2009

GENE DISCOVERY IN SCHIZOPHRENIA THROUGH IDENTIFICATION OF COPY NUMBER VARIATIONS (CNVs) IN DISCORDANT MONOZYGOTIC (MZ) TWINS

Christina Angela Castellani

GENE DISCOVERY IN SCHIZOPHRENIA THROUGH IDENTIFICATION OF COPY NUMBER VARIATIONS (CNVs) IN DISCORDANT MONOZYGOTIC (MZ) TWINS

(Spine title: CNV in monozygotic twins discordant for schizophrenia)

(Thesis format: Monograph)

by

Christina Angela Castellani

Graduate Program in Biology

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science

School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

ABSTRACT

Schizophrenia (MIM 181500) is a complex and devastating human disease that affects greater than 1% of the population (Huxley *et al.*, 1964). The disease has a heritability of 80%, but a concordance rate in monozygotic (MZ) twins of only 48% (Gottesman, 1991). This suggests a role for both genetics and random/environmental factors in the etiology of this complex disorder (Singh *et al.*, 2009). Copy Number Variation (CNV) is now a known candidate for disease associated variation in humans. CNVs have been identified as a common feature covering 12% of the genome in normal, healthy individuals (Feuk *et al.*, 2006; Redon *et al.*, 2006). Although little is known about the causes and consequences of this common genomic phenomenon, CNVs represent causal mutational events in any study of genetic causes of complex diseases. This study focuses on monozygotic twins who represent the best possible genetic match, but are discordant for the disease; this approach allows for a significant reduction of disease heterogeneity. A genome wide analysis of copy number variation on three pairs of monozygotic twins was performed using the Affymetrix Human Array 6.0. The results show that they differ significantly in variation profile and that the affected member of each twin pair had a significantly higher number of CNV when compared to their unaffected co-twin ($p=0.01$). In addition, a set of five *de novo* CNV's were found to be shared in all three unrelated patients. These regions contain genes, many of which play a role in neurodevelopment and have the potential to cause schizophrenia pathogenesis. The results support a role for *de novo* CNVs in the etiology of schizophrenia.

Keywords: copy number variations, schizophrenia, replication error, complex disease, discordance, monozygotic twins

ACKNOWLEDGEMENTS

I would like to first and foremost thank my supervisor Dr. Shiva Singh for allowing me the opportunity, support and funding necessary to pursue this research. I have learned a great deal from this experience and am thankful for the time he has taken to teach me.

To my lab mates, who offered continuous support, assistance and encouragement: Morgan Kleiber, Patrick McDonald, Haroon Sheikh, Melissa Symons, Luke Tapscott, Imran Tayyab and Julia Weng.

Thanks to David Carter from the London Regional Genomics Center at Robarts Research Institute who assisted with our analysis of the data using Partek Genomics Suite. In addition, to Steve Scherer and both Christian Marshall and Chao Lu from Dr. Scherer's group for their assistance in hybridizing our samples to the Affymetrix 6.0 Array at The Center for Applied Genomics in Toronto.

A special thank you to Dr. Richard O'Reilly, who both interviewed and procured the blood samples used in this study. In addition, I wish to acknowledge the patients who so graciously offered their samples for study in this research.

I am gracious for the support and knowledge of my advisory committee members, Dr. Kathleen Hill, Dr. Robert Cumming and Dr. Jim Karagiannis.

Most importantly, to my parents, Bruno and Gina, for their constant support and encouragement in helping me achieve any goals which I set out for myself. Words cannot express how lucky I am to have parents that believe in me like you do.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

°C: Degrees Celsius

ABI: Applied Biosystems

bp: Base pair

.cel file: Cell intensity file, a single intensity value is calculated for every probe on the chip

.cdf file: The chip description file (includes probe level data), provided by Affymetrix

.chp file: Chip file, signal values and presence calls for each probe set on the microarray (created from cel and cdf files)

CN: Copy Number

CNV: Copy Number Variation

C_T: The C_T (cycle threshold) is defined as the number of cycles required for the fluorescent signal to cross the threshold (exceed background level).

Comparative C_T: The quantity of target, normalized to an endogenous control and relative to a reference sample is calculated by $2^{-\Delta\Delta Ct}$

.dat file: The raw image file resulting from a hybridization

DNA: Deoxyribonucleic acid

DSM-IV: Diagnostic and Statistical Manual of Mental Disorders IV

Heritability: The proportion of phenotypic variation in a population that is attributable to genetic variation.

Kb: Kilobase

Mb: Megabase

MIM: Mendelian Inheritance in Man

MZ: Monozygotic Twins

NAHR: Non Allelic Homologous Recombination

NHEJ: *Non Homologous End Joining*

PCR: Polymerase Chain Reaction

QC: Quality Control

SNP: Single Nucleotide Polymorphism

Unique CNV: A copy number variant which is not found to be present in the database of genomic variants or 270 HapMap samples.

CHAPTER 1: INTRODUCTION

1.1 Schizophrenia

Schizophrenia (MIM 181500) is the most serious and debilitating of the functional mental illnesses and affects approximately 1% of the population (Huxley *et al.*, 1964). Schizophrenia has serious individual, family and societal burdens (Singh *et al.*, 2009). The onset of symptoms typically occurs in late adolescence to early adulthood though cases of child-onset schizophrenia are also frequent (Delisi, 2009). Schizophrenia is categorized by major impairments in both the perception and expression of reality (Andreasen, 1995). The disease presents with significant heterogeneity across patients and variability in both presentation and classification of the disease is common. Schizophrenia concordance rates for monozygotic twins average 48% while concordance rates for dizygotic twins average 17% (Gottesman, 1991; McGuffin *et al.*, 1994). There is no cure for schizophrenia though treatment often includes a combination of antipsychotic medications, cognitive therapy, daily support and social skills training. Approximately 10% of schizophrenic patients eventually die from suicide (Pompili *et al.*, 2007). Schizophrenia is a major burden on society, with over 230,000 patients diagnosed in one year in Canada (Goeree *et al.*, 2005). In addition to this, schizophrenia cost $2.02 billion in direct costs in 2004 alone, in Canada (Goeree *et al.*, 2005). Diagnosis takes place using the Diagnostic and Statistical Manual of Mental Disorders IV, Text Revision (DSMV IV-TR). According to the DSMV-IV, to be diagnosed with schizophrenia, a person must display:

a. **Characteristic Symptoms:** Two or more of the following each present for a significant portion of time during a one-month period: delusions, hallucinations, disorganized speech, grossly disorganized behaviour or catatonic behaviour and negative symptoms.

b. **Social/Occupational Dysfunction:** For a significant portion of time since the onset of the disturbance, one or more major areas of functioning such as work, interpersonal relations or self care are below the level achieved prior to the onset.

c. **Duration:** Continuous signs of the disturbance persist for at least six months.

1.2 Schizophrenia Development

There are two schools of thought with regard to schizophrenia development. The first is a neurodegenerative hypothesis and the second is a neurodevelopmental hypothesis, the latter claims that the brain does not develop properly from conception (Singh *et al.*, 2004). The neurodevelopmental hypothesis has, on a whole, received more support from the scientific community.

1.3 Schizophrenia Genetics

Schizophrenia onset typically occurs during adolescence or early adulthood. Environmental factors affecting the disease include maternal malnutrition, viral infections, birth complications, urban/rural living, poverty, winter/spring births as well as many more (Jablensky *et al.*, 2000; Mednick *et al.*, 1988; Torrey *et al.*, 1977). However, the effect of each individual environmental factor appears to be quite small. Physiologically, studies have shown that there is a significant decrease in ventricle size in the brains of affected individuals. In addition to this, schizophrenic patients show whole brain volume reductions of approximately 3% (Keshavan *et al.*, 2007).

Schizophrenia has a heritability estimate of 80% (Sullivan, 2008). It is therefore likely that inheritance is of a genetic predisposition, rather than the disease itself; with environmental factors playing a major role. Figure 1 shows the risk of developing schizophrenia based on relation to an affected patient. The risk of developing the disease increases with increasing genetic relatedness to an affected schizophrenic patient. It is clear however that genetic relatedness alone does not account for the full picture, if it did, we would expect monozygotic twins to be 100% concordant. Currently, the Schizogene online database of schizophrenia candidate genes lists over 1000 genetic studies

performed in the search of schizophrenia genes (Accessed June 2009). Through accepted observations of clinicians, we can say that schizophrenia appears to cluster in some families and be of a sporadic occurrence in other families (Delisi, 2009). Familial and sporadic schizophrenia are very similar and no consistent differences are seen between the two.

A scientific approach that takes into account many genes of small effect will allow us to explain both the non-Mendelian pattern of inheritance seen in the disease and the variability in presentation of symptoms and penetrance (Freedman *et al.*, 2001; Gershon, 2000; Gottesman *et al.*, 1967). The difficulty in the quest to unravel complex disorders calls for a need to reconsider our previous models and approaches. Solving the schizophrenia puzzle will require an interdisciplinary approach as well as advanced genetic techniques. There is a large body of evidence to support a significant role for genetics in schizophrenia, including the clustering of the disease in some families.

The search for a schizophrenia gene or set of genes has been the current focus of schizophrenia research. These efforts have largely focused on linkage and association studies across the genome. Although a large number of genomic regions have been identified which include genes of interest to the disorder, issues of reproducibility have hampered the contribution of these studies. Despite the collective efforts of many years of research, no single gene or set of genes can currently be viewed as the cause of schizophrenia despite the fact that almost every chromosome arm has a reported linkage to schizophrenia (DeLisi, 2000). The search for the underlying causes of schizophrenia remains a distinct challenge of medical science and much larger numbers of patients and controls may be needed to produce more effective results.

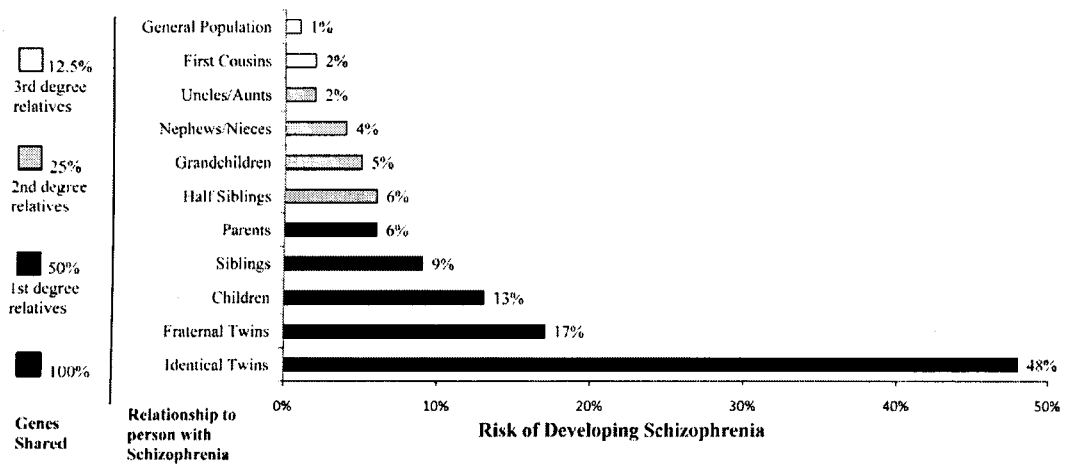


Figure 1. Risk of developing schizophrenia based on relatedness to individual with schizophrenia (Modified from Gottesman *et al*, 1991).

1.4 Copy Number Variation

Structural variants are the middle ground between small variants like SNPs (Single Nucleotide Polymorphisms) and large chromosomal variants that can be recognized on karyotypes (Scherer *et al.*, 2007). This variation includes large insertions, duplications and deletions of DNA, collectively termed Copy Number Variants (CNV) (Sebat *et al.*, 2007). A CNV therefore, can be defined as a piece of DNA, longer than 1 kb, with a variable copy number as compared to a reference genome (Feuk *et al.*, 2006). Two landmark studies by Sebat *et al.* and Iafrate *et al.* in 2004 laid the groundwork for further characterization of copy number variants (Iafrate *et al.*, 2004; Sebat *et al.*, 2004). CNVs are distributed across the genomic landscape, covering greater than 12% of the human genome in individuals with no obvious disease phenotype (Zogopoulos *et al.*, 2007). A copy number polymorphism (CNP) is a CNV that occurs in more than 1% of the population (Scherer *et al.*, 2007). Over 77% of copy number variants are predicted to include genes (Graubert *et al.*, 2007). The presence of large structural variants in the human genome challenges the notion that the diploid state represents a normal copy number for all regions of DNA present in the human genome (Zogopoulos *et al.*, 2007).

Copy Number Variation is now a well established cause of rare genomic disorders (Freeman *et al.*, 2006). CNV is now thought to have an effect in cancer (La Starza *et al.*, 2007), mental retardation (Sharp *et al.*, 2006) and neurological disorders (Lee *et al.*, 2006). It has been reported that a high frequency of *de novo* copy number variation is significantly associated with Autism Spectrum Disorders, with CNV being identified in 10% of patients with sporadic autism (Sebat *et al.*, 2007). In a recent study, Weiss *et al.*(2008) found a recurring microdeletion and a reciprocal microduplication that appear to explain autism susceptibility in approximately 1% of cases (Weiss *et al.*, 2008).

CNVs may also play a role in adaptive evolution. Copy number variation of the salivary amylase gene (AMY1) may stem from human adaptation. Subjects from populations with high starch diets on average have more AMY1 gene copies than those with low starch diets (Perry *et al.*, 2007). This finding suggests a role for copy number variation (CNV) in adaptive evolution. Copy number variation has also been associated

with a number of everyday traits in humans, including height differences (Weedon *et al.*, 2007). CNV may be a cause of explanation relating to both spontaneous disease and heritable disorders whose prevalence in the population seems to defy Darwinian logic. In theory, complex disease could be more susceptible to 'soft' forms of genetic variation – such as variation in noncoding sequences and copy number changes that change gene dosage without completely wiping out gene function (McCarroll *et al.*, 2007). This follows from the finding that genes involved in rapid evolution, including brain development, appear to be enriched in CNVs (de Smith *et al.*, 2007). It has therefore been proposed that structural variation has become a method for our genomes to remain both adaptable and in a state of constant flux.

Structural variants can theoretically affect phenotype through gene dosage effects, position effects, and by deletions that allow recessive phenotypes to emerge. Copy number variation arises both in families, as an inherited phenomenon, and in the individual, as a result of spontaneous or *de novo* events. Highly distributed CNV have already been documented in other species. Significant CNV's have been detected and studied in mice (*Mus musculus*) and in primate species such as chimpanzees (*Pan troglodytes*) (Graubert *et al.*, 2007).

There are a number of mechanisms that have been suggested to contribute to the existence of CNVs in the genome. The current prevailing mechanism for rearrangement is Non Allelic Homologous Recombination (NAHR) (Freeman *et al.*, 2006). NAHR occurs between very similar duplicated sequences and generates deletions, duplications, inversions and translocations. CNVs have been found to occur most often in regions that are reported to contain, or be flanked by, large homologous repeats and it is likely that these low-copy repeats create *hotspots* for NAHR to occur (Freeman *et al.*, 2006; Lupski, 2007). The second mechanism that has been implicated in the creation of copy number variants is Non Homologous End Joining (NHEJ). NHEJ utilizes short microhomologies to guide repair. A number of short (<6 bp) microhomologies have been found to be overrepresented at CNV breakpoint junctions (Arlt *et al.*, 2009). Although our understanding of CNV is in its infancy, the origin, size, genomic locations and population

genetics of CNVs have brought new insights to the structure of the human genome and the cause of complex diseases (Conrad *et al.*, 2007). It is clear that the knowledge we have about human variation to date only encompasses the first chapter of this rapidly unravelling story.

1.4.1 Copy Number Variation and Expression

It has been reported that copy number changes explain approximately 20 per cent of detected gene expression variation in humans (Stranger *et al.*, 2007). In addition to this, genome-wide expression data from six different tissues in mice showed variation in expression of genes situated in and around CNV regions. Copy number variation gene expression differences were found to extend up to a half megabase upstream and downstream (Henrichsen *et al.*, 2009).

Human monozygotic twins arise at a frequency of 1 in every 250 live births. The use of twins in this study helps to significantly reduce patient heterogeneity and are therefore of special interest to genetic research (Singh *et al.*, 2002). At the outset of this research, the CNV profiles of monozygotic twins was not known. Bruder *et al.* (2008) published a study on 19 pairs of concordant and discordant monozygotic twins, which showed copy number variation within pairs of both groups (Bruder *et al.*, 2008). This research follows that of Fraga *et al.* (2005) which showed that as monozygotic twins age, the number of epigenetic differences between them increases (Fraga *et al.*, 2005).

1.5 Schizophrenia and Copy Number Variation

Recently, a number of *de novo* copy number mutations have been found to be risk factors for sporadic cases of childhood neurodevelopmental disorders, including autism (Marshall *et al.*, 2008; Sebat *et al.*, 2007). This recent success has led researchers to believe that copy number variation may be involved in the etiology of schizophrenia; also a neurodevelopmental disorder. In addition to the similarities between these two disorders, schizophrenia is marked by gene expression changes and these types of changes may arise as a result of copy number variation (McInnes *et al.*, 2006; Mirnics *et al.*, 2006; Tsankova *et al.*, 2007; Verveer *et al.*, 2007). Furthermore, the limited success

in identifying genetic modifiers of schizophrenia necessitates both novel approaches and innovative hypotheses.

Copy Number Variation studies have been a recent focus in the field of schizophrenia genetics. Three independent groups recently performed major genome-wide searches for CNV differences. The most notable result of these studies was two copy number variants, 1q21.1 and 15q13.3, which showed associated deletions in two independent studies (Stefansson *et al.*, 2008; The International Schizophrenia Consortium *et al.*, 2008) A summary of three major studies is provided below:

Stefansson *et al.*(2008): 1,433 patients with schizophrenia, 33,250 controls. Patients were from England, Finland, Germany, Iceland, Italy and Scotland. Results: Deletions at chromosome 1q21.1, 15q11.2 and 15q13.3 were highly associated with schizophrenia (Stefansson *et al.*, 2008).

The International Schizophrenia Consortium (2008): 3,391 patients, 3,181 ancestry matched controls. Patients were from Bulgaria, England, Ireland, Portugal (Azores), Scotland and Sweden. Results: 1.15 x increase in CNV in schizophrenic patients (>100 kb CNVs). 1q21.1 and 15q13.3 were associated with schizophrenia (The International Schizophrenia Consortium *et al.*, 2008).

Walsh *et al.*(2008): 150 patients with schizophrenia, 268 ancestrally matched controls. Patients were from USA only. Results: Novel deletions and duplications present in 15% of adult and 20% of childhood onset schizophrenia (COS) patients versus 5% of controls. In childhood onset schizophrenia patients only, 2q31.2, 2p16.3 (NRXN1 gene), and 16p11.2 were associated with the disease. In addition, the genes found to be in the copy number variable regions were significantly overrepresented by neurodevelopmental genes (Walsh *et al.*, 2008).

In addition to these studies, Xu *et al.*(2008) published a ground breaking study in which *de novo* CNVs were found to be approximately 8 times higher in patients with sporadic schizophrenia when compared with unaffected controls (Xu *et al.*, 2008). This

finding lays claim to the possibility that rare mutations may contribute to the etiology of schizophrenia and that multiple rare mutations at many different loci could be responsible for the high level of heterogeneity seen in the disease.

