavage through interaction with the microprocessor complex [163-165],

which consists of the RNase III enzyme Drosha [166] and the RNA-binding protein DGCR8 [167-169]. DGCR8, acting as a molecular spacer, mediates the exact nucleotide at which Drosha cleavage occurs [170]. Cell line knock-outs reveal that DGCR8 is necessary for processing the majority, if not all, miRNAs [171]. Drosha/DGCR8 cleavage of the pri-miRNA produces an approximately 65 nucleotide stem loop structure termed the pre-miRNA. After export from the cytoplasm to the nucleus, the pre-miRNA undergoes a further round of cleavage involving another RNase III enzyme, Dicer, and the RNA binding proteins TRBP and PACT. This leaves an approximately 22 nucleotide double-stranded RNA product that is unwound to a mature single-stranded microRNA that is incorporated into the miRISC complex, which ultimately targets mRNAs for translational repression or degradation.

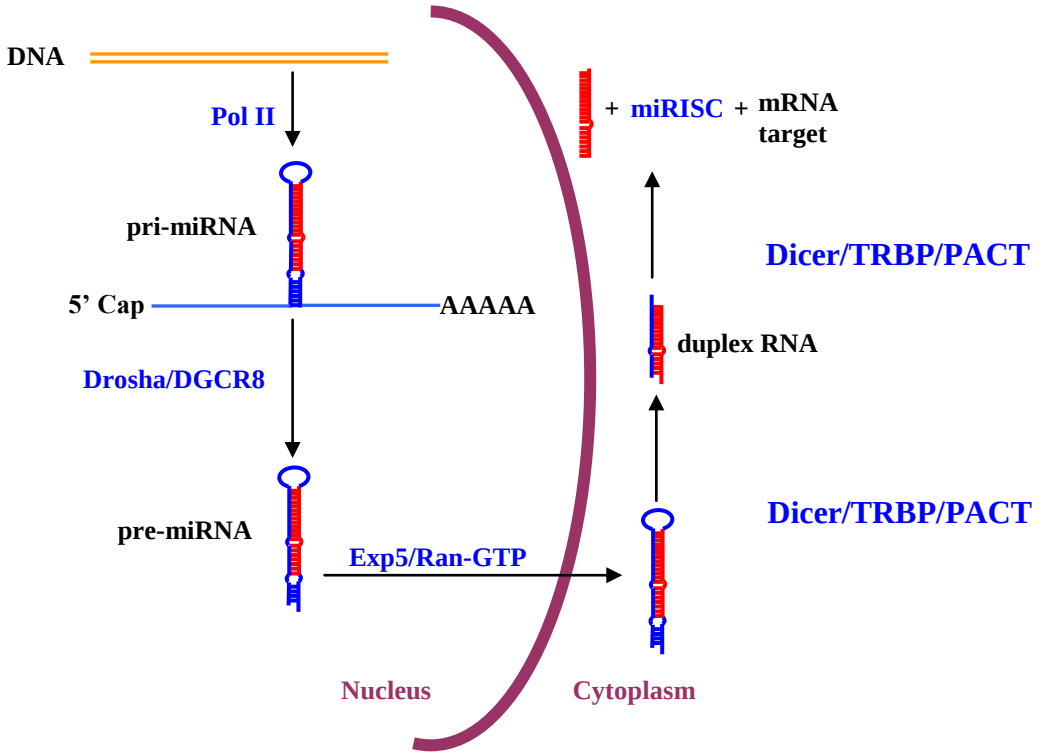

Figure 4.1. The microRNA biogenesis pathway: This figure displays the processing of miRNAs from the Pol II derived primary transcript to the mature miRNA of 19-23 base pairs. There are two main rounds of RNA processing, both involving a combination of an

RNase III enzyme (Drosha or Dicer) and associated RNA binding proteins (DGCR8/TRBP/PACT). All genes indicated are human; nomenclature differs in lower organisms.

Genetic polymorphisms that alter the function or expression of DGCR8 could conceivably alter the expression levels of mature microRNAs. Indeed, subtle alterations in miRNA expression levels can have profound effects on mammalian organ systems [172]. MicroRNAs are abundantly expressed in the developing and adult mammalian brain [173], and misexpression of microRNAs in post-mortem brain samples from patients with schizophrenia has recently been reported [174]. In this study, we hypothesized that genetic polymorphisms in DGCR8 are associated with schizophrenia and/or schizoaffective disorder. The hypothesis was tested in three ethnically similar (Caucasian) collections of pedigrees ascertained for schizophrenia or schizoaffective disorder. Given prior evidence for ZDHHC8 involvement in schizophrenia, and its close proximity to DGCR8, we genotyped a panel of SNPs encompassing a 75 Kb region at chromosome 22q11.2 that includes DGCR8 and ZDHHC8, as well as the intervening genes HTF9C and RANBP1.

Materials and methods:

Subjects

Three samples of schizophrenia families were used for this study: (1) a sample obtained from the NIMH Schizophrenia Genetic Initiative collection (2) a sample obtained from an ongoing study being carried out at McLean Hospital (3) a sample of Canadian subjects of mostly Celtic background previously studied in the Brzustowicz lab [45].

The NIMH sample of families was obtained from the National Institute of Mental Health (NIMH) Schizophrenia Genetic Initiative Collection. Schizophrenia pedigrees were ascertained at three independent sites (Columbia University, Harvard University, Washington University). Nuclear and extended pedigrees were ascertained on the basis of having at least one individual who met DSM-III-R and/or DSM-IV criteria for a diagnosis of schizophrenia or schizoaffective disorder and at least one other subject with either "affected" or "unaffected" (but not "unknown") status under at least one of the phenotype definitions described below. The Diagnostic Interview for Genetic Studies [175-177] was administered to proband subjects and to their relatives. Data on family structure and individual family members, including psychiatric diagnosis and other clinically relevant information, as well as lymphoblastoid cell lines and DNA, are stored, maintained and distributed by the NIMH Center for Genetic Studies.

Families in the McLean sample were ascertained through a proband who met DSM-IV criteria for a diagnosis of schizophrenia or schizoaffective disorder. All living members of the nuclear family were recruited, as were second-degree relatives where other branches of the family were reported to have a history of schizophrenia or schizoaffective disorder. The Structured Clinical Interview for DSM-IV (SCID-I/P) [178] was administered to all subjects by one of four interviewers. The Structured Interview for Schizotypal Symptoms (SISS) [179, 180] was used to assess signs and symptoms of schizotypal, schizoid, and paranoid personality disorder in nonpsychotic family members. Data obtained from personal interviews were supplemented by informant-based information using the Family Informant Schedule & Criteria (FISC) [181] and Kendler's criteria for schizophrenia-related personality disorders [182]. All applicable lifetime

DSM-IV Axis I diagnoses were assigned by a group of four-six senior clinicians based on a review of the SCID, an interview narrative, all available hospital records, and family informant information, using "best-estimate" consensus methods [183-185]. The interviewers and diagnosticians were blind with respect to subject group and family membership (other subjects being interviewed and assessed diagnostically during the same time period included individuals with bipolar disorder, nonpsychiatric controls, and relatives of both of these groups).

The Canadian sample group consists of subjects of mostly Celtic background from extended pedigrees. These pedigrees had been investigated previously by the Brzustowicz lab, and a genome-wide linkage scan for schizophrenia susceptibility loci was performed with this sample group [45]. Details regarding ascertainment criteria and clinical diagnostic methods have been described in previous studies [136, 137].

All data and biological material are stored and distributed without any links to personal identifying information. The individual diagnostic information as well as available non-clinical information was reviewed for each subject in the sample. To reduce the impact of genetic heterogeneity on association with the schizophrenia phenotype, we restricted our analyses to families of Caucasian ancestry. Manifestations of substance abuse or dependence in individuals with psychotic illnesses can affect the diagnostic stability of clinical phenotypes [186, 187]. Individuals with schizophrenia spectrum or affective disorders who also met criteria for co-morbid psychoactive substance abuse or dependence (cocaine, LSD, heroin, inhaled hallucinogens, multiple substances, etc) other than THC and alcohol, as well as subjects with any type of organic pathology (dementia, mental retardation) and non-schizophrenia-related psychotic

conditions were assigned a phenotype of "unknown" (under all of the definitions described below).

Phenotype definitions

Categorical phenotypes

Recent studies have shown that diagnostic stability is about 73-82% for schizophrenia and schizophreniform disorder retrospectively and 92-97% prospectively; corresponding values for bipolar disorder are 76- 85% retrospectively and 80-83% prospectively [186, 188]. The diagnosis of schizoaffective disorder appears to be the most unstable with an approximately equivalent propensity to change to schizophrenia or bipolar disorder in up to 81% of initially diagnosed subjects [187]. To take the element of diagnostic uncertainty into account, we defined three "affected" phenotypes (described below). Since the diagnostic instability implies that some subjects may change their affection status under a particular phenotype definition, the majority of subjects should still continue to meet criteria for being affected under one of three proposed phenotypic categories. Therefore, when analyzing the results of phenotype-genotype association in linkage disequilibrium studies, a conclusion can be formed about the effects of a particular genetic variant on psychotic diseases *per se* and separately for psychotic syndromes with and without an affective component. For both the McLean and NIMH samples, three categorical phenotype definitions were employed:

Phenotype 1: Narrow Schizophrenia. Subjects with a DSM-III-R or DSM-IV diagnosis of schizophrenia were classified as affected. All other subjects, including individuals with schizophrenia spectrum illnesses (paranoid or schizotypal personality

disorders, non-specified psychosis, and delusional disorder), were classified as unknown to preserve the genotyping information for error-checking and reconstruction of genotypes in subjects with unavailable DNA.

Phenotype 2: Broad Schizophrenia. Subjects with a diagnosis of either schizophrenia or any form of schizoaffective disorder (depressive, bipolar, or unspecified type) were classified as affected. All other subjects were classified as unknown.

Phenotype 3: Affective. Individuals with a diagnosis of bipolar disorder or a bipolar type of schizoaffective illness were classified as affected. All other subjects were classified as unknown.

Under all three phenotype definitions, clinically evaluated individuals were classified as unaffected if they had no history of psychosis, broadly defined schizophrenia spectrum disorders, affective pathology other than bipolar disorder, or organic illness. All phenotype assignments of unaffected were made independent of the presence of psychoactive substance abuse or dependence. **Table 4.1** presents descriptive statistics for both the NIMH and McLean samples as well as the distribution of the categorical phenotypes within those families.

Table 4.1. Descriptive statistics of NIMH and McLean samples

Sample	Families	DNA samples	Total subjects	Males	Categorical Phenotypes			
					Pheno1*	Pheno2*	Pheno3*	Unaffected
NIMH	307	457	1,692	930	473	563	76	275
McLean	141	358	568	289	64	156	82	113

*- number of affected individuals

The Canadian sample employed methods of ascertainment and exclusion criteria that differed from those of the McLean or NIMH samples. Therefore, it was not useful to apply the categorical phenotype classifications outlined above. Instead, the diagnostic

classifications "narrow" and "broad" were used. Individuals were considered affected under the narrow diagnostic classification if they were diagnosed with schizophrenia or chronic schizoaffective disorder. Individuals were considered affected under the broad diagnostic classification if they had been diagnosed with one of those disorders or with a nonaffective psychotic disorder, schizotypal personality disorder, or paranoid personality disorder [45]. **Table 4.2** presents descriptive statistics for the Canadian samples as well as the distribution of the categorical phenotypes within those families.

Table 4.2. Descriptive statistics of Canadian sample

Sample	Families	DNA samples	Total subjects	Males	Narrow*	Broad*
Canadian	25	332	455	216	84	133

*- number of affected individuals

Combined Categorical-Quantitative Frontonasal Maxillary Juncture Phenotype

Corresponding to the categorical phenotypes described above, three categorical-quantitative phenotypes were defined within the McLean sample based on the affection status of each individual under the categorical definitions and the individual's juncture score (described below).

Both the face and brain derive from common embryonic primordia and are shaped by shared forces. Therefore, a genetic or environmental insult that disrupts early development could manifest itself as both brain pathology and craniofacial dysmorphology. We have adopted objective quantitative measures and an embryological approach that relates this dysmorphology to brain pathology. Specifically, we have identified a craniofacial phenotype associated with schizophrenia, as described below.

Our studies conform to the embryologic convention of demarcating facial prominences, which are outgrowths from neural crest ectomesenchyme that arise as recognizable, discrete masses. Development of these prominences progresses from the embryonic to the mature face, and anomalies can be classified by the specific primordia from which they derive. Previous studies of schizophrenia have found prominent asymmetries at the interface of the frontonasal and maxillary anlagen; this pattern of anomalies along this border corresponds to anterior-posterior brain midline dysmorphology as predicted by embryologic fate maps [189].

The frontonasal-maxillary juncture variable is an embryologically-derived combination of facial measurements. These are exemplified at its upper and lower boundaries by the orbital and palatal regions, respectively. The juncture score is derived from a combination of individual anomalies along this embryologic interface. To quantify these measures and render them objective, we have employed the procedures of Deutsch and Farkas [190] for operational definitions of individual anomalies. Measurements were performed by raters who were blind to diagnosis and family membership.

Combined categorical-quantitative phenotypes are based on the values of the juncture score and affection status (affected, unaffected, or unknown) under each of the categorical phenotypes described above. Individuals who met criteria for being classified as affected under the corresponding categorical phenotype retained the same qualitative "affected" status under the corresponding combined definitions. For the individuals with "unaffected" or "unknown" status, the juncture score was used as a quantitative phenotype if it was available. When the juncture score was not available, the subject's status remained the same as under the categorical definitions. As a result, the number of

"affected" subjects in the McLean sample remained the same for the combined phenotype as it was for the categorical definitions. Out of 113 categorically "unaffected" subjects, 14 had a categorical phenotype and 99 had both a categorical phenotype and a juncture score. In addition, for 242, 161, and 224 subjects with an "unknown" categorical phenotype, a juncture score was available for the combined categorical-quantitative phenotype 1 (Combined Juncture - Narrow Schizophrenia), 2 (Combined Juncture – Broad Schizophrenia), and 3 (Combined Juncture – Affective) respectively.

TagSNP selection:

This study was undertaken prior to release of the HapMap Phase II. Therefore, an overlapping strategy was employed utilizing linkage disequilibrium (LD) information from three databases: Perlegen genome browser, Applera's SNP browser with Celera SNP content, and HapMap Phase I. First, using the Perlegen Genome Browser (version 1), regions of LD were determined from an approximately 75 kb region spanning 2 kb upstream of the 5' end of DGCR8 to 2 kb downstream of the 3' end of ZDHHC8. In the European-American population, 5 LD bins were identified and 5 tag SNPs were chosen from each LD bin (rs9606240, rs9606241, rs737935, rs1640299, rs720014). We then used SNPBrowser software (version 2.0) to select tagSNPs (in the same 75 kb region) based on SNP density, in units of linkage disequilibrium (LDUs). SNP density was set to 1 SNP per 0.5 LDU for the Caucasian population. Based on these results, 10 SNPs were added to the 5 SNPs already selected (rs11703058, rs7288396, rs443678, rs720012, rs3757, rs1633445, rs885980, rs175162, rs175168, rs175172). 15 SNP panel had no tag SNPs directly within ZDHHC8. Therefore, the HapMap database was used to select 4

more SNPs in and around ZDHHC8, including a ZDHHC8 intronic SNP reported to be genetically and functionally associated with schizophrenia (rs175169, rs175174, rs175175, and rs2292570). The final tagSNP panel contained 19 SNPs, four of which failed to convert to working genotype assays (rs7288396, rs175162, rs175172, rs2292570), yielding a panel of 15 SNPs that was genotyped in all three samples. Linkage disequilibrium among the 15 SNPs was assessed, and 3 SNPs in complete LD with one or more other SNPs were removed from the analysis to avoid redundant testing. We retrospectively assessed the thoroughness of this tagSNP panel using the current version of HapMap (release 22/Phase II April 2007). Using Pairwise Tagger and setting the $r^{\wedge 2}$ cutoff to 0.8 and MAF to 0.05, the current version of HapMap identified 11 LD bins covering 41 SNPs. Our 12 SNP panel successfully tagged eight of these 11 bins, and 37 of the 41 SNPs. Three bins containing four SNPs were not tagged. HapMap contains no genotype information for one of our 12 SNPs, rs175168, which could potentially tag one of the three missed LD bins.

Genotyping

The genotyping method employed multiplexed PCR and the oligonucleotide ligation assay (OLA) in conjunction with the Luminex flow cytometer. Multiplexed PCR primer sets were designed to avoid spurious products [191]. PCR primers were ordered in 96-well plate format (IDT, Coralville, Iowa) and purified using standard desalting. The OLA requires two allele-specific and one common probe for each SNP being assayed. Allele-specific probe pairs consist of a 5' tag sequence and a 3' locus specific portion that differs only at the base-pair containing the SNP. Each allele specific probe contains a

unique 24-base FlexMAP™ tag (Luminex® Corporation, Austin, TX) at the 5' end to allow hybridization to a reverse complement anti-tag coupled to a unique FlexMAP microsphere. Common probes contain a locus specific portion at the 5' end, a universal capture sequence at the 3' end, and are 5' phosphorylated by the manufacturer (Integrated DNA Technologies, Coralville, IA). Genomic DNA was isolated from whole blood using the QIAamp Blood Kit (Qiagen, Santa Clara, CA). Amplifications were performed in a solution (30 μl) containing 1.5 U AmpliTaq Gold DNA polymerase (Applied Biosystems); 1:10 dilution of 10X Buffer II (supplied with AmpliTaq Gold); 200 μM of each dNTP; 166 nM of each primer; 2.5 mM MgCl; and 40 ng genomic DNA. Reactions were initially heated at 94 ˚C for 10 minutes, followed by 40 cycles of 94 ˚C for 40 seconds, 60 ˚C for 30 seconds, ramping from 60 ˚C to 65 ˚C using fifty 0.1 ˚C two second cycles, and 72 ˚C for 2 minutes. Reactions were completed at 72 ˚C for 10 minutes and 20 ˚C for 5 minutes. OLAs were performed in a solution (15 μl) containing 20 mM Tris/HCl buffer pH 7.6; 25 mM KOAc; 10 mM MgOAc; 1 mM NAD^+; 10 mM DTT; 0.1% Triton X-100; 10 nM (200 fmol) of each OLA probe; 2 μl of multiplexed PCR product; and 3 U of Taq DNA ligase (New England Biolabs, Beverly, MA). Reactions were initially heated for 1 minute at 95 ˚C, followed by 32 thermal cycles of 95 ˚C for 15 seconds (denaturation) and 58 ˚C for 2 minutes (annealing/ligation). 50 μl of TMAC hybridization solution (3 M tetramethylammonium chloride, 50 mM Tris-HCl, pH 8.0, 3 mM EDTA, pH 8.0, 0.10% SDS) containing 200 beads from each FlexMAP™ microsphere set was added directly to a well containing a completed 15 μl OLA reaction. Hybridization reactions were then denatured at 95 ˚C for 90 seconds and hybridized at 37 ˚C for 20 minutes. Following hybridization, 6 μl of TMAC

hybridization solution containing 180 ng streptavidin-R-phycoerythrin (Molecular Probes – Invitrogen Corporation, Carlsbad, CA) was added directly to the well containing the hybridization reaction. Labeling reactions were incubated at 37 ˚C for 40 minutes. Finally, 20 of each FlexMap bead included in the multiplexed assay were sorted and quantified using a Luminex 100 or Luminex 200 flow cytometer. All primer and probe sequences are contained within Supplementary Material 1 – Primer and Probe Sets.

