DISRUPTIONS TO THE MICRORNA REGULATORY PATHWAY ARE RESPONSIBLE FOR THE INSCREASED RATE OF SCHIZOPHRENIA IN INDIVIDUALS WITH 22q11.2 DELETION SYNDROME

by

WILLIAM MANLEY

A dissertation submitted to the

Graduate School-New Brunswick Rutgers,

The State University of New Jersey

And

The Graduate School of Biomedical Sciences

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Cell and Developmental Biology

Written under the direction of

Linda Brzustowicz, M.D.

And Approved by

_____

_____

_____

_____

New Brunswick, New Jersey

October, 2016

ABSTRACT OF THE DISSERTATION

Disruptions to the miRNA Regulatory Pathway are Responsible for the Increased Rate of Schizophrenia in Individuals with 22q11.2 Deletion Syndrome

By WILLIAM MANLEY

Dissertation Director:

Linda Brzustowicz, M.D.

Schizophrenia is a complex and poorly understood disease caused by the interplay of many environmental and genetic factors. It occurs in 1% of the general population and causes delusions, hallucinations, and disorganized behavior. The 22q11.2 deletion syndrome is a disorder that is caused by the microdeletion of part of chromosome 22 leading to a 25-fold greater chance of developing schizophrenia in affected individuals versus the general population. The missing 22q11.2 region contains the *DGCR8* gene, which is required for the initial step of miRNA biogenesis. In Dgcr8 knockout mice, the levels of specific miRNAs in the brain are reduced, and these animals show behavioral and cognitive deficits similar to schizophrenia. However, the 22q11.2 deletion itself is not directly the cause of schizophrenia, since 75% of individuals with this deletion do not develop the disease.

We hypothesize that the 22q11.2 deletion increases the risk of schizophrenia through alterations in miRNA regulatory networks via depletion of miRNAs which may serve as a protective buffer against the accumulation of deleterious mutations elsewhere in the genome. In healthy individuals, genes associated with schizophrenia may harbor regulatory mutations, but these mutations are not deleterious since the miRNA regulatory system acts as a buffer that maintains gene expression at the required levels. In 22q11.2 deletion syndrome, the same mutation may have pathogenic effects, since miRNA critical for compensating for these mutations are absent. We believe that these unchecked mutations result in the development of schizophrenia. These experiments attempt to characterize which miRNAs are altered in 22q11.2 deletion syndrome. Based on which miRNAs exhibit altered expression, online resources and predictive software can now be used to determine which genes may have altered expression as a result of disruptions to the miRNA regulatory pathway and these genes could potentially be the targets for diagnostic assays or pharmacological intervention. Lastly, experiments using cellular models and luciferase reporter constructs have been performed to validate how a schizophrenia associated SNP in a specific gene, H3F3B, alters miRNA regulation of one specific miRNA, miR-616. To investigate our hypothesis that the 22q11.2 deletion increases the risk of schizophrenia through alterations in miRNA regulatory networks via depletion of miRNAs which may serve as a protective buffer against the accumulation of deleterious mutations, the experiments described in the following chapters have been performed

Acknowledgment

First and foremost, I would like to thank Linda Brzustowicz, M.D. for allowing me to be a graduate student in her laboratory. I will always be grateful for her guidance over the years. I especially enjoyed the tremendous amount of freedom she allowed me to pursue new projects and the trust she placed in me to supervise a small team of researchers.

I would also like to thank the members of the committee, Jim Millonig, Ph.D., Bonnie Firestein, Ph.D., Steve Buyske, Ph.D., and Anne Bassett, M.D. FRCPC, who have all been extremely helpful in keeping the project focused with their many suggestions over the years. Drs Millonig and Firestein have both generously opened their labs up to me at key points in this research. I would also like to thank Dr. Buyske for all of his direct contributions to the statistical analysis included in this research.

I am also tremendously grateful for the guidance, technical support and training I was able to receive from the RUCDR. Jennifer Moore, Ph.D. and Navindra Sukhram have been absolutely essential to the development of stem cell projects performed as a part of this research.

This work would have not been possible without contributions from the Brzustowicz lab. Past and present members of the lab have created an excellent working environment that has been a pleasure to come to every day. Special thanks to past lab members Mike Moreau, Ph.D., Abby Hare-Harris, Ph.D., Carmen Ramierz, Ray Zimmerman, Tony Marketta and Kate Tobin, PhD as well as current lab members Marco Azaro, PhD, Gillian Davis, Judy Flax, PhD, Christine Gwin, Jared Hayter, Karen Law,

Vaidhy Mahaganapathy, Jaime Messenger, Brenda Patel, and Sherri Wilson for their contributions to this research.

A project like is would never have been possible without help performing the research. Fortunately, I have been lucky enough to have an extremely talented group of hard working undergraduates, graduate students, post-bacs, volunteers, and technicians over the years who have contributed in major ways to this research. Steve Siecinski, Sean Ryan, and Vanessa Coulibaly have contributed countless hours to this project and have been great colleagues to work with in the lab over the years. Additionally, the help of Nkem Ndubuizu, Varun Gupta, Larry Cheng, Nate Stauring, Mark Pinkerton, Yuanchao Zang, Kevin Lu, Priya Kantesaria, Gabby Kleyner, and Natalie Samper has been essential to this research.

Lastly on a more personal note, I would like to thank all of my family and friends who have helped directly or indirectly to complete this work. I would like to extend special thanks to my parents who encouraged me to continue my studies indefinitely and for their unending financial support. I would also like to extend special thanks to Patty for her unconditional encouragement, friendliness, and putting up with my constant need to make a "quick" trip to lab or otherwise scheduling life around the needs of my cells. Thanks you to all of my family and friends for coming to visit during my defense. It was nice to see everyone. Lastly, I would like to thank the Pittsburgh Penguins for providing the best distraction from writing I could ask for. Watching the Penguins win the 2016 Stanley Cup really made writing my thesis this past spring much more tolerable.

Table of Contents

List of Tables:

List of Figures:

General Background:

Schizophrenia:

Schizophrenia is a complex disorder that is believed to be caused by the combined interactions of environmental and genetic risk factors. The Diagnostic and Statistical Manual of Mental Disorders Fifth Edition (DSM-5) is currently the preferred system used for diagnosing schizophrenia in the United States. A detailed list of criteria allows the diagnosis of schizophrenia to be made when symptoms have existed for at least six months. Symptoms of schizophrenia include positive symptoms such as delusions, hallucinations, disorganized speech, and disorganized behavior. Delusions are false beliefs or a misinterpretation of actual experiences. These differ from hallucinations, which are profound distortions in a person's perception of reality, typically accompanied by a powerful sense of reality. Hallucinations may be a sensory experience in which a person can see, hear (most common in schizophrenia), smell, taste, or feel something that is not there (Buchanan and Carpenter).

There are also negative symptoms such as alogia (a lack of speech), avolition or lack of desire, flat affect, and social withdrawal. In order to make the diagnosis, at least two or more of the symptoms must be present for a minimum of one month and the overall disturbance must exist for at least six months. However, only one of the preceding symptoms is required if the person exhibits bizarre delusions or hallucinations with voices having a running commentary or of two separate voices carrying on a conversation. Another additional required symptom for a schizophrenia diagnosis is social or occupational dysfunction. Unlike mood disorders, a schizophrenia diagnosis does not require depression or manic symptoms (Castle et al.; Walker et al.)

Lastly, other medical conditions or common substance abuse problems that result in psychotic symptoms must be considered to avoid a false diagnosis of schizophrenia (Walker et al.) Treatment of schizophrenia focuses primarily on controlling the positive symptoms of the disease with antipsychotic drugs. One of the oldest theories relating to the onset of the disease was hyperactivity in the dopamine system, therefore many of these drugs are antagonists of dopamine D2 receptors (Meltzer and Stahl). Second generations drugs continue to target this pathway and have less side effects, as well as help treat some of the negative symptoms of schizophrenia (Tandon and Halbreich). The improved function of second generation schizophrenia drugs has been attributed both to a more moderate blocking of the D2 receptors and by acting on the 5-HT_{1A} receptors, which ultimately stimulate the release of dopamine in the prefrontal cortex (Meltzer and Sumiyoshi).

Historically, four subtypes of schizophrenia have been described: paranoid, disorganized, catatonic, and undifferentiated. The paranoid subtype is recognizable by a preoccupation with hallucinations and delusions, but without the disorganized behavior or speech. This subtype usually has the most favorable prognosis. The catatonic subtype is classified by postural and movement abnormalities, as well as echolalia, which is the automatic repetition of vocalizations made by another person. The disorganized type exhibits disorganized speech and behavior, yet fails to meet symptoms of the catatonic subtype. This subtype usually leads to the worst prognosis. The final class, the undifferentiated subtype, occurs when the patient meets the overall diagnosis criteria of schizophrenia but does not have all the required aspects of any of the previous subtypes (Walker et al.).

In addition to the clinical symptoms associated with schizophrenia, there are also severe deficits in cognitive functioning as a result of the disease. Schizophrenia causes performance deficits in both simple and complex cognitive tasks including secondary verbal memory, and working memory (Green et al.). Individuals with schizophrenia also have deficits in spatial memory, attention and executive functions such as reasoning and planning (Kuperberg and Heckers). These symptoms are likely the result of structural and functional brain abnormalities caused by schizophrenia. Schizophrenia results in enlarged lateral ventricles, and decreased frontal, temporal, and whole brain volume (Dennert and Andreasen; Lawrie and Abukmeil; Northoff et al.). Functional magnetic resonance imaging has more recently been used to measure brain activity in response to the presentation of stimuli in individuals with schizophrenia who were found to have reductions of activity in the thalamus, the prefrontal cortex, and the parietal lobe relative to healthy individuals (Braus et al.). In addition to these neuroanatomical changes in the brain observed after onset of symptoms, several alterations appear prior to onset of schizophrenia such as a reduction of gray matter in the cingulate gyrus and left parahippocampal, fusiform, orbitofrontal, and cerebellar cortices (Pantelis et al.).

Schizophrenia occurs throughout the world, affecting roughly 1% of the general population (Kulhara and Chakrabarti). The rate is strongly influenced by gender and living in an urban environment. Males and city dwellers have significantly higher risks of developing schizophrenia than females or individuals living in more rural environments (Saha et al.). In addition to an increased risk of developing the disease, males also have an earlier onset of schizophrenia than females. Men typically develop schizophrenia in their late teens or early twenties, while women have peak onset in their late twenties or

early thirties (Aleman, Kahn, and Selten). Although environmental factors do play a role in the risk of developing schizophrenia, the genetic component of the disease is strong with a high heritability estimated at 80% (McGuffin et al.).

Studies examining twins and extended family histories show that the risk for developing schizophrenia is elevated if an individual has a biological relative who also has the disorder. These studies also show that the genetically closer the relative with schizophrenia, the greater the risk that individual has of developing schizophrenia. For instance, the child of a parent with schizophrenia has a 13% chance of developing schizophrenia. This chance of developing schizophrenia jumps to 40% if both parents of an individual have schizophrenia themselves (Davis et al.; Vogel). A first-degree sibling with schizophrenia increases the likelihood of an individual developing schizophrenia 8 times (Keshavan et al.). Perhaps the closest relatives of all, monozygotic twins, have the highest concordance rate for developing schizophrenia with up to 50% of co-twins developing the disease (Beckmann, Franzek, and Stöber; Beckmann and Franzek). Schizophrenia is likely caused by several genetic factors since the patterns of transmission of the disease throughout families do not match the Mendelian pattern that would be expected if schizophrenia could be traced to a single locus (Harrison and Weinberger). A more likely explanation is that a series of genes is acting in concert along with environmental factors to cause the disease (Kato et al.).

22q11.2 Deletion Syndrome

In order to understand why there is a relationship between the development of schizophrenia and 22q11.2 DS (deletion syndrome), it is first necessary to discuss the clinical and genetic features of this microdeletion. 22q11.2 DS is also referred to as

DiGeorge syndrome, velocardiofacial syndrome (VCFS), conotruncal anomalies face syndrome, CATCH 22, and Sedlackova syndrome (R J Shprintzen, Goldberg, Lewin, et al.; Deborah a Driscoll et al.; Karayiorgou et al.; Shprintzen; Bassett; de la Chapelle et al.; Kelley et al.). These were all originally believed to be separate disorders, but were confirmed to all be 22q11.2 deletion syndrome using FISH in the 1990s (Scambler et al.; D a Driscoll et al.). In 90% of cases, 22q11.2 DS is caused by a 2.5-mega-base microdeletion of chromosome 22 at the q11.2 band (Shprintzen; Emanuel; Anne S. Bassett, Marshall, et al.). The remaining 10% of cases are caused by smaller nested deletions with the same proximal breakpoint, but a distal breakpoint that is only 1.5Mb away (Morrow et al.; Edelmann, Pandita, and Morrow; Shaikh et al.). The more common 2.5 Mb deletion includes 90 genes, 46 of which are protein coding. 41 of these genes have been found to be expressed in the human brain. These genes are highly conserved. 37 homologs were found in the zebrafish, 40 homologs were found in mouse, 22 homologs were found in the fruit fly, and 17 homologs were found in the in the worm. Additionally, 17 of these genes were conserved across all tested species (Guna, Butcher, and Bassett).

This disorder has a high prevalence in the general population, occurring in 1 out of every 4000 live births or .025% of the general population (Goodship et al.). In a study just limited to the United States, the 22q11.2 deletion was observed to occur in 1 out of every 2000 individuals making it the most commonly occurring genetic disorder (Shprintzen; Bassett; Costain, Mcdonald-Mcginn, and Bassett; Kaminsky et al.). 22q11.2 DS is most often the result of a spontaneous mutation, although it occurs with an autosomal dominant pattern of inheritance in 25% of cases (R J Shprintzen, Goldberg,

Young, et al.; Leana-Cox et al.; McDonald-McGinn et al.). The penetrance is 100% and cases exhibit extreme variation in expression. The 22q11.2 region is highly mutable, readily susceptible to rearrangement, and contains several large blocks of segmental duplications (Edelmann, Pandita, and Morrow; Shaikh et al.; Saitta et al.). Rearrangements most likely occur as the result of erroneous recombination during stage 1 of meiosis (Bailey et al.; Babcock et al.). At each end of the typical 3Mb deletion are identical sets of low copy repeats that suggest that the mechanism for the deletion is likely an interchromosomal recombination error (Edelmann, Pandita, and Morrow).

22q11.2 deletion syndrome causes a multitude of symptoms that affect nearly every organ or system, with the most prominent effects being cardiac defects, facial dysmorphism, thymic hypoplasia, hypocalcaemia, and cleft palate (Wilson et al.; A S Bassett et al.; McDonald-McGinn et al.). However, of the 180 reported behavioral and physical symptoms, no single one occurs in 100% of cases, and there has yet to be a case that has all or even most of the many symptoms of the deletion (Robinson et al.). Congenital heart disease though does occur in 70% of cases of 22q11.2 DS (Goldmuntz et al.) The severity of the symptoms has an expansive range as well, with some individuals having only the most minimal symptoms, while some have severe life-threatening conditions. Because there are so many symptoms associated with 22q11.2 DS, it may be very difficult to diagnose the disease based purely on clinical presentation. Any diagnosis must be confirmed by looking for the deletion of DNA at q11.2 of chromosome 22. Diagnostic methods are virtually 100% accurate and make use of FISH (Fluorescence In Situ Hybridization) which uses DNA probes to determine if a specific region in the genome is present and quantitative real-time PCR that uses a specifically

designed probe to determine copy number (Robert J Shprintzen et al.; Tomita-Mitchell et al.).

Individuals with 22q11.2 DS have a clearly defined deletion that is a major risk factor for the development of schizophrenia and the greatest molecular risk factor for the development of schizophrenia (Costain and Bassett; Fung et al.). This microdeletion leads to a 25-fold greater chance of developing schizophrenia in affected individuals versus the general population and 1% of individuals with schizophrenia have the 22q11.2 deletion (Sugama et al.; Anne S. Bassett, Costain, et al.; Fung et al.). Therefore, important genomic factors in the deleted region must have some important impact on how the brain develops or functions.

A cluster of seven validated and putative miRNAs in the 22q11.2 deletion region (the highly characterized miR-185, miR-649, miR-1286, miR-1306, miR-3618, miR-4761, and miR-6816) were identified and the targets of these miRNAs were predicted (Stark et al.; Sellier et al.; Zhao et al.). The potential targets of these seven miRNAs were found to be important genes for neuronal processes and schizophrenia candidate genes. The development of schizophrenia could be the result of the reduced expression of these miRNAs as a result of the 22q11.2 deletion instead of or in addition to the *DGCR8* deletion itself (Merico et al.).

22q11.2 DS Alters the Brain

22q11.2 deletion syndrome may cause a number of structural changes in the brain that may increase an individual's susceptibility to develop schizophrenia. Significant reductions in brain volume and myelinated axons in the frontal lobe were observed as

well as an increased volume of cerebral ventricles in the brains of individuals with 22q11.2 DS versus individuals without 22q11.2 DS. The dorsolateral prefrontal cortex, an area known to be involved in schizophrenia, also has reduced overall volume in individuals with 22q11.2 DS (R J Shprintzen, Higgins, et al.; N J Beveridge et al.; Whalley et al.). However, individuals dually affected with 22q11.2 DS and schizophrenia exhibit several abnormal changes in the structure of the brain versus healthy individuals or individuals with 22q11.2 DS without schizophrenia. The left superior temporal gyrus and bilateral gray matter volume in the temporal lobes and superior temporal gyri have been shown to be reduced in individuals with schizophrenia and 22q11.2 DS compared to individuals with 22q11.2 DS alone (Chow et al.).

The deletion also causes functional changes that are also seen in schizophrenia. The 22q11.2 deletion leads to deficits in cognitive tasks dependent on proper prefrontal cortex and hippocampus function such as short term memory and executive function impairments, in individuals with and without schizophrenia (Kiley-Brabeck and Sobin) . Functional magnetic resonance imaging studies of individuals with 22q11.2 have shown that Broca's area, the anterior cingulate, and dorsolateral prefrontal cortex exhibit hypoactivation relative to healthy controls in tests of working memory (Kates et al.). These functional changes in the brain could be signs of schizophrenia predisposition in individuals with 22q11.2 DS (van Amelsvoort et al.).

Importance of *DGCR8*:

The 22q11.2 deletion causes these functional and structural changes to the brain as a result of the disrupted expression of genes in the deleted chromosomal region. Table 1 shows a list of some of the most widely studied genes on the 22q11.2 deletion and their

function. One of the 35 genes in the deleted region is *DGCR8*, which is part of the microprocessor complex required for miRNA production (Dunham et al.). miRNAs are able to bind to specific regions of the 3' UTR of coding mRNAs thereby decreasing expression of their protein product (Bartel). This is relevent since miRNAs are also known to have important roles in brain development and the development of schizophrenia and other psychiatric disorders (Lewis, Burge, and Bartel; Hansen et al.; Bartel). A list of miRNAs previously implicated as important in the development of schizophrenia is shown in Table 2. These miRNAs have all been identified using postmortem brain studies. Additionally, the miRNAs let-7g, miR-17, miR-92a, miR 181b, miR-219-2-3-p, miR 346, and miR-1308 have all also been identified as been dysregulated in individuals with schizophrenia as a result of serum studies (Shi et al.), while miR-137 has also been implicated as a risk allele associated with schizophrenia by the Psychiatric GWAS Consortium (2011).

The list of miRNAs present in the brain with important regulatory functions related to schizophrenia is continuously growing and becoming more complex. However, these studies do not all agree on which miRNAs are dysregulated as a result of schizophrenia. This discrepancy is likely the result of differences in miRNA expression caused by the inherent limitations associated with studies of postmortem samples. miRNA expression has been shown to change based on the age and gender of the postmortem samples Postmortem brains are also generally not dissected at the same time and subject to the same storage conditions. Variables in storage conditions and preparation of samples introduce unwanted variability to the expression profiles of miRNAs in these studies. Samples must be properly stored after death to minimize miRNA degradation. The total

sample storage time, brain pH at death and postmortem interval affect miRNA levels in postmortem tissue. Alcohol, antipsychotic, or antidepressant usage at the time of death, also impact miRNA expression (Moreau et al.). These variables can be problematic since they are outside the control of the researcher and could confound the profiling of miRNAs present in the brain.

miRNAs found in the prefrontal cortex are currently believed to be candidate genes for schizophrenia (Perkins et al.). Postmortem brain tissue samples from individuals with schizophrenia were found to have altered expression of these miRNAs (Wong et al.; Banigan et al.; Xu et al.; Miller et al.; Guella et al.; Natalie J. Beveridge et al.; Morishita et al.; Mor et al.; Moreau et al.). These miRNAs are believed to elevate the risk of developing schizophrenia by mechanisms such as increasing the chances of developing abnormal cytoarchitecture and a reduction in both dendritic spine density and the number of synapses (Davidsson et al.; Arnold, Talbot, and Hahn). Since the 22q11.2 deletion includes the miRNA regulatory gene *DGCR8* and miRNAs have been shown to have reduced expression in schizophrenia, it is conceivable that the deletion impacts schizophrenia via the miRNA regulatory pathway.

Rare copy number variants (CNVS) have also been shown to contribute to the development of neurodevelopmental disorders and schizophrenia (Cook and Scherer; Vorstman et al.; Klaassen et al.). In individuals with schizophrenia, an increase in CNVs that overlap with miRNAs has been observed. This significant increase in the presence of a rare CNV that overlapped with a miRNA was detected in individuals with schizophrenia using high-resolution genome-wide microarrays. Additionally the predicted target genes of CNVs that were found to overlap with miRNAs were genes that

have been previously identified as being important for neurodevelopment (Warnica et al.).

Table 1:

Gene	Base Pair Position on chr22:	Function
DGCR8	20080232-20111871	Potential Candidate Gene for Schizophrenia
PRODH	18912777-18936553	Defects in this gene are a cause of hyperprolinemia type 1 and possibly susceptibility to schizophrenia
DGCR2	19036286-19122454	Plays a role in neural crest cell migration
STK22B	19130808-19132623	Involved in the late stages of spermatogenesis during the reconstruction of the cytoplasm
DGCR14	19130279-19144684	Component of C complex spliceosomes
GSCL	19,148,576-19,150,283	Required for craniofacial development
SLC25A1	19175575-19178830	Involved in citrate-H(+)/malate exchange
CLTCL1	19179473-19291714	Major protein of the polyhedral coat of coated pits and vesicles
HIRA	19330698-19431733	Required for cell cycle exit
MRPL40	19431902-19436075	Involved with protein synthesis in the mitochondrion
UFD1L	19449910-19479215\	Necessary for the degradation of ubiquitinated proteins
CDC45L	19479459-19520612	Required to the initiation of DNA replication
CLDN5	19523027-19524400	Encodes a member of the claudin family
SEPT5	19714464-19723319	Filament-forming cytoskeletal GTPase
GP1BB	19722945-19724771	Part of the GPib-V-IX system that constitutes the receptor for von Willebrand factor
TBX1	19756703-19767334	Transcriptional regulator involved in developmental processes
GNB1L	19783224-19854939	Highly expressed in the heart / contributes to heart defects
TXNRD2	19875522-19941992	Implicated in the defenses against oxidative stress
COMT	19941607-19969975	Localized to postsynaptic neurons where it degrades neurotransmitters following their release

ARVCF	19969896-20016808	Involved in protein-protein interactions at adherins junctions
DGCR8	20080232-20111871	DGCR8 dimer binds pri-miRNAS and is active in triggering pri-miRNA damage
HTF9C	20111875-20117392	Involved in nucleic acid metabolism
RANBP1	20117548-20127181	Inhibits GTP exchange on Ran
ZDHHC8	20131807-20148007	Putative palmitoyltransferase involved in glutamatergic transmission
RTN4R	20241415-20268531	Mediates axonal growth inhibition

Table 1 shows a list of genes present in the human 22q11.2 deletion and some of their important functions

Table 2: miRNA Important for Neuropsychiatric Disorders Identified via Postmortem Study

miRNA	Reference	Changes in miRNA as a result of schizophrenia
let-7e	Beveridge, et al. 2010	Increased expression in superior temporal gyrus in schizophrenia
miR-7	Kim et al. 2010	Elevated expression in dorsolateral prefrontal cortex samples with schizophrenia
	Perkins, Jeffries et al. 2007	Reduced expression levels in prefrontal cortex samples with schizophrenia
miR-15a	Moreau, Bruse, et al. 2011	Reduced expression in prefrontal cortex samples with bipolar disorder
	Beveridge, et al. 2010	Increased expression in superior temporal gyrus in schizophrenia
miR-15b	Beveridge, et al. 2010	Increased expression in superior temporal gyrus in schizophrenia
miR-16	Beveridge, et al. 2010	Increased expression in superior temporal gyrus in schizophrenia
miR-17	Wong, Duncan et al. 2013	Upregulated in dorsolateral prefrontal cortex in schizophrenia
miR-19a	Beveridge, et al. 2010	Increased expression in superior temporal gyrus in schizophrenia
miR-20a	Beveridge, et al. 2010	Increased expression in superior temporal gyrus in schizophrenia
miR-22	Moreau, Bruse, et al. 2011	Reduced expression in prefrontal cortex samples with bipolar disorder
miR-22*	Kim et al. 2010	Elevated expression in dorsolateral prefrontal cortex samples with bipolar disorder
miR-24	Perkins, Jeffries et al. 2007	Reduced expression levels in prefrontal cortex samples with schizophrenia
miR-26b	Beveridge, et al. 2010	Increased expression in superior temporal gyrus in schizophrenia
	Perkins, Jeffries et al. 2007	Reduced expression levels in prefrontal cortex samples with schizophrenia
miR-27b	Moreau, Bruse, et al. 2011	Reduced expression in prefrontal cortex samples with bipolar disorder
miR-29a	Kim et al. 2010	Reduced expression in dorsolateral prefrontal cortex samples with bipolar disorder

	Perkins, Jeffries et al. 2007	Reduced expression levels in prefrontal cortex samples with schizophrenia
miR-29b	Perkins, Jeffries et al. 2007	Reduced expression levels in prefrontal cortex samples with schizophrenia
miR-29c	Banigan, Kao et al. 2013	Elevated levels in the exosomes of prefrontal cortex samples with bipolar disorder
	Perkins, Jeffries et al. 2007	Reduced expression levels in postmortem samples with schizophrenia
miR-30a-5p	Perkins, Jeffries et al. 2007	Reduced expression levels in prefrontal cortex samples with schizophrenia
miR-30b	Perkins, Jeffries et al. 2007	Reduced expression levels in prefrontal cortex samples with schizophrenia
miR-30d	Perkins, Jeffries et al. 2007	Reduced expression levels in prefrontal cortex samples with schizophrenia
miR-30e	Xu, Li et al 2010	Contains SNPs associated with schizophrenia
	Perkins, Jeffries et al. 2007	Reduced expression levels in prefrontal cortex samples with schizophrenia
miR-32	Kim et al. 2010	Reduced expression in dorsolateral prefrontal cortex samples with bipolar disorder
miR-33	Moreau, Bruse, et al. 2011	Reduced expression in prefrontal cortex samples with bipolar disorder
miR-34a	Kim et al. 2010	Elevated expression in dorsolateral prefrontal cortex samples with schizophrenia
miR-92	Perkins, Jeffries et al. 2007	Reduced expression levels in prefrontal cortex samples with schizophrenia
miR-106b	Moreau, Bruse, et al. 2011	Reduced expression in prefrontal cortex samples with bipolar disorder
miR-107	Beveridge, et al. 2010	Increased expression in superior temporal gyrus in schizophrenia
miR-132	Miller, Zeier et al. 2012	Downregulated in dorsolateral prefrontal cortex in schizophrenia
	Kim et al 2010	Elevated expression in dorsolateral prefrontal cortex samples with schizophrenia
miR-132*	Kim et al 2010	Elevated expression in dorsolateral prefrontal cortex samples with schizophrenia
miR-133b	Kim et al. 2010	Elevated expression in dorsolateral prefrontal cortex samples with bipolar disorder

miR-137	Guella et al. 2012	Reduced levels of miR-137 in dorsolateral prefrontal cortex with rs1625579 genotype
miR-138	Moreau, Bruse, et al. 2011	Reduced expression in prefrontal cortex samples with bipolar disorder
miR-140-3p	Kim et al. 2010	Reduced expression in dorsolateral prefrontal cortex samples with bipolar disorder
miR-145	Kim et al. 2010	Elevated expression in dorsolateral prefrontal cortex samples with bipolar disorder
miR-145*	Kim et al. 2010	Elevated expression in dorsolateral prefrontal cortex samples with bipolar disorder
miR-148b	Moreau, Bruse, et al. 2011	Reduced expression in prefrontal cortex samples with bipolar disorder
miR-151	Moreau, Bruse, et al. 2011	Reduced expression in prefrontal cortex samples with bipolar disorder
miR-154*	Kim et al. 2010	Elevated expression in dorsolateral prefrontal cortex samples with schizophrenia
	Kim et al. 2010	Elevated expression in dorsolateral prefrontal cortex samples with bipolar disorder
miR-181a	Moreau, Bruse, et al. 2011	Reduced expression in prefrontal cortex samples with bipolar disorder
miR-181b	Beveridge, et al. 2010	Increased expression in superior temporal gyrus in schizophrenia
	Beveridge, Tooney, et al. 2008	Upregulated in postmortem cortical grey matter from the superior temporal gyrus
miR-186	Moreau, Bruse, et al. 2011	Reduced expression in prefrontal cortex samples with bipolar disorder
miR-190	Moreau, Bruse, et al. 2011	Reduced expression in prefrontal cortex samples with bipolar disorder
miR-192	Moreau, Bruse, et al. 2011	Reduced expression in prefrontal cortex samples with bipolar disorder
miR-193a	Moreau, Bruse, et al. 2011	Reduced expression in prefrontal cortex samples with bipolar disorder
miR-193b	Moreau, Bruse, et al. 2011	Reduced expression in prefrontal cortex samples with bipolar disorder
		Increased expression in prefrontal cortex samples with schizophrenia
miR-195	Shi, Du et al. 2012	Reduced expression in serum from individuals with schizophrenia
	Beveridge, et al. 2010	Increased expression in superior temporal gyrus in schizophrenia

	Perkins, Jeffries et al. 2007	Reduced expression levels in prefrontal cortex samples with schizophrenia
miR-210	Moreau, Bruse, et al. 2011	Reduced expression in prefrontal cortex samples with bipolar disorder
miR-212	Kim et al. 2010	Elevated expression in dorsolateral prefrontal cortex samples with schizophrenia
	Perkins, Jeffries et al. 2007	Reduced expression levels in prefrontal cortex samples with schizophrenia
miR-300	Moreau, Bruse, et al. 2011	Reduced expression in prefrontal cortex samples with bipolar disorder
miR-301	Moreau, Bruse, et al. 2011	Reduced expression in prefrontal cortex samples with bipolar disorder
miR-324.3p	Moreau, Bruse, et al. 2011	Reduced expression in prefrontal cortex samples with bipolar disorder
miR-338	Moreau, Bruse, et al. 2011	Reduced expression in prefrontal cortex samples with bipolar disorder
miR-339	Moreau, Bruse, et al. 2011	Reduced expression in prefrontal cortex samples with bipolar disorder
miR-382	Mor, E., S. I. Kano, et al. 2013	Elevated in cultures derived from individuals with schizophrenia
miR-425	Moreau, Bruse, et al. 2011	Reduced expression in prefrontal cortex samples with bipolar disorder
miR-454*	Kim et al. 2010	Reduced expression in dorsolateral prefrontal cortex samples with bipolar disorder
miR-497	Banigan, Kao et al. 2013	Elevated in the exosomes of prefrontal cortex samples with bipolar disorder
miR-504	Kim et al. 2010	Elevated expression in dorsolateral prefrontal cortex samples with bipolar disorder
miR-520c-3p	Kim et al. 2010	Reduced expression in dorsolateral prefrontal cortex samples with bipolar disorder
miR-544	Kim et al. 2010	Elevated expression in dorsolateral prefrontal cortex samples with schizophrenia
miR-545	Moreau, Bruse, et al. 2011	Reduced expression in prefrontal cortex samples with bipolar disorder
miR-573	Kim et al. 2010	Reduced expression in dorsolateral prefrontal cortex samples with bipolar disorder
miR-639	Moreau, Bruse, et al. 2011	Reduced expression in prefrontal cortex samples with bipolar disorder

miR-767-5p	Kim et al. 2010	Reduced expression in dorsolateral prefrontal cortex samples with bipolar disorder
miR-874	Kim et al. 2010	Reduced expression in dorsolateral prefrontal cortex samples with bipolar disorder
miR-889	Kim et al. 2010	Elevated expression in dorsolateral prefrontal cortex samples with bipolar disorder

Table 2 shows a list of miRNAs that have previously been shown to be important for the development of neuropsychiatric disorders. These studies have all used postmortem brains.

miRNAs have a complex production process. It is our hypotheses that the 22q deletion causes a disruption to this process which increases an individual's risk for developing schizophrenia. Normally, processing of miRNA requires several steps that are used to transform primary miRNA transcripts (pri-miRNA) into active and mature miRNAs (Kim). Genes for miRNA are transcribed by RNA polymerase II to synthesize a pri-miRNA transcript. This leads to pri-miRNA forming during the initial step of miRNA biosynthesis as shown in Figure 1, when RNA polymerase II-dependent transcription occurs (Y Lee et al.). The miRNA biogenesis pathway converts pri-miRNA to mature miRNA through two main steps that involve ribonuclease reactions (Yoontae Lee et al.). Pri-miRNA is located in the nucleus where it is cleaved into a hairpin structure by the DROSHA complex to produce a hairpin structure that is called pre-miRNA.

DROSHA is an RNase III enzyme that forms the microprocessor complex along with DGCR8, the stranded RNA-binding protein. The interaction between DGCR8 and DROSHA occurs when DGCR8 binds to the RNase III domains of DROSHA (Y Lee et al.). Together DROSHA and DGCR8 form the microprocessor that converts pri-miRNA to pre-miRNA (Gregory et al.) . DROSHA requires DGCR8 to specifically bind to the pri-miRNA and recognize the pre-miRNA structure by detecting a terminal loop longer than 10 nucleotides. DROSHA cleavage sites are then determined based on proximity to this loop (Zeng, Yi, and Cullen). This mechanism results in DGCR8 and DROSHA determining the end points of the pre-miRNA. miRNA biogenesis is regulated by an auto feedback loop between DROSHA and DGCR8. In this loop, DGCR8 stabilizes DROSHA protein levels, and DROSHA destabilizes DGCR8 mRNA by cleaving the hairpin structures in DGCR8 mRNA (Han et al.; Yoontae Lee et al.; Y Lee et al.).

The pre-miRNAs formed as a result of interaction with the microprocessor complex go on to become mature miRNAs. This is accomplished by the actions of the RNaseIII DICER. DICER is able to cut pre-miRNA into a 22nt long miRNA duplex. This duplex is made up of two miRNA strands. One of the strands is loaded onto RISC, the RNA-induced silencing complex. Argonaut is a major component of the RISC complex (Chiang et al.). The mature miRNAs can act as guides which target RISC to the appropriate mRNAs, which then can be degraded or destabilized by the Ago proteins (Bartel).

Figure 1:

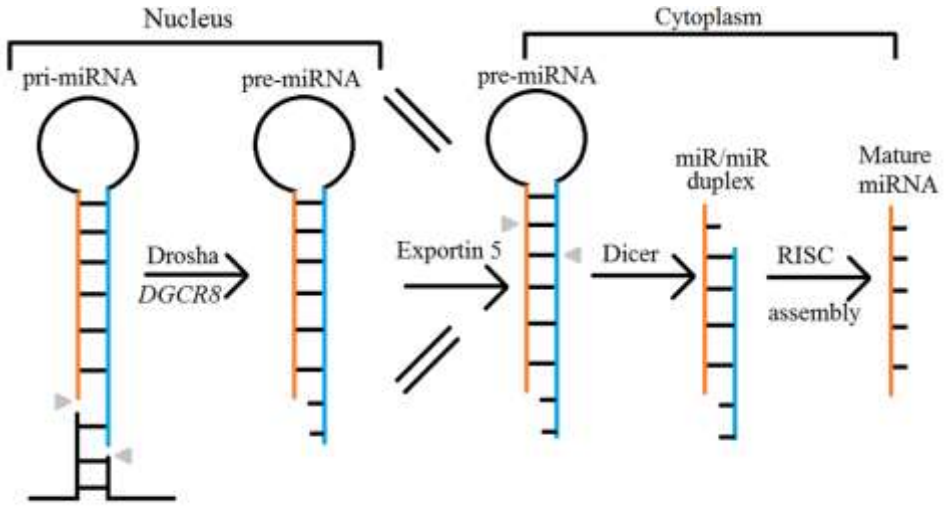


Figure 1 depicts the process of miRNA biogenesis. DROSHA and DGCR8 together form the microprocessor complex that converts pri-miRNA into pre-miRNA in the nucleus. Pre-miRNA is exported to the cytoplasm via exportin 5. DICER then cuts the pre-miRNA into the 22 nucleotide long miRNA duplex. One of these strands is then converted into a mature miRNA by the RISC complex

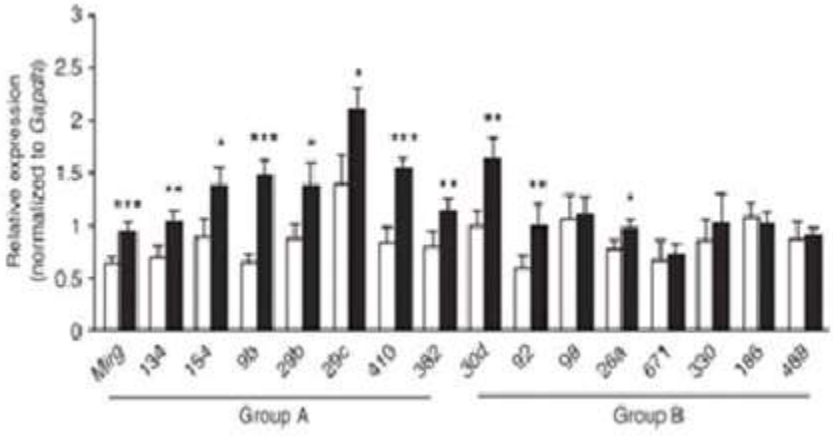
Dgcr8 $^{+/-}$ mice have behaviors analogous to schizophrenia

The complex process of miRNA biogenesis is perturbed as a result of the 22q11.2 deletion, since the deletion causes DGCR8 haploinsufficiency. Haploinsufficiency occurs when only a single functional copy of a gene exists instead of the normal two copies, and the one functioning copy cannot produce sufficient protein by itself. The importance of DGCR8 in miRNA synthesis has been elucidated through the development of a strain of mice with a hemizygous deletion on chromosome 16 that is syntenic to the 22q11.2 deletion in humans and includes 27 genes (Edelmann, Pandita, and Morrow; Stark et al.). Mice with this deletion, as well as *Dgcr8* $^{+/-}$ mice, show low levels of multiple miRNAs and behaviors reminiscent of schizophrenia in humans. Both *Dgcr8* $^{+/-}$ mice and the mice with the partial deletion of chromosome 16 had deficits in synaptic connectivity in the hippocampus as a result of a reduction in the size and number of dendritic spines. Also found were reductions in dendritic complexity of CA1 pyramidal neurons. The mice were found to have hyperactive behavior and increased fear behaviors. These behaviors, like schizophrenia, were found to be more prevalent in males (Stark et al.).

Using miRNA probes that targeted specifically mature miRNAs, transcripts that included miRNAs were found to be differentially expressed in the hippocampus and prefrontal cortex of *Dgcr8* $^{+/-}$ mice (Stark et al.). Knowing that production of mature miRNAs is altered and that the deletion analogous to 22q11.2 contains *Dgcr8,* which is required to process pri-miRNA to mature miRNAs, the logical assumption is that miRNA is not being converted from pri-miRNA to pre-miRNA. Without proper levels of *Dgcr8,* miRNAs cannot be processed, and pri-miRNA specifically should be expected to build up (Stark et al.).

The accumulation of pri-miRNA was confirmed by using pri-form-specific qRT-PCR on miRNA genes that were close to the probe sets that had differential expression in the prefrontal cortex and the hippocampus. In figure 2, the levels of pri-miRNA of two groups were analyzed. Group A consists of genes that were in the vicinity of probe sets differentially upregulated in the prefrontal cortex. Group B consists of genes that were not in the vicinity of probe sets differentially upregulated in the prefrontal cortex. All of the miRNA genes in group A exhibit buildup of pri-miRNAs and a failure of mature miRNAs to be processed as a result of *Dgcr8* deficiency (Stark et al.). However, an increase in pri-miRNA does not always mean that there is a decrease in the mature form of the miRNA. Only one out of six increased pri-miRNAs had a reduction in the mature form of miRNA, suggesting that alternative compensatory mechanisms exist to regulate specific miRNAs. These alternative compensatory mechanisms could consist of positive and negative feedback mechanisms that exist to ensure proper development of mature miRNAs.

Figure 2: The expression levels of pri-miRNA in the prefrontal cortex of *Dgcr8* $^{+/-}$ mice were measured using pri-form-specific qRT-PCR. Group A consists of miRNA that are associated with genes that were found to be differentially upregulated in the mice. Group B consists of miRNA that are associated with genes not found to be differentially upregulated in the mice. Shaded bars are the *Dgcr8* $^{+/-}$ mice. White bars are wild type mice (Stark et al.).



miRNA-based microarrays were used to determine which mature miRNAs were down-regulated as a result of the deletion by comparing the expression of miRNAs in the prefrontal cortex and the hippocampus in mice with and without the deletion. 25 miRNAs were found to have reduced expression in both the PFC and the hippocampus. The predictive software programs TargetScan and PicTar were used to determine potential target sites for these miRNAs in the 3′ UTR of transcripts within the TargetScan or PicTar databases that were differentially expressed in the PFC and the hippocampus. Since miRNAs are known to regulate stability of the transcript, the down-regulation of miRNA should result in the upregulation of their target genes. A higher percentage of transcripts that were upregulated in the PFC and hippocampus had one or more target seed sites for these miRNAs than transcripts that were down regulated in the PFC and hippocampus. This demonstrates that proper brain miRNA processing and regulation of target genes is dependent on *Dgcr8* (Stark et al.). These alterations to the miRNA regulatory pathways have already been shown to have a direct link to schizophrenia. The human equivalent of eight of the miRNAs that were found to have reduced expression of their mature forms in Stark et al. in the prefrontal cortex of the *Dgcr8* $^{+/-}$ mice were also found to have reduced expression in prefrontal cortex postmortem brain samples from individuals with schizophrenia (Stark et al.; Moreau et al.).

miRNA Canalization

miRNAs are abundant in the brain and nearly 70% of all known miRNAs are found in the nervous system where their expression is controlled both spatially and temporally (Krichevsky et al.; Miska et al.). miRNAs are also able to keep gene expression at the required levels by binding to the 3'UTR of genes and down-regulating expression. Many

miRNAs work in concert to regulate the expression of genes so that only a specific and desired developmental process occurs. This process is called canalization, in which developmental pathways are stabilized to increase phenotypic reproducibility (Gibson and Wagner; Hornstein and Shomron). Without miRNAs to ensure proper canalization, mutations in regulatory regions could readily cause over or under expression leading to deleterious effects and disease.

A Hypothesis for 22q11.2 Schizophrenia Risk

It is our hypothesis that the interaction of miRNAs and a constellation of underlying variants contribute to determining why some individuals with 22q11.2 DS develop schizophrenia, yet 75% of individuals with the deletion do not develop the disease. In healthy individuals, genes may accumulate deleterious mutations that alter gene expression. However, these mutations may be harmless if the miRNA regulatory system is able to counteract the effect of the mutation. We believe that the miRNAs with altered expression as a result of the 22q11.2 deletion increase the risk of schizophrenia through a disruption in the miRNA regulatory pathway. This disruption represents just one step in surpassing a threshold level for the onset of schizophrenia and compromising existing compensatory mechanisms. It is important to remember that although this may be one factor involved in the disease, schizophrenia is ultimately triggered by a cumulative effect of multiple mutations and environmental factors causing the illness.

Stark et al. have identified multiple miRNAs that may play a role in the development of behavioral and neuronal deficits as a result of the 22q11.2 deletion in mice. However, the human genome contains different genes and different miRNAs than those present in mouse models. If different miRNAs are differentially expressed, then the target genes

impacted likely will also be different too. The proposed experiments attempt to determine exactly which miRNAs have altered expression as a result of the 22q11.2 deletion in humans. Previous studies have attempted to characterize these miRNAs in humans and how our study improves on existing work will be discussed in depth in Chapter 2 (Sellier et al.; Zhao et al.).

In the hopes of further characterizing disruptions to miRNAs as a result of the 22q11.2 deletion and better understanding how these disruptions to miRNAs could contribute to the development of schizophrenia, Chapter 2 describes a stem cell model that we have developed to characterize these disrupted miRNAs. Chapter 3 presents an in depth description of our stem cell quality control methods that have been developed to ensure the chromosomal fidelity of both our iPSCs and NSCs. Chapter 4 describes our characterization of dysregulated miRNAs via TaqMan® arrays. Lastly, Chapter 5 demonstrates how one specific miRNA, miR-616, can contribute to the development of schizophrenia.

Chapter 2: Importance of Stem Cell Models, Rationale For Study, and Generating Stem Cell Models.

Previous Human Studies of miRNA in 22q11.2 DS

Beyond the previously described postmortem studies of miRNAs in the brain, two recent attempts have been made to characterize the miRNAs that have disrupted expression in individuals with 22q11.2 DS using other models. One study by Sellier et al. used peripheral leukocytes from individuals with and without 22q11.2 DS to compare differences in miRNA expression. One of the major strengths of this study is that it was able to acquire a large number of individuals with 22q11.2 DS. Blood samples from ninety individuals with the 22q11.2 deletion were compared to blood samples from forty otherwise healthy individuals. This study specifically identified seven miRNAs that were found to be dysregulated in blood samples obtained from individuals with 22q11.2 DS. miR-150, miR-185, miR-194, and miR-363 were found to be downregulated, while miR-208, miR-190, and miR-1 were upregulated in individuals with 22q11.2 DS (Sellier et al.). Of these miRNAs, miR-185 and miR-194 were both previously observed to be downregulated in the prefrontal cortex and hippocampus of mice with a deletion equivalent to the 22q11.2 deletion (Stark et al.).

The Sellier study only looked at the 600 miRNAs preselected for their importance in cardiac function on an array. Using small miRNA sequencing would have allowed many more miRNAs to be studied without a bias toward miRNAs believed to be important in cardiac function, so in this respect the scope of the miRNA study is fairly limited. Furthermore, this study looked at miRNAs found in human blood samples. miRNAs have been shown to have altered expression levels in various tissues in the body (Guo et al.), therefore the miRNAs that are present in the blood at altered levels in

individuals with 22q11.2 DS may not be miRNAs that are dysregulated in the brain. The brain has been shown to express more distinct miRNAs than any other tissue, including miRNAs that are unique to the brain (Adlakha and Saini). Since we are interested in determining which miRNAs may play a causative role in the development of schizophrenia as a result of the 22q11.2 deletion, it is crucial to study neural tissue directly.

