UNIVERSITY OF LONDON THESIS

Degree PHD Year 2008 Name of Author CHOUDHURY, KHALID, HUSSAIN.

MOLECULAR GENETICS OF THE CHROMOSOME 11q22-24 SCHIZOPHRENIA SUSCEPTIBILITY REGION

Khalid Hussain Choudhury

University College London

Molecular Psychiatry Laboratory,
Department of Mental Health Sciences,
Royal Free and University College London Medical School,
The Windeyer Institute of Medical Sciences,
46 Cleveland Street, London W1T 4JF.

This thesis is submitted for the degree of PhD in Genetics at
University College London
October 2007

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

ACKNOWLEDGEMENTS

I would like to gratefully acknowledge the enthusiastic supervision of Professor Hugh Gurling during the PhD. I thank Doctor Andrew McQuillin for supervision of laboratory work and Professor Dave Curtis for his advice and help in the statistical analysis of data.

Patients suffering from schizophrenia and volunteers acting as controls are thanked for agreeing to take part in this project. As are those people who contributed to collecting samples from both patients and controls.

Finally, I am forever indebted to my family and friends for their understanding, endless patience and encouragement when it was most required. A special appreciation goes to my nephew Wafi, who helped enormously during the write-up of the thesis.

ABSTRACT

Linkage analyses of multiply affected schizophrenic families have confirmed the involvement of the chromosome 11q22-24 region in the aetiology of schizophrenia, with LOD scores of 3.4 and 3.1. As a result, a genetic association study was performed to fine map this area by detecting linkage disequilibrium between DNA markers and the genetic disorder. This was done using the University College London (UCL) sample of 496 cases and 488 supernormal controls. Seven microsatellite or SNP markers localised within or near the FXYD6 gene showed empirically significant allelic associations with schizophrenia. Confirmation was then performed by the study of an Aberdeen sample consisting of 858 cases and 591 controls. Here, two of the SNPs replicated association with schizophrenia. Sequencing of the FXYD6 gene led to the discovery of several mutations including a rare non-database SNP, which was found at a significantly higher frequency in the selected high risk haplotype schizophrenic patients, when compared to the randomly chosen control individuals. However, a genetic association test of this SNP was found to be negatively associated to schizophrenia when genotyping in the whole UCL case-control sample. FXYD6 encodes for the protein phosphohippolin and functions by modulating the kinetic properties of Na,K-ATPase to the specific requirements of a given tissue, cell type or physiological state. Phosphohippolin is predominantly expressed in the brain, with strong expression in the hippocampus and the cerebellum. The prominent levels of FXYD6 expression in regions of the brain thought to be involved in schizophrenia provide strong support that the FXYD6 gene increases genetic susceptibility to schizophrenia. Further confirmations involving FXYD6 to schizophrenia are essential in other case-control samples. In addition, aetiological base pair changes in FXYD6 or in associated promoter/control regions that cause an abnormal function or expression of phosphohippolin require detection.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

AIMS OF THESIS

1) To screen the chromosome 11q22-24 region that has been implicated by genetic linkage studies in families. The method employed would use the case-control approach by detecting evolutionary determined linkage disequilibrium to fine map a novel schizophrenia susceptibility gene.

2) To replicate any genes found to be associated with schizophrenia in an independent Aberdeen case-control sample.

3) To find potential aetiological base pair changes by resequencing the exons, slice site junctions and promoter region (5 kb), as well as the 5' and 3' untranslated regions, in any genes found within or close to regions implicated by linkage disequilibrium.

4) To genotype any potential aetiological base pair changes in a large case-control sample, so as to confirm or reject genetic involvement in susceptibility to schizophrenia.

1.0 INTRODUCTION TO THE MOLECULAR GENETICS OF SCHIZOPHRENIA

1.1 GENERAL INTRODUCTION

Schizophrenia is a common psychiatric disorder that covers a broad range of cognitive, affective and behavioural abnormalities. It includes one or more symptoms of personality change, delusions, thought disorder, lack of self belief and auditory hallucinations (Andreasen 1995). These symptoms lead towards considerable social and occupational dysfunction. There is no known chemical pathology. However, abnormal brain morphology has been reported in over a hundred studies.

Schizophrenia can have a good prognosis with spontaneous recovery. Other cases have associated cognitive deficits with a downward spiral of increasing severe symptoms. The episodes of illness can be brief where there is relative recovery but long term overall decline. Some cases experience very good recovery in between illnesses, even without treatment. It seems likely that 60-70% of schizophrenics require long term antipsychotic medication.

The lifetime prevalence of schizophrenia is approximately 0.85% in the United Kingdom (Hafner & an der Heiden 1997). Worldwide, this can range between 0.03% in the Far East to as high as 3% in certain isolates of northeast Finland (Hovatta et al. 1997). Onset of symptoms typically occurs during late adolescence and early adulthood. These symptoms can be subtle with sudden appearance (Castle et al. 1991). The mean age of onset for schizophrenia is 25 years (Gottesman 1991), with females (27 years) developing the disease slightly later in life than males (23 years).

The relatively early age of onset in schizophrenia is a critical period in a young adult's development. It is common to attribute problems at this age to social

isolation, long term unemployment, poverty, homelessness and life events generally. The large social and economic burden of schizophrenia makes it one of the world's most pressing health problems. Schizophrenia is associated with a decrease in life expectancy of approximately 80-85% compared to the general population, as shown in a study of over 168,000 Swedish citizens undergoing psychiatric treatment (Hannerz et al. 2001). This is mainly due to a suicide rate of 10% and to the adverse effects of the illness on medical health.

1.2 HISTORICAL BACKROUND

Descriptions of schizophrenia-like symptoms date back to as far as 200 BC. However, writings in the Book of Hearts of the Eber papyrus and other ancient texts depict symptoms that vaguely relate to the modern criteria for schizophrenia (Evans et al. 2003). The first recognised descriptions of schizophrenia were by two physicians, John Haslam (1764-1844) in England and Phillipe Pinel (1745-1826) in France. Bénédict Morel (1809–1873) then introduced the term "dementia praecox" in 1852 to describe dementia in the young. This emphasised on the deterioration in mental abilities of patients and the early age of onset for the disease.

In 1893, Emil Kraepelin (1856-1926) made a distinction between the psychotic disorders dementia praecox and manic depression. Kraepelin believed that dementia praecox was primarily a disease of the brain and should be distinguished from other forms of mental illness such as manic depression (Kraepelin 1907; Hansen & Atchison 2000). He also included other kinds of psychosis within the dementia praecox category including catatonia (muscular rigidity and stupor), paranoia (systemised delusions and baseless suspicion of other people's motives) and hebephrenia (inappropriate behaviour and affect with delusional hallucinations). These have now been used for classifying different subtypes of schizophrenia.

Finally in 1908, the Swiss psychiatrist Eugen Bleuler (1857-1939) proposed the name schizophrenia to describe deterioration in the functioning of personality, thinking, memory and perception (Liddell & Scott 1980). This highlighted the

disruption of the thought processes and mood of the patient. Bleuler realised that the illness was not a dementia because some of his patients had improved rather than deteriorated.

1.3 DIAGNOSIS OF SCHIZOPHRENIA

Modern psychiatrists use operational criteria for diagnosis that are updated regularly. This has helped to improve the diagnosis of schizophrenia throughout the world using reliable and valid systems. Diagnosis is usually based on the patient's self-reported experiences and observed abnormalities in speech and behaviour, as described by family members and friends. This is usually followed up with secondary observations by a psychiatrist. Operational criteria are then applied, which are known to be sufficiently sensitive as well as specific.

The most widely used criteria for diagnosing schizophrenia are attained from the American Psychiatric Association's Diagnostic and Statistical Manual of Mental Disorders, currently DSM-IV (APA 1994), and the World Health Organization's International Statistical Classification of Diseases and Related Health Problems, currently ICD-10 (WHO 1992). Both diagnostic schemes are system based on a set of criteria that may or may not be fulfilled by the observations during clinical examination. Particular emphasis is placed on affective and psychotic symptoms as a defining feature because it has been shown that mood incongruent features are typical of schizophrenia, whereas mood congruent features are typical of bipolar disorder.

To diagnose schizophrenia under DSM-IV, a person must display the following (APA 1994):

1) Two or more of the subsequent symptoms, present for a significant portion of time during a one month period:
    a. Delusions.
    b. Hallucinations.
    c. Disorganised speech.
    d. Disorganised/Catatonic behaviour.
    e. Negative symptoms.

2) Exhibit social/occupational dysfunction for a significant portion of time since the onset of disturbance e.g. work, interpersonal relations or self-care.

3) Show duration with continuous signs of disturbance persisting for at least six months including one month of symptoms.

Other diagnostic systems for schizophrenia include Schneider (1959) and RDC, the Research Diagnostic Criteria (Spitzer & Endicott 1977). DSM-IV, ICD-10 and RDC have all been found to be reliable and valid as diagnostic instruments as those used in other branches of medicine.

1.3.1 POSITIVE AND NEGATIVE SYMPTOMS

Positive (productive) and negative (deficit) symptoms are terms often used to describe schizophrenia as a way of classifying them into groups (Sims 2002). Positive symptoms include auditory hallucinations, delusions, thought disorder and disorganised behaviour. Negative symptoms include qualities such as alogia (poverty of speech), blunted emotion, avolition (lack of desire and motivation) and anhedonia (inability to experience pleasure such as eating and social interactions). Other attempts to divide schizophrenia include the factor analytic syndromes of

"disorganisation", "confusion" and "psychomotor retardation" (Liddle 1987; Peralta & Cuesta 2001).

1.3.2 SUBTYPES OF SCHIZOPHRENIA

DSM-IV contains five sub-classifications of schizophrenia (WHO 1992; APA 1994):

Catatonic:

The dominant features of catatonic schizophrenia are major psychomotor disturbances e.g. reduced mental awareness, immobility, freezing, waxy flexibility of the limbs and sudden motor excitement. Patients are often uncommunicative and may be stuporose. Sometimes bizarre motor postures are maintained for long periods. Episodes of violent excitement may be a striking feature of this condition, even during periods of immobility (DSM code 295.2/ICD code F20.2).