Statistical analysis and error checking

In this study, genetic association analysis was performed using the posterior probability of linkage disequilibrium (PPLD), a variation of the PPL. The PPL assesses the evidence for linkage and LD and estimates LD between the SNP marker and the disease phenotype [192]. The PPL was developed as a Bayesian alternative to traditional linkage analysis [193, 194]. Despite the general robustness of parametric linkage analysis [195, 196], power can be lost due to model misspecification. The PPL differs from both lod scores and "non-parametric" methods in that it directly measures the probability of linkage, given the data, and it incorporates prior genomic information, including the prior probability of linkage. The PPLD is a version of the PPL that also allows for genetic heterogeneity within and between data sets, and thus is ideal for analyses involving different collections of pedigrees. The PPLD gives a probability of LD. The prior probability of LD is set to 2%. Values below 2% represent evidence against LD, whereas values above 2% represent evidence for LD. Simulations (in one sample) indicate that 95% of PPLD values are below 4% and 99% are below 10%. Mendelian errors were

assessed using the program PedCheck. Hardy-Weinberg equilibrium was assessed using the program Pseudomarker.

Results

Genotyping of 15 tagSNPs was performed in three samples of pedigrees segregating schizophrenia (predominantly) as well as schizoaffective and bipolar disorders. The tagSNPs covered a 75 kB region on chromosome 22q11.2 containing four identified protein coding genes: DGCR8, HTF9C, RANBP1, and ZDHHC8 (**Figure 4.2**).

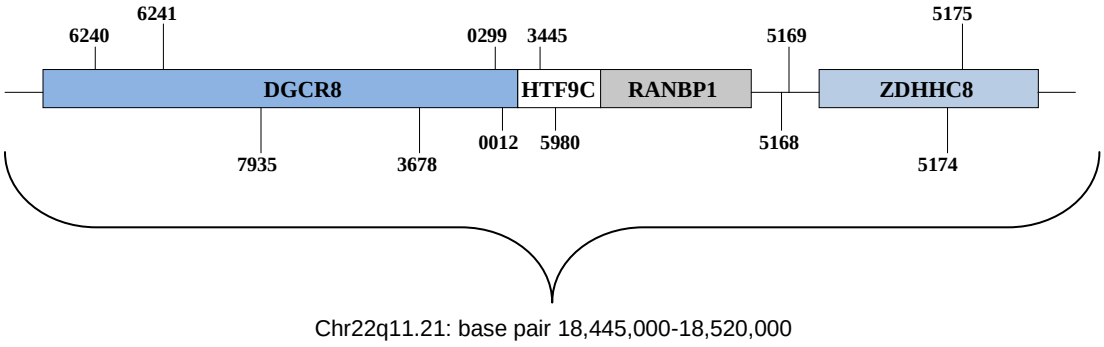

Figure 4.2. TagSNPs in chromosome 22 candidate gene region: This figure shows 12 tagSNPs covering four genes in a 75 kb region on chr22q11.21. Only the last four digits of the dbSNP rs numbers are shown (full rs numbers are given in Materials and Methods). tagSNPs cover the entire region shown here, from 2 kb upstream of DGCR8 to 2 kb downstream of ZDHHC8.

All SNPs tested were in Hardy-Weinberg equilibrium (data not shown). Patterns of LD were assessed, and the three SNPs that showed complete LD (in all three samples) with other SNPs were removed from further analysis. Thus, 12 tagSNPs without significant inter-SNP LD were assessed for genetic association. Genetic association analysis was performed using the PPLD [192]. Three categorical and three combined categorical-quantitative phenotype classifications were used for analyses (for full details see

Materials and Methods). Of note, for categorical phenotypes 1 (narrow schizophrenia) and 2 (broad schizophrenia) all patients with schizophrenia are categorized as affected, while for phenotype 3 (affective) only those with bipolar disorder and bipolar type schizoaffective disorder are considered affected. Thus, phenotype 1 corresponds to schizophrenia only, phenotype 3 to a variety of disorders with an affective component, and phenotype 2 corresponds to schizophrenia with a significant affective component. Samples sets were analyzed individually and as a whole by sequential updating.

Under the null hypothesis, the posterior probability of LD is assumed to be less than 2%; thus any value above 2% indicates significant evidence for LD. PPLD results for the three categorical phenotypes in the NIMH sample set are shown in **Figure 4.3**.

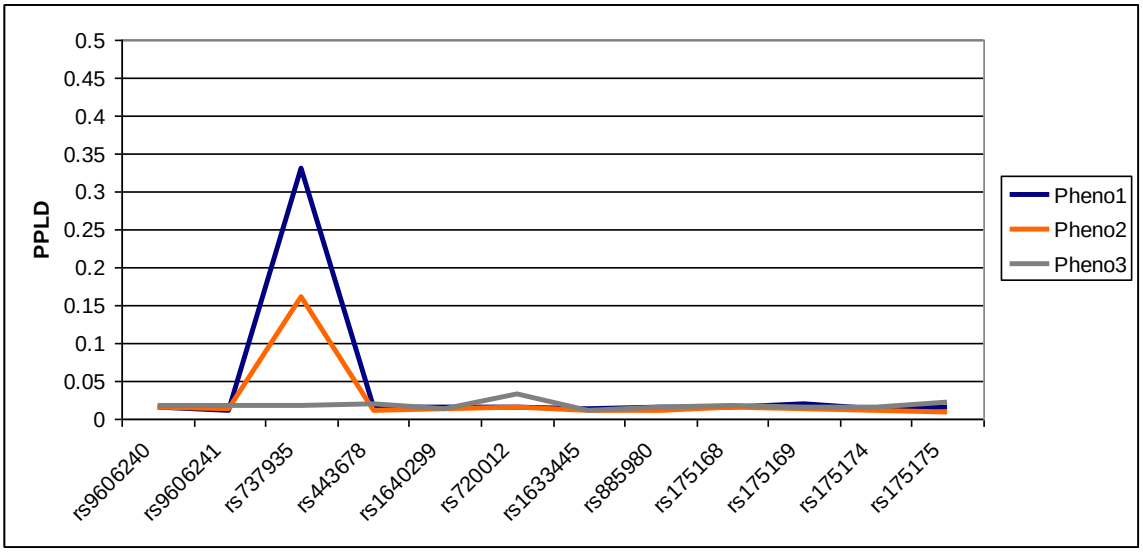

Figure 4.3. PPLD scores using categorical phenotypes in the NIMH sample: The figure shows PPLD scores for 12 SNPs analyzed under the three categorical definitions of phenotype. rs737935 is significant for phenotypes 1 and 2, and rs720012 is significant for phenotype 3.

Phenotypes 2 and 1 show significant evidence for association with rs737935, with PPLDs of 16% and 33%, respectively. Phenotype 3 is associated with rs720012, showing weak association with a PPLD of 3.4%. Thus, in this sample, rs737935 is associated with schizophrenia and schizoaffective disorder while rs720012 is weakly associated with disorders that contain a bipolar component (i.e., bipolar disorder and the bipolar subtype schizoaffective disorder).

The McLean sample allows endophenotypes to be incorporated into genetic studies of schizophrenia, potentially increasing the power of genetic studies [197]. Such bivariate phenotypes may not only reduce the high risk of false negatives but also limit the impact of genetic heterogeneity . One of these endophenotypes involves anomalies at the frontonasal-maxillary juncture, which have been associated with schizophrenia (Deutsch et al. 2000, 2007a), and were thus included in the association analyses of the combined phenotypes (as described in Materials and Methods). **Figure 4.4** shows the PPLD results analyzing the three combined categorical-quantitative phenotypes for association with DGCR8 SNPs.

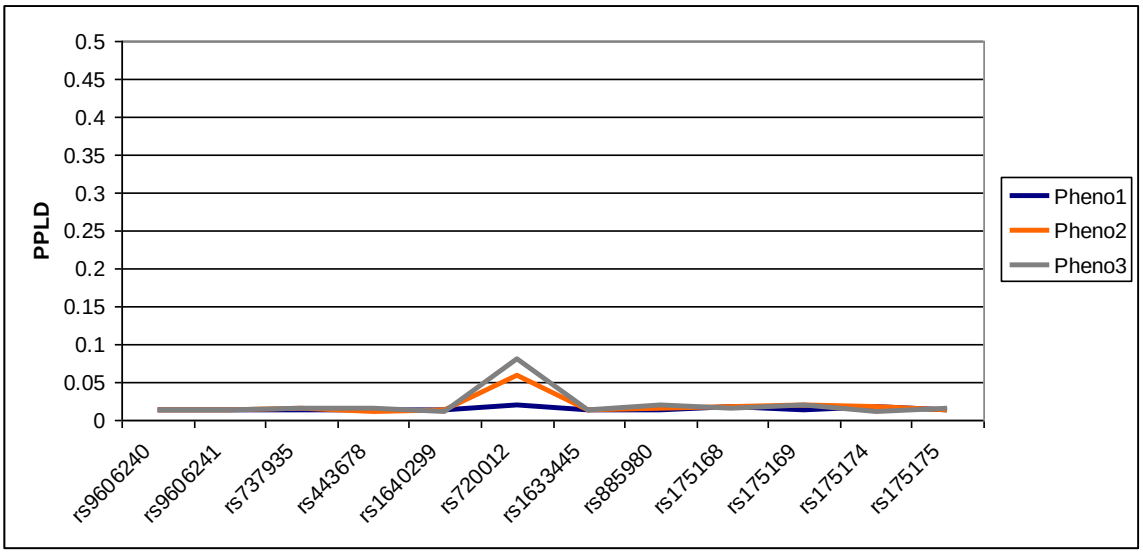

Figure 4.4. PPLD scores using combined categorical-quantitative phenotypes in the McLean sample: The figure shows PPLD scores for 12 SNPs analyzed under three definitions of the combined categorical-quantitative phenotype. rs720012 is significant under phenotypes 2 and 3.

Both phenotypes 2 and 3 show association with rs720012. Phenotype 2 has a PPLD of 6.0% and phenotype 3 has a PPLD of 8.1%. Notably, the PPLD for the combined categorical-quantitative phenotype 3 was over two times higher than the PPLD of 3.4% seen for the clinical phenotype alone in the NIMH sample. These findings provide further support for the association between rs720012 and psychotic illness that has an affective component. When applying the categorical phenotypes only to the McLean sample there was no evidence for genetic association (data not shown).

Figure 4.5 presents the results of sequential updating across both samples, utilizing the categorical phenotypes in the NIMH sample and the combined categorical-quantitative phenotypes in the McLean sample. There is evidence for association of rs737935 with phenotypes 1 and 2 (PPLDs of 15% and 24%, respectively) and rs720012 with phenotype 2 and 3 (PPLDs of 6% and 24 %, respectively.

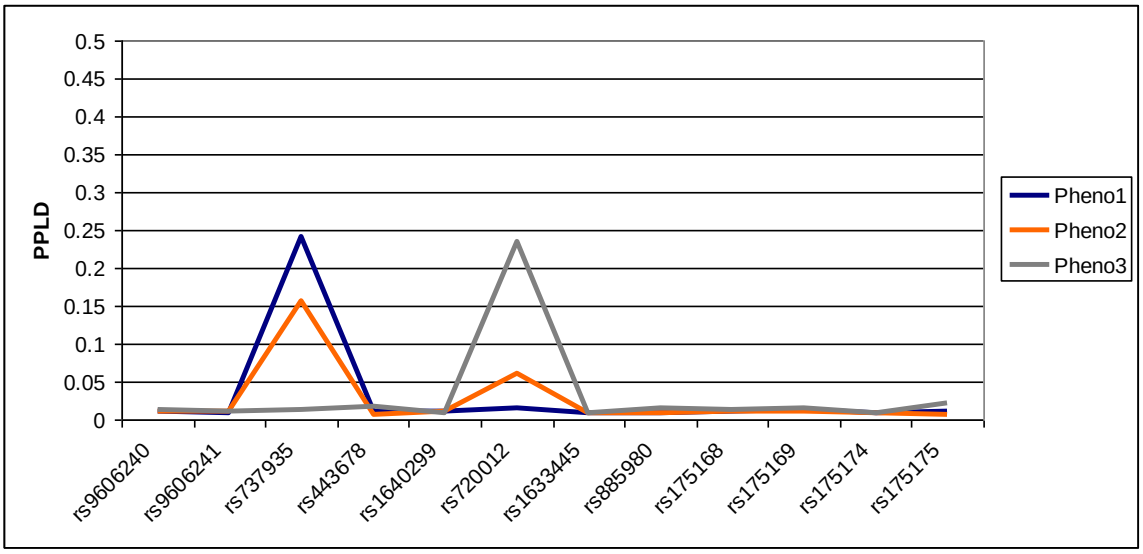

Figure 4.5. PPLD scores from sequential updating of NIMH and McLean samples using categorical and combined categorical-quantitative phenotypes, respectively: The figure shows PPLD scores for 12 SNPs analyzed via sequential updating in the NIMH and McLean samples under categorical (NIMH) and categorical-quantitative (McLean) phenotypes. rs737935 is significant under phenotypes 1 and 2 in the NIMH sample, but not in the McLean sample and rs720012 is significant under phenotypes 2 and 3 in the McLean sample.

Genotyping was performed in a third sample set, ascertained in Canada and containing subjects of Celtic and German ancestry [36]. The Canadian pedigrees showed no evidence for linkage to chromosome 22q11 in a previous study [45]. Notably, this particular sample was purposefully screened to exclude individuals with any clinical features of 22q11DS and to exclude individuals with bipolar illness. Results of PPLD analysis are shown in **Figure 4.6**. There is no evidence for association with any of the chromosome 22 SNPs in this sample

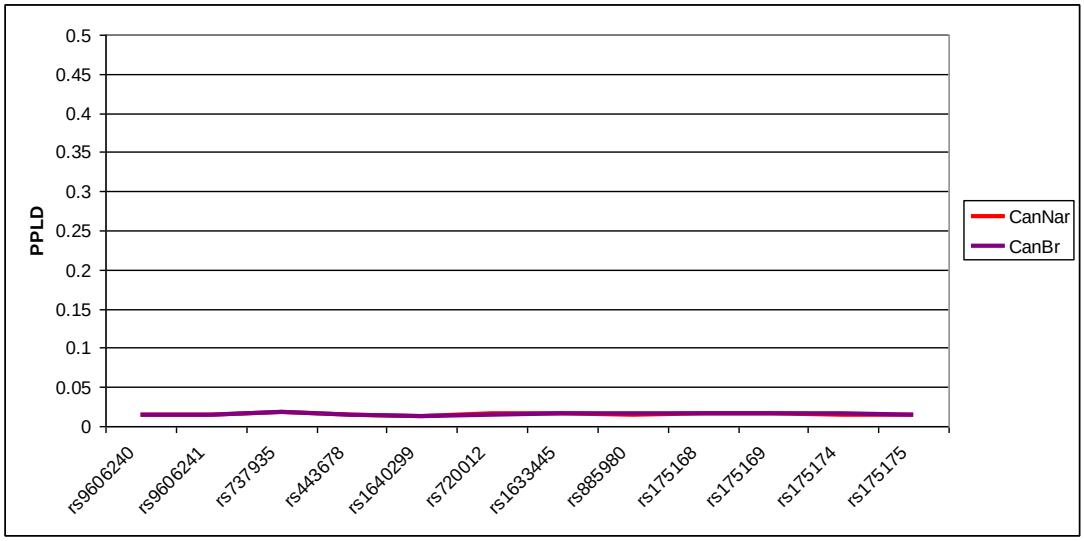

Figure 4.6. PPLD scores using narrow and broad phenotypes in the Canadian sample: The figure shows PPLD scores for 12 SNPs analyzed in the Canadian sample. None of the PPLDs for these SNPs was significant.

Discussion:

In this study twelve tag SNPs, spanning four genes (75 kb total) within the 22q11DS region, were genotyped and analyzed for association to schizophrenia, schizoaffective disorder, and bipolar disorder. The deletion region contains ~40-50 protein coding genes, and has been implicated in dramatically increased risk for schizophrenia and other psychiatric disorders, including bipolar illness [198]. Our findings implicate rs737935 in susceptibility to schizophrenia in one sample but do not confirm this association in the other two samples. Another SNP, rs720012, was associated with psychotic illness with an affective component in two pedigree collections. Both associated SNPs are within the DGCR8 gene.