The second study that attempted to characterize the miRNAs that have disrupted expression in individuals with 22q11.2 DS and schizophrenia used a stem cell model (Zhao et al.). The stem cells used in this research were induced pluripotent stem cells (iPSCs). iPSCS are cells that have been reprogrammed from existing differentiated cells, in this case fibroblasts, to become pluripotent. A pluripotent cell is a cell that can differentiate into any cell type in the body. Reprogramming of fibroblasts can be achieved by transfecting the Yamanaka factors (Oct3/4, Sox2, c-Myc, and Klf4) into the cells (Liu et al.).

The iPSCs and ultimately neurons used by Zhao et al were derived from cells lines generated from individuals with 22q11.2 DS and schizophrenia. The 22q11.2 deletion was confirmed in these cell lines via FISH. The Zhao et al. study used small miRNA sequencing using the Illumina HiSeq2500 Massively Parallel Sequencing platform as single end 50 bp read length to examine the miRNA expression levels in neurons from individuals with the 22q11.2 deletion and compare them to otherwise healthy individuals. DESeq2 was used to identify the differentially expressed miRNAs. The miRNAs that were found to have significantly downregulated expression in individuals with schizophrenia and 22q11.2 DS included miR-185 and miR-491, which

were previously found to be downregulated as result of the 22q11.2 DS in mouse models (Stark et al.). Additionally, miR-185 and mir-194 were both found to be downregulated in both mouse models and in human peripheral leukocytes (Stark et al.; Sellier et al.). The additional significantly downregulated miRNAs in the Zhao et al. study included miR-3175, miR-486, and miR-4804. Several miRNAs were found to be increased in samples with both 22q11.2 DS and schizophrenia, however there was only a nominally significant increase reported. Many of these miRNAs were further found to be involved in neurological disorders and psychiatric diseases after DAVID functional annotation Gene Ontology analysis was performed suggesting that they are important even if they did not reach statistical significance in this study (Zhao et al.).

These two studies used live cells to identify miRNAs that are dysregulated as a result of the 22q11.2 deletion and are potentially important for the development of schizophrenia. Only one miRNA found to be dysregulated in either of these studies, miR-190, was previous reported in postmortem studies (Moreau et al.). This likely reflects the inherent difficulties in studies postmortem brains that were previously described in Chapter 1. Although these two studies eliminate the complications associated with postmortem brain study, the study by Sellier et al. and the study by Zhao et al. have flaws of their own. The Sellier et al. study did not use neural tissue as a model. In order to characterize the miRNAs that are important to the development of schizophrenia in the brain it is imperative that some type of brain matter be studied. Although studying blood can be useful to potentially develop a diagnostic test, blood samples are no substitute for neural tissue. The second study by Zhao et al. made other experimental choices that were not ideal. The sample size was small and the samples that were sequenced were

heterogeneous populations of neurons.

The neurons used for miRNA-seq in this study were day 14 neurons and consisted of a mixed population of glutamatergic and GABAergic neurons (Zhao et al.). The use of a mixed population of neurons will be further addressed in our analysis of 22q11.2 DS neural stem cells and will be further discussed in this chapter.

Although neurons would at first glance appear to be the ideal cellular model to study the dysregulated miRNAs in individuals with 22q11.2 DS and schizophrenia, further consideration shows that this is not the case. In the Zhao et al study, they further differentiated their NSCs into neurons. These neurons were then used for their small miRNA sequencing. The fundamental problem here is that they differentiated their NSCs into a mixed population of neurons, rather than just limiting their profiling to one subtype of neuron. They observed that they obtained a mixed population of glutamatergic and GABAergic neurons. This is problematic because miRNAs are known to vary between neuronal subtypes (He et al.). Therefore it would be impossible to determine if differences in miRNA expression levels are a result of the differences between samples with 22q11.2 DS and schizophrenia and control samples or are actually just as a result of differences in the amount of a particular neuronal subtype present in a particular sample.

Neural Stem Cell Rationale

Although there are a number of faults present in the Zhao et al study, it does make use of stem cells which appear to be the best experimental model available for characterizing the dysregulated miRNAs in individuals with 22q11.2 DS. Unlike postmortem samples, which are difficult to obtain, stem cell lines can be generated from living individuals from easily obtainable fibroblast or lymphoblast samples. Stem cells

are a renewable source of tissue, eliminating the issue of the limited amount of tissue available from postmortem brains. Furthermore, neural stem cells can be grown in controlled conditions, eliminating the previously described environmental issues associated with postmortem samples such as improper storage conditions of the brain, and drug and alcohol usage prior to death. Neurons generated from iPSCs derived from individuals with schizophrenia have already been shown to exhibit a phenotype consistent with schizophrenia (Brennand et al.). However, this promising new avenue of research is not without limitations. NSCs are challenging to make, and differentiating neural stem cells into neurons can result in a mixed population of neurons that may not accurately represent the neurons found in a brain region of interest. Even though they more closely model human brain tissue, NSCs still lack the complex architecture and developmental history of a brain.

Cell Line Selection:

All of the samples used by Zhao et al. were submitted to the Rutgers University Cell and DNA Repository. We were able to obtain all of the lines with 22q11.2 DS and schizophrenia, as well as four of the control lines. The following 9 cell lines that are a subset of the Zhao et al. served as the basis for the initial phase of this research (Control lines: MH0159019, MH0159020, MH0159021 MH0159022)(22q11.2 Lines: MH0159023, MH0159024, and MH0159027)(Schizoaffective lines MH0159025, MH0159026). These lines were selected initially since RUCDR was able to distribute these lines as iPSCS. All other lines were only available as fibroblasts. iPSCS for all other lines were made as part of this research.

Every opportunity was taken to find additional 22q11.2 lines that were already

collected by known cell repositories and available for distribution. Finding the best suitable control lines for each identified 22q11.2 line was also a top priority. Every effort was made to match each 22q11.2 line to a relative that was age and gender matched. However, preference was given for including relatives of individuals with 22q11.2 DS wherever available even if there age and gender differed.

Table 3:

The following table shows the lines that were used by Zhao et al as well as other cell lines used in this research. “Zhao ID” is the identification number used by Zhao et al. “RUID” is the identification number issued by the RUCDR and used for our research. Nine of these twelve lines served as the initial sample for our study (MH0159019, MH0159020, MH0159021, MH0159022, MH0159023, MH0159024, MH0159025, MH0159026, and MH0159027). All four of the 22q11.2 DS / schizophrenia lines that were previously used in the Zhao study were obtained. Lines shown in the table that do not have an RUID number were not used in our study. iPSCs for lines MH0159019, MH0159020, MH0159021, MH0159022, MH0159023, MH0159024, MH0159025, MH0159026, and MH0159027 were obtained from RUCDR. iPSCS from MH0161608 were not viable. However, fibroblasts samples of MH0161608 were obtained and iPSCs were able to be generated for MH0161608. All other iPSC lines with RUID numbers were generated from fibroblast or iPSC samples distributed by RUCDR

RUID	Zhao ID	Gender	Age	Race	Ethnicity	Diagnosis	Cell Type Acquired
MH0159019	ctrl_iPSC1	Female	29	White	Hispanic or Latino	control	iPSC
MH0159020	ctrl_iPSC2	Male	58	White	Not Hispanic or Latino	control	iPSC
MH0159021	ctrl_iPSC5	Male	32	White	Hispanic or Latino	control	iPSC
MH0159022	ctrl_iPSC6	Male	46	White	Not Hispanic or Latino	control	iPSC
MH0159023	SZ_22q11-30	Female	41	White	Not Hispanic or Latino	schizoaffective disorder / VCFS	iPSC
MH0159024	SZ_22q11-10	Male	37	White	Not Hispanic or Latino	schizoaffective disorder / VCFS	iPSC
MH0159027	SZ_iPSCS15	Male	31	White	Not Hispanic or Latino	schizoaffective disorder / VCFS	iPSC
MH0161608	SZ_22q11-60	Male	25	White	Not Hispanic or Latino	schizoaffective disorder / VCFS	iPSC (not viable) / Fibroblast
MH0159025	n/a	Male	48	White	Not Hispanic or Latino	schizophrenia, paranoid type	iPSC
MH0159026	n/a	Male	60	White	Not Hispanic or Latino	schizophrenia undiff type	iPSC
MH0162499	n/a	Male	31	White	Not Hispanic or Latino	22q11.21 del / childhood onset SZ	Fibroblast

MH0162508	n/a	Female	23	White	Not Hispanic or Latino	22q11.21 del / childhood onset SZ	Fibroblast
MH0162519	n/a	Female	26	White	Not Hispanic or Latino	22q11.21 del / childhood onset SZ	Fibroblast
MH0162627	n/a	Female	12	Black	Not Hispanic or Latino	22q11.21 del / childhood onset SZ	Fibroblast
MH0162762	n/a	Male	4	White	Not Hispanic or Latino	22q11.2 deletion / Autism	Fibroblast
MH0162509	n/a	Male	28	White	Not Hispanic or Latino	Control	Fibroblast
MH0162626	n/a	Female	37	Black	Not Hispanic or Latino	Control	Fibroblast
MH0162577	n/a	Female	15	White	Not Hispanic or Latino	Control	Fibroblast
MH0167170	n/a	Male	31	White	Not Hispanic or Latino	Control	iPSC
MH0161602	n/a	Female	18	White	Not Hispanic or Latino	Control	iPSC

Generating Stem Cell Models

Fibroblast Culture

All of the cells lines used for the characterization of miRNAs in individuals with 22q11.2 DS were originally obtained from the Rutgers University Cell and DNA Repository (RUCDR). The cells were obtained as either induced pluripotent stem cells (iPSCs) or fibroblasts. In cases where iPSCs were not submitted to RUCDR or the submitted iPSCs were not viable (MH0161608), the iPSCs used were generated from fibroblast samples that were submitted to RUCDR (Table 3). All fibroblasts were cultured in fibroblast medium consisting of DMEM with high glucose, L-glutamine, phenol red, and sodium pyruvate (Thermo Fisher Scientific) supplemented with 15% fetal bovine serum (Thermo Fisher Scientific), 1% penicillin-streptomycin (10,000 U/mL) (Thermo Fisher Scientific), and 1% non-essential amino acids solution (100X) (NEAA) (Hyclone). Fibroblasts were grown in 6 well plates coated with 50 µl/cm^2 Matrigel (BD Biosciences). When fibroblasts reached 80%-90% confluency, they were passaged at ratio 1:3 with accutase cell detachment solution (Innovative Cell Technologies, Inc.) onto Matrigel (BD Biosciences) coated plates and given fresh medium until cells reach 80% confluency (~3-4 days)

Conversion of Fibroblasts to iPSCS

The fibroblasts that we obtained from RUCDR were converted to iPSCs using an RUCDR protocol that used episomal vectors. The Invitrogen episomal iPSC reprogramming vectors (Thermo Fisher Scientific) contain a mixture of three vectors designed to generate transgene-free and virus-free iPSCs in a feeder-free environment

(Yu et al.). The oriP/EBNA1 vectors include six reprogramming factors (Oct4, Sox2, Nanog, Lin28, Klf4 and lMyc). These vectors have been proven to be effective at reprogramming a variety of cell types including fibroblasts. The vectors are inserted into the cells via electroporation and replicate extra chromosomally once per cell cycle. However, it is estimated that 5% of cells lose the episomes during each passage. Episomal reprogramming vectors were used instead of lentivurus, which is known to contain transgenes that can integrate into the host genome.

Three days prior to transfections $1.5x10^6$ human fibroblast cells were plated in 50 $\mu l/cm^2$ Matrigel (BD Biosciences) coated T75 flasks with 10ml of fibroblast medium. The fibroblasts in the T75 flask were grown at 37°C, in a 5% CO2 incubator. Spent fibroblast medium was exchanged every other day with fresh fibroblast medium until the fibroblasts in the T75 culture flask reached 75%-90% confluency.

Confluent fibroblasts were removed from the plates with accutase cell detachment solution (Innovative Cell Technologies, Inc.) The cells were washed with PBS (Thermo Fisher Scientific) and a pellet was collected via centrifugation at 3000rpm for 3 min in a microcentrifuge tube. Transfections of fibroblasts were performed using the Amaxa NHDF Nucleofector kit (Lonza) following the Lonza standard protocol. Approximately one million fibroblast cells was transfected per well. One million fibroblasts were resuspended in were resuspended in 100ul of NHDF buffer (Lonza). 8.5µl of Invitrogen Episomal vectors were added to the resuspended fibroblasts in NDHF buffer. The mixture was gently mixed using plastic transfer pipettes. The cells in NHDF buffer were then transferred to AMAXA nucleovette tubes (Lonza). Electroporation was performed

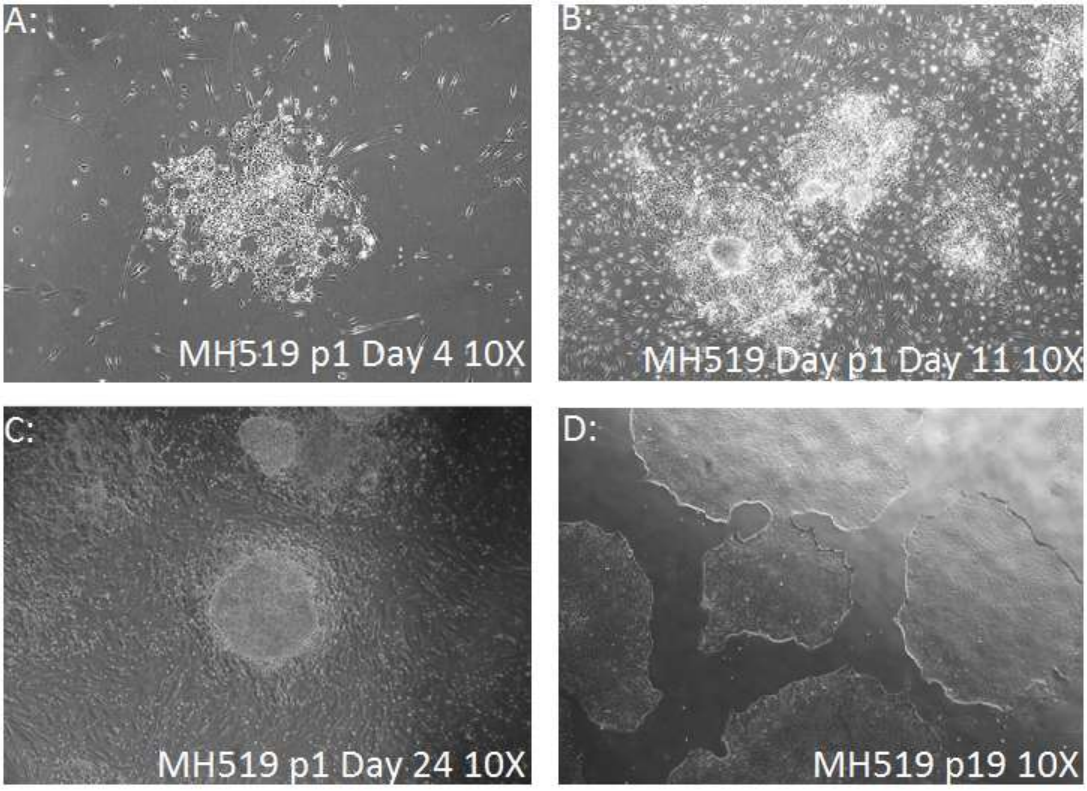
using the AMAXA transfection System following the NHDF Nucleofector kit protocol (Lonza).

Newly transfected cells were plated onto plates coated with 50 µl/cm^2 Matrigel (BD Biosciences). A complex reprogramming medium consisting of 1X NEAA (Hyclone), 1X N2 (Thermo Fisher Scientific), 1X B27 (Thermo Fisher Scientific), 0.1mM Beta-Mercaptoethanol (Thermo Fisher Scientific), 100ng/ml fibroblast growth factor (Peprotech), 1X GLUTAMAX (Thermo Fisher Scientific), 0.5uM PD0325901 (Stemgent), 3uM CHIR99021 (Stemgent), 0.5uM A8301 (Stemgent), 1000units/ml Human Recombinant human leukemia inhibitory factor(Millipore), 10uM HA100 (Santa Cruz Biotechnology), and 94.478 ml of DMEM/F12 per 100 ml of media (Thermo Fisher Scientific) was used for the newly transfected cells. The cells were treated with fresh reprogramming media every other day for two weeks' time. After two weeks the cells were fed with mTeSR (StemCell) media every other day. iPSCs colonies formed over the course of the following 30 days.

In the first week following electroporation cells were observed to form aggregate colonies with morphology that differed from the neighboring fibroblasts (Figure 3 Panel A). During the second week the colonies of cells grew denser and more visually distinct from the neighboring fibroblasts (Figure 3 Panel B). By the third week following transfection, the colonies began to show boarders that were clearly defined which separated the colonies from the surrounding fibroblasts (Figure 3 Panel C). Larger colonies were able to be picked and replated using manual scraping techniques. The colonies were detached using fire blown glass picks. After several weeks the colonies were largely free of differentiated cells (Figure 3 Panel D), although at even given

passage approximately 10% of the cells were observed to differentiate into fibroblasts. Additional examples of morphology observed following iPSC transfections are shown in supplemental figure 1.

Figure 3: The following figure shows the induction process of fibroblasts to iPSCs. Newly formed iPSC colonies are shown at 4, 11, and 24 days following transfection (Panel A-C). The colonies are also shown at several passages after transfection (Panel D).



Alkaline Phosphatase Stain:

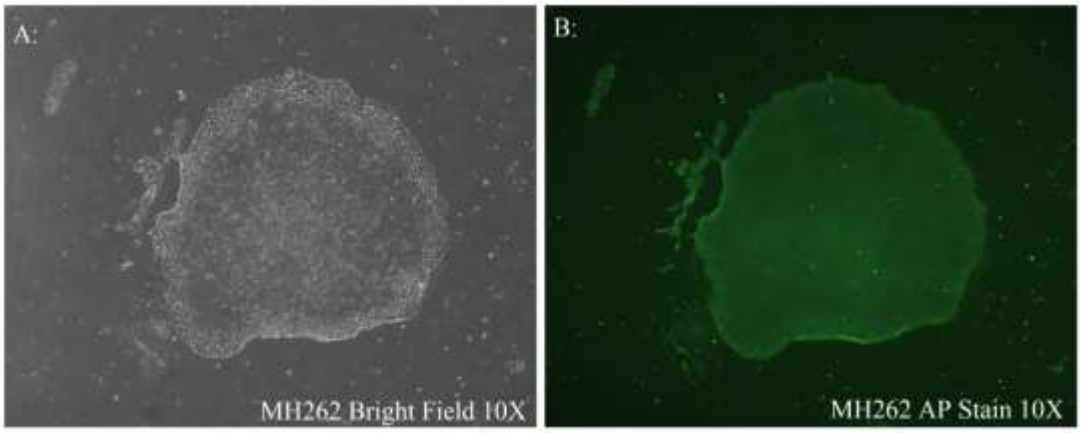
In order to confirm that iPSCs colonies were pluripotent, alkaline phosphatase staining was performed on iPSC lines using alkaline phosphatase live stain 500X (Thermo Fisher Scientific). Alkaline phosphatase (AP) is a phenotypic marker of pluripotent stem cells, including induced pluripotent stem cells (iPSCs). Although alkaline phosphatase is expressed in most cell types, the expression of alkaline phosphatase is highly elevated in pluripotent stem cells (Quintanilla, Asprer, and Lakshmipathy; Martí et al.).

In order to stain iPSCs with the AP live stain, mTeSR was first removed from cells growing in 6 well plates. The iPSCs were twice washed with pre-warmed DMEM/F-12 (Thermo Fisher Scientific) for 2–3 minutes. 1X AP Live Stain working solution was prepared by diluting the 500X stock solution in DMEM/F-12 (Thermo Fisher Scientific). 2 mls of the 1X AP Live Stain solution was applied directly to the iPSCs in 6 well plates. The iPSCs were incubated for 30 minutes. The AP Live Stain was aspirated and the iPSCs were washed twice with DMEM/F-12 for 5 minutes per wash. After the final wash fresh mTeSR was added prior to the visualization of fluorescent-labeled colonies.

Alkaline phosphatase live stain is a cell-permeable green fluorescent substrate for alkaline phosphatase. A major advantage to using alkaline phosphatase staining over other types of pluripotent stem cell markers is that it is non-toxic to cells. The protocol allows visualization of iPSC colonies in thirty minutes and diffuses out of the cells after about two hours. The fact that alkaline phosphatase staining is non-toxic is especially important, since this stain can be performed early in reprogramming on the first iPSC

passage to test for pluripotency without sacrificing valuable iPSC colonies. A representative photo of the alkaline phosphatase staining that was observed after staining iPSC colonies is shown in figure 4. The glowing green iPSC colony confirms the pluripotency of this colony.

Figure 4: The figure below shows a newly formed iPSC colony (Panel A). The same colony fluoresces green after stained for alkaline phosphatase (Panel B).



RUCDR Quality Control:

Additional quality control was performed on iPSC lines obtained directly from RUCDR. RUCDR performed their own test of Alkaline Phosphatase expression. The RUCDR Quality Assurance processes also include tests to demonstrate that the cell lines distributed are free of microbiological contamination. Mycoplasma are a type of bacteria that commonly contaminate cell lines in culture. Unlike other common bacterial contamination, mycoplasma contamination is very difficult to observe and does not change the color of the culture medium. Mycoplasma are especially problematic because they can often go undetected altering growth rates of the cells, causing morphological changes, and triggering chromosomal aberrations. Mycoplasma testing was performed using medium removed from the cell cultures after it had sat on the cells for 24 hours (RUCDR).

RUCDR has also confirmed the pluripotency of the distributed iPSCs using fluorescence-activated cell sorting (FACS), which is a type of flow cytometry that is used to measure fluorescence intensity produced by fluorescent-labeled antibodies. The staining protocol uses a single cell suspension which is then treated with fluorochrome-labeled antibodies, Octamer-4 (Oct4) (Millipore) and Tra-1-60 (Millipore). Oct-4 is a member of the POU family of transcription factors. Oct-4 is found in pluripotent cells of the inner cell mass during embryogenesis and is an accepted marker of human iPSC pluripotency (Tadeu et al.). After cell differentiation, Oct-4 expression decreases. Tra-1-60 is highly expressed upon the surface of human embryonic stem cells (ES) and has been shown to also be a reliable marker for human iPSC pluripotency (Ban et al.).

RUCDR quality control reports for all distributed cell lines showed that the cell lines exhibited normal cell morphology before and after cryopreservation. All iPSCs distributed by RUCDR were successfully stained for alkaline phosphatase before distribution. Oct/4Tra-1-60 FACs analysis was performed on all RUCDR iPSCS and showed that the distributed iPSCs express these two markers of pluripotency. An example of the RUCDR quality control report is shown in supplemental figure 2.

SNP Identity Panel

The cell lines used throughout this research were all obtained directly from RUCDR. However, all of the cell lines were deposited to RUCDR by several different research groups. Due to the scarcity of known 22q11.2 deletion / schizophrenia lines every attempt was made to use any 22q11.2 DS / schizophrenia line that could be obtained. The possibility exists that multiple cell lines were created and submitted to RUCDR from the same subject by different investigators. In order to ensure that all of our cell lines did indeed represent unique subjects, SNP identity panels were run by RUCDR on all the cell lines that were directly submitted to RUCDR. Each cell line was found to have a unique SNPTrace profile on file with no matches to any other samples in the database, except for expected instances where iPSC lines matched to fibroblast or lymphoblastoid lines known to be generated from the same subject. Genotyping of the samples was performed by RUCDR using SNPtrace™ Panel Genotyping Assays which uses 96 SNPs to provide unique information about the identity and quality of human genomic DNA samples. The panel includes highly polymorphic SNPs that are gender and ethnicity specific. An example of the SNPtrace™ Panel Genotyping Assays results is shown in supplementary figure 3.

Neural Induction Protocol

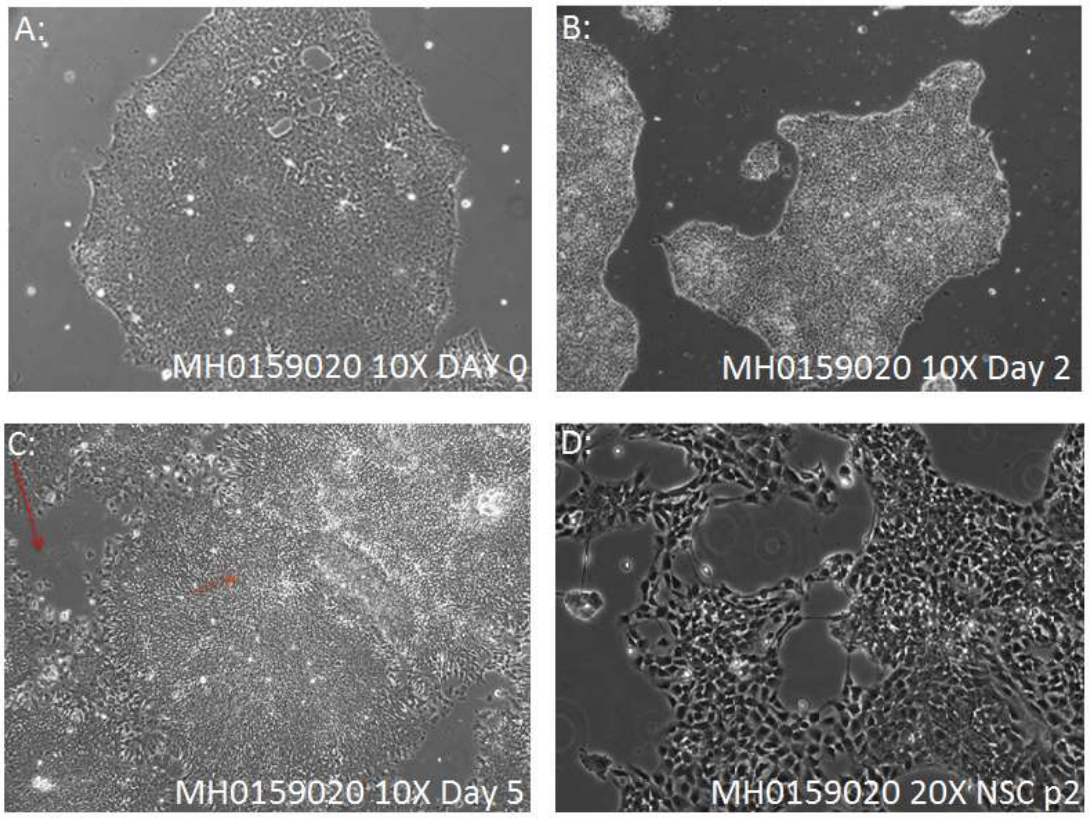
All iPSC lines were converted to NSCs using PSC Neural Induction Medium consisting of 98 mls of Neurobasal Medium and 2 mls of Neural Supplement (Thermo Fisher Scientific). PSC Neural Induction Medium is a serum-free medium that that is used for the induction of neural human pluripotent stems cells (PSCs). The neural induction protocol started with iPSCs growing in 6 well plates that had reached ~50% confluency. mTeSR was aspirated from the iPSCs and the cells were treated with Neural Induction Medium. Neural Induction Medium was changed daily.

The PSC neural induction medium is preferable to other induction methods since this protocol is fast, taking only seven days to generate NSCs. Cell lines that were slower growing were treated with neural induction media for an additional week. PSC Neural Induction Medium also does not require the intermediary step of embryoid body (EB) formation, which adds time, labor, and variability. During the induction process the morphology of the iPSC colonies can visibly be observed to change. The normally round colonies become less round. The tint of the colonies changes as well. Normally iPSCs have a very flat appearance. The NSCs are shinier than the iPSC colonies as shown in Panels B-D.

Once the NSCs were generated they were passaged using cell detachment solution (Innovative Cell Technologies, Inc.) The cells were exposed to accutase for 5 mins. After the first 45 secs, detached cells were aspirated since this fraction of cells would generally contain differentiated cells. The cells were retreated with fresh accutase for the remaining four minutes and scraped. The cells were washed with PBS (Thermo Fisher Scientific) and a pellet was collected via centrifugation at 3000 rpm for 5 min in a

centrifuge tube. The pellet was resuspended in neural expansion medium consisted of 49 mls of Neural Basal Medium (Thermo Fisher Scientific), 49 mls of Advanced DMEM (Thermo Fisher Scientific), and 2 mls of Neural Supplement (Thermo Fisher Scientific). The cells were grown on 6 well plates coated with Matrigel (BD Biosciences). 1 µl of 5 mM ROCK Inhibitor Y27632 (Sigma-Aldrich) was added to each well of the 6 well plate containing cells.

Figure 5: This figure shows the induction process of NSCs using an iPSC line. Panel A shown the starting iPSC line treated with induction media. Panels B and C show the induction at 2 and 5 days after treatment with induction media. Panel D shows the newly formed NSCs on their second passage.



Immunofluorescent Staining for Neural Markers

In order to confirm that NSCs generated from iPSCs were actually NSCs and not another subtype of cell, immunofluorescent staining was used for markers specifically expressed in NSCs. The NSCs were stained with β-Tubulin, Musashi, Nestin, and Pax-6. All of these proteins have been shown to be excellent markers of neural stem cells (Strojnik et al.).

Immunofluorescent Staining Methods:

In order to stain the NSCs, the NSCs were plated on a 4-well 50 µl/cm^2 Matrigel (BD Biosciences) coated chambered slide at a density of $0.5x10^5$ – $1x10^5$ cells/cm^2. NSCs were plated in NEM. 0.5 µl of 5 mM ROCK Inhibitor Y27632 (Sigma-Aldrich) was added to each well of the chamber slide. The NSCs received fresh NEM (Thermo Fisher Scientific) every other day until they reached 85% confluency. After reaching confluency, the NSCs were fixed with 4% paraformaldehyde (Sigma-Aldrich) and stored in PBS at 4°C. The cells were treated with 250 µl of 0.1% Triton X-100 (Thermo Fisher Scientific) and incubated for 15 minutes at room temperature for up to one week. The Triton X-100 was aspirated and 250 µl of 4% normal goat serum/PBS was added to each well and incubated at room temperature for one hour (Thermo Fisher Scientific). The goat serum was aspirated and the primary antibody diluted in 4% Normal Goat Serum/PBS was added to the chamber slides. The cells were incubated in the primary antibody wrapped in Parafilm (Bemis NA) overnight at 4°C. The next day the primary solution was aspirated and the cells were washed three times with PBS. The secondary antibody was prepared in 4% Normal goat serum/PBS. 250 µl of the appropriate secondary antibody mix was added to each well of the chamber slides. The slides were

then incubated at room temperature in the dark. After one hour, the secondary was aspirated and the cells were washed in PBS. 250 µl of DAPI (Thermo Fisher Scientific) was added to the slides. DAPI was allowed to incubate for 5 minutes in the dark at room temperature. The cells were then washed with PBS again.

After completion of the immunofluorescent staining protocol, the chambers were carefully removed from the chamber slides while the slides were submerged in ddH_20. Glass coverslips were mounted onto the slides using Fluoromount-G (Thermo Fisher Scientific). The slides were transferred to a flat container and stored in the dark overnight. 24-48 hours later the slides were sealed with clear polish and kept in the dark. The slides were examined and photographed using an Olympus IX71 inverted fluorescence microscope and QCapture 2.9.13 Software (Quantitative Imaging Corporation).

Antibodies Used:

The immunofluorescent staining was performed on the NSCs using three separate sets of antibodies on three separate coverslips per cell line. All samples were stained with DAPI, a blue control marker for DNA (Thermo Fisher Scientific). The first combination included DAPI, a primary of B-Tubulin (TUJ1) monoclonal antibody(1:1000) (BioLegend) and a secondary of Alexa-Fluor®488 goat anti-mouse IgG2a(y2a) (1:500) (Thermo Fisher Scientific).

The second combination included DAPI, a primary of Mus-R (1:250) (Abcam) and a secondary of Alexa-Fluor® 594 goat anti-rabbit (Thermo Fisher Scientific) (1:500).

The third combination consisted of primary antibodies pax-6 polyclonal antibody (1:300) (BioLegend) and anti-nestin monoclonal (1:200) (Millipore) along with

secondary antibodies Alexa-Fluor® 488 goat anti-mouse IgG2a(y2a) (1:500) (Thermo Fisher Scientific) and Alexa-Fluor® 594 goat anti-rabbit (1:500) (Thermo Fisher Scientific).

Immunofluorescent Staining For Neural Markers Results:

β-Tubulin

Mouse IgG2a β-Tubulin was used to stain NSCs (BioLegend). Microtubules are made up of tubulin. Tubulin is primarily expressed in neurons and is commonly used as a marker of NSCs. It has been shown to be important for neuronal cell proliferation and differentiation (Von Bohlen Und Halbach). β-Tubulin staining is shown in Figure 6 Panel A. All NSCs generated exhibited similar staining for β-tubulin. NSCs were also stained with DAPI (Panel B) which was found to colocalize with β-Tubulin (Panel C).

Musashi

Rabbit monoclonal Musashi (Abcam) was used to stain the NSCs. Musashi is an RNA binding protein that regulates the expression of target mRNAs at the translation level. Musashi is expressed by neural progenitor cells during CNS development. It has also been shown to play a role in the proliferation and maintenance of stem cells in the central nervous system (Wilson and Stice). Representative staining of Musashi exhibited in all NSCs is shown (Panel D). NSCs were stained with DAPI (Panel E) and DAPI staining was observed to colocalize with Musashi (Panel F).

Nestin

Mouse monoclonal Nestin was used to stain the NSCs (Millipore). Nestin has been shown to have important roles in organization of the cytoskeleton, cell signaling, organogenesis and cell metabolism (Fuchs and Cleveland). Class VI nestin is expressed

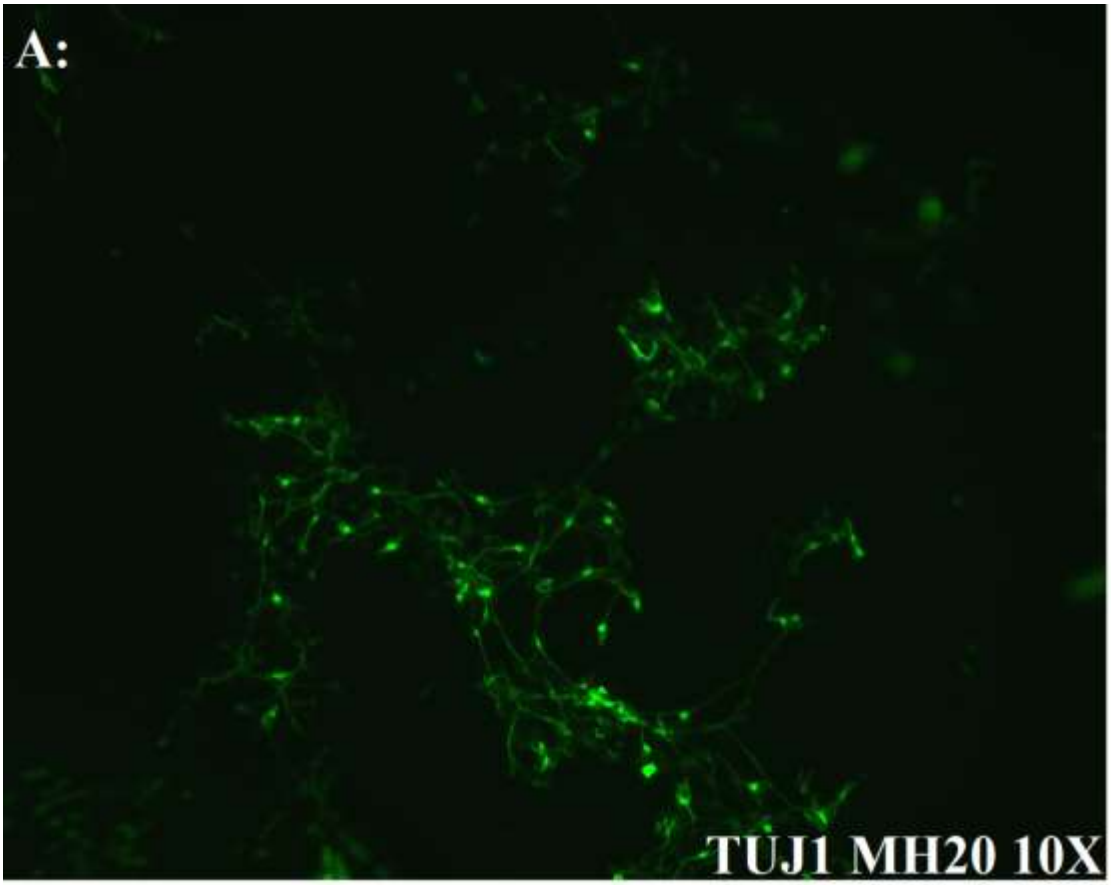
by neural progenitor cells during CNS development. The expression of class VI nesting is prevalent during early embryogenesis in neuroepithelial stem cells and later is absent in mature CNS cells (Lendahl, Zimmerman, and McKay). Nestin been shown to be excellent marker of neural stem cells and well suited for immunofluorescence labeling (Strojnik et al.). Nestin was readily observed to be present in the NSCs that were generated and found to colocalize with DAPI (Figure 6 Panel G-I).

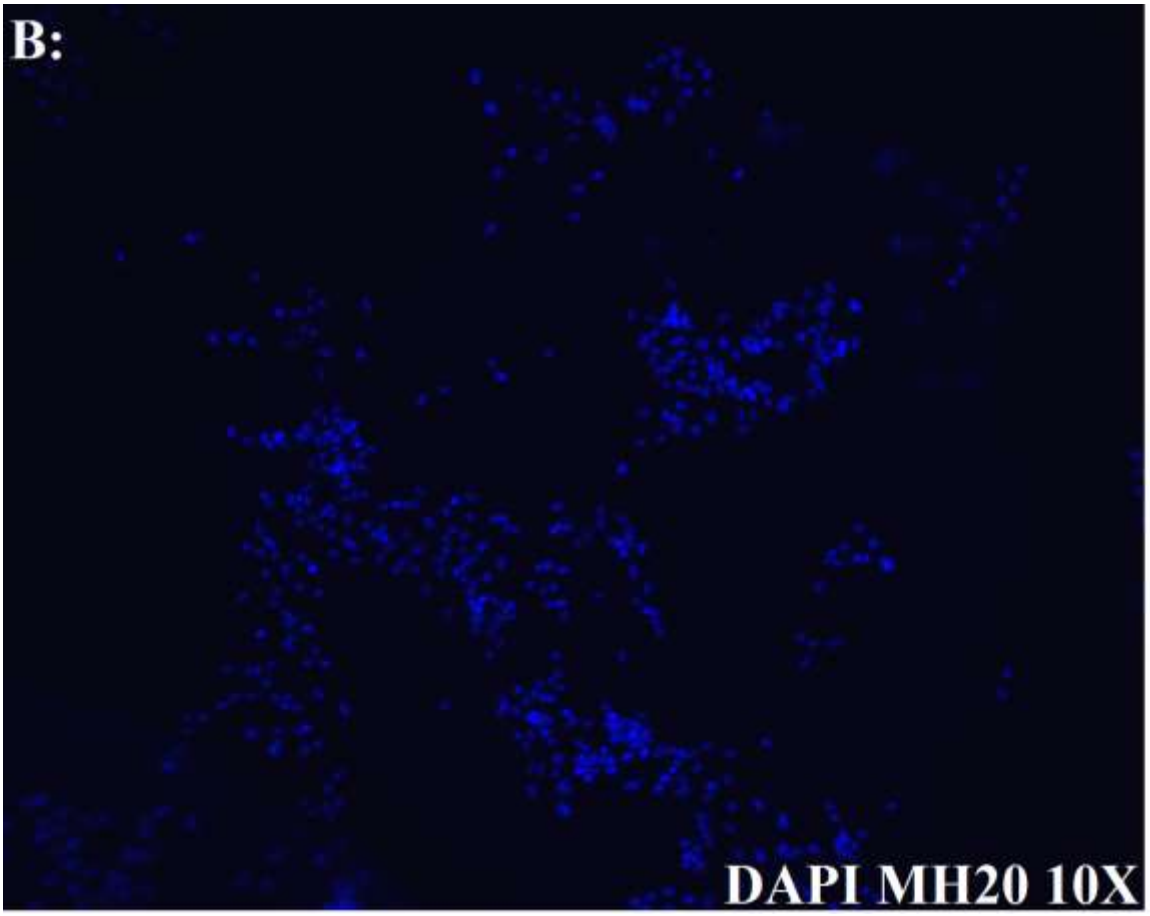
Pax-6

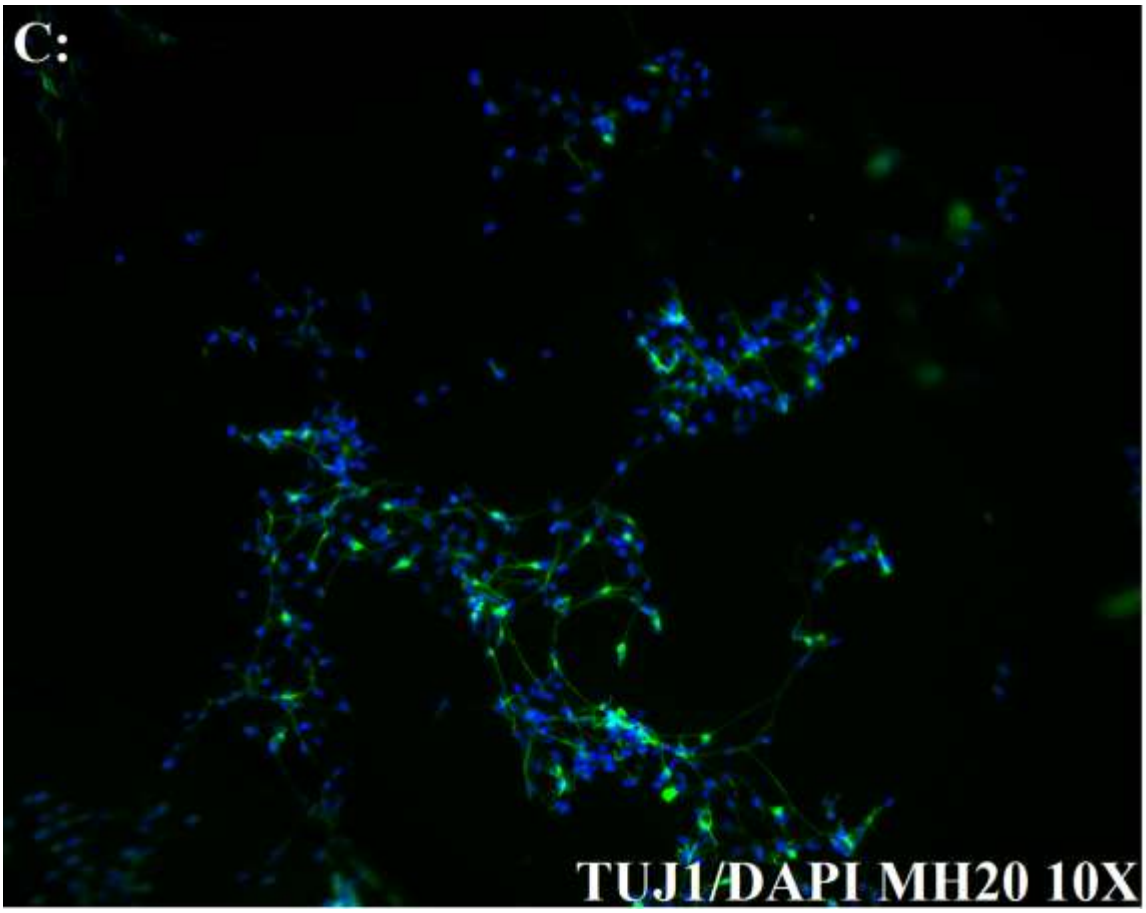
Rabbit Polyclonal Pax-6 was used to stain the NSCs (BioLegend). PAX genes encode nuclear transcription factors which are regarded as major controllers of developmental processes. Pax-6 has also been shown to be a surface market of NSCs(Yuan et al.).

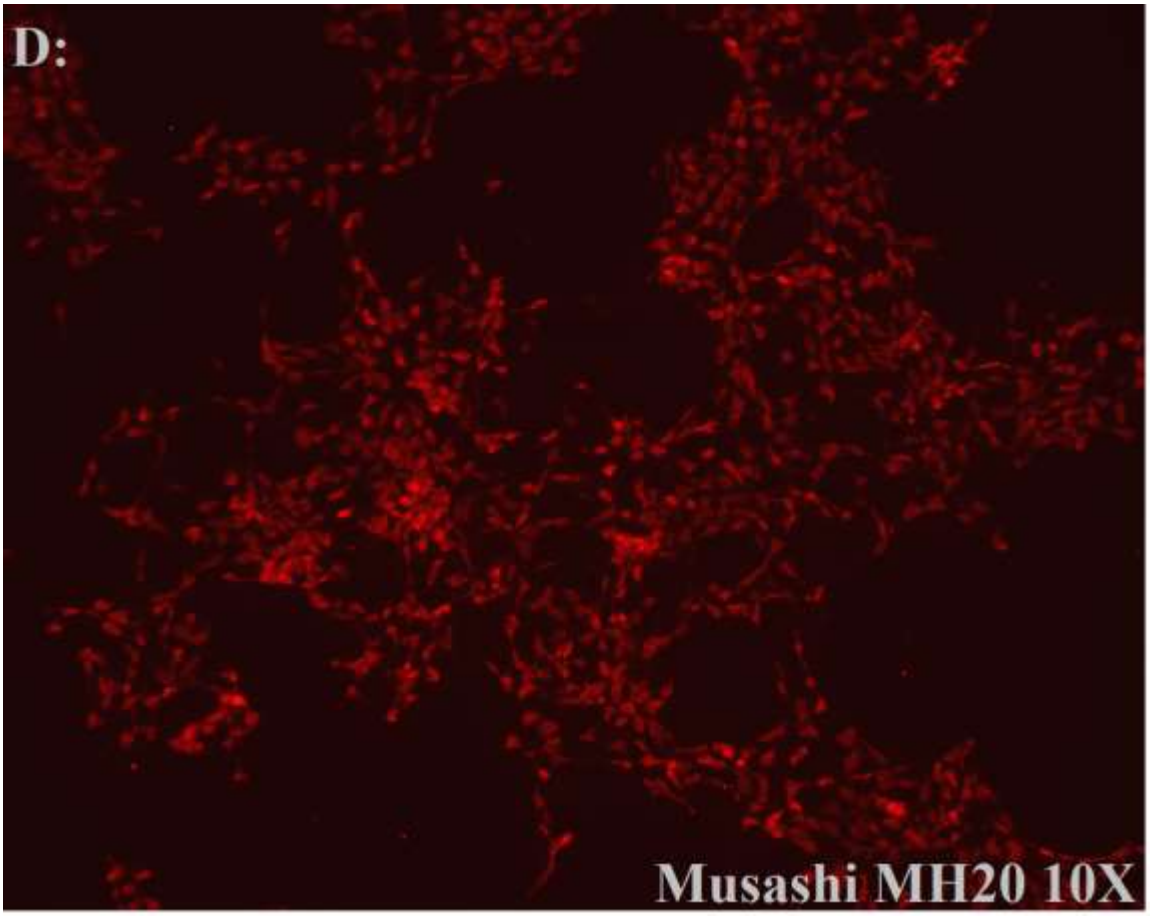
Purified Rabbit Polyclonal Pax-6 antibody was used to label the NSCs that were generated (BioLegend). A representative photo of Pax-6 staining observed in all NSCs in shown (Figure 6 Panel J &L). Pax-6 staining colocalized with DAPI staining on the NSCs (Figure 6 Panel K).