Hebephrenic:

A form of schizophrenia characterised by prominent affective changes, loss of volition and severe thought disorder. The mood is shallow and often accompanied by giggling or self-satisfaction, while thought is sometimes grossly disorganised with neologisms, thought blocking, knight's move thinking derailment and incomprehensible speech. There is a tendency to remain solitary as behaviour seems empty of purpose and feeling. The age of onset is usually early (15-25 years) and prognosis is generally poor due to the rapid development of negative symptoms, largely flattening, loss of affect and loss of volition. There is sometimes a preoccupation with religion, philosophy or sex. However, this is not specific to hebephrenic schizophrenia (DSM code 295.1/ICD code F20.1).

Paranoid:

This is the generally the most common type of schizophrenia and is typified by paranoid delusions. It is sometimes accompanied by auditory hallucinations. Disturbances of affect, speech and desire are not prominent. Examples of paranoid

delusions include ideas of persecution, being the centre of a conspiracy, being undermined and being targeted by religious or criminal forces. Hallucinatory voices that threaten the patient or give commands are typically resented and unwanted. Non-verbal auditory hallucinations such as whistling, humming or laughter and olfactory hallucinations of smell and sexual sensation are common. The course of paranoid schizophrenia maybe episodic, with partial and complete remissions, or chronic, with symptoms persisting over many years making it difficult to distinguish discrete episodes. The age of onset tends to be later than hebephrenic and catatonic forms. A subtype of paranoid schizophrenia can have a very good prognosis (DSM code 295.3/ICD code F20.0).

Residual:

In residual schizophrenia, the patient does not exhibit prominent psychotic symptoms but has a previous history of at least one episode of schizophrenia with prominent psychotic symptoms. Two or more characteristic symptoms are continued to be suffered such as social withdrawal, eccentric behaviour and illogical thinking. Positive symptoms are only present at low intensity (DSM code 295.6/ICD code F20.5).

Undifferentiated:

When general criteria of schizophrenia are met but do not conform to any form of subtypes, patients are categorised as undifferentiated. It also includes sufferers who exhibit features of more then one subtype without a clear predominant set of diagnostic characteristics to put them into that group. Often, patients who were catatonic or hebephrenic become undifferentiated over time. (DSM code 295.9/ICD code F20.3).

A further two subtypes are recognised in ICD-10 (WHO 1992):

Depression superimposed on schizophrenia or post-schizophrenic depression:
This is where a depressive episode, which may be prolonged, arises in the aftermath of a schizophrenic illness. Some symptoms may still be present but at low level. It is often difficult to decide which of the patient's symptoms are due to depression and which are due to antipsychotic drugs or to the blunted affect of schizophrenia itself. An increased risk of suicide is associated with this depressive disorder. This category is not to be confused with schizoaffective psychosis, which has a different conceptual basis (ICD code F20.4).

Simple schizophrenia:
A disorder where there is a subtle but gradual development of prominent negative symptoms e.g. inability to meet the demands of society. There is no history of psychotic episodes with delusions and hallucinations. The increasing social isolation causes individuals to lose any means of livelihood and they often lose all motivation/direction in life (ICD code F20.6).

1.4 TREATMENT

Treatment was revolutionised in the mid 1950s with the development and introduction of the antipsychotic drug chlorpromazine (Turner 2007). Although antipsychotic treatments are currently the mainstay for treating schizophrenia, research has shown that cognitive and psychosocial interventions carried out after successful drug treatments can enhance clinical improvements. Almost all schizophrenic patients benefit from antipsychotic drugs. The UK national institute of clinical excellence (NICE) guidelines suggest that all in patients should have ten cognitive therapy sessions. This is rarely, if ever, achieved. Some severe cases of schizophrenia who are relapsing may deteriorate even whilst being treated with clozapine and other modern antipsychotic drugs. However, even these cases seem to stabilise and eventually respond to treatment after relapses.

Antipsychotic drugs can be very successful in treating a proportion of schizophrenics. However, a certain amount of patients respond less well. Measures of effective treatment can include scales to score positive and negative symptoms (PANSS) and a host of other methods. About a third of schizophrenics have a severe deteriorating prognosis and only partially respond to antipsychotic treatments.

1.4.1 HOSPITALISATION

Hospitalisation can be voluntary or involuntary, known as a hospital treatment section or order in the UK. In hospitals, a specific treatment plan is created for each individual and future discharge is planned early on. The patient and his/her family members are involved as much as possible. Following hospital admission, available support services can include day hospitals, drop-in centres, support groups, sheltered employment and visits from various community mental health teams e.g. nurses, occupational therapists, social workers and counsellors. Patients can often remain out of hospital and the quality of their daily lives can improve steadily if they continue to retake antipsychotic drug treatment (Sadock & Sadock 2003).

1.4.2 MEDICATION

The main treatment for schizophrenia is antipsychotic medication. This gets rid of many symptoms within two to three weeks. Auditory hallucinations are the first to disappear and then thought disorder, followed by delusions. Risperidone is a common atypical antipsychotic drug that is usually preferred for initial treatment over the older typical antipsychotic drugs. It is generally well tolerated and is less likely to induce weight gain and obesity related diseases, which are associated with olanzopine and clozapine (Lieberman et al. 2005). Failure of symptoms to respond satisfactorily to at least two different antipsychotics is generally termed treatment-resistant schizophrenia (Meltzer 1997). Patients in this category may be prescribed clozapine, a more effective medication but potentially lethal due to the side affects e.g. agranulocytosis and myocarditis (Haas et al. 2007).

1.5 AETIOLOGY

The cause of the schizophrenia is now known to be largely genetic. Psychological and sociological factors were once thought to be relevant in the aetiology, but have fallen out of fashion. Genes have consistently been proven to be the most important biological risk factor for the development of schizophrenia (Section 1.6). Furthermore, schizophrenia-like psychoses may rarely develop after head injuries or as a result of demonstrable organic diseases (Johnstone et al. 1987). Some forms of epilepsy have also shown to be associated with schizophrenia, but these may also have a genetic origin.

Numerous attempts have been made to implicate different environmental risk factors as a possible cause for the disease. These include winter-spring season of birth, urban birth and rearing, prenatal or obstetrical complications and viral infection. However, none of these factors have shown consistent or large effects on the risk of the disease (Done et al. 1991).

1.5.1 BRAIN MORPHOLOGY OF SCHIZOPHRENIA

Schizophrenia is not only characterised by psychopathological features but also by alterations in physiology and anatomy. Brain changes were demonstrated following the introduction of non-invasive imaging by Johnstone et al. (1976). These have consistently shown enlarged ventricular spaces and reduced cortical volumes, particularly involving temporal lobe structures (Shenton et al. 1992). A meta-analysis of regional brain volumes found differences between schizophrenic patients and normal controls (Wright et al. 2000). The patients all had smaller cerebral volume, a bilaterally reduced volume of medial temporal lobe structures and a greater ventricular volume. Differences in temporal lobes, hippocampus and frontal lobes, as observed using brain imaging technologies, are heavily linked to the neurocognitive deficit associated with schizophrenia (Green 2006). This is specifically in areas of memory, attention, problem solving, decision making and social cognition.

More recently, positron emission tomography (PET), with oxygen isotopes, has suggested hypofrontality (Liddle et al. 1992). However, this was not well replicated using labelled glucose. Thompson et al. (2001) used high resolution magnetic resonance images (MRI) to show accelerated grey matter loss in very early onset schizophrenia. In addition, repeated MRI scans over a five year period found deficits progressing interiorly into temporal lobes. In some schizophrenics, this engulfed the sensorimotor and dorsolateral prefrontal cortices plus the frontal eye fields.

The overwhelming evidence in the literature show a volumetric reduction in the temporal lobe structures of schizophrenics. Furthermore, studies of the unaffected relatives of patients with schizophrenia have also shown the presence of similar brain abnormalities. This is thought to be partly due to the shared genetic liability (Lawrie et al. 2001).

The structural brain changes of patients previously identified as having a high genetic risk of developing schizophrenia have been studied by Job et al. (2005). Compared to findings two to three years before they were diagnosed as schizophrenic, these individuals showed a different spatial pattern of reductions in grey matter density in the left temporal lobe and right cerebellum.

The histological findings in schizophrenia are less convincing but include aberrantly located or clustered neurons, particularly in lamina II of the entorhinal cortex and in the neocortical white matter (Jakob & Beckmann 1986; Akbarian et al. 1996). These abnormalities may suggest an early neurodevelopmental anomaly affecting neuronal migration, survival and connectivity. Kovelman and Scheibel (1984) also found abnormally arranged pyramidal cells in the hippocampus. Other studies have considered that the cell bodies of pyramidal neurones in the hippocampus and neocortex are smaller (Zaidel et al. 1997; Pierri et al. 2001).

In summary, the neuropathology of schizophrenia has been proposed to consist of alterations in various neural microcircuitries ranging from the dendritic tree to the

synaptic terminal (Harrison & Weinberger 2005). These reported changes in synaptic products include reduced glutamate and γ-aminobutyric acid synaptosome release and reduced synaptophysin in the prefrontal cortices of patients with schizophrenia (Sherman et al. 1991; Glantz & Lewis 1997). By linking all these findings together, some investigators explain schizophrenia as a disorder of the synapse (McGlashan & Hoffman 2000; Frankle et al. 2003). Others have linked the neuropathological findings with genes recently implicated in the disease, as further evidence of schizophrenia being a disorder of the synapse (Owen et al. 2005). Several of these implicated genes are known to be directly involved in synaptic function (Section 1.9.2). However, the genes may encode for proteins with multiple functions that are not specifically concerned with the synapse e.g. NRG1 is known to encode for around 15 proteins, each with a diverse range of function within the brain. These include cell-cell signalling, receptor interactions, axon guidance, synaptogenesis, glial differentiation, myelination and neurotransmission (Section 1.9.2.4).

The most plausible synthesis of the genetic, neuropathological and epidemiological evidence is that schizophrenia is a disorder where genetically timed abnormalities are expressed during development (Weinberger 1995).