SNP rs720012, located in the 3' UTR of DGCR8, was associated with psychotic conditions having an affective component in the NIMH sample set. This SNP was also significantly associated with the combination of psychotic conditions having an affective

component (primarily schizoaffective disorder) and a quantitative measure of craniofacial dysmorphology in the McLean sample set. The association with craniofacial dysmorphology was observed in a phenotype that included schizoaffective disorder of the bipolar type and bipolar disorders, but not with a phenotype restricted to narrowly defined schizophrenia. The specific craniofacial dysmorphology assessed, which is based on a pattern of quantifiable anomalies at the frontonasal-maxillary juncture, has previously been associated with schizophrenia and corresponds to anterior-posterior brain midline dysmorphology in patients with schizophrenia and in their clinically unaffected first-degree relatives.

This research does not answer whether rs720012 is a causative mutation or merely in linkage disequilibrium with a causative mutation. rs720012 is an A:G SNP, with a minor allele (the A allele) frequency of approximately 10% in the Caucasian population. In this study, the A allele was more likely to be found in affected individuals. Intriguingly, rs720012 is located in the 3' UTR of DGCR8, within a predicted (using RNAhybrid; http://bibiserv.techfak.uni-bielefeld.de/rnahybrid/) conserved binding site for hsa-mir-221, and thus could potentially impact the expression of DGCR8 by altering the hsa-mir-221 targeting of this gene. The A allele replaces an unfavourable G:U wobble with a more favourable A:U pairing (**Figure 4.7**). A hypothesis for further testing is that the A allele results in reduced levels of DGCR8 expression, at some point in development, through an increased binding affinity of hsa-mir-221 for the target site in the 3' UTR of DGR8.

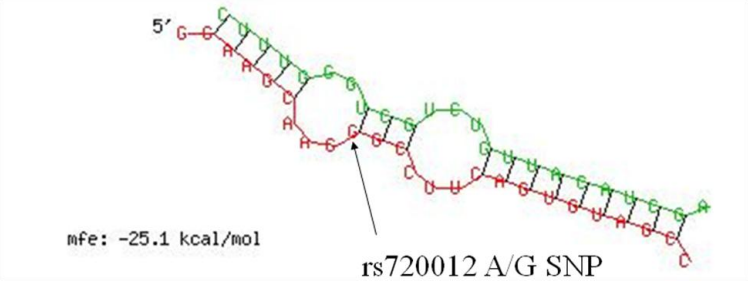

Figure 4.7. Predicted hybridization of hsa-mir-221 to 3' UTR of DGCR8: This figure shows the location of SNP rs720012 within the predicted hybridization site of hsa-mir-221 (green) with the DGCR8 3' UTR (red). The MiRanda target prediction algorithm was used to predict targeting of hsa-mir-221 to DGCR8. RNAHybrid was used to generate the image and to calculate the minimum free energy (mfe) of hybridization.

SNP rs737935, located within exon 9 of DGCR8, is associated with schizophrenia in the NIMH sample set and in combined analysis of the NIMH and McLean sample sets. A search for functional elements (using the UCSC genome browser) reveals no obvious elements that rs737935 impacts. It may be that the finding of genetic association for this SNP is due to LD between rs737935 and some as yet unknown functional variant. However, using the most recent HapMap data release (release 21a, phase II, January 2007) neither rs737935 nor rs720012 share LD above $r^{\wedge 2}$ =0.80 with any other SNP in the 75 KB region studied. While it is true that both rs737935 and rs720012 were selected as tagSNPs for this study, both were contained within singleton LD bins.

In conclusion, this study represents the first genetic evidence for association of the miRNA biogenesis pathway with psychiatric disorders. A previous report has identified misexpressed miRNAs in post-mortem samples from patients with schizophrenia and schizoaffective disorder [174]. Through a comparison of the expression of intronic miRNAs and the mRNAs in which they are located, the authors of that study hypothesized that altered biogenesis of miRNAs might be the mechanism by which miRNA misexpression is induced. Our data support that hypothesis.

Chapter 5

MicroRNA Expression Profiling in Post-Mortem Brain Samples from Patients with Schizophrenia and Bipolar Illness

Introduction

We previously have identified genetic polymorphisms in the microRNA processing gene DGCR8 that are associated with both schizophrenia and bipolar disorder. Genetic polymorphisms which alter the function or expression of DGCR8 could conceivably alter the expression levels of mature microRNAs (miRNAs), and in this study we assessed the expression levels of specific miRNAs in a collection of post-mortem brain samples from patients with schizophrenia, bipolar illness, and psychiatrically normal controls. MiRNAs are abundantly expressed in the developing and adult mammalian brain [173], and there are a number of reasons to hypothesize why miRNA expression may be altered in brain tissue of individuals with neuropsychiatric disorders (reviewed in [199]). Schizophrenia is thought to be a developmental disorder [200], and miRNAs have emerged as important regulators of animal development [72]. An *in vitro* model of mouse embryonic stem cell derived neurogenesis identified specific microRNAs important in progenitor differentiation to neurons or astrocytes [71], and *in vitro* results correlated with expression patterns of *in vivo* mouse embryonal neurogenesis [70, 73] . The dre-mir-430 family has been shown in zebrafish to partially rescue brain morphogenesis defects observed in Dicer knockouts [75]. It has been demonstrated that mir-134 negatively regulates dendritic spine width in cultured rat hippocampal neurons through targeting of the LIMK1 gene [74], and alterations in dendritic spine density and morphology have been noted in human post-mortem studies and pharmacological animal models of

schizophrenia [24, 201]. Schizophrenia is also hypothesized to involve defects in synaptic transmission, notably in glutamate signaling through the NMDA receptor [202], and a report of computationally predicted microRNA targets took note of the fact that multiple microRNAs are predicted to target the NMDA receptor complex, including three NMDA receptor subunits, all PSD-95 family members (DLG1-5), and other PSD-95 associated genes [76]. It is clear from these studies that microRNAs are abundant and powerful effector molecules in the central nervous system that could potentially play a role in schizophrenia.

In this study, we performed TaqMan based expression profiling of post-mortem brain samples obtained from patients with schizophrenia, bipolar disorder, and psychiatrically normal controls. A small pilot experiment was performed to assess the expression of 160 microRNAs in 6 post-mortem brain samples (schizophrenia=3, control=3). Potentially interesting microRNAs identified in the pilot study were then followed up in a larger collection of post-mortem brain samples (schizophrenia=35, bipolar=35, control=35). Samples in the follow-up expression study were genotyped with the same tagSNP panel that was utilized in the DGCR8 genetic study, allowing us to investigate correlations between associated DGCR8 SNPs and miRNA expression levels.

Materials and methods

Post-mortem brain samples:

Tissue samples were obtained from three sources: the Human Brain and Spinal Fluid Resource Center (VA Medical Center, Los Angeles, California), the Harvard Brain Tissue Resource Center (Harvard, Boston, Massachusetts), and the Stanley Medical

Research Institute (SMRI, Chevy Chase, Maryland). The SMRI samples consist of post-mortem brain specimens from 35 individuals with schizophrenia, 35 individuals with bipolar disorder, and 35 psychiatrically normal controls. This sample contains many more individual subjects than is typical for post-mortem studies of psychiatric disorders, and samples were collected in a standardized fashion with an emphasis on obtaining high-quality RNA for expression studies. Diagnoses were made by two senior psychiatrists, using DSM-IV criteria, based on medical records, and when necessary, telephone interviews with family members. Diagnoses of unaffected controls were based on structured interviews by a senior psychiatrist with family member(s) to rule out Axis I diagnoses. Specimens were collected, with informed consent from next-of-kin, by participating medical examiners between January 1995 and June 2002. The specimens were all collected, processed, and stored in a standardized way. Exclusion criteria for all specimens included: 1) significant structural brain pathology on post-mortem examination by a qualified neuropathologist, or by pre-mortem imaging; 2) History of significant focal neurological signs pre-mortem; 3) History of central nervous system disease that could be expected to alter gene expression in a persistent way; 4) Documented IQ < 70; or 5) Poor RNA quality. Additional exclusion criteria for unaffected controls included age less than 30 (thus, still in the period of maximum risk) and substance abuse within one year of death or evidence of significant alcohol-related changes in the liver. Relevant demographic information for the Stanley samples is contained in **Table 5.1**.

Table 5.1. Demographic variables of patients comprising Stanley post-mortem brain sample set

	Unaffected controls	Schizophrenia	Bipolar Disorder
Number	**35**	**35**	**35**
Mean age and range	**44.1 (31-59)**	**42.6 (19-59)**	**45.3 (19-64)**
Race	**35 white**	**35 white**	**33 white** **1 Black** **1 Native American**
Sex	**26M, 9F**	**26M, 9F**	**17M, 18F**
Mean PMI and range (hrs)	**29.4 (9-58)**	**31.3 (9-80)**	**37.9 (12-84)**
Mean pH and range	**6.6 (6.0-7.0)**	**6.5 (5.9-6.9)**	**6.4 (5.8-7.0)**
Side of brain	**19 right, 16 left**	**18 right, 17 left**	**16 right, 19 left**

Purification of RNA from tissue samples:

For the pilot study samples (LA and Harvard), ~300-500 mg sections of frozen tissue were ground to a powder in liquid nitrogen. Total RNA was purified from a portion of this powder using the mirVana miRNA isolation kit (Ambion, Austin, Texas) following the manufacturer's recommended protocol. The remainder of the powder was

stored at -80C for future use. Quality and quantity of total RNA was assessed using an Agilent 2100 Bioanalyzer and NanoDrop ND-1000 spectrophotometer.

For Stanley samples region BA9, 300-500 mg sections of frozen tissue were added to 3mL of chilled RNALaterICE (Ambion) and stored at -20C for a minimum of 16 hours. To prepare samples for RNA purification, RNALaterICE was decanted, and the tissue sample lightly blotted to wick away excess RNALaterICE. Next, 3mL of chilled mirVana lysis buffer was added to the tissue and the tissue homogenized with a Fisher PowerGen 35 handheld homogenizer. The mirVana miRNA isolation kit was used to purify total RNA from a 750μL aliquot of lysate. For Stanley samples region BA46 samples, we received total RNA from the SMRI. Yield and purity were determined using a NanoDrop ND-1000 spectrophotometer.

TaqMan miRNA profiling in pilot study:

For the pilot experiment, we measured the relative expression of 160 microRNAs included in the Applied Biosystems TaqMan MicroRNA Assays Human Panel (Early Access Kit version 1.0). TaqMan miRNA assays have a linear dynamic range of greater than 7-logs; this is a key advantage of the TaqMan miRNA assay over competing technologies, given that miRNA expression varies from less than 1 copy to more than 50,000 copies per cell. Primers and probes were delivered in a 96-well plate format. Simplexed RT reactions were performed using the High-Capacity cDNA Archive Kit (Applied Biosystems). For a 1x reaction, RT was performed in a solution (15uL) containing 0.15uL dNTP (100mM), 1uL MultiScribe Reverse Transcriptase (50U/uL), 1.5uL 10x RT Buffer, 0.188uL RNase Inhibitor (20U/uL), 4.126 uL nuclease-free water,

3uL stem-loop RT primer and 5uL (50 ng/uL) input RNA. RT reactions were performed in 96-well plates under the following cycling conditions: 16˚C for 30 minutes, 42˚C for 30 minutes, and 85˚C for 5 minutes. Simplexed PCR reactions were performed in quadruplicate using a 7900HT Real-Time PCR System (9600 emulation mode, 95˚C for 10 minutes followed by 40 cycles of 95˚C for 15 seconds and 60˚C for 60 seconds). No RT and no template controls were routinely performed.

TaqMan miRNA profiling in Stanley sample set:

For follow-up profiling in Stanley samples BA9 region, we saved on reagent costs by performing multiplexed RTs and reducing the PCR volume to 5uL. Both multiplexed RTs and reduced volume PCRs were assessed to ensure that these modifications did not affect the linear dynamic range of the assay. At very high concentrations of input RNA (>1.5ug) we did observe compression of the linear range for a significant portion of miRNAs tested. However, for all assays tested, input RNA amounts of 1µg or less showed excellent linearity, and for more highly expressed microRNAs we were able to achieve a linear dynamic range of 7-logs (1µg to 0.01ng of input RNA). The worst performing assays, which were often those expressed at lower levels, showed a linear dynamic range of no less than 4-logs. Multiplexed RT reactions were performed using the TaqMan MicroRNA Reverse Transcription Kit (Applied Biosystems). For a 1x reaction, RT was performed in a solution (20µL) containing 0.40µL dNTP (100mM), 4µL MultiScribe Reverse Transcriptase (50U/uL), 2uL 10x RT Buffer, 0.25uL RNase Inhibitor (20U/uL), 5.35 uL nuclease-free water, 4uL stem-loop RT primer (62.5 nM) and 4uL input RNA (18.75 ng/uL). Simplexed PCRs were performed in quadruplicate using a

7900HT Real-Time PCR System (Applied Biosystems). Reactions were performed in solution (5uL) containing 1.89uL nuclease-free water, 2.5uL TaqMan Universal PCR Master Mix (No AmpErase UNG), 0.25uL 20x probe/primer mix, and 0.36uL RT product. All PCRs were performed in quadruplicate.

Control RNAs for normalization of miRNA expression:

At the time of our initial experiments using the LA and Harvard samples, small RNA housekeeping genes were not available, but Applied Biosystems had in-house data showing that mir-16 might be an appropriate endogenous control (personal communication with Applied Biosystems). We assessed the appropriateness of mir-16 as a housekeeper by normalizing it to the total expression of all 160 microRNAs in six samples, and saw very little sample-to-sample variation in mir-16 expression levels. Thus, mir-16 was used to normalize the pilot experiment data.

For the follow up experiment in the Stanley samples, we performed geNorm analysis [203] on a panel of recently available TaqMan assays for small nucleolar RNAs. GeNorm analysis is a common method used to assess the suitability of putative housekeeping genes as normalization controls. Evaluation of putative housekeeping genes requires prior knowledge of a stably expressed housekeeper with which to normalize the putative housekeeper, thus creating a circular problem. One way to solve this problem is to utilize non-normalized expression data, by relying on the principal that the expression ratio of two ideal control genes should be identical from sample to sample. The expression of a set of putative housekeepers can be measured across a number of samples, and the average expression stability, M, can be determined. M is defined as the

average pairwise variation of a particular gene with all other control genes. Thus a lower M equals a more stable control gene. For BA9 GeNorm analysis, 150ng of total RNA from each of 16 samples consisting of a mixture of phenotypes (schizophrenia, bipolar, control) was subjected to two-step RT-PCR. The 16 samples were split into groups of 8 so that the RT and PCR were performed separately for each group of 8, resulting in a replication of the GeNorm analysis to improve confidence in our results. RT and PCR were performed as outlined above. Based on results of the GeNorm analysis (**Figure 5.1, below**) we used RNU24 and RNU44 to normalize expression data from the Stanley samples.

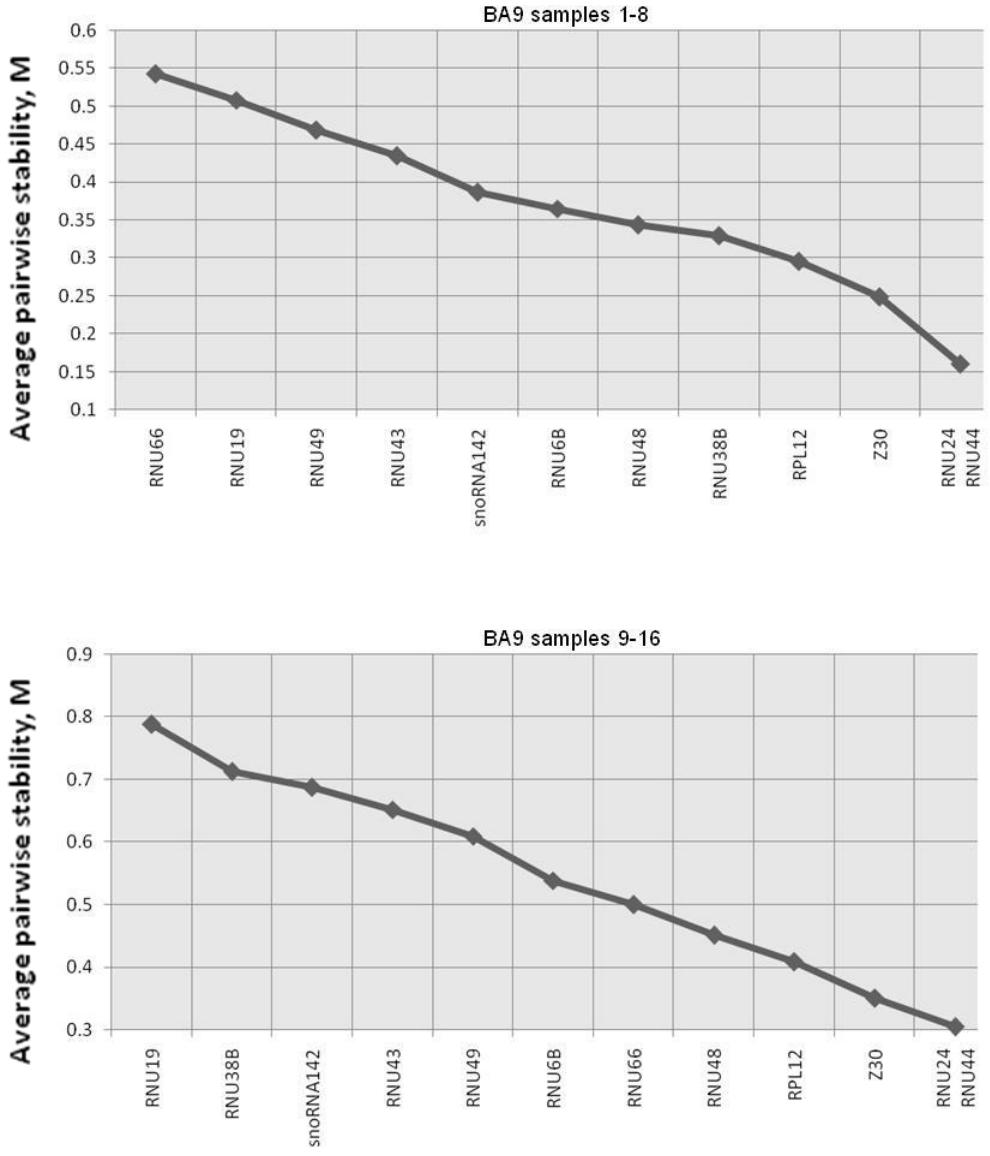

Figure 5.1. Expression stability of small nucleolar RNAs as determined by GeNorm analysis: This figure shows the output from GeNorm software, indicating the expression stability of the 12 putative housekeeping genes assessed. Each graph represents an independent experiment performed on different samples. In both experiments, the top four most stable putative housekeepers were RNU44, RNU24, Z30, and RPL12, strongly indicating that these four small nucleolar RNAs are suitable housekeepers. Based on these results, we selected RNU24 and RNU44 to use as normalization controls.