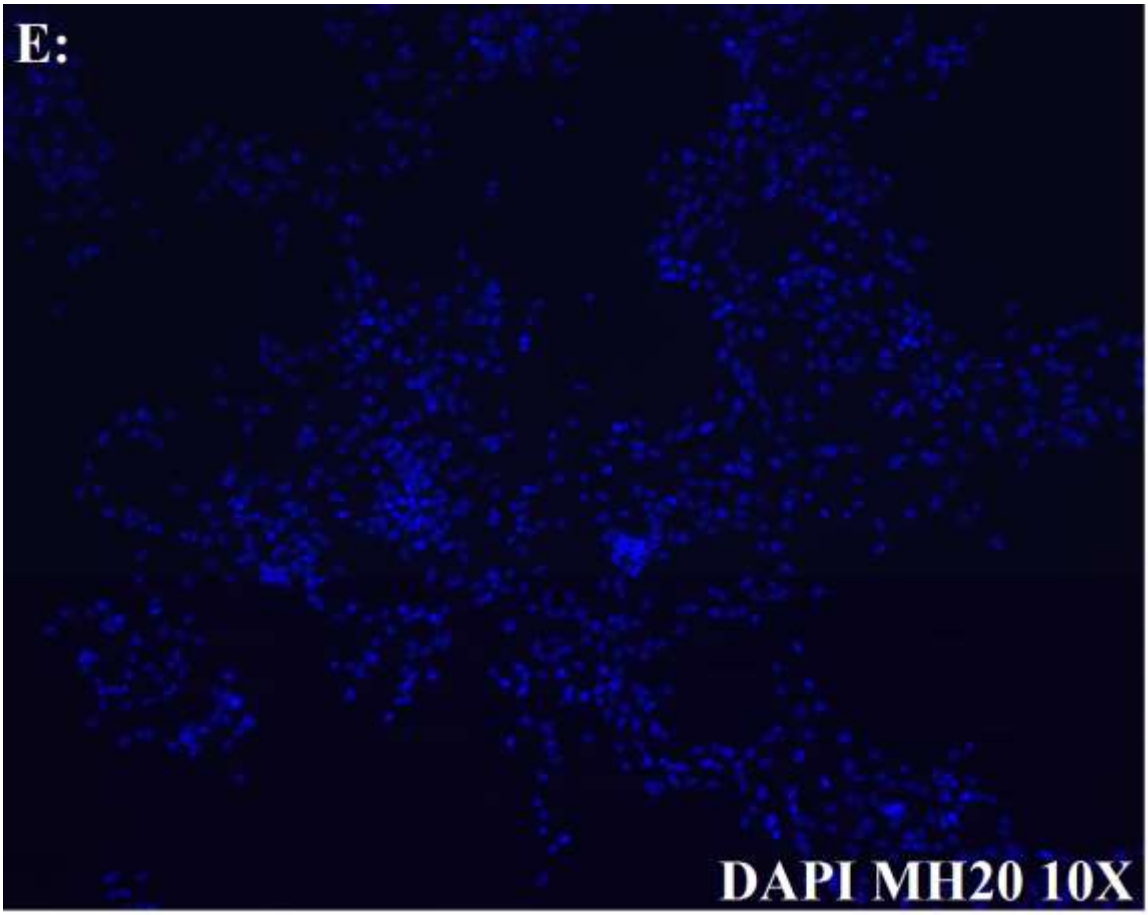
Figure 6: The following panels show immunofluorescent staining that was performed on the NSCs using antibodies for Pax6 [(Rabbit) Red 488], Nestin [(IgG1) Green 594], Tuji594 [(IgG2A) Green 594], and Mushai488 [(Rabbit) Red 488].All samples were stained with DAPI, a blue control marker for DNA

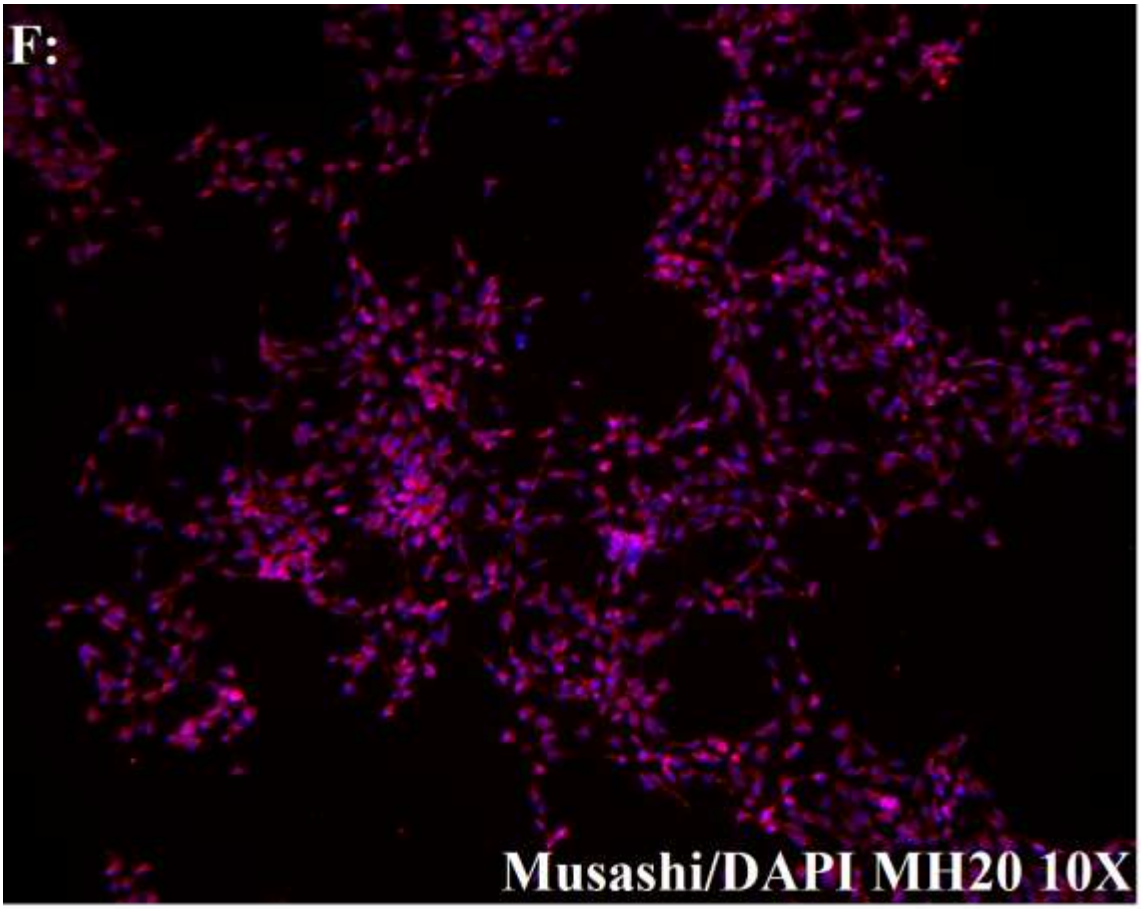


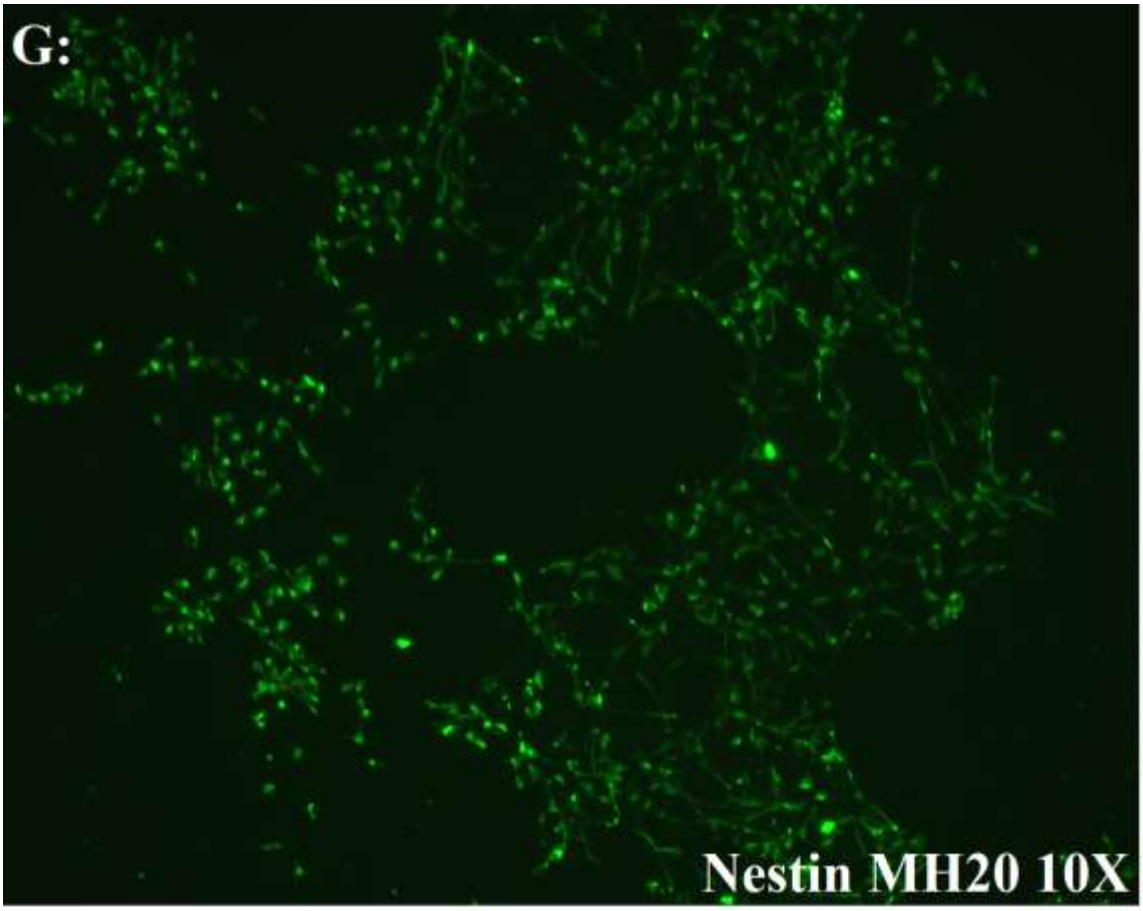

B:
DAPI MH20 10X

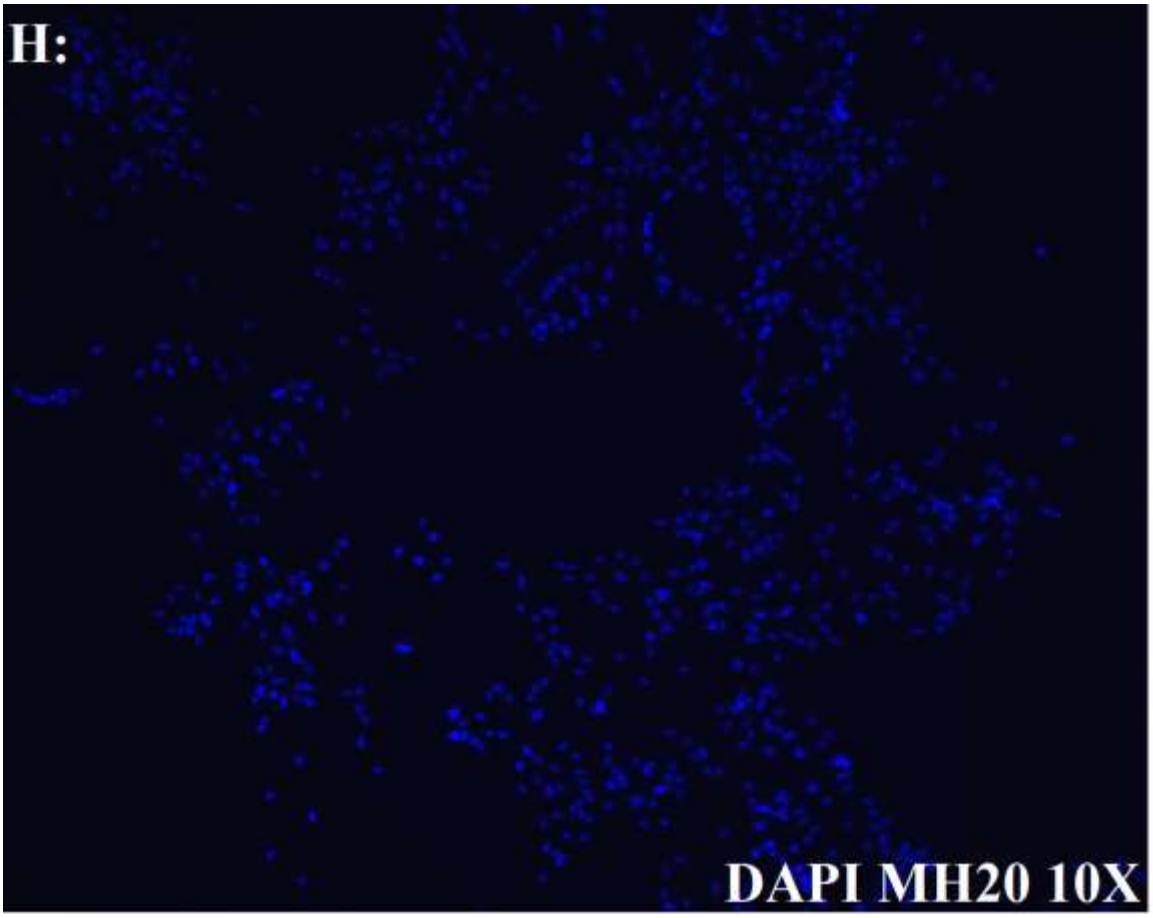

C:
TUJ1/DAPI MH20 10X

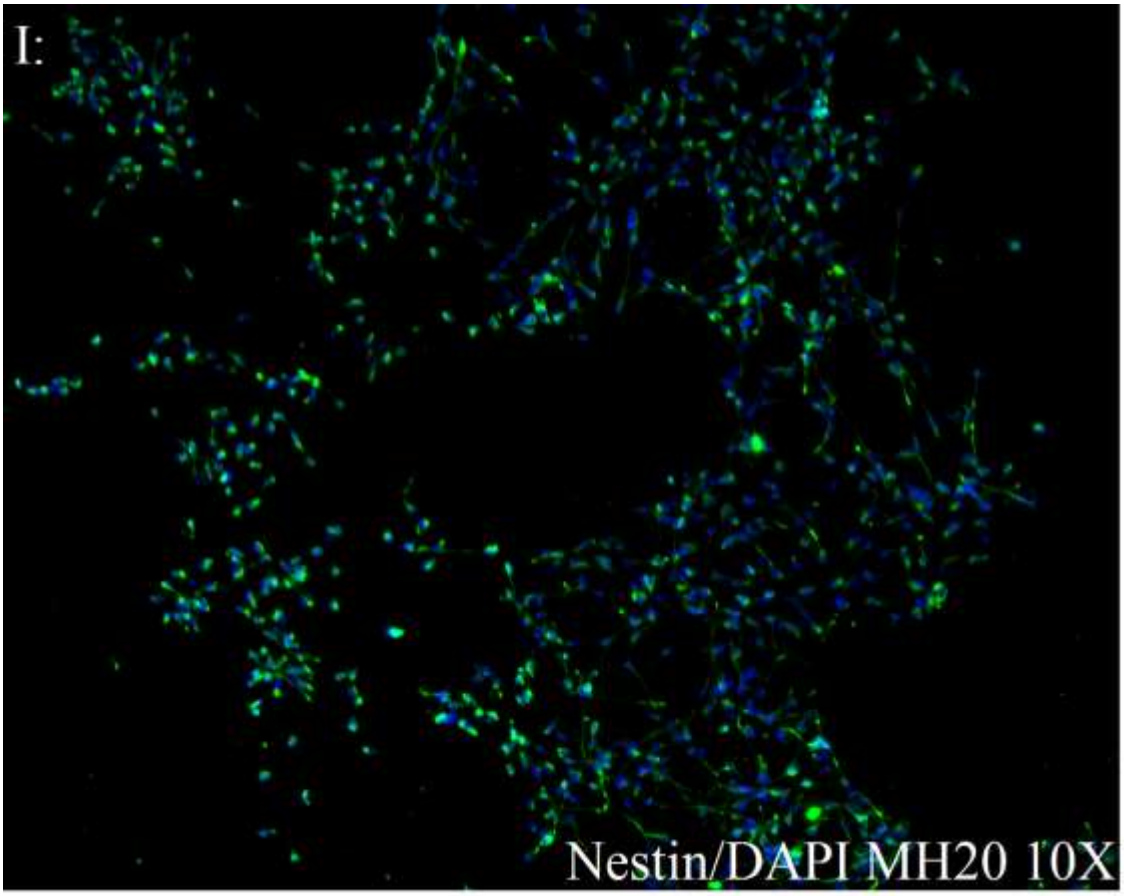

D:
Musashi MH20 10X

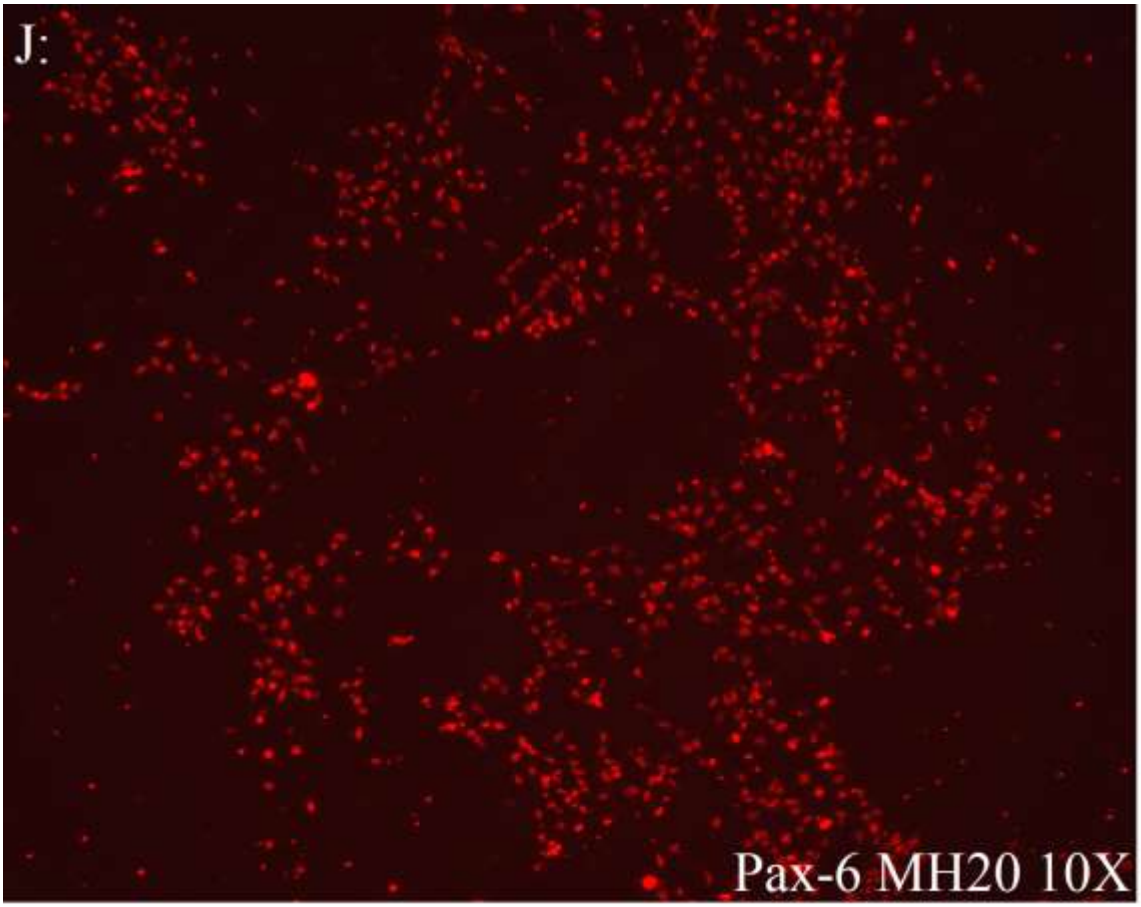

E:
DAPI MH20 10X

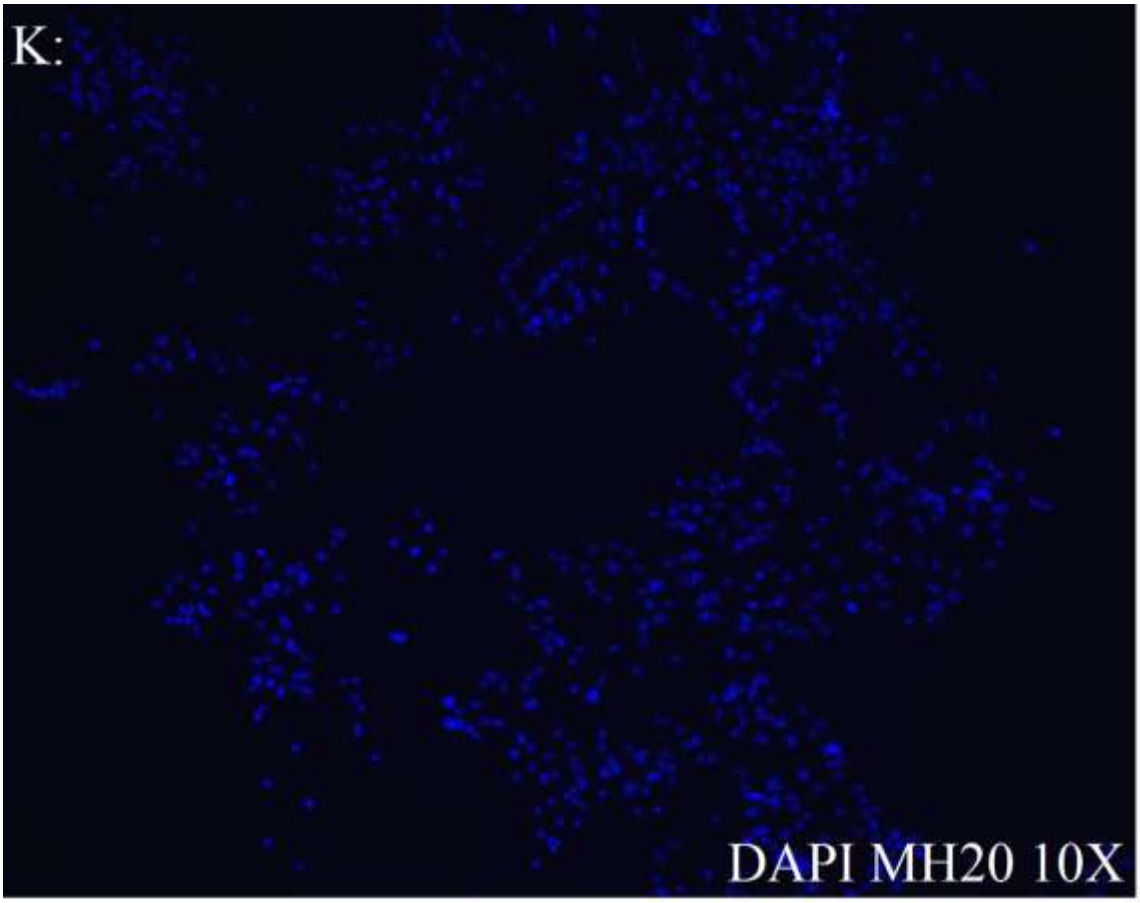

F:
Musashi/DAPI MH20 10X

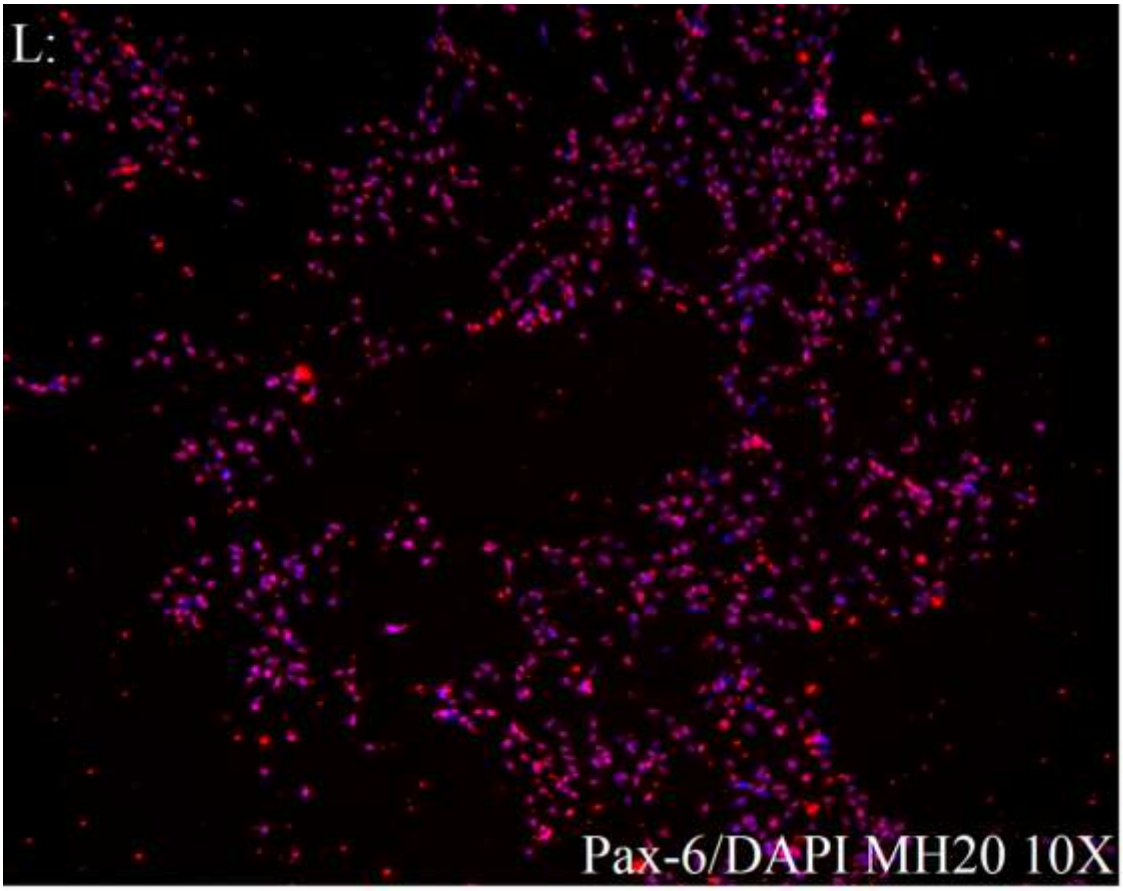

G:
Nestin MH20 10X


H:
DAPI MH20 10X


I:
Nestin/DAPI MH20 10X


J:
Pax-6 MH20 10X


K:
DAPI MH20 10X


L:
Pax-6/DAPI MH20 10X

Limitations of Neural Stem Cells

Attempting to generate neurons resulted in a mixed population of neuronal subtypes. Even if it were possible to isolate a particular neuronal subtype, it is difficult to determine which neuronal subtype would be the most sensible to use for further characterization. Characterization of miRNAs by both TaqMan® Human miRNA array card and small miRNA sequencing is extremely expensive and time consuming. It would be unrealistic given both budgetary and time constraints to profile multiple neuronal subtypes. Also, generating neurons as opposed to NSCs is adds additional expense. The neurons are also more difficult to work with than NSCs since they are not renewable and must be remade each passage. Rather than profiling miRNAs using mixed populations of neurons, our work focuses on profiling neural stem cells. Neural stem cell populations are more homogenous than the neuronal populations that we were able to obtain. Although we recognize that our neural stem cell populations do contain a small percentage of differentiated cells, this population is minute. For these reasons we have decided to perform all of our miRNA characterization on the NSC and not on differentiated neurons.

Identification and Development of Additional 22q11.2 iPSC Lines

In addition to the aforementioned twenty lines, additional 22q11.2 LCLs (lymphoblastoid cell lines) were found using SNP data which was downloaded from the National Institute of Mental Health (NIMH) Repository and Genomics Resource website. This SNP array data from 6135 subjects in the NIMH Repository and Genomics Resource Schizophrenia collection was screened for high homozygosity (near 100%) within base pairs 17,256,416 to 19,795,674 on chromosome 22 (the base pairs comprising the 22q11.2 deletion) (Table 4). DNA samples from eleven lines in the

Molecular Genetics of Schizophrenia (MGS) and three lines from the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) clinical trial were detected with homozygosity very close to or at 100% (Linda Brzustowicz, Personal Communication, Nov 12 2015). Additionally, SNPTrace data on all of these LCLs confirmed that these lines were all different from any of the existing 22q11.2 lines that were already part of this research (RUCDR).

DNA from those fourteen cell lines was obtained from RUCDR (Table 5). TaqMan® copy number variant assays were performed on these lines using the *DGCR8* probe and RNase P control probe. Further description of the TaqMan® Copy Number Assay principal and protocol is described in depth in Chapter 3. Ten of the fourteen lines were observed to have 1 copy of *DGCR8* after a minimum of three independent copy number assays were run each in quadruplicate (Figure 7). The presence of only one copy of *DGCR8* in these LCL lines suggested that these LCLs could have the 22q11.2 deletion. LCLs for these lines were obtained from RUCDR for further study.

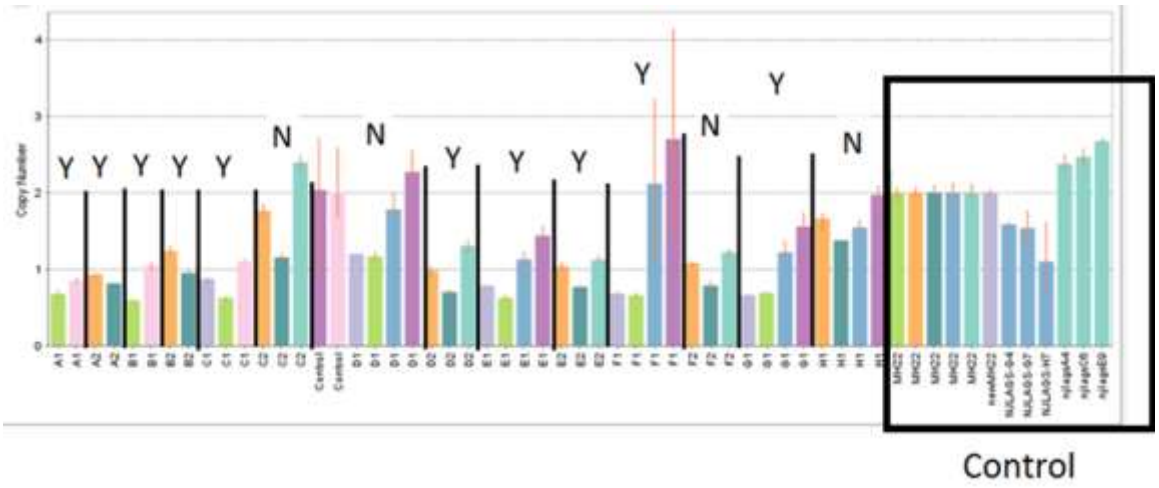
Table 4:

RUID	Percent Homozygosity	Origin
03C15933	0.942	CATIE
06C52254	0.921	MGS
06C58964	0.909	MGS
02C12950	0.976	CATIE
90C02141	0.98	MGS
03C21822	1	CATIE
06C56851	0.906	MGS
04C27348	0.915	MGS
04C24683	0.933	MGS
04C27536	0.868	MGS
06C57389	0.9	MGS
04C37291	0.904	MGS
05C39413	0.978	MGS
05C44882	0.883	MGS

Table 5: This table lists all of the LCL lines that were acquired. These lines are known to have schizophrenia and TaqMan® Copy Number assays suggest that they have the 22q11.2 deletion.

Obtained From:	**Cell Type:**	**Affected Status:**	Line:
The Molecular Genetics of Schizophrenia **(MGS)** and the Clinical Antipsychotic Trials of Intervention Effectiveness **(CATIE)** trial	Lymphoblastic Cell Lines	Presumed 22q11.2DS, confirmed schizoaffective cell lines	03C15933 (A1)
			06C52254 (B1)
			02C12950 (C1)
			03C21822 (D1)
			04C27348 (E1)
			04C27536 (F1)
			04C37291 (G1)
			05C39413 (H1)
			05C44882 (A2)
			06C58964 (B2)
			90C02141 (C2)
			06C56851 (D2)
			04C24683 (E2)
			06C57389 (F2)

Figure 7: This figure shows the results of TaqMan® Copy Number Assays performed on the previously extracted DNA obtained from the LCLs in Table 4.The DNA was compared to control DNA from an unrelated autism sample and the previously described MH22 line. All samples were probed for *DGCR8*. LCL lines that had only 1 copy of *DGCR8* were used for further study. Lines C2, D1, F2, and H1 seemed to have more than one copy of *DGCR8* and were dropped from this panel.



LCL Cell Culture:

Lymphoblastoid cell lines (LCL) and fibroblast lines were maintained using standard cell culture techniques and grown in Dulbecco's Modified Eagle's Medium (Thermo Fisher Scientific) supplemented with 10% Fetal Bovine Serum (Corning) and 1% Penicillin Streptomycin (Thermo Fisher Scientific).

Lymphoblastoid Cell Line to iPSC Transfection:

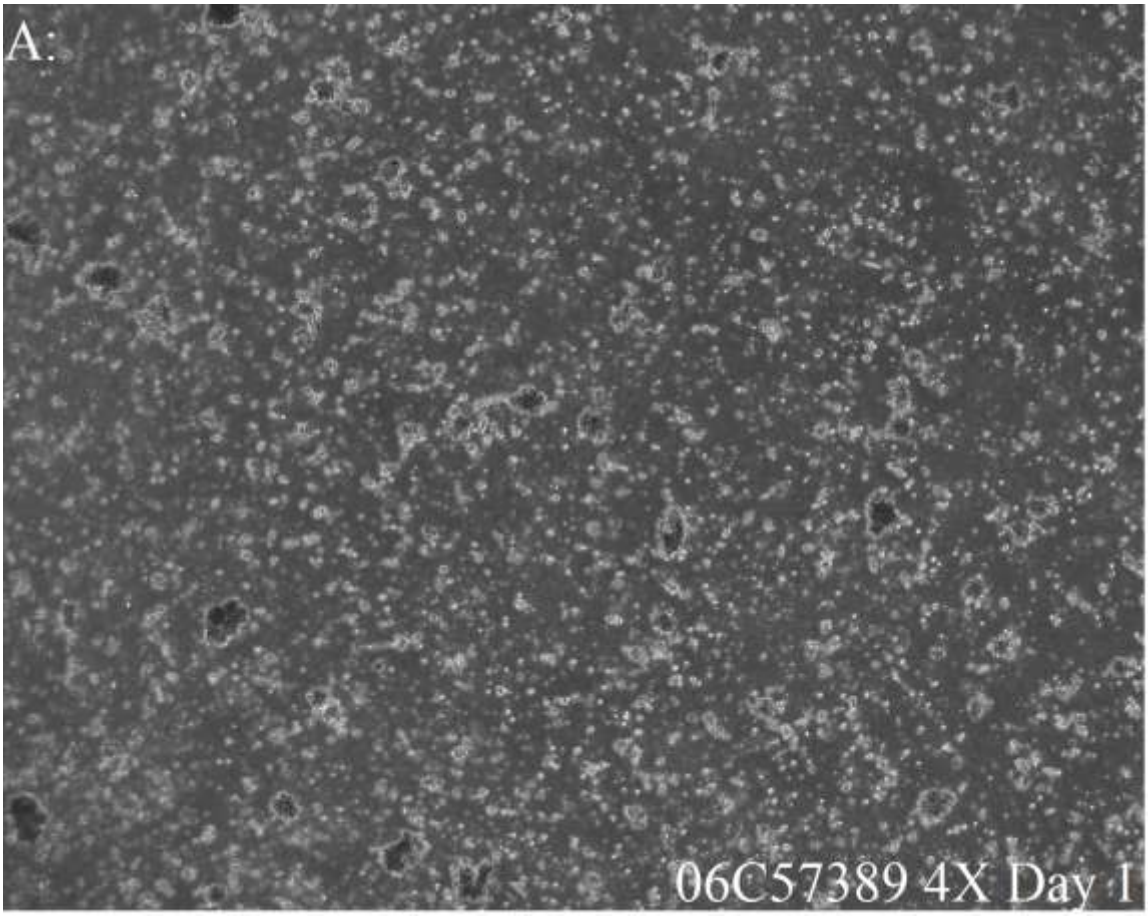
iPSCs were generated by reprograming LCLs using the Epi5 Episomal iPSC Reprogramming Kit (Thermo Fisher Scientific) according to "Conversion of LCL to iPSC by Epi5 Vectors" protocol (RUCDR) Prior to transfection, LCLs were cultured in 15% Roswell Park Memorial Institute medium (Thermo Fisher Scientific), 10% FBS (Corning) and 1% l-glutamine (Thermo Fisher Scientific) and grown in 75mm culture flasks. LCLs were collected in RPMI medium (Thermo Fisher Scientific) in a 15 ml centrifuge tube and then spun down at 300 RPM and 1 X 10^6 cells were resuspended in 100 µl of Amaxa Nucleofector Kit V buffer (Lonza) and transferred to a microcentrifuge tube. 1 µl of Epi5 Preprogramming vectors (1µg/µl) and 1 µl of Epi5 p53 (1µg/µl) and EBNA (1µg/µl) (Thermo Fisher Scientific) were added to the microcentrifuge. The mixture was transferred to a cuvette (Lonza). Electroporation was performed using the AMAXA Transfection System following the Nucleofector kit protocol (Lonza).

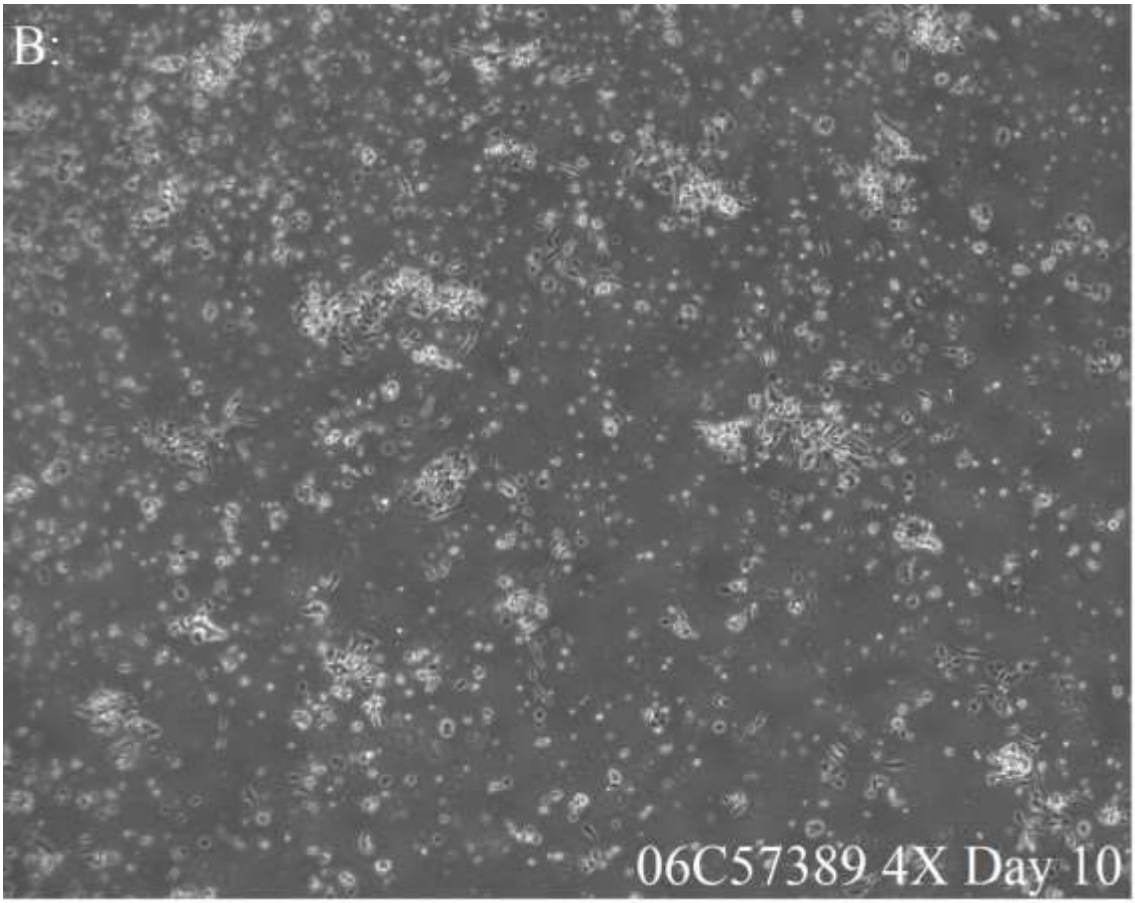
After electroporation, cells were washed in RPMI medium and plated onto a 12 well plate without Matrigel (BD Biosciences). On day 2 transfected LCLs were placed on 50 µl/cm^2 Matrigel (BD Biosciences)coated plates and given 1 ml of N2B27 medium (RUCDR). For the first week, the medium was not removed. Additionally 750 µl of N2B27 medium was added every other day. On day 8, all of the medium was aspirated

and fresh N2B27 Medium was added. Between day 8 and 15, N2B27 medium was replaced every other day. From day 16 onwards, the medium was replaced with 1/3 mTeSR and the rest of the media was N2B27. The mTeSR percentage was increased by 1/3 every other day until it reached 100% mTeSR. Colonies formed around day 21. Undifferentiated iPSC colonies were picked and placed into fresh Matrigel (BD Biosciences) coated plates. iPSCs colonies were than cultures as previously described earlier in chapter 2.

The ten lines that had only one copy of *DGCR8* detected in the TaqMan® Copy Number Assays were obtained as LCLs. Currently, iPSCs from one of these lines (06C57389) has been successfully reprogrammed into an iPSC line (Figure 8). TaqMan® Copy Number Assays have been performed on the iPSCs generated from passage 3 of LCL 06C57389. Only one copy of *DGCR8* was detected in 06C57389 (Table 6). Further characterization of this iPSC line, generation of NSCs from this cell line, and characterization of the miRNAs present in this line, are all planned for future experiments. Future experiments also include generating iPSC lines from the remaining LCLs.

Figure 8: Induction of iPSCs from LCLs was performed over the course of three weeks using the Epi5 Episomal iPSC Reprogramming Kit (Thermo Fisher Scientific) in conjunction with the Nucleofector Kit V (Lonza) and the AMAXA Transfection System (Lonza). Panel A shows the LCLs from line 06C57389 one day following transfection. The LCLs are not yet attached to the Matrigel coated plates. Panel B shows the same LCLs after 10 days following transfection. At this point the LCLs have started to attach to the bottom of the Matrigel coated plate. Small colonies are observed. The colonies are too small to be picked at this point. Panel C shows the cells from 06C57389 16 days following transfection. iPSC colonies can clearly been seen. These colonies are large enough that they are suitable for manual picking to a new plate. The iPSCs were picked from this plate and grown successfully on a new plate.


A:
06C57389 4X Day 1


B:
06C57389 4X Day 10

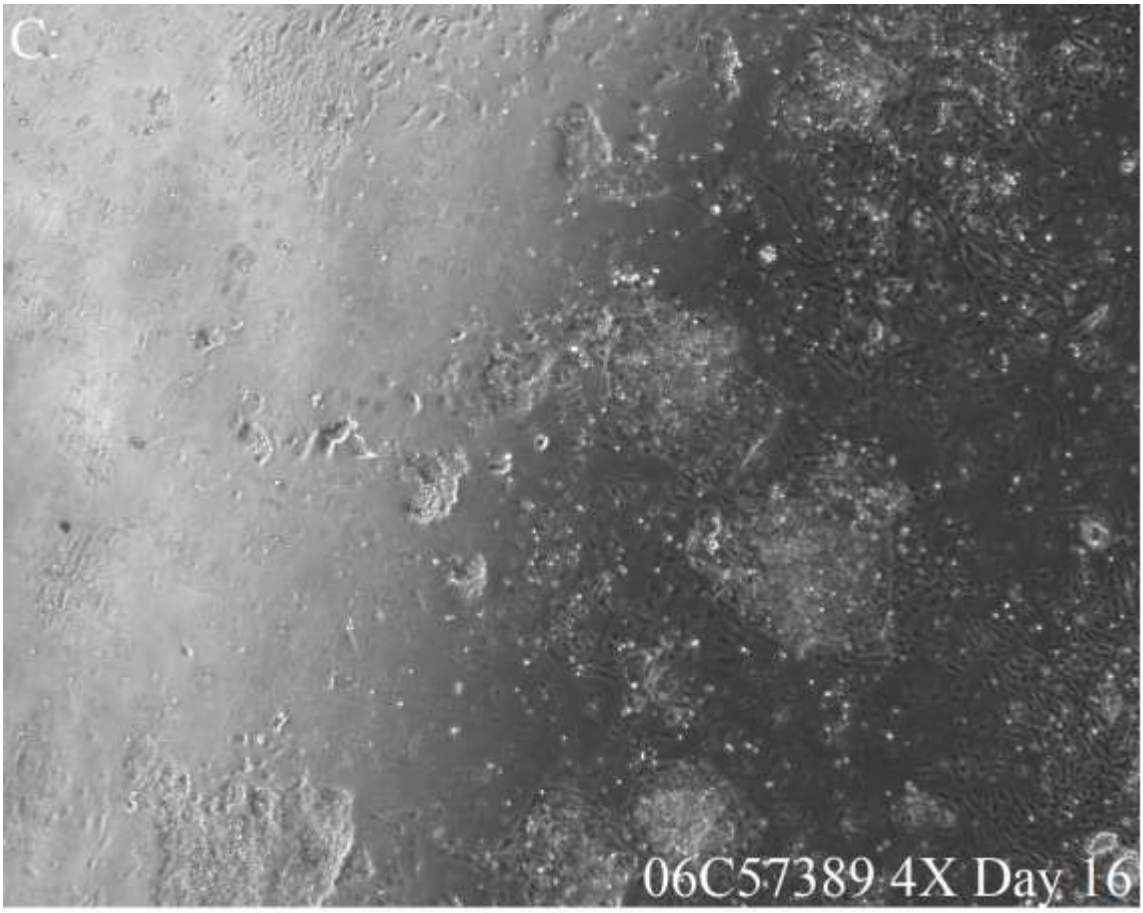

C:
06C57389 4X Day 16

Table 6: TaqMan® Copy Number Assays have been performed on iPSCs generated from LCL 06C57389. This assay probed for *DGCR8*. Cell line MH 509 was used as a control and does not contain the 22q11.2 deletion. Four replicates of this assay were performed on DNA isolated from passage 3 of iPSC 06C57389.