1.6 GENERAL GENETICS

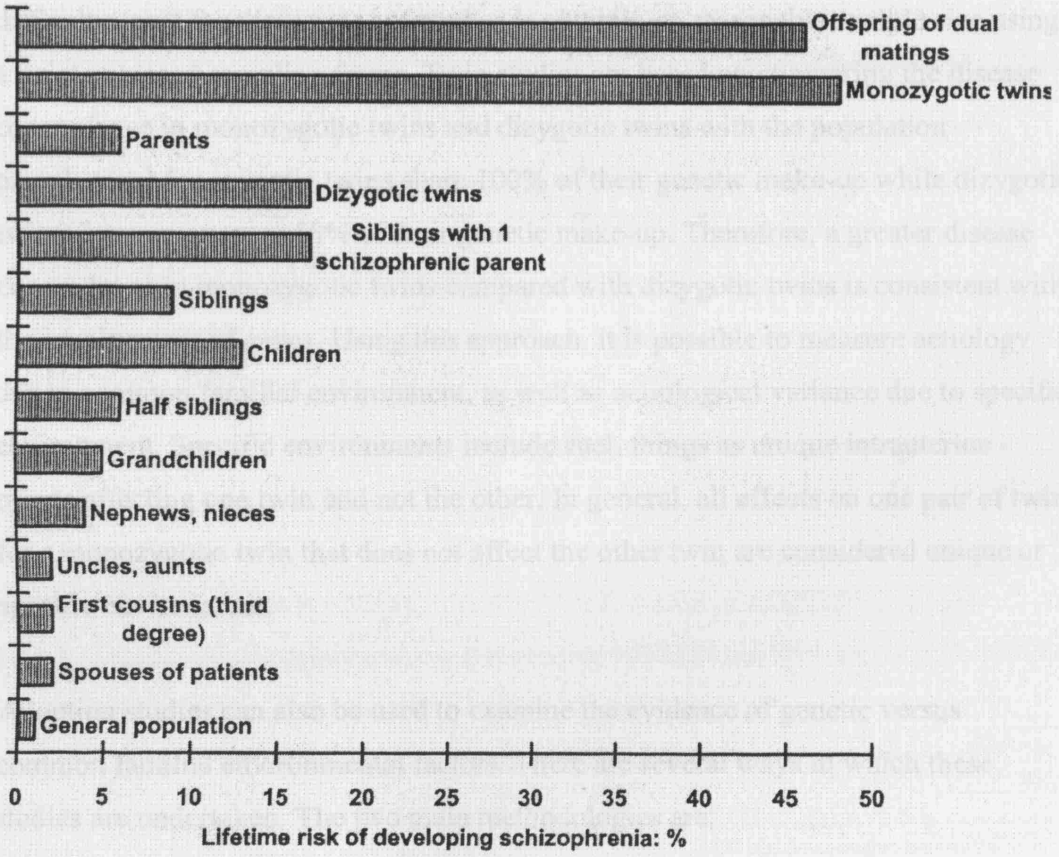
The evidence for genetic transmission of a susceptibility to schizophrenia is compelling. The proportion of variance due to genetic effects of the disease "heritability" has been calculated to above 80% (Sullivan et al. 2003).

1.6.1 FAMILY STUDIES

The genetic hypothesis of schizophrenia was based on early observations that schizophrenia tends to run in families. The first systematic family study was published by Ernst Rüdin in 1916 (Kendler & Zerbin-Rudin 1996). He found out that what was then known as dementia praecox was more common among siblings of affected individuals then in the general population. A larger study of over 1000

individuals by Franz Kallman (1953) showed that both siblings and offspring had increased rates of the disorder. These early studies were flawed due to the unclear diagnostic criteria and the fact that the risk statistics were not age-corrected. This is because many of the siblings/relatives had not completely passed the maximum risk period of 18-35 years. However, Gottesman (1991) calculated the lifetime morbid risk of schizophrenia corrected for age from previously published data using modern diagnostic criteria (Figure 1.1). The results show that the risk for schizophrenia among relatives of an affected individual increases as a function of familial relatedness.

Figure 1.1: Lifetime risk of developing schizophrenia (Gottesman 1991).



Family studies have provided an important validation of Kraepelin's original theory that dementia praecox and manic depressive psychosis are separate disorders. These studies have repeatedly shown that the biological relatives of parents with

schizophrenia have increased risk of schizophrenia and schizophrenia spectrum disorders, whereas the biological relatives of parents with major affective disorder have an increased risk for the affective disorder. The separation is not complete but is supportive of the two disorders as independent disease entities. Clearly, there are some families where schizophrenia comes from one side of the family and bipolar from the other. This can be due to assortative mating by social homogamy for mental diorders.

1.6.2 TWIN AND ADOPTION STUDIES

Twin and adoption studies aim to provide the opportunity to separate the role of gene's common familial environment and specific or unique environment. The most difficult aspect for these types of studies is obtaining a reasonable sample size using a strict unbiased sampling frame. Twin studies are based on comparing the disease concordance in monozygotic twins and dizygotic twins with the population prevalence. Monozygotic twins share 100% of their genetic make-up while dizygotic twins share on average 50% of their genetic make-up. Therefore, a greater disease concordance in monozygotic twins compared with dizygotic twins is consistent with the involvement of genes. Using this approach, it is possible to measure aetiology due to common familial environment, as well as aetiological variance due to specific environment. Specific environments include such things as unique intrauterine events effecting one twin and not the other. In general, all effects on one pair of twin for a monozygotic twin that does not affect the other twin are considered unique or specific environment.

Adoption studies can also be used to examine the evidence of genetic versus common familial environmental factors. There are several ways in which these studies are undertaken. The two main methodologies are:

Parent as proband:

In these studies, rates of illness are compared among adopted offspring whose parents (either mother or father) have the disease in question, with those adopted

offspring whose parents are healthy. The historical study by Heston (1966) found dramatic differences in rates of schizophrenia among adoptees whose mothers were suffering from schizophrenia (5 out of 47) compared to adoptees whose non-biological mothers were suffering from schizophrenia (0 out of 50).

Adoptee as proband:
In these studies, adopted children who become ill are termed as cases. The rates of illness are then compared between their biological and adoptive families.

1.6.2.1 Twin Studies in Schizophrenia

Early German twin studies conducted before the Second World War, found concordance in monozygotic twins to be in excess of 65% (Kringlen 2000). These investigations were criticised for having potential sources of bias. The figures for dizygotic twins ranged from about 6% to 20%. Further studies performed more recently have reproduced the finding of greater concordance in monozygotic twins than dizygotic twins. The Maudsley Twin Study of Gottesman and Shields (1972) has been reassessed using a variety of modern methods and produced an estimate of 67% concordance for monozygotic twins (Farmer et al. 1984; McGuffin et al. 1984). Even this is an underestimate because it is now known that so called "normal" co-twins have volumetric abnormalities at scanning, as well as abnormal eye tracking, EEG and mild thought disorder.

Using modern diagnostic criteria, five recent methodically established studies reported monozygotic concordances between 41% to 65% compared with dizygotic concordances that were between 0% to 28%. This produces an estimated broad heritability of 85% (Cardno & Gottesman 2000). Broad heritability reflects all possible genetic contributions to a population's phenotypic variance and is calculated using the equation shown in Figure 1.2.

Figure 1.2: Calculation of broad heritability.

$$H^2 = \frac{Var(G)}{Var(P)}$$

H^2 is the Broad heritability, Var (G) is the genetic variance of the phenotype and Var (P) is the total phenotypic variance.

1.6.2.2 Adoption Studies in Schizophrenia

Classic adoption studies looked at schizophrenic parents who had their offspring adopted away into adoptive families (Heston 1966). The rates of illness in the adoptees of these schizophrenic probands were compared to that in adoptees of non-schizophrenic parents. The increased rates of illness in children of schizophrenic probands clearly illustrated that transmission of the disease within families was under considerable genetic control (Gottesman & Shields 1967; Kety et al. 1971; Rosenthal et al. 1971; Gottesman & Shields 1976; Kety et al. 1976).

A study researching the effect of being raised by a psychotic adoptive parent was undertaken by Wender et al. (1974). In this cohort study, adoptees with a schizophrenic biological parent but raised by normal adoptive parents, were compared with a group of adoptees born to normal parents but adopted by parents who subsequently became schizophrenic. The adoptees that have schizophrenic biological parents had high rates of spectrum disorders, whereas those who have normal biological parents had low rates.

As mentioned previously, one of the difficulties with adoption studies is that any child of an affected parent adopted shortly after birth will still have experienced the prenatal and perinatal environment provided by their biological mother. In an attempt to overcome the criticism of failing to exclude the possibility of unique common environmental influences in utero, Kety et al. (1976) studied the rate of illness in a group of paternal half-siblings of schizophrenia adoptees. They showed an increased incidence of schizophrenia in paternal half-siblings not attributable to prenatal and perinatal effects. Kety et al. (1994) then replicated their earlier work

and found that schizophrenia was exclusively in the biological relatives of adoptees with the disorder. The prevalence overall was shown to be 10 times greater than that in the adoptive relatives of schizophrenic adoptees.

Confirmation of earlier classical studies was recently confirmed in a long-term study report of adoptees from Finland (Tienari et al. 2000). Further analysis of the Finnish subjects have found that adoptees at high genetic risk of schizophrenia experienced more adverse family effects compared to adoptees at low genetic risk for schizophrenia or the schizophrenia-spectrum disorders (Tienari et al. 2004). This supports both genetic and environmental effects, but makes the assumption that the adverse family environment was not caused by the presence of the schizophrenic adoptee in the adoptive family.

1.7 GENETIC TRANSMISSION MODELS

Family, twin and adoption studies demonstrate a substantial genetic contribution to the aetiology of schizophrenia. Once familial aggregation with a probable genetic aetiology has been established for a trait, evaluation of whether minor or major genes contribute to the phenotypic expression is required. Segregation analysis is considered to be one of the most established methods for use of this purpose. It aims to determine the transmission pattern of the trait within families and to test this pattern against predictions from specific genetic models. Maximum likelihood statistics are used to compare the observed pattern occurring compared to the population base rate.

Segregation analyses for schizophrenia have shown both major gene and polygenic effects. The most plausible models that fit the data from the majority of studies are the generalised single major locus model and the multifactorial polygenic model. The relative validity of these different models has been investigated in several studies and the findings of these have been reviewed extensively (Baron 1986; Kendler et al. 1993). In a more recent review, Baron (2001) states that a single major locus is unlikely as a common mode of inheritance, whereas oligogenic or polygenic

models are plausible alternatives. However, because of the fact that schizophrenia has been proven to be linked to many chromosomal loci, a phenomenon known as locus heterogeneity, the studies have produced results that must be deemed as invalid (Winokur et al. 1982; McGue et al. 1985).