Statistical analyses of expression profiling:

For the pilot study, we used a combination of ΔΔCt analysis and the Significance Analysis of Microarray (SAM) software package from Stanford University to identify differentially expressed microRNAs [204, 205]. The ΔΔCt method utilizes normalization to housekeeping genes to assess relative expression differences between samples. It is widely accepted as an appropriate method for TaqMan assays, where PCR efficiencies are near 100%. Briefly, SAM is a statistical method and free software that performs permutation-based significance analysis for microarray data. It reports a list of significant genes and the False Discovery Rate associated with that analysis. For our experiment, we have a two-class unpaired design (post mortem brain samples from schizophrenic or control individuals). SAM calculates an observed and an expected d value for each gene. $d_{observed}$ represents the observed difference in mean expression levels between the classes being compared (i.e. schizophrenic and control) divided by the variance in the data and a fudge factor. $d_{expected}$ is calculated by permutating the class labels but not the data over all possible permutations. The difference between $d_{observed}$ and $d_{expected}$ is termed delta. The experimenter sets delta to an arbitrary significance threshold, and a false discovery rate is calculated for this particular threshold. The false discovery rate is calculated by estimating how many genes would be called significant in the permutated data set (on average) for this particular value of delta. The false discovery rate is the average number of significant genes in the permutated data set divided by the number of significant genes in the real data set. The desired false discovery rate determines the delta value you wish to set. SAM analysis offers a less conservative, and hence more powerful, solution to the problem of corrections for multiple testing when testing many genes.

For analysis of miRNAs assessed in the Stanley sample set BA9 region, we used multiple imputation to fill in the missing covariate data, giving ten complete datasets, then used Bayesian model averaging for each dataset and each miRNA. Combining the results across datasets gave estimated probabilities that a given covariate belonged in the model for a given miRNA. For each dataset, the log of the relative expression of each miRNA was regressed against covariates with greater than 10% probability of inclusion and the schizophrenia/bipolar/control diagnosis. In addition, for each dataset the log of the relative expression of each miRNA was regressed against the DGCR8 SNP genotypes. Results were combined across imputed datasets. For Stanley sample BA46 the log of relative expression were regressed against subject covariates as well as diagnostic groups, but we did not test for association of miRNA expression with DGCR8 SNP genotypes, given the very low expression values for most miRNAs assessed in BA46.

Results:

For the pilot study, TaqMan microRNA assays were used to assess expression of 160 microRNAs in post-mortem samples from patients with schizophrenia (n=3) and psychiatrically normal controls (n=3). All samples were from Brodmann area 9 in the dorsolateral prefrontal cortex. Of the 160 microRNAs tested in the pilot study, 157 showed measurable (average Ct <35) levels of expression, though many at the lower level of expression require very high input RNA (500ng) to achieve acceptable technical replicates. Extrapolating from published data on the average copy number per cell of hsa-mir-16 in adult brain tissue [206], a rough (order of magnitude) estimate was made of the copy number per cell for all 160 microRNAs. For ease of discussion, the levels of

expression have been categorized from very high to very low, and results are summarized in **Table 5.2**.

Table 5.2. Rough copy number estimate of 160 miRNAs in BA9 of the adult human brain

Copies/cell	Relative expression level	# of miRNAs
>10,000	Very high	14
1,000 to 10,000	High	42
100 to 999	Medium	54
1 to 99	Low	31
<1	Very low	16
No expression		3

A review paper from 2005 showed that 50% (125/250) of known microRNAs were expressed at some level in the brain [207], and these data add a significant number to the list of brain-expressed and brain-enriched microRNAs. The authors cite 12 microRNAs that are enriched in the brain, and these data generally agree with the published findings. Ten of the 12 microRNAs were identified as brain-enriched are in the highest expression group, 1 in the second highest, and 1 in the third highest. Below (**Table 5.3**) is data from our study showing the 10 most highly expressed miRNAs in the dorsolateral prefrontal cortex. Additionally, **Table 5.3** shows the percentage - relative to total miRNA expression of all 160 miRNAs measured - of each miRNA. It can be seen that the 10 most highly expressed miRNAs account for over 60% of the total miRNA fraction.

Table 5.3. 10 most highly expressed miRNAs in BA9

microRNA	% of total
miR-124a	10.97%
miR-29c	9.74%
miR-29b	8.67%
miR-125b	6.72%
miR-9	6.18%
miR-29a	5.74%
miR-26a	5.69%
miR-124b	3.03%
miR-30c	2.74%
let-7b	2.33%
Sum	61.82%

ΔΔCt analysis was performed to assess relative expression differences between the control and disease samples. **Table 5.4** shows the 15 most under- and over-expressed miRNAs (schizophrenia relative to control), along with their relative expression levels as indicated by our rough copy number analysis.

Table 5.4. 15 most under- and over-expressed miRNAs in pilot study of BA9 as given by ΔΔCt analysis

	Under	Level	miRNA	Over	Level
miR-144	3.52	very low	miR-371	11.37	very low
miR-198	2.09	very low	miR-372	7.34	very low
miR-193	2.04	low	miR-373	3.57	very low
miR-199a	2.04	low	miR-122a	2.20	very low
miR-190	1.96	low	miR-150	1.80	low
miR-189	1.92	very low	miR-125a	1.77	high
miR-199b	1.78	low	miR-339	1.62	low
miR-211	1.76	low	miR-10a	1.60	low
miR-105	1.73	low	miR-30d	1.55	medium
miR-182	1.66	low	miR-107	1.55	low
miR-200a	1.65	low	miR-149	1.53	medium
miR-183	1.62	very low	miR-133b	1.50	low
miR-199a*	1.61	low	miR-328	1.46	medium
miR-224	1.58	very low	miR-29b	1.46	high
miR-338	1.58	medium	miR-130b	1.44	low

Statistical Analysis of Microarrays (SAM) was applied to assess the significance of expression differences between diseased and control samples. **Table 5.5** shows the results of the SAM analysis, with fold differences and q values indicated. Additionally, we indicate the expression level of each miRNA as estimated by our rough copy number analysis.

Table 5.5. SAM analysis of BA9 pilot study

Gene Name	Fold Change	q-value (%)	Level
hsa-mir-92	1.41	0	Medium
hsa-mir-17-5p	2.42	0	Medium
hsa-mir-151	1.5	0	Medium
hsa-mir-25	1.28	21.66	Medium
hsa-mir-342	1.33	28.87	Medium
hsa-mir-328	1.66	28.87	Medium
hsa-mir-30d	1.7	37.12	Medium
hsa-mir-222	1.59	37.12	Medium
hsa-mir-125a	2.15	37.12	High
hsa-mir-130b	1.6	37.54	Low
hsa-mir-191	1.22	37.54	Medium
hsa-mir-372	9.23	37.54	Very low
hsa-mir-296	2.06	37.54	Low
hsa-mir-149	2.05	37.54	Medium
hsa-mir-31	1.36	37.54	Low
hsa-mir-133b	2.08	37.54	Low
hsa-mir-150	1.64	38.5	Low
hsa-mir-373	3.36	38.5	Very low
hsa-mir-213	1.48	43.31	Low
hsa-mir-182	0.53	45.04	Low
hsa-mir-126	0.5	45.04	Medium
hsa-mir-193	0.31	45.04	Low
hsa-mir-330	1.26	45.28	Low
hsa-let-7b	1.84	45.28	High
hsa-mir-28	0.51	46.64	Medium

Three “over-expressed” microRNAs (mir-372, mir-373, and mir-17-5p) were observed to share seed sequence (bp 2-8 AAGUGCU) with the dre-mir-430 cluster,

which in zebrafish has been implicated in proper brain morphogenesis [206]. Two other "over-expressed" microRNAs (mir-25 and mir-92) are located within clusters that contain miRNAs with dre-mir-430 seed sequence. **Table 5.6** shows human miRNA clusters that contain dre-mir-430 like miRNAs, as determined by similarity of seed sequence.

Table 5.6. Human genomic locations of dre-mir-430 like miRNA clusters

microRNA	Chromosome	microRNA	Chromosome
25	7	371	19
93	7	372	19
106b	7	373	19
17	13	363	X
18a	13	106a	X
19a	13	18b	X
19b-1	13	19b-2	X
20a	13	20b	X
92-1	13	92-2	X

Thus, in our follow-up study in the Stanley sample set, we chose to test the expression of most human microRNAs containing the dre-mir-430 seed sequence (AAGUGCU), as well as miRNAs associated with them by virtue of a clustered genomic location. The results of the ΔΔCt and SAM analysis, irrespective of relationship to dre-mir-430 miRNAs, were also utilized to select miRNAs for inclusion in the follow up study. Thus, miRNAs were selected for inclusion in the follow up study based on two criteria: sequence relationship to the dre-mir-430 family of miRNAs, and observed misexpression in the pilot study.

Results of ΔΔCt expression analysis for the follow up study in BA9 of the Stanley sample set are shown in **Figures 5.2**. Interestingly, over-expression of mir-372 was observed in samples from patients with both schizophrenia and bipolar disorder. Mir-372 is a human homologue of dre-mir-430 miRNAs. We also assessed mir-372 expression (along with 9 other microRNAs of interest) in a separate dorsolateral prefrontal cortex region, BA46, derived from the same Stanley patients from which the BA9 samples were obtained. We again observed overexpression of mir-372 (1.5 to 2.2 fold) in both schizophrenic and bipolar samples (**Figure 5.3**). It should be noted that mir-372 is expressed at very low levels in the adult human brain, raising questions as to the biological relevance of the observed over-expression (see Discussion).

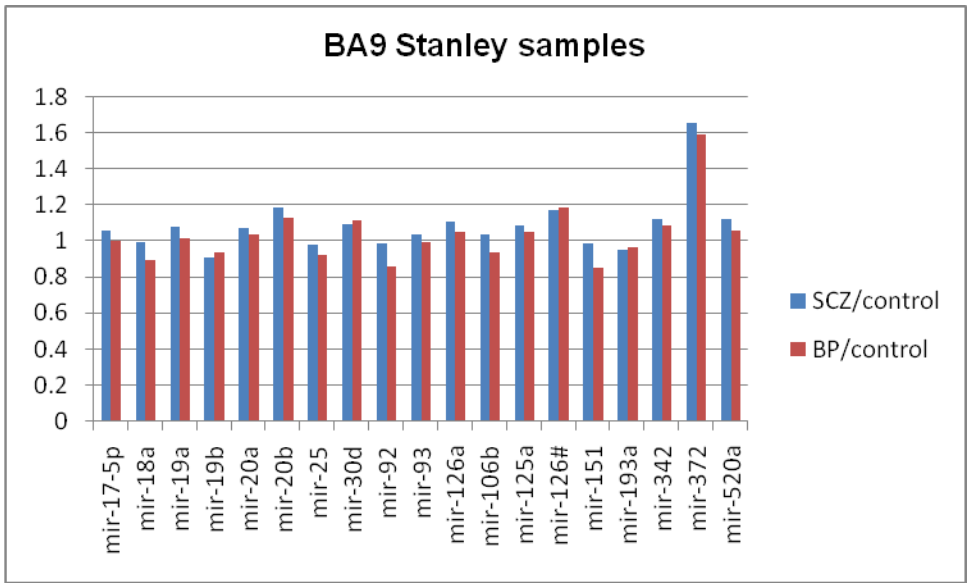

Figure 5.2. Expression differences between diagnostic groups in Stanley BA9 experiment: Figure shows results of ΔΔCt analysis for the 19 follow up miRNAs (X-axis) in the Stanley sample set. Blue bars indicate fold differences (Y-axis) between patients with schizophrenia and controls, red bars indicate fold differences between patients with bipolar disorder and controls.

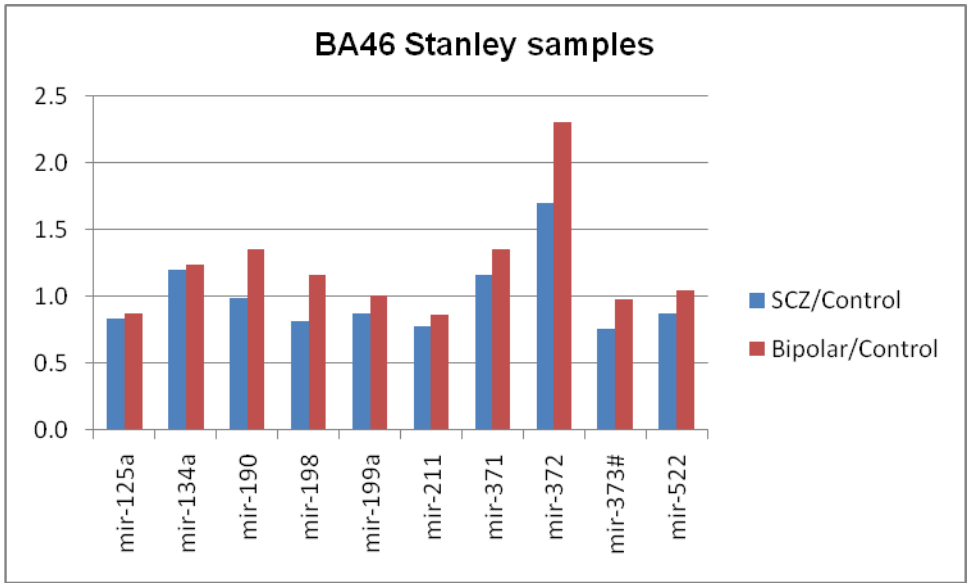

Figure 5.3. Expression differences between diagnostic groups in Stanley BA46 experiment: Figure shows results of ΔΔCt analysis for the 19 follow up miRNAs (X-axis) in the Stanley sample set. Blue bars indicate fold differences (Y-axis) between patients with schizophrenia and controls, red bars indicate fold differences between patients with bipolar disorder and controls.

It is conceivable that treatment with antipsychotic medication could be responsible for the increased expression of mir-372. To address this issue, we utilized the fact that the bipolar samples in the Stanley collection can be divided into three patient groups: never treated with antipsychotics (NT), treated in the past but not at time of death (TIP), and treated at time of death (TAD). ANOVA was performed to assess whether the bipolar samples differed in mir-372 expression based on treatment category. Though the results were not significant ($p=0.77$), there is a trend of increased expression based on treatment category. **Figure 5.4** shows the relative expression of mir-372 in the three bipolar treatment groups and in psychiatrically normal controls.

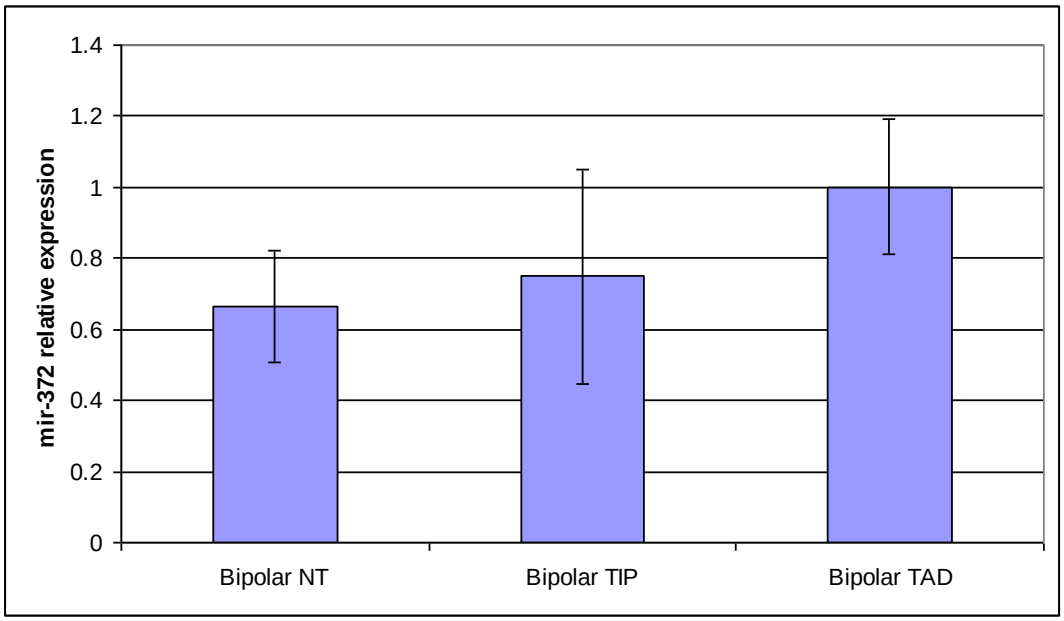

Figure 5.4. Effect of antipsychotic drug treatment on mir-372 expression: Figure represents the expression of mir-372 in BA9 (Stanley sample), relative to control, of patients with bipolar disorder divided by category of antipsychotic medication use. NT = never treated with antipsychotics, TIP = treated in past, TAD = treated at time of death. This graph indicates a non-significant (0.77, ANOVA) trend of increased expression of mir-372 with more recent antipsychotic use. Error bars represent the standard error.

ΔΔCt analysis only indicates relative expression differences, and does not assess the statistical significance of those differences. Indeed, when regression analysis was performed on the miRNAs assessed in BA9 of the Stanley sample set, accounting for covariates such as antipsychotic treatment, the observed mir-372 overexpression was not statistically significant. **Table 5.7** shows all covariates assessed and those deemed likely to impact mir-372 expression (indicated with an X).