	DGCR8 Copy Number in 06C57389	Standard Deviation
Replicate 1	1.2	0.035118846
Replicate 2	1.18	0.179350681
Replicate 3	1.25	0.094295634
Replocate 4	1.21	0.132507861

Chapter 3: Stem Cell Quality Control (TaqMan® Assays, Low Coverage Sequencing, FISH)

Coriell Fibroblasts

Previous attempts to collect existing 22q11.2 cell lines for karyotyping have revealed that a number of lines found in cell repositories have had additional chromosomal abnormalities beyond the 22q11.2 deletion making these lines unsuitable for this research. The initial sample used as a model were nine cell lines obtained as fibroblasts (Coriell)(Table 7). Due to a scarcity of 22q11.2 lines, every effort was made to use these lines despite repeated problems in getting these lines to grow with normal morphology. Three of these lines were submitted for karyotyping and FISH analysis (Cell Line Genetics). GM03577 and GM05401 both had extensive chromosomal abnormalities (Table 7). These cell lines have all been passaged extensively before they were distributed to us. These chromosomal abnormalities were pre-existing, since we tested the earliest passages of these lines that we were able to obtain. Furthermore, the cell line GM03577 was found to have two copies of the 22q11.2 (TUPLE) region. This was inconsistent with its status as a DiGeorge Syndrome (22q11.2 deletion) line. GM05401 was found to have only one copy of TUPLE1 consistent with its status as a DiGeorge Syndrome (22q11.2 deletion) line. The control line GM22202 was found to have both a normal karyotype and normal FISH results consistent with 2 copies of 22q11.2 as expected.

Further experiments were suspended on the Coriell fibroblasts for several reasons. First, the chromosomal abnormalities detected in the karyotypes of the 22q11.2 deletion lines made these lines unsuitable for further study. It would be impossible to claim that

any miRNAs are dysregulated as a result of the 22q11.2 deletion specifically if there are a multitude of other chromosomal abnormalities that could be having an effect as well. Second, the 22q11.2 deletion was not detected in one of the fibroblast lines that was tested and was previously reported to have the 22q11.2 deletion. Although this could have been the result of an error in our cell culture techniques, all other cell lines that have been submitted for FISH analysis over the course of a 2 year period have matched the disease status that we have expected to see upon submission. This is the only inconsistency that we have observed in detecting the 22q11.2 deletion. It is possible that the line was contaminated before distribution with otherwise healthy cells. It is also possible that the deletion of the 22q11.2 region in this cell line did not contain TUPLE1, which is the probe used for FISH analysis. If the deletion in this line did not contain TUPLE1, no deletion would be able to be detected. Another possibility is that the original sample was mosaic and contained both cells with and without the 22q11.2 deletion. After multiple passages it would then be possible for cells without the deletion in the original sample to overtake the culture. Additionally, the Coriell cell lines were all very difficult to grow. The fibroblasts grew with unusual morphology and were observed to have unusually slow growth rates. Coriell is an external site and there was also difficulty in shipping the cells to our laboratory. Lastly, these lines were all fibroblasts. In order to study neural stem cells (NSCs) generated from these lines, we would have had to make the induced pluripotent stem cells (iPSCs) ourselves for all of the Coriell lines. The availability of preexisting iPSCS from RUCDR served as a much more attractive option for future research. Studies on the Coriell lines were suspended in 2014.

We were also cognizant of the fact that the cell lines, especially iPSCs, are prone to spontaneous chromosomal abnormalities. Every effort was made to identify these abnormalities and discard passages expressing abnormal chromosomal structure beyond the 22q11.2 deletion**.**

Table 7: The following table provides information about the cell lines obtained from Coriell. The Coriell ID Number, the phenotype, and the results from any karyotype or FISH analysis performed are shown below.

Coriell ID Number	Phenotype	Karyotype Results	TUPLE1 FISH Results
GM02944	DiGeorge Syndrome	n/a	n/a
GM03479	DiGeorge Syndrome	n/a	n/a
GM03577	DiGeorge Syndrome	46,XY,add(17)(q21),-20,inv(21)(q11.2q22.1),+mar[1] 47,XY,inv(21)(q11.2q22.1),+mar[1]	Normal
GM05401	DiGeorge Syndrome	46,XY,der(4)t(4;22)(q35;q11.2),der(6)t(4;6)(q31.3;q27),-22[20]	Abnormal 22q11.2 (TUPLE1)
GM10382	DiGeorge Syndrome	n/a	n/a
GM05399	Apparently Healthy	n/a	n/a
GM05756	Apparently Healthy	n/a	n/a
GM22202	Apparently Healthy	Normal	Normal
GM22252	Apparently Healthy	n/a	n/a

iPSC/NSC Quality Control

A major focus of this project has been generating iPSCs and NSCs and ensuring their high quality. Throughout all experiments it was imperative that the integrity of the cell lines be preserved. A rigorous effort was made to ensure that the lines were not comingled. Also, we have implemented multiple levels of repetitive testing to ensure that lines that were originally sourced as 22q11.2 deletion lines continue to display the deletion throughout the progression through our quality control workflow (Table 8).

Table 8 outlines the series of tests that was performed on all cell lines used in this research. Both the iPSCs and NSCs are extensively tested to confirm their respective cell types and check for the presence of the 22q11.2 deletion and other chromosomal abnormalities.

iPSCs	NSCs
AP Stain	Staining for NSC Markers
TaqMan® Gene Expression Assay	TaqMan® Gene Expression Assay
TaqMan® CNV Assay	TaqMan® CNV Assay
Karyotype	Low Coverage Sequencing
FISH	

Karyotyping

All of the iPSC lines obtained from RUCDR or generated from fibroblasts obtained from RUCDR were sent to an external laboratory for karyotyping (Cell Line Genetics). Cell Line genetics is a cytogenetic facility that specializes in performing G-band karyotyping on stem cells. G banding uses the Giemsa stain to stain condensed chromosomes and produce a visible karyotype. Chromosomes in metaphase are treated with trypsin to allow the Giemsa stain to take effect. Gene-poor regions rich in adenine and thymine stain more darkly, while condensed regions containing more guanine and cytosine stain less darkly.

Cell Line Genetics performed the karyotyping on all iPSC lines. Due to occasional difficulties in shipping live cultures, samples occasionally had to be resubmitted to Cell Line Genetics. Cell Line Genetics examined 20 cells per line. Karyotype results were obtained that were consistent with the results of other additional experiments characterizing the line in question for all cell lines.

It was important to submit the iPSCs for karyotyping to ensure chromosomal integrity. iPSCs are known to frequently acquire chromosomal anomalies in culture that can lead to invasive cell populations. Common abnormalities arising in cultured iPSCs include trisomies of chromosome 1, chromosome 8, chromosome 12, and chromosome 20 (Taapken et al.). These aneuploidies increase in rate as the passage of the cells increases. However, normal karyotypes at high passages and abnormal karyotypes at lower passages are possible and were observed in this research as later discussed. It was important that these iPSCs cell lines maintained a stable karyotype so karyotyping was

performed in conjunction with routine TaqMan® Copy Number Assays, which check for the chromosomal aneuploidies that are most common in iPSCs.

A representative example of a normal karyotype (MH 509) from a cell line that serves as a healthy control is shown as well as a normal karyotype from a cell line with 22q11.2 DS (MH 046) (Supplemental Figure 4).

FISH

All of the iPSCs lines used were sent to Cell Line Genetics for FISH (Fluorescence in Situ Hybridization) to be performed. FISH uses fluorescent probes that target a specific region of the chromosome. The probes are tagged with fluorophores, as well as targets for antibodies. The binding of these probes to a specific region of the chromosome is used to quantify the number of chromosomes that have that particular sequence. Chromosomes are tested in cells that are in interphase. This technique is useful for detecting specific deletions, trisomies, or other segmental aberrations (Langer-Safer, Levine, and Ward).

22q11.2 DS is confirmed by detecting a deletion in TUPLE1 (Weksberg et al.). In order to detect the 22q11.2 deletion N85A3(22q13.33) and TUPLE1 (22q11.21) provers were used. FISH results from all of the cell lines known to not have the 22q11.2 deletion showed that there were two copies of TUPLE1 present in a majority of the cells tested. FISH results from all 22q11.2 DS cell lines were found to have the deletion in a majority of the cells tested. There were some instances where the non-chromosomal aberrations were detected that are believed to be normal artifacts of cell culture (Supplemental Figure 4).

CNV Assays

In order to test for the deletion and common trisomies (Taapken et al.), TaqMan® copy number assays (Thermo Fisher Scientific) were routinely performed on all cell lines. TaqMan® copy number assays used DNA isolated using the PureLink® Genomic DNA Mini Kit (Invitrogen). The concentration of genomic DNA was measured using a NanoDrop 2000 UV-Vis Spectrophotometer.

A duplex real-time PCR is used to run both TaqMan® Copy Number Assays and TaqMan® Copy Number Reference Assays simultaneously (Thermo Fisher Scientific). The Copy Number Assay is the probe that is targeting the gene of interest. An additional Copy Number Assay is performed, called the reference assay, which targets a sequence that is known to have two copies of DNA for a specific region and act as a control. The copy number of each gene of interest is then determined by relative quantification using the CT (ΔΔCT) method, which measures the CT difference (ΔCT) between target and reference sequences. The ΔCT values of the test samples can then be compared to a calibrator sample, which is a sample that is known to have two copies of gene of interest, in this case a sample that does not express the 22q11.2 deletion.

The copy number assay was performed by adding into each well of a 384 well plate: 10 ng of purified genomic DNA, .5µl of the TaqMan® copy number assay that contains two primers and a FAM™ dye labeled MGB probe to detect the genomic DNA target sequence of interest, .5µl of the TaqMan® copy number reference assay, which contains two primers and a VIC® dye-labeled TAMRA™ probe to detect the genomic DNA reference sequence, 5µl of the TaqMan® Genotyping Master Mix, which contains AmpliTaq Gold® DNA Polymerase, UP (Ultra-Pure) and the necessary dNTPs required for the PCR reactions. This recipe was repeated for each different copy number probe and

each different DNA sample that was used. Each copy number assay was performed in quadruplicate for every run in a 384 well plate. Each cell line and cell type was tested a minimum of three separate times, consisting of three independent copy number assays performed on different DNA samples from the same cell line over multiple passages, each run in quadruplicate.

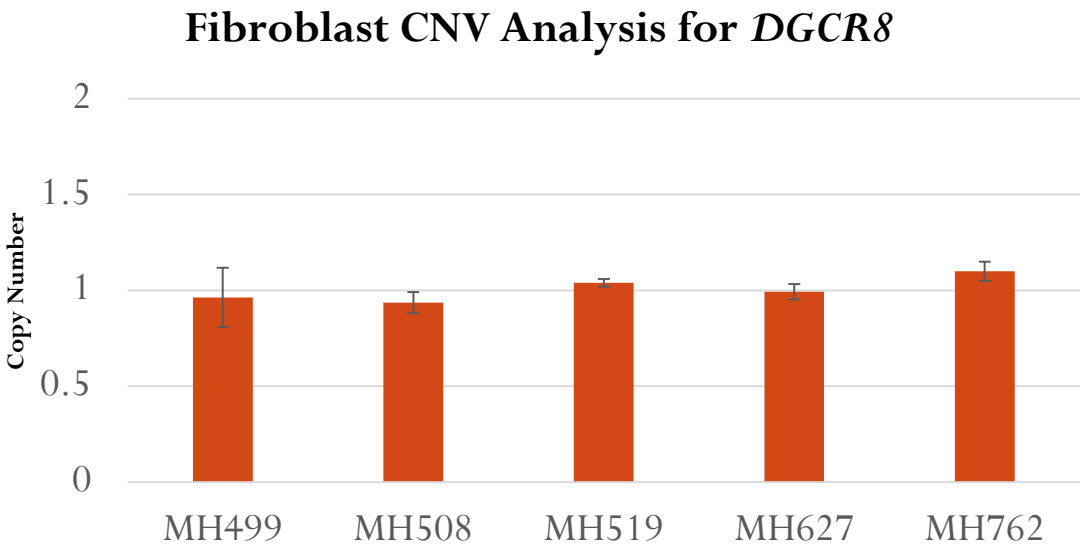
TaqMan® Copy Number Assays (Thermo Fisher Scientific) were performed on DNA extracted from fibroblasts, iPSCs, and NSCs. Copy number assays were performed using six probes known to be located copy number stable regions in chromosomes that are commonly susceptible to chromosomal abnormalities as a result of iPSC culture: *BCL2L1* (Hs05567439_cn)(Chr.22:18207651), *DGCR8* (Hs00173864_cn) (Chr.22:20073855), *GDF6* (Hs00803962_cn)(Chr.8:97154900), *LAD2* (Hs00194326_m1)(Chr.1: 201349966), *SMURF2*(Hs05483324_cn) (Chr.17:62594890), and *SDC4*(Hs07181769_cn)(Chr.20:43971608)(Kent et al.). The Copy Number Reference Assay detects a sequence that is known to exist in two copies in a diploid genome, in this case the Ribonuclease P RNA component H1 (H1RNA) gene (RPPH1) on chromosome 14, cytoband 14q11.2. The assay location is chr.14:20811565 on NCBI build 37. It has an 87 bp amplicon that maps within the single exon RPPH1 gene.

The real-time PCR reaction was run using a 7900HT System using 384 well plates sealed with MicroAmp® Optical Adhesive Film (Thermo Fisher Scientific). Results from the copy number analysis were analyzed using Copy Caller Software (Life Technologies). In lines known the have the 22q11.2 deletion, *DGCR8* was observed to be present at only one copy. In cell lines known to not have the 22q11.2 deletion *DGCR8*, 2 copies of *DGCR8*

were detected. All of the other probes were chosen because they target chromosomes that are known to develop spontaneous trisomies as a result of iPSC culture conditions.

TaqMan® Copy Number Assays were performed on the fibroblast samples that were obtained from RUCDR. Since fibroblasts are not known to generate spontaneous trisomies in culture, the TaqMan® Copy Number Assays were performed using only the *DGCR8* probe. This was done to validate that the lines did indeed express the 22q11.2 deletion before time and resources were spent generating iPSCs of these lines. All assays were run in quadruplicate as previously described. Figure 9 shows that ~1 copy of *DGCR8* was detected using DNA extracted from 22q11.2 fibroblast lines. Values slightly higher or lower than 1 are the result of the inherent variability in the assay.

Figure 9: The figure below shows the results of TaqMan® CNV Analysis performed on DNA extracted from fibroblasts that were obtained from RUCDR. The DNA was probed for *DGCR8*. All 5 lines were distributed as 22q11.2 deletion lines. Samples were run in quadruplicate and error bars represent the standard deviation between these replicates.



TaqMan® Copy Number Assays were performed on all 20 of the iPSCs lines that were used in this research. All assays were run in quadruplicate as previously described. Figure 10 shows a representative assays showing that 1 copy of *DGCR8* was detected using DNA extracted from 22q11.2 iPSC lines, while 2 copies of *DGCR8* were detected in lines known to not have the 22q11.2 deletion.

Control, 22q11.2, and schizophrenia lines generally have values of 2 for the TaqMan® Copy Number Assays performed on iPSCs for probes LAD2, *BCL2L1*, *SDC4*, *GDF6*, and *SMURF2* (Figures 12-16). There were several examples where probes were found to have an elevated copy number of 2.5 or 3 and reductions in copy number to 1.5 as shown throughout Figures 12-16. These instances were not detected in karyotypes, FISH, or low coverage sequencing. Furthermore, they were not detected in later analysis of the NSCs. These variations could be the results of the combined effects of natural biological variability within the samples, the differences in sample passages and plating density for which the samples were collected, or minor differences in RNA quality, or reverse transcription efficiency.

Figure 10: This figure below shows the CNV assay results using a TaqMan® probe for *DGCR8* (Chromosome 22) on DNA extracted from iPSCs. Control lines are shown in green. 22q11.2 lines are shown in red. Schizophrenia only lines are shown in blue. All control and schizophrenia lines had values of ~2 which indicated that two copies of *DGCR8* are present in the control cell lines. All 22q11.2 lines have values of ~1 indicated that these lines only have one copy of DGCR8. Samples were run in quadruplicate and error bars represent the standard deviation between these replicates.

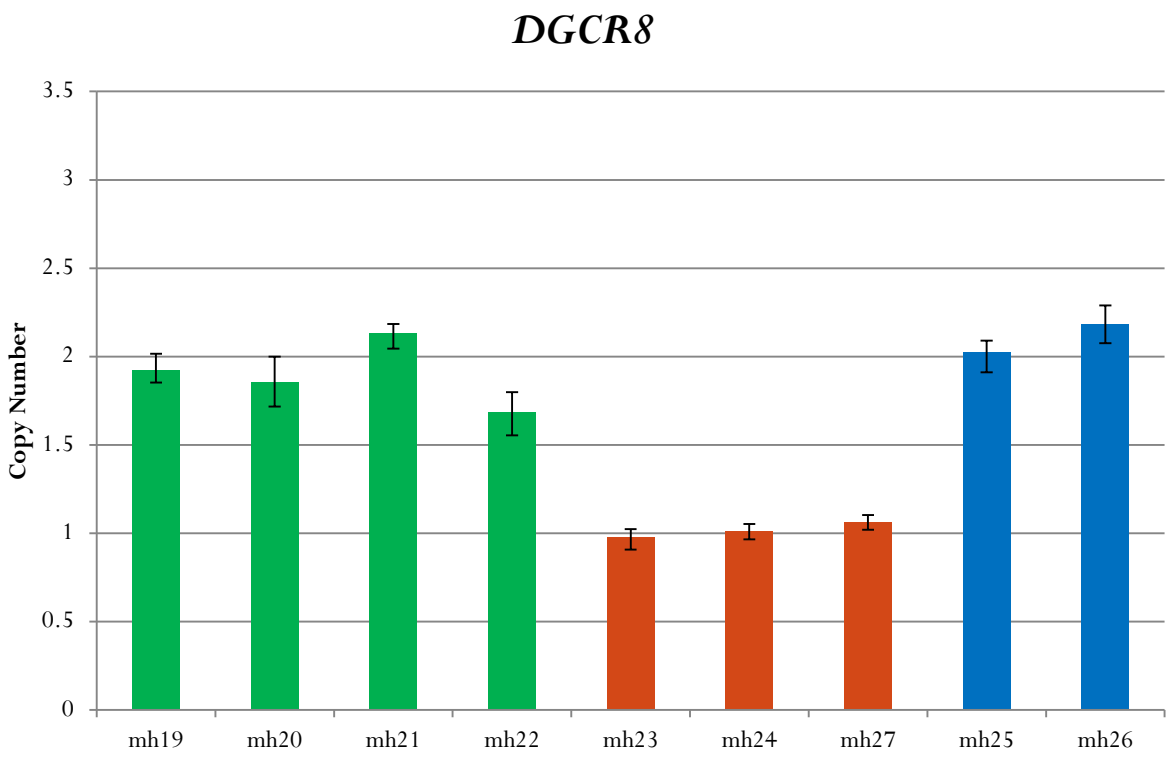


Figure 11: This figure below shows the CNV assay results using a TaqMan® probe for LAD2 (Chromosome 1) testing DNA isolated from iPSCs. Control lines are shown in green. 22q11.2 lines are shown in red. Schizophrenia only lines are shown in blue. All control, 22q11.2, and schizophrenia lines have values of ~2 indicated that two copies of LAD2 are present in all cell lines. Elevations in the copy numbers in MH 20 and MH 27 were not detected in later NSCs profiled for these lines. Samples were run in quadruplicate and error bars represent the standard deviation between these replicates.

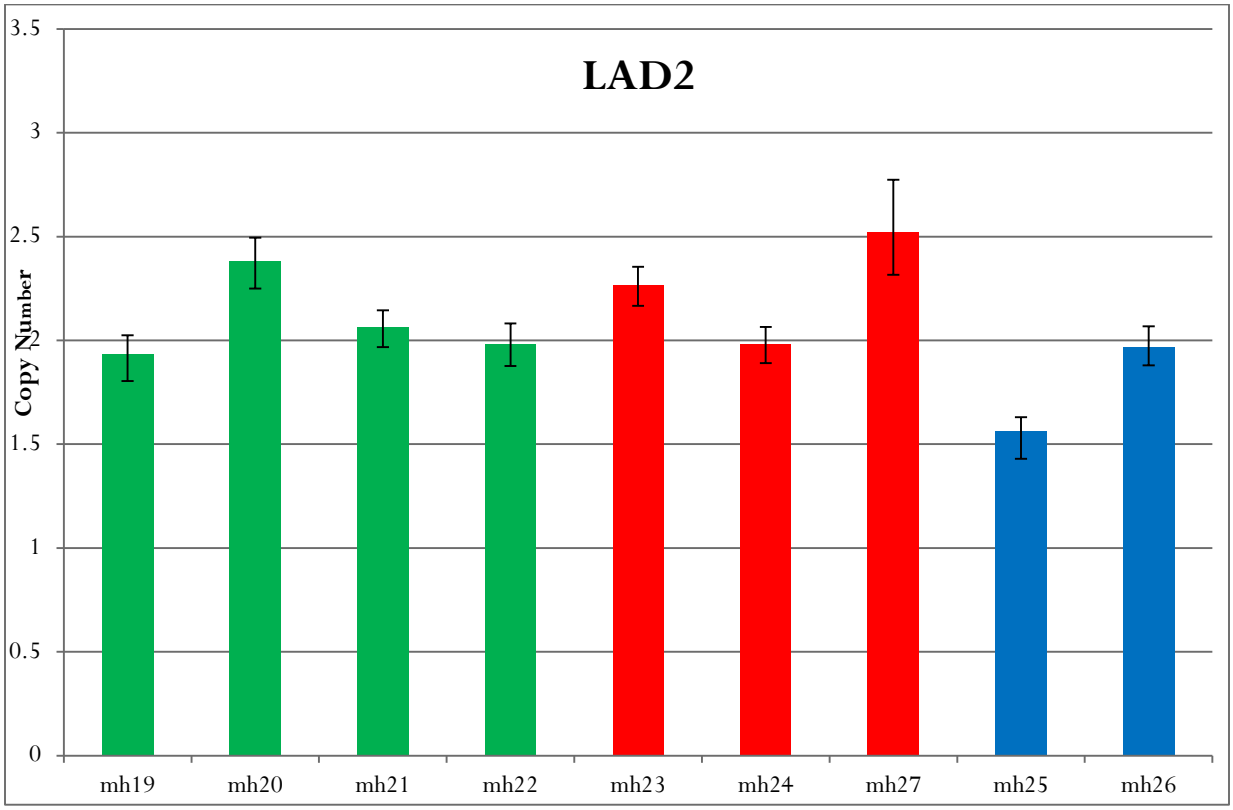


Figure 12: This figure below shows the CNV assay results using a TaqMan® probe for *BCL2L1* (Chromosome 20) on DNA isolated from iPSCs. Control lines are shown in green. 22q11.2 lines are shown in red. Schizophrenia only lines are shown in blue. Most cell lines have the expected 2 copies of *BCL2L1*. Cell lines MH 19, MH 21, and MH 26 do have elevated copies of *BCL2L1*. These elevated levels of *BCL2L1* were not detected in later analysis of the NSCs. Samples were run in quadruplicate and error bars represent the standard deviation between these replicates.

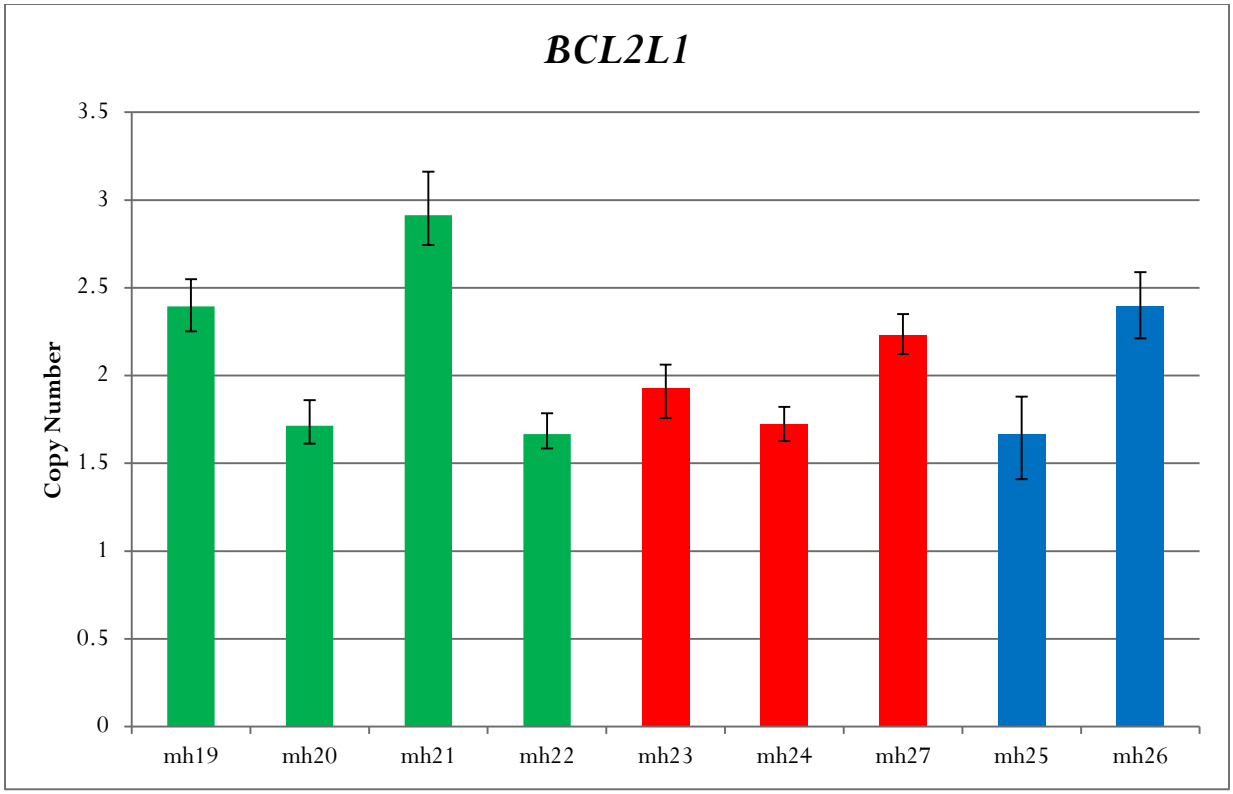


Figure 13: This figure below shows the CNV assay results using a TaqMan® probe for *SDC4* (Chromosome 20) on DNA isolated from iPSCs. Control lines are shown in green. 22q11.2 lines are shown in red. Schizophrenia only lines are shown in blue. All cell lines have the expected 2 copies of *SDC4*. The unexpected reduction in MH25 is likely the result of biological variability. No alterations in MH25 were detected in later characterization of NSCs derived from this line. Samples were run in quadruplicate and error bars represent the standard deviation between these replicates.

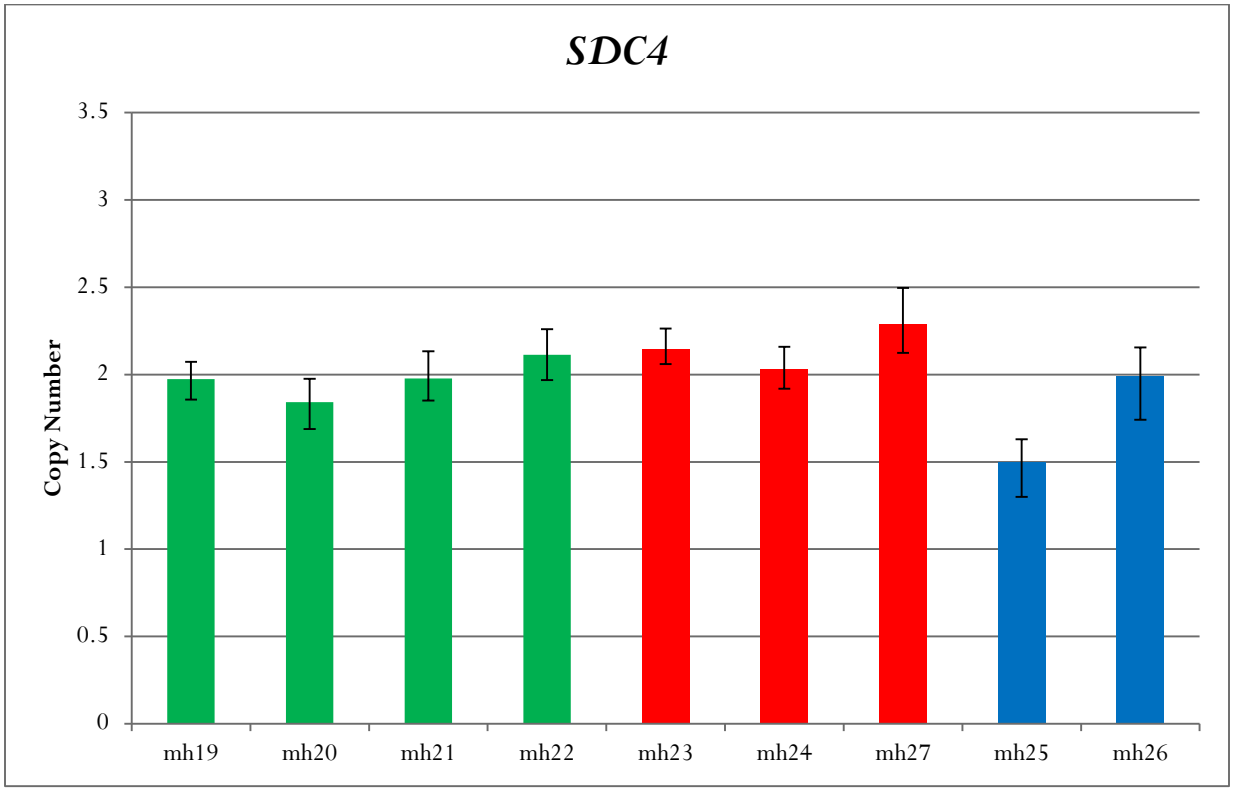


Figure 14: This figure below shows the CNV assay results using a TaqMan® probe for *GDF6* (Chromosome 8) on DNA isolated from iPSCs. Control lines are shown in green. 22q11.2 lines are shown in red. Schizophrenia only lines are shown in blue. All cell lines have the expected 2 copies of *GDF6*. Elevations in MH 23 and MH 27 were not detected in later characterization of the NSCs derived from these lines. Samples were run in quadruplicate and error bars represent the standard deviation between these replicates.

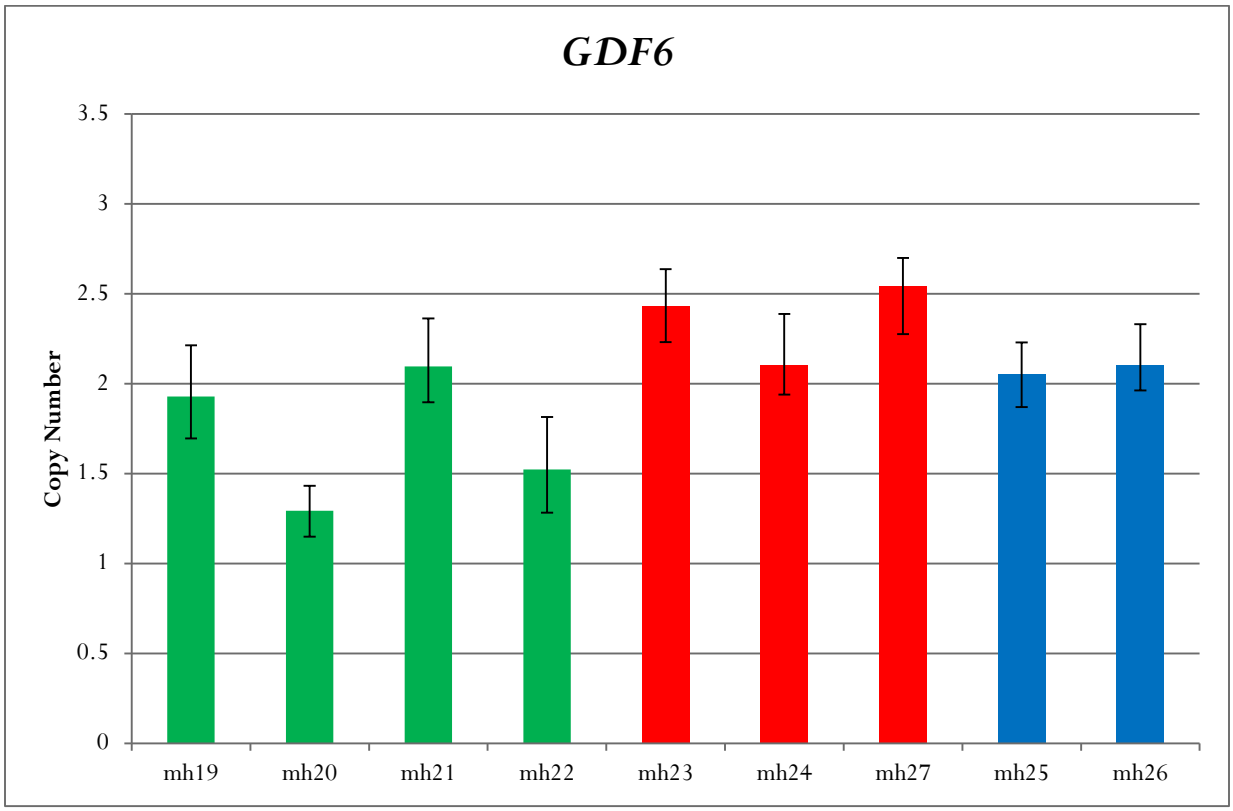


Figure 15: This figure below shows the CNV assay results using a TaqMan® probe for *SMURF2* on DNA isolated from iPSCs (Chromosome 17). Control lines are shown in green. 22q11.2 lines are shown in red. Schizophrenia only lines are shown in blue. All cell lines have the expected ~2 copies of *SMURF2*. The increases in levels of *SMURF2* for MH 19 and MH 21 as well as reductions in MH 20 and MH 22 were not detected in later characterization of this line as an NSC. Samples were run in quadruplicate and error bars represent the standard deviation between these replicates.

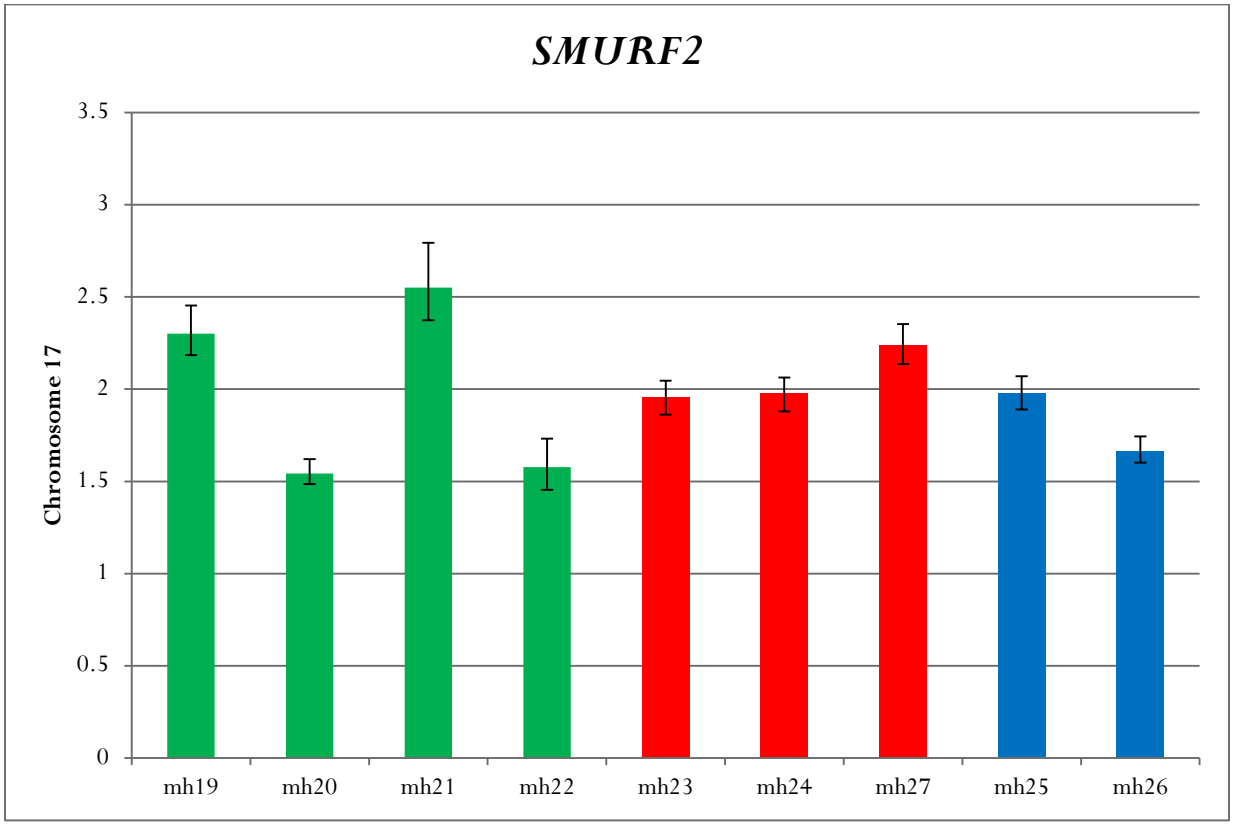
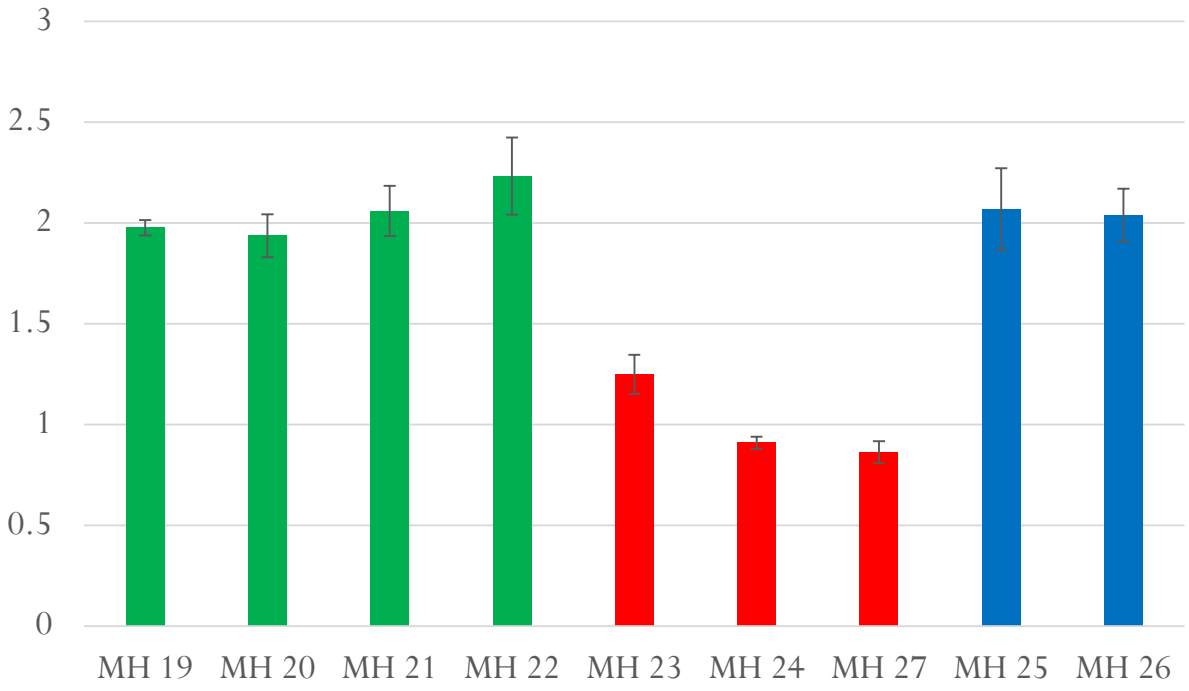


Figure 16: This figure below shows the CNV assay results using a TaqMan® probe for *DGCR8* using multiple NSC lines. Control lines are shown in green. 22q11.2 lines are shown in red. Schizophrenia lines are shown in blue. Samples were run in quadruplicate and error bars represent the standard deviation between these replicates.



TaqMan® CNV assays were performed routinely for all cell lines and cell types. Copy number assays were performed on NSCs in addition to the iPSCs to confirm that there was no change in the copy number between cell types. The results of copy number assays testing *DGCR8* levels in NSCs generated from the previously described iPSCs is shown (Figure 16). 22q11.2 lines continue to have only one copy of *DGCR8*. Control lines and schizophrenia-only lines continue to have 2 copies of *DGCR8*.

Gene Expression Assays

Another test used to determine if the cell lines contain the 22q11.2 deletion is the TaqMan® Gene Expression Assay. TaqMan® Gene Expression Assays target protein coding transcripts and allow their expression to be quantified. Each Gene Expression Assay consists of two unlabeled primers (20X), a 6-FAM™ dye-labeled MGB probe (20X), and one 6-VIC® dye-labeled TaqMan® MGV probe (20X). TaqMan® probes were used for *DGCR8* (Chromosome 22), *RPL9* (Chromosome 4), *TNP1* (chromosome 2) and *TNP2* (chromosome 16). These probes were predesigned real-time PCR assays. *RPL9*, *TNP1*, and *TNP2* were chosen as endogenous controls. They were used to normalize for differences in sample RNA added to the reaction.

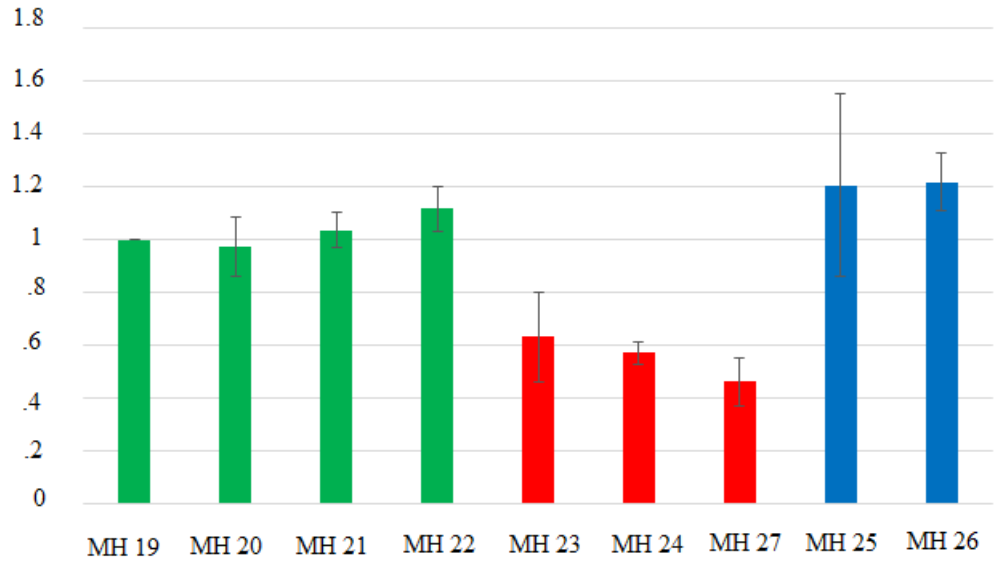
The Gene Expression assays were performed using RNA that was isolated from cells as a template for the synthesis of cDNA, using the RNeasy Mini kit (Qiagen). The total RNA used was 2 µg (for a 20-µL reaction). The concentration of extracted RNA was estimated using a NanoDrop 2000 UV-Vis Spectrophotometer. The extracted RNA was used to generate cDNA using a High Capacity cDNA Reverse Transcription Kit (Thermo Fisher Scientific).Using the cDNA Reverse Transcription Kit, a 2X RT master mix (per 20-µL reaction) was made that consisted of 2 µL of 10X RT Buffer µL .8 µL 25X dNTP Mix (100 mM), 2 µL of 10X RT Random Primers, 1 µL of MultiScribe™ Reverse Transcriptase, 1 µL of RNase Inhibitor and 3.2 µL of nuclease-free H2O Reverse transcription was performed in 96 well plates using a PTC-200 Peltier thermal cycler (BioRad) and following the manufacturer's protocol (Thermo Fisher Scientific).

The cDNA was then used to run the real-time PCR reaction. 100 ng of cDNA was used per 20- µL amplification reaction. Each well of the real-time PCR reaction consisted of 1 µL of a 20X TaqMan® Gene Expression Assay (either *DGCR8* or one of the endogenous controls), 10 µL of 2X TaqMan® Gene Expression Master Mix, 4 µL of cDNA template and 5 µL of RNase-free water. Gene expression was determined by measuring the expression of the fluorescent TaqMan® probes.

The real-time PCR reaction was run using a 7900HT System using 384 well plates sealed with MicroAmp® Optical Adhesive Film (Thermo Fisher Scientific). Gene expression values were calculated and analyzed using the program SDS 2.3 and Data Assist vs 2.0 programs (Thermo Fisher Scientific). Relative quantification using the Comparative CT method was used to calculate the expression of *DGCR8* compared to the relative expression of *TNP1*, *TNP2*, and *RPL9*. This method compares the Ct value of the target gene to that of control genes (using the formula: $2\Delta\Delta CT$) in order to quantify the expression of the target gene relative to the expression of the control gene.

Gene Expression Assays performed on control lines showed that expression of *DGCR8* was ~1. MH1059020, MH1059021, and MH1059022 differed only slightly from MH1059019 which was set at 1. The expression levels of *DGCR8* were ~50% in 22q11.2 DS cell lines. This indicates that there is only one copy of *DGCR8* present in these cell lines and the expression of *DGCR8* in the remaining copy has not been increased as part of a compensatory measure. The expression levels of *DGCR8* in schizophrenia-only cell ines was observed to be ~1. This indicates that the expression of *DGCR8* is not significantly different between non-deletion cell lines regardless of schizophrenia status (Figure 17).