In July 2009, three papers published by separate groups in the journal *Nature* concluded that thousands of alleles of very small effect size likely contribute to disease pathology in schizophrenia (Shi *et al.*, 2009; Stefansson *et al.*, 2009; The International Schizophrenia Consortium *et al.*, 2009). One of these groups, *Shi et al.* (2009) performed a meta-analysis of over 8,000 cases and 19,000 controls (Shi *et al.*, 2009). This study represents the largest sample size used to-date in a schizophrenia study of its kind.

1.6 Hap Map Control Samples

The Hap Map samples, which were used as a control group in this study, are the result of an international effort to identify human haplotypes and find a standard set of common allele SNPs from healthy, normal individuals. The project is a partnership of researchers from Canada, China, Japan, Nigeria, the United Kingdom and the United States. The goal is a public resource which will help scientists find genes associated with human disease.

The Hap Map 270 samples include:

- 30 CEPH trios (Caucasians)
- 30 Yoruban trios
- 45 unrelated Han Chinese
- 45 unrelated Japanese samples

1.7 Database of Genomic Variants

DNA can vary widely and several extra copies of genes can exist or be missing completely in healthy individuals. The Database of Genomic Variants is an effort by The Center for Applied Genomics (TCAG) (Toronto, Canada) to provide a complete record of the structural variation found in normal, healthy humans (http://projects.tcag.csa/variation/) (Iafrate *et al.*, 2004). Variation in the database includes any reported segments that are larger than 1kb. The database serves as a

representation of variation in healthy control samples for use in any genetic study. A CNV in the database can be assumed to be non-pathogenic and therefore is not reported as novel variation in a discovery or association study, including this study. The number of CNVs currently listed in the Database of Genomic Variants is well over 31,000 (May 2009) (Iafrate *et al.*, 2004). This number can be contrasted to the number of variants in the database in May of 2008, which was just over 11,000 (http://projects.tcag.ca/variation/). It is clear that we have only touched the surface when it comes to a full understanding of the copy number variation found in healthy adults.

1.8 Hypothesis

Copy number variation (CNV) is involved in the causation of schizophrenia. Copy number differences in monozygotic twins discordant for schizophrenia will identify the best potential regions for schizophrenia contributing loci.

1.9 Objectives

i. Identify CNV differences between MZ twin pairs discordant for schizophrenia using Affymetrix's Human SNP Array 6.0. The use of twins addresses the issue of heterogeneity in schizophrenia patients allowing us to reduce inter-individual variability.

ii. Analyze chip data using Affymetrix Genotyping Console and Partek Genomics Suite. Compare SNP 6.0 hybridizations with 270 Hap Map Samples (Chip data provided by Affymetrix) and >25,000 CNVs catalogued in the Database of Genomic Variants.

iii. Characterization of CNVs based on potential involvement in schizophrenia, with genomic location and function to provide "pathogenic clues".

iv. Confirm CNVs that are novel (not identified as a CNV in the normal population by the Database of Genomic Variants and not found in 270 Hap Map samples) and shared (among affected twins) using Real Time PCR.

CHAPTER 2: MATERIALS AND METHODS

2. 1 Ethics Review

This research was undertaken following approval from the Research Ethics Board of The University of Western Ontario (Appendix A).

2.2 Collection of Monozygotic Discordant Patients

Through the assistance of Dr. Richard O'Reilly, Regional Mental Health Care, London, ON, our lab acquired whole blood samples from three pairs of monozygotic twins discordant for the schizophrenia phenotype and two 22q11 deletion syndrome experimental controls. Deletion syndrome patients were used as experimental controls because they provided the best possible way to assess known copy number variation. This allowed for direct evaluation of our experimental design. Dr. O'Reilly received consent from the individuals and was also the physician responsible for the diagnosis of each patient using DSM-IV criteria. Whole blood samples were collected and immediately stored at -80°C.

2.2.1 Sample Biographies

- **Twin Pair 01:** Female. 11 years passed between diagnosis and sample collection. The twins were firstborn children to a 21 year old woman, the twins were born vaginally 23 minutes apart. The unaffected twin was born normally while the affected twin was born breeched. Both twins were significantly underweight and were therefore kept in the hospital until they gained enough weight to go home. The affected twin has separately been diagnosed with both hypoglycaemia and hypothyroidism, for which she took Synthroid but was not taking the drug during the time of sample collection. The affected patient of this twin pair was said to have been "almost always 2 to 6 weeks behind" the development of her identical twin. Family history is negative for psychiatric disorders with the exception of one second degree relative treated for a crisis-induced brief, but severe, depression. Approximately 6 months prior to onset of her illness, the affected

twin's family recalls occasional unusual behaviour including irrational fears and premonitions and complaints of headaches.

Affected Twin 01 Disease Presentation: Presentation of the disease begun with inappropriate giggling and bizarre behaviour as well as very awkward posturing and mute behaviour. She was also noted to be withdrawn and suspicious with suicidal ideation. Initial diagnosis was catatonic schizophrenia with paranoid features. She suffered from visual, auditory, and somatic hallucinations, incoherent speech, confusion, flat affect and grossly agitated/bizarre behaviour and was treated with antipsychotics (IM Prolixin at the time of testing). The affected twin was hospitalized on many occasions. (Dr. O'Reilly, RMHC London).

- **Twin Pair 02:** Male. 18 years passed between diagnosis and sample collection. The twins were the second and third born children to a 26 year old mother following a normal, full term pregnancy. The twins were both born breech 4 minutes apart. Developmental milestones seemed to be both on time and consistent for both twins. Both twins were noted to have bilateral inguinal hernias as infants although surgery was only necessary on the affected schizophrenic twin; the surgery occurred at age 1. Family history is positive for alcoholism and a genetic metabolic disease in a primary relative. Approximately six months prior to his 21st birthday the affected twin started to show symptoms of his illness. It is of note that the well twin had more birth trauma, had history of a head injury and abused street drugs more heavily.

  Affected Twin 02 Disease Presentation: Symptoms began with an episode of mutism with catatonic behaviour. He was hospitalized on many occasions after being diagnosed with catatonic schizophrenia. He was prescribed antipsychotic medications to control the symptoms. He suffered from paranoid and bizarre delusions, loose associations, delusional thinking, auditory and visual hallucinations and flattened affect. (Dr. O'Reilly, RMHC London)

- **Twin Pair 03:** Male. 25 years passed between diagnosis and sample collection. The twins were first born children to a 22 year old woman following an uncomplicated pregnancy. The twins were both born vertex, 8 minutes apart. For both twins, developmental milestones were within normal limits. Family history is negative for mental illness. As early as kindergarten, the affected twin's teacher noted that he could not tie his shoes, had a strange gait and was more withdrawn than his brother. The affected twin was referred for psychological and neurological testing. The only abnormal result was some occipital lobe trauma in the brain scan. The affected twin was also sensitive to loud noises as a child. In grade 4, the affected twin began having outbursts where he would "scream or cry" in the classroom as well as trouble concentrating and a withdrawn affect. He was treated with Dexedrine briefly (which made symptoms worse) and Dilantin (which improved behaviour modestly) for one year. The affected twin was involved in a physical fight at age 20 which caused brief unconsciousness, a concussion and hospitalization. The affected twin feels that paranoid ideation began at this point.

  Affected Twin 03 Disease Presentation: At age 24, he became increasingly paranoid and would occasionally hurt himself. At age 26, he was diagnosed with schizophrenia. He has never been hospitalized, but has been treated with many antipsychotic medications; of them, low dose haloperidol has been the most successful in controlling his paranoia. His symptoms included outbursts of screaming and crying, delusional thinking, auditory hallucinations, mildly flattened affect. In addition to this, the affected twin reports having had a period of depression lasting 1 year (Dr. O'Reilly, RMHC London).

- **Del22q11 Experimental Control 1 (Schizo 07):** Confirmed 22q11 deletion by chromosome karyotype via FISH analysis of blood at the Regional Cytogenetics Laboratory in London, ON. DOB: 1989/10/14. The patient presented with a developmental delay and facial dysmorphic features as well as cardiac problems *(Dr. O'Reilly, RMHC London).*

- **Del 22q11 Experimental Control 2 (Schizo 08):** Confirmed 22q11 deletion by chromosome karyotype via FISH analysis of blood at Credit Valley Hospital in Oakville. DOB: 1985/12/27. The patient presented with odd behaviours, cranio-facial dysmorphology and developmental disability. Strong family history of developmental disability (Dr. O'Reilly, RMHC London).

NOTE: Patients from this study were all individuals whose permanent residency was Southwestern Ontario. The confirmed diagnosis of schizophrenia was made by Dr. O'Reilly (Regional Mental Health Care, London ON). Experimental controls were also assessed by Dr. O'Reilly for both schizophrenia and related disorders. Previous twin studies have established that the vast majority of twin pairs that become concordant for schizophrenia will do so within 5 years of initial onset of the illness by the first affected twin (Belmaker *et al.*, 1974).

2.3 DNA Extraction and Quality Control

DNA was extracted using the QIAamp DNA Blood Micro Kit (Qiagen, Mississauga) and the Archive Pure DNA Blood kit (5Prime, Germany) according to the manufacturer's instructions (Extraction ratio: 75µL whole blood to every 25µL buffer). Both kits are optimized for the purification of genomic DNA from small volumes of blood.

Polymerase Chain Reaction: 30 cycles

Reaction Setup:

- Buffer 2.5 µL
- $MgCl_2$ 1 µL
- Taq 0.15 µL (added last)
- D2-TIA Primers 0.5 µL each
- dNTP 0.1 µL
- DNA template 1 µL
- dH_2O 19.25 µL

Cycling Conditions:

- 94°C 4 min
- 95°C for 30', 60°C for 30', 72°C for 30'
- 72°C for 7 minutes

Quantity of DNA was checked using a spectrophotometer; quality of DNA was assessed by 260/280 calculations. A 6% polyacrylamide gel was also used to check quality of isolated DNA.

2.4 Hybridization of genomic DNA to Affymetrix Human SNP 6.0

All 8 extracted genomic DNA samples were hybridized to the Affymetrix Human SNP 6.0 (Santa Clara, CA) at The Center for Applied Genomics (TCAG) in Toronto, ON with the assistance of Dr. Steve Scherer and his team according to Affymetrix's protocol.

In addition, the DNA from 270 HapMap Samples used in this study were hybridized by Affymetrix to the Affymetrix Human SNP 6.0 Chip (Santa Clara, CA) and subsequent data files were provided free of charge to our lab for use in this study.

2.4.1 Choice of Platform

The Affymetrix Human SNP 6.0 Microarray was chosen for this study. This chip is currently the most comprehensive chip available in the field. The average number of CNVs identified per genome is 2-25 when using the 5.0 SNP array, 30-100 when using the Illumina 1M array and 50-150 when using the Affymetrix 6.0 platform.

The use of Single Nucleotide Polymorphisms to assess copy number variants is efficient because more is known about SNPs and they are easier to classify and work with than copy number variants. In the presence of copy number variability in a genome, individuals may theoretically harbour any number of copies of two or more SNP alleles if the CNV overlaps it.

2.4.2 Affymetrix Human SNP 6.0

The Affymetrix Human SNP 6.0 Array is a high density array with a median intermarker distance of less than 700 bases. The chip includes 1.8 million markers for

genetic variation and therefore represents more genetic variation on a single array than ever before. The chip includes 906,600 SNP probes and 946,000 CNV probes. Of the copy number variation probes on the chip, 202,000 target known regions of variation from the Database of Genomic Variants (Toronto, ON). One major advantage to the use of the SNP 6.0 chip is the identification of accurate breakpoints. Each SNP or CNV on the chip is assessed using a perfect match (PM) method and 4 probes per SNP are present on the array.

The Affymetrix Human SNP 6.0 Array requires a minimum concentration of 50 ng/µl of genomic DNA per chip. Genomic DNA is first digested with Nsp and Sty restriction enzymes in separate reactions and ligated to adaptors that recognize the four base pair overhangs generated by the digest. All of the fragments that result from restriction enzyme digestion are substrates for adaptor ligation (despite the size of the fragment). A generic primer that recognizes the adaptor sequence is then used to amplify adaptor-ligated DNA fragments. PCR conditions are optimized to preferentially amplify only those fragments in the size range of 200 to 1,100 base pairs. Amplification products from each digest are pooled, the DNA is end-labelled with terminal deoxynucleotidyl transferase and hybridized to the Genome Wide Human SNP 6.0 Array. Fluidics station 450 was used to wash and stain the array and Genechip scanner 3000 was used to scan the samples for fluorescence. Hybridization and scanning produced .dat and .exp files for each chip. Lastly, GeneChip Operating Software (GCOS) was used for processing and .cel file creation. One .cel file is created for each array.

2.4.3 Copy Number Analysis

To assess and report copy number variation regions, the copy number analysis and segment reporting tool in Affymetrix Genotyping Console 2.1 was used. In addition, Partek Genomics Suite was used to confirm the results found using Genotyping Console 2.1. Reference copy number state was assumed to be 2, except for the Y chromosome where the copy number reference state used was 1.

Affymetrix Genotyping Console 2.1 analysis thresholds:

- Reported CNV regions were a minimum of 100 kb in length
- Regions had a minimum of 10 probes/segment
- Use of Birdseed V2 Genotyping
- Hidden Markov Model (HMM) to calculate copy number statistics
- CN State: 0, 1, 2, 3, 4+
- Prior probabilities: 0.2
- Beta: .001
- Alpha 0.01
- Smoothing: log 2 ratio
- Quantile normalization was used, which seeks to help overcome technical variance in microarray studies in order to maximize the detection of biological variance. Array hybridization data was adjusted for GC content, fragment length and probe sequence. (Quantile normalization is the standard for association studies)

Partek Genomics Suite analysis thresholds:

- Reported CNV regions were a minimum of 100 kb in length
- Regions had a minimum of 10 probes/segment
- Max probability: 0.95
- Genomic Decay: 0
- Sigma: 1

Library files (which contain information about the probe array design, layout, probe use and content as well as scanning parameters), reference model file and browser annotation files were provided by Affymetrix and downloaded from www.affymetrix.org.

Copy number variation analysis workflow:

1. Data set addition (.cel file)
2. Perform quality control
3. Perform genotyping
4. Run the copy number algorithm
5. Run the segment reporting tool (identifies copy number segments within each sample)

Note: The copy number algorithm estimates copy number intensity for each marker (SNP and CNV) across all samples in the batch and an added .ppw reference file.

This analysis used a reference .ppw sample set created from 270 HapMap samples hybridized to the Affymetrix Human SNP 6.0 Array. Using the Genotyping Console 2.1 CN/LOH Reference Model File Creator a .ppw reference file was created and used in this analysis.

2.5 Characterization of CNVs based on shared regions and potential involvement in schizophrenia

Manually characterized results from all samples for shared regions that met all previously identified thresholds were selected for further analysis. Novel copy number variants that were shared by two or more affected patients were further characterized. The known function of any genes in the shared copy number variable regions was researched. In addition, any prior research to show involvement of the identified genes in the pathology of schizophrenia or related disorders was used as selection criteria for more in-depth analysis.

2.6 Real Time PCR using Applied Biosystem's Custom Copy Number Assays

One gene from each novel and shared (between all three unrelated, affected twins) copy number variant was assessed for copy number using Applied Biosystem's TaqMan Gene Copy Number Assays.

2.6.1 TaqMan probe design

Primers and probes for selected genes in shared regions identified as potential schizophrenia modifiers were designed using Applied Biosystem's File Builder 3.1 (Table 1). Custom Applied Biosystems probes were found to be 100% consistent with karyotyping by the company.

2.6.2 Real Time PCR using Applied Biosystem's Step One Real Time PCR System

This analysis used a Comparative C_T method for Real Time PCR analysis using the Applied Biosystem (California, USA) Step One Real-Time PCR System (48 well platform, three colour system). Comparative C_T ($\Delta\Delta C_T$), collects data following each extension step of PCR. Fluorescent ROX was used as a passive reference dye. Fluorescent dye FAM was used to label each gene specific assay and fluorescent dye VIC was used to label the endogenous control gene, RNAseP. Reactions were run in replicates of four and negative controls (where H_20 replaced template) were included with each run. Samples were run in multiplex. The Comparative C_T method determines the relative target quantity in samples. The StepOne software records amplification of both the target and the endogenous control in a **sample** and a **reference sample**. In our case, the reference sample used was each affected patient's identical twin and is used to directly compare amplification from the sample and the reference in order to determine relative expression values. Since the threshold cycle number (C_T) is a function of the amount of starting template the relative amount can be used to identify the copy number (Wilke *et al.*, 2000).

Table 1. Forward (Fwd) and Reverse (Rev) primer sequences as well as TaqMan probe sequences for the four genes of interest. Primers and probes were designed using Applied Biosystem's File Builder 3.1.

Gene Name	Primer Sequences	FAM labelled TaqMan Probe Sequence
ATXN2	FWD 5'-GCCCCGGGACCGTAT-3' REV 5'-GAGCTCTGCCGGGAGG-3'	5'-CCCTCCGCCGCCCCT-3'
UNC5C	FWD 5'-TCTGCGGGCGACAGC -3' REV5'-CACGAGCATTTGCAGCAAGTAT-3'	5'-CCCAGTCCCAGTCCGC-3'
CLSTN1	FWD 5'-GCCGGACCCGAGAGAGA-3' REV5'-GCCTCAGAGCAGCGTCTT-3'	5'-CCGCCGCCATCTTA-3'
NR1D1	FWD 5'-GGTGGACGCCCAGTGA-3' REV5'-GAGAGAGAAGTGCAGAGTTCGATTC-3'	5'-TTCTGCCGCTGCCCC-3'

All measurements were normalized using the endogenous control, RNAseP. The relative quantity of each target in each sample was found by comparing normalized target quantity in each tube to the normalized target quantity in the reference sample. RNAseP is present in exactly two copies in any human diploid genome; it is a type of ribonuclease which cleaves RNA.

Reaction setup and run method

- 2X TaqMan Universal PCR Master Mix 12.5 µl / reaction
- 20X VIC RNaseP Assay 1.25 µl
- 20X FAM labelled gene specific assay1.25 µl
- Genomic DNA 10 µl (20 ng/µl DNA)

Total Reaction Volume = 25 µl

Run Method:

50°C for 2 minutes (HOLD)

95°C for 10 minutes (HOLD)

95°C for 15 sec, 60°C for 60 sec (40 CYCLES)

Reactions were run in **standard** mode with a **genomic DNA** template.

Relative Quantity (RQ) values are determined by the ΔΔCt ((FAM Ct – VIC Ct)$_{SAMPLE}$ – (FAM Ct- VIC Ct)$_{CALIBRATOR}$) method. Gene copy number is double the relative quantity (RQ) value. An RQ value of 0.5 = 1 copy, 1.0 = 2 copies, 1.5 = 3 copies and so on. In other words, the actual gene copy number is two times the value of the relative quantity. Threshold Cycle (C_T) is the PCR cycle number at which the fluorescence meets the threshold in the amplification plot.

2.7 Sample Size and Power

The use of twins addresses the issue of heterogeneity in schizophrenic patients allowing us to significantly reduce inter-individual variability. There is little precedent for the sample size of twins needed to detect CNVs, however, any variation between monozygotic twins should be seen as unique, that is, that statistical analysis should center on differences within twin pairs. Twins are of distinct interest when interrogating the genetics behind complex human diseases as twins provide a natural match and are able to help us reduce the effect of many potential environmental factors, for example, maternal nutrition and gestation length.