Table 5.7. Significant mir-372 covariates controlled for in regression analysis

Covariate	
Age	
Sex	
Brain PH	X
Left Brain	
Lifetime Alcohol Use	
Lifetime Drug Use	
Smoking At TOD	
Psychotic Feature	X
Antipsychotics at death	X
Mood Stabilizer at death	X
Antidepressants at death	
Anticholinergic at death	
Storage	
Refrigerator Interval	X
Post-mortem interval	
Lifetime Antipsychotics	X
Brain Weight (kg)	

Those covariates were thus controlled for when regressing mir-372 expression against diagnostic category (bipolar, schizophrenic, psychiatrically normal control). Regression analysis reveals no statistically significant expression differences (**Table 5.8**), based on either diagnostic category or any of the covariates, for any of the 19 miRNAs tested in BA9. Included in these negative results were microRNAs reported to be differentially expressed in a report of DLPFC expression profiling of schizophrenic samples [174]. Specifically, mir-92, mir-30d, mir-20b, and mir-106b, all reported as differentially expressed, showed no evidence for misexpression in our sample set. Regression results for the 10 miRNAs assessed in BA46 of the Stanley sample set also failed to identify

significant misexpression of any miRNA based on diagnostic category or any of the covariates.

We were also interested in whether DGCR8 genotype, especially those polymorphisms associated with schizophrenia and bipolar illness in our earlier genetic study impacted the expression of any of the microRNAs assessed in the larger Stanley sample set. Genotyping of the same 12 SNPs was performed on the 105 samples in the Stanley sample set. DGCR8 genotype was regressed individually against expression of the 19 miRNAs assessed in BA9, and also against a global measure of miRNA expression which averaged the expression of all 19 miRNAs. Full results are given as p-values in **Table 5.8**. No statistically significant relationship between miRNA expression and DGCR8 genotype was observed after correcting for multiple testing. As stated earlier, there were also no significant differences based on diagnostic category.

Table 5.8. Full regression results assessing miRNA expression against both diagnosis and DGCR8 genotypes

	mir.92	mir.193a	mir.151	mir.17.5p	mir.18a	mir.19a	mir.19b
dx[+]	0.232	0.869	0.172	0.887	0.539	0.447	0.420
rs9606240	0.497	0.647	0.435	0.809	0.121	0.686	0.910
rs9606241	0.506	0.419	0.682	0.682	0.779	0.968	0.904
rs737935*	0.387	0.574	0.561	0.936	0.528	0.679	0.848
rs443678	0.843	0.361	0.696	0.458	0.810	0.201	0.116
rs1640299	0.408	0.225	0.722	0.132	0.374	0.488	0.666
rs720012**	0.313	0.193	0.696	0.447	0.287	***0.007***	***0.007***
rs1633445	0.698	**0.096**	0.701	0.780	0.822	0.993	0.987
rs885980	0.700	0.361	0.455	0.775	0.716	0.487	0.361
rs175168	0.626	0.608	0.593	0.480	0.383	0.585	0.694
rs175169	0.745	0.209	0.598	0.598	0.788	0.199	0.149
rs175174	0.748	0.252	0.483	0.733	0.796	0.356	0.224
rs175175	0.801	0.258	0.699	0.204	0.627	**0.076**	**0.033**
	mir.20a	mir.20b	mir.25	mir.30d	mir.93	mir.106a	mir.106b
dx	0.724	0.944	0.909	0.949	0.777	0.919	0.405
rs9606240	0.904	0.744	0.503	0.820	0.718	0.731	0.802
rs9606241	0.895	0.920	0.602	0.933	0.587	0.567	0.433
rs737935*	0.896	0.544	0.638	0.729	0.668	0.918	0.670
rs443678	0.240	0.338	0.673	0.401	0.599	0.892	0.704
rs1640299	0.374	0.264	0.500	0.620	0.320	0.346	0.300
rs720012**	**0.097**	**0.061**	0.488	0.164	**0.088**	0.299	0.617
rs1633445	0.868	0.356	0.389	0.299	0.618	0.516	0.761
rs885980	0.733	0.536	0.521	0.449	0.633	0.531	0.361
rs175168	0.652	0.890	0.778	0.886	0.983	0.661	0.711
rs175169	0.272	0.260	0.371	0.172	0.323	0.326	0.426
rs175174	0.565	0.459	0.459	0.310	0.523	0.429	0.357
rs175175	0.116	0.189	0.706	0.622	0.392	0.759	0.480
	mir.125a	mir.126.	mir.151.1	mir.342	mir.372	mir.520a	mir.gm[++]
dx	0.905	0.472	0.172	0.835	0.776	0.321	0.412
rs9606240	0.554	0.563	0.435	0.798	0.193	0.726	0.836
rs9606241	0.734	0.985	0.682	0.796	0.915	0.462	0.888
rs737935*	0.405	0.651	0.561	0.885	0.940	0.395	0.810
rs443678	0.841	0.317	0.696	0.442	0.297	0.445	0.210
rs1640299	0.385	0.418	0.722	0.503	0.285	0.102	0.678
rs720012**	0.355	0.161	0.696	**0.099**	0.993	**0.063**	0.163
rs1633445	0.301	0.504	0.701	0.499	0.416	0.426	0.393
rs885980	0.608	0.709	0.455	0.352	0.720	0.914	0.315
rs175168	0.792	0.873	0.593	0.847	0.326	0.615	0.867
rs175169	0.311	0.388	0.598	0.190	0.578	0.846	**0.078**
rs175174	0.510	0.676	0.483	0.284	0.697	0.970	0.251
rs175175	0.926	0.550	0.699	0.426	0.211	0.498	0.169

* associated with SCZ in genetic study **associated with BP in genetic study
[+]diagnosis [++]global miRNA expression
Nominal significance at 0.10 = **bold**, at 0.05 = **bold**, at 0.01 = ***bold italicized***

Discussion

Expression profiling of 160 miRNAs was performed in post-mortem brain samples from the dorsolateral prefrontal cortex of patients with schizophrenia and psychiatrically normal controls. Highly sensitive and quantitative TaqMan miRNA assays were utilized to assess expression in a small set of samples, and follow-up expression profiling of a subset of miRNAs was performed in a larger, better characterized sample set. At the time of this work, no publicly available data existed showing the expression patterns of miRNAs in the adult human brain. Novel and useful baseline information was obtained from this work, including which miRNAs are expressed, and at what level. Surprisingly, virtually all miRNAs (157/160) were detectable at some level of expression in the adult human brain. Using a rough estimate of copy number, we determined that nearly 1/3 of those tested are expressed at >1000 copies per cell.

The purpose of the study was to identify miRNAs differentially expressed between patients with schizophrenia, bipolar disorder, and psychiatrically normal controls. In both the pilot study and the follow-up studies in the Stanley sample set, differential expression of individual miRNAs was observed when assessing fold differences using the ΔΔCt method. However, an assessment of the statistical significance of these differences by regression analysis reveals no significant differentially expressed miRNAs, including miRNAs reported in the literature to be differentially expressed. The power to detect a significant difference in miRNA expression was assessed in the Stanley samples, using a significance level corrected for 560 miRNAs (roughly the total number of identified human miRNAs) and 2 comparisons per miRNA (BP and SCZ versus control). **Table 5.9** shows the required difference (on

the log base 10 scale) and m-fold increase (in relative expression) for 80%, 90%, and 95% power. Power calculations indicate that we had a 95% probability of detecting true difference of 3.5 fold or greater, and an 80% probability of detecting true differences 2.9 fold or greater.

Table 5.9. Power to detect expression differences in the Stanley sample set

Power	80%	90%	95%
Difference (log base 10 scale)	0.46	0.50	0.54
m-fold increase	2.92	3.21	3.47

Of note, there was a consistent observation of mir-372 overexpression in samples from patients with both schizophrenia and bipolar illness, in both the pilot study and in two separate brain regions (BA9 and BA46) of the Stanley sample set. While it appears that this overexpression may be the result of treatment with antipsychotic medication, we must be cautious in interpreting the trend in overexpression associated with antipsychotic drug treatment. While the trend may be a drug effect, it is also plausible that patients with more severe disease are more likely to be treated with antipsychotic medications, and thus it is severity of disease that is causing the trend towards mir-372 overexpression. Alternatively, mir-372 overexpression might be a compensatory mechanism against illness, and drug treatment may mimic this compensatory mechanism. It should be noted that mir-372 is expressed at very low levels in adult human brain, and we do find it difficult to believe that a gene whose expression is near the limit of detection of TaqMan assays is of biological relevance. Additionally, it is a known phenomenon that assays at

the limit of detection of real-time PCR are subject to stochastic variation at the RT step, resulting in less quantitative assays near the limit of detection. In order to ensure that the observed mir-372 overexpression was not a technical artifact influenced by its extremely low expression, we assessed the variation of mir-372 expression in biological replicates by analyzing Stanley BA46 (n=105) using two amounts of input RNA at the RT step (200 ng and 1ug). The Pearson correlation coefficient of mir-372 expression using two different RNA input amounts was $r^2 > 0.95$, suggesting we are accurately quantifying this very lowly expressed microRNA. We also note that mir-372 is strongly predicted to target the 3' UTR of the NOS1AP gene, a gene implicated in schizophrenia through both genetic and expression studies. Mir-372 may be an important regulator of NOS1AP at an earlier developmental time point when mir-372 is more highly expressed.

Previously, we have identified polymorphisms in the gene DGCR8 that are associated with both schizophrenia and bipolar disorder. DGCR8 is involved in the processing of miRNAs, and it is conceivable that functional polymorphisms in this gene could impact mature miRNA expression levels. We therefore genotyped the Stanley sample set with the same set of markers used in the DGCR8 candidate gene study. We used regression analysis to determine if any correlations existed between DGCR8 genotype and miRNA expression for the 19 miRNAs assessed in BA9 of the Stanley sample set. rs720012, a SNP significantly associated with bipolar type schizoaffective disorder and bipolar disorder, showed nominally significant association with hsa-mir-19a and hsa-mir-19b expression (both at p=0.007). Additionally, this SNP gave was nominally significant at the p=0.10 level for five other miRNAs. No other SNP showed this pattern, and it is of interest given the previous genetic finding. However, after

correcting for multiple testing none of the values remained significant. Additionally, no SNP was significantly correlated with a global measure of miRNA expression obtained by averaging the expression of all 19 miRNAs measured in BA9 of the Stanley sample.

Chapter 6

Conclusion

Review of major findings

The work presented here focused on genetic studies of schizophrenia, by examining both the DNA and RNA of schizophrenia candidate genes in samples obtained from human patients. We began by developing an inexpensive, multiplexed genotyping method that allowed us to carry out SNP analysis on human genomic DNA obtained from collections of pedigrees segregating schizophrenia. The candidate gene studies focused on two major hypotheses of schizophrenia etiology: the well known glutamate hypotheses and a novel investigation into a potential role for miRNAs in this disease. Panels of tagSNPs were tested in three candidate genes for association with schizophrenia. Those genes included SNAP-25 and ZDHHC8, both with established roles in glutamate signaling at the NMDA receptor. Additionally the DGCR8 gene, a gene involved in global miRNA processing, was assessed. The expression of many candidate miRNAs was also measured in post-mortem brain samples from patients with schizophrenia. Although the primary focus was on schizophrenia, the collections of samples utilized in these studies allowed us to assess genetic associations with schizoaffective and bipolar disorders as well.

At the time these studies were undertaken, inexpensive multiplexed genotyping technologies were not readily available. Therefore, an in-house SNP typing assay was developed, built on existing methods, that was both high-throughput and inexpensive. A two-color flow cytometer (Luminex, Austin, Texas) was used in conjunction with oligonucleotide ligation assay chemistry and polystyrene beads to develop an assay

capable of genotyping 50 SNPs simultaneously at a cost of less than $0.09 per genotype. Three innovations helped reduce the cost and time of the assay, relative to previously published work. First, we established through extensive experimental work that far fewer beads than previously reported could be used to obtain accurate and reproducible genotypes. Second, a biotinylated universal oligonucleotide was used to offset the cost of labeling OLA products. Third, reagent concentrations were optimized such that no washing of the polystyrene beads was necessary during the entire course of the genotyping assay, in contrast to all previously published Luminex-based genotyping protocols. While new methods are continually being developed and old methods are becoming less expensive, our method remains a very inexpensive and reliable method that is widely useful to scientists performing targeted genetic studies.

We next utilized the genotyping assay to perform a candidate gene study of SNAP-25, a t-SNARE protein involved in presynaptic vesicle docking and fusion with the lipid membrane. The function of SNAP-25, along with expression studies and its genomic location at 20p12-p11.2, make it an intriguing candidate gene for schizophrenia. No thorough assessment of SNAP-25 as a schizophrenia candidate gene has previously been published. A tagSNP approach was utilized to select SNPs for genotyping, and we assessed genetic association with schizophrenia in two collections of pedigrees of Caucasian ancestry. This study provided evidence of association with several SNPs within SNAP-25 and suggested that a causal mutation is present in the 3’ half of the gene. The presence of several nominally significant markers near exon 5 suggests that alternative splicing of this exon and the resultant SNAP-25 isoforms may be involved in schizophrenia etiology, possibly due to the altered capacity of SNAP-25 to regulate

synaptic vesicle fusion. This represents the first genetic association between the SNAP-25 gene and schizophrenia.

A study was also performed evaluating multiple genes within a 75 kb region on chromosome 22q11.2. This region of the genome is very likely to harbor schizophrenia susceptibility loci, with evidence from two sources. First, patients who are suffering from 22q11DS – a disorder involving a common 3 Mb hemizygous deletion – have a greatly increased risk of neuropsychiatric disorders including schizophrenia. Second, genetic studies, including linkage and association analyses, have repeatedly implicated this region in schizophrenia. Our study focused on the chromosome 22 genes ZDHHC8 and DGCR8, genes with prior evidence (ZDHHC8) and no prior evidence (DGCR8) of association with schizophrenia. We failed to replicate previous findings of genetic association with the ZDHHC8 gene. We did, however, find convincing evidence that two SNPs within the DGCR8 gene are associated with psychotic disorders. SNP rs737935, located within exon 9 of DGCR8, was associated with schizophrenia in a large sample set collected by the NIMH, and in combined analysis of the NIMH sample and a separate sample from McLean Hospital. SNP rs720012, located in the 3' UTR of DGCR8, was associated with psychotic disorders with a significant affective component in the NIMH sample set. Additionally, rs720012 was associated with a quantitative measure of craniofacial dysmorphology in subjects with affective psychoses in the McLean sample. The specific dysmorphology used to define the quantitative phenotype is based on a series of anomalies at the frontonasal-maxillary juncture, and has previously been linked to schizophrenia and correlated with anterior-posterior brain midline dysmorphology.

Finally, we performed expression analysis of miRNAs in post-mortem prefrontal cortex tissue from patients with schizophrenia, bipolar disorder, and psychiatrically normal controls. An initial study measured the expression of 160 miRNAs in a small number of samples, followed by screening of individual miRNAs of interest in a much larger sample set. We experimentally determined appropriate housekeeping genes for miRNA expression studies of the human brain, and profiled the expression of a large number of miRNAs in the prefrontal cortex of adult human brain. No statistically significant expression differences were observed based on the diagnostic category in the larger sample set, including four miRNAs reported as misexpressed by another group. We tested, and failed to find supporting evidence for, the hypothesis that a family of human miRNAs related to the zebrafish mir-430 family were misexpressed in samples from patients with schizophrenia or bipolar disorder. We do note, however, that we observed consistent overexpression of one mir-430 like miRNA, hsa-mir-372, in samples from patients with both schizophrenia and bipolar disorder. There are two major caveats to this observation; it appears that this overexpression is correlated with the use of antipsychotic medication, and the absolute expression level is quite low. Additionally, we assessed whether miRNA expression was associated with DGCR8 SNP genotypes using the same SNPs examined in our previous candidate gene study. While no statistically significant correlation was found between DGCR8 genotype and miRNA expression, there was a trend of nominally significant p-values observed for multiple miRNAs when regressing their expression against rs720012 genotype.

Future Work

Investigation into the role that miRNAs play in major neuropsychiatric disorders is a novel, and potentially powerful, means by which to understand these enigmatic diseases. Our miRNA expression work was initially limited by the lack of available profiling methods in early 2004, when the work was first proposed. The importance of miRNAs in higher organisms was only just being realized, and large scale expression studies of human diseases were non-existent. While we struggled to use miRNA expression profiling methods that were still being developed during 2004 and 2005, the advent of commercially available assays in 2005 - 2006 allowed us to achieve accurate and reproducible expression profiling of miRNAs in post-mortem human tissue samples. Still, of the total number of miRNAs now known to exist, only a portion were included in first generation assays. In 2006 - 2007, there was rapid development of commercially available miRNA expression platforms, finally allowing comprehensive expression profiling. The Brzustowicz lab is currently profiling all known human miRNAs, in various tissues and at various stages of human development, using highly sensitive and quantitative TaqMan miRNA assays. This dataset will be invaluable not only for understanding the role of miRNAs in schizophrenia, but will represent the first truly quantitative global expression atlas of miRNAs in the human brain.

This expression atlas will also be useful in expanding the work presented in this dissertation. Global miRNA expression patterns can be correlated with the disease-associated DGCR8 polymorphisms identified in Chapter 4. While a similar analysis was presented in Chapter 5, it was limited by the small number of miRNAs profiled. Additionally, mRNA expression arrays have been performed on the same samples that

are currently being profiled for miRNA expression. Joint analysis of these datasets will allow assessment of miRNA impact on predicted mRNA targets in-vivo; no such studies are currently reported in the literature. This is one method, though not the only, that will be used to begin to identify the protein coding mRNA targets of miRNAs misexpressed in neuropsychiatric disorders, with the ultimate aim of identifying the core molecular pathways disrupted in schizophrenia.