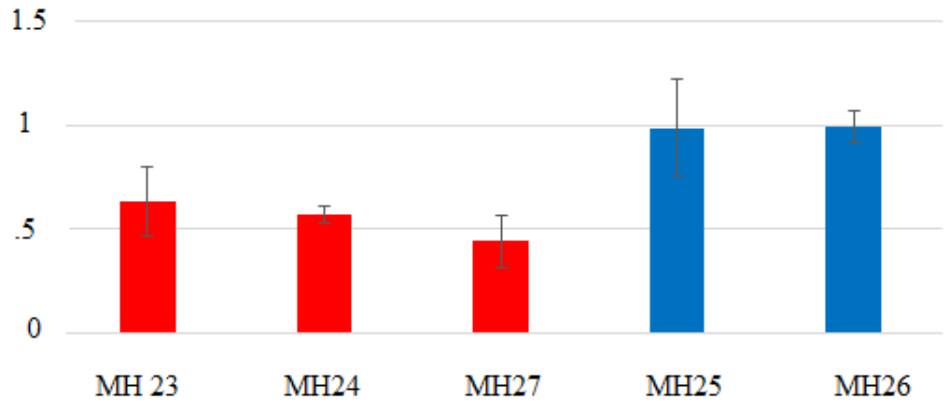
Figure 17: This figure shows results of gene expression assays analyzed using the relative quantification (RQ) method (2ΔΔCT). Gene Expression Assays were performed on cell lines MH0159023 MH0159024 MH0159025 MH0159026 and MH0159027. MH015919, MH0159020, MH0159021, and MH0159022 served as a control for this assay and are shown as green bars. Blue bars represent cell lines with only schizophrenia. Red bars represent cell lines with 22q11.2 DS. All of the cells were normalized to MH1059019, which was set to the fixed value of "1". TNP01, TNP02, and *RPL9* served as control TaqMan® Probes. *DGCR8* gene expression for all 22q11.2 lines was observed to be ~50%. Schizophrenia only lines had *DGCR8* expression levels at 1, indicating no change from the control lines. Samples were run in quadruplicate and error bars represent the standard deviation between these replicates.



The results of gene expression assays can also be analyzed using absolute quantification (shown in Figure 18). In absolute quantifi0cation a standard curve is generated using control measurements. The experimental measurement is than interpreted based on fitting to the standard curve. Lines known to have the 22q11.2 deletion were observed to have *DGCR8* expression at ~0.5. Expression levels ~0.5 represent a 50% reduction in *DGCR8* expression.

DGCR8 was expected to be expressed at ~50% in 22q11.2 DS cell lines. This is because *DGCR8* is located on the deleted region of chromosome 22. The presence of only one copy of *DGCR8* should be observed as *DGCR8* having expression levels of ~50% using absolute quantification. The expression levels of *DGCR8* in cell lines with schizophrenia were observed to be ~1 using absolute quantification. This indicates that there are two copies of *DGCR8* being expressed in these cell lines.

Figure 18: The following shows the results of Gene Expression Assays performed on cell lines MH0159023 MH0159024 MH0159025 MH0159026 and MH0159027 using absolute quantification. MH015919, MH0159020, MH0159021, and MH0159022 served as a control for this assay. A standard curve was generated using the control expression levels of *DGCR8*. Blue bars represent cell lines with only schizophrenia. Red bars represent cell lines with 22q11.2 DS. The expression levels of *DGCR8* in the 22q111.2 deletion lines and the lines known to have schizophrenia were compared to the control lines. TNP01, TNP02, and *RPL9* served as control TaqMan® Probes. *DGCR8* gene expression for all 22q11.2 lines was observed to be ~50%. Schizophrenia only lines had *DGCR8* expression levels at 1, indicating no change from the control lines. Samples were run in quadruplicate and error bars represent the standard deviation between these replicates.



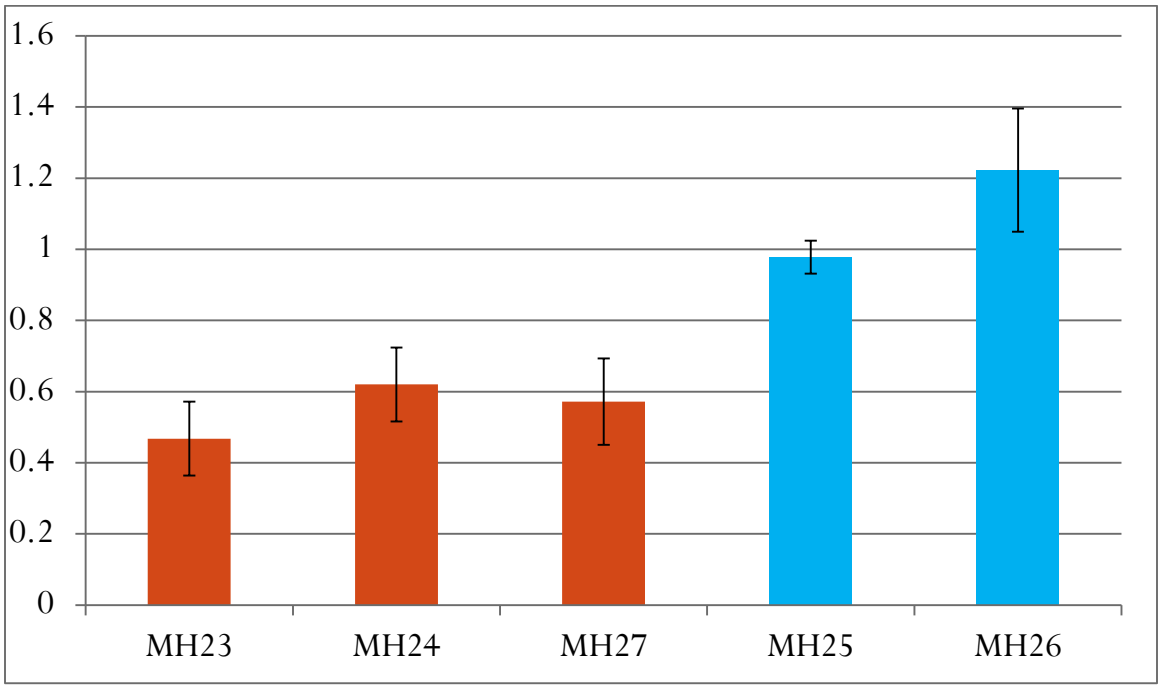
Gene expression assays were performed on NSC lines in addition to the previously discussed iPSC lines. The purpose of the gene expression assays was twofold: to confirm the results of the copy number assays and also to rule out the possibility of *DGCR8* expression increasing in the 22q11.2 deletion lines as part of compensatory mechanism. The expression levels of *DGCR8* were consistent between iPSCs and NSCs indicating that the expression levels of *DGCR8* is maintained throughout the 22q11.2 deletion cell lines regardless of cell types. Cell lines with only schizophrenia were found to have consistent levels of *DGCR8* expression in both iPSCs and NSCs. Elevation of expression in schizophrenia only lines over 1 is a result of experimental variation that was observed throughout the gene expression assays.

Gene expression assays and copy number assays can be performed in conjunction to further confirm the genomic integrity of cell lines. The gene expression assays are generally consistent with the results obtained from copy number assays. Having only one copy of *DGCR8* reduces the expression of *DGCR8* by ~50%. However, the gene expression assays seem to exhibit greater experimental variation than the copy number assays. This variation consistently appears to ~ 30% between measurements. This variation could be due to small changes in gene expression caused by unknown variability in cells or culture conditions. It could also be the result of inherent difficulties of working with unstable RNA.

While this assay demonstrates a lack of compensatory regulatory mechanisms that allow a single copy of *DGCR8* to express at levels similar to two copies, inferring gene copy number from gene expression level is more difficult than from DNA copy number

assays. As a result of this variability, DNA copy number assays were used as the common diagnostic to confirm the 22q11.2 deletion.

Figure 19: The figure below shows the results of TaqMan® Gene Expression assays measuring *DGCR8* expression levels in NSCs. Blue bars represent cell lines with schizophrenia only. Red bars represent cell lines with 22q11.2 DS. All 22q11.2 DS lines have reduced expression of *DGCR8*. The expression levels of *DGCR8* in these NSC lines are consistent with the expression levels of *DGCR8* observed in these lines as iPSCs. Samples were run in quadruplicate and error bars represent the standard deviation between these replicates.



Low Coverage Sequencing for Chromosomal Abnormalities

Low Coverage Sequencing was performed using Ion Torrent karyotyping in order to test the NSC lines for chromosomal abnormalities (RUCDR). This was performed as a pilot experiment with RUCDR as a potential replacement for traditional karyotyping/FISH analysis. The Ion Torrent karyotyping detects gains or losses of whole chromosomes or subchromosomal alterations tens of Mb in size. The Low Coverage Sequencing used detection by low-pass whole genome sequencing on the Ion PGM system. Whole genome amplification was performed on the NSCs, which provided ~80,000 reads per sample and 1,000-7,000 reads per chromosome. Additional details about the quality of the run are provided in Supplemental Figure 6.

Nine iPSC lines and nine corresponding NSC lines were submitted to RUCDR for Low Coverage Sequencing using Ion Torrent Karyotyping. Eight of these lines were found to have 2 copies of each chromosome as expected (MH 19, MH 20, MH 21, MH 22, MH 23, MH 25, MH 26, and MH 27). A second sample of MH 25 was submitted, MH 25 passage 23. In later passages of MH25, chromosomal abnormalities in this line were detected during routine CNV analysis, and this line was observed to be growing much faster than normal. It was hoped that this line would serve as a control to confirm our ability to detect chromosomal abnormalities using our Taqman® CNV diagnostic panel. As expected the MH 25 passage 23 sample did contain chromosomal abnormalities. Trisomies on chromosome 1 and 11 were detected in MH 25 passage 23. The only unexpected anomaly was a trisomy of chromosome 11 in line MH 24 p7 (Tables 9). NSCs from MH 24 had already been successfully generated that were free of chromosomal abnormalities. These NSCs were generated using an earlier iPSC passage,

therefore there was no reason to consider any additional experiments with the iPSCs for MH 24.

Nine different NSCs were submitted for low coverage sequencing as well to check for chromosomal abnormalities. Eight of these lines were found to have 2 copies of each chromosome as expected (MH 19, MH 20, MH 21, MH 22, MH 23, MH 24, MH 25, and MH 27). A second sample of MH 20 was submitted, MH 20 p 28, since it was suspected that at later passages MH 20 might have chromosomal abnormalities. It was observed to have inconsistent results as a result of testing with our Taqman® CNV diagnostic panel and this line was observed to be growing at a much higher rate at later passages than at early passages. As expected MH 20 passage 28 was found to contain chromosomal abnormalities. 10 copies of chromosome 2 were observed in MH 20 passage 28. Additionally, an unexpected chromosomal abnormality was detected in MH 26 passage 13. MH 26 passage 13 was observed to have five copies of chromosome 7, four copies of chromosome 12, eight copies of chromosome 16, seven copies of chromosome 21, and three copies of the X chromosome (Table 10).

The abnormalities detected on MH 26 were problematic since having chromosomal abnormalities renders this line unsuitable for further sequencing or miRNA array analysis. In order to address the problems of chromosomal abnormalities detected in MH 26, the iPSCs were used to remake NSCs for MH 26. The iPSCs for MH26 were found to be free of any known chromosomal abnormality in the iPSC low coverage sequencing.

MH 26 is a cell line known to have schizophrenia. The iPSCs were sent for traditional karyotyping and FISH and were found to be free of any chromosomal

abnormalities. Additionally, the results of Taqman® CNVs using extracted DNA from MH 26 failed to detect any chromosomal abnormalities. The iPSCs that passed the quality control metrics were used to remake NSCs for MH 26. Taqman® CNV assays were then repeated on the NSCs to ensure that the newly generated NSC lines were free of chromosomal abnormalities (Figure 20). Ultimately, MH 26 has not been used for further microarray or sequencing projects since they have focused on control and 22q11.2 DS lines, rather than lines that just have schizophrenia only.

Overall, this was an interesting project to see if it was technically possible to use low coverage next generation sequencing as a replacement for traditional karyotyping. However, the traditional karyotyping proved to be quicker and more cost effective making it much more attractive for use in future projects.

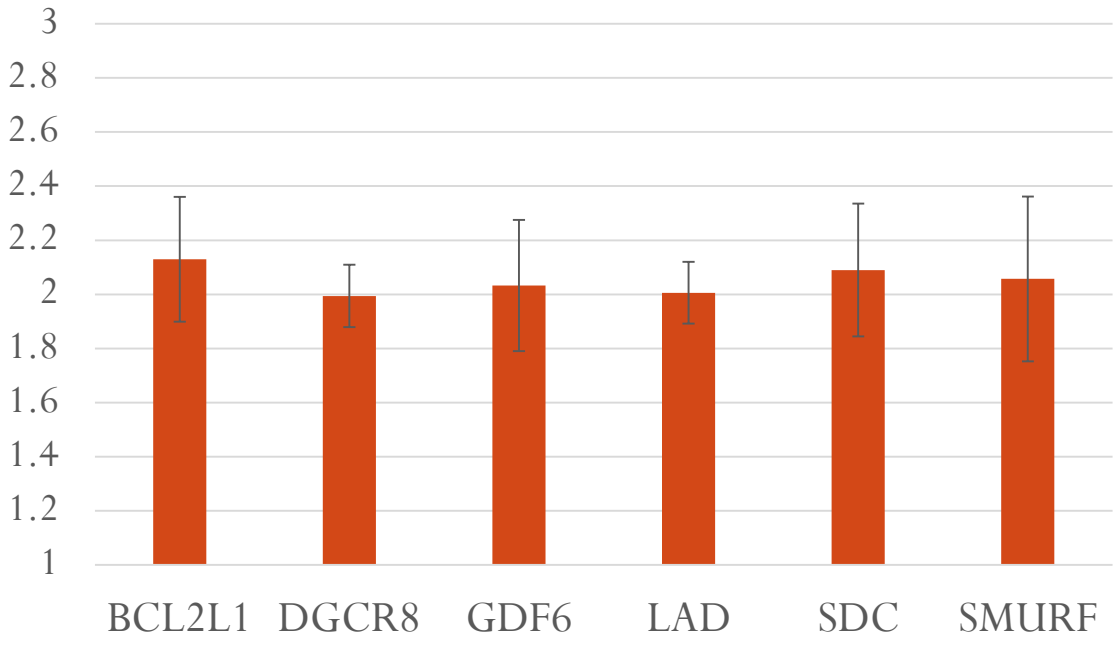
Table 9: The results of the Low Coverage sequencing using the Ion Torrent karyotyping on iPSC samples are shown below. All iPSC lines were found to have 2 copies of each chromosome except for MH 24 passage 7 and MH 25 passage 23.

Sample	iPSC Chromosome Copy Number
IPSC_MH19_P28	2
IPSC_MH20_P19	2
IPSC_MH21_P26	2
IPSC_MH22_P15	2
IPSC_MH23_P14	2
IPSC_MH24_P07	Trisomy Chromosome 11
IPSC_MH25_P09	2
IPSC_MH25_P23	Trisomy Chromosome 1 and 11
IPSC_MH26_P23	2
IPSC_MH27_P14	2

Table 10: The results of the Low Coverage sequencing using the Ion Torrent karyotyping on NSC lines are shown below. All NSC lines were found to have 2 copies of each chromosome except for MH 26 p 13 and MH 20 p 28.

Sample	NSC Chromosome Copy Number
NSC_MH19_P13	2
NSC_MH20_P17	2
NSC_MH20_P28	10 copies of Chr 2
NSC_MH21_P10	2
NSC_MH22_P12	2
NSC_MH23_P08	2
NSC_MH24_P13	2
NSC_MH25_P12	2
NSC_MH26_P13	5 Copies of 7, 4 of 12, 8 of 16, 7 of 21, 3 of X
NSC_MH27_P20	2

Figure 20: Taqman® CNV assays were performed on newly generated NSCs for MH 26. All of the probes show that there are 2 copies of each gene present in MH 26 as expected. The NSCs for this line were remade in response to the results of the low coverage sequencing. Samples were run in quadruplicate and error bars represent the standard deviation between these replicates.



Low Coverage Sequencing to Detect the 22q11.2 Deletion

In addition to testing for unexpected chromosomal abnormalities, it was hoped that Low Coverage sequencing using the Ion Torrent karyotyping could be used to screen for the presence of the 22q11.2 deletion. Unfortunately, the 3 Mb 22q11.2 deletion was too small to be detected using this method, since the Ion Torrent karyotyping algorithm was not designed to detect CNVs smaller than 20Mb.

Since the sequencing was performed at a slightly higher coverage (67,412,950 reads) than usually used for Ion Torrent karyotyping, it was initially hoped that the 22q11.2 deletion could be detected by reducing the tile size (the minimum region to estimate raw copy number change). However, even with an increase in the coverage, it was not possible to reduce the tile size enough to detect a deletion as small as the 22q11.2

The deletion was unable to be detected even after reducing the tile size because reducing the tile size would require repeating the standard deviation training and repeating the baseline reads for the new tile size. The standard deviation training required for Ion Torrent karyotyping is normally done with a 2MB tile size. If a smaller tile size were used, the standard deviation training would have to be repeated for coverage of 200KB tiles. Furthermore, since the initial low coverage sequencing was performed using the default baseline, there was not a way to change the tile size without repeating the sequencing entirely.

The Low Coverage Sequencing used for the karyotypic analysis would theoretically require at least a 10X increase in coverage to generate the statistical power that would make 200kb tiles able to detect the 22q11.2 deletion. As a result of potential GC bias known to be present in 2MB tiles a 15X increase in coverage may be needed to

detect the 22q11.2 deletion. Therefore, the tile size was set to 2MB. 2 MB is the minimum size of regions that are suitable for estimating changes in copy numbers. In order to generate a confident call usually at least 10 reads are needed within a given tile when looking at the whole genome. The low coverage of the run makes the average coverage of a smaller region (smaller than the tile size of 2MB) unreliably low, which is why the tile size must be set to 2MB. Ultimately, this means that the Low Coverage Sequencing can be reliably used to replace karyotyping but cannot FISH for detecting the 22q11.2 deletion.

The Importance of Quality Control and Multiple Testing

It seems that throughout the field other researchers screening iPSCs for chromosomal deletions are verifying their chromosomal deletions with a single karyotype or a Taqman® assay. We have found this to be entirely insufficient for studying chromosomal abnormalities in iPSCs known to develop spontaneous chromosomal aberrations. We have observed normal cell lines that have developed chromosomal abnormalities after multiple passages, and we have observed abnormal lines that have stabilized over the course of multiple passages. There does not seem to be an ideal consensus passage number for profiling all iPSC and NSC lines that we have used in this research.

For example, MH 24 is a line with schizophrenia and the 22q11.2 deletion. The line was submitted for karyotyping at passage 6 and the karyotype was obtained for a cell line with 22q11.2 DS. An abnormal FISH signal pattern was detected for the TUPLE1 probe confirming the presence of the 22q11.2 deletion. However, Taqman® copy number assays performed using DNA from iPSC passage 20 started to show some

chromosomal instability in the line, and 3 copies of chromosome 17 were detected. Passage 20 was abandoned and the line was resumed from an earlier passage. Passage 24 of MH 24 was submitted for low coverage sequencing and a trisomy of chromosome 11 was detected (Table 9). This line represents an example of an iPSC line where earlier passages were found to have normal chromosome profiles, but the iPSC line appeared to destabilize as the passage number increased. This demonstrates the importance of continually profiling the iPSCs throughout the duration of the research.

Alternatively, we have observed lines that looked abnormal at earlier passages that have stabilized and improved at later passages. MH 577 is a cell line that was first generated from a white non-Hispanic or Latino 15 year old female that was observed to have no known other chromosomal abnormalities. This otherwise healthy line was selected for this research to use as a control line versus other lines with the 22q11.2 deletion. Passage 3 of MH 577 was submitted for karyotyping and was observed to have abnormal karyotypic results. Of the 20 cells analyzed, 7 were found to have had non-clonal chromosomal aberrations. This culture seemed to be unstable. As expected, no deletion of 22q11.2 was detected using FISH, but unexpectedly 6% of the cells exhibited a trisomy of 22q11.21 and 11.2% of the cells exhibited a trisomy of 22q13.3 indicating that this passage was a mosaic (Supplementary Figure 5 Panel A).

In response to the abnormal karyotype and FISH results obtained from passage 3 of MH 577, MH 577 was resubmitted for a karyotype and FISH at passage 15. At this later passage MH 577 had stabilized and was exhibiting a normal karyotype. However, the abnormal FISH pattern for chromosome 22 still persisted. 8% of the cells exhibited a trisomy of 22q11.21 and 11.2% of the cells exhibited a trisomy of 22q13.3. The 22q11.21

trisomy had increased 2% from the previous passage, while the 22q13.3 trisomy remained unchanged over the 12 passages (Supplementary Figure 5 Panel B).

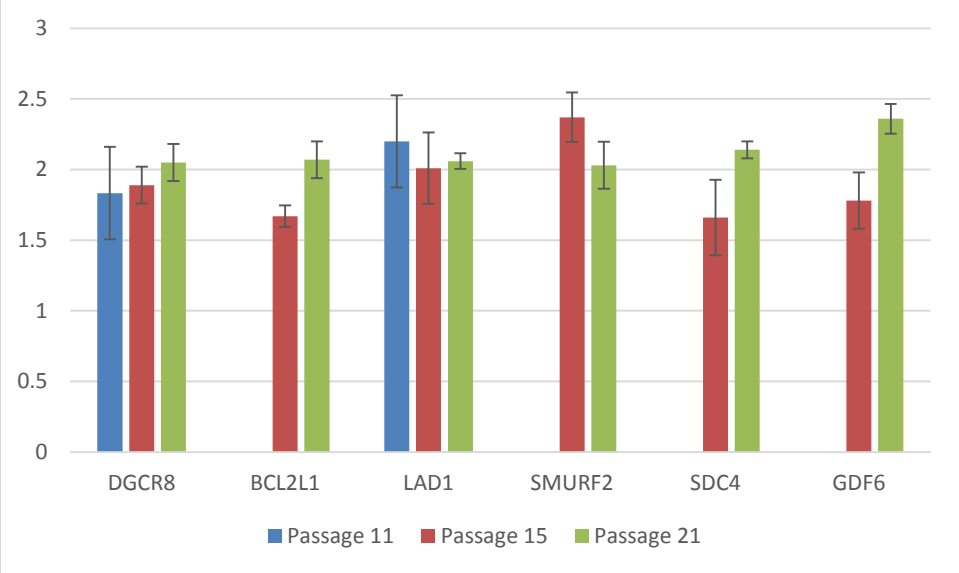
MH 577 was resubmitted for karyotypic and FISH analysis again at passage 21. Although the karyotype failed as a result of low mitotic index in the sample submitted, no elevations in chromosome number were seen for all of the TaqMan® Copy Number Assays performed on MH 577 after passage 15. The MH 577 line appeared to stabilize by passage 21 (Figure 21). The 22q11.21 trisomy decreased to 2.5%, while the 22q13.3 trisomy decreased to 4.9% (Supplementary Figure 5 Panel C). These results were consistent with measurements of two copies of chromosome 22q11.2. By the time the cell line had reached passage 21, both a normal karyotype and the expected FISH results showing two copies of chromosome 22 had been obtained, and MH 577 was able to be used for further analysis. The anomalies detected in the first karyotype and FISH show why multiple testing is needed when generating new iPSC lines.

Overall, these examples of anomalies detected in earlier passages of MH 577 and at later passages of MH 24 demonstrate some of the inherent difficulty in studying a chromosomal deletion in iPSCs. The iPSC lines used in this research were observed to differ in their general susceptibility to chromosomal abnormalities. Some lines, like MH 19, were free of chromosomal abnormalities consistently in TaqMan® Copy Number Assays, in karyotypes, and in low coverage sequencing. Other lines like MH 26 were regularly found to exhibit unexpected chromosomal abnormalities (Table 11). The cell lines each had their own unique growth rates and generally more differentiation was observed in the 22q11.2 deletion lines than the otherwise healthy control lines. It was also observed that iPSC lines made inhouse from fibroblast samples submitted to RUCDR

grew more quickly and had less differentiation than lines obtained directly as iPSCs from RUCDR that were transfected by third parties. These inherent differences in individual cell lines and the risk of developing chromosomal abnormalities at essentially any passage illustrate the important of repeated testing of the chromosomal integrity of all iPSC and NSC cell lines throughout the research.

.

Figure 21: The average results of TaqMan® Copy Number Assays performed on passage 15 and 21 of the control line MH 577 are shown below. ~2 copies of each probe were detected in samples from MH 577 and the control line MH 509. Copy Caller Software estimated that the copy number for all probes measured was ~2 even if there is some variability within the assay or mosaicism within the cell lines.



BCL2L1 Probe Irregularities

The Taqman® probe chosen for Chromosome 20 was fBCL2L1. However, this probe consistently proved unreliable (Figure 22). A probe for SDC4 which is also on Chromosome 20 was found to be consistently reliable and used throughout this research. The questionable data obtained from BCL2L1 can be illustrated by looking at MH 26. MH 26 was probed for *DGCR8*, *BCL2L1*, *SDC4*, *LAD2*, and *GDF6*. Cell line MH 19 and 20 were used as a control and do not contain the 22q11.2 deletion. Four replicates of this assay were performed for each passage.

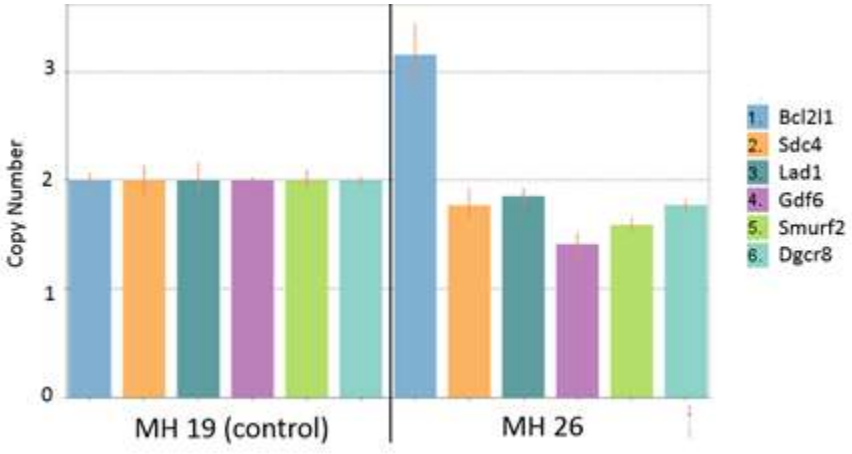
The copy number for the probe *BCL2L1* which is in a region known to commonly develop trisomies was observed to change over multiple passages of this cell line. Taqman® copy number assays performed on passage 19 (Figure 22 Panel A) and passage 20 (Figure 22 Panel B) of MH 26 showed elevated levels of *BCL2L1* that suggested that their might be something unexpected occurring on chromosome 20 in this cell line due to elevations of *BCL2L1*. However, *SDC4* is also present on chromosome 20 and was observed to have only 2 copies. These results are conflicting and suggest that something abnormal is occurring in either this cell line or with the probes for chromosome 20.

Additional cell lines tested with *BCL2L1*, demonstrated that the probe had higher variability that the other probes. Furthermore, no abnormalities were detected on chromosome 20 in MH 26 in the karyotype that was performed. In Panel C, passage 24 shows values for each probe that the Copy Caller Software reported as 2 copies. The elevated *BCL2L1* from the previous passages ws reduced. There were also fluctuations to the other probes due to biological variability. Panel C also shows passage 27 of MH 26.

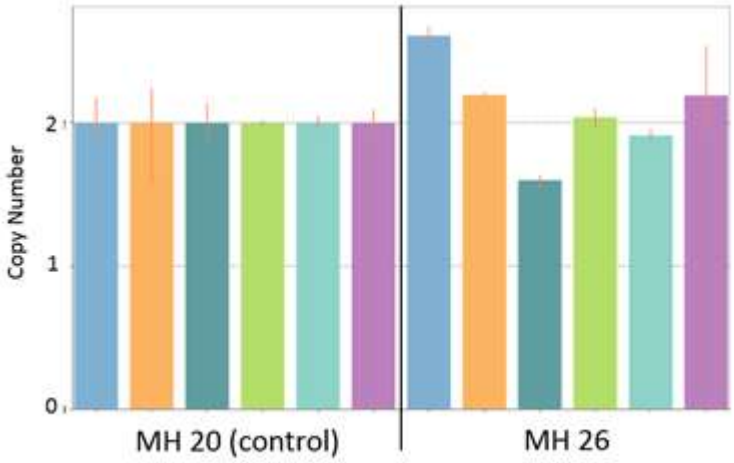
BCL2L1 was elevated again and 3 copies of *BCL2L1* were detected. The chromosomal integrity of MH 26 around passage 24 was confirmed by submitting passage 23 for low coverage sequencing. Low coverage sequencing detected 2 copies of each chromosome in iPSC passage 23 (Figure 22 Panel C). No abnormalities on chromosome 20 were detected in the low coverage sequencing results. The consistent variability in Taqman® *BCL2L1* data is therefore likely the result of a bad probe and not of biological importance.

Figure 22: TaqMan® Copy Number Assays were performed on iPSC line MH 26 for multiple passages. These assays probed for *DGCR8*, *BCL2L1*, *GDF6*, *LAD2*, and S*DC4*. Cell line MH 19 and 20 were used as a control and do not contain the 22q11.2 deletion. Four replicates of this assay were performed for each passage. Panel A shows passage 19, Panel B shows passage 20, and Panel C shows passages 24 and 27.

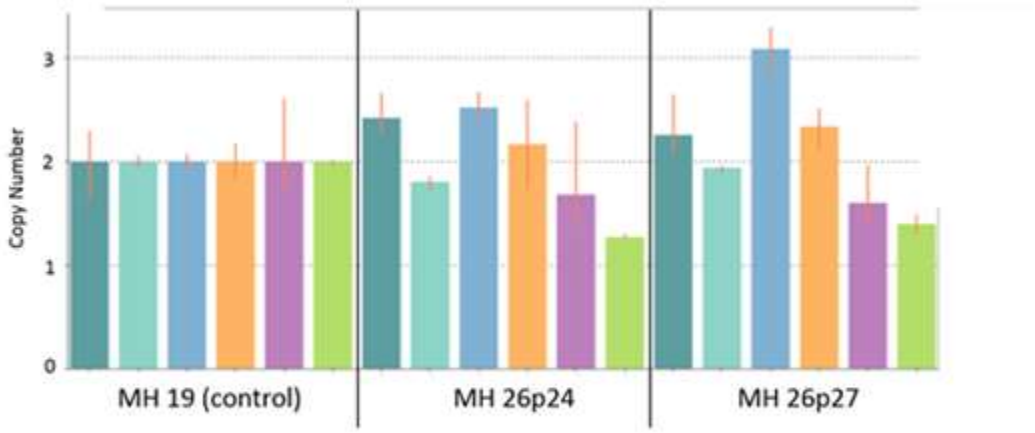
Panel A:



Panel B:



Panel C:



In order to generate high quality iPSCs and NSCs, multiple experiments are performed on each line. iPSCs are confirmed using alkaline phosphatase stains. The iPSCs were tested for the 22q11.2 deletion using Taqman® gene expression assays and Taqman® CNV assays (Thermo Fisher Scientific). The deletion was further tested for using FISH (Cell Line Genetics). Karyotypes were performed to confirm chromosomal integrity (Cell Line Genetics). The NSCs were also rigorously tested. All NSCs were stained for using NSC markers. The 22q11.2 deletion in the NSCs was tested for again using Taqman® gene expression assays and Taqman® CNV assays (Life Technologies).

If unusual results persisted, the particular passage of the cell line in question was abandoned and previous passages of the same cell line were used. A summary of all abandoned passages where chromosomal abnormalities were detected using either Taqman® CNV assays, karyotyping/FISH, or low coverage sequencing is shown in Table 11. These previous passages were then tested until a passage exhibiting the expected CNV profile was obtained. If a passage with the expected profile could not be obtained, NSCs/iPSCs were remade from the existing stock of deposited fibroblasts.

Table 11: The following table lists all of the abandoned passages for iPSC lines obtained from RUCDR (n=20) where chromosomal abnormalities were detected using either Taqman® CNV assays, karyotyping/FISH, or low coverage sequencing.

RUID	Diagnosis	Completed Taqman® miRNA Arrays	Anomalies Detected by Taqman® CNV Assay	Anomalies Detected by Karyotype & FISH (TUPLE1)	Anomalies Detected by Low Coverage Sequencing
MH0159019	Control	Yes	None	None	None
MH0159020	Control	Yes	P 20 = 3 copies of Chr 20	P10 = partial duplication of the long arm of Chr 1 between bands 1q21 and 1q42; trisomy Chr 12; P 20 = deletion of long arm of Chr 5 at band q33; unbalanced translocation between long arm of Chr 21 at band 22 and the long arm of Chr 1 at band q12 resulting in partial trisomy of Chr 1	P 28 = 10 copies of Chr 2
MH0159021	Control	Yes	P 19 = 4 copies of Chr 20	None	None
MH0159022	Control	Yes	P 21 = 3 copies of Chr 17	P 20 = extra copy of 22q13.3	None

MH0159023	Schizo-affective disorder / VCFS	Yes	P 18 = 4 copies of Chr 20 P 17 = 4 copies of Chr 20 P 12 = 4 copies of Chr 20 P 11 = 3 copies of Chr 20 P 8 = 6 copies of Chr 23,	P 24 = isochromosome of the long arm of Chr 20, i.e., 3 copies of the Chr 20 long (g) arm and 1 copy of the short (p) arm	None
MH0159024	Schizo-affective disorder / VCFS	Yes	P 20 = 3 copies of Chr 17	None	P 24 = trisomy Chr 11
MH0159025	schizophrenia, paranoid type		None	P 12 = multiple large gains and losses (>10Mb)	P 23 = trisomy Chrs 1 and 11
MH0159026	schizophrenia, paranoid type		P 27 = 3 copies of Chr 20 P 20 = 3 copies of Chr 20 P 19 = 3 copies of Chr 20 P 16 = 3 copies of Chr 20	P 17 = gain of marker chromosome of unknown origin	P 13 = 5 copies of Chr 7, 4 copies of Chr 12, 8 copies of Chr 16, 7 copies of Chr 21, 3 copies of X

			P 15 = 3 copies of Chr 20		
MH0159027	schizoaffective disorder / VCFS	Yes	P 18 = 4 copies of Chr 20	None	None
MH0161608	schizoaffective disorder / VCFS		None	None	N/A
MH0162499	22q11.21 del / childhood onset SZ	Yes	P 7 = 3 copies of Chr 20 P 4 = 3 copies of Chr 8 P 3 = 3 copies of Chr 17	None	N/A
MH0162508	22q11.21 del / childhood onset SZ		None	None	N/A
MH0162519	22q11.21 del / childhood onset SZ	Yes	P 23 = 3 copies of Chr 8 P 8 = 3 copies of Chr 1	None	N/A
MH0162627	schizoaffective disorder / VCFS		P 5 = 3 copies of Chr 1	None	N/A
MH0162762	22q11.2 deletion / Autism		P 12 = 3 copies of Chr 8 P 11 = 3 copies of Chr 8 P 5 = 3 copies of Chr 8	None	N/A

			P 3 = 3 copies of Chr 17, 4 copies of Chr 8		
MH0162509	Control	Yes	None	None	N/A
MH0162626	Control		P 17 = 3 copies of Chr 1 P 12 = 4 copies of Chr 17 P 10 = 6 copies of Chr 17, 3 copies of Chr 8 P 9 = 5 copies of 17	None	N/A
MH0162577	Control	Yes	None	P 15 = trisomy 22q13.3 and 22q11.21; P 3 = multiple numerical and structural aberrations in 7/20 cells tested	N/A
MH0167170	Control		None	P 9 = trisomy Chr 8	N/A
MH0161602	Control		P 21 = 8 copies of Chr 17 P 15 4 copies of Chr 1, 5 copies of Chr 17	None	N/A

Chapter 4: Characterizing miRNAs in NSCs via TaqMan® Arrays

In order to characterize the miRNAs that are expressed in NSCs from individuals with and without the 22q11.2 deletion two very different yet equally important techniques have been used: TaqMan® Array Cards and small miRNA sequencing. Although small miRNA sequencing is the newer technique that is en vogue, it is not without limitations. Both of these techniques have their own strengths and weaknesses. It is our opinion that the best possible results come from using these techniques in parallel.

Cell Line Selection for TaqMan® Arrays

Ideally, all of the 22q11.2 cell lines that have been obtained (n=8) and their corresponding control lines would be analyzed using TaqMan® Arrays and small miRNA sequencing. However, only the NSCs have been analyzed using TaqMan® Arrays and submitted for small miRNA sequencing since the NSCs are a neuronal subtype. Studying a neuronal cell is important, since the goal of this research is to understand more about miRNA dysregulation in schizophrenia, a neuropsychiatric disorder. At this point there are no plans to study the iPSCs directly using the TaqMan® Arrays and small miRNA sequencing. Resources spent on profiling the iPSCs would likely be better spent on developing and analyzing additional NSC lines that have 22q11.2 DS.

Eventually, all of the 22q11.2 cell lines that have been gathered will be profiled using TaqMan® Arrays and submitted for small RNA sequencing. However, at this point the miRNA profiling has only been completed for a subset of the available lines. Due to time and budget constraints six of the best characterized 22q11.2 deletion lines and six of the corresponding control lines have been profiled using TaqMan® Arrays and submitted

for small miRNA sequencing. This subset of lines includes control lines MH 19, MH 20, MH 21, MH 22, MH 509, and MH 577 and 22q11.2 deletion lines MH 23, MH 24, MH 27, MH 499, and MH 519.

miRNA Extraction:

miRNA was isolated from NSCs (n=12) using the *mir*VANA miRNA Isolation Kit according to the manufacturer's protocol (Thermo Fisher Scientific). miRNAs range from 15-30 nucleotides in length and require a specific kit for their isolation. The *mir*VANA miRNA isolation procedure uses an organic extraction followed by immobilization of RNA on glass fiber filters to purify miRNAs. Collecting just the small miRNAs rather than total RNA requires the use of chaotropic salts and phenol-chloroform to inactivate RNases and purify RNA.

Cell samples consisting of ~1 X 10^6 cells were first lysed in 600 µl of denaturing lysis solution to stabilize the RNA and inactivate RNases. The lysate was extracted with acid phenol chloroform to remove most of the other cellular components following the *mir*VANA miRNA Isolation Kit protocol. The semi-pure RNA sample was purified over a glass fiber filter to yield an enriched portion of miRNAs. The glass fiber filter was treated with kit solutions designed specifically for miRNA retention and prevention of miRNA degradation as directed in the enrichment procedure for small RNAs (Thermo Fisher Scientific). The concentration of miRNA was estimated using a NanoDrop 2000 UV-Vis Spectrophotometer.

TaqMan® Array Cards Methods

TaqMan® miRNA arrays were used to compare the expression profiles of miRNA from miRNA extracted from 22q11.2 cell lines versus the miRNA extracted

from control cell lines. These cards function as microarrays that use small RNAs extracted from cells and apply them to a series of preselected human probes and have been shown to be an effective way of measuring miRNA expression in human samples (Moreau et al.). The TaqMan® human microRNA array contains dried TaqMan® primers and probes for 768 unique assays specific for human miRNAs. The miRNAs contained within both Human Pool A vs 2.0 and Human Pool B vs 3.0 are shown in Supplementary Figure 7 (Thermo Fisher).

Megaplex™ Reverse transcription was performed. Megaplex™ Pools for microRNA expression analysis were used to prepare the samples for the TaqMan® Array Cards (Thermo Fisher Scientific) Megaplex™ Pools can detect and quantify up to 380 miRNAs per pool. The Megaplex™ Pool consists of stem looped reverse transcription primers that are used to synthesize cDNA for mature miRNAs. Two Megaplex™ Pools were used to run the arrays, Human Pool A vs 2.0 and Human Pool B vs 3.0. The Megaplex™ Pools consist of a matched set of primer pool and TaqMan® Array Cards.

The Megaplex™ Reverse transcription was performed using the TaqMan® miRNA reverse transcription kit and the Megaplex™ RT primers to generate single stranded cDNA from the miRNA samples. 600 ng of RNA was used as a starting point for each sample. The reverse transcription reaction used a RT reaction mix of 4.5µl consisting of 0.80 µl of Megaplex™ RT primers (10X), 0.20 µl of dNTPs with dTTP (100mM), 1.50 µl of MultiScribe™ reverse transcriptase (50 U/µl), 0.80 µl of 10X RT buffer, 0.90 µl of $MgCl_2$ (25 mM), 0.10 µl of RNase inhibitor (20 U/µl), and 0.20 µl nuclease free water for one sample. The Megaplex™ Reverse transcription was run in

MicroAmp 8-Tube strips using a PTC-200 Peltier thermal cycler (BioRad) and following the manufacturer's protocol (Thermo Fisher Scientific).

Additionally, Megaplex™ PreAmp Primers were used. The Megaplex™ PreAmp Primers consists of two sets of gene specific forward and reverse primers that are designed to work with very small quantities of starting material. These primers allow the miRNA cDNA target to be preamplified prior to running the TaqMan® arrays. The preamplification was performed since preamplification is recommended for all samples of cDNA ranging from 1 to 250 ng. The preamplification reaction was performed using 2.5 µl of RT product generated from the Megaplex™ Reverse transcription and 22.5 µl of PreAmp reaction mix per sample. The PreAmp reaction mix consisted of 12.5 µl of TaqMan® PreAmp Master Mix (2X), 2.5 µl of Megaplex™ PreAmp Primers (10X), and 7.5 µl of nuclease free water. The Megaplex™ Preamplification Reaction was run in MicroAmp 8-Tube strips using a PTC-200 Peltier thermal cycler (BioRad) and following the manufacturer's protocol (Thermo Fisher Scientific).

Following the preamplification reaction the sample was diluted in 75 µl of 0.1X TE pH 8.0. The RT-PCR master mix for one array consisted of 450 µl of TaqMan® Universal PCR Master Mix (No Amperase® UNG, 2X), 9 µl of diluted PreAmp product, and 441 µl of nuclease free water. The TaqMan® arrays were run using 100 µl of RT-PCR reaction mix per port of the TaqMan® array. The arrays were sealed and run using a 7900HT System with automation accessory (Thermo Fisher Scientific).

TaqMan® Array Cards Analysis Methods

A full TaqMan® miRNA profile represents two Megaplex™ RT reactions (Pool A and Pool B), two preamplification reactions (Pool A and Pool B), and two TaqMan®

microarrays per sample. These TaqMan® probes measured the expression levels of a panel of miRNAs present in the cell lines shown as miRNA levels expressed as a relative fold change using the $2^{-\Delta\Delta Ct}$ method. ExpressionSuite vs 1.0.3 was used to generate and analyze the RQ values of the miRNAs (Thermo Fisher Scientific).

Preliminary Analysis of miRNA Variability Between Control and 22q11.2 Deletion Groups in NSCs from Developed from RUCDR Samples

A preliminary analysis was performed using the array data taken from arrays ran using MH 20 and MH 22 (both control lines)(n=2) and lines MH 23 and MH 24 (both 22q11.2 lines) (n=2). A plot comparing the log of the standard deviation of $\log_2$(RQ) for the miRNA RQ values was used to compare the variability of miRNAs within control cell lines versus comparing the control cell lines to the 22q11.2 cell lines. The results of this analysis produced four clusters of points, with each point representing a specific miRNA (Figure 23).

Due to the small number of cell lines and replicates within cell lines in this preliminary data, once the variability across cell lines drops too far below the variability within cell lines, the model started to fail and gave very small estimates for the variability across cell lines. Some of the clusters represented low variability in miRNAs across cell lines in figure 23 are those where the model collapsed, but these represent miRNAs that likely are among the least interesting, since they show no change between control and 22q11.2 deletion groups and are likely stably expressed miRNAs.

The top left cluster represents 65 miRNAs with higher variability within the control cell liens tested than between the control cell lines and the 22q11.2 lines. This

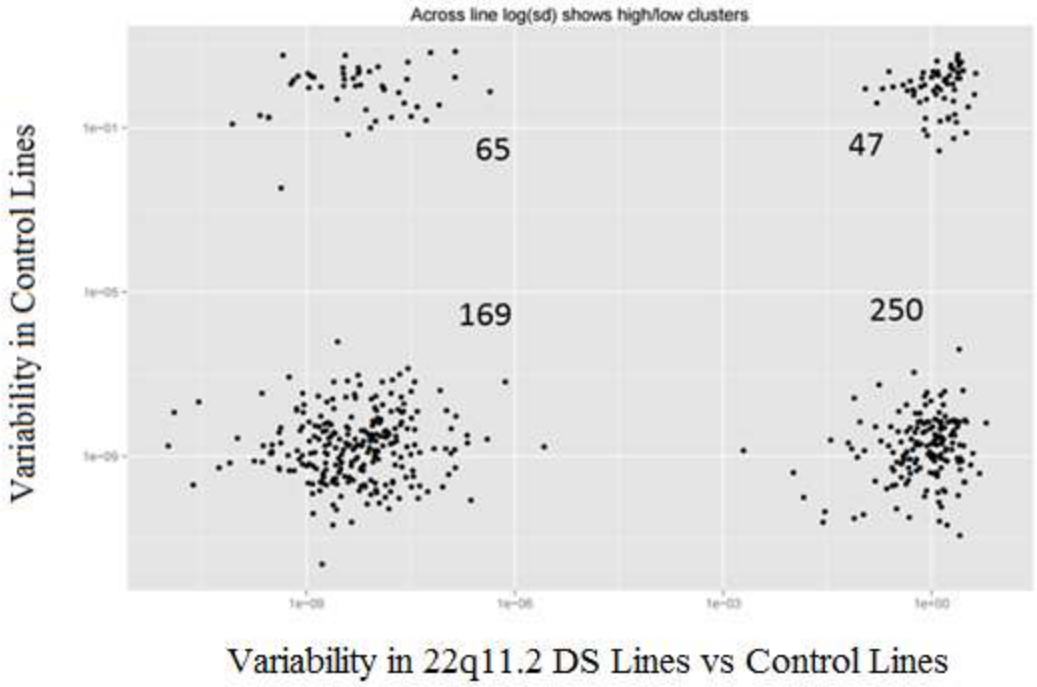
result could be caused by unreliable probes or variability that is not related to the 22q11.2 deletion and schizophrenia.

The bottom left cluster contains 169 miRNAs that were found to have low variability within the controls and between the control cell lines and the 22q11.2 lines. At some points within this cluster the model starts to collapse but these miRNAs are stably expressed and not related to 22q11.2.

The top right cluster consists of 47 miRNAs which were found to have high variability within the control cell lines and between the control cell lines and the 22q11.2 lines. These are miRNAs with high variability in expression where the variability does not appear related to the existence of the 22q11.2 deletion.

The cluster in the bottom right corner contains 250 miRNAs that all were found to have greater variability between the 22q11.2 lines to the controls than within the control lines. These represent miRNAs that are most interesting to this research as they appear to be showing expression differences as a result of the 22q11.2 deletion.

Figure 23: These clusters represent the variability between found to exist in individual miRNA probes between control and 22q11.2 lines using the TaqMan® Human miRNA Cards. The numbers next to each of the 4 clusters represent the number of miRNAs (represented as dots in the cluster) that make up each of the four clusters.