1.7.1 SINGLE MAJOR LOCUS (SML) MODEL

This model is often based on the assumption that the inheritance of a disorder is a consequence of a single locus with two alleles (Elston & Campbell 1970). The concepts of reduced penetrance and phenocopies are introduced to account for the deviations from classical Mendelian inheritance. Several analyses have demonstrated that the SML model is sufficient to correctly predict the data from twin and family studies (Slater 1971; Kidd 1973), whilst other analyses found it insufficient (Baron et al. 1982; Risch & Baron 1984; McGue et al. 1985).

1.7.2 POLYGENIC / OLIGOGENIC MODELS

A polygenic or oligogenic model is one where two or more disease alleles at two or more distinct loci are needed before a disease is expressed. Such models can be used to explain the recurrence of many common diseases. However, methodology in the area is poorly developed. Further understanding of how two or more loci interact will depend on the accurate identification of each locus. The oligogenic model assumes that only a few may act additively, interactively or multiplicatively on the aetiology of the illness.

Risch (1990a) assessed the compatibility of multilocus models with the observed recurrence risks in schizophrenia and suggested that there should be multiplicative effects from at least three loci acting on the risk for this illness. This is based on the assumption of a single subtype of schizophrenia with an equal effect size from oligogenes in every individual. The evidence to date is incompatible with this belief. Therefore, multiple models of transmission are more likely with heterogeneity in the oligogenes that increase susceptibility to schizophrenia.

1.7.3 MULTIFACTORIAL – POLYGENIC (MFP) MODEL

Falconer's method of partitioning genetic liability has been widely adopted for the multifactorial-polygenic model. Here, the trait is assumed to be the result of many genes that have additive effects and the unspecified environmental factors amount to a normal distribution in the general population (Falconer 1965). The disease is manifested when the liability exceeds a certain threshold, where all individuals above this point are affected and those below are normal. Relatives have a greater risk for a genetic disorder than that of the general population and their mean liability is higher. This model has produced a good fit with the observed risk in several family studies of schizophrenia, especially when environmental factors are taken into account (Gottesman & Shields 1967; Kidd 1973; Rao et al. 1981; McGue et al. 1985). However, it has also been rejected in several other studies (Matthysse & Kidd 1976; Baron et al. 1982; Tsuang et al. 1983).

1.7.4 MIXED MODEL

While early segregation analyses considered SML or polygenic models, more recent approaches have considered a mixed model where the phenotype may result from a combination of these two (Lalouel et al. 1983). These models are also invalidated in the presence of locus heterogeneity.

1.7.5 OTHER MODELS

Other models include the two-locus theory, where the phenotype results from two separate loci interacting, and a polygenic model, with graduated gene effect to allow variable liability contributions from the different loci (Matthysse et al. 1979).

Segregation analyses have produced controversial results among the different studies. This is mainly due to limitations in the genetic models examined as well as diagnostic uncertainties. None of the models used take into account the proven presence of locus heterogeneity. Attempts to model assortative mating, reduced fertility and social isolation have been made. The limitations in these analytical methodologies are further compounded by variable/unknown penetrance,

phenocopies, diagnostic difficulties, sampling bias, ascertainment bias, mortality, variable age of onset and the lack of enough family data to estimate the large number of unknown parameters required to accurately model the complex trait.

It is clear that schizophrenia is a complex disorder that does not always show a clear pattern of Mendelian inheritance in families. It is possible that both dominant and recessive transmission is plausible for different subtypes. Yet further subtypes may be oligogenic. As for other common disorders, a mixture of different genetic and non-genetic subtypes is highly likely. Some of these subtypes may be influenced by the environment or by multiple genes. Classical segregation analysis does not have the power to identify a definitive mode of transmission for this disorder and will never do so in the future due to the drawbacks mentioned. In order to overcome the limitations of segregation analysis, geneticists turned to linkage and association/linkage disequilibrium (LD) studies using several different types of genetic markers. LD is the non-random association of alleles at two or more loci, not necessarily on the same chromosome.

A disease may well be the result of interactions between different genes. These interactions can be direct and additive, multiplicative or epistatic where a major gene effect is modified by another locus. The interactions could cause difficulties in finding genes involved in a disease, particularly complex ones. If the effect of one locus is altered or masked by the effects at another locus, power to detect the first locus is likely to be reduced. Thus, elucidation of the joint effects at the two loci will be hindered by their interaction. If more than two loci are involved, the situation is likely to be further complicated by the possibility of complex multiway interactions among some or all of the contributing loci (Cordell 2002).

1.8 METHODS OF MAPPING

1.8.1 GENETIC MARKERS

A genetic marker is a known DNA sequence that can be identified by a simple assay. They can be used to study the relationship between an inherited disease and its genetic cause e.g. a particular mutation of a gene that results in a defective protein. It is known that pieces of DNA that lie near each other on a chromosome tend to be inherited together. This property enables the use of a marker, which can then be used to determine the precise inheritance pattern of the gene that has not yet been exactly localised.

Genetic markers have to be easily identifiable and associated with a specific locus. They also have to be highly polymorphic because homozygotes do not provide any information. The DNA sequence of a genetic marker may be of short e.g. SNP, or long e.g. microsatellites.

1.8.1.1 Restriction Fragment Length Polymorphisms (RFLPs)

RFLPs were discovered in 1978 (Kan 1978). They became the first genetic markers to be successfully used in finding the cause of a disease when they linked Huntington's disease to the tip of the short arm on chromosome 4 (Gusella et al. 1983). RFLPs are centred on a single base pair change that creates or destroys a cleavage site for a specific restriction enzyme. This results in variation between individuals that can be detected by DNA digestion when using the appropriate restriction enzyme. The RFLPs are inherited as simple Mendelian co-dominant markers that can be readily identified in families. A disadvantage of RFLPs is the fact that they are biallelic. Therefore, they are not very informative with low heterozygosity in linkage studies (usually < 0.4).

1.8.1.2 Variable Number of Tandem Repeats (VNTRs)

The heterozygosity of genotyping markers was greatly increased in 1987 through the identification of polymorphic repeat sequence known as VNTRs (Nakamura et al.

1987). These markers are made up from a specific set of consensus sequences that vary between 14-100 bp in length. They are remarkably polymorphic, with a high rate of heterozygosity in the population (usually > 0.6). However, there are only a small number of VNTRs available and they have limited distribution in the genome, often tending to cluster towards the telomere.

1.8.1.3 Minisatellite Polymorphisms

A second class of repeat sequence, known as a "minisatellite", is less complex than a VNTR (Jeffreys 1985). They are more polymorphic than VNTRs and are therefore used to identify individuals uniquely in forensic studies. Minisatellites consist of repetitive, generally GC-rich, variant repeats that range in length from 10 to over 100 bp.

1.8.1.4 Short Tandem Repeats (STRs) or Microsatellite Polymorphisms

STRs or microsatellites are distributed widely and evenly throughout the genome (Litt & Luty 1989; Weber & May 1989). They sometimes have high heterozygosities (usually > 0.7) and are relatively easy to score in the laboratory. The number of repeats at any one locus is limited and mainly consists of two base (dinucleotide) repeats. The dinucleotide (CA)n, or it's complementary sequence (GT)n, is the most common repeat with a highly polymorphic form of the repeat occurring approximately every 0.4 cM. Following the success of dinucleotide repeat markers, other polymorphic markers were isolated and characterised. These include tri- (three) and tetra- (four) nucleotide repeat markers.

1.8.1.5 Single Nucleotide Polymorphisms (SNPs)

In recent years, attention has focused on the use of SNPs as genetic markers. They are the most frequent form of human DNA variation and represent a position in which two alternative bases occur at a considerable frequency (>1%) in the human population (Wang et al. 1998). On average they occur at a rate of 1 per 300 to 1000 base pairs (Collins et al. 1997). Although individual SNPs are less informative than typical multi-allelic simple sequence length polymorphisms, they are more abundant

and their genotyping can be automated with the use of DNA chip-based microarrays (Hacia et al. 1996). This allows for rapid analysis of a very large numbers of SNPs.

Huge collections of mapped SNPs have been developed to provide a powerful tool for human genetic studies. The International HapMap project is one of several such ventures (HapMap 2003). By collating SNP data from researchers in Canada, China, Japan, Nigeria, United Kingdom and the USA, the International HapMap Project provides a haplotype map of the whole human genome with genotyping data on sets of individuals from around the world. This information gives LD patterns across the genome for the different ethnic groupings, enabling researchers to undertake association studies using so called "tagging SNPs". These tagging SNPs provide the same information as many other SNPs in adjacent regions of high linkage disequilibrium. Therefore, it has become possible to identify genetic variation without genotyping every single SNP in a chromosomal region. Tagging SNPs enable investigators to feasibly search the whole genome for candidate genes in large scale association studies.

1.8.2 POSITIONAL CLONING / FINE MAPPING

Positional cloning involves identifying disease genes through their genomic position rather than through their function (Collins 1992). The power of positional cloning lies in the fact that it is conducted in the absence of any suggestive knowledge as to the disease pathogenesis. It is seen by some as a "pure" genetics approach (Gurling et al. 1989). Positional cloning is preceded by fine mapping. The methods adopted include:

1) Linkage followed by association analysis e.g. fine mapping, which is used to detect linkage disequilibrium between the disease and a genetic marker.

2) Cytogenetic analysis where genes are identified through the discovery of a chromosomal abnormality significantly associated with the disease e.g. DISC1 and DISC2 (Section 1.9.2.1).

1.8.3 LINKAGE ANALYSIS

Linkage analysis relies upon the ability to detect the co-segregation of marker alleles with those of the disease gene. For a marker to be "informative" for a particular meiosis, the individual concerned must be heterozygous at both the marker and disease locus. Therefore, the usefulness of a given genetic marker in linkage analysis depends upon the frequency with which it is heterozygous within a population. Microsatellites give investigators the use of regularly spaced polymorphic markers across the whole genome. Linkage studies investigate the departure from independent assortment of a marker and the disorder, by estimating the amount of recombination between the marker and the disease. This recombination results from crossover between the homologous pair of chromosomes during meiosis. Loci where crossover has occurred are known as recombinants, whereas loci without the crossover are known as non-recombinants.