Ultimately, functional analysis will be necessary to fully understand the role of miRNAs involved in schizophrenia. Cell culture transfection experiments are one way, of many, to address the functional importance of misexpression. Transfection experiments of hsa-mir-372, the only misexpressed miRNA identified in this work, have recently been performed by a member of the Brzustowicz lab. Synthetic hsa-mir-372 was transfected into rat primary cortical precursors, to mimic the overexpression we observed in post-mortem samples. In this experiment, overexpression of hsa-mir-372 resulted in a significant increase in neurite outgrowth, the only notable phenotype that was observed. I will end this discussion with an image of that experiment (**Figure 6.1**), to demonstrate the power of a single misexpressed miRNA as well as to indicate potential future directions of this research.

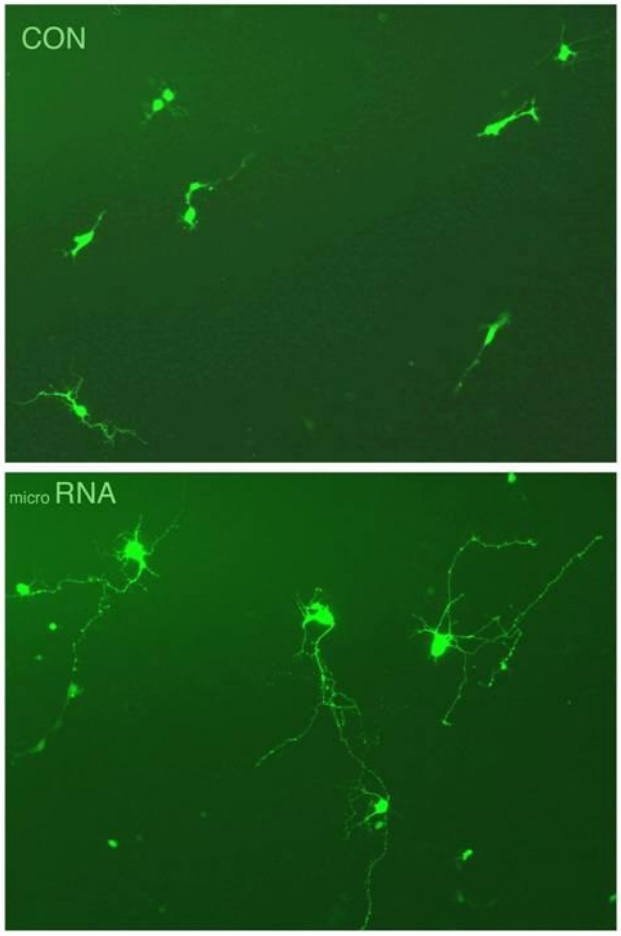

Figure 6.1. Misexpression of mir-372 in rat cerebral cortical precursors enhances neuritic outgrowth: Effects of miR-372 were examined by transfecting the pre-miR-372 vector (Ambion) into cortical precursors. Pre-miR vectors are synthetic microRNAs designed to mimic endogenous mature miRNA molecules. Pre-miR-372 was transfected at 3, 30 and 100nMol concentrations that were co-transfected with a reporter construct, EGFP (CLONTECH). Morphological differences in cortical precursor neurite outgrowth at 30nMol pre-miR-372 were observed compared to cells transfected with EGFP vector alone, while there was prominent cell death at the highest concentration. The figure demonstrates EGFP control transfection alone (top) and with pre-miR-372 (bottom). The processes on pre-miR-372 transfected cells were frequently longer than controls and had more branch points. (Work performed by Michael Moreau in collaboration with Emmanuel DiCicco-Bloom)

References

1. Goldner, E.M., et al., *Prevalence and incidence studies of schizophrenic disorders: a systematic review of the literature.* Can J Psychiatry, 2002. **47**(9): p. 833-43.
2. Jablensky, A., et al., *Schizophrenia: manifestations, incidence and course in different cultures. A World Health Organization ten-country study.* Psychol Med Monogr Suppl, 1992. **20**: p. 1-97.
3. Avila, M., G. Thaker, and H. Adami, *Genetic epidemiology and schizophrenia: a study of reproductive fitness.* Schizophr Res, 2001. **47**(2-3): p. 233-41.
4. Jablensky, A., *The 100-year epidemiology of schizophrenia.* Schizophr Res, 1997. **28**(2-3): p. 111-25.
5. Mueser, K.T. and S.R. McGurk, *Schizophrenia.* Lancet, 2004. **363**(9426): p. 2063-72.
6. Castle, D., et al., *The incidence of operationally defined schizophrenia in Camberwell, 1965-84.* Br J Psychiatry, 1991. **159**: p. 790-4.
7. Harding, C.M., et al., *The Vermont longitudinal study of persons with severe mental illness, II: Long-term outcome of subjects who retrospectively met DSM-III criteria for schizophrenia.* Am J Psychiatry, 1987. **144**(6): p. 727-35.
8. Harrison, G., et al., *Recovery from psychotic illness: a 15- and 25-year international follow-up study.* Br J Psychiatry, 2001. **178**: p. 506-17.
9. Ustun, T.B., et al., *Multiple-informant ranking of the disabling effects of different health conditions in 14 countries. WHO/NIH Joint Project CAR Study Group.* Lancet, 1999. **354**(9173): p. 111-5.
10. Auquier, P., et al., *Mortality in schizophrenia.* Pharmacoepidemiol Drug Saf, 2007. **16**(12): p. 1308-1312.
11. Cardno, A.G. and Gottesman, II, *Twin studies of schizophrenia: from bow-and-arrow concordances to star wars Mx and functional genomics.* Am J Med Genet, 2000. **97**(1): p. 12-7.
12. Sullivan, P.F., K.S. Kendler, and M.C. Neale, *Schizophrenia as a complex trait: evidence from a meta-analysis of twin studies.* Arch Gen Psychiatry, 2003. **60**(12): p. 1187-92.
13. McGuffin, P., *Nature and nurture interplay: schizophrenia.* Psychiatr Prax, 2004. **31 Suppl 2**: p. S189-93.
14. Risch, N. and K. Merikangas, *The future of genetic studies of complex human diseases.* Science, 1996. **273**(5281): p. 1516-7.
15. Lewis, C.M., et al., *Genome scan meta-analysis of schizophrenia and bipolar disorder, part II: Schizophrenia.* Am J Hum Genet, 2003. **73**(1): p. 34-48.
16. O'Donovan, M.C., N.M. Williams, and M.J. Owen, *Recent advances in the genetics of schizophrenia.* Hum Mol Genet, 2003. **12 Spec No 2**: p. R125-33.
17. Harrison, P.J. and D.R. Weinberger, *Schizophrenia genes, gene expression, and neuropathology: on the matter of their convergence.* Mol Psychiatry, 2004.
18. Pearlson, G.D. and L. Marsh, *Structural brain imaging in schizophrenia: a selective review.* Biol Psychiatry, 1999. **46**(5): p. 627-49.
19. Nelson, M.D., et al., *Hippocampal volume reduction in schizophrenia as assessed by magnetic resonance imaging: a meta-analytic study.* Arch Gen Psychiatry, 1998. **55**(5): p. 433-40.
20. Konick, L.C. and L. Friedman, *Meta-analysis of thalamic size in schizophrenia.* Biol Psychiatry, 2001. **49**(1): p. 28-38.
21. Lawrie, S.M. and S.S. Abukmeil, *Brain abnormality in schizophrenia. A systematic and quantitative review of volumetric magnetic resonance imaging studies.* Br J Psychiatry, 1998. **172**: p. 110-20.
22. Harrison, P.J., N. Freemantle, and J.R. Geddes, *Meta-analysis of brain weight in schizophrenia.* Schizophr Res, 2003. **64**(1): p. 25-34.
23. Garey, L.J., et al., *Reduced dendritic spine density on cerebral cortical pyramidal neurons in schizophrenia.* J Neurol Neurosurg Psychiatry, 1998. **65**(4): p. 446-53.
24. Glantz, L.A. and D.A. Lewis, *Decreased dendritic spine density on prefrontal cortical pyramidal neurons in schizophrenia.* Arch Gen Psychiatry, 2000. **57**(1): p. 65-73.
25. Rosoklija, G., et al., *Structural abnormalities of subicular dendrites in subjects with schizophrenia and mood disorders: preliminary findings.* Arch Gen Psychiatry, 2000. **57**(4): p. 349-56.

26. Kovalenko, S., et al., *Regio entorhinalis in schizophrenia: more evidence for migrational disturbances and suggestions for a new biological hypothesis.* Pharmacopsychiatry, 2003. **36 Suppl 3**: p. S158-61.
27. Bagary, M.S., et al., *Gray and white matter brain abnormalities in first-episode schizophrenia inferred from magnetization transfer imaging.* Arch Gen Psychiatry, 2003. **60**(8): p. 779-88.
28. Davis, K.L., et al., *White matter changes in schizophrenia: evidence for myelin-related dysfunction.* Arch Gen Psychiatry, 2003. **60**(5): p. 443-56.
29. Javitt, D.C. and S.R. Zukin, *Recent advances in the phencyclidine model of schizophrenia.* Am J Psychiatry, 1991. **148**(10): p. 1301-8.
30. Kegeles, L.S., et al., *NMDA antagonist effects on striatal dopamine release: positron emission tomography studies in humans.* Synapse, 2002. **43**(1): p. 19-29.
31. Krystal, J.H., et al., *Subanesthetic effects of the noncompetitive NMDA antagonist, ketamine, in humans. Psychotomimetic, perceptual, cognitive, and neuroendocrine responses.* Arch Gen Psychiatry, 1994. **51**(3): p. 199-214.
32. Coyle, J.T. and G. Tsai, *NMDA receptor function, neuroplasticity, and the pathophysiology of schizophrenia.* Int Rev Neurobiol, 2004. **59**: p. 491-515.
33. Pietraszek, M., *Significance of dysfunctional glutamatergic transmission for the development of psychotic symptoms.* Pol J Pharmacol, 2003. **55**(2): p. 133-54.
34. Clinton, S.M. and J.H. Meador-Woodruff, *Abnormalities of the NMDA Receptor and Associated Intracellular Molecules in the Thalamus in Schizophrenia and Bipolar Disorder.* Neuropsychopharmacology, 2004. **29**(7): p. 1353-62.
35. Natchkebia , N.N.a.M.D., *Effect of phencyclidine, MK-801 and dextrorphan on the viability of neuronal/glial cells and expression of NMDA-receptor subunits, PSD-95, nNOS and Ras-GRF1.* The Journal of Biological Physics and Chemistry, 2003. **3**(1): p. 18-22.
36. Brzustowicz, L.M., et al., *Linkage disequilibrium mapping of schizophrenia susceptibility to the CAPON region of chromosome 1q22.* Am J Hum Genet, 2004. **74**(5): p. 1057-63.
37. Mukai, J., et al., *Evidence that the gene encoding ZDHHC8 contributes to the risk of schizophrenia.* Nat Genet, 2004. **36**(7): p. 725-31. Epub 2004 Jun 06.
38. El-Husseini Ael, D., et al., *Synaptic strength regulated by palmitate cycling on PSD-95.* Cell, 2002. **108**(6): p. 849-63.
39. Stefansson, H., et al., *Neuregulin 1 and schizophrenia.* Ann Med, 2004. **36**(1): p. 62-71.
40. Bao, J., et al., *Activity-dependent transcription regulation of PSD-95 by neuregulin-1 and Eos.* Nat Neurosci, 2004. **7**(11): p. 1250-8.
41. Talbot, K., et al., *Dysbindin-1 is reduced in intrinsic, glutamatergic terminals of the hippocampal formation in schizophrenia.* J Clin Invest, 2004. **113**(9): p. 1353-63.
42. Schumacher, J., et al., *Examination of G72 and D-amino-acid oxidase as genetic risk factors for schizophrenia and bipolar affective disorder.* Mol Psychiatry, 2004. **9**(2): p. 203-7.
43. Polesskaya, O.O., et al., *Novel putative nonprotein-coding RNA gene from 11q14 displays decreased expression in brains of patients with schizophrenia.* J Neurosci Res, 2003. **74**(1): p. 111-22.
44. Walikonis, R.S., et al., *Identification of proteins in the postsynaptic density fraction by mass spectrometry.* J Neurosci, 2000. **20**(11): p. 4069-80.
45. Brzustowicz, L.M., et al., *Location of a major susceptibility locus for familial schizophrenia on chromosome 1q21-q22.* Science, 2000. **288**(5466): p. 678-82.
46. Fire, A., et al., *Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans.* Nature, 1998. **391**(6669): p. 806-11.
47. Catalanotto, C., et al., *Involvement of small RNAs and role of the qde genes in the gene silencing pathway in Neurospora.* Genes Dev, 2002. **16**(7): p. 790-5.
48. Hamilton, A.J. and D.C. Baulcombe, *A species of small antisense RNA in posttranscriptional gene silencing in plants.* Science, 1999. **286**(5441): p. 950-2.
49. Zamore, P.D., *Ancient pathways programmed by small RNAs.* Science, 2002. **296**(5571): p. 1265-9.
50. Carninci, P., et al., *The transcriptional landscape of the mammalian genome.* Science, 2005. **309**(5740): p. 1559-63.
51. Mehler, M.F. and J.S. Mattick, *Non-coding RNA in the nervous system.* J Physiol, 2006.

52. Lee, R.C., R.L. Feinbaum, and V. Ambros, *The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14.* Cell, 1993. **75**(5): p. 843-54.
53. Lau, N.C., et al., *An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans.* Science, 2001. **294**(5543): p. 858-62.
54. Lee, R.C. and V. Ambros, *An extensive class of small RNAs in Caenorhabditis elegans.* Science, 2001. **294**(5543): p. 862-4.
55. Bentwich, I., et al., *Identification of hundreds of conserved and nonconserved human microRNAs.* Nat Genet, 2005. **37**(7): p. 766-70.
56. Lee, Y., et al., *MicroRNA genes are transcribed by RNA polymerase II.* Embo J, 2004. **23**(20): p. 4051-60.
57. Rodriguez, A., et al., *Identification of mammalian microRNA host genes and transcription units.* Genome Res, 2004. **14**(10A): p. 1902-10.
58. Giraldez, A.J., et al., *Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs.* Science, 2006. **312**(5770): p. 75-9.
59. Standart, N. and R.J. Jackson, *MicroRNAs repress translation of m7Gppp-capped target mRNAs in vitro by inhibiting initiation and promoting deadenylation.* Genes Dev, 2007. **21**(16): p. 1975-82.
60. Grimson, A., et al., *MicroRNA targeting specificity in mammals: determinants beyond seed pairing.* Mol Cell, 2007. **27**(1): p. 91-105.
61. Cai, X., C.H. Hagedorn, and B.R. Cullen, *Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs.* Rna, 2004. **10**(12): p. 1957-66.
62. Xie, Z., K.D. Kasschau, and J.C. Carrington, *Negative feedback regulation of Dicer-Like1 in Arabidopsis by microRNA-guided mRNA degradation.* Curr Biol, 2003. **13**(9): p. 784-9.
63. Landthaler, M., A. Yalcin, and T. Tuschl, *The Human DiGeorge Syndrome Critical Region Gene 8 and Its D. melanogaster Homolog Are Required for miRNA Biogenesis.* Curr Biol, 2004. **14**(23): p. 2162-7.
64. Lund, E., et al., *Nuclear export of microRNA precursors.* Science, 2004. **303**(5654): p. 95-8.
65. Hutvagner, G., et al., *A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA.* Science, 2001. **293**(5531): p. 834-8.
66. Pasquinelli, A.E. and G. Ruvkun, *Control of developmental timing by micrornas and their targets.* Annu Rev Cell Dev Biol, 2002. **18**: p. 495-513.
67. Poy, M.N., et al., *A pancreatic islet-specific microRNA regulates insulin secretion.* Nature, 2004. **432**(7014): p. 226-30.
68. Esau, C., et al., *MicroRNA-143 regulates adipocyte differentiation.* J Biol Chem, 2004. **279**(50): p. 52361-5. Epub .
69. Suh, M.R., et al., *Human embryonic stem cells express a unique set of microRNAs.* Dev Biol, 2004. **270**(2): p. 488-98.
70. Krichevsky, A.M., et al., *A microRNA array reveals extensive regulation of microRNAs during brain development.* Rna, 2003. **9**(10): p. 1274-81.
71. Krichevsky, A.M., et al., *Specific microRNAs modulate embryonic stem cell-derived neurogenesis.* Stem Cells, 2006. **24**(4): p. 857-64.
72. Hornstein, E. and N. Shomron, *Canalization of development by microRNAs.* Nat Genet, 2006. **38 Suppl**: p. S20-4.
73. Miska, E.A., et al., *Microarray analysis of microRNA expression in the developing mammalian brain.* Genome Biol, 2004. **5**(9): p. R68.
74. Schratt, G.M., et al., *A brain-specific microRNA regulates dendritic spine development.* Nature, 2006. **439**(7074): p. 283-9.
75. Giraldez, A.J., et al., *MicroRNAs regulate brain morphogenesis in zebrafish.* Science, 2005. **308**(5723): p. 833-8.
76. John, B., et al., *Human MicroRNA targets.* PLoS Biol, 2004. **2**(11): p. e363.
77. Abelson, J.F., et al., *Sequence variants in SLITRK1 are associated with Tourette's syndrome.* Science, 2005. **310**(5746): p. 317-20.
78. Jin, P., R.S. Alisch, and S.T. Warren, *RNA and microRNAs in fragile X mental retardation.* Nat Cell Biol, 2004. **6**(11): p. 1048-53.
79. Comery, T.A., et al., *Abnormal dendritic spines in fragile X knockout mice: maturation and pruning deficits.* Proc Natl Acad Sci U S A, 1997. **94**(10): p. 5401-4.