TaqMan® Array Cards Analysis:

The expression levels of all miRNAs were measured relative to a stably expressed endogenous control gene, mammalian RNU48 (Thermo Fisher Scientific). All arrays from control groups were pooled into one control group that was compared to all arrays from 22q11.2 deletion lines that was pooled into a 22q11.2 DS group. All arrays were run in triplicate. A list of miRNAs was generated that had increases in fold changes greater than 2 or fold changes that were decreased by greater than 0.5 for miRNAs on either the A or B Arrays. Using this criterion 142 decreased miRNAs and 64 increased miRNAs were found for the A Plates, while 69 decreased miRNAs and 160 increased miRNAs were detected on the B plates. Table 12 outlines this procedure. Individual miRNAs are listed in tables 13-16. miRNAs with an undetermined CT value were indicative of poor quality data that was flagged. These miRNAs were removed from the analysis, leaving 106 decreased miRNAs and 53 increased miRNAs for the A Plates, and 29 decreased miRNAs and 87 increased miRNAs for the B plates. Finally, the list of remaining miRNAs was further refined by retaining only miRNAs with a nominally significant p-value <0.05. This left 25 decreased miRNAs and two increased miRNAs detected using the A Plates, while only one decreased miRNAs remained in the B plates (Table 17). The fold changes and p-values were generated using ExpressionSuite vs 1.0.3 (Thermo Fisher Scientific).

The final miRNAs that had a p-value < 0.05 consisted of 26 miRNAs that were decreased and only two miRNAs that were increased (Table 17). This is expected since the deletion is believed to interfere with the production of mature miRNAs. The miRNAs that were observed to increase could be the result of compensatory measures increasing

the production of a miRNA that performs a similar function to a miRNA that was not able to be produced. The reason for datable increases in miRNAs remains unclear and is an interesting avenue for future research.

Of the 26 remaining miRNAs, 25 of them were miRNAs that were on Plate A. Only one miRNA was from Plate B. Card A contains miRNAs that have been more highly characterized, while Card B contains more recently discovered miRNAs that may not be as important. Additionally, Plate A always had hundreds fewer flags than Plate B. A majority of the miRNA assays on Plate B failed or were otherwise undetectable in our miRNA samples. It was expected and confirmed that Plate A would perform better than Plate B (Thermo Fisher Scientific).

Of the 26 miRNAs that were observed to be decreased with p-values <0.05, 6 were miRNAs that had been previously implicated as being important for the development of neuropsychiatric disorders (miR-185, miR-212, miR-339, miR-382, miR-425, and miR-889). miR-185 was also observed to be downregulated in leukocytes with 22q11.2 deletion syndrome (Sellier et al.), neurons derived from iPSCs with the 22q11.2 deletion (Zhao et al.), and in studies of the prefrontal cortex and hippocampus of mice with a deletion equivalent to the 22q11.2 deletion (Stark et al.). miR-212 was found to be elevated in dorsolateral prefrontal cortex samples with schizophrenia (Kim et al.) and to have reduced expression levels in the prefrontal cortex of samples with schizophrenia (Perkins et al.). Here miR-212 was observed to be reduced. miR-339 was also observed to have reduced expression in prefrontal cortex samples with bipolar disorder (Moreau et al.). miR-382 was observed to be elevated in cultures derived from individuals with schizophrenia (Mor et al.), while here it was observed to have decreased expression.

miR-425 was observed to have decreased expression in the prefrontal cortex of samples with bipolar disorder (Moreau et al.). miR-889 was found to have elevated expression in dorsolateral prefrontal cortex samples with bipolar disorder, while here miR-889 was observed to be decreased in expression (Kim et al.).

The results from the TaqMan® Array Cards detected one miRNA that was significantly reduced that also was observed to be reduced in the Sellier et al. lymphocyte study, miR-185. Other miRNAs found to be reduced in the leukocytes were miR-150, miR-194, and miR-363 (Sellier et al.). No reduction in these miRNAs was observed in the TaqMan® Array Cards data. miR-205, miR-190, and miR-1 were reported to have significantly increased expression in leukocytes (Sellier et al.). No increase in these miRNAs was detected. These differences could be the result of differences between profiling miRNAs in leukocytes and profiling miRNAs in NSCs, since miRNA expression has been shown to be tissue specific (Adlakha and Saini). There are a number of other explanations for this discrepancy as well. There could be differences in the expression of the miRNAs between these cell lines as a result of age, gender, and ethnic background. The differences in miRNAs could also be the result of differences in small RNA extraction technique, handling of the small RNA, or variations in miRNA quantification methods. As a result of the multitude of differences between experimental parameters, these differences were expected.

The Zhao et al. study of neurons derived from iPSCs reported a reduction of miR-185, miR-486, miR-491, miR-3175, and 4804 (Zhao et al.). miR-185 was also reported as being downregulated by our TaqMan® Array Cards, the Sellier et al study, and the Stark et al. study (Zhao et al.; Sellier et al.; Stark et al.). This evidence a meaningful reduction

of miR-185 in samples with 22q11.2 DS and the potential importance of miR-185 to schizophrenia is mounting. miR-486 was observed to be reduced but not significantly so in the TaqMan® Arrays. No reduction in miR-491 was detected in the arrays. This difference could be the result of our study using a homogenous mixture of NSCs and the Zhao et al study using a mixture of poorly characterized neuronal subtypes. miR-3175 and miR-4804 were miRNAs that were not included on the TaqMan® Array Cards. This illustrates one of the major weaknesses with the TaqMan® Array Cards and strengths of small RNA sequencing, which was used by Zhao et al. Small RNA sequencing is able to detect miRNAs genome wide, unlike the TaqMan® Array Cards, which are limited to profiling only a limited number of preselected miRNAs.

Table 12: The table below outlines the process by which the relative quantification or fold change for the TaqMan® miRNA Array probes data was analyzed. The first box shows the number of probes that were found to be reduced less than 0.5 fold or increased greater than 2 fold on either the A or B plates. Next, the miRNA probes that had undetermined CT Values or other flags were omitted. Probes with undetermined CT values are unreliable. The last box shows probes that had a p-value < 0.05 after an unpaired, two-tailed t-test of sample normalization with an assumption of unequal variance was performed.

A Plates	B Plates
142 Probes Initially Downregulated 64 Probes Initially Upregulated	69 Probes Initially Downregulated 160 Probes Initially Upregulated
Omitted Probes With Undetermined CT Values or Otherwise Flagged	
A Plates	B Plates
106 Downregulated 53 Upregulated	29 Downregulated 90 Upregulated
Retained only probes with p<0.05	
A Plates	B Plates
23 Downregulated 2 Upregulated	1 Downregulated

Table 13: This table lists all of the TaqMan® miRNA probes that were found to have relative quantification (RQ) values that were reduced greater than 0.5 fold on the 22q11.2 arrays versus the control arrays for TaqMan® Human miRNA Plate A. The Ct (cycle threshold) is the number of cycles it takes the fluorescent signal to cross the threshold and exceed the background level in qRT-PCR. The p-value was obtained by using an unpaired, two-tailed t-test of sample normalization with an assumption of unequal variance.

Target Name	Ct	RQ	P-Value
hsa-miR-541-002201	35.719	0.022	0.003
hsa-miR-487b-001285	28.201	0.101	0.006
hsa-miR-376c-002122	27.896	0.111	0.006
hsa-miR-655-001612	32.262	0.265	0.008
hsa-miR-758-001990	35.182	0.117	0.009
hsa-miR-433-001028	24.269	0.401	0.009
hsa-miR-494-002365	27.217	0.314	0.010
hsa-miR-510-002241	34.168	0.288	0.013
hsa-miR-127-000452	26.800	0.302	0.014
hsa-miR-431-001979	34.220	0.051	0.016
hsa-miR-323-3p-002227	27.107	0.328	0.017
hsa-miR-339-3p-002184	24.565	0.341	0.020
hsa-miR-369-5p-001021	31.681	0.284	0.021
hsa-miR-493-002364	29.795	0.467	0.024
hsa-miR-889-002202	30.354	0.228	0.025
hsa-miR-212-000515	25.728	0.464	0.025
hsa-miR-185-002271	27.885	0.324	0.026
hsa-miR-200c-002300	27.169	0.385	0.028
hsa-miR-382-000572	24.620	0.364	0.030
hsa-miR-616-002414	35.157	0.088	0.037
hsa-miR-425-5p-001516	23.843	0.400	0.038
hsa-miR-429-001024	38.948	0.047	0.045
hsa-miR-146b-001097	22.375	0.354	0.051
hsa-miR-891a-002191	34.68	0.116	0.054
hsa-miR-191-002299	17.42	0.311	0.074

hsa-miR-410-001274	27.129	0.248	0.076
hsa-miR-33b-002085	27.963	0.417	0.078
hsa-miR-331-5p-002233	28.193	0.332	0.081
hsa-miR-486-3p-002093	29.937	0.475	0.084
hsa-miR-29b-000413	35.145	0.097	0.095
hsa-miR-454-002323	20.532	0.432	0.101
hsa-miR-299-5p-000600	38.685	0.063	0.102
hsa-miR-190-000489	29.820	0.108	0.121
hsa-miR-147b-002262	34.798	0.173	0.122
hsa-miR-133b-002247	28.191	0.277	0.138
hsa-miR-628-5p-002433	25.904	0.303	0.138
hsa-miR-325-000540	31.842	0.102	0.141
hsa-miR-329-001101	38.066	0.077	0.155
hsa-miR-125a-5p-002198	19.403	0.382	0.161
hsa-miR-342-5p-002147	28.870	0.423	0.166
hsa-miR-328-000543	24.403	0.301	0.166
hsa-miR-146b-3p-002361	35.746	0.316	0.170
hsa-miR-370-002275	26.467	0.164	0.182
hsa-miR-184-000485	25.028	0.334	0.199
hsa-miR-22-000398	29.713	0.489	0.215
hsa-miR-186-002285	22.020	0.126	0.222
hsa-miR-10a-000387	24.145	0.482	0.234
hsa-miR-674-002021	31.198	0.302	0.237
hsa-miR-589-002409	32.267	0.313	0.241
hsa-miR-1-002222	26.702	0.324	0.246
hsa-miR-342-3p-002260	18.295	0.414	0.252
hsa-miR-197-000497	19.710	0.402	0.254
hsa-miR-876-3p-002225	36.627	0.279	0.268
hsa-miR-142-3p-000464	32.831	0.249	0.268
hsa-miR-484-001821	17.425	0.334	0.274
hsa-miR-149-002255	20.945	0.392	0.279
hsa-miR-519e-002370	25.062	0.102	0.280
hsa-miR-485-3p-001277	28.335	0.104	0.287
hsa-miR-873-002356	32.743	0.164	0.288
hsa-miR-499-3p-002427	37.998	0.003	0.294
hsa-miR-138-002284	27.417	0.256	0.313
hsa-miR-876-5p-002205	39.383	0.019	0.315
hsa-miR-105-002167	32.357	0.262	0.327
hsa-miR-545-002267	30.212	0.121	0.328
hsa-miR-339-5p-002257	27.055	0.472	0.329
hsa-miR-298-002190	36.999	0.440	0.340

hsa-miR-500-002428	29.145	0.305	0.342
hsa-miR-487a-001279	32.969	0.059	0.343
hsa-miR-346-000553	30.376	0.294	0.353
hsa-miR-501-3p-002435	31.539	0.401	0.361
hsa-miR-548d-5p-002237	35.403	0.309	0.367
hsa-miR-483-5p-002338	21.779	0.448	0.367
hsa-miR-561-001528	30.418	0.091	0.374
hsa-miR-885-5p-002296	24.011	0.493	0.378
hsa-miR-516b-001150	33.937	0.298	0.398
hsa-miR-139-5p-002289	25.694	0.384	0.404
hsa-miR-345-002186	22.543	0.405	0.437
hsa-miR-135b-002261	24.469	0.465	0.438
hsa-miR-139-3p-002313	28.166	0.382	0.443
hsa-miR-517c-001153	29.544	0.429	0.463
hsa-miR-501-001047	25.828	0.312	0.485
hsa-miR-95-000433	28.929	0.299	0.491
hsa-miR-548d-001605	35.081	0.202	0.500
hsa-miR-519a-002415	29.093	0.308	0.533
hsa-miR-140-3p-002234	27.095	0.468	0.545
hsa-miR-532-001518	26.028	0.321	0.569
hsa-miR-133a-002246	26.309	0.287	0.572
hsa-miR-518b-001156	32.035	0.281	0.629
hsa-miR-520g-001121	31.919	0.284	0.648
hsa-miR-217-002337	32.746	0.418	0.726
hsa-miR-618-001593	29.331	0.353	0.728
hsa-miR-362-3p-002117	30.398	0.476	0.747
hsa-miR-532-3p-002355	23.953	0.499	0.751
hsa-miR-182-002334	30.105	0.256	0.828
hsa-miR-214-002306	20.309	0.414	0.867
hsa-miR-660-001515	24.826	0.456	0.883
hsa-miR-183-002269	30.748	0.416	0.940
hsa-miR-126-002228	23.863	0.423	0.970
hsa-miR-519d-002403	33.607	0.060	0.984

Table 14: This table lists all of the TaqMan® miRNA probes that were found to have relative quantification (RQ) values that were increased greater than 2.0 fold on the 22q11.2 arrays versus the control arrays for TaqMan® Human miRNA Plate A. The Ct (cycle threshold) is the number of cycles it takes the fluorescent signal to cross the threshold and exceed the background level in qRT-PCR. The p-value was obtained by using an unpaired, two-tailed t-test of sample normalization with an assumption of unequal variance.

Target Name	Ct	RQ	P-Value
hsa-miR-519c-001163	26.425	57.066	0.027
hsa-miR-515-3p-002369	30.896	16.258	0.031
hsa-miR-520a-001167	23.72	2.557	0.146
hsa-miR-199b-000500	28.923	5.511	0.169
hsa-miR-10b-002218	24.18	4.663	0.173
hsa-miR-520e-001119	29.13	25.104	0.202
hsa-miR-199a-000498	30.841	31.679	0.207
hsa-miR-672-002327	21.201	18.443	0.216
hsa-miR-220c-002211	26.929	6.353	0.230
hsa-miR-143-002249	24.579	2.300	0.236
hsa-miR-155-002623	21.94	7.650	0.237
hsa-miR-521-001122	29.621	5.904	0.254
hsa-miR-145-002278	22.411	2.456	0.269
hsa-miR-885-3p-002372	23.735	2.526	0.270
hsa-miR-330-5p-002230	26.161	4.532	0.280
hsa-miR-513-5p-002090	27.58	3.982	0.297
hsa-miR-299-3p-001015	35.444	9.673	0.303
ath-miR159a-000338	30.906	3.907	0.338
hsa-miR-886-3p-002194	23.712	2.375	0.338
hsa-miR-21-000397	22.435	3.703	0.339
hsa-miR-548a-001538	28.281	3.729	0.346
hsa-miR-296-3p-002101	29.375	2.973	0.358
hsa-miR-506-001050	27.018	6.382	0.373
hsa-miR-629-002436	30.876	5.226	0.376
hsa-miR-200a-000502	25.107	3.230	0.379
hsa-miR-517b-001152	23.833	14.173	0.386

hsa-miR-337-5p-002156	25.061	3.210	0.392
hsa-miR-219-2-3p-002390	26.272	2.073	0.393
hsa-let-7d-002283	21.406	2.229	0.399
hsa-miR-582-3p-002399	31.546	2.290	0.403
hsa-miR-548c-001590	31.556	2.045	0.413
hsa-miR-524-5p-001982	34.842	140.542	0.422
hsa-miR-129-000590	32.172	2.008	0.465
hsa-miR-202-002363	23.687	2.451	0.469
hsa-miR-196b-002215	26.883	30.453	0.491
hsa-miR-27a-000408	27.438	2.274	0.501
hsa-miR-520f-001120	25.873	2.158	0.512
hsa-miR-520b-001116	28.439	2.490	0.527
hsa-miR-34c-000428	27.597	3.636	0.585
hsa-miR-20b-001014	20.028	6.052	0.597
hsa-miR-130a-000454	21.465	2.260	0.635
hsa-let-7c-000379	22.778	4.393	0.678
hsa-miR-576-3p-002351	29.435	2.599	0.683
hsa-miR-194-000493	27.98	3.391	0.683
hsa-let-7a-000377	23.838	6.022	0.715
hsa-miR-424-000604	25.771	2.333	0.722
hsa-miR-515-5p-001112	31.74	6.901	0.731
hsa-miR-490-001037	29.635	4.088	0.748
hsa-miR-20a-000580	18.6	3.146	0.756
hsa-miR-27b-000409	26.893	2.049	0.866
hsa-miR-518a-3p-002397	28.288	8.477	0.931
hsa-miR-517a-002402	24.509	11.676	0.990

Table 15:

This table lists all of the TaqMan® miRNA probes that were found to have relative quantification (RQ) values that were reduced greater than 0.5 fold on the 22q11.2 arrays versus the control arrays for TaqMan® Human miRNA Plate B. The Ct (cycle threshold) is the number of cycles it takes the fluorescent signal to cross the threshold and exceed the background level in qRT-PCR. The p-value was obtained by using an unpaired, two-tailed t-test of sample normalization with an assumption of unequal variance.

Target Name	Ct	RQ	P-Value
hsa-miR-770-5p-4395189	37.792	0.126	0.029
hsa-miR-766-4395177	36.683	0.290	0.127
hsa-miR-942-4395298	35.349	0.271	0.192
hsa-miR-92b*-4395454	31.569	0.295	0.197
hsa-miR-767-5p-4395182	33.166	0.344	0.204
hsa-miR-135a*-4395343	37.642	0.046	0.272
hsa-miR-182*-4378066	20.968	0.310	0.325
hsa-miR-936-4395290	34.65	0.068	0.366
hsa-miR-613-4380989	32.984	0.101	0.392
hsa-miR-631-4380971	34.752	0.280	0.395
hsa-miR-32*-4395222	32.993	0.100	0.488
hsa-miR-600-4380963	30.047	0.448	0.520
hsa-miR-519e*-4378084	23.826	0.075	0.594
hsa-miR-622-4380961	24.934	0.141	0.605
hsa-miR-555-4380933	25.685	0.402	0.626
hsa-miR-30e*-4373057	34.886	0.321	0.630
hsa-miR-624*-4380964	34.206	0.275	0.651
hsa-let-7b*-4395515	27.166	0.259	0.654
hsa-miR-18b*-4395421	37.737	0.325	0.665
hsa-miR-593-4395522	23.484	0.423	0.690
hsa-miR-18a*-4395534	33.61	0.317	0.700
hsa-miR-202*-4395473	29.976	0.091	0.713
hsa-miR-640-4386743	25.068	0.125	0.840
hsa-miR-30d-4373059	39.31	0.126	0.858
hsa-miR-937-4395291	31.187	0.332	0.892
hsa-miR-585-4381027	29.339	0.049	0.925
hsa-miR-302c*-4373277	33.695	0.339	0.937

hsa-miR-30e-4395334	32.624	0.347	0.965
hsa-miR-497*-4395479	33.039	0.245	0.992

Table 16: This table lists all of the TaqMan® miRNA probes that were found to have relative quantification (RQ) values that were increased greater than 2.0 fold on the 22q11.2 arrays versus the control arrays for TaqMan® Human miRNA Plate B.
The Ct (cycle threshold) is the number of cycles it takes the fluorescent signal to cross the threshold and exceed the background level in qRT-PCR. The p-value was obtained by using an unpaired, two-tailed t-test of sample normalization with an assumption of unequal variance.

Target Name	Ct	RQ	P-Value
hsa-miR-626-4380966	30.542	15.336	0.087
hsa-miR-181a*-4373086	31.51	5.035	0.125
hsa-miR-562-4380939	25.417	51.598	0.185
hsa-miR-606-4380974	32	9.131	0.187
hsa-miR-584-4381026	29.879	17.374	0.24
hsa-miR-649-4381005	31.551	180.903	0.241
ath-miR159a-4373390	29.265	23.903	0.263
hsa-miR-30c-2*-4395221	28.955	7.879	0.278
hsa-miR-588-4380952	31.374	41.108	0.280
hsa-miR-621-4381001	37.33	19.580	0.281
hsa-miR-633-4380979	24.479	155.904	0.316
hsa-miR-145*-4395260	35.359	7.499	0.332
hsa-miR-380*-4373021	28.594	2.321	0.338
hsa-miR-801-4395183	28.376	47.382	0.338
hsa-miR-632-4380977	30.432	4.000	0.341
hsa-miR-596-4380959	26.736	3.514	0.344
hsa-miR-596-4380959	24.947	3.514	0.344
hsa-miR-30a-4373061	26.002	2.700	0.356
hsa-miR-892b-4395325	32.599	2.028	0.359
hsa-miR-22*-4395412	38.526	2.766	0.363
hsa-miR-768-3p-4395188	29.683	105.605	0.363
hsa-miR-374b*-4395502	29.632	4.178	0.371
hsa-miR-566-4380943	23.287	5.931	0.374
hsa-miR-92a-2*-4395249	36.699	2.422	0.375

hsa-miR-302d-4373063	22.08	3.873	0.386
hsa-miR-498-4373223	32.879	7.112	0.388
hsa-miR-595-4395178	30.71	4.715	0.389
hsa-miR-744*-4395436	30.509	16.985	0.391
hsa-miR-644-4380999	31.261	8.560	0.397
hsa-miR-518c*-4378082	29.481	196.545	0.400
hsa-let-7d*-4378108	31.58	266.541	0.401
hsa-miR-541*-4395311	28.422	4.484	0.408
hsa-miR-519b-3p-4395495	29.771	5.633	0.414
hsa-miR-122*-4395241	27.146	28.847	0.422
hsa-miR-609-4380978	38.335	3.634	0.424
hsa-miR-181c*-4395444	27.272	4.070	0.429
hsa-miR-623-4386740	37.504	340.790	0.429
hsa-miR-943-4395299	27.131	6.750	0.439
hsa-miR-497-4373222	29.184	2.771	0.449
hsa-miR-518f*-4395498	29.851	6.258	0.453
hsa-miR-571-4381016	24.987	2.121	0.466
hsa-miR-599-4380962	34.7	5.689	0.474
hsa-miR-29a*-4395558	34.212	58.498	0.487
hsa-miR-144*-4395259	28.525	3.149	0.489
hsa-miR-206-4373092	28.252	11.696	0.489
hsa-miR-565-4380942	32.969	3.103	0.499
hsa-miR-378-4395354	34.622	2.203	0.525
hsa-miR-646-4381002	31.613	3.485	0.527
hsa-miR-935-4395289	23.435	2.343	0.531
hsa-miR-603-4380972	32.028	2.447	0.546
hsa-miR-769-5p-4395186	23.092	2.728	0.547
hsa-miR-552-4380930	34.186	6.500	0.550
hsa-miR-10b*-4395426	30.737	28.975	0.555
hsa-miR-587-4380950	31.379	5.641	0.569
hsa-miR-33a*-4395247	29.387	9.101	0.573
hsa-miR-923-4395264	33.019	41.748	0.574
hsa-miR-10a*-4395399	28.597	15.469	0.582
hsa-miR-27a*-4395556	25.703	34.445	0.585
hsa-miR-154*-4378065	30.837	7.483	0.586
hsa-miR-567-4380944	28.491	3.603	0.586
hsa-miR-126*-4373269	28.045	6.679	0.588
hsa-miR-650-4381006	31.536	8.188	0.594
hsa-miR-553-4380931	26.976	2.823	0.595
hsa-miR-551b*-4395457	29.634	2.715	0.596
hsa-miR-573-4381018	28.181	11.367	0.602

hsa-miR-934-4395288	30.238	5.776	0.602
hsa-miR-675-4395192	37.743	12.165	0.606
hsa-miR-610-4380980	29.041	11.609	0.628
hsa-miR-24-1*-4395551	28.219	4.172	0.650
hsa-miR-509-3p-4395347	33.396	3.996	0.679
hsa-miR-645-4381000	34.222	3.082	0.710
hsa-miR-662-4381010	32.139	5.261	0.717
hsa-miR-20b*-4395422	22.943	13.672	0.722
hsa-miR-452*-4395441	31.454	25.493	0.729
hsa-let-7a*-4395418	28.533	2.312	0.751
hsa-miR-222*-4395208	25.344	18.787	0.798
hsa-miR-575-4381020	25.59	5.477	0.803
hsa-miR-34b*-4373037	20.772	30.41	0.845
hsa-miR-944-4395300	30.364	3.673	0.872
hsa-miR-583-4381025	34.767	6.797	0.878
hsa-miR-769-3p-4395190	33.378	7.516	0.882
hsa-miR-190b-4395374	28.891	3.689	0.901
hsa-miR-921-4395262	30.237	3.799	0.901
hsa-miR-143*-4395257	26.16	25.599	0.909
hsa-miR-661-4381009	31.743	6.404	0.912
hsa-miR-924-4395265	25.747	11.105	0.933
hsa-miR-518e*-4395482	25.388	24.794	0.937

Table 17: Table 17 shows a list of miRNAs found to have nominally significantly different fold changes on Taqman® Human miRNA cards between miRNA isolated from control cell lines versus 22q11.2 and schizophrenia lines. miRNAs that have been previously implicated in the development of psychiatric disorders.

Target Name	Ct	RQ	P-Value	Previous Evidence
hsa-miR-127-000452	26.800	0.302	0.014	
hsa-miR-185-002271	27.885	0.324	0.026	Downregulated in leukocytes with 22q11.2 deletion syndrome (Sellier et al.); Downregulated in prefrontal cortex and hippocampus of mice with a deletion equivalent to the 22q11.2 deletion (Stark et al.). Downregulated in neurons derived from iPSCs (Zhao et al.)
hsa-miR-200c-002300	27.169	0.385	0.028	
hsa-miR-212-000515	25.728	0.464	0.025	Elevated in dorsolateral prefrontal cortex samples with schizophrenia (Kim et al.); Reduced expression levels in the prefrontal cortex of samples with

				schizophrenia (Perkins et al.)
hsa-miR-323-3p-002227	27.107	0.328	0.017	
hsa-miR-339-3p-002184	24.565	0.341	0.020	Reduced expression in prefrontal cortex samples with bipolar disorder (Moreau et al.)
hsa-miR-369-5p-001021	31.681	0.284	0.021	
hsa-miR-376c-002122	27.896	0.111	0.006	
hsa-miR-382-000572	24.620	0.364	0.030	Elevated in cultures derived from individuals with schizophrenia (Mor et al.)
hsa-miR-425-5p-001516	23.843	0.400	0.038	miR-425 was also observed to have decreased expression in the prefrontal cortex of samples with bipolar disorder (Moreau et al.)
hsa-miR-429-001024	38.948	0.047	0.045	
hsa-miR-431-001979	34.220	0.051	0.016	
hsa-miR-433-001028	24.269	0.401	0.009	

hsa-miR-487b-001285	28.201	0.101	0.006	
hsa-miR-493-002364	29.795	0.467	0.024	
hsa-miR-494-002365	27.217	0.314	0.010	
hsa-miR-510-002241	34.168	0.288	0.013	
hsa-miR-515-3p-002369	30.896	16.258	0.031	
hsa-miR-519c-001163	26.425	57.066	0.027	
hsa-miR-541-002201	35.719	0.022	0.003	
hsa-miR-616-002414	35.157	0.088	0.037	
hsa-miR-655-001612	32.262	0.265	0.008	
hsa-miR-758-001990	35.182	0.117	0.009	
hsa-miR-770-5p-4395189	37.792	0.126	0.029	

hsa-miR-889-002202	30.354	0.228	0.025	Elevated expression in dorsolateral prefrontal cortex samples with bipolar disorder, while here is was observed to be decreased in expression (Kim et al.).

RNA Extraction for Small miRNA Sequencing:

Frozen pellets consisting of 5-10 million cells collected from NSC lines were submitted to RUCDR for small miRNA sequencing. The miRNA extraction for small miRNA sequencing was performed using automated solid phase extraction and bead based chemistries, which allows for the isolation/enrichment of micro RNA species (miRNA) in an aliquot separate from the total RNA. Analytical quality control was then performed on the isolated miRNA. The Qiagen BioRobot Universal fully automated nucleic acid extractor was used isolate small miRNA from the cultured cell pellets. The Universal uses a unique solid-phase extraction technology that utilizes a filter plate to trap and isolate nucleic acids.

Small miRNA Sequencing:

The RUCDR performed Illumina TruSeq Small RNA sequencing. TruSeq Small RNA Library Preparation Kits were used to generate small RNA libraries directly from total RNA. The modified adapters that are included targeted miRNAs that are generated by Dicer processing. Multiplexed sequencing was performed using 48 unique indexes so that miRNAs could be profiled using Illumina sequencing. The indexes were added using a universal amplification reaction, so that ligation bias would be reduced.

Strengths and Weaknesses of Arrays and RNA Sequencing:

Arrays and RNA sequencing both have their own unique advantages and disadvantages. TaqMan® assays have a proven track record consisting of nearly 25 years of use (Holland et al.). Conversely RNA sequencing is still a very new technology, and small RNA sequencing newer still (Motameny et al.). Running TaqMan® probes using arrays is something that can be readily performed in house with cheaper reagents and less expensive equipment. The Illumina small miRNA sequencing requires more expensive equipment and reagents that are not readily available within out lab. This is a large project that must be performed by the RUCDR. The startup costs for running the TaqMan® arrays is less expensive and less labor intensive than setting up an RNA sequencing facility. Furthermore, the data analysis for the TaqMan® arrays is relatively straightforward, while the data analysis for the Illumina small miRNA sequencing is complicated and requires licensed software.

However, the cost of submitting samples for small RNA sequencing to RUCDR and having them perform the sequencing is actually less expensive and less labor intensive than running the replicate number of TaqMan® array assays that are required. Furthermore, the Illumina small miRNA sequencing provides a much more detailed look at all of the miRNAs being expressed within the sample genome wide and can even detect novel miRNAs. RNA sequencing offers increased specificity and sensitivity for the detection of miRNAs (Git et al.). Unlike the TaqMan® arrays, there is no design bias and the results are not limited to a select number of probes that were preselected for inclusion on the arrays.

Small RNA sequencing is not without limitations though. It can be difficult to quantify which signals are background noise, especially when looking at miRNAs that

are present in very small amounts. RNA-seq has also been shown to have difficulties handling repetitive sequences and differentiating between alternative spliced isoforms (Robert and Watson). miRNA sequencing library preparations also have been shown to have preferential presentation of the miRNA complement, which could lead to misleading estimates of miRNA abundance (Linsen and Cuppen). There is also concern about the replicability of RNA sequencing. Since RNA sequencing is still so new, there is not a perfect consensus of how quality control should be performed on RNA sequencing data. This can lead to difficulty replicating the results of RNA sequencing depending on how the studies were performed (Ballouz and Gillis).

The ideal study would involve the usage of small miRNA sequencing, TaqMan® array assays, and individual TaqMan® assays. Running the small miRNA sequencing first seems like the broadest way to start looking at miRNA expression changes. Arrays can be then be custom made using the miRNA sequencing data to confirm the expression levels of those miRNAs. Lastly, individual TaqMan® miRNA probes can be used to validate the expression levels of the most interesting results. .

Chapter 5: Genetic Association and Validations of a microRNA Target Site Polymorphism in H3F3B with Schizophrenia

ABSTRACT

MicroRNAs (miRNAs) are endogenous ~21 nucleotide molecules that play important roles in gene repression. Single nucleotide polymorphisms (SNPs) that disrupt miRNA binding sites make up a new class of functional genetic polymorphisms. A pattern-based miRNA target prediction algorithm called miRSNiPer, capable of predicting target site enhancement or disruption by SNPs, was previously used to construct a panel of 48 putative functional SNPs, which were genotyped and tested for association to schizophrenia spectrum disorders. One SNP (rs1060120) within the 3'UTR of H3F3B, was previously found to be associated with schizophrenia. The predicted miRNA/SNP interaction between the 3' UTR of H3F3B and a targeting miRNA, miR-616, was then functionally validated using luciferase reporter assays.

Introduction

miRNAs

miRNAs are small ~21 nucleotide single-stranded molecules that silence the expression of 20-30% of human genes by binding to the three prime untranslated regions (3'UTRs) of messenger RNA. The 3'UTR is the section of mRNA that immediately follows the translation termination codon (Lewis, Burge, and Bartel). Each 3'UTR is made of a sequence of nucleotides, which in turn can contain single nucleotide polymorphisms (SNPs) (Xie et al.). These SNPs within the 3'UTR are known as miRSNPs and can increase the functionality as well as the genetic diversity of the 3'UTR (Mishra et al.).

H3F3B and miR-616

A target prediction algorithm called miRSNiPer, was previously developed by Moreau et al. (Moreau, 2009) and used to detect target site enhancement or disruption by SNPs in 3'UTRs that are associated with schizophrenia. A SNP (rs1060120) in the 3'UTR of H3F3B was found that was associated with schizophrenia (Figure 24)(Moreau, 2009). H3F3B encodes a core histone protein that regulates chromatin architecture during development(Albig et al.). This SNP (rs1060120) was predicted to enhance the degree of miR-616 binding (Moreau, 2009).

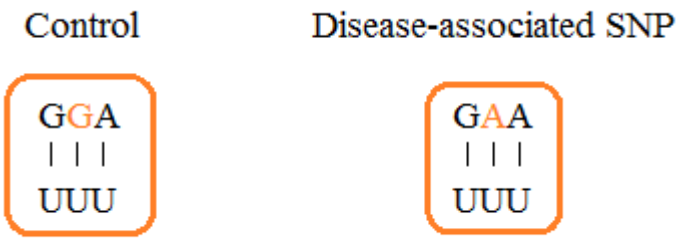
Validating miRSNPs identified using Luciferase Assays

miRNA target prediction software is not sufficient to confirm miRNA/miRSNP interactions. In order to experimentally validate the effects a predicted miRSNP might have on miRNA binding, functional studies must be performed. The predicted

miRNA/miRSNP interactions between H3F3B and rs1060120were validated using 3'UTR reporter assays in which luciferase constructs containing the 3'UTR of H3F3B were co-transfected into cells along with synthetic miRNA to confirm the miRNA/miRSNP interaction. 3'UTR reporter assays have been shown to reliably validate the interaction between miRNAs and their targets(Iliopoulos, Rotem, and Struhl; Shu et al.; Moffatt and Lamont). Constructing 3'UTRs containing the predicted miRSNP allows the functionality of the SNP to be studied directly and the predictions of miRSNiPer to be validated.

The predicted miRNA/SNP interaction between the 3' UTR of H3F3B and a targeting miRNA, miR-616, was tested using reporter assays containing RenSP. RenSP is a luciferase developed by SwitchGear Genomics specifically to function as a reporter gene assay. RenSP has a half-life of ~1 hour enabling a detailed analysis of kinetic responses with a highly robust signal and can be detected using LightSwitch Luciferase Assay Reagents (Figure 25). The Cypridina TK control construct (pTK-CLuc Vector) contained the Cypridina luciferase reporter gene driven by an HSV-TK constitutive promoter (Figure 26). Its signal was detected using the BioLux® Cypridina Luciferase Assay Kit developed by New England Biolabs. Different luciferase genes are used for the Cypridina TK control construct and LightSwitch GoClone experimental constructs. They use different substrates, so there is no cross-reaction between the two reporter genes or their substrates. The Cypridina reporter protein is secreted into the culture media. Using this dual luciferase assay, miR-616 was found to have a stronger interaction with the 3'UTR containing the schizophrenia-associated A allele of rs1060120 than with the control 3'UTR which contained the G allele of rs1060120.

Figure 24: A schematic of the control SNP and the disease-associated SNP (rs1060120) which was previously found to be linked to schizophrenia are shown in the schematic below



Methods:

Cell culture:

HT-1080 fibrosarcoma cells (ATCC) were cultured in Eagle's Minimum Essential Medium Alpha, supplemented with fetal bovine serum to a final concentration of 10% and 50 to 100 I.U./mL penicillin and 50 to 100 (µg/mL) streptomycin. 96-well plates were seeded with HT-1080 fibrosarcoma cells (ATCC) 24 hours prior to transfection to achieve 80% confluence at the time of transfection.

Dual Luciferase Reporter Assays:

Luciferase reporter constructs containing the human reference 3′-untranslated region (3′UTRs) for H3F3B and the 3'UTR for H3F3B containing the A allele at SNP rs1060120, were custom built by Switchgear Genomics (Figure 25 and 28). The LightSwitch Dual Assay System (Switchgear Genomics) was used to normalize for variation between transfection replicates in the H3F3B lines. LightSwitch Assay Reagents were used as directed (Switchgear Genomics).

The co-transfection assay consisted of the following components for a single replicate in a single 96-well plate well: 0.18 µl of FuGENE HD* Transfection Reagent (Promega), 2.15 µl of Opti-MEM (serum free media) (Thermo Fisher Scientific), 1.67 µl of GoClone construct (30ng/ µl)(Switchgear Genomics) and 1.00 µl of Cypridina control construct. The Cypridina solution was prepared so that for 100 samples, 50 µl of the reconstituted substrate (100X solution) was added to 5 ml of BioLux Cypridina Luciferase Assay Buffer (New England Biolabs).This mixture was allowed to sit for 30 minutes and then gently dripped on to the seeded cells.

Wells were co-transfected with either the miR-616-3p mimic or the non-targeting scramble miRNA (Switchgear Genomics) (Figure 27) to yield a final concentration of 20 nM in a total volume of 100 µl/well or an equivalent volume of media as a control. Each transfection condition was performed in triplicate. Plates were incubated at 37 °C for 24 hours post transfection. After transfection the plates were briefly frozen, then allowed to return to room temperature.

On the third day 20 µl of media was collected from each well and 50 µl of the 1X BioLux® Assay added following the manufacturer's protocol (New England Biolabs). The Cypridina reporter was measured using a Veritas Microplate luminometer (Turner Biosystems) to determine the baseline luciferase value.

The RenSP reporter protein remains within the cell during the duration of the protocol. LightSwitch Assay Reagents were used as directed and 100 µl of 1X LightSwitch Assay Reagent was added to each well (Switchgear Genomics). The plates were incubated at room temperature for 30 minutes while protected from light and gently agitated. After 30 minutes, luciferase expressions levels were measured on a Veritas Microplate luminometer (Turner Biosystems).

Figure 25: Both the reference sequence of the H3F3B 3'UTR and the H3F3B 3'UTR containing the A allele of SNP rs1060120 were cloned into the pLightSwitch_3UTR Reporter Vector (Switchgear) shown below. This plasmid also contains RenSP a luciferase reporter gene for measuring expression of the 3'UTR. The plasmid also contains a constitutive promoter.

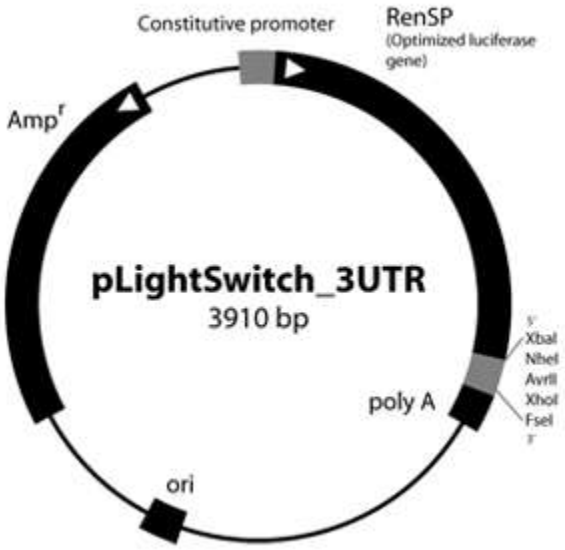


Figure 26: The following figure shows the Cypridina TK control construct (pTK-CLuc Vector). This vector contains the Cypridina luciferase reporter gene driven by an HSV-TK constitutive promoter. Its signal was detected using the BioLux® Cypridina Luciferase Assay Kit developed by New England Biolabs.

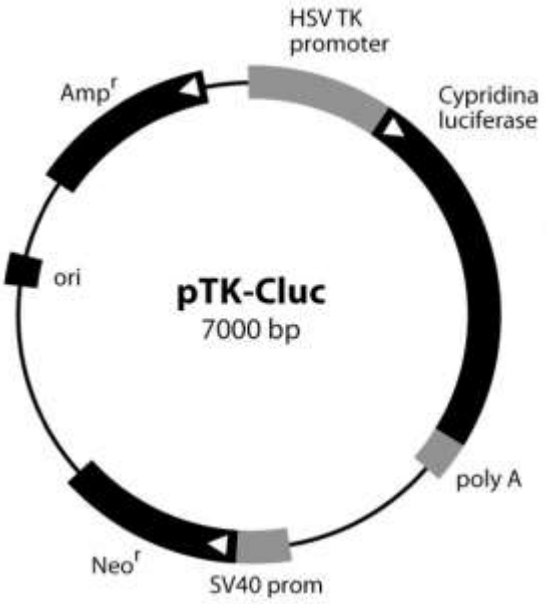


Figure 27: The following figure shows the sequence and structure of the stem loop for miR-616.

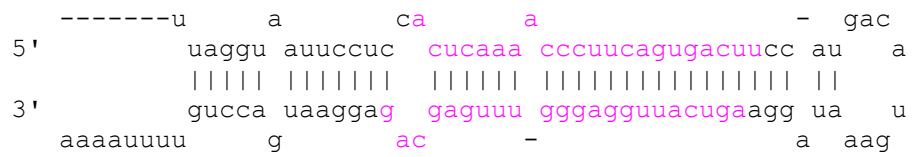


Figure 28: The sequence for the H3F3B 3'UTR is shown below.

Product ID: S811081

Vector name: pLightSwitch_3UTR

Gene symbol: H3F3B

Insert length: 2327

Restriction pair: NHE1XHO1

NOTE: The sequence below is sequence from the reference human genome.

Sequence:

CTAAGAGAGTCACCATCATGCCCAAAGACATCCAGTTGGCTCGCCGGATACGGGGAGAGA
GAGCTTAAGTGAAGGCAGTTTTTATGGCGTTTTGTAGTAAATTCTGTAAAATACTTTGGT
TTAATTTGTGACTTTTTTTTGTAAGAAATTGTTTATAATATGTTGCATTTGTACTTAAGTC
ATTCCATCTTTCACTCAGGATGAATGCGAAAAGTGACTGTTCACAGACCTCAGTGATGTG
AGCACTGTTGCTCAGGAGTGACAAGTTGCTAATATGCAGAAGGGATGGGTGATACTTCTT
GCTTCTCATGATGCATGTTTCTGTATGTTAATGACTTGTTGGGTAGCTATTAAGGTACTA
GAGTTGATAAATGTGTACAGGGTCCTTTTGCAATAAAACTGGTTATGACTTGATCCAAGT
GTTTAACAATTGGGGCTGTTAAGTCTGACCATACATCACTGTGATAGAATGTGGGCTTTT
TCAAGGGTGAAGATACAAGTCTTAACCACAGTGTAACTTACAGTTTCCTTTAAAAAAAAA
AAAAGTAAACCTGGCAGCTATAGAATACACTATGTGCATTTATAATAGCTATTTTATATA
TTGTAGTATCAACATTTTTAAATTAAATGTTTTACATTCACAAGTGGTGGGGAGTCTTGT
CATTAAGGTGTGTGTAATTTAGAGTCCAGTTGGTTTTCTTCTGACTGCACTTGTTCTCAT
AGTAGTAAAATGCTATGCGCATTTATACCTTGCATAAGTCCTCATTCTACCACATGTTAA
CCCTCTAGCTGATAATGCAAACACTAACTGGGGGATTTTATTTATAAGGGCTCTAGAAAA
AACGAGTTATTCACACCAGCATCATCTTAACTAACATTCTGAACTAGTTAGTGCAGCTTT
TCATTGTGTTGTGTGGTTGGTCTCATAACTAGGTTGAGTTTTTCTCCTCTGCTGAGGAAA
CAGTACCGAAGTTCTTTTTCTTGTGGCATTTGTATTATAAAAACTTGGTGTGGGGGAGGA
GCACAAAACTCCAGCCCACTGAACCTCTGCCAATTAAGATGGTGTTGGGTTAGGTTACAT
CTGGTTACTGTCCTGGGAAAATCATTTTTATAGAGATGGCCTTCCAAGTGGTTTTAAAAT
TTACTGAAGTTTTTAGGTCAATTATGTATGTTGACTAAATTTACAAATAAACTTGTTTAT
CCAACTAAGTGTCCAAAACCTAAATTGAATGTACTAAGTTTTCACATGTCCCATTATCTA
GGTCCTTGTATACTAATGTTTTGAACTTAGATCATTTCAGGTGTTGTTTGGTGGATAAAG
GAACCTTTTATTTATAAAGATACTGTAGAAAGCATGTGAACAGCTCTCTGCTTGATTAAG
ATGCCATAATAGTGCTGTATTTGCAGTGTGGGCTAAGACAAAGTATATTAATAAGCTTTT
CAGCCCCCCCACTCCCGTTCCGTAGTGTAGAAGCCCACAGGTGTAGAACTCAGTCTTAAA
CTTCAGTATGAAACCAGTTTCCTTGTGCGATGATGGCCACTAAAGCATAGTACGTGGATG
TCAGTGAGACAGCATGAGAGCCAGCAGTCATCAAAGCGTTCCACGTTTGAAGTTAGCAAC
TGCTTAAAGTTATACCCCATTAAAATTGCTTTCTCAAAAGTTTGGGTTAGTTTCAAATGT
GATATTTTGGAGGGAAGGTAAAGTAGGTATCTTTCAGGTCGTGATAATGAGCTCCTATGA
AAGGATGCAATATAATGACCCGCTTTTCTAGAAAGTTCATAATCAGCTCTGGAACAAGCA
CACTTGATTCCTCACTGTGCTTCAGAATGAGATTAAGATCAGATGTTGGAACGTGCTATG
CTGTAGCGTGTCTGGAAACAAAGTACACAAACCTGGCTACGGTGATGAGTTAGCTTCTGC
TTACTACCTGTGACAACCCAAGTGGGTGACACTAGTGAACCTTCTCCAGTCTGCAGGCTG
GCATAGAAGGCTCTTAGATTATATTGGGCAGCTTGCAATCTGCCGAAGCAGTGACTTGCA
TTTCCACACTTGGCTTGAGCACTCAACCCAGAAGGCGAAGATAGCTTTTGGTTGTAGGCG

GCTTCCTGTATGGGATATCCCTCGGTAAGGGTAAAGGAGCAGAGGCAAAGGAGAAAAGCA
GAAGTTGCAGCTGATGCAGGTATCCTATGCCCTTGATGGATGAGACTAAAATAAAATTTT
TGAAGTTAAGATTTTCTCATTTCTCAATGGTTTTCATTCAAATGATAATGTCTTTCAAAA
AGTTTGGGGCAATGTCCAAAAAGCACTGACTTCTGATTCATCTGGGG

Statistical Methods Dual Luciferase Assay:

A randomized complete block design was used to analyze these results. An ANOVA was used using anova(lm(y ~ trial + trt, data =)) in R: A language and environment for statistical computing. Tukey's HSD (honest significant difference) test was also performed using TukeyHSD(aov(y ~ trial + trt))

Results:

The 3'UTR for H3F3B contains the SNP rs1060120. The A allele of rs1060120 has been previously shown to be associated with schizophrenia, while the G allele of rs1060120 is not associated with illness (Moreau, 2009). miR-616 binds to the H3F3B 3'UTR, and it is believed that the A allele found in individuals with schizophrenia increases the strength of this binding. To confirm this prediction, reporter constructs (Switchgear Genomics) were used that contain the 3'UTR for H3F3B with either the A allele or the G allele. Initially single luciferase assays were performed. While overall results with this method were similar to the dual luciferase assays, the run-to-run variability in magnitude of the single assay results were much larger, complicating the analysis of these data. The methods and results for these experiments are however contained in the Supplementary Methods and Supplementary Results section.