Linkage involves calculating how much the recombination fraction, the number of recombinants divided by the total number of offspring, is significantly different from 0.5, the expected value on the null hypothesis of no linkage.

If two loci are on different chromosomes or are very widely separated on the same chromosome, then independent assortment takes place and the recombination fraction is 0.5 or ½. Thus, there would be a 50:50 chance of loci being either recombinants or non-recombinants. When two loci are close together, or linked, the assortment is no longer independent and recombination fractions of less than 50% are seen. The unit of linkage is known as the centiMorgan (cM) and is referred to as the distance between two loci determined by the frequency with which recombination occurs between them. Two loci are said to be 1 cM apart if recombination is observed between them in 1% of meioses. The number of base pairs this corresponds to varies widely across the genome because different regions of a chromosome have different tendencies towards crossover. In humans, 1 cM on average physically represents a distance of about 7.5×10^5 bp (Scott et al. 2004).

Morton (1955) showed that calculating the LOD (log of the odds ratio) score represents the most efficient statistic for evaluating pedigrees for linkage. A LOD score is the common log of the likelihood that the recombination fraction has a certain value, θ', divided by the likelihood that θ is ½:

$$\text{LOD}(\theta') = \log_{10}(\text{likelihood}(\theta = \theta') / \text{likelihood}(\theta = 0.5)$$

Traditionally a LOD ≥ 3 is taken as "significant" evidence for linkage although this does not equate to the meaning of statistical significance in other contexts.

Linkage analysis was conventionally undertaken using a two-point analysis where the co-segregation of the disease and a single marker was studied. This method was extended to multipoint linkage analysis where multiple markers lying reasonably close together on the same chromosome were studied (Ott 1991). Thus, allowing for greater precision in the location of the disease locus. However, it is more prone to false negative results from genotyping errors.

A linkage study relies on analysing sets of related affected subjects and is capable of detecting a disease gene over a relatively large range. Approximately 400 highly polymorphic microsatellite markers are sufficient to carry out screening of the whole genome. This means that linkage studies can be used to provide initial localisation of disease genes even without prior information on the type of gene that may be involved or its chromosomal location.

1.8.3.1 Parametric Methods

Standard LOD score linkage analysis is undertaken using a parametric method. It requires a precise genetic model that details the mode of inheritance, disease and marker allele frequencies and also disease penetrance. The analysis is carried out either as two-point, the marker and the disease, or multipoint, the disease with more than one marker. Parametric methods have been accused of causing problems for

non-Mendelian conditions, but simulation of two locus transmission and uncertain penetrance have shown it to be a robust approach (Vieland & Huang 2003). The parametric system has a lot more power to detect linkage than the non-parametric sib-pair method, which is advocated when penetrance is unknown.

The classical threshold for establishing significant linkage, LOD $\geq$ 3, was derived for single marker tests. Modifications to this are required when genomic screening with a greater number of microsatellite markers. There is no general agreement on the most suitable approach (Schmidt 2006). Lander and Kruglyak (1995) suggested using LOD scores and P values that are slightly more stringent, as evidence for linkage of Mendelian traits, when multiple markers are typed. They proposed four levels of stringency in estimating the robustness of linkage for complex traits: suggestive, significant, highly significant and confirmed. However, these thresholds have been disputed (Curtis 1996; Witte et al. 1996).

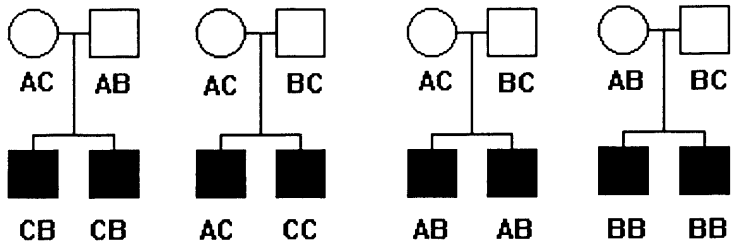
More recently, a "model free" parametric linkage method (MFLINK) was developed by Curtis and Sham (1995). This approach assumes a fixed recombination fraction and estimates a penetrance that maximises the LOD score. The method was used for the linkage analyses that led to the chromosome 11q research in this thesis (Gurling et al. 2001).

1.8.3.2 Non-Parametric Methods

Model free or non-parametric methods of linkage analysis test the hypothesis that alleles or chromosomal segments are shared by affected individuals (Risch 1990b; Risch 1990c; Risch 1990d). The simplest example of this is the sib-pair method whereby allele sharing between pairs of affected relatives is examined (Figure 1.3). Taking pairs of affected siblings, we would expect by chance that two alleles of a DNA marker would be shared 25%, one allele 50% and no alleles 25% of the time. However, if the marker is linked to the disease gene then alleles will be shared between affected sib-pairs more often than expected. If parents are also genotyped, then the inheritance of the marker alleles can be studied directly. This is known as

identity by descent (IBD). Even if the parents are unavailable, one can use population allele frequencies to estimate whether increased allele sharing is occurring. The strength of evidence in favour of linkage can be calculated with a chi-squared statistic or by a maximum likelihood score (MLS), the latter being similar to a LOD score.

Figure 1.3: Sib-pair analysis.



The second sib-pair shares one allele IBD, while all the others share both alleles. This suggests that the gene is closely linked to the marker. Figure courtesy of Dave Curtis, http://www.smd.qmul.ac.uk/statgen/dcurtis.html.

A well recognised drawback of affected sib-pair analysis is the difficulty in calculating the power of a proposed study. Such calculations require knowledge about the overall genetic effect to be observed as well as the role of individual genes, their interactions with each other and the interactions with environmental factors. This is further complicated if the disease under consideration shows locus heterogeneity. The presence of multiple disease susceptibility loci may make it difficult to collect a large enough sample of affected sib-pairs that is adequate to detect linkage of disease susceptibility loci to marker loci (Sribney & Swift 1992). Therefore, it is hard to estimate in advance the number of affected sib-pairs required for recruitment (Hauser 2006).

1.8.4 ALLELIC AND HAPLOTYPIC ASSOCIATION STUDIES

Case-control and family based association studies represent the final common pathway for all genetic linkage studies. They are necessary for identifying alleles that confer susceptibility to specific phenotypes. Association studies play a critical role in analysing genetically complex traits by fine mapping a region already

implicated through linkage studies. These studies are based on finding an allelic association between a marker allele and a disease allele. The observation of allelic association, or those between two adjacent markers, is due to linkage disequilibrium, which is defined as the non-random association of alleles at linked loci.

Linkage disequilibrium between alleles mainly reflects the recombination history in the haplotypic population. Recently acquired mutations, or those in founder or isolated populations with limited chromosome diversity, are thus likely to show linkage disequilibrium extending over long distances. However, as a result of a number of different factors including regional variability in recombination patterns, recent population admixture and local chromosomal composition, linkage disequilibrium may vary significantly within and between populations.

Comparisons are made between the frequency of marker phenotypes in a sample of patients and a sample of "healthy" controls. The statistical analysis used is a 2 x 2 contingency table (Table 1.1).

Table 1.1: Marker-disease association in a population.

Marker	n affected	n unaffected
Present	*a*	*b*
Absent	*c*	*d*

It is also possible to calculate the χ^2 statistic from the 2 x 2 contingency table. This is done using the formula shown if Figure 1.4, with a one degree of freedom (df).

Figure 1.4: Calculation of the χ^2 statistic using a 2 x 2 contingency table.

$$\chi^2 = \frac{(ad-bc)^2(a+b+c+d)}{(a+b)(c+d)(b+d)(a+c)}$$

Reasons for finding association include:

Direct causation:

This is where having an allele results in susceptibility to the disease. Possession alone may not necessarily lead to the development of the disease, but it increases the likelihood.

Population stratification:

The population contains several genetically distinct subsets where both the disease and a particular allele happen to be more frequent in one subset. Lander and Schork (1994) gives an example of this association in the San Francisco bay area between HLA-A1 and the ability to eat with chopsticks. HLA-A1 is more frequent among the Chinese than Caucasians.

Type I error:

Association studies test a large number of markers. Even without true effects 5% will be significant at the P = 0.05 level.

Linkage disequilibrium:

LD between the marker and a disease locus truly exists.

1.8.5 THE PROBLEM OF GENETIC HETEROGENEITY

In the search for genes that cause complex diseases, what originally appeared to be a single disease entity may be caused by multiple different mutations. This phenomenon is referred to as genetic heterogeneity. The heterogeneity may be found at the same genetic locus (allelic or intralocus heterogeneity) or different loci may be involved in producing the same disease entity (non-allelic or interlocus heterogeneity).

The problem of locus heterogeneity has been a cause of difficulty in identifying regions of importance within families for linkage studies. There may be several

different loci where mutations result in similar phenotypes and each mutation may have different modes of transmission. Example of this include Charcot-Marie-Tooth, tuberous sclerosis and retinitis pigmentosa, which all result from a number of distinct mutations in genes from different chromosomes. Charcot-Marie-Tooth exists in an X-linked form with two autosomal dominant forming the loci on different chromosomes. If homogeneity had been assumed and linkage analysis carried out on a collection of families containing a mixture of the two autosomal varieties, the results would have been misleading and linkage might have been overlooked. In retinitis pigmentosa, the identification of mutations in over 30 genes has resulted in genetic reclassification of the retinal degeneration syndrome (Rivolta et al. 2002).

Investigators were initially concerned with heterogeneity when searching for mutations using linkage analysis. The adoption of parametric methods in large families has helped to overcome these difficulties. This enabled studies to maximise the power to find linkage within a single homogenous family. Theoretical work has shown that the parametric LOD approach is capable of detecting linkage in a two locus polygenic disease (Vieland et al. 1992a; Vieland et al. 1992b). Linkage can therefore be detected even when the inheritance is complex and includes both genetic heterogeneity and variable penetrance (Gershon et al. 1990). Non-parametric linkage methods have the advantages of not specifying the mode of transmission, but the disadvantages of not being able to detect heterogeneity.