The probability of a neurodevelopmental gene being affected by copy number variation in the human genome is 9.24%. The probability of the same gene being affected in all three affected patients by chance alone is therefore 0.078%. A copy number variant present in all three affected patients and affecting the same neurodevelopmental gene would therefore represent a statistically rare event.

CHAPTER 3: RESULTS

The results generated by this study are presented below and fall under the broad categories of DNA isolation, hybridization, copy number analysis and data collection, real time analysis and bioinformatic characterization of genomic regions.

3.1 DNA Isolation

Genomic DNA was isolated from whole blood and the quality of DNA was assessed using a 6% polyacrylamide gel stained with ethidium bromide (Figure 2). The DNA from all 8 samples was found to be of high quality and in high enough quantity to proceed with hybridization of the DNA to the Affymetrix Human SNP 6.0 Array.

3.2 Affymetrix Human SNP 6.0 Array

Quality control (QC) was assessed using Affymetrix Genotyping Console 2.1. Quality control was measured using 3022 signature SNPs on the array. Quality control call rates were all found to be greater than 98%. A chip with a quality control call rate greater than 86% is considered to be valid for use according to the manufacturer (Table 2).

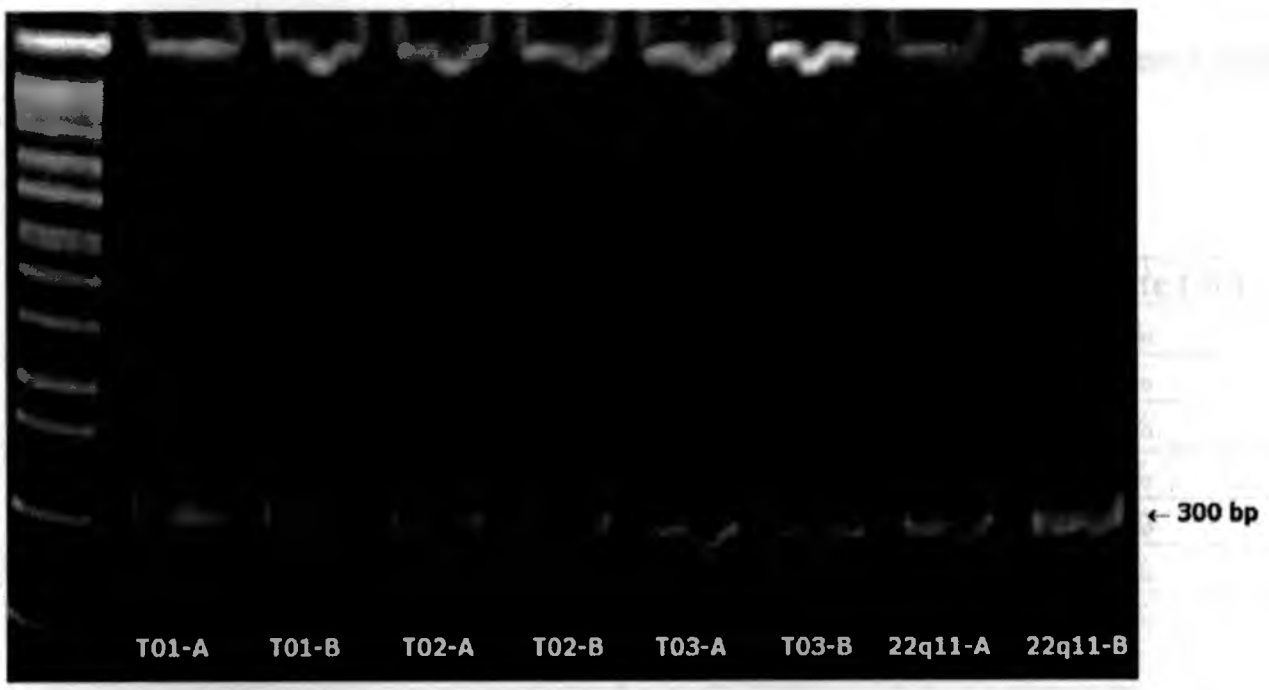


Figure 2. Genomic DNA extracted with the QIAamp DNA Micro Kit and amplified by D2-TIA forward and reverse primers. A 6% PAG stained with ethidium bromide was used. 1 Kb plus DNA ladder. Expected amplicon = 300 bp.

Table 2. Quality control call rates for all 8 samples, from Affymetrix Genotyping Console 2.1.

Sample Name	QC Call Rate (%)
Twin 1a	98.58%
Twin 1b	98.81%
Twin 2a	98.64%
Twin 2b	98.61%
Twin 3a	98.18%
Twin 3b	98.11%
22q11 1	98.38%
22q11 2	98.05%

All of our samples were found to be *in bounds* by Affymetrix Genotyping Console 2.1. For an array to be considered *in bounds,* it must have a Contrast QC that is greater than 0.4 and a MAPD value that is less than 0.4. Contrast QC, which is the ability to resolve SNP signals into clusters was greater than 0.4 for all of the samples in this study. In addition, MAPD, which measures the median of the absolute values of all pairwise differences between log2 ratios, was less than 0.4 for all of the samples in our study.

3.2.1 SNP Twin Concordance Test

The Affymetrix Human SNP 6.0 Array along with Affymetrix Genotyping Console 2.1 allows for conformation of zygosity of .cel file samples using an internal SNP concordance feature that utilizes 72 signature SNPs.

The results of the zygosity concordance check for our twin pairs are as follows:

- TWIN PAIR 01: 99.80% concordance
- TWIN PAIR 02: 99.69% concordance
- TWIN PAIR 03: 99.38% concordance

These values were sufficient to confirm the assumed zygosity of the monozygotic twin pair in this study.

3.2.2 22q11 Experimental Control Results

Our experimental controls allowed us to assess a known copy number variable region in two patients. Both of the 22q11 deletion syndrome patients used as experimental controls in this study had been previously assessed by karyotype in a clinical lab to have known 22q11 deletions. Default copy number state in this region is 2; an expected deletion in this region would therefore show as a copy number state of 1 or 0. Our experimental controls both showed a copy number state of 1 for a known 22q11 deletion region, though some minor variability in breakpoints was seen. The breakpoints that are variable in this region according to the literature are: 17063468 – 20207821,

17063468 – 18725000, 17383947 – 18725000, and 17579237 – 19600433. Existence of one or more of the aforementioned breakpoint regions causes 22q11 deletion syndrome. The deletions identified in our experimental controls by Affymetrix Genotyping Console 2.1 in this study are shown in Figure 3. A deletion was identified in Sample 07 (Experimental Control 1) from 17270046 – 19941337 on chromosome 22. A deletion was also identified in Sample 08 (Experimental Control 2) from 17256416 – 20056498 on chromosome 22. Minor variability in breakpoints is most likely due to probe positioning (Figure 3).

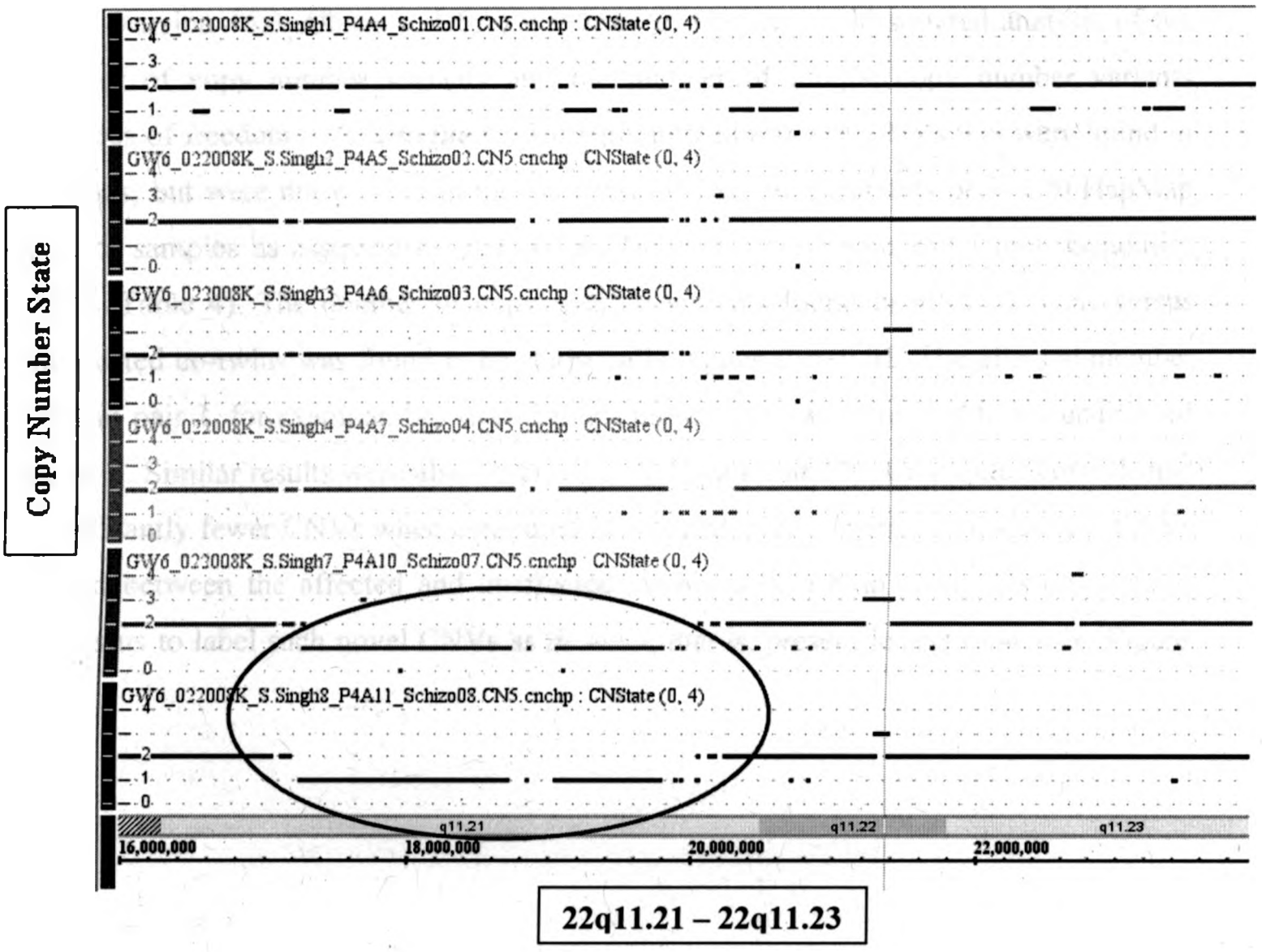


Figure 3. 22q11 deletion syndrome experimental control results. Samples Singh7 and Singh8 represent 22q11 deletion syndrome individuals.

3.2.3 Copy Number Analysis

The affected member of each of the three twin pairs had a significantly higher number of unique CNVs as compared to their unaffected co-twin (p=0.01, chi squared). SPSS Version 16.0 (Chicago, USA) was used to perform a chi squared analysis of the number of copy number variants and the number of unique copy number variants (degrees of freedom = 2). ***Unique CNVs*** were defined as those CNVs that were found in samples, but were not present in the Database of Genomic Variants or in 270 HapMap control samples as assessed by Affymetrix Genotyping Console and Partek Genomics Suite (Table 4). The number of copy number variants found in affected twins versus unaffected co-twins was found to be statistically higher at p=0.01. The affected member of twin pair 2, for example, has over 5 times more CNVs as compared to his unaffected co-twin. Similar results were also observed in twin pair 1 and 3. Unaffected controls had significantly fewer CNVs when compared to affected twins. There are few unique CNVs shared between the affected and unaffected members of a twin pair. This observation allows us to label such novel CNVs as *de novo*, that is, present in one twin only (Figure 5).

Table 3. Unique CNVs in our discovery study distributed by size. Unique CNVs are defined as any variants that are found in samples, but are not found in the Database of Genomic Variants or 270 Hap Map Samples.

Novel CNVs	100 Kb	250 Kb	500 Kb	750 Kb	1 Mb
Twin 1A	152	40	8	1	1
Twin 1B	0	0	0	0	0
Twin 2A	37	4	0	0	0
Twin 2B	7	1	1	1	1
Twin 3A	318	148	42	19	4
Twin 3B	146	55	11	2	0
Del22q11 1	1	0	0	0	0
Del22q11 2	1	0	0	0	0
Affected patient CNVs not shared with co-twin	**100 Kb**	**250 Kb**	**500 Kb**	**750 Kb**	**1 Mb**
Twin 1A	152	40	8	1	1
Twin 2A	34	4	0	0	0
Twin 3A	281	122	32	12	2

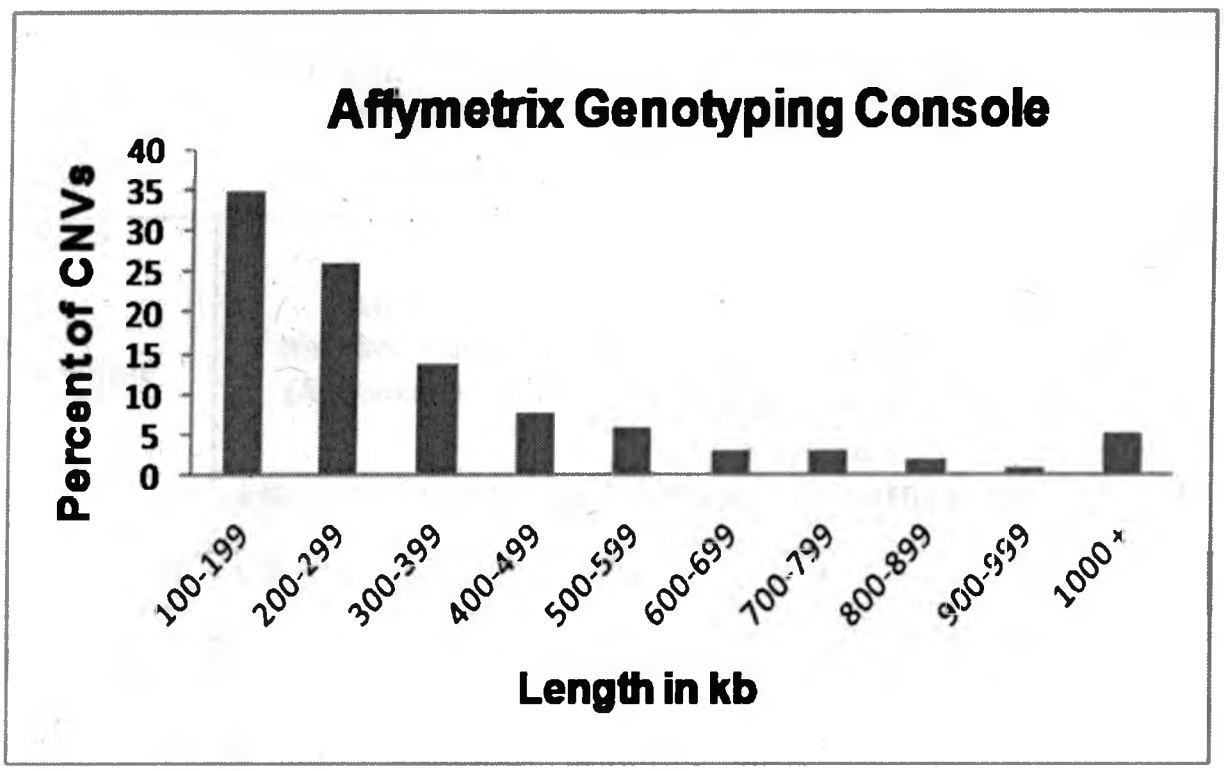
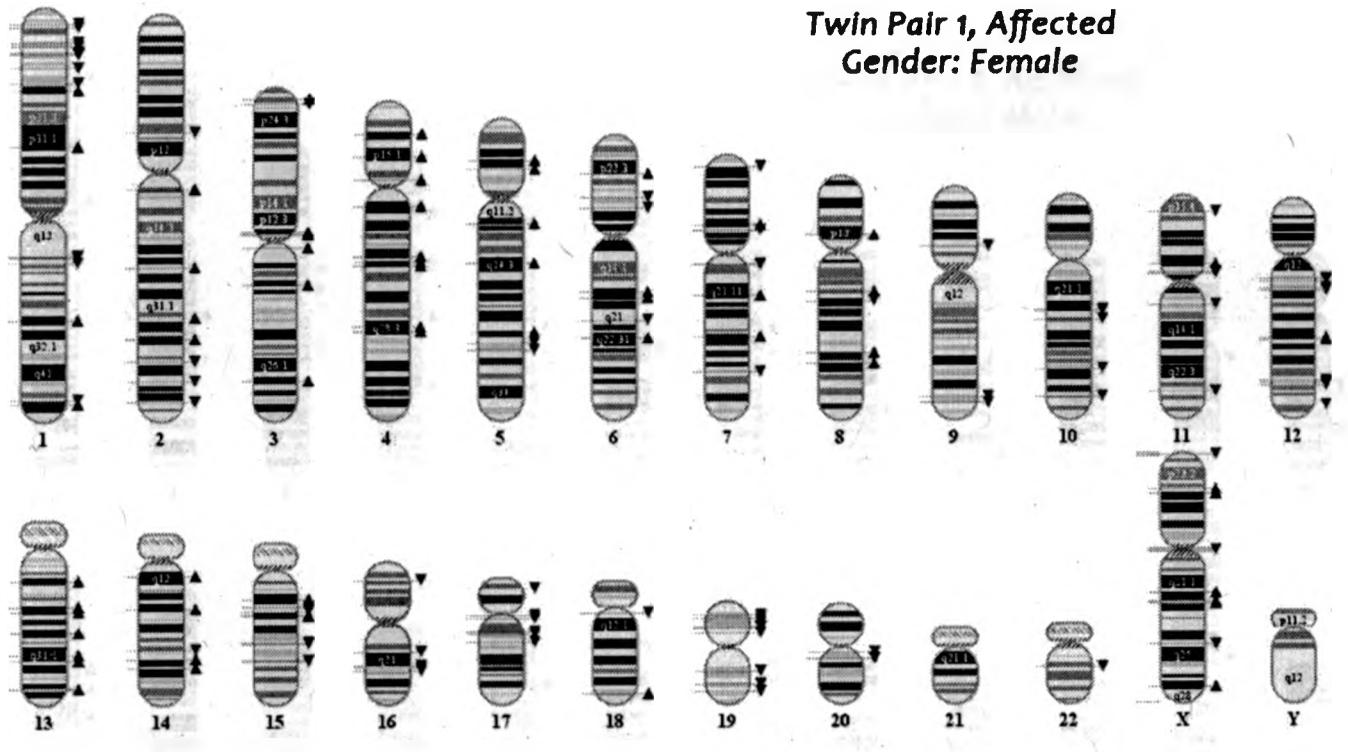


Figure 4. Distribution of Copy Number Variant sizes as identified in this study by Affymetrix Genotyping Console 2.1.

Table 4. Copy Number Variants (>100 kb) in MZ twin pairs discordant for schizophrenia. Unique CNVs are those that are found in samples but were not found in the Database of Genomic Variants (TCAG) or 270 Hap Map samples.