The LightSwitch Dual Assay System (Switchgear Genomics) was used to normalize for variation between experimental replicates in the H3F3B lines due to transfection variability by utilizing a co-transfection control. Each construct was co-transfected with either miR-616 or a nonspecific scrambled miRNA mimic (Switchgear

Genomics). Three independent transfections were performed consecutively over the course of several weeks. The results of the dual luciferase assay are shown in Table 18.

A randomized complete block design was used to analyze these results. The results of the ANOVA and Tukey's HSD (honest significant difference) test are shown below. There was evidence of a treatment effect (Table 19 and 20). A significantly greater reduction of luciferase expression was observed when the A construct was co-transfected with miR-616 than when the G construct was co-transfected with mir-616 (Table 18, Figure 29).

Table 18: This table shows the results of the dual luciferase assay, which used RenSP Luciferase in addition to pTK-CLuc for normalization. The results from each RnSP Luciferase measurement and each pTK-CLuc measurement are shown below. Additionally, the normalized values for each trial are shown.

Trial 1 RenSP Luciferase:

A	A SCR	A 616	G	G SCR	G 616
212079	72988	185724	176742	45032	32472
202808	61982	125444	214614	79177	35527
65181	104980	36767	141431	13299	25078

Trial 1 pTK-CLuc:

A	A SCR	A 616	G	G SCR	G 616
65065	17503	241201	45202	11429	15316
47606	14054	137850	56927	20891	13061
23271	29655	54876	29775	4067	10071

Normalized Trial 1:

	A	A SCR	A 616	G	G SCR	G 616
	3.26	4.17	0.77	3.91	3.94	2.12
	4.26	4.41	0.91	3.77	3.79	2.72
	2.8	3.54	0.67	4.75	3.27	2.49
Average	3.44	4.04	0.78333	4.14333	3.66667	2.4433333
STDEV	0.746458304	0.4493328	0.12055	0.53003	0.35162	0.30271

Trial 2 RenSP Luciferase:

A	A SCR	A 616	G	G SCR	G 616
174389	58446	35191	33229	7061	57737
177200	51133	30564	28058	11717	196491
146418	46985	45985	33165	10868	120315
199489	70382	53129	31291	17050	151639

Trial 2 pTK-CLuc:

A	A SCR	A 616	G	G SCR	G 616
56404	19940	56029	10384	1929	24569
35603	12872	31315	5622	2858	77359
45642	12498	38339	8333	2794	75197
44535	26184	37465	9203	4971	64254

Normalized Trial 2:

	A	A SCR	A 616	G	G SCR	G 616
	3.09	2.93	0.63	3.2	3.66	2.35
	4.97	3.97	0.98	4.99	4.1	2.54
	3.21	3.76	1.2	3.98	3.89	1.6
	4.48	2.69	1.42	3.4	3.43	2.36
Average	3.9375	3.34	1.0575	3.8925	3.77	2.2125
STDEV	0.932358121	0.622863	0.33689	0.80297	0.28925	0.4175624

Trial 3 RenSP Luciferase:

A	A SCR	A 616	G	G SCR	G 616
72981	112184	101763	161608	94142	89569
65575	60240	77790	83080	86792	61560
76170	56075	71553	88805	91144	55136
71173	89235	58025	66094	79071	65645

Trial 3 pTK-CLuc:

A	A SCR	A 616	G	G SCR	G 616
21785	26710	107119	26065	26296	43906
19230	26892	57622	18022	25527	22884
16380	14953	58650	27241	27370	23462
14436	19108	30864	15264	16823	29569

Normalized Trial 3:

	A	A SCR	A 616	G	G SCR	G 616
	3.35	4.2	0.95	6.2	3.58	2.04
	3.41	2.24	1.35	4.61	3.4	2.69
	4.65	3.75	1.22	3.26	3.33	2.35
	4.93	4.67	1.88	4.33	4.7	2.22

Average	4.085	3.715	1.35	4.6	3.7525	2.325
STDEV	0.822415143	1.0526316	0.39064	1.21499	0.64038	0.2745299

Figure 29: The normalized values for all three trials of the dual luciferase assay are shown below. The reduction of luciferase activity as a result of miR-616 treatment was greater in the A allele than in the G allele. The scramble miRNA slightly reduced luciferase expression, but was not significantly different from the untreated constructs. Error bars represent standard deviation between the three trials performed using the dual luciferase assay.

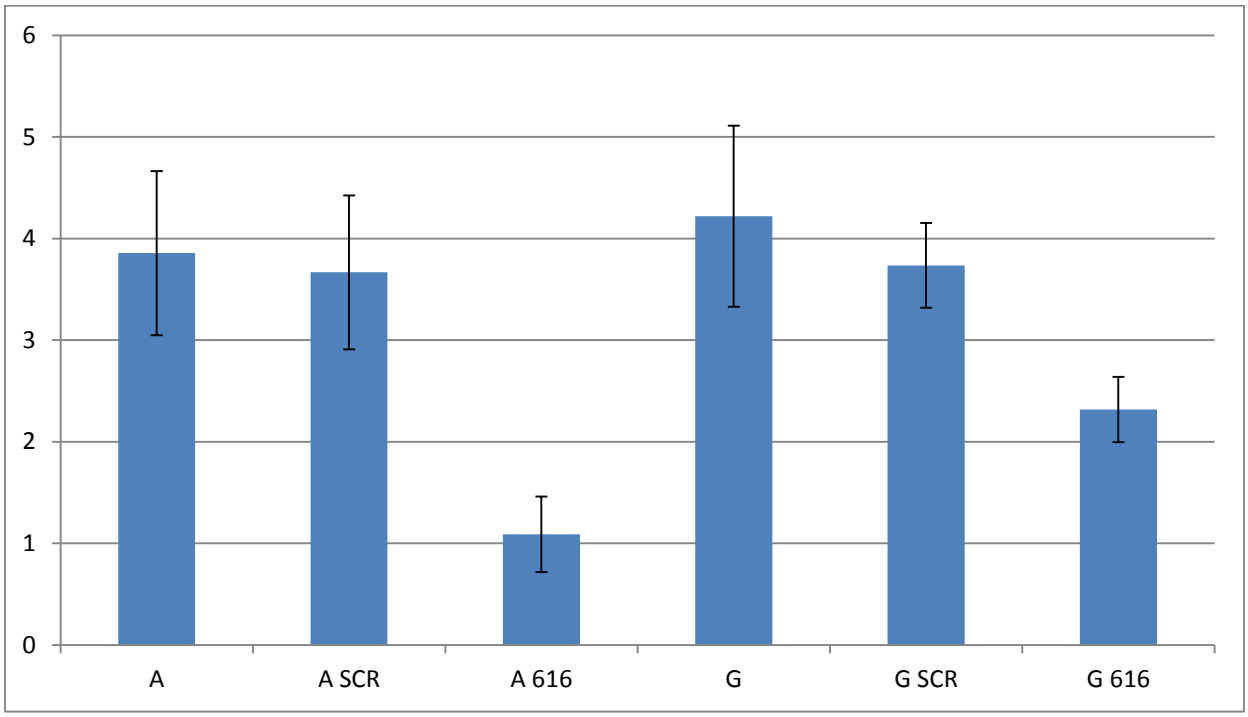


Table 19: Analysis of Variance table. The following table shows that there was a highly significant treatment effect, since the observed F value was greater than the critical value. There was not a significant difference between the trials, since the observed F value was less than the critical value. There was no evidence of a trial effect.

	Df	Sum Sq	Mean Sq	F Value	Pr(>F)
trial	2	0.966	0.483	1.1972	0.3094
treatment	5	79.093	15.8187	39.2095	<2e-16
Residuals	58	23.399	0.4034		

Table 20: Tukey's HSD test was performed on the data collected from the dual luciferase assay to find trial and condition means that are significantly different from each other.

Conditions Compared	difference	lower	upper	p adjusted
A.616-A	-2.76636364	-3.5645496	-1.9681776	0
A.SCR-A	-0.18909091	-0.9872769	0.6090951	0.9814351
G-A	0.36272727	-0.4354587	1.1609133	0.7620481
G.616-A	-1.53909091	-2.3372769	-0.7409049	0.0000066
G.SCR-A	-0.12	-0.918186	0.678186	0.997735
A.SCR-A.616	2.57727273	1.7790867	3.3754587	0
G-A.616	3.12909091	2.3309049	3.9272769	0
G.616-A.616	1.22727273	0.4290867	2.0254587	0.000412
G.SCR-A.616	2.64636364	1.8481776	3.4445496	0
G-A.SCR	0.55181818	-0.2463678	1.3500042	0.3344472
G.616-A.SCR	-1.35	-2.148186	-0.551814	0.0000843
G.SCR-A.SCR	0.06909091	-0.7290951	0.8672769	0.999845
G.616-G	-1.90181818	-2.7000042	-1.1036322	0
G.SCR-G	-0.48272727	-1.2809133	0.3154587	0.4848827
G.SCR-G.616	1.41909091	0.6209049	2.2172769	0.0000336

Discussion:

A SNP (rs1060120) in the H3F3B 3'UTR has previously been shown to be associated with schizophrenia (Moreau, 2009). This schizophrenia associated allele was thought to increase the binding of miR-616 versus the allele not associated with illness. This prediction was validated using 3'UTR reporter assays in which luciferase constructs containing the 3'UTR of H3F3B were co-transfected into cells along with synthetic mir-616 to confirm this interaction (Switchgear Genomics).

3'UTRs for both the reference H3F3B G allele and the schizophrenia associated A allele of rs1060120 were each separately transfected into HT-1080 cells along with the Cypridina control vector. Transfections of the 3'UTRs and the miR-616 miRNA were performed in three independent consecutive trials. A nonspecific scrambled miRNA mimic with a nonsense sequence was used as a transfection control in a series of replicates and co-transfected with each UTR. The luciferase expression varied between each trial, likely due to transfection efficiency of the 3'UTR constructs. The control Cypridina was used to normalize for these transfection differences. Initially, the dual luciferase assay was not used and the single luciferase assay was performed. The results of these trials that illustrate the differences between trials in transfection efficiency are shown in the supplemental results.

The normalized values for all three trials of the dual luciferase assay are shown (Table 18). The untreated A and G 3' UTR constructs transfected alone have the highest expression of luciferase. The scramble miRNA reduced luciferase expression, but it was not significantly different from the untreated constructs. This minor reduction can likely

be attributed to the slightly toxic effect of including miRNA mimic in the transfection. The death of luciferase expressing cells early in the transfection would reduce luciferase expression. It is also possible, although less likely, the result of non-specific miRNA targeting. The reduction of luciferase as a result of miR-616 treatment was greater in the G allele than in the A allele ($P = 0.000412$; Figure 29 and Table 19 and 20).

In order to better analyze the dual luciferase data a randomized complete block design was used. The values between each trial varied enough that this model was used rather than simply taking the mean of each trial. A randomized complete block design is a model that is ideal for estimating and comparing treatment means. Blocking was used to diminish the effects of variation among units. A unit here represents the luciferase values measured for each condition and replicate. Due to unexpected variation observed in the data between trials, it was necessary to diminish the effects of variation between each trial in order to analyze the results of multiple trials. The variation between trials appears to be the result of differences in transfection efficiency between trials.

Blocks are groups of units that are formed so that units within the blocks are as nearly homogenous as possible. In this case blocks represent the three experimental trials of an individual condition. The six different conditions consisting of each 3'UTR with or without the SNP alone and each 3'UTR individually treated with miR-616 or the scramble miRNA are the treatments. The treatments are then randomly assigned to units within the blocks. Treatment effects are fixed, but the block effects are considered to be random because the blocks in the experiment represent a small subset of the larger set of blocks. Inferences about the treatments are then made about a larger set of blocks than what is actually represented in the measured data for the three trials. This model that

contains a randomized block design and fixed effects for treatment contributions and random effects for block contributions it was constitutes a mixed model and adjusts for random effects that may have occurred throughout these three experimental trials.

Tukey's honest significance test was used to compare the means of each possible pair of conditions tested in the dual luciferase assays. Tukey's honest significance is a single-step multiple comparison procedure used in conjunction with an ANOVA. The set of all pairwise comparisons is then used to find means that are significantly different from each other by looking for differences between two means that is greater than the expected standard error. The Tukey method is also ideal when there are unequal sample sizes between trials. The results of the ANOVA and the Tukey's honest significance test show that there is significant difference between the treatments applied to each condition, but not a significant difference between the trials conducted for each experiment.

The greater reduction by miR-616 versus the scramble miRNA observed in both the RenSP luciferase assays and the dual luciferase assays demonstrate that miR-616 is specially targeting the H3F3B 3'UTR and not randomly targeting a 3'UTR. The reduction of luciferase expression when the schizophrenia-associated A allele of rs1060120 was present is greater than the reduction observed when the G allele was present. This suggests that the A allele could be altering the function of miR-616 in individuals with schizophrenia. It is therefore possible that an alteration of this miRNA/3'UTR interaction could be contributing to the development of schizophrenia. This altered interaction would lead to a decrease in H3F3B which would likely impact the function of histones and their epigenetic modifications, which have already been implicated as being important for the development of schizophrenia(Guidotti et al.; Huang et al.; Morishita et al.).

Since this miRNA/3'UTR interaction has only been tested in cellular models, more work is needed to determine if this interaction is indeed occurring in individuals with schizophrenia. miRNAs have long been believed to play a role in the development of schizophrenia (Perkins et al.; Natalie J. Beveridge et al.). Further understanding the relationship between this miRNA and its target 3'UTR will extend knowledge of miRNA targeting interactions, which could lead to greater understanding about the development of schizophrenia.

Chapter 6: Conclusions

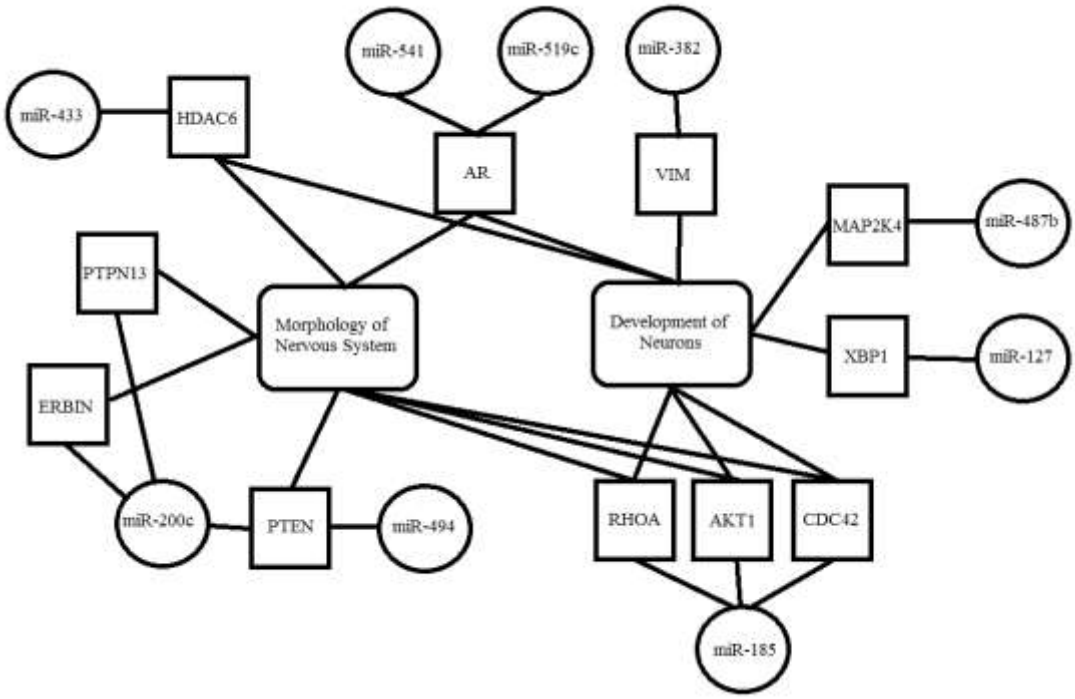
Schizophrenia is a common complex and poorly understood disease. The genetic factors that contribute to the development of schizophrenia are still not fully understood. The 25% increase in the risk of schizophrenia as a result of 22q11.2 therefore is an important risk factor for the development of schizophrenia worthy of continued study. At the outset of this research, miRNAs were found to be disrupted in a mouse model of 22q11.2 as a result of DGCR8 for the first time (Stark et al.). Human studies of post mortem samples have shown that a number of miRNAs have been found to be dysregulated in individuals with schizophrenia (Table 1). These studies have expanded into examining the role of miRNAs in the relationship between schizophrenia and 22q11.2 deletion syndrome in human studies. Studies of human circulating blood and neural stem cell studies, in addition to our own, have revealed miRNAs that are disrupted as a result of the 22q11.2 deletion and are now implicated as having important roles in the development of schizophrenia(Sellier et al.; Zhao et al.).

We have refined existing protocols for developing iPSCs and NSCs. We have acquired all of the publically available 22q11.2 deletion sample cell lines, as well as identified additional 22q11.2 lines within the NIMH genomics repository that have been previously unreported. Using these cell lines we have made new iPSC and NSC lines that contain the 22q11.2 deletion. A major part of this research has also been developing a rigorous quality control workflow to ensure that the cell lines involved in this study are properly characterized and maintained. The risk of developing chromosomal abnormalities in stem cell lines was observed to be a very real and continuous threat to the proper study of chromosomal abnormalities in human stem cell models. We believe

we have developed careful methods to monitor the emergence of these unwanted chromosomal abnormalities so that problematic cell lines could be excluded and so prevented from confounding the miRNA profiling data.

We have also further characterized the miRNAs that are dysregulated in these NSC cell lines as a result of the 22q11.2 deletion syndrome. We have found a small subset of miRNAs that are significantly dysregulated. This subset of miRNAs contains six miRNAs that have been previously been reported as being important for schizophrenia (Table 17). This subset of miRNAs was analyzed using Ingenuity Pathway Analysis (IPA) to perform an analysis to determine what genes have been confirmed as targets of these miRNAs (Qiagen). The list of genes that these miRNAs were confirmed to target was then analyzed using an IPA core analysis. Ten of these miRNAs, again including miR-185, were found to target genes that are important for the development of neurons and the morphology of the nervous system. Since schizophrenia is a neurodevelopmental disorder, these miRNAs could play a role in the development of schizophrenia.

Figure 30: This figure shows the results of an Ingenuity Pathways Analysis (IPA) Core Analysis of miRNAs and their confirmed gene targets. All of these miRNAs were found to have significantly altered expression in the TaqMan® Arrays performed using human NSCs from individuals diagnosed with 22q11.2 DS and schizophrenia. The pathways shown below are ones with potential importance for schizophrenia and contain genes important for the morphology of the nervous system and the development of neurons.



Lastly, this research has validated a predicted miRNA target interaction that is important for schizophrenia. The importance of a particular miRSNP (rs1060120) in H3F3B has been validated by using luciferase assays and constructs containing this miRSNP. A microRNA target site containing a polymorphism in H3F3B that was previously found to be associated with schizophrenia (Moreau, 2009) was interrogated using the Switchgear Dual Luciferase assays and the schizophrenia-associated allele was observed to significantly enhance the regulatory effects of miR-616.

The newly predicted miRNAs that have been found to be dysregulated as a result of disruptions to miRNA biogenesis as a result of the 22q11.2 deletion will need to be studied further to identify their potential targets. These miRNA target interactions will have to be verified. Additionally, it will be important to search for SNPs in the 3'UTRs of the genes regulated by these dysregulated miRNAs to see if any of these SNPs themselves are associated with schizophrenia. Any of these SNPs that are found will need to be validated using the dual luciferase assays that were used to confirm the H3F3B rs1060120 miR-616 interaction.

References:

Adlakha, Yogita K, and Neeru Saini. "Brain microRNAs and Insights into Biological Functions and Therapeutic Potential of Brain Enriched miRNA-128." *Molecular Cancer* 13.1 (2014): 33. Web.

Albig, W et al. "The Human Replacement Histone H3.3B Gene (H3F3B)." *Genomics* 30.2 (1995): 264–72. Web.

Aleman, Andre, René S. Kahn, and Jean-Paul Selten. "Sex Differences in the Risk of Schizophrenia: Evidence from Meta-Analysis." *Archives of General Psychiatry* 60.6 (2003): 565–71. Web.

Arnold, Steven E., Konrad Talbot, and Chang G. Hahn. "Neurodevelopment, Neuroplasticity, and New Genes for Schizophrenia." *Progress in Brain Research* 2004: 319–345. Web.

Babcock, Melanie et al. "Shuffling of Genes within Low-Copy Repeats on 22q11 (LCR22) by Alu-Mediated Recombination Events during Evolution." *Genome Research* 13.12 (2003): 2519–2532. Web.

Bailey, Jeffrey a et al. "Human-Specific Duplication and Mosaic Transcripts: The Recent Paralogous Structure of Chromosome 22." *The American Journal of Human Genetics* 70 (2002): 83–100. Web.

Ballouz, Sara, and Jesse Gillis. "AuPairWise: A Method to Estimate RNA-Seq Replicability through Co-Expression." *PLoS computational biology* 12.4 (2016): e1004868. Web.

Ban, H. et al. "Efficient Generation of Transgene-Free Human Induced Pluripotent Stem Cells (iPSCs) by Temperature-Sensitive Sendai Virus Vectors." *Proceedings of the National Academy of Sciences* 108.34 (2011): 14234–14239. Web.

Banigan, Meredith G. et al. "Differential Expression of Exosomal microRNAs in Prefrontal Cortices of Schizophrenia and Bipolar Disorder Patients." *PLoS ONE* 8.1 (2013): n. pag. Web.

Bartel, David P. "MicroRNAs: Target Recognition and Regulatory Functions." *Cell* 2009: 215–233. Web.

Bassett, A S et al. "22q11 Deletion Syndrome in Adults with Schizophrenia." *American journal of medical genetics* 81.4 (1998): 328–37. Web.

Bassett, Anne S., Gregory Costain, et al. "Clinically Detectable Copy Number Variations in a Canadian Catchment Population of Schizophrenia." *Journal of Psychiatric Research* 44.15 (2010): 1005–1009. Web.

Bassett, Anne S., Christian R. Marshall, et al. "Copy Number Variations and Risk for Schizophrenia in 22q11.2 Deletion Syndrome." *Human Molecular Genetics* 17.24 (2008): 4045–4053. Web.

Bassett, As. "Practical Guidelines for Managing Patients with 22q11. 2 Deletion Syndrome." *J Pediatr.* 17.2 (2011): 281–294. Web.

Beckmann, H, and E Franzek. "The Genetic Heterogeneity Of 'schizophrenia.'" *The world journal of biological psychiatry : the official journal of the World Federation of Societies of Biological Psychiatry* 1.1 (2000): 35–41. Web.

Beckmann, H, E Franzek, and G Stöber. "Genetic Heterogeneity in Catatonic Schizophrenia: A Family Study." *American journal of medical genetics* 67.3 (1996):

289–300. Web.
Beveridge, N J et al. “Schizophrenia Is Associated with an Increase in Cortical microRNA Biogenesis.” *Molecular psychiatry* 15 (2010): 1176–1189. Web.
Beveridge, Natalie J. et al. “Dysregulation of miRNA 181b in the Temporal Cortex in Schizophrenia.” *Human Molecular Genetics* 17.8 (2008): 1156–1168. Web.
Braus, D. F. et al. “Functioning and Neuronal Viability of the Anterior Cingulate Neurons Following Antipsychotic Treatment: MR-Spectroscopic Imaging in Chronic Schizophrenia.” *European Neuropsychopharmacology* 12.2 (2002): 145–152. Web.
Brennand, Kristen J et al. “Modelling Schizophrenia Using Human Induced Pluripotent Stem Cells.” *Nature* 473 (2011): 221–225. Web.
Buchanan, R W, and W T Carpenter. “Domains of Psychopathology: An Approach to the Reduction of Heterogeneity in Schizophrenia.” *The Journal of nervous and mental disease* 182.4 (1994): 193–204. Print.
Castle, D. et al. “The Incidence of Operationally Defined Schizophrenia in Camberwell, 1965-84.” *British Journal of Psychiatry* 159.DEC. (1991): 790–794. Web.
Chiang, H. Rosaria et al. “Mammalian microRNAs: Experimental Evaluation of Novel and Previously Annotated Genes.” *Genes and Development* 24.10 (2010): 992–1009. Web.
Chow, E W et al. “Structural Brain Abnormalities in Patients with Schizophrenia and 22q11 Deletion Syndrome.” *Biol Psychiatry* 51.3 (2002): 208–215. Web.
Cook, Edwin H, and Stephen W Scherer. “Copy-Number Variations Associated with Neuropsychiatric Conditions.” *Nature* 455.October (2008): 919–923. Web.
Costain, Gregory, and Anne S. Bassett. “Clinical Applications of Schizophrenia Genetics: Genetic Diagnosis, Risk, and Counseling in the Molecular Era.” *Application of Clinical Genetics* 2012: 1–18. Web.
Costain, Gregory, Donna M. Mcdonald-Mcginn, and Anne S. Bassett. “Prenatal Genetic Testing with Chromosomal Microarray Analysis Identifies Major Risk Variants for Schizophrenia and Other Later-Onset Disorders.” *American Journal of Psychiatry* 2013: 1498. Web.
Davidsson, Pia et al. “The Synaptic-Vesicle-Specific Proteins rab3a and Synaptophysin Are Reduced in Thalamus and Related Cortical Brain Regions in Schizophrenic Brains.” *Schizophrenia Research* 40.1 (1999): 23–29. Web.
Davis, Kenneth L. et al. “Dopamine in Schizophrenia: A Review and Reconceptualization.” *American Journal of Psychiatry* 148.11 (1991): 1474–1486. Web.
de la Chapelle, a et al. “A Deletion in Chromosome 22 Can Cause DiGeorge Syndrome.” *Human genetics* 57.3 (1981): 253–6. Web.
Dennert, J W, and N C Andreasen. “CT Scanning and Schizophrenia: A Review.” *Psychiatr.Dev.* 1 (1983): 105–122. Print.
Driscoll, D a et al. “Prevalence of 22q11 Microdeletions in DiGeorge and Velocardiofacial Syndromes: Implications for Genetic Counselling and Prenatal Diagnosis.” *Journal of medical genetics* 30.10 (1993): 813–817. Web.
Driscoll, Deborah a et al. “Deletions and Microdeletions of 22q11.2 in Velo-Cardio-Facial Syndrome.” *American Journal of Medical Genetics* 44.2 (1992): 261–268. Web.

Dunham, I et al. “The DNA Sequence of Human Chromosome 22.” *Nature* 402.6761 (1999): 489–495. Web.

Edelmann, Lisa, Raj K. Pandita, and Bernice E. Morrow. “Low-Copy Repeats Mediate the Common 3-Mb Deletion in Patients with Velo-Cardio-Facial Syndrome.” *The American Journal of Human Genetics* 64 (1999): 1076–1086. Web.

Emanuel, Beverly S. “Molecular Mechanisms and Diagnosis of Chromosome 22q11.2 Rearrangements.” *Developmental Disabilities Research Reviews* 2008: 11–18. Web.

Fuchs, E., and D W Cleveland. “A Structural Scaffolding of Intermediate Filaments in Health and Disease.” *Science* 279.5350 (1998): 514–519. Web.

Fung, W L A et al. “Elevated Prevalence of Generalized Anxiety Disorder in Adults With 22q11.2 Deletion Syndrome.” *Developmental disabilities research reviews* 167.8 (2010): 998. Web.

Gibson, Greg, and Günter Wagner. “Canalization in Evolutionary Genetics: A Stabilizing Theory?” *BioEssays* 2000: 372–380. Web.

Git, Anna et al. “Systematic Comparison of Microarray Profiling, Real-Time PCR, and next-Generation Sequencing Technologies for Measuring Differential microRNA Expression.” *RNA (New York, N.Y.)* 16.5 (2010): 991–1006. Web.

Goldmuntz, Elizabeth et al. “Frequency of 22q11 Deletions in Patients with Conotruncal Defects.” *Journal of the American College of Cardiology* 32.2 (1998): 492–498. Web.

Goodship, J et al. “A Population Study of Chromosome 22q11 Deletions in Infancy.” *Archives of disease in childhood* 79.4 (1998): 348–351. Web.

Green, Michael F et al. “Neurocognitive Deficits and Functional Outcome in Schizophrenia: Are We Measuring The ‘right Stuff’?” *Schizophrenia Bulletin* 26.1 (2000): 119–136. Web.

Gregory, Richard I et al. “The Microprocessor Complex Mediates the Genesis of microRNAs.” *Nature* 432.7014 (2004): 235–240. Web.

Guella, Ilaria et al. “Analysis of miR-137 Expression and rs1625579 in Dorsolateral Prefrontal Cortex.” *Journal of Psychiatric Research* 47.9 (2013): 1215–1221. Web.

Guidotti, A et al. “Decrease in Reelin and Glutamic Acid decarboxylase67 (GAD67) Expression in Schizophrenia and Bipolar Disorder: A Postmortem Brain Study.” *Archives of general psychiatry* 57 (2000): 1061–1069. Web.

Guna, Alina, Nancy J Butcher, and Anne S Bassett. “Comparative Mapping of the 22q11.2 Deletion Region and the Potential of Simple Model Organisms.” *Journal of neurodevelopmental disorders* 7.1 (2015): 18. Web.

Guo, Zhiyun et al. “Genome-Wide Survey of Tissue-Specific microRNA and Transcription Factor Regulatory Networks in 12 Tissues.” *Scientific reports* 4 (2014): 5150. Web.

Han, Jinju et al. “Molecular Basis for the Recognition of Primary microRNAs by the Drosha-DGCR8 Complex.” *Cell* 125.5 (2006): 887–901. Web.

Hansen, Thomas et al. “Brain Expressed microRNAs Implicated in Schizophrenia Etiology.” *PLoS ONE* 2.9 (2007): n. pag. Web.

Harrison, P J, and D R Weinberger. “Schizophrenia Genes, Gene Expression, and Neuropathology: On the Matter of Their Convergence.” *Molecular psychiatry* 10.1 (2005): 40–68; image 5. Web.

He, Miao et al. “Cell-Type-Based Analysis of MicroRNA Profiles in the Mouse Brain.”

Neuron 73.1 (2012): 35–48. Web.
Holland, P M et al. "Detection of Specific Polymerase Chain Reaction Product by Utilizing the 5'----3' Exonuclease Activity of Thermus Aquaticus DNA Polymerase." *Proc Natl Acad Sci U S A* 88.16 (1991): 7276–7280. Web.
Hornstein, Eran, and Noam Shomron. "Canalization of Development by microRNAs. TL - 38 Suppl." *Nature genetics* 38 Suppl V.May (2006): 4. Web.
Huang, Hsien-Sung et al. "Prefrontal Dysfunction in Schizophrenia Involves Mixed-Lineage Leukemia 1-Regulated Histone Methylation at GABAergic Gene Promoters." *Journal of Neuroscience* 27.42 (2007): 11254–11262. Web.
Iliopoulos, Dimitrios, Asaf Rotem, and Kevin Struhl. "Inhibition of miR-193a Expression by Max and RXRα Activates K-Ras and PLAU to Mediate Distinct Aspects of Cellular Transformation." *Cancer Research* 71 (2011): 5144–5153. Web.
Kaminsky, Erin B et al. "An Evidence-Based Approach to Establish the Functional and Clinical Significance of CNVs in Intellectual and Developmental Disabilities." *Genetic Medicine* 13.9 (2013): 777–784. Web.
Karayiorgou, M et al. "Schizophrenia Susceptibility Associated with Interstitial Deletions of Chromosome 22q11." *Proceedings of the National Academy of Sciences of the United States of America* 92.17 (1995): 7612–7616. Web.
Kates, Wendy R. et al. "The Neural Correlates of Non-Spatial Working Memory in Velocardiofacial Syndrome (22q11.2 Deletion Syndrome)." *Neuropsychologia* 45.12 (2007): 2863–2873. Web.
Kato, Chieko et al. "Molecular Genetic Studies of Schizophrenia: Challenges and Insights." *Neuroscience Research* 2002: 295–304. Web.
Kelley, Richard I. et al. "The Association of the DiGeorge Anomalad with Partial Monosomy of Chromosome 22." *The Journal of Pediatrics* 101.2 (1982): 197–200. Web.
Kent, W. J. et al. "The Human Genome Browser at UCSC." *Genome Research* 12.6 (2002): 996–1006. Web.
Keshavan, Matcheri S. et al. "Premorbid Indicators and Risk for Schizophrenia: A Selective Review and Update." *Schizophrenia Research*. Vol. 79. N.p., 2005. 45–57. Web.
Kiley-Brabeck, Karen, and Christina Sobin. "Social Skills and Executive Function Dericits in Children with the 22q11 Deletion Syndrome." *Applied Neuropsychology* 13.4 (2006): 258–268. Web.
Kim, Albert H. et al. "MicroRNA Expression Profiling in the Prefrontal Cortex of Individuals Affected with Schizophrenia and Bipolar Disorders." *Schizophrenia Research* 124 (2010): 183–191. Web.
Kim, V Narry. "MicroRNA Biogenesis: Coordinated Cropping and Dicing." *Nature reviews. Molecular cell biology* 6.5 (2005): 376–85. Web.
Klaassen, Petra et al. "Explaining the Variable Penetrance of CNVs: Parental Intelligence Modulates Expression of Intellectual Impairment Caused by the 22q11.2 Deletion." *American Journal of Medical Genetics, Part B: Neuropsychiatric Genetics* 2016. Web.
Krichevsky, Anna M et al. "A microRNA Array Reveals Extensive Regulation of microRNAs during Brain Development." *RNA (New York, N.Y.)* 9.10 (2003): 1274–81. Web.

Kulhara, Parmanand, and Subho Chakrabarti. “Culture and Schizophrenia and Other Psychotic Disorders.” *Psychiatric Clinics of North America* 2001: 449–464. Web.

Kuperberg, Gina, and Stephan Heckers. “Schizophrenia and Cognitive Function.” *Current Opinion in Neurobiology* 2000: 205–210. Web.

Langer-Safer, P R, M Levine, and D C Ward. “Immunological Method for Mapping Genes on Drosophila Polytene Chromosomes.” *Proceedings of the National Academy of Sciences of the United States of America* 79.14 (1982): 4381–4385. Web.

Lawrie, Stephen M., and Suheib S. Abukmeil. “Brain Abnormality in Schizophrenia. A Systematic and Quantitative Review of Volumetric Magnetic Resonance Imaging Studies.” *British Journal of Psychiatry* 1998: 110–120. Web.

Leana-Cox, Julie et al. “Familial DiGeorge/velocardiofacial Syndrome with Deletions of Chromosome area22q11.2: Report of Five Families with a Review of the Literature.” *American Journal of Medical Genetics* 65.4 (1996): 309–316. Web.

Lee, Y et al. “MicroRNA Genes Are Transcribed by RNA Polymerase II.” *EMBO J* 23.20 (2004): 4051–4060. Web.

Lee, Yoontae et al. “The Nuclear RNase III Drosha Initiates microRNA Processing.” *Nature* 425.6956 (2003): 415–419. Web.

Lendahl, Urban, Lyle B. Zimmerman, and Ronald D.G. McKay. “CNS Stem Cells Express a New Class of Intermediate Filament Protein.” *Cell* 60.4 (1990): 585–595. Web.

Lewis, Benjamin P., Christopher B. Burge, and David P. Bartel. “Conserved Seed Pairing, Often Flanked by Adenosines, Indicates That Thousands of Human Genes Are microRNA Targets.” *Cell* 120 (2005): 15–20. Web.

Linsen, Sam E V, and Edwin Cuppen. “Methods for Small RNA Preparation for Digital Gene Expression Profiling by next-Generation Sequencing.” *Methods in Molecular Biology* 822 (2012): 205–217. Web.

Liu, Xiaosong et al. “Yamanaka Factors Critically Regulate the Developmental Signaling Network in Mouse Embryonic Stem Cells.” *Cell research* 18.12 (2008): 1177–1189. Web.

Martí, Mercè et al. “Characterization of Pluripotent Stem Cells.” *Nature protocols* 8.2 (2013): 223–53. Web.

McDonald-McGinn, D M et al. “Phenotype of the 22q11.2 Deletion in Individuals Identified through an Affected Relative: Cast a Wide FISHing Net!” *Genetics in medicine : official journal of the American College of Medical Genetics* 3.1 (2001): 23–29. Web.

McGuffin, P. et al. “The Strength of the Genetic Effect. Is There Room for an Environmental Influence in the Aetiology of Schizophrenia?” *British Journal of Psychiatry* 1994: 593–599. Web.

Meltzer, H Y, and S M Stahl. “The Dopamine Hypothesis of Schizophrenia: A Review.” *Schizophrenia bulletin* 2.1 (1976): 19–76. Web.

Meltzer, Herbert Y., and Tomiki Sumiyoshi. “Does Stimulation of 5-HT1A Receptors Improve Cognition in Schizophrenia?” *Behavioural Brain Research* 2008: 98–102. Web.

Merico, Daniele et al. “MicroRNA Dysregulation, Gene Networks and Risk for Schizophrenia in 22q11.2 Deletion Syndrome.” *Frontiers in Neurology* 5.NOV

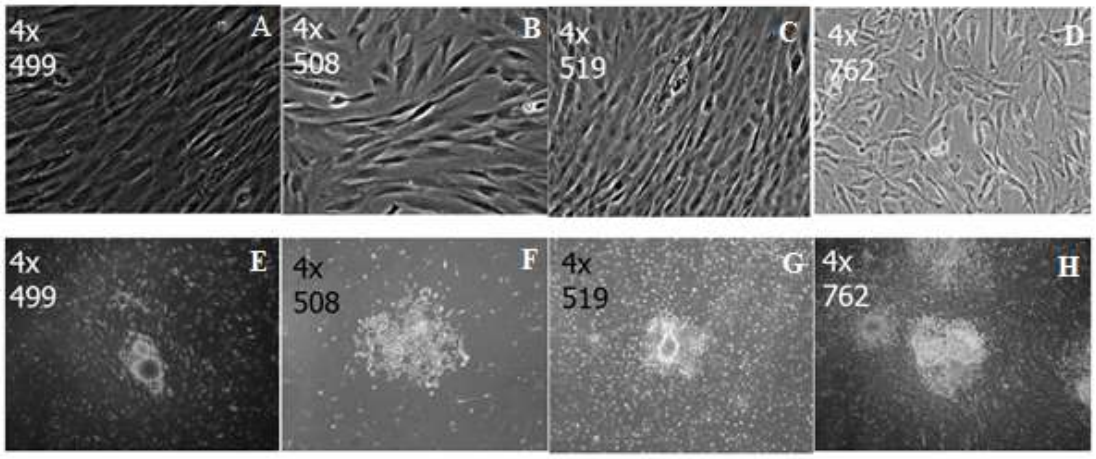
(2014): n. pag. Web.
Miller, B. H. et al. “MicroRNA-132 Dysregulation in Schizophrenia Has Implications for Both Neurodevelopment and Adult Brain Function.” *Proceedings of the National Academy of Sciences* 109.8 (2012): 3125–3130. Web.
Mishra, Prasun J. et al. “MiRSNPs or MiR-Polymorphisms, New Players in microRNA Mediated Regulation of the Cell: Introducing microRNA Pharmacogenomics.” *Cell Cycle* 2008: 853–858. Web.
Miska, Eric A et al. “Microarray Analysis of microRNA Expression in the Developing Mammalian Brain.” *Genome biology* 5.9 (2004): R68. Web.
Moffatt, Catherine E., and Richard J. Lamont. “Porphyromonas Gingivalis Induction of microRNA-203 Expression Controls Suppressor of Cytokine Signaling 3 in Gingival Epithelial Cells.” *Infection and Immunity* 79 (2011): 2632–2637. Web.
Mor, Eyal et al. “MicroRNA-382 Expression Is Elevated in the Olfactory Neuroepithelium of Schizophrenia Patients.” *Neurobiology of Disease* 55 (2013): 1–10. Web.
Moreau, Michael P. et al. “Altered MicroRNA Expression Profiles in Postmortem Brain Samples from Individuals with Schizophrenia and Bipolar Disorder.” *Biological Psychiatry* 69.2 (2011): 188–193. Web.
Morishita, Hirofumi et al. “Interneuron Epigenomes during the Critical Period of Cortical Plasticity: Implications for Schizophrenia.” *Neurobiology of Learning and Memory* 2015: 104–110. Web.
Morrow, B et al. “Molecular Definition of the 22q11 Deletions in Velo-Cardio-Facial Syndrome.” *American journal of human genetics* 56.6 (1995): 1391–403. Web.
Motameny, Susanne et al. “Next Generation Sequencing of miRNAs - Strategies, Resources and Methods.” *Genes* 1.1 (2010): 70–84. Web.
Northoff, G et al. “Reduced Activation and Altered Laterality in Two Neuroleptic-Naive Catatonic Patients during a Motor Task in Functional MRI.” *Psychological medicine* 29.4 (1999): 997–1002. Web.
Pantelis, Christos et al. “Early and Late Neurodevelopmental Disturbances in Schizophrenia and Their Functional Consequences.” *Australian and New Zealand Journal of Psychiatry* 2003: 399–406. Web.
Perkins, Diana O et al. “microRNA Expression in the Prefrontal Cortex of Individuals with Schizophrenia and Schizoaffective Disorder.” *Genome biology* 8.2 (2007): R27. Web.
Quintanilla, Rene H, Joanna Asprer, and Uma Lakshmipathy. “Novel Markers to Distinguish Pluripotent and.” *Life Technologies* Primary and Stem Cell Systems (2013): 92008. Web.
Robert, Christelle, and Mick Watson. “Errors in RNA-Seq Quantification Affect Genes of Relevance to Human Disease.” *Genome biology* 16.1 (2015): 177. Web.
Robinson, Peter N. et al. “Shprintzen-Goldberg Syndrome: Fourteen New Patients and a Clinical Analysis.” *American Journal of Medical Genetics* 135 A.3 (2005): 251–262. Web.
Saha, Sukanta et al. “A Systematic Review of the Prevalence of Schizophrenia.” *PLoS Medicine* 2005: 0413–0433. Web.
Saitta, Sulagna C. et al. “Aberrant Interchromosomal Exchanges Are the Predominant Cause of the 22q11.2 Deletion.” *Human Molecular Genetics* 13.4 (2004): 417–428.

Web.
Scambler, Peter J. et al. “Microdeletions within 22q11 Associated with Sporadic and Familial DiGeorge Syndrome.” *Genomics* 10.1 (1991): 201–206. Web.
Sellier, Chantal et al. “Decreased DGCR8 Expression and miRNA Dysregulation in Individuals with 22q11.2 Deletion Syndrome.” *PLoS ONE* 9.8 (2014): n. pag. Web.
Shaikh, T H et al. “Chromosome 22-Specific Low Copy Repeats and the 22q11.2 Deletion Syndrome: Genomic Organization and Deletion Endpoint Analysis.” *Hum Mol Genet* 9.4 (2000): 489–501. Web.
Shi, Wenting et al. “Aberrant Expression of Serum miRNAs in Schizophrenia.” *Journal of Psychiatric Research* 46.2 (2012): 198–204. Web.
Shprintzen, R J, R B Goldberg, M L Lewin, et al. “A New Syndrome Involving Cleft Palate, Cardiac Anomalies, Typical Facies, and Learning Disabilities: Velo-Cardio-Facial Syndrome.” *The Cleft palate journal* 15.1 (1978): 56–62. Web.
Shprintzen, R J, R B Goldberg, D Young, et al. “The Velo-Cardio-Facial Syndrome: A Clinical and Genetic Analysis.” *Pediatrics* 67.2 (1981): 167. Print.
Shprintzen, R J, A M Higgins, et al. “Velo-Cardio-Facial Syndrome.” *Current Opinion in Pediatrics* 17.6 (2005): 725–730. Print.
Shprintzen, Robert J et al. “Velo-Cardio-Facial Syndrome.” *Current opinion in pediatrics* 17.6 (2005): 725–730. Web.
Shprintzen, Robert J. “Velo-Cardio-Facial Syndrome: 30 Years of Study.” *Developmental Disabilities Research Reviews* 2008: 3–10. Web.
Shu, Jingmin et al. “Genomewide microRNA down-Regulation as a Negative Feedback Mechanism in the Early Phases of Liver Regeneration.” *Hepatology* 2011: 609–619. Web.
Stark, Kimberly L et al. “Altered Brain microRNA Biogenesis Contributes to Phenotypic Deficits in a 22q11-Deletion Mouse Model.” *Nature genetics* 40.6 (2008): 751–60. Web.
Strojnik, Tadej et al. “Neural Stem Cell Markers, Nestin and Musashi Proteins, in the Progression of Human Glioma: Correlation of Nestin with Prognosis of Patient Survival.” *Surgical Neurology* 68.2 (2007): 133–143. Web.
Sugama, S et al. “Psychiatric Inpatients and Chromosome Deletions within 22q11.2.” *Journal of neurology, neurosurgery, and psychiatry* 67.6 (1999): 803–6. Web.
Taapken, Seth M et al. “Karyotypic Abnormalities in Human Induced Pluripotent Stem Cells and Embryonic Stem Cells.” *Nature biotechnology* 29.4 (2011): 313–4. Web.
Tadeu, Ana Mafalda Baptista et al. “Transcriptional Profiling of Ectoderm Specification to Keratinocyte Fate in Human Embryonic Stem Cells.” *PLoS ONE* 10.4 (2015): n. pag. Web.
Tandon, R., and U. Halbreich. “The Second-Generation ‘Atypical’ Antipsychotics: Similar Improved Efficacy but Different Neuroendocrine Side Effects.” *Psychoneuroendocrinology* 2003: 1–7. Web.
Tomita-Mitchell, Aoy et al. “Multiplexed Quantitative Real-Time PCR to Detect 22q11.2 Deletion in Patients with Congenital Heart Disease.” *Physiological genomics* 42A.1 (2010): 52–60. Web.
van Amelsvoort, T et al. “Brain Anatomy in Adults with Velocardiofacial Syndrome with and without Schizophrenia: Preliminary Results of a Structural Magnetic Resonance

Imaging Study." *Arch Gen Psychiatry* 61.11 (2004): 1085–1096. Web.
Vogel, F. "Schizophrenia Genesis: The Origins of Madness." *American journal of human genetics* 48.6 (1991): 1218. Web.
Von Bohlen Und Halbach, O. "Immunohistological Markers for Staging Neurogenesis in Adult Hippocampus." *Cell and Tissue Research* 2007: 409–420. Web.
Vorstman, Jacob A S et al. "Association of the PIK4CA Schizophrenia-Susceptibility Gene in Adults with the 22q11.2 Deletion Syndrome." *American Journal of Medical Genetics, Part B: Neuropsychiatric Genetics* 150.3 (2009): 430–433. Web.
Walker, Elaine et al. "Schizophrenia: Etiology and Course." *Annual review of psychology* 55 (2004): 401–430. Web.
Warnica, William et al. "Copy Number Variable Micrornas in Schizophrenia and Their Neurodevelopmental Gene Targets." *Biological Psychiatry* 77.2 (2015): 158–166. Web.
Weksberg, Rosanna et al. "Molecular Characterization of Deletion Breakpoints in Adults with 22q11 Deletion Syndrome." *Human Genetics* 120.6 (2007): 837–845. Web.
Whalley, Heather C et al. "Impact of a microRNA MIR137 Susceptibility Variant on Brain Function in People at High Genetic Risk of Schizophrenia or Bipolar Disorder." *Neuropsychopharmacology* 2012: 2720–2729. Web.
Wilson, D I et al. "DiGeorge Syndrome: Part of CATCH 22." *Journal of medical genetics* 30.10 (1993): 852–6. Web.
Wilson, P G, and S S Stice. "Development and Differentiation of Neural Rosettes Derived from Human Embryonic Stem Cells." *Stem Cell Rev* 2.1 (2006): 67–77. Web.
Wong, Jenny et al. "Expression of NPAS3 in the Human Cortex and Evidence of Its Posttranscriptional Regulation by miR-17 during Development, with Implications for Schizophrenia." *Schizophrenia Bulletin* 39.2 (2013): 396–406. Web.
Xie, Xiaohui et al. "Systematic Discovery of Regulatory Motifs in Human Promoters and 3' UTRs by Comparison of Several Mammals." *Nature* 434 (2005): 338–345. Web.
Xu, Yong et al. "MicroRNAs and Target Site Screening Reveals a Pre-microRNA-30e Variant Associated with Schizophrenia." *Schizophrenia Research* 119.1-3 (2010): 219–227. Web.
Yu, Junying et al. "Human Induced Pluripotent Stem Cells Free of Vector and Transgene Sequences." *Science (New York, N.Y.)* 324.5928 (2009): 797–801. Web.
Yuan, Shauna H. et al. "Cell-Surface Marker Signatures for the Isolation of Neural Stem Cells, Glia and Neurons Derived from Human Pluripotent Stem Cells." *PLoS ONE* 6.3 (2011): n. pag. Web.
Zeng, Yan, Rui Yi, and Bryan R Cullen. "Recognition and Cleavage of Primary microRNA Precursors by the Nuclear Processing Enzyme Drosha." *The EMBO journal* 24.1 (2005): 138–48. Web.
Zhao, Dejian et al. "MicroRNA Profiling of Neurons Generated Using Induced Pluripotent Stem Cells Derived from Patients with Schizophrenia and Schizoaffective Disorder, and 22q11.2 Del." *PLoS ONE* 10.7 (2015): n. pag. Web.