Common variants contributing to the genetic risk for common diseases are known to exist e.g. APOEε4 in Alzheimer's disease. Thus, giving the common disease-common variant (CD-CV) hypothesis. However, some diseases have been shown to be due to a heterogeneous collection of rare mutations. This latter type of mutation is referred to as the rare allele hypothesis and occurs through the presence of multiple "private" mutations that are not found in sequence databases.

In a review of breast cancer, two familial breast cancer genes, BRCA1 and BRCA2, showed at least 1,200 and 1,400 different mutations of large effect respectively.

Each mutation was rare except in the founder populations (Wright et al. 2003). However, only one of the several common SNPs located within the BRAC1 and BRAC2 coding regions exerts a marginal increase (1.3-fold) in breast cancer risk. Therefore, only the rare mutations in the BRCA1 and BRCA2 have an effect and the common SNPs in the coding region have little or no functional effect on the disease.

Reich (2001) looked at the population genetics and various examples of common diseases e.g. Alzheimer's disease (APOEε4), to conclude that the CD-CV model is a good predictor of disease alleles in complex disorders. However, studies of hypercholesterolemia favoured multiple rare variants as being aetiological in three genes (Lohmueller et al. 2003).

1.9 LINKAGE AND ASSOCIATION RESULTS IN SCHIZOPHRENIA

1.9.1 LINKAGE RESULTS

Meta-analysis combine the results of several independent studies that address a set of related research hypotheses. There have so far been two meta-analyses of linkage studies undertaken in schizophrenia (Badner & Gershon 2002; Lewis et al. 2003). The rank-based genome scan meta-analysis (GSMA) method used by Lewis et al. (2003) is set out in Levinson et al. (2003). Here, markers were placed within bins covering every 30 cM of the genome excluding the sex chromosomes. No assumptions were made about the model of inheritance or of genetic heterogeneity. The method used by Badner and Gershon (2002) is known as multiple scan probability (MSP). This is where P values from published genome scans of schizophrenia were combined after correcting for size of the linkage area. The region showing significant LOD scores where both studies agree upon are chromosome 8p and chromosome 22q (Table 1.2).

Table 1.2: Meta-analyses of genomewide linkage studies in schizophrenia.

Lewis et al. (2003)	Badner and Gershon (2002)
1p13.3-q23.3	
1q23.3-31.1	
2p12-q22.1	
2q22.1-q23.3	
3p25.3-22.1	
5q23.2-34	
6pter-22.3	
6p22.3-21.1	
6q15-23.2	
8p22-21.1	8p
10pter-14	
11q22.3-24.1	
	13q
14pter-q13.1	
15q21.3-26.1	
16p13-q12.2	
18q22.1-ter	
20p12.3-11	
22pter-q12.3	22q

The analysis by Lewis et al. (2003) most strongly favoured a 2p12-q22.1 region, but also found that the number of loci meeting the aggregate criteria for significance was much greater than the number of loci expected by chance ($p < 0.001$). Support was also obtained for regions, in weighted rank order, on chromosome 5q23.2-34, 3p25.3-22.1, 11q22.3-24.1, 6pter-22.3, 2q22.1-23.3, 1p13.3-q23.3, 22pter-q12.3, 8p22-21.1, 6p22.3-21.1, 20p12.3-11 and 14pter-q13.1. The greater consistency of linkage results than in previous recognised studies suggests that some or all of these areas contain loci that increase susceptibility to schizophrenia in diverse populations. It is worth mentioning that the region most strongly supported by Lewis et al. (2003), chromosome 2q, is not one that has received strong support previously. The analysis by Badner and Gershon (2002) found the strongest evidence for susceptibility loci on chromosome 8p, 13q and 22q.

1.9.2 ASSOCIATION RESULTS

1.9.2.1 Chromosome 1

DISC1 and DISC2 (disrupted in schizophrenia 1 and 2):

A Robertsonian translocation involving chromosome 1 and 11, t(1;11)(q42.1; q14.3), was identified in a large Scottish pedigree (St Clair et al. 1990). Segregation of this translocation was observed through an extended pedigree (Section 1.10). Clinical studies found schizophrenia, major depression and bipolar disorder in translocation carriers (Blackwood et al. 2001). Two novel genes within the chromosome 1q42.1 region, DISC1 and DISC2, were found to be directly disrupted by the translocation (Millar et al. 2000). DISC1 is associated with numerous cytoskeletal proteins involved in centrosomal and microtubule function. DISC2 transcribes a non-coding RNA molecule that is antisense to DISC1. It is thought that this antisense RNA may regulate expression of the sense gene. Suggestions are made that DISC1 and DISC2 may confer susceptibility to psychiatric illnesses such as schizophrenia.

RGS4 (regulator of G-protein signalling 4):

Transcription of RGS4 was shown to decrease in schizophrenic patients' post-mortem brains following a microarray based gene expression study (Mirnics et al. 2001). Genetic analysis was then used to evaluate the possible role of the chromosome 1q23.3 RGS4 gene in schizophrenia. Significant associations involving four SNPs were identified within a 10 kb span of RGS4 (Chowdari et al. 2002). Several independent replications have been reported to display haplotypic association (Chen et al. 2004; Morris et al. 2004; Williams et al. 2004b). Studies that have failed to confirm support for allelic and haplotypic association have also been published and a recent meta-analysis was inconclusive (Sobell et al. 2005; Rizig et al. 2006; Talkowski et al. 2006). RGS4 is biologically plausible as a schizophrenia gene. It functions as a GTPase activator that negatively modulates G protein-mediated signalling via dopamine and glutamate receptors.

CAPON (officially known as NOS1AP, nitric oxide synthase 1 adaptor protein):
Brzustowicz et al. (2004) first reported an association of novel candidate gene CAPON on chromosome 1q23.3. Using a sample of 24 Canadian familial-schizophrenia pedigrees, significant evidence of linkage disequilibrium between schizophrenia and eight markers were produced. All of the markers exhibiting significant LD to schizophrenia fall within the genomic boundary of the CAPON gene, making it a prime positional candidate for schizophrenia. CAPON encodes a protein nitric oxide synthase 1 adaptor protein and is thought to be involved in signal transduction within the N-Methyl-D-Aspartate receptor system. The significant association finding was not replicated by Puri et al. (2006).

UHMK1 (U2AF homology motif kinase 1):
Puri et al. (2007) reported that the chromosome 1q23.3 UHMK1 gene increases susceptibility to schizophrenia. Positive allelic association with schizophrenia was found with three microsatellite markers and two SNPs. These markers are localised within the 700 kb region between RGS4 and CAPON, in the UHMK1 gene. UHMK1 is a serine threonine kinase nuclear protein that plays a role in phosphorylating certain proteins.

PLXNA2 (plexin A2):
Mah et al. (2006) recently undertook a genomewide association scan with the use of over 25,000 SNPs, located within approximately 14,000 genes. Association to schizophrenia with SNP rs752016 was found on chromosome 1q32.2. This marker is situated within the PLXNA2 gene, thus making it a novel schizophrenia candidate gene. PLXNA2 is a member of the transmembrane semaphorin receptor family involved in axonal guidance during development. It is thought to modulate neuronal plasticity and regeneration.

1.9.2.2 Chromosome 5

Epsin 4 (officially known as CLINT1, clathrin interactor 1):

Pimm et al. (2005) found significant evidence of LD with schizophrenia and four markers located at the 5' end of the chromosome 5q33.3 Epsin 4 gene. Replication using 308 Han Chinese trios showed significant haplotypic association (Tang et al. 2006). In another replication study, this time involving a case-control sample, significant allelic association with schizophrenia was found with SNP rs1186922 (Liou et al. 2006). A haplotypic association that survived permutation correction was also positive. The Epsin 4 gene encodes for the clathrin-associated protein, enthoprotin. This molecule functions in the transport and stability of neurotransmitter vesicles, at the synapses and within neurons.

1.9.2.3 Chromosome 6

Dysbindin (officially known as DTNBP1, dystrobrevin binding protein 1):

Straub et al. (2002) performed family based association analysis and found SNPs within the chromosome 6p22.3 Dysbindin gene that were associated with schizophrenia. This finding has been replicated by several other groups (Schwab et al. 2003; Kirov et al. 2004; Numakawa et al. 2004; Williams et al. 2004a). However, a recent review of Dysbindin association studies in independent European samples found inconsistencies in the alleles and haplotypes associated with schizophrenia (Mutsuddi 2006). The authors cast doubt on the fine mapping of this locus concluding that evidence for association is vague and insufficient. Dysbindin is thought to be involved in trafficking of receptors and signal transduction proteins.

1.9.2.4 Chromosome 8

NRG1 (neuregulin 1):

A genomewide scan in Icelandic families showed that a schizophrenia locus maps to chromosome 8p21-22 (Stefansson et al. 2002). Extensive fine mapping of this locus and haplotypic association analysis identified NRG1, on chromosome 8p21-22, as a novel schizophrenia candidate gene (Stefansson et al. 2002). Further individual studies have also described positive association of NRG1 with schizophrenia

(Stefansson et al. 2003; Williams et al. 2003b; Yang et al. 2003b; Bakker et al. 2004; Corvin et al. 2004; Li et al. 2004b; Zhao et al. 2004; Petryshen et al. 2005). Recently, a meta-analysis of studies showing association with NRG1 found no evidence for the gene being associated with schizophrenia (Munafo et al. 2006), while two association studies reported negative findings (Iwata et al. 2004; Thiselton et al. 2004). Law et al. (2006) have described an interesting mRNA variation in post-mortem hippocampi associated with a single SNP from a risk haplotype. NRG1 is expressed at CNS synapses and has a clear role in the expression and activation of neurotransmitter receptors (Stefansson et al. 2002).

PPP3CC (protein phosphatase 3, catalytic subunit, gamma isoform):
A transmission disequilibrium study in a large sample of affected families detected association of the PPP3CC gene with schizophrenia (Gerber et al. 2003). PPP3CC, located at chromosome 8p21.3, encodes a calcineurin gamma catalytic subunit. It is thought that alterations in calcineurin signalling may contribute to the development of schizophrenia. This is because it plays a key role in the downstream regulation of dopaminergic signal transduction. However, association of this potential schizophrenia susceptibility gene was not replicated in another study (Kinoshita et al. 2005).