Sample	All Copy Number Variants (Affymetrix GC)		Unique Copy Number Variants (Not present in HapMap 270 or Database of Genomic Variants)		Unique Copy Number Variants not shared with twin	Deletions and Duplications (Unique CNVs)
	Affected	Unaffected	Affected	Unaffected	Affected	
Twin Pair 1	438	18	152	0	152	**Dels:** 88 **Dups:** 64
Twin Pair 2	107	35	37	7	34	**Dels:** 11 **Dups:** 23
Twin Pair 3	844	414	318	146	281	**Dels:** 196 **Dups:** 85
22q11 1	35		1		N/A	**Dels:** 0 **Dups:** 1
22q11 2	23		1		N/A	**Dels:** 0 **Dups:** 1



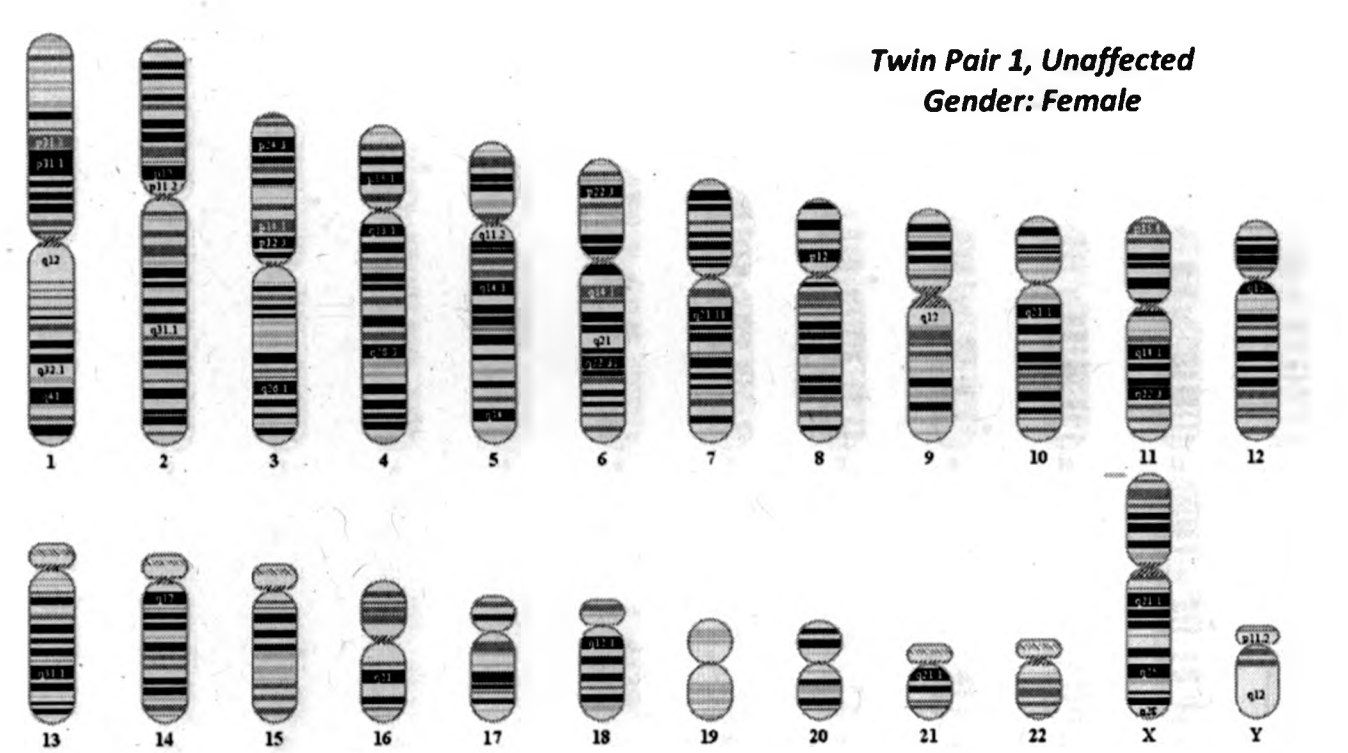
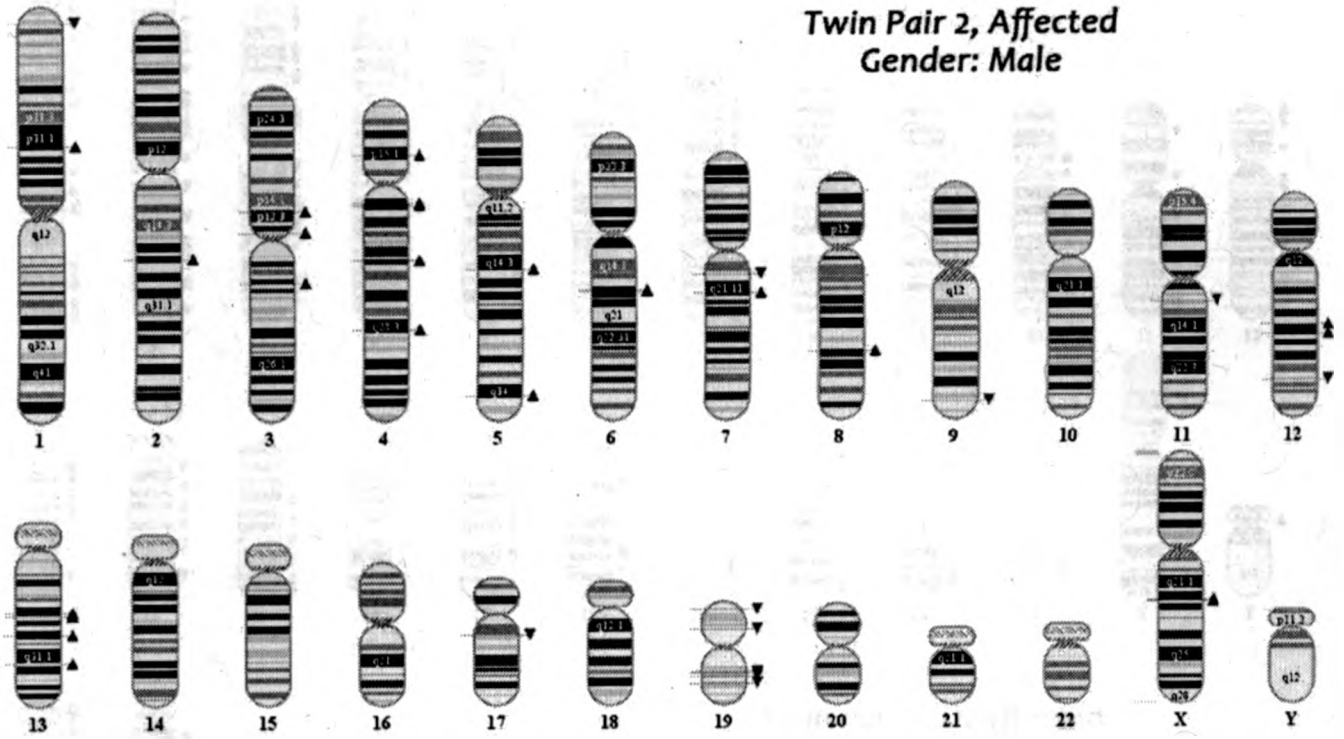


Figure 5a. Unique CNV distribution in Twin Pair 1. Unique CNVs are those that are found in samples, but are not found in the Database of Genomic Variants or 270 HapMap control samples. Triangles pointing down represent a loss and triangles pointing up represent a gain.



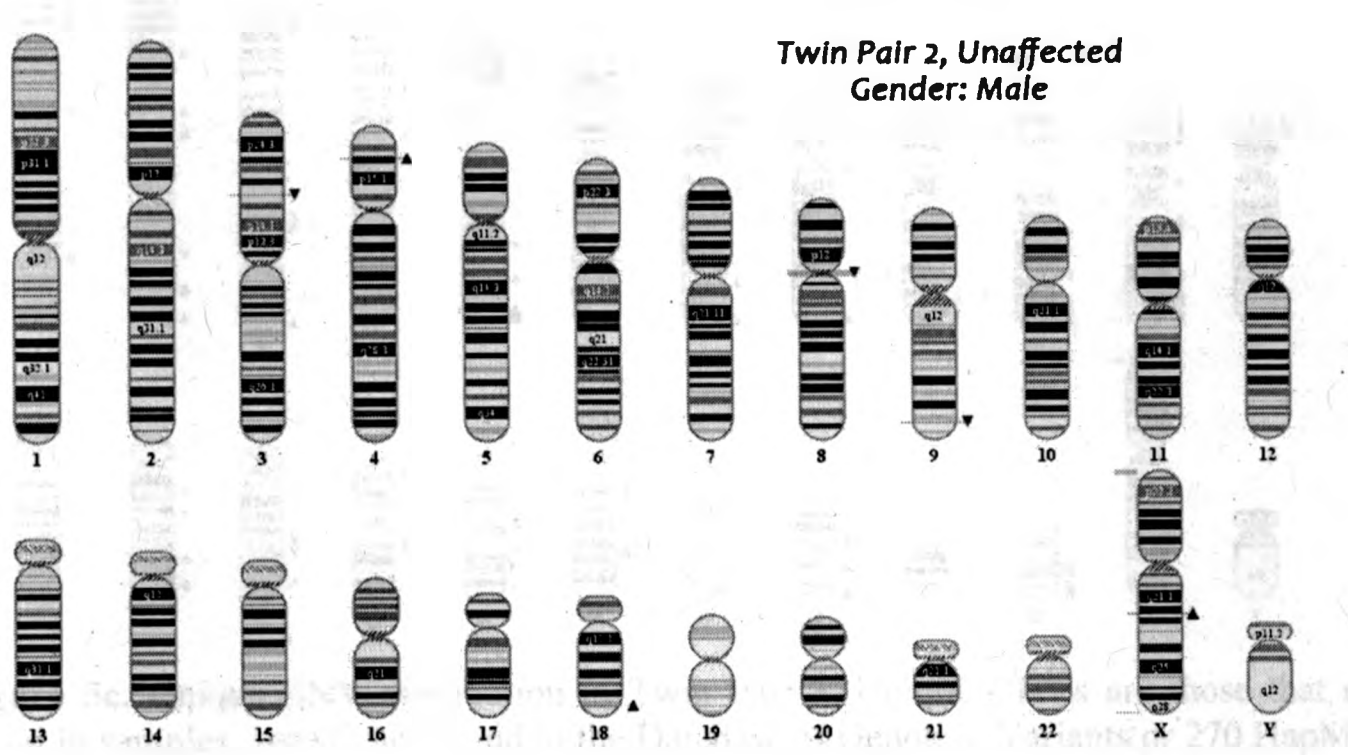


Figure 5b. Unique CNV distribution in Twin Pair 2. Unique CNVs are those that are found in samples, but are not found in the Database of Genomic Variants or 270 HapMap control samples. Triangles pointing down represent a loss and triangles pointing up represent a gain.

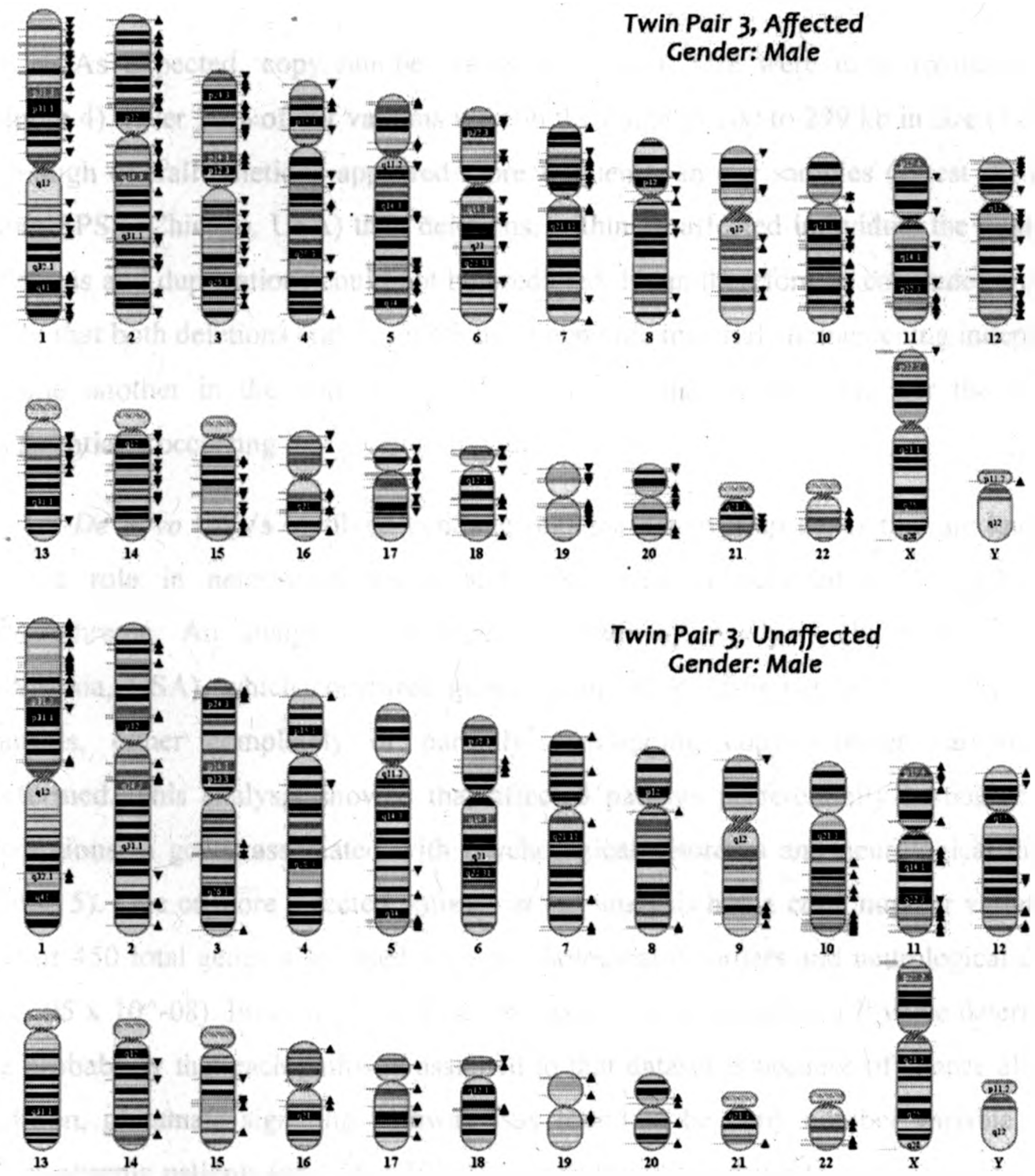


Figure 5c. Unique CNV distribution in Twin Pair 3. Unique CNVs are those that are found in samples, but are not found in the Database of Genomic Variants or 270 HapMap control samples. Triangles pointing down represent a loss and triangles pointing up represent a gain.

As expected, copy number variants of small size were most frequently seen (Figure 4). Over 50% of our variants were in the range of 100 to 299 kb in size (Table 3). Although overall deletions appeared more frequently in our samples (Z-test performed using SPSS, Chicago, USA) than deletions, within an affected individual the number of deletions and duplications could not be predicted. It can therefore be concluded from our study that both deletions and duplications of genomic material are happening independent of one another in the human genome. That is, that neither one nor the other is preferentially occurring in a given individual.

De novo CNVs involved genomic regions that overlap genes that are known to play a role in neurodevelopment and other systems believed to be affected in schizophrenia. An analysis with Ingenuity Pathway Analysis (Ingenuity Software, California, USA), which compared genes disrupted in cases versus those disrupted in controls, either completely or partially overlapping copy number variants, was performed. This analysis showed that affected patients preferentially harboured CNV disruptions in genes associated with psychological disorders and neurological diseases (Table 5). One or more affected patients in our analysis had a copy number variation in 110 of 450 total genes associated with psychological disorders and neurological disease ($p=3.05 \times 10^{-08}$). Ingenuity uses Fischer's exact test to calculate a *P* value determining the probability that each pathway assigned to that dataset is because of chance alone. In addition, glutamate signaling pathway was found to be copy number variable in our schizophrenic patients ($p=3.25 \times 10^{-03}$), a result that is similar to that found by Wilson *et al.* in 2006. Only genes included in Ingenuity's Pathway Knowledge Base were included in this analysis.

Table 5. Ingenuity Pathway Analysis (Version 7.5) list of functions and diseases highly associated with CNV within affected patients in this study.

Relevant Functions and Diseases	p-value	Number of Genes
Cardiovascular Disease	6.96 x 10^-13	88
Genetic Disorder	1.35 x 10^-12	175
Skeletal and Muscular Disorders	2.96 x 10^-08	96
Neurological Disorders	3.05 x 10^-08	110
Connective Tissue Disorders	6.52 x 10^-08	63
Immunological Disease	6.52 x 10^-08	82
Inflammatory Disease	6.52 x 10^-08	86
Metabolic Disease	1.25 x 10^-06	79
Nervous System Development and Function	5.89 x 10^-05	15
Psychological Disorders	2.49 x 10^-04	43

3.3 Characterization of Copy Number Variants

Unique CNVs were assessed to identify any shared CNV between unrelated, affected twins. This analysis identified five unique CNVs that were shared between the affected patient in three unrelated MZD twin pairs (Table 6). Of these five regions, only four contained known genes. Interestingly, genes contained in four of the five CNV regions are known to be involved in neurodevelopment, degeneration and neural functions and therefore have the potential to be pathogenic in schizophrenia. The probability (based on the number of known neurodevelopmental genes in the human genome) of finding four distinct neurodevelopmental genes disrupted in three unrelated individuals is 3.87 x 10^-11.

Thirteen copy number variations were found to be common between 2 of 3 unrelated, affected patients in this study (Table 7). These variants were not found in the Database of Genomic Variants, 270 HapMap samples, unaffected twins or experimental controls.

Figures 6 to 8 show copy number variant regions shared by all three unrelated, affected patients. These aberrations showed identical to near identical breakpoints between patients.

In addition to this, a CNV was found to be present in all three affected patients at 1q21.1, a region that was previously implicated by two independent schizophrenia studies by Stefansson *et al.* (2008) and The International Schizophrenia Consortium Group (2008). These variants were all in the same region, however, they did not overlap one another in unrelated, affected patients.

Table 6. Common gains and losses of the three affected patients representing the three unrelated pairs of MZ twins discordant for schizophrenia. Underlined genes were chosen for downstream confirmation by Real Time PCR based on function.

Location	Gain/Loss	Genes Present	Functions
1p36.22	Loss	*CLSTN1;* *CTNNBIP1;* *LZIC;* *NMNAT1.*	Calcium-mediated postsynaptic signals; Interacts with ß-catenin; Leucine Zipper; Involved in Ca(2+) signalling pathways.
4q22.3	Gain	*UNC5C*	Netrin receptors. Netrins are secreted proteins that direct axon extension and cell migration during neural development.
12q24.12	Loss	*ATXN2;* *BRAP.*	Mutations in *ATXN2* cause spinocerebellar ataxia type 2 (a neurodegenerative disorder); Cancer associated gene.
13q21.1 – 13q21.2	Gain	Does not overlap any known genes	N/A
17q21.1	Loss	*WIPF2;* *RAPGEFL1;* *CASC3;* *MSL-1;* *MED24;* *THRA;* *NR1D1.*	Actin formation associated protein; No confirmed function; Cancer susceptibility candidate; Acetylation; Gene expression mediator complex; Thyroid Hormone Receptor; Nuclear Receptor (found in brain among other locations)

Table 7. Common gains and losses of two affected patients from the three unrelated pairs of MZ twins discordant for schizophrenia.