80. John, B., et al., *Human MicroRNA targets.* PLoS Biol, 2004. **2**(11): p. e363. Epub 2004 Oct 05.
81. Berezikov, E., et al., *Diversity of microRNAs in human and chimpanzee brain.* Nat Genet, 2006. **38**(12): p. 1375-7.
82. Craddock, N., M.C. O'Donovan, and M.J. Owen, *Genes for schizophrenia and bipolar disorder? Implications for psychiatric nosology.* Schizophr Bull, 2006. **32**(1): p. 9-16.
83. Maier, W., et al., *Genetic models of schizophrenia and bipolar disorder: overlapping inheritance or discrete genotypes?* Eur Arch Psychiatry Clin Neurosci, 2005. **255**(3): p. 159-66.
84. Craddock, N., M.C. O'Donovan, and M.J. Owen, *The genetics of schizophrenia and bipolar disorder: dissecting psychosis.* J Med Genet, 2005. **42**(3): p. 193-204.
85. Berrettini, W., *Bipolar disorder and schizophrenia: convergent molecular data.* Neuromolecular Med, 2004. **5**(1): p. 109-17.
86. Maier, W., A. Zobel, and M. Wagner, *Schizophrenia and bipolar disorder: differences and overlaps.* Curr Opin Psychiatry, 2006. **19**(2): p. 165-70.
87. Murray, R.M., et al., *A developmental model for similarities and dissimilarities between schizophrenia and bipolar disorder.* Schizophr Res, 2004. **71**(2-3): p. 405-16.
88. Giacomini, K.M., et al., *The pharmacogenetics research network: from SNP discovery to clinical drug response.* Clin Pharmacol Ther, 2007. **81**(3): p. 328-45.
89. Thomas, F.J., H.L. McLeod, and J.W. Watters, *Pharmacogenomics: the influence of genomic variation on drug response.* Curr Top Med Chem, 2004. **4**(13): p. 1399-409.
90. Lu, J., et al., *MicroRNA expression profiles classify human cancers.* Nature, 2005. **435**(7043): p. 834-8.
91. Dunbar, S.A., et al., *Quantitative, multiplexed detection of bacterial pathogens: DNA and protein applications of the Luminex LabMAP system.* J Microbiol Methods, 2003. **53**(2): p. 245-52.
92. Flagella, M., et al., *A multiplex branched DNA assay for parallel quantitative gene expression profiling.* Anal Biochem, 2006. **352**(1): p. 50-60.
93. Zhang, A., et al., *Small interfering RNA and gene expression analysis using a multiplex branched DNA assay without RNA purification.* J Biomol Screen, 2005. **10**(6): p. 549-56.
94. Landegren, U., et al., *A ligase-mediated gene detection technique.* Science, 1988. **241**(4869): p. 1077-80.
95. Iannone, M.A., et al., *Multiplexed single nucleotide polymorphism genotyping by oligonucleotide ligation and flow cytometry.* Cytometry, 2000. **39**(2): p. 131-40.
96. McNamara, D.T., et al., *Diagnosing infection levels of four human malaria parasite species by a polymerase chain reaction/ligase detection reaction fluorescent microsphere-based assay.* Am J Trop Med Hyg, 2006. **74**(3): p. 413-21.
97. Mehlotra, R.K., et al., *Prevalence of CYP2B6 alleles in malaria-endemic populations of West Africa and Papua New Guinea.* Eur J Clin Pharmacol, 2006. **62**(4): p. 267-75.
98. Macdonald, S.J., et al., *A low-cost open-source SNP genotyping platform for association mapping applications.* Genome Biol, 2005. **6**(12): p. R105.
99. Dunbar, S.A., *Applications of Luminex xMAP technology for rapid, high-throughput multiplexed nucleic acid detection.* Clin Chim Acta, 2006. **363**(1-2): p. 71-82.
100. Bortolin, S., et al., *Analytical validation of the tag-it high-throughput microsphere-based universal array genotyping platform: application to the multiplex detection of a panel of thrombophilia-associated single-nucleotide polymorphisms.* Clin Chem, 2004. **50**(11): p. 2028-36.
101. Pickering, J.W., et al., *Flow cytometric assay for genotyping cytochrome p450 2C9 and 2C19: comparison with a microelectronic DNA array.* Am J Pharmacogenomics, 2004. **4**(3): p. 199-207.
102. Taylor, J.D., et al., *Flow cytometric platform for high-throughput single nucleotide polymorphism analysis.* Biotechniques, 2001. **30**(3): p. 661-6, 668-9.
103. Ye, F., et al., *Fluorescent microsphere-based readout technology for multiplexed human single nucleotide polymorphism analysis and bacterial identification.* Hum Mutat, 2001. **17**(4): p. 305-16.
104. Wang, H.Y., et al., *A genotyping system capable of simultaneously analyzing >1000 single nucleotide polymorphisms in a haploid genome.* Genome Res, 2005. **15**(2): p. 276-83.
105. Ahmadian, A., et al., *Single-nucleotide polymorphism analysis by pyrosequencing.* Anal Biochem, 2000. **280**(1): p. 103-10.
106. Ronaghi, M., M. Uhlen, and P. Nyren, *A sequencing method based on real-time pyrophosphate.* Science, 1998. **281**(5375): p. 363, 365.

107. Lovmar, L., et al., *Silhouette scores for assessment of SNP genotype clusters.* BMC Genomics, 2005. **6**(1): p. 35.
108. Jacobson, J.W., et al., *Analysis of individual data from bead-based assays ("bead arrays").* Cytometry A, 2006. **69**(5): p. 384-90.
109. Toulopoulou, T., et al., *The relationship between volumetric brain changes and cognitive function: a family study on schizophrenia.* Biol Psychiatry, 2004. **56**(6): p. 447-53.
110. Mallard, E.C., et al., *Ventriculomegaly and reduced hippocampal volume following intrauterine growth-restriction: implications for the aetiology of schizophrenia.* Schizophr Res, 1999. **40**(1): p. 11-21.
111. Sheng, Z.H., et al., *Calcium-dependent interaction of N-type calcium channels with the synaptic core complex.* Nature, 1996. **379**(6564): p. 451-4.
112. Washbourne, P., et al., *Genetic ablation of the t-SNARE SNAP-25 distinguishes mechanisms of neuroexocytosis.* Nat Neurosci, 2002. **5**(1): p. 19-26.
113. Blouin, J.L., et al., *Schizophrenia susceptibility loci on chromosomes 13q32 and 8p21.* Nat Genet, 1998. **20**(1): p. 70-3.
114. Coon, H., et al., *Genomic scan for genes predisposing to schizophrenia.* Am J Med Genet, 1994. **54**(1): p. 59-71.
115. Coon, H., et al., *Evidence for a chromosome 2p13-14 schizophrenia susceptibility locus in families from Palau, Micronesia.* Mol Psychiatry, 1998. **3**(6): p. 521-7.
116. DeLisi, L.E., et al., *Genome-wide scan for linkage to schizophrenia in a Spanish-origin cohort from Costa Rica.* Am J Med Genet, 2002. **114**(5): p. 497-508.
117. DeLisi, L.E., et al., *A genome-wide scan for linkage to chromosomal regions in 382 sibling pairs with schizophrenia or schizoaffective disorder.* Am J Psychiatry, 2002. **159**(5): p. 803-12.
118. Garver, D.L., et al., *Schizophrenia spectrum disorders: an autosomal-wide scan in multiplex pedigrees.* Schizophr Res, 2001. **52**(3): p. 145-60.
119. Gurling, H.M., et al., *Genomewide genetic linkage analysis confirms the presence of susceptibility loci for schizophrenia, on chromosomes 1q32.2, 5q33.2, and 8p21-22 and provides support for linkage to schizophrenia, on chromosomes 11q23.3-24 and 20q12.1-11.23.* Am J Hum Genet, 2001. **68**(3): p. 661-73.
120. Kaufmann, C.A., et al., *NIMH Genetics Initiative Millenium Schizophrenia Consortium: linkage analysis of African-American pedigrees.* Am J Med Genet, 1998. **81**(4): p. 282-9.
121. Levinson, D.F., et al., *Genome scan of schizophrenia.* Am J Psychiatry, 1998. **155**(6): p. 741-50.
122. Lindholm, E., et al., *A schizophrenia-susceptibility locus at 6q25, in one of the world's largest reported pedigrees.* Am J Hum Genet, 2001. **69**(1): p. 96-105.
123. Moises, H.W., et al., *An international two-stage genome-wide search for schizophrenia susceptibility genes.* Nat Genet, 1995. **11**(3): p. 321-4.
124. Paunio, T., et al., *Genome-wide scan in a nationwide study sample of schizophrenia families in Finland reveals susceptibility loci on chromosomes 2q and 5q.* Hum Mol Genet, 2001. **10**(26): p. 3037-48.
125. Schwab, S.G., et al., *A genome-wide autosomal screen for schizophrenia susceptibility loci in 71 families with affected siblings: support for loci on chromosome 10p and 6.* Mol Psychiatry, 2000. **5**(6): p. 638-49.
126. Straub, R.E., et al., *Genome-wide scans of three independent sets of 90 Irish multiplex schizophrenia families and follow-up of selected regions in all families provides evidence for multiple susceptibility genes.* Mol Psychiatry, 2002. **7**(6): p. 542-59.
127. Williams, N.M., et al., *A two-stage genome scan for schizophrenia susceptibility genes in 196 affected sibling pairs.* Hum Mol Genet, 1999. **8**(9): p. 1729-39.
128. Augustine, G.J., et al., *Proteins involved in synaptic vesicle trafficking.* J Physiol, 1999. **520 Pt 1**: p. 33-41.
129. Bennett, M.K. and R.H. Scheller, *A molecular description of synaptic vesicle membrane trafficking.* Annu Rev Biochem, 1994. **63**: p. 63-100.
130. Sanderfoot, A.A. and N.V. Raikhel, *The specificity of vesicle trafficking: coat proteins and SNAREs.* Plant Cell, 1999. **11**(4): p. 629-42.
131. Martin, T.F., *Tuning exocytosis for speed: fast and slow modes.* Biochim Biophys Acta, 2003. **1641**(2-3): p. 157-65.

132. Pabst, S., et al., *Rapid and selective binding to the synaptic SNARE complex suggests a modulatory role of complexins in neuroexocytosis.* J Biol Chem, 2002. **277**(10): p. 7838-48.
133. Sutton, R.B., et al., *Crystal structure of a SNARE complex involved in synaptic exocytosis at 2.4 A resolution.* Nature, 1998. **395**(6700): p. 347-53.
134. Zhang, Y., et al., *A partially zipped SNARE complex stabilized by the membrane.* J Biol Chem, 2005. **280**(16): p. 15595-600.
135. Tachikawa, H., et al., *Polymorphism of the 5'-upstream region of the human SNAP-25 gene: an association analysis with schizophrenia.* Neuropsychobiology, 2001. **43**(3): p. 131-3.
136. Bassett, A.S., et al., *Positive and negative symptoms in families with schizophrenia.* Schizophr Res, 1993. **11**(1): p. 9-19.
137. Bassett, A.S. and W.G. Honer, *Evidence for anticipation in schizophrenia.* Am J Hum Genet, 1994. **54**(5): p. 864-70.
138. Halldorsson, B.V., et al., *Optimal haplotype block-free selection of tagging SNPs for genome-wide association studies.* Genome Res, 2004. **14**(8): p. 1633-40.
139. Goring, H.H. and J.D. Terwilliger, *Linkage analysis in the presence of errors IV: joint pseudomarker analysis of linkage and/or linkage disequilibrium on a mixture of pedigrees and singletons when the mode of inheritance cannot be accurately specified.* Am J Hum Genet, 2000. **66**(4): p. 1310-27.
140. Thompson, P.M., A.C. Sower, and N.I. Perrone-Bizzozero, *Altered levels of the synaptosomal associated protein SNAP-25 in schizophrenia.* Biol Psychiatry, 1998. **43**(4): p. 239-43.
141. Fatemi, S.H., et al., *Altered levels of the synaptosomal associated protein SNAP-25 in hippocampus of subjects with mood disorders and schizophrenia.* Neuroreport, 2001. **12**(15): p. 3257-62.
142. Thompson, P.M., S. Egbufoama, and M.P. Vawter, *SNAP-25 reduction in the hippocampus of patients with schizophrenia.* Prog Neuropsychopharmacol Biol Psychiatry, 2003. **27**(3): p. 411-7.
143. Thompson, P.M., et al., *Elevated cerebrospinal fluid SNAP-25 in schizophrenia.* Biol Psychiatry, 2003. **53**(12): p. 1132-7.
144. Numakawa, T., et al., *Evidence of novel neuronal functions of dysbindin, a susceptibility gene for schizophrenia.* Hum Mol Genet, 2004. **13**(21): p. 2699-708.
145. Bark, I.C., et al., *Differential expression of SNAP-25 protein isoforms during divergent vesicle fusion events of neural development.* Proc Natl Acad Sci U S A, 1995. **92**(5): p. 1510-4.
146. Ross, C.A., et al., *Neurobiology of schizophrenia.* Neuron, 2006. **52**(1): p. 139-53.
147. Gogos, J.A. and D.J. Gerber, *Schizophrenia susceptibility genes: emergence of positional candidates and future directions.* Trends Pharmacol Sci, 2006. **27**(4): p. 226-33.
148. Bulayeva, K.B., et al., *Genome-wide linkage scan of schizophrenia: a cross-isolate study.* Genomics, 2007. **89**(2): p. 167-77.
149. Hamshere, M.L., et al., *Genomewide linkage scan in schizoaffective disorder: significant evidence for linkage at 1q42 close to DISC1, and suggestive evidence at 22q11 and 19p13.* Arch Gen Psychiatry, 2005. **62**(10): p. 1081-8.
150. Williams, N.M., et al., *A systematic genomewide linkage study in 353 sib pairs with schizophrenia.* Am J Hum Genet, 2003. **73**(6): p. 1355-67.
151. Weksberg, R., et al., *Molecular characterization of deletion breakpoints in adults with 22q11 deletion syndrome.* Hum Genet, 2007. **120**(6): p. 837-45.
152. Arinami, T., *Analyses of the associations between the genes of 22q11 deletion syndrome and schizophrenia.* J Hum Genet, 2006. **51**(12): p. 1037-45.
153. Liu, H., et al., *Genetic variation in the 22q11 locus and susceptibility to schizophrenia.* Proc Natl Acad Sci U S A, 2002. **99**(26): p. 16859-64.
154. Mukai, J., et al., *Evidence that the gene encoding ZDHHC8 contributes to the risk of schizophrenia.* Nat Genet, 2004. **36**(7): p. 725-31.
155. Demily, C., et al., *ZDHHC8 single nucleotide polymorphism rs175174 is not associated with psychiatric features of the 22q11 deletion syndrome or schizophrenia.* Psychiatr Genet, 2007. **17**(5): p. 311-2.
156. Glaser, B., et al., *Analysis of ProDH, COMT and ZDHHC8 risk variants does not support individual or interactive effects on schizophrenia susceptibility.* Schizophr Res, 2006. **87**(1-3): p. 21-7.

157. Faul, T., et al., *ZDHHC8 as a candidate gene for schizophrenia: analysis of a putative functional intronic marker in case-control and family-based association studies.* BMC Psychiatry, 2005. **5**: p. 35.
158. Otani, K., et al., *The ZDHHC8 gene did not associate with bipolar disorder or schizophrenia.* Neurosci Lett, 2005. **390**(3): p. 166-70.
159. Glaser, B., et al., *No association between the putative functional ZDHHC8 single nucleotide polymorphism rs175174 and schizophrenia in large European samples.* Biol Psychiatry, 2005. **58**(1): p. 78-80.
160. Saito, S., et al., *No association was found between a functional SNP in ZDHHC8 and schizophrenia in a Japanese case-control population.* Neurosci Lett, 2005. **374**(1): p. 21-4.
161. Ambros, V., *The functions of animal microRNAs.* Nature, 2004. **431**(7006): p. 350-5.
162. Bracht, J., et al., *Trans-splicing and polyadenylation of let-7 microRNA primary transcripts.* Rna, 2004. **10**(10): p. 1586-94.
163. Denli, A.M., et al., *Processing of primary microRNAs by the Microprocessor complex.* Nature, 2004. **432**(7014): p. 231-5.
164. Gregory, R.I., T.P. Chendrimada, and R. Shiekhattar, *MicroRNA biogenesis: isolation and characterization of the microprocessor complex.* Methods Mol Biol, 2006. **342**: p. 33-47.
165. Gregory, R.I., et al., *The Microprocessor complex mediates the genesis of microRNAs.* Nature, 2004. **432**(7014): p. 235-40.
166. Lee, Y., et al., *The nuclear RNase III Drosha initiates microRNA processing.* Nature, 2003. **425**(6956): p. 415-9.
167. Han, J., et al., *The Drosha-DGCR8 complex in primary microRNA processing.* Genes Dev, 2004. **18**(24): p. 3016-27.
168. Yeom, K.H., et al., *Characterization of DGCR8/Pasha, the essential cofactor for Drosha in primary miRNA processing.* Nucleic Acids Res, 2006. **34**(16): p. 4622-9.
169. Shiohama, A., et al., *Molecular cloning and expression analysis of a novel gene DGCR8 located in the DiGeorge syndrome chromosomal region.* Biochem Biophys Res Commun, 2003. **304**(1): p. 184-90.
170. Han, J., et al., *Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex.* Cell, 2006. **125**(5): p. 887-901.
171. Wang, Y., et al., *DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal.* Nat Genet, 2007. **39**(3): p. 380-5.
172. Zhao, Y., et al., *Dysregulation of cardiogenesis, cardiac conduction, and cell cycle in mice lacking miRNA-1-2.* Cell, 2007. **129**(2): p. 303-17.
173. Kim, J., et al., *Identification of many microRNAs that copurify with polyribosomes in mammalian neurons.* Proc Natl Acad Sci U S A, 2004. **101**(1): p. 360-5.
174. Perkins, D.O., et al., *microRNA expression in the prefrontal cortex of individuals with schizophrenia and schizoaffective disorder.* Genome Biol, 2007. **8**(2): p. R27.
175. Cloninger, C.R., et al., *Genome-wide search for schizophrenia susceptibility loci: the NIMH Genetics Initiative and Millennium Consortium.* Am J Med Genet, 1998. **81**(4): p. 275-81.
176. Faraone, S.V., et al., *Diagnostic accuracy and confusability analyses: an application to the Diagnostic Interview for Genetic Studies.* Psychol Med, 1996. **26**(2): p. 401-10.
177. Nurnberger, J.I., Jr., et al., *Diagnostic interview for genetic studies. Rationale, unique features, and training. NIMH Genetics Initiative.* Arch Gen Psychiatry, 1994. **51**(11): p. 849-59; discussion 863-4.
178. First, M.B., et al., *Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition (SCID-I/P).* 2002, New York: Biometrics Research, New York State Psychiatric Institute.
179. Kendler, K.S., *Structured Interview for Schizotypal Symptoms (SISS version 1.5)*. 1989, Richmond, Virginia: Department of Psychiatry, Virginia Medical College.
180. Kendler, K.S., J.A. Lieberman, and D. Walsh, *The Structured Interview for Schizotypy (SIS): a preliminary report.* Schizophr Bull, 1989. **15**(4): p. 559-571.
181. Mannuzza, S., et al., *Family Informant Schedule and Criteria (FISC)*. 1985, New York: Anxiety Disorders Clinic, New York State Psychiatric Institute.
182. Kendler, K.S., et al., *A family history study of schizophrenia-related personality disorders.* Am J Psychiatry, 1984. **141**(3): p. 424-427.