FZD3 (frizzled homolog 3):
The FZD3 protein is a transmembrane receptor essential in neurodevelopment. Human FZD3 gene is mapped to chromosome 8p21.1 and was shown to be associated to schizophrenia using 246 schizophrenic family trios of Chinese Han descent (Yang et al. 2003a). Nevertheless, the finding that FZD3 may confer susceptibility to the disease was not replicated by Wei et al. (2004) and Hashimoto et al. (2005).

PCM1 (pericentriolar material 1):
Identification of a chromosome 8p susceptibility gene was carried out by finding linkage disequilibrium between genetic markers and schizophrenia in a UK case-

control sample and in a USA trio sample (Gurling et al. 2006). Significant allelic association were found with markers in the chromosome 8p21.3-22 PCM1 gene. In the same study, cases who had inherited a PCM1 genetic susceptibility showed a significant relative reduction in the volume of orbitofrontal cortex grey matter compared to patients with non-PCM1 associated schizophrenia. Further confirmation of this gene being implicated in susceptibility to schizophrenia has been obtained from an Aberdeen sample (Datta et al. 2008).

1.9.2.5 Chromosome 11

DRD2 (dopamine receptor D2):

DRD2 receptors are G protein-coupled receptors localised on postsynaptic dopaminergic neurons. They have been identified as targets for antipsychotic drugs used to treat schizophrenia. A family based association study between a functional polymorphism in the promoter region of the DRD2 gene and schizophrenia was investigated in a Portuguese population (Schindler et al. 2002). 78 trios were analysed revealing evidence of association to schizophrenia with an insertion allele. However, no mutations in the coding sequence and splice sites of the DRD2 gene, localised to chromosome 11q23.1-23.2, were associated with schizophrenia in an earlier study (Gejman et al. 1994).

BDNF (brain-derived neurotrophic factor):

BDNF, on chromosome 11p14.1, was studied as a risk factor for schizophrenia in a Scottish population that included 321 schizophrenia probands and 350 controls (Neves-Pereira et al. 2005). A val66met SNP showed significant association with schizophrenia. Haplotypic analysis of this SNP with a dinucleotide repeat polymorphism located in the promoter region also revealed highly significant association to schizophrenia. Therefore, although the val66met polymorphism alters gene function, the genetic risk may be dependent on the haplotypic background on which the val66met variant is carried. Krebs et al. (2000) also reported that BDNF was associated with schizophrenia susceptibility but several other groups have failed to replicate the findings (Tochigi et al. 2006; Watanabe et al. 2006).

1.9.2.6 Chromosome 13

G72 (officially known as DAOA, D-amino acid oxidase activator):

Chumakov et al. (2002) obtained good evidence of association to schizophrenia on chromosome 13q33.2. The positive markers localised in and around the G72 gene. This enzyme activator is expressed in the human brain and functions by activating the oxidation of D-serine, a potent activator of N-methyl-D-aspartate-type glutamate receptor. Thus, activation of D-amino acid oxidase by G72 implicates N-methyl-D-aspartate receptor regulation pathway in the genetic aetiology of schizophrenia. A large case-control study found further evidence for association of selected SNPs and a haplotype to schizophrenia (Schumacher et al. 2004), while another study also reported significant association (Korostishevsky et al. 2004).

1.9.2.7 Chromosome 22

Velocardiofacial syndrome (VCFS) is associated with variable clinical symptoms including learning disability, cardiac anomalies, typical facial appearance and in approximately 1-5% of cases, schizophrenic psychosis. It usually arises from small, relatively frequent deletions on the long arm of chromosome 22. Although the exact boundaries of the critical region remain uncertain, most patients (~90%) have a 3 Mb deletion encompassing 30 genes. The associated phenotype is highly variable with congenital heart defects occurring in approximately three-quarters of patients. Pulver et al. (1994) found a high rate of psychosis among patients diagnosed with VCFS and their relatives provided the first possible suggestion that there may be a gene associated with schizophrenia on chromosome 22q11 or that a DNA rearrangement in this region may be important to the aetiology of some form of the disease. The relative risk of schizophrenia in people with 22q11 deletion syndromes is around 2.5-3.0. Through linkage and association studies, two genes in the area stand out as possible candidates for psychiatric outcomes. These genes are COMT and PRODH, both located on chromosome 22q11.2.

COMT (catechol-O-methyltransferase):

COMT is enzyme that degrades catecholamines and their metabolites. It modulates the neurotransmitter function of dopamine by altering the concentration of catecholamines in different parts of the brain. A study showing highly significant association between schizophrenia and a COMT haplotype, using Ashkenazi Jew case-control samples, revealed lower expression of COMT in the human brain (Shifman et al. 2002). However, genotyping errors in this study have now been reported (Munafo et al. 2005).

PRODH (proline dehydrogenase 1):

PRODH is involved in the degradation of the amino acid proline. It catalyses the conversion of proline to pyrroline-5-carboxylate. Liu et al. (2002) was among the first investigators to find positive association between PRODH and schizophrenia. Li et al. (2004a) also reported a positive association using schizophrenic patients from the Sichuan province of China. The sample, comprising of 528 family trios and sib-pairs, gave evidence of association between PRODH markers and schizophrenia in two different haplotypes. However, several investigators including Williams et al. (2003a) have reported negative findings.

APOL1-6 (apolipoprotein L, 1-6):

Other candidates within the region of chromosome 22q11 include the six members of the APOL family of genes, all located in close proximity to each other on chromosome 22q11.2-13.2. The high density lipoprotein family plays a central role in cholesterol transport. The cholesterol content of membranes is important in cellular processes such as modulating gene transcription and signal transduction, both in the adult brain and during neurodevelopment.

1.10 CHROMOSOME 11q22-24 AS A SCHIZOPHRENIA SUSCEPTIBILITY REGION

Chromosome 11 was the site of a cytogenetic abnormality, which was a Robertsonian translocation, t(1;11)(q42.1; q14.3) found in a single large Scottish

pedigree containing 21 cases of mental and/or behavioural disorders (St Clair et al. 1990; Blackwood et al. 1998; Millar et al. 2000). Using cytogenetic analysis, 34 people within the 77 extended family members were found to carry the balanced translocation. Schizophrenia and other psychiatric disorders were then recorded in 16 of the 34 members with the translocation compared to 5 of the 43 individuals without it. Analysis of linkage between the translocation and the disease locus produced LOD scores between 3.1 and 6.0, depending on the disease model applied. The LOD score (against chance linkage of the translocation with mental illness) was greatest when the mental disorders in the phenotype were restricted to schizophrenia, schizoaffective disorder and recurrent major depression. Due to the balanced translocation significantly co-segregating with the disease, St. Clair et al. (1990) implied that the chromosome 11 translocation breakpoint region may be the site of a gene predisposing to major mental illness. However, other linkage studies within the area have been found to be negative (Gill et al. 1993; Wang et al. 1993; Mulcrone et al. 1995).

BDNF and DRD2 have both been suggested as plausible candidate genes for schizophrenia on chromosome 11 (Section 1.9.2.5). Several linkage studies have been undertaken examining markers at these candidate genes and across other aspiring regions of the chromosome (Kalsi et al. 1995; Moises et al. 1995; Faraone et al. 1998; Kaufmann et al. 1998). In one of these linkage studies, thirty individuals from two Japanese pedigrees were genotyped at eight different loci (Nanko et al. 1992). Data analysed using a range of single gene model, from near dominant to intermediate, found positive LOD scores between 1 and 1.5. The modest linkage signal for schizophrenia was reported with the polymorphic marker D11S35. This was confirmed in a separate study where one of four Canadian pedigrees produced a maximum LOD score of 3.41 at the same D11S35 locus (Maziade et al. 1995). Further evidence of linkage to schizophrenia on chromosome 11q23.3-24 was published by Gurling et al. (2001). Here, a genetic linkage analysis of thirteen very large multiply affected families, in which members of three or more generations suffered schizophrenia, was performed. A 3-point LOD score of 3.1 with markers

D11S925 and D11S934 was found. There was also a LOD score of 3.2 in a single large Icelandic pedigree using marker D11S934.

The chromosome 11q22-24 region has been shown to be one of the most well established linkages to schizophrenia by a meta-analysis of twenty genome scans. Here, non-parametric rank-order statistics was employed to show that evidence for linkage at this region was non-random (Lewis et al. 2003). Chromosome 11q22.3-q24.1 was ranked third out of all regions and when taken into account the different sample sizes of the twenty schizophrenia genome scans, it was ranked fourth. This provided considerable confidence that a schizophrenia susceptibility locus was likely to exist within the chromosome 11q22-24 vicinity. Due to significant genomewide evidence for linkage, it was justifiable to investigate this area to detect linkage disequilibrium. Therefore, an association study was applied in order to localise a schizophrenia susceptibility gene in the chromosome 11q22-24 region.

2.0 MATERIALS AND METHODS

2.1 RECRUITING BLOOD SAMPLES

The UCL case-control sample consists of 496 unrelated cases and 488 ancestrally matched controls. All bloods were collected in 8 ml EDTA or citrate plastic tubes and stored at -80°C until DNA extraction. For confidentiality, patient names were kept anonymous and samples were identified through a numbering system. Research subjects were selected only if both parents were of English, Scottish or Welsh descent, with at least three grandparents having the same origins. The fourth grandparent was permitted if he/she was of another Caucasian European origin but not of Jewish or pre-2004 non-European Union ancestry. These data were recorded in an ancestry questionnaire, with confirmation from family histories noted on medical records. The ancestry selection criteria is not meant to be discriminatory, but was carried out to ensure that the observed genetic differences are disease related and not ancestry related.

U.K. National Health Service (NHS) multicentre and local research ethics committee (LREC) approval was obtained and all participating subjects signed an approved consent form after reading the information sheet. Each schizophrenic research subject had received a diagnosis and assessment by NHS psychiatrists as part of their routine clinical diagnosis and treatment. Those with short-term drug-induced psychoses, psychoses with either learning disability or head injury and other symptomatic psychoses were excluded. Schizophrenic subjects were recruited on the basis of having an International Classification of Diseases version 10 (ICD10) diagnosis of schizophrenia recorded in medical case history notes after clinical interview by NHS psychiatrists. The diagnoses were confirmed by a senior psychiatrist, usually within one week. Schedule for Affective Disorders and Schizophrenia-Lifetime Version (SADS-L) interview was completed for all cases by a research psychiatrist (Spitzer & Endicott 1977). Schizophrenic subjects were then chosen on the basis of having received a diagnosis at the "probable level" of RDC,

the Research Diagnostic Criteria (Spitzer et al. 1978). Patients with schizoaffective bipolar disorder or schizomania were excluded.