Location	Gain/Loss	Patients from Twin Pairs	Genes Present
1p31.1	Gain	1, 2	*LPHN2*
3p11.2	Gain	1, 2	No Known Genes
3q13.31	Gain	1, 2	No Known Genes
4q13.1	Gain	1, 2	*LPHN3*
4q28.3	Gain	1, 2	No Known Genes
5q34	Gain	2, 3	No Known Genes
6q16.1	Gain	1, 2	No Known Genes
7q21.11	Gain	1, 2	No Known Genes
11q13.1	Loss	1, 2	*BBS1, DPP3, PELI3, MRPL11, NPAS4*
12q21.31	Gain	1, 2	No Known Genes
19p13.11	Loss	1, 3	*NWD1, TMEM38A, MED26, SLC35E1, CHERP, TPM4, RAB8A, HSH2D, CIB3, AP1M1, EPS15L7, CALR3*
19q13.32	Loss	2, 3	*SYMPK, RSHL1*
Xq21.31	Gain	1, 2	*PCDH11X*

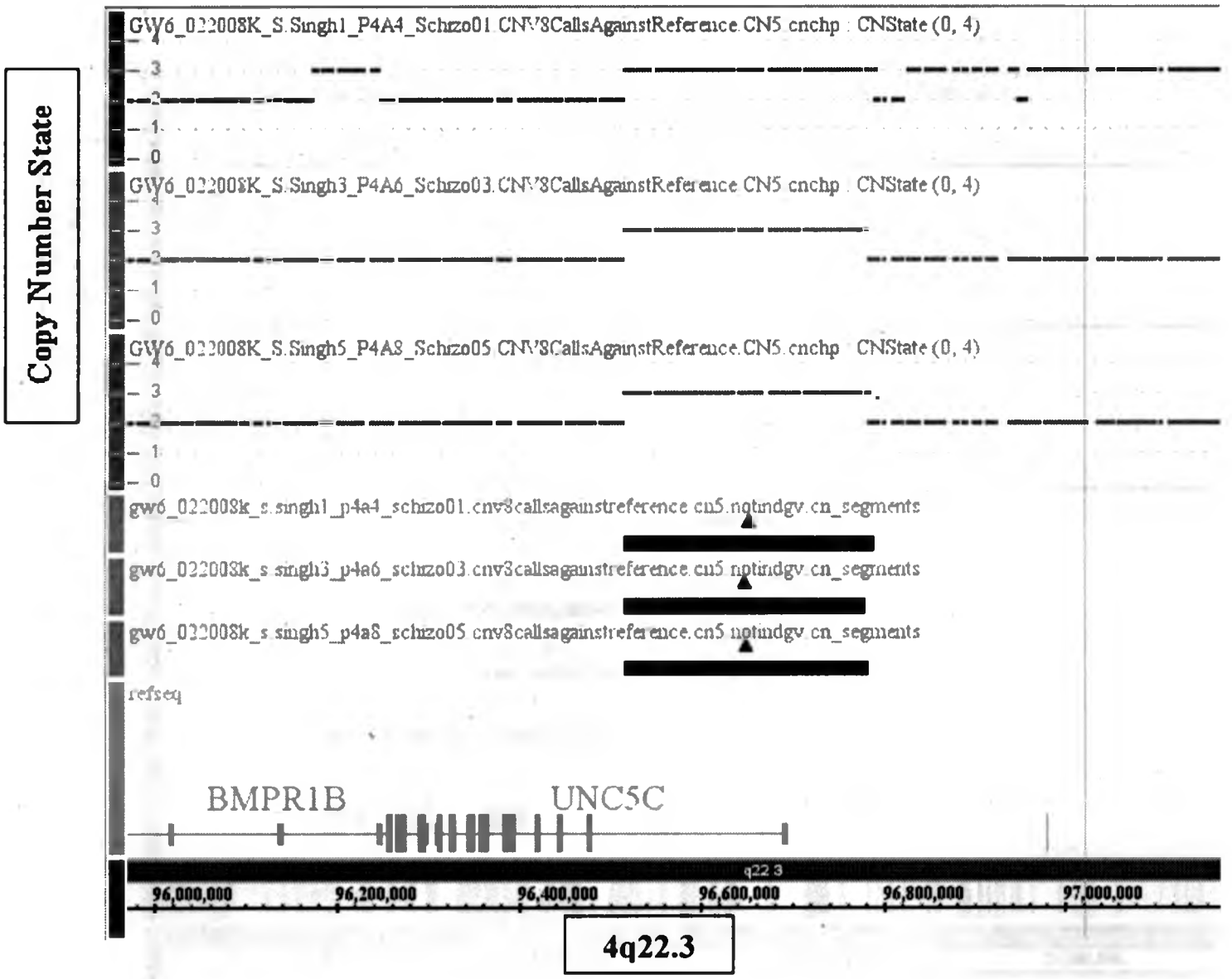


Figure 6. Shared copy number variable region at 4q22.3 overlapping the *UNC5C* gene. This variant is shared between all three affected, unrelated patients in the study.

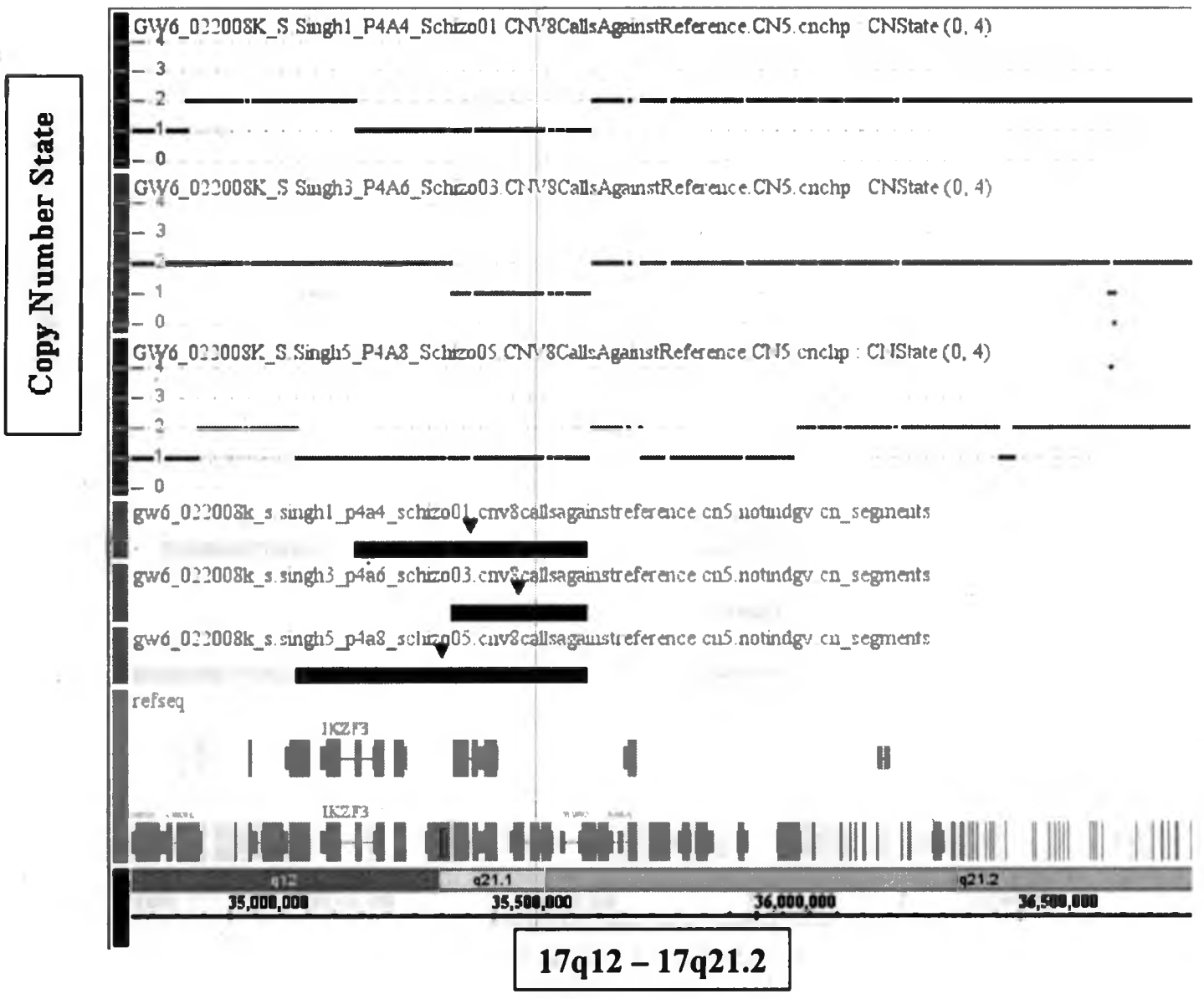


Figure 7. Shared copy number variable region at 17q21.1 overlapping 7 genes. This variant is shared between all three affected, unrelated patients in the study.

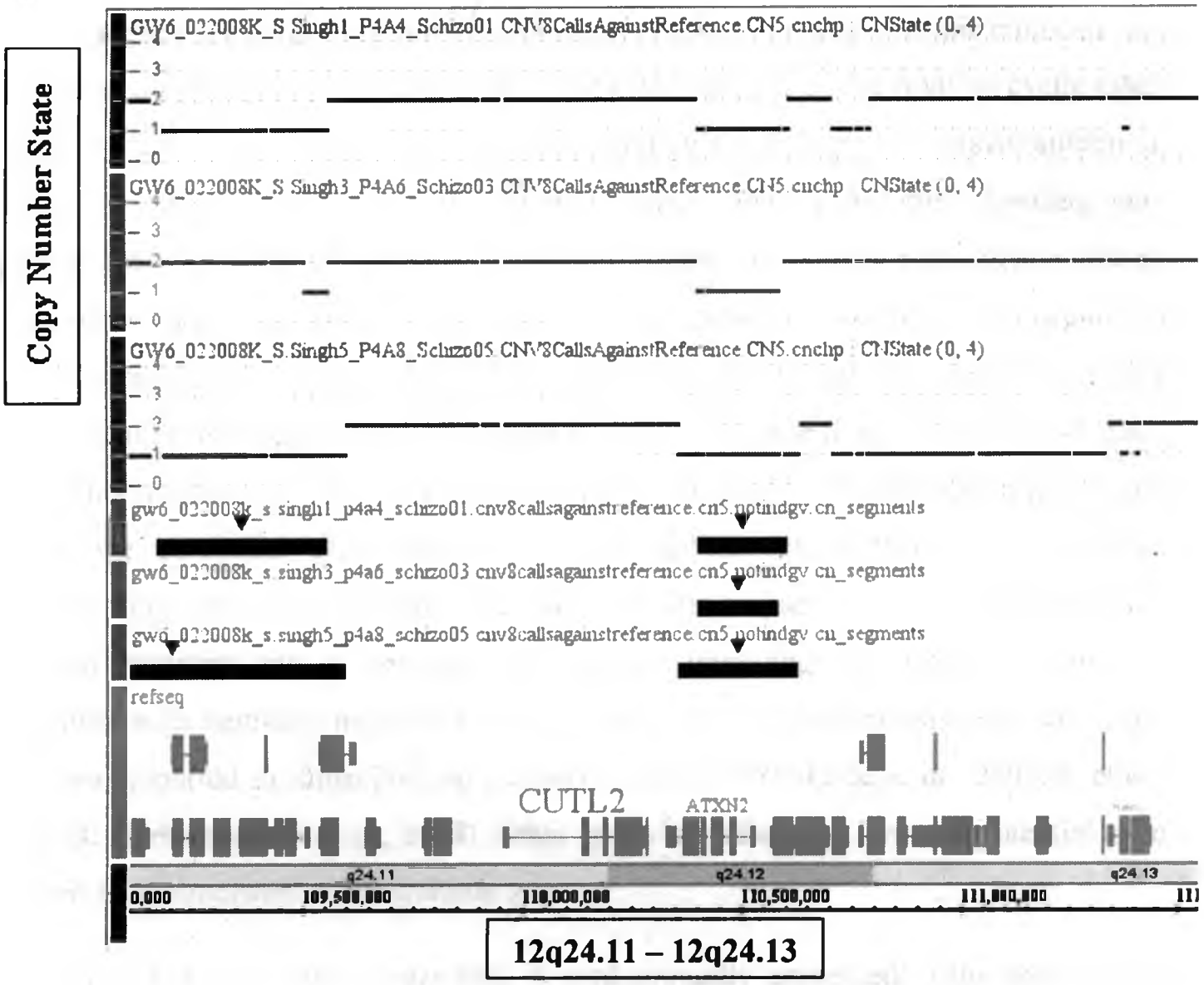


Figure 8. Shared copy number variable region at 12q24.12 overlapping *ATXN2* and *BRAP*. This variant is shared between all three affected, unrelated patients in the study.

The *UNC5C* gene product belongs to the UNC-5 family of netrin receptors. Netrins are a family of secreted molecules that have critical functions in axon guidance and cell migration during neuronal development. Netrin-1 is a chemotropic molecule and also acts as a survival factor (Tang *et al.*, 2008). The ratio of cyclic AMP to cyclic GMP activities sets the polarity of netrin-1-induced axon guidance: high ratios favor attraction, whereas low ratios favor repulsion. Further, cyclic AMP and GMP signaling may modulate calcium channel activity in growth cones, and signal transduction during bidirectional axon guidance (Nishiyama *et al.*, 2003), which is important in neurodevelopment. A critical role for this gene in neuronal migration during cerebellar development is also supported by studies on the *UNC5H3* gene in mice (Przyborski *et al.*, 1998). This transcript is expressed early (embryonic day 8.5) in the hindbrain region and later in the cerebellar primordia. Also, the establishment of the rostral cerebellar boundary may rely on chemorepulsive signaling that requires *UNC5H3* expressed by cerebellar neurons and a genomic abnormality involving the *UNC5C* would be compatible with neuronal migration abnormalities including cerebrovascular pathology, which was reported in schizophrenia (Akbarian *et al.*, 1993; Ende *et al.*, 2005; Kiehl *et al.*, 2008; Kyriakopoulos *et al.*, 2008). Other genes identified that have the potential to be involved in schizophrenia pathogenesis include:

(1) *CLSTN1* (Loss), a gene that is predominantly expressed in the brain and is involved in learning and memory (Cheng *et al.*, 2006), (2) *NMNAT1* (Loss), which is involved in calcium signaling and prevents axonal degeneration caused by stress and reduces reactive oxygen species levels (Press *et al.*, 2008), (3) *ATXN2* (Loss), which causes neurodegeneration, leading to a form of spinocerebellar ataxia (Al-Ramahi *et al.*, 2007), (4) *WIPF2* (Loss), which plays a role in organization of the actin cytoskeleton, (5) *NR1D1* (Loss), a transcriptional repressor involved in development and regulation in many tissues, including brain, and (5) *MED24* (Loss), which encodes a component of a transcriptional coactivator complex that is thought to be required for gene expression.

Another notable gene was Methylenetetrahydrofolate (*MTHFR*) on chromosome 1, listed as one of the top 20 schizophrenia candidate genes in the SchizoGene Database (Accessed June 2009.), *MTHFR* was found to have regions of the gene deleted in the affected patients of twin pairs 2 and 3 in this analysis.

Interestingly, all samples, except the unaffected twins in Twin Pair 1 and Twin Pair 3, showed a gain at Xq21.31 with near identical breakpoints. It is possible that this region could lead to schizophrenia susceptibility or predisposition in both patients and unaffected controls.

3.4 Real Time Analysis

One gene from each shared copy number variant between affected, unrelated patients was chosen for confirmation by Real Time Analysis of genomic DNA. The Real Time results are summarized in Figure 11 and Table 8. Genes were chosen based on known function. Genes whose functions were most likely to be involved in the causation of schizophrenia were chosen from each region. An example of Real Time delta delta C_T loss and gain calculations are shown in Figures 9 and 10, respectively.

The following genes were chosen for confirmation, included beside each gene is the region of that gene that was found to be copy number variable:

CLSTN1: 9778163 → 9807137

ATXN2: 110521863 → 110400664

UNC5C: 96689185 → 96515818

NR1D1: Entire Gene was found in copy number variable region

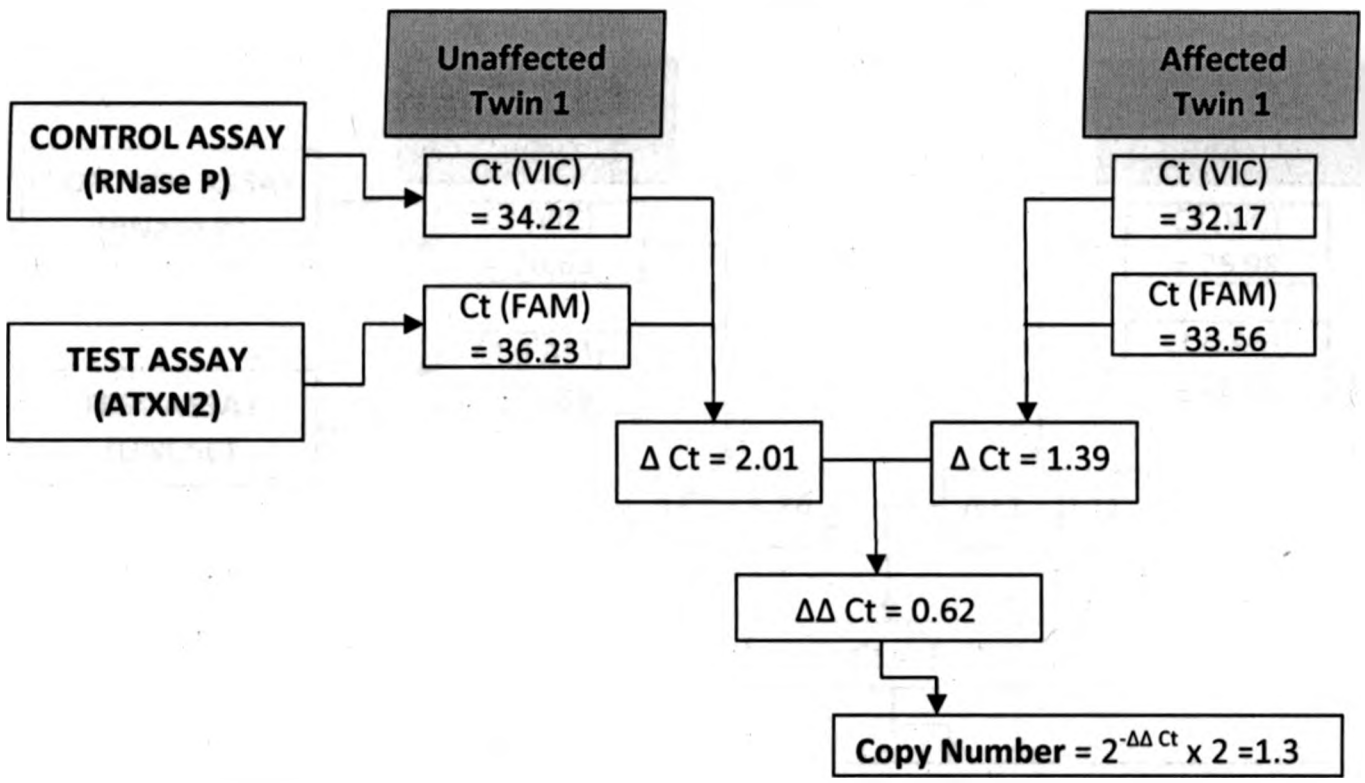


Figure 9. An example of Copy Number loss calculation using the delta delta Ct method. Calculation shown here is for the ATXN2 gene of the affected patient of Twin Pair 1.