183. Kendler, K.S., C.C. Masterson, and K.L. Davis, *Psychiatric illness in first-degree relatives of patients with paranoid psychosis, schizophrenia and medical illness.* Br J Psychiatry, 1985. **147**: p. 524-31.
184. Leckman, J.F., et al., *Best estimate of lifetime psychiatric diagnosis: a methodological study.* Arch Gen Psychiatry, 1982. **39**(8): p. 879-83.
185. Carey, G. and Gottesman, II, *Reliability and validity in binary ratings: areas of common misunderstanding in diagnosis and symptom ratings.* Arch Gen Psychiatry, 1978. **35**(12): p. 1454-9.
186. Whitty, P., et al., *Diagnostic stability four years after a first episode of psychosis.* Psychiatr Serv, 2005. **56**(9): p. 1084-8.
187. Chen, Y.R., A.C. Swann, and B.A. Johnson, *Stability of diagnosis in bipolar disorder.* J Nerv Ment Dis, 1998. **186**(1): p. 17-23.
188. Schwartz, J.E., et al., *Congruence of diagnoses 2 years after a first-admission diagnosis of psychosis.* Arch Gen Psychiatry, 2000. **57**(6): p. 593-600.
189. Deutsch, C.K., et al., *Skewing of the brain midline in schizophrenia.* Neuroreport, 2000. **11**(18): p. 3985-8.
190. Deutsch, C.K. and L.G. Farkas, *Quantitative methods of dysmorphology diagnosis*, in *Anthropometry of the head and face* L.G. Farkas, Editor. 1994, Raven: New York. p. 151-158.
191. Wang, H.Y., et al., *A genotyping system capable of simultaneously analyzing >1000 single nucleotide polymorphisms in a haploid genome.* Genome Res, 2005. **15**(2): p. 276-83.
192. Yang, X., et al., *The posterior probability of linkage allowing for linkage disequilibrium and a new estimate of disequilibrium between a trait and a marker.* Hum Hered, 2005. **59**(4): p. 210-9.
193. Vieland, V.J., *Bayesian linkage analysis, or: how I learned to stop worrying and love the posterior probability of linkage.* Am J Hum Genet, 1998. **63**(4): p. 947-54.
194. Vieland, V.J., K. Wang, and J. Huang, *Power to detect linkage based on multiple sets of data in the presence of locus heterogeneity: comparative evaluation of model-based linkage methods for affected sib pair data.* Hum Hered, 2001. **51**(4): p. 199-208.
195. Greenberg, D.A., et al., *Power, mode of inheritance, and type I error in LOD scores and affecteds-only methods: reply to Kruglyak.* Am J Hum Genet, 1998. **62**(1): p. 202-4.
196. Vieland, V.J., D.A. Grenberg, and S.E. Hodge, *Adequacy of single-locus approximations for linkage analyses of oligogenic traits: extension to multigenerational pedigree structures.* Hum Hered, 1993. **43**: p. 329-336.
197. Matthysse, S. and J. Parnas, *Extending the phenotype of schizophrenia: implications for linkage analysis.* J Psychiatr Res, 1992. **26**(4): p. 329-44.
198. Papolos, D.F., et al., *Bipolar spectrum disorders in patients diagnosed with velo-cardio-facial syndrome: does a hemizygous deletion of chromosome 22q11 result in bipolar affective disorder?* Am J Psychiatry, 1996. **153**(12): p. 1541-7.
199. Perkins, D.O., C. Jeffries, and P. Sullivan, *Expanding the 'central dogma': the regulatory role of nonprotein coding genes and implications for the genetic liability to schizophrenia.* Mol Psychiatry, 2005. **10**(1): p. 69-78.
200. Lewis, D.A. and P. Levitt, *Schizophrenia as a disorder of neurodevelopment.* Annu Rev Neurosci, 2002. **25**: p. 409-32.
201. Selemon, L.D., et al., *Amphetamine sensitization alters dendritic morphology in prefrontal cortical pyramidal neurons in the non-human primate.* Neuropsychopharmacology, 2007. **32**(4): p. 919-31.
202. Harrison, P.J. and D.R. Weinberger, *Schizophrenia genes, gene expression, and neuropathology: on the matter of their convergence.* Mol Psychiatry, 2005. **10**(1): p. 40-68; image 5.
203. Vandesompele, J., et al., *Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes.* Genome Biol, 2002. **3**(7): p. RESEARCH0034.
204. Livak, K.J. and T.D. Schmittgen, *Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method.* Methods, 2001. **25**(4): p. 402-8.
205. Tusher, V.G., R. Tibshirani, and G. Chu, *Significance analysis of microarrays applied to the ionizing radiation response.* Proc Natl Acad Sci U S A, 2001. **98**(9): p. 5116-21.
206. Chen, C., et al., *Real-time quantification of microRNAs by stem-loop RT-PCR.* Nucleic Acids Res, 2005. **33**(20): p. e179.

207. Kosik, K.S. and A.M. Krichevsky, *The Elegance of the MicroRNAs: A Neuronal Perspective.* Neuron, 2005. **47**(6): p. 779-82.

CURRICULUM VITAE: SHANNON E. BRUSE

4214 State Road 73
Deerfield, WI 53531
Phone (h) 608-762-2802
Phone (c) 908-329-4469
bruse@biology.rutgers.edu

CURRENT POSITION

Senior Scientist
Product Development
Mirus Bio Corporation
Madison, Wisconsin

EDUCATION

2002-2007	Ph.D. Microbiology and Molecular Genetics (GPA 3.985)	Rutgers University
1994-1998	B.A. Biology (GPA 3.76, graduated Highest Distinction)	University of Iowa
1986-1989	HS Diploma	Dubuque High School

PROFESSIONAL EXPERIENCE

January 2004- present Graduate Student, Genetics
Rutgers University, Piscataway, NJ

Shannon worked towards his PhD in Microbiology and Molecular Genetics under the mentorship of Dr. Linda Brzustowicz. His dissertation work has focused on three areas: (1) development of an inexpensive multiplexed genotyping method (2) application of this method to family-based candidate gene studies of major neuropsychiatric disorders (3) investigation of possible role of microRNAs in schizophrenia through candidate gene and expression studies. Shannon developed a medium throughput, inexpensive SNP typing method based on the Luminex platform (Bruse et al, in preparation), and this method was used to identify a major gene involved in autism (Benayed et al, 2005) and was used to identify a microRNA processing gene contributing to the risk for schizophrenia (Bruse et al, in preparation). Shannon has also performed extensive TaqMan and Panomics based expression studies of both microRNAs and mRNAs in post-mortem brain samples from patients with schizophrenia and bipolar disorder, potentially identifying a microRNA misexpressed in both schizophrenic and bipolar samples. Shannon contributed significantly to grant application that was successfully funded to continue work related to microRNAs and schizophrenia.

Methods/skills acquired: Design and implementation of family-based genetic association studies, analysis of family-based genetic association studies, design and implementation of post-mortem expression studies using TaqMan and Luminex-based expression platforms, operation of liquid handling stations (Biomek and Janus), Luminex-based SNP typing, highly multiplexed PCR.

April 2002-January 2004 Graduate Student, Genetic Epidemiology

Case Western Reserve University (CWRU), Cleveland, OH

Shannon began his graduate studies in Genetic Epidemiology at CWRU (later transferring to Rutgers due to a job opportunity for his wife). While at CWRU, Shannon worked on the human genetics of HIV infection, investigating the impact of polymorphisms in the HIV co-receptor, CCR5. In vivo experiments demonstrated that promoter polymorphisms known to be associated with rate of progression to AIDS indeed altered the expression of CCR5 and impacted HIV propagation (Salkowitz et al, 2003). Further experiments in a model system demonstrated that a synthetic CCR5 antagonist (PSC-RANTES) could effectively block HIV transmission regardless of CCR5 polymorphisms (Kawamura et al, 2004). Additional work included a case-control genetic association study of CCR5 polymorphisms and perinatal transmission of HIV. Unpublished work included identification of rhesus macaque CCR5 polymorphisms, to help identify a genetic basis of drug resistance in this important model organism.

Methods/skills acquired: Design, implementation, and analysis of case/control genetic associations studies, use of statistical packages for data analysis (SAS, S-Plus, R).

1998-2001 Research Assistant II, Department of Internal Medicine
University of Iowa, Iowa City, Iowa

Shannon worked as a research assistant in a molecular parasitology lab, performiing RNAi experiments to generate conditional genetic knockouts in *Trypanosoma brucei* (LaCount et al, 2000), the parasite that causes African sleeping sickness. At the time, it was notoriously difficult to generate knockouts to investigate gene function. Shannon's work contributed to establishment of the RNAi knockout system in *T. brucei*, as well as investigation of the function of the FLA1 gene using this method (LaCount et al, 2000). In addition to work on *T. brucei*, he participated in research aimed at developing a DNA vaccine for *Leishmania chagasi*.

Methods/skills acquired: DNA cloning, microsatellite typing, SSCP-based SNP discovery, recombinant protein expression and purification, prokaryotic cell transfection and cell culture, use of molecular biology software packages.

1996-1998 Research Assistant I, Department of Internal Medicine
University of Iowa, Iowa City, Iowa

Shannon worked in two labs during this time period, one lab focusing on cardiovascular physiology/heart disease and the other a yeast lab focused on vesicle trafficking. He worked on developing a semi-quantitative RT-PCR for the measurement of iNOS expression. Additionally, work was performed on a yeast genetic screen to identify genes involved in vesicular trafficking.

Methods/skills acquired: Western blot, immunoprecipitation, pulse-chase, FACS analysis, yeast cell transfection, RT-PCR, protein purification, biochemical assays (Griess, ELISA, B-gal, and Bradford), SDS-PAGE (silver and Coomassie blue staining), agarose and acrylamide gels for RNA and DNA, RNA and DNA purification, restriction enzyme analysis of plasmid constructs, basic cell culture, and limited small animal surgery.

1995 summer Lab Assistant, Department of Microbiology
University of Iowa, Iowa City, Iowa

A summer internship was spent in a lab studying phage virus. Basic laboratory practices were learned.

MENTORING AND TEACHING

Spring & Fall 2006 Mentored Andrew Hoffman (undergraduate student)
Spring 2007 Project: Candidate gene study of NTNG1 and schizophrenia
Dept. of Genetics, Rutgers University, Piscataway, NJ

Spring 2006	Mentored Michael Moreau (undergraduate student) Project: Candidate gene study of SNAP25 and schizophrenia Dept. of Genetics, Rutgers University, Piscataway, NJ
Fall 2005	Teaching Assistant Human Genetics course (graduate level) Dept. of Genetics, Rutgers University, Piscataway, NJ
Spring & Fall 2005	Mentored Lili Buzaki (undergraduate student) Project: Extraction of small RNAs from human post-mortem brain tissue Dept. of Genetics, Rutgers University, Piscataway, NJ
Spring 2003	Mentored Nishant Joshi (dental student) Project: Identification of single nucleotide polymorphisms in beta-defensin Department of Global Health, CWRU, Cleveland, OH

PUBLICATIONS AND PRESENTATIONS

Manuscripts:

Lacount DJ, **Bruse SE**, Hill KL, Donelson JE. Double-stranded RNA interference in Trypanosoma brucei using head-to-head promoters. Mol Biochem Parasitol. 2000 Nov;111(1):67-76.

Salkowitz JR, **Bruse SE**, Meyerson H, Valdez H, Mosier DE, Harding CV, Zimmerman PA, Lederman MM. CCR5 promoter polymorphism determines macrophage CCR5 density and magnitude of HIV-1 propagation in vitro. Clin Immunol. 2003 Sep; 108(3):234-40.

Kawamura T, **Bruse SE**, Abraha A, Sugaya M, Hartley O, Offord RE, Arts EJ, Zimmerman PA, Blauvelt A. PSC-RANTES blocks R5 human immunodeficiency virus infection of Langerhans cells isolated from individuals with a variety of CCR5 diplotypes. J Virol. 2004 Jul; 78(14):7602-9.

Benayed R, Gharani N, Rossman I, Mancuso V, Lazar G, Kamdar S, **Bruse SE**, Tischfield S, Smith BJ, Zimmerman RA, Dicicco-Bloom E, Brzustowicz LM, Millonig JH. Support for the homeobox transcription factor gene ENGRAILED 2 as an autism spectrum disorder susceptibility locus. Am J Hum Genet. 2005 Nov; 77(5):851-68. Epub 2005 Oct 5.

Bruse SE, Moreau M, Azaro MA, Hoffman A, Xu B, Wratten N, Zimmerman R, Brzustowicz LM. An inexpensive bead-based oligonucleotide ligation assay for SNP genotyping. (In preparation)

Bruse SE, Moreau M, DiCicco-Bloom E, Buyske S, Brzustowicz LM. Candidate gene study of DGCR8 and microRNA expression profiling in post-mortem brain samples from patients with schizophrenia (In preparation)

Poster Presentations:

N. Joshi, **S.E. Bruse**, P.A. Zimmerman, A. Weinberg. Single nucleotide polymorphisms in the human beta defensin 3 gene. 2nd Annual Research Showcase, Case Western Reserve University. April 2, 2004.

S.E. Bruse, L. Buzaki, L.M. Brzustowicz. Multiplexed SNP Genotyping of Schizophrenia Candidate Genes. 2nd Annual Planet XMap Symposium, Austin, Texas. April 25-27, 2005.

S.E. Bruse, M.A. Azaro, I.V. Tereschenko, L.A. Goff, R.P. Hart, L. M. Brzustowicz. Investigation of DGCR8 and the miRNA pathway in schizophrenia. American Society of Human Genetics, Salt Lake City, Utah. October 25-29, 2005.

I.V. Tereschenko, L. Serrano, C.L. Ayala, **S.E. Bruse**, J.A. Tischfield. Major regions of mitotic recombination of heterozygous mice fibroblasts located on chromosome 8. American Society of Human Genetics, Salt Lake City, Utah. October 25-29, 2005.

N. S. Wratten, **S. E. Bruse**, M.A. Azaro, J. Simone, J. E. Hayter, L. M. Brzustowicz. Identifying schizophrenia-associated non-coding variants in the CAPON gene. American Society of Human Genetics, Salt Lake City, Utah. October 25-29, 2005.

S.E. Bruse, N. McGregor, M.A. Azaro, B. Xu, L.M. Brzustowicz. An Inexpensive Bead-Based Oligonucleotide Ligation Assay for SNP Genotyping. American Society of Human Genetics, New Orleans, Louisiana. October 9-13, 2006.

S.E. Bruse, M. Moreau, E. DiCicco-Bloom, L.M. Brzustowicz. MicroRNA Expression in Post-Mortem Samples from Patients with Schizophrenia. Frontiers in Biopharmaceutical Sciences, Piscataway, NJ. 5:00-6:00 P.M., Life Sciences Building Auditorium. November 30, 2006.

S.E. Bruse, M. Moreau, E. DiCicco-Bloom, L.M. Brzustowicz. MicroRNA Expression in Post-Mortem Samples from Patients with Schizophrenia. 1st Annual Graduate Student Symposium, Piscataway, NJ. 4:00-5:00 P.M., Life Sciences Building Auditorium. February 23, 2006.

S.E. Bruse, M. Moreau, E. DiCicco-Bloom, M.A. Azaro, L.M. Brzustowicz. Schizophrenia and microRNAs: expression profiling and candidate gene study.CHI Third Annual MicroRNA Conference. World Trade Center, Boston, Massachusetts. March 29-30, 2007.

Oral Presentations:

S.E. Bruse. CCR5 Promoter Polymorphisms in HIV Infection. Graduate Student Seminar Series, Department of Genetic Epidemiology, Case Western Reserve University, Cleveland, OH. Sept 19, 2003.

S.E. Bruse. MicroRNA Expression Profiling of Post-Mortem Brain Samples from Individuals with Schizophrenia. Synergy Seminar Series hosted by Applied Biosystems, Bridgewater, NJ. December 5, 2006. (Invited talk)

S.E. Bruse. A Bead-Based Oligonucleotide Ligation Assay for Multiplexed SNP Genotyping. Thomas Jefferson University, Philadelphia, PA. March 7, 2007. (Invited talk)

S.E. Bruse. MicroRNA Expression Profiling of Post-Mortem Brain Samples from Individuals with Schizophrenia. Sanofi-Aventis, Bridgewater, NJ. May 5, 2007. (Invited talk)

S.E. Bruse. MicroRNAs and Schizophrenia. Annual Department of Genetics Faculty Retreat. Rutgers University, Piscataway, New Jersey. May 12, 2007.

S.E. Bruse. MicroRNAs and Schizophrenia. NY Psychiatric Genetics Summer Forum 2007. Cold Spring Harbor Laboratory, New York. June 27, 2007.

AWARDS

2005-2006	Predoctoral Fellowship "Molecular and developmental basis of mental health and aging" Competitive fellowship awarded through Dept. of Neuroscience and Cell Biology, Rutgers University

PROFESSIONAL MEMBERSHIPS AND AFFILIATIONS

2005-2007	American Society of Human Genetics