Supplementary Figures:

Supplemental Figure 1: The following figure shows additional examples of induction process of fibroblasts to iPSCS. Newly transfected fibroblasts are shown in Panels A-D during the first 2-5 days following transfection. Newly formed iPSC colonies are shown in Panel E-F during the second week following transfection.



Supllemental Figure 2:

Attached below is an example of the RUCDR cell line report. Basic identifying information about the cell line as well as photos to confirm the iPSC has normal cell morphology are provided. Karyotypes and CNVs were not performed by RUCDR. Alkaline phosphatase staining was performed by RUCDR on all distributed iPSCs. Details of cryopreservation for the individual samples that were distributed by RUCDR are also shown. Photos of the iPSCs after thawing show that the iPSCs had otherwise normal morphology. The Oct/4Tra-1-60 FACs analysis also showed that the iPSCs express these two markers of pluripotency.

Identity and growth

Date: 08/14/2015

Cell line ID - MH0161602

Lot numbers - R139030944 & R149451830

Cell line submitter - NIMH

RUID SNP panel - available upon request

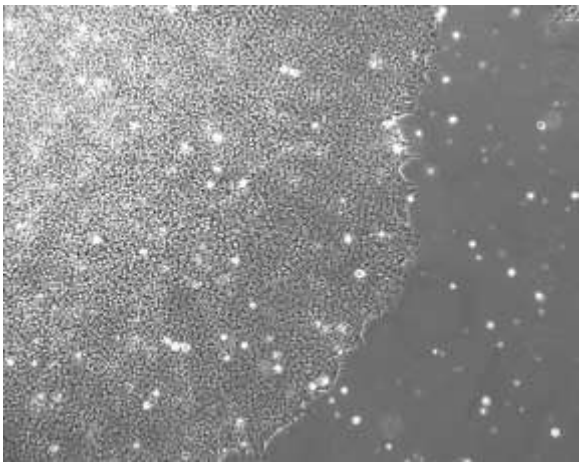
Source cell type - EPIPSC Reprogramming method - N/A

Growth conditions - mTeSR and Matrigel

Passage method - Dispase

Pre-freeze mycoplasma testing – pass

Morphology (10X) – pass

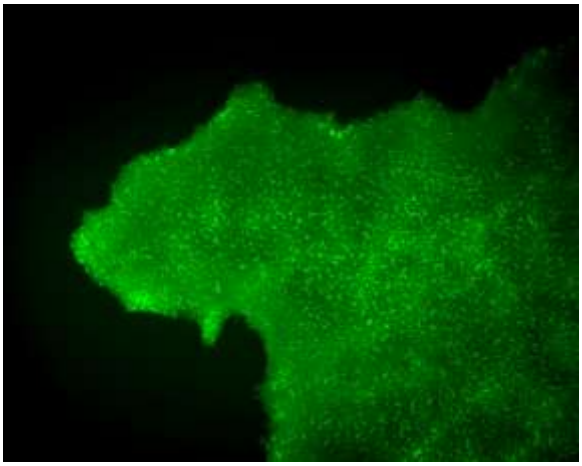


Karyotype - N/A

CNV Analysis - N/A

Pluripotency

Alkaline Phosphatase Staining – pass



Immunohistochemistry - N/A Oct4/Tra-1-60/Dapi

Cryopreservation

Cryopreservation date - 02/18/2014

Cryopreservation method - mFreSR

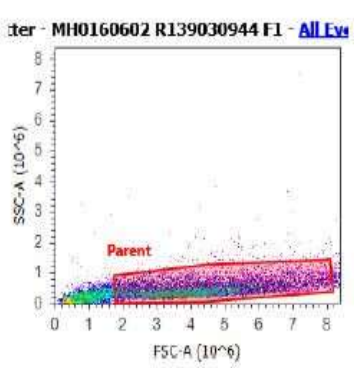
Passage number -10

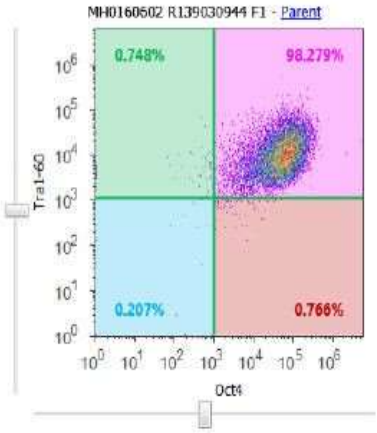
Post thaw mycoplasma testing – negative

Post thaw FACS analysis- 95.0% 98.3%% Oct4/Tra-1-60 double positive

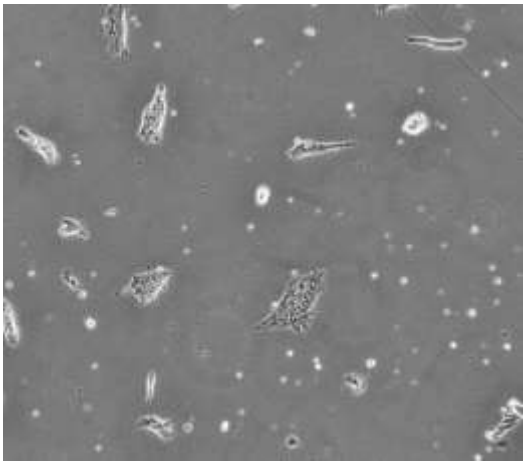
Post thaw morphology - pass

	R139030944	R149451830
Recomended thaw density per 9.6cm2 (1 well of a 6 well plate)	1000000	1000000
At least 10 colonies after 3 days (+/-)	+	+
At least 50% pluripotent based on morphology (+/-)	+	+





10x magnification 3 days after (R139030944):



Supplementary figure 3 shows the results of RUCDR's SNPtrace™ Panel Genotyping Assays that have been completed for cell lines used in these experiments.

RUID	02C1295 0	03C1593 3	03C2182 2	04C2734 8	04C2753 6	04C3729 1	05C3941 3	05C4488 2	04C2468 3	06C5225 4	06C5685 1	06C5738 9	06C5896 4
SNPTrace status (call rate %)	st reques	st reques	st reques	st reques	st reques	st reques	st reques	st request	100%	97.92%	100.00%	100.00%	100.00%
subject ID	76-28288	76-54915	76-74517	65-1122-0	01-0114-0	01-0217-0	02-1241-0	08-2184-0	01-0088-0	05-1346-0	01-0060-0	07-2438-0	04-1480-0
project	017 SZ	017 SZ	017 SZ	029 SZ	029 SZ	029 SZ	029 SZ	029 SZ	029 SZ	029 SZ	029 SZ	029 SZ	029 SZ
Matches?	pending	pending	pending	pending	pending	pending	pending	pending	none	none	none	none	none
RUID (match)													
subject ID													
project													
RUID (match)													
subject ID													
project													

RUID	MH0159 019 (IPSC)	MH0159 020 (IPSC)	MH0159 021 (IPSC)	MH0159 022 (IPSC)	MH0159 024 (IPSC)	MH0159 025 (IPSC)	MH0159 026 (IPSC)	MH0159 027 (IPSC)	MH0162 499 (FCL)	MH0162 519 (FCL)	MH0162 627 (FCL)	MH0162 762 (FCL)
SNPTrace status (call rate %)	97.92%	98.96%	98.96%	97.92%	100.00%	100.00%	100.00%	98.96%	98.96%	100.00%	100.00%	100.00%
subject ID	03-IPS1B	03-IPS2C	03-IPS5C	03-IPS6C	22Q11-1	03-S259C	03-S249C	3-IPS15B	499-537	499-1220	499-3189	734-4276
project	125 22Q	125 22Q	125 22Q	125 22Q	125 22Q	125 22Q	125 22Q	125 22Q	132 PSY	132 PSY	132 PSY	115 PMF
Matches?	none	none	none	none	none	none	none	none	none	none	none	none
RUID (match)												
subject ID												
project												
RUID (match)												
subject ID												
project												

RUID	90C02141	MH015902 (IPSC)	MH0162508 (FCL)	MH0162764 (IPSC)
SNPTrace status (call rate %)	98.96%	100.00%	100.00%	100.00%
subject ID	32-10465	393-22Q11-30C1	499-1275	734-2172.4
project	NIMH Coriell	125 22Q	132 PSY	115 PMF
Matches?	10C111829	MH0161611BZ0000045	MH0167168	MH0162756
RUID (match)	10C111829	MH0161611 (FCL)	MH0167168 (IPSC)	MH0162756 (FCL)
subject ID	262-12514	393-SF030	392-1275C9	732-2172
project	076 SZ	125 22Q	125 22Q	115 PMF
RUID (match)	10C111829	820000045	MH0167168 (IPSC)	MH0162756 (FCL)
subject ID	262-12514	NSC_MH23_P08	392-1275C9	732-2172
project	076 SZ	ANEUPLOIDY (LB lab	125 22Q	115 PMF

Supplemental Figure 4:

Panel A:

This figure shows the karyotype and FISH results from a control line. The karyotype is listed at “NORMAL” This karyotype was performed on 20 G-banded metaphase cells. All 20 of the cells demonstrated a “NORMAL” karyotype. The analysis is consistent with previous information known about this individual cell

Panel B:

This figure shows the karyotype and FISH results from a 22q11,2 DS line. The karyotype is listed at “NORMAL” This karyotype was performed on 20 G-banded metaphase cells. All 20 of the cells demonstrated a “NORMAL” karyotype. The analysis is consistent with previous information known about this individual cell line.

Panel A:

Cell Line GENETICS

Cell Line Characterization

Cell Line ID: 509
Passage #: 8
Specimen Type: Human iPSC Culture
Indication for Study: Routine Culture QC

Test Code: 100,201
Account #: NA
PO #: 2188340

Date Received: 3/17/2016
Date Reported: 4/8/2016
Time in Culture: 1 Day

Banding Technique: GTL **Band Resolution:** Poor
Metaphases Counted: 20 Analyzed: 7 Karyotyped: 3

Lab #: CLG-23102
PI: Linda Brzustowkz
Contact Person: William Manley
Email: billman119@gmail.com
Address:
Rutgers University
Human Genetics Institute
145 Bevier Road
Piscataway, New Jersey 08854
Additional copies sent to:

RESULTS: 46,XY[20] Apparently NORMAL Human Male Karyotype

FISH Results: See Interpretation Below

Non-clonal Aberrations: None

INTERPRETATION:

Cytogenetic analysis was performed on twenty G-banded metaphase cells from human cell line 509 p8 and all twenty cells demonstrated an apparently normal male karyotype.

Two hundred interphase nuclei were examined by fluorescence in situ hybridization (FISH) using the probe set listed below in order to detect deletion of chromosome 22q11.21.

Probe set used:
N85A3 (22q13.33) / TUPLE1 (22q11.21)

The signal pattern is consistent with two copies of chromosome 11q11.2.

(Continued on page 2)

Cell Line GENETICS

CLG-23102
509 p8

Normal Cut-off Values

Probe Set: **22q11.21(Red) / 22q13.33(Green)**

Signal Pattern	Observed % of cells w/signal pattern (Your Results)	Normal Cut-off value at 99% Confidence
2R/2G	81.0%	NA
2R/1G	11.0%	13.6%
3R/2G	6.0%	6.9%
3R/3G	2.0%	4.9%

Please note that signal patterns that fall below the normal cut-off value most likely represent false-positive signal patterns, but the possibility that these may represent a small emerging cell population cannot be ruled out.

Panel B:

Cell Line Characterization

Cell Line ID: 046
Passage #: 8
Specimen Type: Human iPSC Culture
Indication for Study: Routine Culture QC

Test Code: 100,201
Account #: NA
PO #: 2251529

Date Received: 4/5/2016
Date Reported: 4/22/2016
Time in Culture: 1 Day

Banding Technique: GTL **Band Resolution:** Fair
Metaphases Counted: 20 Analyzed: 7 Karyotyped: 2

Lab #: CLG-23291
PI: Linda Brzustowkz
Contact Person: William Manley
Email: billman119@gmail.com
Address:
Rutgers University
Human Genetics Institute
145 Bevier Road
Piscataway, New Jersey 08854
Additional copies sent to:

RESULTS: 46,XX[18] Apparently NORMAL Human Female Karyotype

Non-clonal Aberrations: 45,XX,-2(one cell)
45,XX,-9(one cell)

INTERPRETATION:

Cytogenetic analysis was performed on twenty G-banded metaphase cells from human cell line 046 p8. Eighteen cells demonstrated an apparently normal female karyotype, while two cells demonstrated non-clonal chromosome aberrations (listed above) which are most likely artifacts of culture.

Two hundred interphase nuclei were examined by fluorescence in situ hybridization (FISH) using the probe set listed below.
Probe set used:

N85A3 (22q13.3) / TUPLE1 (22q11.21)

The presence of a deletion on chromosome 22q11.21 was detected in 86.0% of the cells in the N85A3 (22q13.3)/TUPLE1(22q11.21) probe set.

RUO - Research Use Only

Cell Line GENETICS

CLG-23291
046 p8

Normal Cut-off Values

Probe Set: **22q11.21(Red) / 22q13.3(Green)**

Signal Pattern	Observed % of cells w/signal pattern (Your Results)	Normal Cut-off value at 99% Confidence
2R/2G	13.5%	NA
1R/2G	60.0%	8.8%
1R/1G	23.0%	4.9%
1R/3G	3.0%	2.3%
2R/1G	0.5%	13.6%

Please note that signal patterns that fall below the normal cut-off value most likely represent false-positive signal patterns, but the possibility that these may represent a small emerging cell population cannot be ruled out.

Supplemental Figure 5: The following panels show the karyotypic history of MH 577. Panel A shows the abnormal results obtained from the karyotype and FISH performed on passage 3 of MH 577. Panel B shows the normal results from the karyotype and the abnormal results of the FISH performed on passage 15 of MH 577. Panel C shows the normal FISH results that were obtained from passage 21 of MH 577.

Panel A:

Cell Line GENETICS

Cell Line Characterization

Cell Line ID: MH577
Passage #: 3
Specimen Type: Human iPSC Culture
Indication for Study: Routine Culture QC

Test Code: 100,201
Account #: NA
PO #: 2188340

Date Received: 10/21/2015
Date Reported: 10/30/2015
Time in Culture: 1 Day

Banding Technique: GTL **Band Resolution:** Poor
Metaphases Counted: 20 Analyzed: 9 Karyotyped: 5

Lab #: CLG-21672
PI: Linda Brzustowicz
Contact Person: William Manley
Email: billman119@gmail.com
Address:
Rutgers University
Human Genetics Institute
145 Bevier Road
Piscataway, New Jersey 08854
Additional copies sent to:

RESULTS: 46,XX[13] See Interpretation Below

FISH Results: See Interpretation Below

Non-clonal Aberrations: 45,X,-X[1] 46,XX,-13,+mar[1] 47,XX,add(10)(q22),+mar[1]
47,XX,t(15;17)(q22;q21),+mar[1] 46,XX,del(9)(q13)[1]
45,XX,-13[1] 45,XX,-17[1]

INTERPRETATION:

Cytogenetic analysis was performed on twenty G-banded metaphase cells from human cell line MH577 p3. Thirteen cells demonstrated an apparently normal female karyotype, while seven cells demonstrated non-clonal chromosome aberrations (listed above). Non-clonal changes such as random chromosome gain/loss or structural aberrations are generally considered artifacts of culture, but having seven cells with multiple numerical and structural aberrations suggest that this culture is unstable.

Two hundred interphase nuclei were examined by fluorescence in situ hybridization (FISH) using the probe set listed below in order to detect deletion of chromosome 22q11.21 and 22q13.3.

Probe set used:
N85A3 (22q13.3) / TUPLE1 (22q11.21)

The presence of a deletion of 22q11.2 was not detected.

RUO - Research Use Only

Cell Line GENETICS

CLG-21672
MH577 p3

Normal Cut-off Values

Probe Set: **22q11.21(Red) / 22q13.3(Green)**

Signal Pattern	Observed % of cells w/signal pattern (Your Results)	Normal Cut-off value at 99% Confidence
2R/2G	91.5%	NA
3R/3G	4.0%	4.9%
2R/1G	1.5%	13.6%
3R/2G	2.0%	6.9%
2R/3G	1.0%	6.3%

Please note that signal patterns that fall below the normal cut-off value most likely represent false-positive signal patterns, but the possibility that these may represent a small emerging cell population cannot be ruled out.

Panel B:

Cell Line GENETICS

Cell Line Characterization

Cell Line ID: MH577
Passage #: 15
Specimen Type: Human iPSC Culture
Indication for Study: Routine Culture QC

Test Code: 100,201
Account #: NA
PO #: 2251529

Date Received: 1/15/2016
Date Reported: 1/19/2016
Time in Culture: 1 Day

Banding Technique: GTL **Band Resolution:** Good
Metaphases Counted: 20 Analyzed: 7 Karyotyped: 2

Lab #: CLG-22386
PI: Linda Brzustowkz
Contact Person: William Manley
Email: billman119@gmail.com
Address:
Rutgers University
145 Bevier Road
Piscataway, New Jersey 08854

Additional copies sent to:

RESULTS: 46,XX[19] Apparently NORMAL Human Female Karyotype

FISH Results: ABNORMAL FISH Signal Pattern for chromosome 22

Non-clonal Aberrations: 45,XX,del(14)(q24),-22(one cell)

INTERPRETATION:

Cytogenetic analysis was performed on twenty G-banded metaphase cells from human cell line MH577 p15. Nineteen cells demonstrated an apparently normal female karyotype, while one cell demonstrated non-clonal chromosome aberrations (listed above) which are most likely artifacts of culture.

Two hundred interphase nuclei were examined by fluorescence in situ hybridization (FISH) using the probe set listed below.

Probe set used:
N85A3 (22q13.3) / TUPLE1 (22q11.21)

Trisomy 22q13.3 and 22q11.21 together was detected in 5.5% of the cells and trisomy 22q13.3 was detected in 8.5% of the cells in the N85A3 (22q13.3) / TUPLE1 (22q11.21) probe set.

Cell Line GENETICS

CLG-22386
MH577 p15

Normal Cut-off Values

Probe Set: **22q11.21(Red) / 22q13.3(Green)**

Signal Pattern	Observed % of cells w/signal pattern (Your Results)	Normal Cut-off value at 99% Confidence
2R/2G	74.5%	NA
2R/1G	7.0%	13.6%
3R/3G	5.5%	4.9%
2R/3G	8.5%	6.3%
3R/2G	2.5%	6.9%
1R/1G	1.0%	4.9%
1R/2G	1.0%	8.8%

Please note that signal patterns that fall below the normal cut-off value most likely represent false-positive signal patterns, but the possibility that these may represent a small emerging cell population cannot be ruled out.

Panel C:

Cell Line GENETICS

Cell Line Characterization

Cell Line ID: MH577
Passage #: 21
Specimen Type: Human iPSC Culture
Indication for Study: Routine Culture QC

Test Code: 100,201
Account #: NA
PO #: 2251529

Date Received: 3/24/2016
Date Reported: 4/8/2016
Time in Culture: 1 Day

Banding Technique: GTL **Band Resolution:** NA
Metaphases Counted: NA Analyzed: NA Karyotyped: NA

Lab #: CLG-23184
PI: Linda Brzustowkz
Contact Person: William Manley
Email: billman119@gmail.com
Address:
Rutgers University
Human Genetics Institute
145 Bevier Road
Piscataway, New Jersey 08854
Additional copies sent to:

RESULTS: Karyotype Failure: Low Mitotic Index

FISH Results: See Interpretation Below

Non-clonal Aberrations: NA

INTERPRETATION:

Despite our best efforts, cytogenetic analysis could not be performed on human cell line MH577 p21 because of a poor mitotic index. We recommend that you resubmit this cell line for characterization in order to ensure that you are working with a cytogenetically normal cell line.

Two hundred interphase nuclei were examined by fluorescence in situ hybridization (FISH) using the probe set listed below in order to detect deletion of chromosome 22q11.21.

Probe set used:
N85A3 (22q13.33) / TUPLE1 (22q11.21)

The signal patterns are consistent with two copies of chromosome 22q11.2.

(Continued on Page 2)

Cell Line GENETICS

CLG-23184
MH577 p21

Normal Cut-off Values

Probe Set: **22q11.21(Red) / 22q13.33(Green)**

Signal Pattern	Observed % of cells w/signal pattern (Your Results)	Normal Cut-off value at 99% Confidence
2R/2G	86.0%	NA
1R/1G	1.0%	4.9%
2R/1G	8.5%	13.6%
3R/3G	1.5%	4.9%
1R/2G	2.0%	8.8%
3R/2G	1.0%	6.9%

Please note that signal patterns that fall below the normal cut-off value most likely represent false-positive signal patterns, but the possibility that these may represent a small emerging cell population cannot be ruled out.

Supplemental Figure 6: The following figure shows the Run Summary of the Low Coverage Sequencing was performed using Ion Torrent karyotyping and analyzed using the Ion Reporter software (IGV). Panel A shows all of the samples that were run, the reads per sample, and the mean read length per sample. Quality control information about the run is shown in panels B-C. Panel D provides information about the test fragment for the run. Panel E shows the consensus key and the library peak. The alignment summary is shown in Panel F. Additional descriptive statistics are shown in panel G. Panels H and I provide further information about the number of reads and the read lengths. Additional analysis details are shown in panel J.

A:

Barcode Name	Sample	Bases	≥*Q*20	Reads	Mean Read Length
No barcode	none	240,348,340	173,123,417	1,964,934	122 bp
IonXpress 001	BZ0000029	261,064,026	188,444,917	2,141,856	121 bp
IonXpress 002	BZ0000028	383,064,197	279,447,207	3,085,268	124 bp
IonXpress 003	BZ0000027	315,246,590	227,149,030	2,572,881	122 bp
IonXpress 004	BZ0000026	461,819,114	337,652,449	3,693,907	125 bp
IonXpress 005	BZ0000025	281,647,546	203,767,613	2,288,705	123 bp
IonXpress 006	BZ0000024	338,939,627	243,459,577	2,797,008	121 bp
IonXpress 007	BZ0000023	314,619,793	222,081,302	2,656,446	118 bp
IonXpress 008	BZ0000022	289,026,343	211,650,849	2,333,591	123 bp
IonXpress 009	BZ0000021	650,779,048	479,938,453	5,180,567	125 bp
IonXpress 010	BZ0000030	528,395,157	387,493,377	4,270,010	123 bp
IonXpress 011	BZ0000049	221,075,182	157,281,904	1,852,748	119 bp
IonXpress 012	BZ0000048	500,429,155	368,829,627	4,013,617	124 bp
IonXpress 013	BZ0000047	449,178,138	323,908,289	3,643,624	123 bp
IonXpress 014	BZ0000045	409,386,835	292,454,480	3,372,003	121 bp
IonXpress 015	BZ0000044	409,779,815	299,062,809	3,325,363	123 bp
IonXpress 016	BZ0000046	542,310,954	393,868,866	4,366,056	124 bp
IonXpress 017	BZ0000043	521,203,148	378,804,796	4,223,548	123 bp
IonXpress 018	BZ0000042	414,677,215	303,312,102	3,392,116	122 bp
IonXpress 019	BZ0000041	293,296,197	217,499,919	2,353,692	124 bp

IonXpress 020	BZ0000050	482,965,194	362,447,216	3,879,847	124 bp

B:

Addressable Wells	148,603,956
With ISPs: 91,824,653	61.80%
Live: 91,081,427	99.20%
Test Fragment: 1,579,478	1.70%
Library: 89,501,949	98.30%

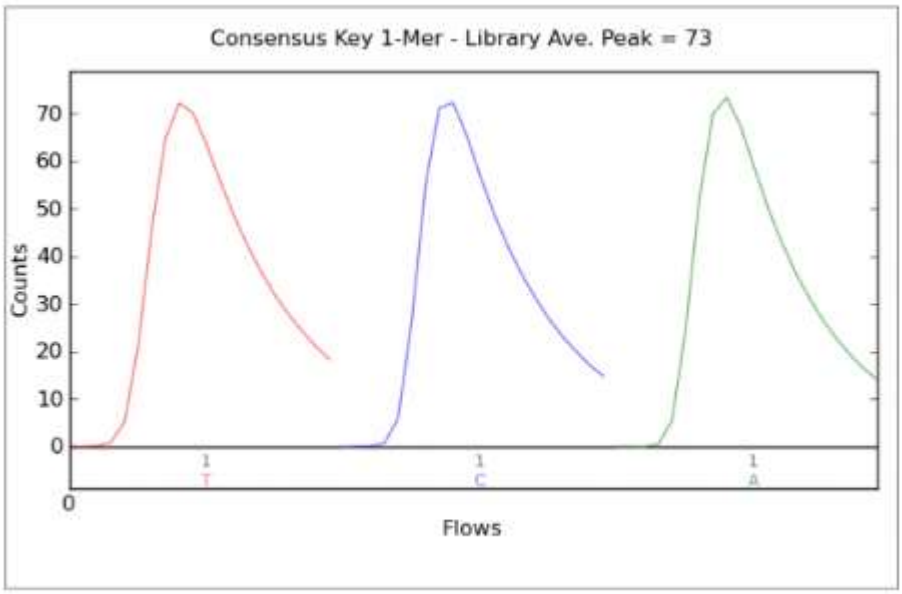
C:

Library ISPs	89,501,949
Filtered: Polyclonal 18,361,719	20.50%
Filtered: Low Quality 4,037,567	4.50%
Filtered: Primer Dimer 10,439	0.00%
Final Library ISPs 67,412,950	75.30%

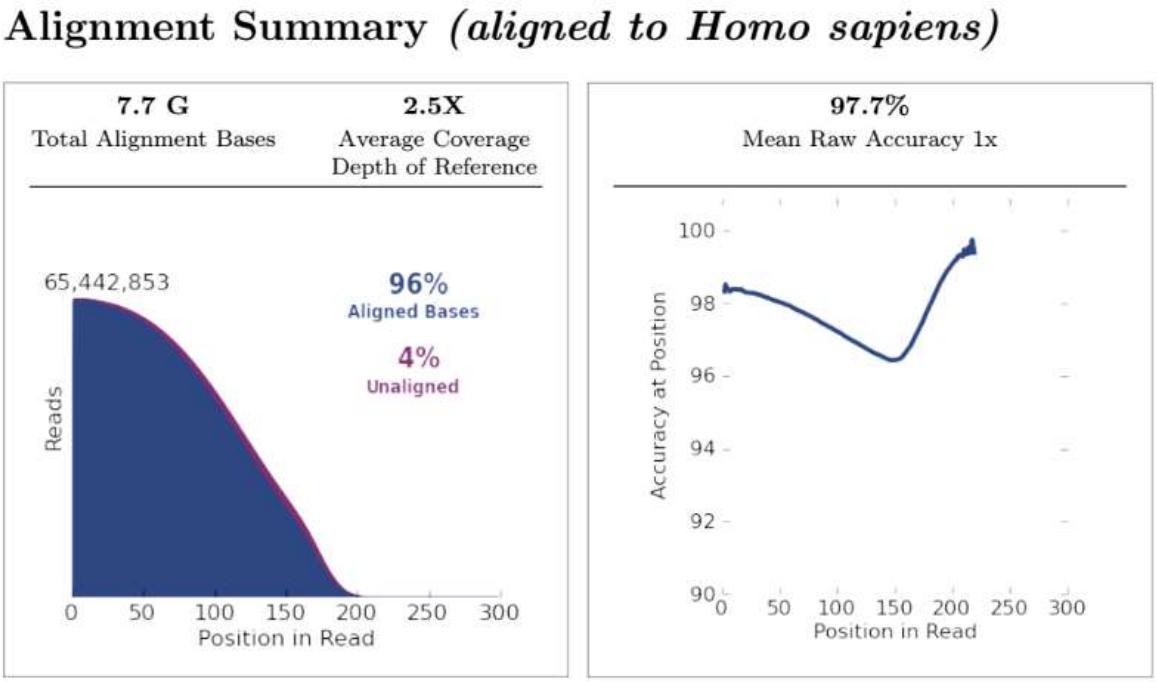
D:

Test Fragment	Reads	Percent 50AQ17	Read Length Histogram
TF C	**1,139,707**	**8%**	

E:



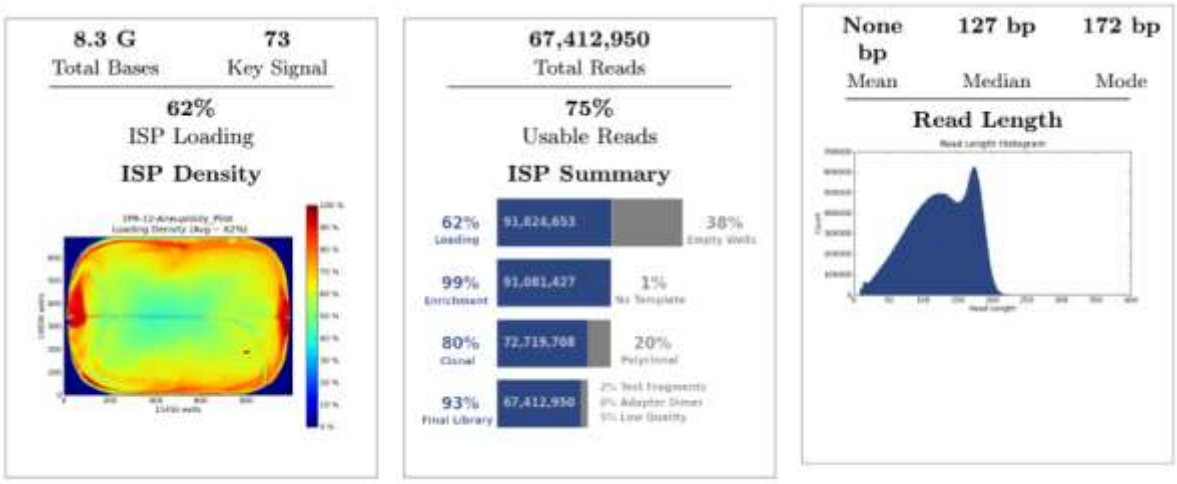
F:



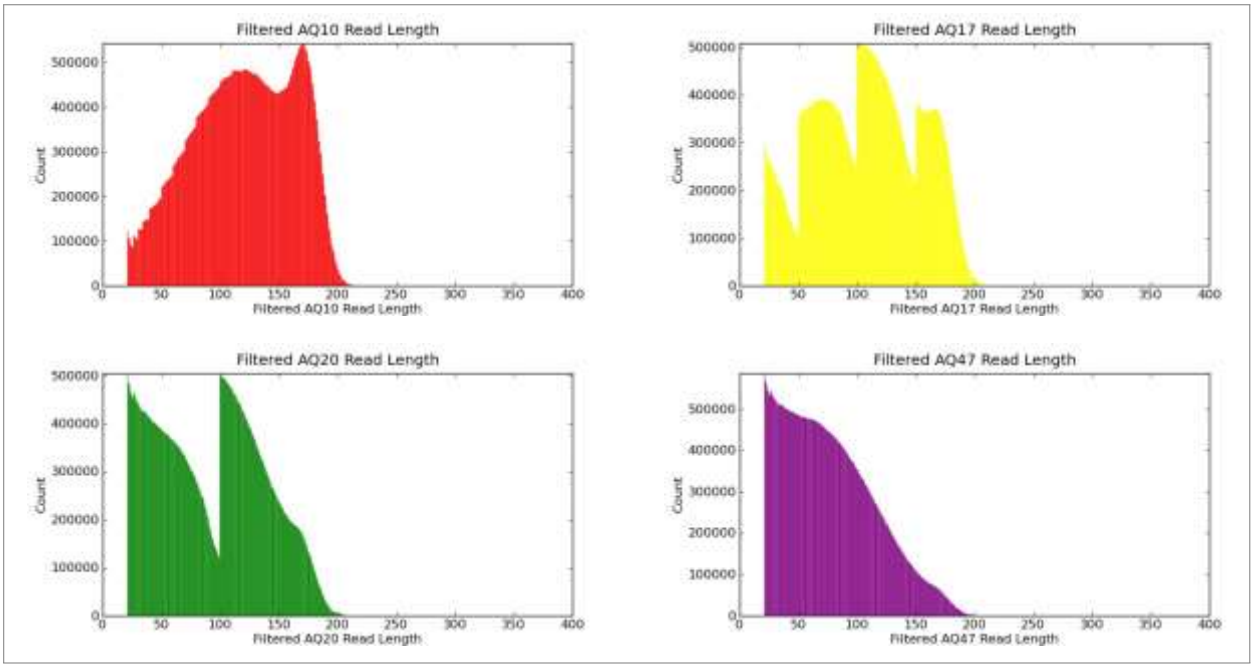
G:

	AQ17	AQ20	Perfect
Total Number of Bases [Mbp]	6 G	4.8 G	3.8 G
Mean Length [bp]	107	92	76
Longest Alignment [bp]	231	231	219
Mean Coverage Depth	2	1.6	1.3

H:


8.3 G
Total Bases
73
Key Signal
62%
ISP Loading
ISP Density
67,412,950
Total Reads
75%
Usable Reads
ISP Summary
62%
Loading
91,824,653
38%
99%
Enrichment
91,081,427
1%
80%
Clonal
72,719,768
20%
93%
Final Library
67,412,950
None bp
Mean
127 bp
Median
172 bp
Mode
Read Length

I:

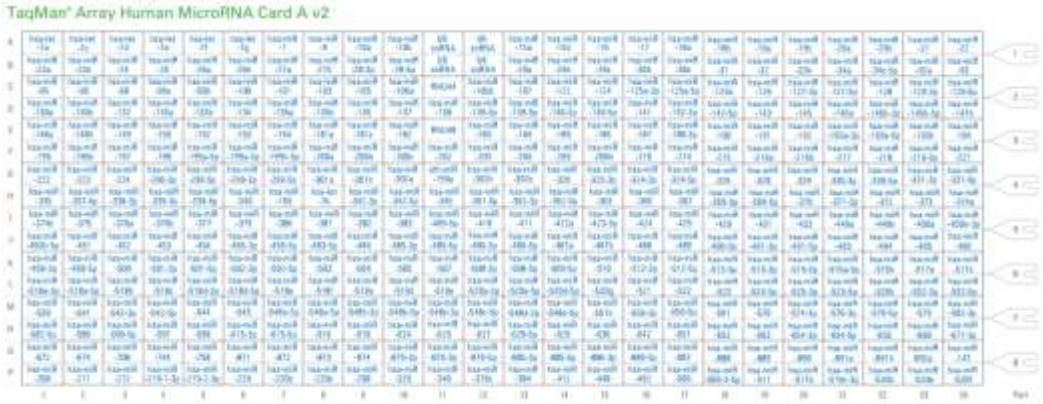

Filtered AQ10 Read Length
Filtered AQ17 Read Length
Filtered AQ20 Read Length
Filtered AQ47 Read Length
Count
Filtered AQ10 Read Length
Filtered AQ17 Read Length
Filtered AQ20 Read Length
Filtered AQ47 Read Length

J:

Analysis Details			
Run Name	R 2015 04 15 10 33 35 user 1PR-12-Aneuploidy Pilot	Torrent Suite	4.2.1
Run Date	April 15, 2015, 10:37 a.m.	host	56JMQW1
Run Flows	520	ion-analysis	2.18-1
Projects		ion-chefupdates	4.2.0
PGM	1proton	ion-dbreports	2.22-1
Flow Order	TACGTACGTCTGAGCATCGATCGATGTACAGC	ion-gpu	2.2-1
Library Key	TCAG	ion-pipeline	2.12-1
TF Key	ATCG	ion-plugins	2.28-1
Chip Check	Passed	ion-protonupdates	4.2.2
Chip Type	P1.1.17	ion-torrentr	2.1-1
Chip Data	tiled	LiveView	1835
Barcode Set	IonXpress	DataCollect	2929
Analysis Name	Auto user 1PR-12-Aneuploidy Pilot 57	OIA	4201
Analysis Date	April 15, 2015, 8:07 p.m.	OS	21
Analysis Flows	0	Graphics	34
runID	7GWHA		

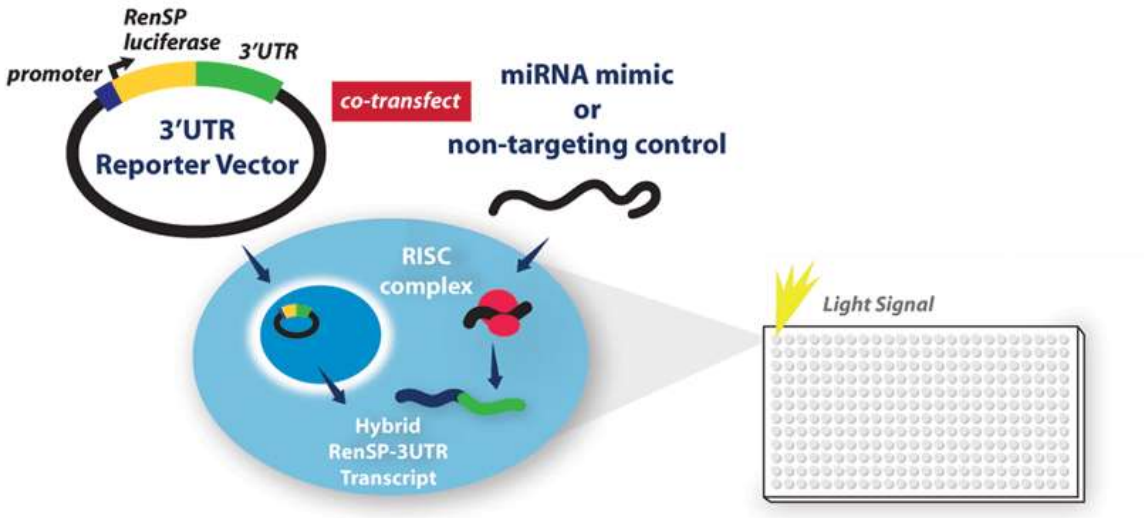
Supplementary Figure 7: The following figures show the probes used on the TaqMan® human miRNA cards. Panel A shows Card A vs 2.0. Panel B shows Card B version 3.0.

A:



B:

Supplementary Figure 8: A schematic of the Switchgear Genomics Luciferase Assay transfection is shown below (Switchgear Genomics). The 3' UTR reporter vector contains a 3'UTR specific for H3F3B, the luciferase gene, and a promoter. This plasmid is co-transfected into the cells with either miR-616 mimic or a non-targeting control.



Supplementary Methods:

Luciferase Reporter Assays

To investigate whether the SNP rs1060120 has an effect on the ability of miR-616 to regulate H3F3B expression, luciferase reporter constructs containing human 3′-untranslated region (3′UTRs) for H3F3B and the 3'UTR for H3F3B containing the SNP rs1060120, were custom built by Switchgear Genomics.

100 ng of each 3'UTR reporter construct was co-transfected with 0.15 µl of DharmaFECT DUO transfection reagent (Dharmacon) and sufficient miRNA mimic (Switchgear Genomics) to yield a final concentration of 20 nM in a total volume of 100 µl/well. Wells were co-transfected with either the miR-616-3p mimic, the non-targeting scramble miRNA, or an equivalent volume of media as a control. Each transfection condition was performed in triplicate. Plates were incubated at 37 °C for 24 hours post transfection. After transfection the plates were briefly frozen, then allowed to return to room temperature.

RenSP is a luciferase developed by SwitchGear Genomics specifically to function as a reporter gene assays. RenSP has a half-life of ~1 hour enabling a detailed analysis of kinetic responses with a highly robust signal and be detected using LightSwitch Luciferase Assay Reagents. LightSwitch Assay Reagents were used as directed (Switchgear Genomics). The plates were incubated at room temperature for 30 minutes while protected from light and gently agitated. After 30 minutes, luciferase expressions levels were measured on a Veritas Microplate luminometer (Turner Biosystems).

Statistical Methods:

95% confidence intervals were constructed for the difference between the mean of the 3’UTR and the 3’UTR co-transfected with miR-616 for both the wild type allele and the schizophrenia risk allele (rs1060120). A 95% confidence interval was also calculated between the difference between the mean of the 3’UTR from the schizophrenia risk allele and the 3’UTR from the schizophrenia risk allele co-transfected with miR-616 and the difference between the mean of the 3’UTR from the wild type allele and the 3’UTR from the wild type allele co-transfected with miR-616.

Statistical inference was based on tests of contrasts under a linear mixed effects model, with condition as a fixed effect, trial as a random effect, and log2 of value as the response variable.

Supplementary Results:

The raw luciferase reporter results of those transfections are shown (Supplementary Table 1). Reduction of luciferase values were observed in constructs co-transfected with miR-616 for both the A allele and the G allele versus constructs that were not co-transfected with miR-616. This is shown in the plot of the average luciferase expression for the A and G alleles versus the average luciferase expression for each allele co-transfected with miR-616 (Supplementary Figure 9).

95% confidence intervals that estimated the difference between the A and G alleles and those co-transfected with miR-616 as well as the reductions between the A allele co-transfected with mir-616 versus the G allele co-transfected with miR-616 were calculated (Supplementary Table 2). A linear mixed effects model was used to test for significance between the reduction of the A allele when co-transfected with miR-616 and the reduction of the G allele when co-transfected with miR-616 (Table 3). The response variable was the log base 2 of all the raw experimental values. The predictors were the experimental condition as a fixed effect and the trial as a random effect. The reduction of the A allele when co-transfected with miR-616 versus the reduction of the G allele when co-transfected with miR-616 was found to be statistically significant (P value = 0.0215*) In order to better standardize the results across multiple trials, the averaged results for each trail are show (Figure 3) using the log2 residuals which have been plotted for each condition after the mean for each trial has been subtracted.

with miR-616.

Supplementary Table 1: Luciferase Values for miR-616 treatments on constructs containing either the wild type allele or the schizophrenia risk allele for the RenSP Luciferase assay are shown in the following table.

	A	A SCR	A 616	G	G SCR	G 616
Trial 1	541283	341089	450348	883213	654278	595157
Trial 1	504427	422809	375780	696570	638289	572227
Trial 1	493486	468903	253428	677310	537360	547474
Trial 2	59192	76075	47661	71288	97346	74258
Trial 2	57941	60878	47323	99900	79378	62142
Trial 2	65011	56657	36578	74258	71178	54707
Trial 3	517405	422116	268555	264754	366799	234000
Trial 3	427192	418991	313907	329410	299834	219931
Trial 3	517405	341899	255863	347685	242673	249074
Trial 4	42916	24052	11721	15228	13791	14535
Trial 4	39336	36631	17122	19347	17412	12412
Trial 4	25086	26332	18737	17582	14890	11257
Trial 4	31464	29670	14240	24245	14535	16567
Trial 4	21242	21912	15702	14521	12412	15869
Trial 4	24686	27594	9690	15772	17000	16685
Trial 5	24965	22619	16097	33032	20875	17853
Trial 5	33140	24809	12594	23018	26969	19175
Trial 5	39823	21801	11756	21574	23080	15871
Trial 6	68412	81764	31326	86242	66801	53123
Trial 6	63284	78487	45491	91955	83130	50676
Trial 6	59881	48236	26242	81861	80465	35368
Trial 6	56705	73387	44321	102979	73505	46015
Trial 6	49275	71050	28403	83285	91934	45017
Trial 6	99661	42574	42751	75920	76182	33006

Supplementary Table 2 Statistical Analysis of RenSP Luciferase Data: 95% confidence intervals (CIs) were calculated for the mean of condition A616 minus the mean of A, the mean of condition G616 minus the mean of condition G, and the difference between these two groups. All values are on the scale of log base 2 of the original experimental values.

Comparison Group	Description	Estimate	Lower Limit (95% CI)	Upper Limit (95% CI)
A/G Double Difference	A Difference – G Difference	-0.33722	-0.62479	-0.04966
A Difference	Mean of Condition A616 minus Mean of A	-0.8695	-1.0729	-0.6662
G Difference	Mean of Condition G616 minus Mean of G	-0.5323	-0.7356	-0.329

Supplementary Table 3:
A linear mixed effects model was used to test for significance between A Difference and G Difference. The response variable was the log base 2 of all the raw experimental values. The predictors were the experimental condition as a fixed effect and the trial as a random effect.
Fit: lme.formula (fixed = log2(Value) ~ Condition, random = ~1 | Trial)
Linear Hypotheses:

	Estimate	Standard Error	z Value	Pr(>IzI)	P Value
A/G Double Difference	0	-0.3372	0.1467	-2.298	0.0215*

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Adjusted p values reported -- single-step method)

Figure 9: Treatment of constructs containing either the wild type allele (G) or the schizophrenia risk allele (A) with miR-616 or scramble control miRNA shown for six trials. The log2 residuals have been plotted for each condition after the mean for each trial has been subtracted to normalize the results between trials.