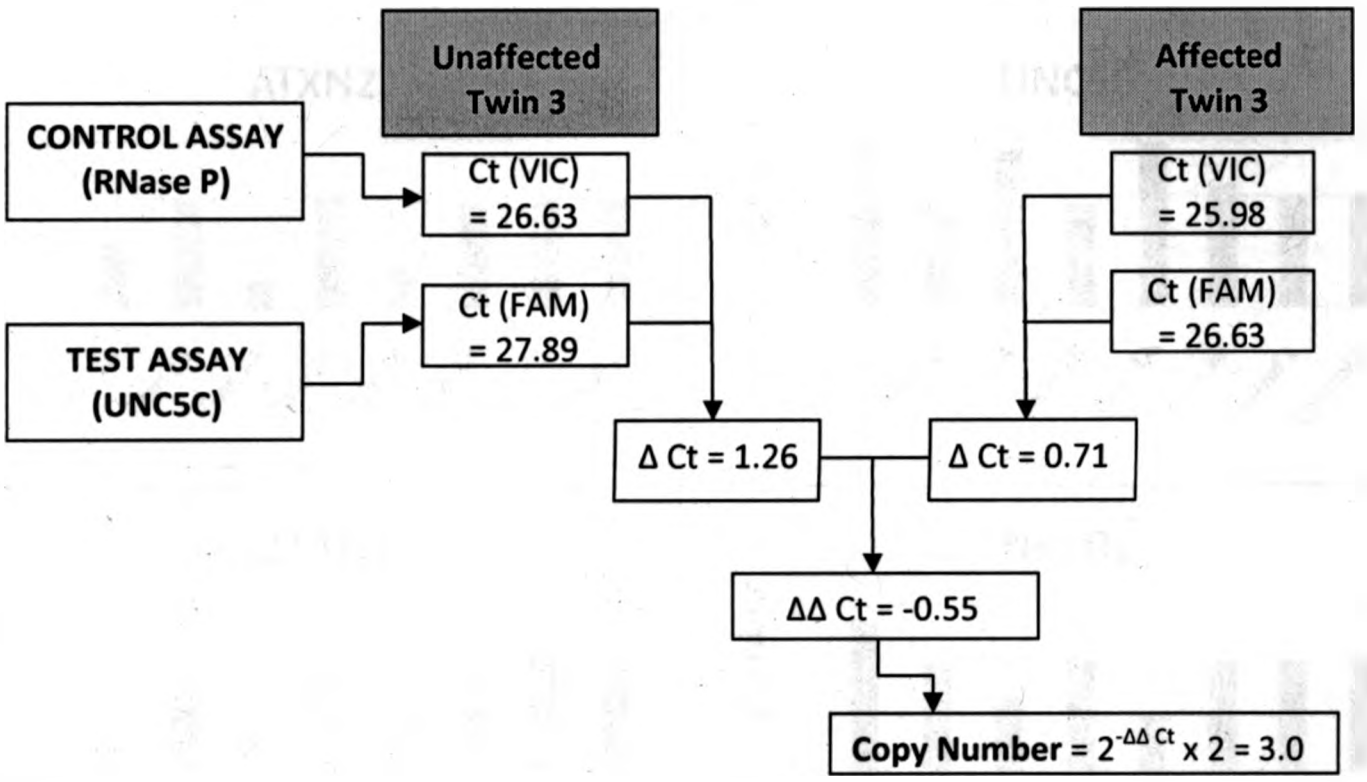


Figure 10. An example of Copy Number gain calculation using the delta delta Ct method. Calculation shown here is for the UNC5C gene of the affected patient of Twin Pair 3.

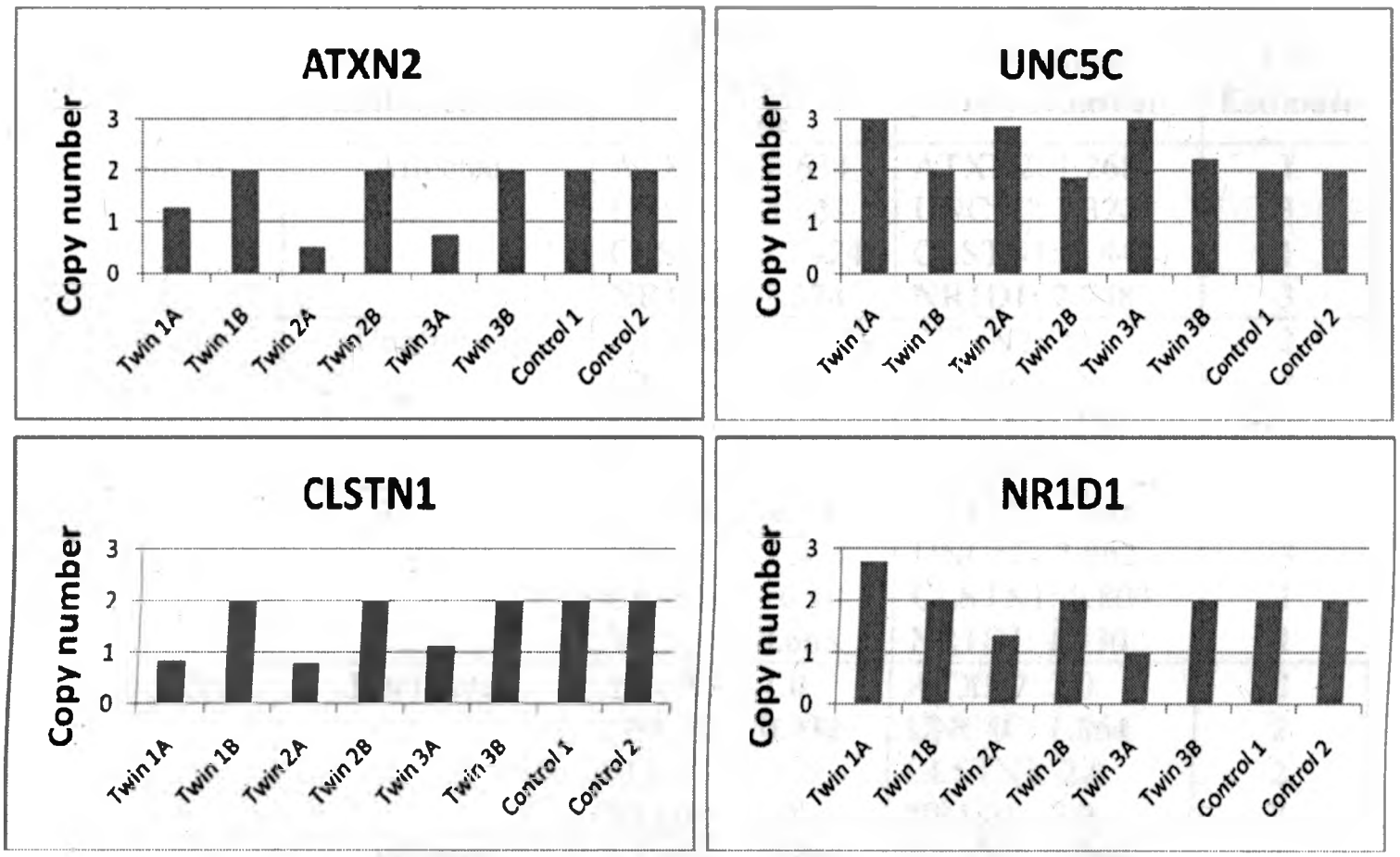


Figure 11. Real Time copy number quantification of candidate genes using Applied Biosystem's Custom Real Time Copy Number Assays. In each case A represents the affected twin and B represents the unaffected twin.

Table 8. Relative Quantity Values are shown from Real Time Analysis with Applied Biosystem's Step One Instrument. Copy Number is double the RQ value. Bolded RQ value shows an unexpected result for Twin Pair 1, affected.

Sample	Affected or Unaffected Twin	Relative Quantity (RQ) Value	Absolute Copy Number	CN Estimate
Twin 1a	Affected	ATXN2 – 0.634 UNC5C – 1.661 CLSTN1 – 0.424 **NR1D1 – 1.374**	ATXN2: 1.268 UNC5C: 3.322 CLSTN1: 0.848 **NR1D1: 2.748**	1 3 1 3
Twin 1b	Unaffected	ATXN2 – 1.0 UNC5C – 1.0 CLSTN1 – 1.0 NR1D1 – 1.0	ATXN2: 2.0 UNC5C: 2.0 CLSTN1: 2.0 NR1D1: 2.0	2 2 2 2
Twin 2a	Affected	ATXN2 – 0.197 UNC5C – 1.431 CLSTN1 – 0.40 NR1D1 – 0.665	ATXN2: 0.394 UNC5C: 2.862 CLSTN1: 0.800 NR1D1: 1.330	0 3 1 1
Twin 2b	Unaffected	ATXN2 – 1.0 UNC5C – 0.932 CLSTN1 – 1.0 NR1D1 – 1.0	ATXN2: 2.0 UNC5C: 1.864 CLSTN1: 2.0 NR1D1: 2.0	2 2 2 2
Twin 3a	Affected	ATXN2 – 0.384 UNC5C – 1.576 CLSTN1 – 0.653 NR1D1 – 0.506	ATXN2: 0.768 UNC5C: 3.152 CLSTN1: 1.306 NR1D1: 1.012	1 3 1 1
Twin 3b	Unaffected	ATXN2 – 1.0 UNC5C – 1.109 CLSTN1 – 1.0 NR1D1 – 1.0	ATXN2: 2.0 UNC5C: 2.218 CLSTN1: 2.0 NR1D1: 2.0	2 2 2 2
Del22q11 Experimental Control 1	Unaffected	ATXN2 – 1.0 UNC5C – 1.0 CLSTN1 – 1.0 NR1D1 – 1.0	ATXN2: 2.0 UNC5C: 2.0 CLSTN1: 2.0 NR1D1: 2.0	2 2 2 2
Del22q11 Experimental Control 2	Unaffected	ATXN2 – 1.0 UNC5C – 1.0 CLSTN1 – 1.0 NR1D1 – 1.0	ATXN2: 2.0 UNC5C: 2.0 CLSTN1: 2.0 NR1D1: 2.0	2 2 2 2

CHAPTER 4: DISCUSSION

4.1 Summary of Conclusions

This study involved the use of monozygotic twins discordant for schizophrenia for the discovery of copy number variable regions in schizophrenic patients. The use of monozygotic twins considerably reduces the heterogeneity that exists in the disorder. The existence of copy number variation as widespread phenomena in the human genome has been the focus of many researchers in the past five years. The results generated in this study have focused on unique copy number variations, that is, variation that is not found in the database of genomic variants, 270 HapMap samples or shared between co-twins. The results generated through this project are discussed below.

The following main conclusions can be drawn from the research in this study:

i. Monozygotic twins discordant for schizophrenia have different copy number variation profiles.
ii. There is a significant increase in copy number variation present in affected twins versus their unaffected co twins and control samples. Affected twins show a minimum 4x increase in the number of unique CNVs when compared to their unaffected co-twin.
iii. The results of this study suggest that multiple and rare individually specific CNVs with major effects on genes from neurodevelopmental pathways are involved in the etiology of schizophrenia.
iv. A number of common copy number variable regions were found to share near identical breakpoints among unrelated, affected patients.

The implications of each conclusion will be presented below.

4.1.1 Monozygotic twins discordant for schizophrenia have different copy number variation profiles.

At the outset of this study, the copy number variation profiles of monozygotic twins were not known. However, a recent study by Bruder *et al.* (2008) found that CNV differences between co-twins are common in monozygotic twin pairs. This study also challenges the assumed genetic equivalence of MZ twins including interpretation of the extensive body of published results on twins. I suggest that the term "identical twin," in a genetic context is inappropriate. Although the twins in this study were monozygotic and in theory should have identical genomes, they differed considerably in terms of copy number variations. This result will make the task of separating the genetic and environmental factors contributing to complex disease considerably more difficult. In addition, knowledge of the genetic differences seen in monozygotic twins with regard to copy number variation detection allows for a whole new area of genetic research as we untangle the cause, development and novelty of CNVs in monozygotic twins.

4.1.2 There is a significant increase in copy number variation present in affected twins versus unaffected co-twins and control samples.

The increased CNVs in the affected as compared to unaffected member of each twin pair implicates large numbers of copy number variants in the genetics of schizophrenia. These results specifically support Walsh *et al.* (2008) and Xu *et al.* (2008), who reported a significantly higher number of CNVs in patients with schizophrenia as compared to unaffected controls (Walsh *et al.*, 2008; Xu *et al.*, 2008). Walsh's group searched for structural variants greater than 100kb in DNA from 150 patients and 268 controls and showed that the frequency of common copy number variants was not significantly different (Walsh *et al.*, 2008). In this study, novel copy number variable regions that altered known genes occurred four times more frequently in cases (onset before age 18) versus controls (Walsh *et al.*, 2008). Similarly, Xu *et al.* (2008) carried out a study in 1077 triads (probands and biological parents) as well as 152 sporadic cases (Xu

et al., 2008). *De novo* copy number changes were seen in 10% of sporadic cases of schizophrenia and in 1.5% of control samples (Xu *et al.*, 2008).

Recently, three independent groups found that many alleles of small effect size are likely contributing to the etiology of schizophrenia, a finding that is consistent with the results found in this study (Shi *et al.*, 2009; Stefansson *et al.*, 2009; The International Schizophrenia Consortium *et al.*, 2009). In the case of this analysis, the matched controls are unaffected MZ co-twins making this observation much more causative rather than statistical. In addition, our 22q11 deletion syndrome experimental control patients did not show an increase in copy number variation, a finding consistent with the results recently found by *Bassett et al* (Bassett *et al.*, 2008).

From this, it can be concluded that the genomes of affected twins have extensive copy number variation when compared to the twin's unaffected co-twin. It is possible that the extensive repeats present in the human genome provide the groundwork for DNA replication errors that have the potential to occur at each cycle of meiosis and mitosis. Although the exact mechanism is not clearly understood at this time, copy number variation showers provide a likely contributor to the pathology of schizophrenia (Singh *et al.*, 2009). The extensive showering of the human genome with structural variants could be the result of a predisposition for DNA replication errors that are present in affected twins and not seen in unaffected twins. It is also probable to assume that a two-hit hypothesis can be applied to discordant monozygotic twins in the study of schizophrenia genetics. For example, a genetic predisposition for schizophrenia could be present in both twins and yet only one of the twins might receive either a second genetic hit (perhaps a *de novo* copy number variation) or an environmental hit that leads to the full development of the disease. Following a similar logic, Guidry and Kent hypothesized that if the inheritance of a single recessive mutated allele of a gene crucial in brain development is followed by a somatic mutation in the normal allele during critical periods of brain development this could result is developmental abnormalities that are expressed behaviourally as schizophrenia (Guidry *et al.*, 1999). The somatic mutation may not always occur, possibly explaining the situation seen in discordant monozygotic twins.

In addition, the observed increase in *de novo* CNVs in schizophrenia could potentially explain the ~1% life time risk of the disease in the general population (Odegaard, 1972).

4.1.3 The results of this study suggest that multiple and rare individually specific CNVs with major effects on genes from neurodevelopmental pathways are involved in the etiology of schizophrenia.

In addition, this analysis suggests that different patients may have different subsets of a larger set of schizophrenia related genes affected by copy number variation. This finding may also explain the extensive heterogeneity of schizophrenia in the population and the variable manifestation of this disease across patients. Therefore, the genetic basis for schizophrenia may be a succession of multiple, *de novo* mutational events, with these events favouring the neurodevelopmental and neurofunctional pathways known to be affected in schizophrenia. This finding agrees with Walsh and colleagues who found that mutations in affected patients disproportionally affected genes from networks controlling neurodevelopment (Walsh *et al.*, 2008).

4.1.4 A number of common copy number variable regions were found to share near identical breakpoints among unrelated, affected patients.

A common set of genes were found to be copy number variable across unrelated, affected patients in our study. The probability, based on known neurodevelopmental genes in the human genome, of finding a shared copy number variant affecting a neurodevelopmental gene in three unrelated schizophrenic patients is 0.078%. The existence of multiple shared regions therefore represents a statistically rare event that may play a causative role in schizophrenia. Shared copy number variable regions between unrelated schizophrenic patients could implicate shared sequence similarity in the identified regions leading to copy number variation hotspots whose aberration cause the pathological symptoms of the disease. In addition, these results suggest that a common set of rare variants may work together to contribute to schizophrenia.

Interestingly, the shared regions were not seen in unaffected twins or control samples. This suggests that copy number variants themselves likely play more of a causative role rather than serve to cause a predisposition to the disorder.

4.2 Caveats

4.2.1 Are the observed variants an effect of illness or treatment?

The results presented here are likely not an incidental finding, as it cannot explain the increased number of CNVs seen in each of the three affected MZD twins or in individuals with schizophrenia when compared with familial and unrelated controls (Walsh *et al.*, 2008; Xu *et al.*, 2008). A caveat of this research, however, is the fact that most schizophrenic patients have been treated with anti-psychotic medication by the time they are included in a study of this kind. It is possible that results of schizophrenia experiments are due in part to an effect of the medication being used to treat the disorder. However, brain morphology studies have shown that the differences observed in schizophrenic patients versus controls are seen both at first episode and after the use of anti-psychotics, indicating that anti-psychotic medications do not affect the primary causes of schizophrenia (Vita *et al.*, 2006). In this study, all affected twins, but not the unaffected co-twin, had significant lifetime exposure to anti-psychotic medication. While not explicitly stated, patients studied by Walsh *et al.* (2008) and Xu *et al.* (2008) likely received anti-psychotic medication for their illness. However, the possibility that the observed increase in CNVs is an effect of medication is not compatible with several facts. First, the novel CNVs detected in patients are highly likely to affect genes and/or pathways implicated and affected in schizophrenia. Secondly, CNVs are more frequent in sporadic as compared to familial schizophrenia, but the two types of schizophrenia are treated the same way and the lifestyle associated with each type of schizophrenia does not differ (Xu *et al.*, 2008).

The possibility that the increase in CNVs are an effect of illness is difficult to dismiss, as individuals with schizophrenia are more likely to have poor nutrition and to smoke than individuals free from the illness, which could both lead to increases in CNVs.

4.2.2 The use of blood to study a brain-based disorder

One current limitation of copy number variation analyses is that the genotype of one tissue analyzed, in this case blood, might not always be relevant to the genotype of the tissue responsible for the disease. For this reason, our lab is currently working on a parallel study that will be testing copy number variation profiles in multiple tissues taken from the same individual (whole blood and buccal smear samples).

4.3 Future Studies

Follow up studies to the one presented here should first assess any involvement of confirmed copy number variants in the causation of schizophrenia based on their distribution in discordant and normal twin pairs. Further, characterization of the identified genes of interest in sib pairs, trios and population samples should be performed to assess the frequency of copy number variations overlapping the candidate genes presented in this study. The aforementioned study was a discovery study, but it should not stand alone and it should be followed up with an association study in order to assess the correlation between genotype and phenotype. In addition to this, given how little is currently known in the field with regard to the effect of copy number variation on gene expression it would be important to identify the expression of the genes of interest in schizophrenic patients. Further to this, the copy number variation profile seen in various tissues of the same individual is not currently known. To this end, another member of the Singh lab is currently assessing the copy number variation profiles in varying sample types from the same patient.

4.4 Final Conclusion

These results argue that the human genome is not static. We do not simply acquire parental genomes and pass them on to the next generation; rather, the genome undergoes a variety of alterations at every DNA replication cycle. One of these alterations may result from the semi-conservative replication of the DNA, particularly in repetitive

regions, resulting in gross numbers of copy number variants scattered across the genome. It is likely that copy number variants are a way for our genomes to remain both fluid and malleable, however, errors in the mechanisms leading to this evolution are likely to be detrimental. It is known that there is a strong genetic aspect to this complex, multifactorial disease with multiple genes interacting, thus traditional approaches have had little success in identifying the cause of schizophrenia and alternative methods, like the ones undertaken in this study, need be used. In principle, complex diseases like schizophrenia may be more likely to be affected by "soft" forms of variation, like copy number changes that can alter gene expression without completely eliminating gene function (McCarroll *et al.*, 2007).

This study adds to the growing body of knowledge in the field of schizophrenia genetics, a field that has posed a great challenge to researchers for many years. This analysis has identified a set of genes whose simultaneous aberration through both deletions and duplications, may have been needed for the development of schizophrenia in the three patients studied. These genes may belong to a larger group of genes that are affected in schizophrenia and related disorders. In addition, we offer evidence supporting the examination of CNVs in monozygotic twins as an effective means to identify genes involved in complex disorders

Furthering our knowledge of schizophrenia will help to lead to better diagnosis, treatment and understanding of this complex disorder. Studies of this kind may contribute to earlier diagnosis of the disease, which could assist in disease management. In addition, further characterization of this disorder may help to better identify subgroups of schizophrenic patients, which will assist in the classification, treatment and scientific study of the disorder. It is possible that in the future, CNV analysis could be considered routine in clinical testing for schizophrenia.

In conclusion, the following significant findings were found in this study:

- Monozygotic discordant twins showed differences in copy number variation profiles despite their genetic similarity.

- Significant increases in the number of copy number variants seen in affected versus unaffected co-twins were seen in all three twin pairs assessed. These increases argue for a hypothesis that includes the showering of the human genome with DNA replication errors leading to copy number variants.

- Affected schizophrenia patients in this study harboured copy number variants that preferentially overlapped known neurodevelopmental genes, a finding supported by the literature.

- Five copy number variable regions at 1p36.22, 4q22.3, 12q24.12, 13q21.1 and 17q21.1 were shared between unrelated twins affected with schizophrenia. With the exception of the copy number variant at 13q21.1, which did not overlap any genes, the other four regions overlapped at least one neurodevelopmental gene.

- The results presented here implicate copy number variation showers causing multiple rare variants of small effect, which cause aberrant expression of neurodevelopmental genes leading to schizophrenia.

REFERENCES

Akbarian, S., Bunney, W.E.,Jr, Potkin, S.G., Wigal, S.B., Hagman, J.O., Sandman, C.A., and Jones, E.G. (1993). Altered distribution of nicotinamide-adenine dinucleotide phosphate-diaphorase cells in frontal lobe of schizophrenics implies disturbances of cortical development. Arch. Gen. Psychiatry *50,* 169-177.

Al-Ramahi, I., Perez, A.M., Lim, J., Zhang, M., Sorensen, R., de Haro, M., Branco, J., Pulst, S.M., Zoghbi, H.Y., and Botas, J. (2007). dAtaxin-2 mediates expanded Ataxin-1-induced neurodegeneration in a Drosophila model of SCA1. PLoS Genet. *3,* e234.

Andreasen, N.C. (1995). Symptoms, signs, and diagnosis of schizophrenia. Lancet *346,* 477-481.

Arlt, M.F., Mulle, J.G., Schaibley, V.M., Ragland, R.L., Durkin, S.G., Warren, S.T., and Glover, T.W. (2009). Replication stress induces genome-wide copy number changes in human cells that resemble polymorphic and pathogenic variants. Am. J. Hum. Genet. *84,* 339-350.

Bassett, A.S., Marshall, C.R., Lionel, A.C., Chow, E.W., and Scherer, S.W. (2008). Copy number variations and risk for schizophrenia in 22q11.2 deletion syndrome. Hum. Mol. Genet. *17,* 4045-4053.

Belmaker, R., Pollin, W., Wyatt, R.J., and Cohen, S. (1974). A follow-up of monozygotic twins discordant for schizophrenia. Arch. Gen. Psychiatry *30,* 219-222.

Bruder, C., and Piotrowski, A.e.a. (2008). Phenotypically Concordant and Discordant Monozygotic Twins Display Different DNA Copy Number Variation Profiles. Am. J. Hum. Genet. *82,* 1-9.

Cheng, X.R., Zhou, W.X., and Zhang, Y.X. (2006). The family of calsyntenins: learning and memory related genes. Sheng Li Ke Xue Jin Zhan *37,* 205-210.

Conrad, D.F., and Hurles, M.E. (2007). The population genetics of structural variation. Nat. Genet. *39,* S30-6.

de Smith, A.J., Tsalenko, A., Sampas, N., Scheffer, A., Yamada, N.A., Tsang, P., Ben-Dor, A., Yakhini, Z., Ellis, R.J., Bruhn, L. *et al.* (2007). Array CGH analysis of copy number variation identifies 1284 new genes variant in healthy white males: implications for association studies of complex diseases. Hum. Mol. Genet. *16,* 2783-2794.

Delisi, L.E. (2009). Searching for the true genetic vulnerability for schizophrenia. Genome Med. *1,* 14.

DeLisi, L.E. (2000). Critical overview of current approaches to genetic mechanisms in schizophrenia research. Brain Res. Brain Res. Rev. *31,* 187-192.

Ende, N., Ende, M., Chen, R., Coakley, K., and Reddi, A.S. (2005). Prevention of atherosclerosis in LDL receptor-mutant mice by human umbilical cord blood cells. Res. Commun. Mol. Pathol. Pharmacol. *117-118,* 125-136.

Feuk, L., Marshall, C.R., Wintle, R.F., and Scherer, S.W. (2006). Structural variants: changing the landscape of chromosomes and design of disease studies. Hum. Mol. Genet. *15 Spec No 1,* R57-66.

Fraga, M.F., Ballestar, E., Paz, M.F., Ropero, S., Setien, F., Ballestar, M.L., Heine-Suner, D., Cigudosa, J.C., Urioste, M., Benitez, J. *et al.* (2005). Epigenetic differences arise during the lifetime of monozygotic twins. Proc. Natl. Acad. Sci. U. S. A. *102,* 10604-10609.

Freedman, R., Leonard, S., Olincy, A., Kaufmann, C.A., Malaspina, D., Cloninger, C.R., Svrakic, D., Faraone, S.V., and Tsuang, M.T. (2001). Evidence for the multigenic inheritance of schizophrenia. Am. J. Med. Genet. *105,* 794-800.

Freeman, J.L., Perry, G.H., Feuk, L., Redon, R., McCarroll, S.A., Altshuler, D.M., Aburatani, H., Jones, K.W., Tyler-Smith, C., Hurles, M.E. *et al.* (2006). Copy number variation: new insights in genome diversity. Genome Res. *16,* 949-961.

Gershon, E.S. (2000). Bipolar illness and schizophrenia as oligogenic diseases: implications for the future. Biol. Psychiatry *47,* 240-244.

Goeree, R., Farahati, F., Burke, N., Blackhouse, G., O'Reilly, D., Pyne, J., and Tarride, J.E. (2005). The economic burden of schizophrenia in Canada in 2004. Curr. Med. Res. Opin. *21,* 2017-2028.

Gottesman, I.I. (1991). Schizophrenia genesis: the origin of madness.(New York.: W.H. Freeman and Co.).

Gottesman, I.I., and Shields, J. (1967). A polygenic theory of schizophrenia. Proc. Natl. Acad. Sci. U. S. A. *58,* 199-205.

Graubert, T.A., Cahan, P., Edwin, D., Selzer, R.R., Richmond, T.A., Eis, P.S., Shannon, W.D., Li, X., McLeod, H.L., Cheverud, J.M., and Ley, T.J. (2007). A high-resolution map of segmental DNA copy number variation in the mouse genome. PLoS Genet. *3,* e3.

Guidry, J., and Kent, T.A. (1999). New genetic hypothesis of schizophrenia. Med. Hypotheses *52,* 69-75.

Henrichsen, C.N., Vinckenbosch, N., Zollner, S., Chaignat, E., Pradervand, S., Schutz, F., Ruedi, M., Kaessmann, H., and Reymond, A. (2009). Segmental copy number variation shapes tissue transcriptomes. Nat. Genet. *41,* 424-429.

Huxley, J., Mayr, E., Osmond, H., and Hoffer, A. (1964). Schizophrenia as a Genetic Morphism. Nature *204,* 220-221.

Iafrate, A.J., Feuk, L., Rivera, M.N., Listewnik, M.L., Donahoe, P.K., Qi, Y., Scherer, S.W., and Lee, C. (2004). Detection of large-scale variation in the human genome. Nat. Genet. *36,* 949-951.

Jablensky, A., McGrath, J., Herrman, H., Castle, D., Gureje, O., Evans, M., Carr, V., Morgan, V., Korten, A., and Harvey, C. (2000). Psychotic disorders in urban areas: an overview of the Study on Low Prevalence Disorders. Aust. N. Z. J. Psychiatry *34,* 221-236.

Keshavan, M.S., Prasad, K.M., and Pearlson, G. (2007). Are brain structural abnormalities useful as endophenotypes in schizophrenia? Int. Rev. Psychiatry. *19,* 397-406.

Kiehl, T.R., Chow, E.W., Mikulis, D.J., George, S.R., and Bassett, A.S. (2008). Neuropathologic Features in Adults with 22q11.2 Deletion Syndrome. Cereb. Cortex

Kyriakopoulos, M., Vyas, N.S., Barker, G.J., Chitnis, X.A., and Frangou, S. (2008). A diffusion tensor imaging study of white matter in early-onset schizophrenia. Biol. Psychiatry *63,* 519-523.

La Starza, R., Crescenzi, B., Pierini, V., Romoli, S., Gorello, P., Brandimarte, L., Matteucci, C., Kropp, M.G., Barba, G., Martelli, M.F., and Mecucci, C. (2007). A common 93-kb duplicated DNA sequence at 1q21.2 in acute lymphoblastic leukemia and Burkitt lymphoma. Cancer Genet. Cytogenet. *175,* 73-76.

Lee, J.A., and Lupski, J.R. (2006). Genomic rearrangements and gene copy-number alterations as a cause of nervous system disorders. Neuron *52,* 103-121.

Lupski, J.R. (2007). Genomic rearrangements and sporadic disease. Nat. Genet. *39,* S43-7.

Marshall, C.R., Noor, A., Vincent, J.B., Lionel, A.C., Feuk, L., Skaug, J., Shago, M., Moessner, R., Pinto, D., Ren, Y. *et al.* (2008). Structural variation of chromosomes in autism spectrum disorder. Am. J. Hum. Genet. *82,* 477-488.

McCarroll, S.A., and Altshuler, D.M. (2007). Copy-number variation and association studies of human disease. Nat. Genet. *39,* S37-42.

McGuffin, P., Asherson, P., Owen, M., and Farmer, A. (1994). The strength of the genetic effect. Is there room for an environmental influence in the aetiology of schizophrenia? Br. J. Psychiatry *164,* 593-599.

McInnes, L.A., and Lauriat, T.L. (2006). RNA metabolism and dysmyelination in schizophrenia. Neurosci. Biobehav. Rev. *30,* 551-561.

Mednick, S.A., Machon, R.A., Huttunen, M.O., and Bonett, D. (1988). Adult schizophrenia following prenatal exposure to an influenza epidemic. Arch. Gen. Psychiatry *45,* 189-192.

Mirnics, K., Levitt, P., and Lewis, D.A. (2006). Critical appraisal of DNA microarrays in psychiatric genomics. Biol. Psychiatry *60,* 163-176.

Nishiyama, M., Hoshino, A., Tsai, L., Henley, J.R., Goshima, Y., Tessier-Lavigne, M., Poo, M.M., and Hong, K. (2003). Cyclic AMP/GMP-dependent modulation of Ca2+ channels sets the polarity of nerve growth-cone turning. Nature *423,* 990-995.

Odegaard, O. (1972). The multifactorial theory of inheritance in predisposition to schizophrenia. Kaplan AR (Ed): Genetic Factors in Schizophrenia. 256-257-275.

Perry, G.H., Dominy, N.J., Claw, K.G., Lee, A.S., Fiegler, H., Redon, R., Werner, J., Villanea, F.A., Mountain, J.L., Misra, R. *et al.* (2007). Diet and the evolution of human amylase gene copy number variation. Nat. Genet. *39,* 1256-1260.

Pompili, M., Amador, X.F., Girardi, P., Harkavy-Friedman, J., Harrow, M., Kaplan, K., Krausz, M., Lester, D., Meltzer, H.Y., Modestin, J. *et al.* (2007). Suicide risk in schizophrenia: learning from the past to change the future. Ann. Gen. Psychiatry. *6,* 10.

Press, C., and Milbrandt, J. (2008). Nmnat delays axonal degeneration caused by mitochondrial and oxidative stress. J. Neurosci. *28,* 4861-4871.

Przyborski, S.A., Knowles, B.B., and Ackerman, S.L. (1998). Embryonic phenotype of Unc5h3 mutant mice suggests chemorepulsion during the formation of the rostral cerebellar boundary. Development *125,* 41-50.

Redon, R., Ishikawa, S., Fitch, K.R., Feuk, L., Perry, G.H., Andrews, T.D., Fiegler, H., Shapero, M.H., Carson, A.R., Chen, W. *et al.* (2006). Global variation in copy number in the human genome. Nature *444,* 444-454.

Scherer, S.W., Lee, C., Birney, E., Altshuler, D.M., Eichler, E.E., Carter, N.P., Hurles, M.E., and Feuk, L. (2007). Challenges and standards in integrating surveys of structural variation. Nat. Genet. *39,* S7-15.

Schizophrenia Gene Database, Accessed June 2009. http://www.schizophreniaforum.org.proxy1.lib.uwo.ca:2048/.

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

Sebat, J., Lakshmi, B., Troge, J., Alexander, J., Young, J., Lundin, P., Maner, S., Massa, H., Walker, M., Chi, M. *et al.* (2004). Large-scale copy number polymorphism in the human genome. Science *305,* 525-528.

Sharp, A.J., Hansen, S., Selzer, R.R., Cheng, Z., Regan, R., Hurst, J.A., Stewart, H., Price, S.M., Blair, E., Hennekam, R.C. *et al.* (2006). Discovery of previously unidentified genomic disorders from the duplication architecture of the human genome. Nat. Genet. *38,* 1038-1042.

Shi, J., Levinson, D.F., Duan, J., Sanders, A.R., Zheng, Y., Pe'er, I., Dudbridge, F., Holmans, P.A., Whittemore, A.S., Mowry, B.J. *et al.* (2009). Common variants on chromosome 6p22.1 are associated with schizophrenia. Nature

Singh, S.M., Castellani, C.A., and O'Reilly, R.L. (2009). Copy number variation showers in schizophrenia: an emerging hypothesis. Mol. Psychiatry *14,* 356-358.

Singh, S.M., McDonald, P., Murphy, B., and O'Reilly, R. (2004). Incidental neurodevelopmental episodes in the etiology of schizophrenia: an expanded model involving epigenetics and development. Clin. Genet. *65,* 435-440.

Singh, S.M., Murphy, B., and O'Reilly, R. (2002). Monozygotic twins with chromosome 22q11 deletion and discordant phenotypes: updates with an epigenetic hypothesis. J. Med. Genet. *39,* e71.

Singh, S.M., and O'Reilly, R. (2009). (Epi)genomics and neurodevelopment in schizophrenia: monozygotic twins discordant for schizophrenia augment the search for disease-related (epi)genomic alterations. Genome *52,* 8-19.

Stefansson, H., Ophoff, R.A., Steinberg, S., Andreassen, O.A., Cichon, S., Rujescu, D., Werge, T., Pietilainen, O.P., Mors, O., Mortensen, P.B. *et al.* (2009). Common variants conferring risk of schizophrenia. Nature

Stefansson, H., Rujescu, D., Cichon, S., Pietilainen, O.P., Ingason, A., Steinberg, S., Fossdal, R., Sigurdsson, E., Sigmundsson, T., Buizer-Voskamp, J.E. *et al.* (2008). Large recurrent microdeletions associated with schizophrenia. Nature

Stranger, B.E., Forrest, M.S., Dunning, M., Ingle, C.E., Beazley, C., Thorne, N., Redon, R., Bird, C.P., de Grassi, A., Lee, C. *et al.* (2007). Relative impact of nucleotide and copy number variation on gene expression phenotypes. Science *315,* 848-853.

Sullivan, P.F. (2008). Schizophrenia genetics: the search for a hard lead. Curr. Opin. Psychiatry. *21,* 157-160.

Tang, X., Jang, S.W., Okada, M., Chan, C.B., Feng, Y., Liu, Y., Luo, S.W., Hong, Y., Rama, N., Xiong, W.C., Mehlen, P., and Ye, K. (2008). Netrin-1 mediates neuronal survival through PIKE-L interaction with the dependence receptor UNC5B. Nat. Cell Biol. *10,* 698-706.

The International Schizophrenia Consortium, Manuscript preparation, Purcell, S.M., Wray, N.R., Stone, J.L., Visscher, P.M., O'Donovan, M.C., Sullivan, P.F., Sklar, P., Data analysis *et al.* (2009). Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature

The International Schizophrenia Consortium, Manuscript preparation, Stone, J.L., O'Donovan, M.C., Gurling, H., Kirov, G.K., Blackwood, D.H., Corvin, A., Craddock, N.J., Gill, M. *et al.* (2008). Rare chromosomal deletions and duplications increase risk of schizophrenia. Nature

Torrey, E.F., Torrey, B.B., and Peterson, M.R. (1977). Seasonality of schizophrenic births in the United States. Arch. Gen. Psychiatry *34,* 1065-1070.

Tsankova, N., Renthal, W., Kumar, A., and Nestler, E.J. (2007). Epigenetic regulation in psychiatric disorders. Nat. Rev. Neurosci. *8,* 355-367.

Verveer, J., Huizer, K., Fekkes, D., and van Beveren, N.J. (2007). Gene-expression profiling in schizophrenia: an overview. Tijdschr. Psychiatr. *49,* 7-16; discussion 17-9.

Vita, A., De Peri, L., Silenzi, C., and Dieci, M. (2006). Brain morphology in first-episode schizophrenia: a meta-analysis of quantitative magnetic resonance imaging studies. Schizophr. Res. *82,* 75-88.

Walsh, T., McClellan, J.M., McCarthy, S.E., Addington, A.M., Pierce, S.B., Cooper, G.M., Nord, A.S., Kusenda, M., Malhotra, D., Bhandari, A. *et al.* (2008). Rare Structural Variants Disrupt Multiple Genes in Neurodevelopmental Pathways in Schizophrenia. Science

Weedon, M.N., Lettre, G., Freathy, R.M., Lindgren, C.M., Voight, B.F., Perry, J.R., Elliott, K.S., Hackett, R., Guiducci, C., Shields, B. *et al.* (2007). A common variant of HMGA2 is associated with adult and childhood height in the general population. Nat. Genet. *39,* 1245-1250.

Weiss, L.A., Shen, Y., Korn, J.M., Arking, D.E., Miller, D.T., Fossdal, R., Saemundsen, E., Stefansson, H., Ferreira, M.A., Green, T. *et al.* (2008). Association between microdeletion and microduplication at 16p11.2 and autism. N. Engl. J. Med. *358,* 667-675.

Wilke, K., Duman, B., and Horst, J. (2000). Diagnosis of haploidy and triploidy based on measurement of gene copy number by real-time PCR. Hum. Mutat. *16,* 431-436.

Wilson, G.M., Flibotte, S., Chopra, V., Melnyk, B.L., Honer, W.G., and Holt, R.A. (2006). DNA copy-number analysis in bipolar disorder and schizophrenia reveals aberrations in genes involved in glutamate signaling. Hum. Mol. Genet. *15,* 743-749.

Xu, B., Roos, J.L., Levy, S., van Rensburg, E.J., Gogos, J.A., and Karayiorgou, M. (2008). Strong association of de novo copy number mutations with sporadic schizophrenia. Nat. Genet. *40,* 880-885.

Zogopoulos, G., Ha, K.C., Naqib, F., Moore, S., Kim, H., Montpetit, A., Robidoux, F., Laflamme, P., Cotterchio, M., Greenwood, C. *et al.* (2007). Germ-line DNA copy number variation frequencies in a large North American population. Hum. Genet. *122,* 345-353.

APPENDIX B: UNIQUE VARIATION RAW DATA TABLE

Unique Copy Number Variants present in **affected patients** 1, 2 and 3. Highlighted samples represent variants which are shared by the patient's identical twin.

Dose	Chr	Size (kb)	Start Breakpoint	End Breakpoint
AFFECTED PATIENT TWIN PAIR 1				
1	1	140	35941848	36081995
3	14	140	46584175	46724599
3	2	141	103527434	103668922
1	2	142	206555025	206696550
3	13	142	86418986	86561112
3	13	142	103912893	104055224
1	1	143	234251677	234394633
3	2	143	150919694	151062297
3	6	143	121313696	121456308
1	12	143	48445979	48588726
3	13	144	83266990	83411321
1	18	144	19214459	19357969
3	7	145	83912677	84057475
3	7	145	108597007	108742375
1	10	145	70349125	70493668
1	12	145	112068420	112213825
1	16	145	67894904	68039688
1	17	145	34108140	34253331
3	4	146	46588523	46734072
1	7	146	6436698	6582639
3	8	146	69614345	69760244
1	15	146	73070222	73216575
1	19	146	56088392	56234603
3	3	147	85540476	85687749
1	22	147	27460012	27607011
1	17	148	26035556	26184034
3	3	150	6797467	6947746
3	4	150	92951929	93101529
3	14	150	26376109	26526221
1	10	151	75287479	75438654
3	5	152	62456472	62608494
3	4	153	19776551	19929432

1	11	153	117977645	118130590
1	6	154	37129238	37283645
1	8	157	74913605	75070367
1	10	158	121278491	121436693
1	10	159	104497146	104655778
1	12	160	55739188	55899475
1	15	160	42330455	42490110
1	X	160	118498146	118658561
1	3	161	9855617	10016384
3	15	161	44221215	44382179
3	1	162	49153270	49315177
3	11	162	41926007	42088475
3	1	163	82200267	82362886
1	2	163	231189442	231352115
3	13	163	36700041	36862892
1	1	164	23316047	23480090
3	3	164	87479383	87643832
3	3	164	175253217	175416967
1	12	164	123898908	124063008
3	8	165	35622835	35788028
3	4	166	98748005	98913662
1	14	166	72525137	72690931
3	1	167	237141304	237308592
3	3	169	87050127	87219238
3	5	170	128030017	128200501
3	5	171	30144134	30315239
1	11	171	9406953	9577980
1	11	171	65866016	66037084
3	X	173	22599023	22772189
3	4	174	137197409	137371898
3	18	174	69402130	69576602
3	12	176	83299820	83475386
3	14	176	82943806	83119956
3	14	177	78398694	78575401
3	3	181	95394436	95575371
1	20	183	30403166	30585730
1	19	184	51638501	51822652
1	7	185	44461469	44646753
1	9	185	36414428	36599046
1	10	189	69827276	70016512
3	X	190	92764162	92954279

3	2	191	193611095	193801842
3	6	191	97869838	98060454
3	8	192	106109040	106301304
3	13	192	55632039	55824272
1	19	193	19222261	19415234
3	2	194	180688512	180882111
1	12	194	54439264	54632817
3	4	195	32739085	32934574
3	5	197	86932276	87129565
1	17	197	40456357	40653321
1	1	199	44959872	45158845
3	4	199	134534784	134733668
1	7	200	129327625	129527522
1	11	200	47181807	47382199
1	20	200	34933472	35133811
1	6	201	111179259	111380489
1	1	206	11846557	12052169
1	12	207	110400664	110607667
1	12	208	48800360	49008396
3	8	210	111905627	112115160
3	7	211	41056861	41267549
3	3	213	118009920	118223314
1	9	214	126551707	126765380
1	16	214	65280534	65494071
1	7	215	65064572	65279453
1	19	216	8448005	8664224
1	2	217	70038612	70255257
1	19	217	13037996	13255175
1	15	219	62790107	63009260
1	1	220	148466043	148685774
3	1	220	186686615	186906636
1	19	225	16557937	16783097
3	6	227	94263187	94489702
3	13	227	52803308	53030285
3	13	228	68959402	69187191
3	15	229	45408259	45637618
3	6	231	22817409	23048769
1	9	231	130073233	130304604
1	11	231	46677336	46908179
1	16	253	55849968	56103237
1	5	256	133908218	134164232

3	4	260	61839120	62099188
1	7	263	44675946	44939421
1	1	265	21206063	21470723
1	17	265	7126970	7391749
1	7	266	65600819	65867070
3	X	268	91384292	91651970
3	4	274	96515818	96789981
1	15	289	62229576	62518093
1	15	291	39239917	39530718
1	16	292	11574650	11866352
1	6	293	42855215	43147929
3	X	302	86275822	86577847
1	1	303	26451650	26755020
1	9	306	129574115	129880384
1	1	319	149067771	149387100
1	2	329	218760791	219090096
1	15	330	73467785	73798190
3	15	336	35006061	35342552
1	12	343	49744769	50087674
1	1	351	27589811	27940397
1	20	366	34449710	34815484
3	X	370	144600143	144970363
1	5	383	138582582	138965878
1	12	390	109170533	109560206
3	12	402	84829113	85230776
3	X	406	25485368	25891814
1	19	419	11130341	11549696
1	1	427	9661373	10087888
1	17	443	35239142	35681720
1	19	454	43367699	43821394
3	5	501	25089165	25590081
1	19	509	16022621	16531660
1	1	527	152115620	152642862
1	17	535	23779286	24314288
1	X	626	1131508	1757545
1	16	662	66092164	66754346
1	1	679	28336765	29015567
1	X	3278	58520670	61798820

AFFECTED PATIENT TWIN PAIR 2

1	1	180	9778163	9957792
1	7	140	73090014	73230072
3	1	141	82222358	82362886
3	7	143	83912677	84055959
1	19	144	44316630	44460510
2	X	145	91506481	91651970
1	19	148	46263744	46411544
3	8	150	106618262	106768118
3	13	150	87127756	87277965
3	13	155	55668814	55824272
1	11	158	65902868	66061025
3	5	161	91637283	91798669
3	5	162	166463726	166625348
3	3	163	87481267	87643832
3	4	164	61923121	62087073
1	19	164	43508366	43671958
3	3	167	118062246	118228951
3	3	168	74933526	75101320
1	9	168	130088581	130256322
2	X	172	91834183	92005980
2	X	176	91057319	91232842
3	4	179	32741664	32920832
3	4	179	137197409	137376822
1	12	185	110400664	110585773
1	19	185	51005101	51189609
3	12	187	83300279	83487631
3	6	195	94686904	94881698
3	4	198	63837162	64034922
3	13	205	58031544	58236907
1	19	206	5224746	5430522
3	2	207	146809578	147016667
1	19	225	17045549	17270289
3	12	229	77756186	77984927
1	17	259	35423126	35681720
3	4	266	96515818	96781483
3	13	268	69962380	70230477
3	6	371	94273915	94644665

AFFECTED PATIENT TWIN PAIR 3

1	1	561	9591661	10152235
1	6	140	86289694	86430082
1	2	141	157913467	158054901
1	13	141	31907274	32048447
3	17	141	11447453	11588180
1	4	142	140266254	140407893
1	14	142	23700850	23842762
1	1	143	152010745	152153905
1	4	143	140100790	140244116
1	5	143	56124612	56267942
1	12	143	49374090	49517482
1	3	144	16459126	16602810
1	12	144	3802327	3946116
3	1	145	235262729	235408039
3	19	146	33164406	33309917
3	2	147	30937261	31084008
1	2	148	233771687	233919614
1	3	148	32454591	32603027
1	4	148	100970546	101118072
3	5	148	7576690	7724849
1	6	148	109781158	109929134
1	19	148	14510264	14657884
3	1	149	212003660	212152450
3	7	149	28735588	28884876
1	9	149	5492365	5641082
1	10	149	120790810	120940124
1	14	149	91246465	91394995
1	1	150	26038272	26188023
3	2	150	138276574	138426407
3	16	151	13348171	13499560
1	19	151	44093550	44244062
1	6	152	158115683	158267312
1	8	152	74917978	75070367
1	2	153	196156925	196309550
3	3	153	183214609	183368058
3	2	154	102795114	102949489
1	3	154	137779150	137933283
3	12	154	112521672	112675877
1	15	154	61199527	61353755

1	18	154	19243906	19397737
1	19	155	47005668	47161162
1	14	156	70482470	70638613
1	1	159	39202766	39361993
1	1	159	220822810	220981603
3	2	159	21593075	21752097
3	3	159	26591799	26750767
3	7	161	67105860	67266598
1	3	162	142664374	142826602
3	5	162	166514243	166676280
1	7	162	43623475	43785847
1	19	163	13039787	13202317
1	21	163	37471009	37634079
1	3	164	45607225	45771661
1	6	164	137100744	137264364
1	14	166	72525137	72690931
1	2	167	197954870	198122260
1	14	167	44485974	44652702
3	18	167	41158256	41325091
3	2	168	114933177	115100863
1	12	168	123898908	124066724
1	1	169	234005297	234173864
3	10	170	34437644	34608038
1	14	170	60922546	61092097
1	17	170	25638880	25808508
3	22	170	31470411	31640129
1	3	172	47824940	47996838
1	3	172	48366512	48538274
1	8	172	41899502	42071053
3	11	173	119882018	120055165
1	2	177	68756206	68932783
3	2	177	122580871	122757625
1	3	177	178421704	178598500
3	10	177	123835145	124011811
1	21	177	29220209	29397477
3	10	178	43686012	43864461
1	14	178	70819219	70997120
3	16	178	81013186	81191554
2	Y	179	8865592	9044308
1	4	180	74133097	74313364
3	10	180	34103865	34283816

1	2	181	234858881	235040094
1	5	181	132275109	132456285
1	11	181	101701760	101883053
3	4	182	21002075	21184311
1	1	183	70407173	70590165
1	1	183	173075216	173258413
1	1	185	242998794	243183734
1	3	185	151689929	151874479
3	5	185	38054031	38239133
1	9	185	73518742	73703650
3	13	185	58579180	58764568
3	18	185	71641081	71825868
3	16	186	63294523	63480068
1	4	187	83364956	83552260
3	9	187	80338349	80525288
1	12	188	54444323	54632817
3	15	188	85069861	85257800
1	3	190	51410594	51600122
3	1	192	208797145	208989609
3	3	192	10838089	11029885
1	3	193	113614892	113808344
1	15	193	77955337	78148060
1	16	193	3599460	3792423
1	2	194	177728614	177922576
1	3	195	4826236	5020869
1	1	196	224930638	225126965
1	2	196	101140713	101336739
1	8	196	29990821	30186827
3	2	197	15721845	15919252
1	2	198	37003811	37201579
1	6	198	74185007	74382957
1	7	199	50212241	50411389
1	14	200	57788577	57988778
1	3	202	171288456	171490700
3	4	203	183173429	183376813
3	7	203	69521684	69725051
1	2	204	65128994	65333370
1	2	205	108610364	108815860
1	19	205	5509263	5714533
1	21	208	32980726	33188785
1	17	209	44593864	44803230

1	1	210	88966952	89176746
1	5	210	56385177	56594995
1	6	211	37103513	37314330
1	14	211	35131471	35342210
1	15	212	63737570	63949814
1	18	212	59015193	59227073
1	20	212	3730305	3942633
1	1	213	15693987	15906788
1	2	213	85893290	86106004
1	1	215	27454501	27669101
3	3	215	128191892	128406672
1	14	215	20703467	20918749
1	6	216	11383973	11599550
3	18	217	42197368	42414476
1	2	218	42174880	42392456
1	2	218	84851397	85068937
1	7	218	23380532	23598892
1	10	218	104428067	104646197
1	19	218	4628570	4846662
1	5	219	98178433	98396954
1	14	220	88903876	89124229
1	15	220	73467785	73688009
1	7	221	139371211	139592522
1	16	221	55744347	55965147
1	2	222	55259871	55481417
1	12	222	44908536	45130428
1	6	223	90496190	90719602
1	17	223	55742647	55965707
1	22	223	40572966	40795781
1	16	224	4412208	4636655
1	5	225	133938960	134164232
3	7	225	40604377	40829641
1	2	229	196649712	196878701
3	2	230	21258669	21489075
1	12	230	26885658	27116119
3	9	231	117233148	117464051
1	4	234	153583372	153817174
1	10	235	98399720	98635005
1	15	235	91117752	91352350
3	10	236	86649591	86886030
3	12	238	96117429	96355377

1	13	239	44700607	44939357
1	10	242	75979266	76221508
3	11	242	120056836	120298597
1	9	243	113878648	114122087
1	10	244	69702528	69946707
1	2	245	68213067	68458094
1	3	246	123617898	123863807
1	2	249	174930664	175179941
1	10	249	63340231	63589398
3	5	251	136424307	136675003
1	10	252	121243589	121495887
3	7	253	30925527	31178236
1	16	254	27107894	27361577
3	5	255	146580403	146835438
1	5	257	156425016	156682224
1	10	259	35393759	35652397
1	8	260	125362216	125622288
3	10	260	72323799	72583730
3	12	261	2323621	2585036
1	15	265	39252637	39517538
1	1	267	101336240	101602779
3	2	267	3869538	4136798
3	17	267	12365264	12631916
3	4	268	96515818	96783465
3	15	270	85881938	86151532
3	1	272	175984190	176256267
1	12	273	110358459	110631490
3	18	273	43074272	43347121
1	1	275	44938458	45213904
1	6	279	13535313	13814341
1	9	280	36298384	36578101
1	1	281	111785010	112066063
3	7	281	155434211	155715304
1	9	282	99646935	99928524
1	6	286	90856975	91143183
1	1	290	21159601	21449383
1	20	290	34557483	34847369
1	12	292	95162838	95454472
1	4	294	2873083	3167446
1	2	296	191430242	191726336
1	14	301	74559747	74861141

1	22	301	37207504	37508647
3	10	302	87673335	87975231
1	6	303	41871828	42174391
1	1	304	36026767	36330570
3	4	304	16117951	16422245
1	6	305	111171099	111475841
3	15	310	35032208	35342552
3	15	310	93491492	93801398
3	13	311	35027630	35338404
1	3	314	36898367	37211962
1	15	314	49853481	50167028
1	18	316	2499180	2815488
1	5	317	133506428	133823088
3	3	319	8452617	8772095
1	3	319	71409296	71728702
1	4	319	38613450	38932061
1	15	320	50471296	50791108
1	20	320	5447379	5767844
1	X	323	1415800	1738663
3	5	325	151698756	152023408
3	20	327	57502963	57829958
1	10	332	69163584	69495985
3	6	335	39242146	39576714
1	11	338	18281695	18620126
1	5	340	42928219	43268629
1	10	340	111809719	112149634
1	17	342	24977717	25319401
3	18	343	5628443	5971146
1	12	346	88504356	88850127
3	10	348	43947174	44295413
1	1	351	92945370	93296587
1	2	361	70014339	70375522
3	5	361	5966192	6327436
1	12	368	14315345	14683601
3	7	370	3912703	4282770
1	2	371	197514145	197885417
1	5	373	138589927	138962496
1	12	373	67071040	67444016
1	17	373	24523266	24896622
1	19	373	16884574	17257761
1	5	376	96025282	96401607

3	7	377	41056861	41433786
1	18	380	2916377	3296329
1	13	382	45474634	45856856
3	2	386	217373660	217759543
3	18	386	47323458	47709706
1	11	391	116240998	116631943
3	1	392	50199129	50590674
3	2	395	118951806	119346777
1	12	399	9985910	10384737
1	8	402	67593619	67995419
1	13	406	48863105	49269415
1	20	406	34933472	35339719
3	20	418	19204240	19622256
1	13	420	29764701	30184370
3	3	424	73520749	73944942
3	8	425	49505148	49930186
3	12	425	124920288	125345716
3	16	426	64212617	64638210
1	11	428	46404226	46832509
3	19	428	36745505	37173086
3	16	431	49619553	50051052
3	3	432	125195683	125627289
3	11	443	119140699	119583496
1	17	446	23888693	24334196
3	7	448	33724570	34172193
1	17	450	4872883	5322591
1	12	454	49642898	50097019
3	10	464	106455000	106919303
3	20	471	24282704	24753601
1	19	481	50569829	51050518
3	1	486	47976577	48462816
1	6	494	42547032	43040813
1	12	496	118988717	119484748
1	9	502	33855624	34358085
1	3	504	102602277	103105972
1	1	513	224041894	224554552
3	5	514	167077326	167591180
1	11	524	127585715	128109442
3	3	529	65226614	65755909
1	2	534	218759559	219293061
1	13	539	98462417	99001574

3	5	546	168103153	168649122
3	13	551	52432673	52984085
1	17	552	35130176	35681720
1	20	592	32298049	32889861
3	22	612	25838858	26450707
1	1	619	152187016	152805729
3	14	628	32648669	33276491
3	3	640	69817050	70456880
3	2	661	71600743	72261888
1	11	662	72847609	73509828
1	1	677	27759600	28436346
3	16	692	50205808	50897341
1	7	701	44412461	45113181
3	1	707	60370502	61077353
3	12	750	5069410	5819619
3	17	762	47083999	47846336
3	1	776	48913275	49689671
1	19	776	16022621	16798263
3	18	781	7141858	7922414
3	10	788	117663132	118450781
1	12	789	108812014	109601472
3	1	821	33932345	34753764
3	19	825	35428309	36253237
3	8	835	136404904	137239867
3	1	847	162621657	163468994
1	11	854	117565770	118419405
3	11	865	112044994	112909947
3	20	868	52879888	53748040
1	16	886	65947156	66833361
1	14	1011	49052167	50062686
3	17	1066	28403584	29469444
1	10	1138	74451311	75589696
1	15	1495	61493768	62989044