All screened control individuals were also interviewed by a psychiatrist with the SADS-L schedule, specifically for the study. They were included only if they had no present or lifetime history of any RDC-defined mental disorder and for not having a family history of schizophrenia, alcoholism or bipolar disorder according to self-report by the research subject and/or recorded case notes.

The UCL case-control sample has been used in studies of other chromosomal regions and has implicated the chromosome 5q33.3 epsin 4 (CLINT1) gene encoding enthoprotin, the chromosome 8p21.3-22 PCM1 gene and the chromosome 1q23.3 UHMK1 gene as the cause of schizophrenia (Pimm et al. 2005; Gurling et al. 2006; Puri et al. 2007). The sample has also been used to exclude association between schizophrenia and markers at the chromosome 1q23.3 RGS4 and CAPON (NOS1AP) genes (Puri et al. 2006; Rizig et al. 2006).

2.2 CHEMICALS AND REAGENTS

The following section contains information on all chemicals and regents used in experiments.

2.2.1 MILLI-Q SYSTEM

Ultra pure DNase- and RNase-free water (Sigma Cat. No. W4502) was used in all PCR master mixes. Otherwise the water used to make up reagents was pure water produced from the Milli-Q system. This will now be referred to as Milli-Q water. The Milli-Q system works by taking in mains tap water and filtering it. The first step of filtration involves reverse osmosis, were the ions and main contaminants present in tap water are removed. This distilled (deionised) water is then passed into a main holding tank that houses an ultraviolet (UV) light source. The UV light destroys and prevents microorganisms from contaminating the water and also breaks down any organic compounds present via photo-oxidation. Water then moves through another

resin filter to remove any remaining water contaminants. A final small membrane filter (0.22 μm) polishes the water before it is ready for dispensing. A resistivity value greater than 18.2 MΩ cm (25°C) is desirable for purified deionised water.

2.2.2 REAGENTS FOR DNA EXTRACTION

0.5 M ethylene diamine tetraacetate acid (EDTA; BDH 100935V):
186.1 g EDTA was added to 800 ml Milli-Q water and dissolved by adjusting to pH8.0, using sodium hydroxide pellets (Sigma S5881). The solution was sterilised by autoclaving at 15 pounds per square inch (121°C) for 15 minutes. An autoclave tape was used as an indicator that the process has taken place.

2 M tris hydrochloride (Tris-HCL):
121.1 g tris base (BDH 103156X) was dissolved in 400 ml Milli-Q water and adjusted to pH 8.0 using concentrated hydrochloric acid (BDH 10125). The final solution was brought up to 500 ml before being sterilised by autoclaving.

TE buffer:
This buffer solution was made up by combining 10 mM Tris-HCL with 1 mM EDTA. Hence, 5 ml of 2 M Tris-HCL and 2 ml of 0.5 M EDTA were mixed and the solution was brought up to 1000 ml with Milli-Q water. This was then sterilised by autoclaving.

Proteinase K buffer:
50 mM Tris-HCL, 50 mM EDTA and 100 mM NaCl (Sigma S7653) is required to make this buffer. Hence, 25 ml of 2 M Tris-HCL, 100 ml of 0.5 M EDTA and 25 ml of 4 M NaCl were mixed and the solution was brought up to 1000 ml with Milli-Q water before being sterilised by autoclaving.

10X lysis buffer:
5.84 g NaCl was mixed with 37.22 g EDTA in 1000 ml Milli-Q water. The solution was then autoclaved (EDTA is an important component as it removes magnesium

ions that are essential for preserving the structure of cell envelope and inhibits cellular enzymes that could degrade DNA).

3 M sodium acetate:
246.09 g of anhydrous sodium acetate (Sigma S2889) was dissolved in 800 ml Milli-Q water. The pH was adjusted to 5.2 with glacial acetic acid (BDH 10001). The volume was brought up to 1000 ml and then autoclaved.

10% sodium dodecyl sulphate (SDS):
This was made by dissolving 10 g SDS (Sigma L4509) in 100 ml Milli-Q water.

20 mg/ml proteinase K enzyme solution (BDH 39509):
Buffered phenol (Sigma P4457) was mixed with a few flakes of indicator dye 8-hydroxyquinoline (Sigma H6878). This dyes the phenol yellow and becomes brown when the phenol is oxidised, indicating that is no longer useable.

Chloroform (BDH 100776B)

Isoamyl Alcohol (IAA; Sigma S5881), also known as 3-Methyl-1-butanol

Polyvinylpolypyrrolidine (PVPP; Sigma P-6755)

Ethanol (BDH 10107)

Virkon:
This was used to deactivate blood and cleanse work surfaces. The disinfectant was made by dissolving 10 g Virkon (VWR 148-0202) in 1 L water. Once mixed with waste blood, it was safe to dispose of down the sink after being left to stand for 5 minutes.

2.2.3 REAGENTS FOR DNA QUANTIFICATION

Molecular probes picogreen (Invitrogen P7581)

Calf thymus DNA (Sigma D3664), 50 µg of dehydrated DNA

2.2.4 REAGENTS FOR PCR AMPLIFICATION AND SEQUENCING REACTIONS

PCR amplification:

The reagents were supplied by bioline in a BioTaq red polymerase kit (BIO-21041) and included

1) 10X solution NH_4 reaction buffer.

2) 50 mM magnesium chloride.

3) 25 mM solution dNTPs.

4) Taq polymerase red 1 U/µl.

5 M Betaine (Sigma B2629), 29.29 g dissolved in 50 ml of ultra pure water (for stock)

Ultra pure water (Sigma W4502)

For genotyping, M13 Forward (M13F) IRD 700/800 was also added into the master mix (MWG, Ebersberg, Germany).

Sequencing reaction:

microCLEAN (Microzone 2MCL-50) was used to remove master mix reagents, enzymes and primer dimmer, leaving clean DNA template.

Sequencing reagents were obtained from the SequiTherm EXCEL™ II kit (Microzone SE9101LC) and included

1) 3.5X SequiTherm Buffer

2) SequiTherm Taq Polymerase

3) Termination mix di-deoxyNTP

4) Ultra pure water

Clean DNA template with 10 pmol/µl of both M13F 700 and M13R 800 (MWG) were also added into the master mix for each PCR reaction.

2.2.5 REAGENTS FOR GEL ELECTROPHORESIS

Low resolution electrophoresis:
1% agarose (Sigma A9539) with 1X TBE (National diagnostics EC-860)

SequaGel XR (National Diagnostics EC-842) polyacrylamide gels for high resolution genotyping and sequencing were created as follows:
10% ammonium persulphate (APS; Sigma A9164) mixed with a set volume of the buffered acrylamide solution, to allow polymerisation (800 µl of 10% APS to every 100 ml buffered Sequagel solution). The gel was then caste within the 2 mm gap between the two glass plates (25 cm length for genotyping or 44 cm length for sequencing) and was allowed to set for approximately two hours.

Tris-Borate-EDTA (TBE; National diagnostics EC-860):
Used as an electrolyte buffer.

Fuchsin:

This loading buffer was used for both sequencing and genotyping. It was made with 0.05 g Fuchsin (Sigma, S5881) and 50 ml Formamide (Sigma F9037).

Ethidium bromide (Sigma E215):

Used in agarose gels to intercalate with double stranded DNA. This was in order to visualise DNA bands under UV light.

Hyperladder IV (Bioline, London, UK; BIO-33029):

Molecular weight used for agarose gels sizes from 100 bp to 1000 bp, increasing in 100 bp increments.

MicroSTEP20a (Microzone 1N700/80020a):

Molecular weight for polyacrylamide gels. Both IRD700 and IRD800 molecular weights contain markers at 70, 90, 100, 120, 140, 160, 180, 200, 220, 240, 260, 280, 300, 310, 330, 340, 350, 360, 380 and 400 bp.

2.3 DNA EXTRACTION FROM WHOLE BLOOD CELLS

Genomic DNA was extracted from whole blood samples by use of standard cell lysis, proteinase K digestion and a phenol/chloroform ethanol precipitation procedure. Most blood samples were stored at -80°C, but on some occasions the DNA was extracted from fresh blood when it was more practical to do so. The extraction takes two days and the first day included isolating the leukocytes, which were then lysed so as to release the genomic material. This was followed by overnight digestion of proteins with proteinase K. The second day involves isolating and extracting the genomic DNA.

Day 1:

If frozen, blood was removed from cold storage (-80°C) and then gently thawed in a water bath (30-35°C). This was done to stop cells from lysing due to shock and to also prevent the release of enzymes that damage genomic DNA e.g. DNAase.

Fully thawed blood was transferred into a 50 ml centrifuge tube, which was topped to 50 ml with lysis buffer. The tube was inverted several times to ensure mixing. This allowed the lysis buffer to break up whole red blood cells (erythrocytes), whilst leaving white blood cells (leukocytes) intact within the DNA.

The lysate was spun using a balanced centrifuge at 3000 rpm (2000 g) for 15 minutes (4°C). This pellets and isolates white blood cells. The supernatant was disposed of into Virkon disinfectant, leaving the white blood cell pellet behind. This pellet was resuspended in 50 ml of lysis buffer to remove any remaining red blood cells and again centrifuged at 3000 rpm for 15 minutes (4°C). The supernatant was once more disposed of in Virkon.

The cleaned white blood cell pellet was resuspended in 500 µl 10% SDS, which breaks open the cells by disrupting the lipid membrane. This releases its contents, including the genomic DNA, into the solution. In addition 10 ml proteinase K buffer was added to provide stable optimum conditions for the proteinase enzyme. To the lysate, 50 µl of 20 mg/ml proteinase K was mixed in. The presence of this enzyme helps break down complex proteins, allowing for their efficient removal. It also deactivates any damaging enzymes such as DNAase, which would harm the targeted DNA. The lysate was then incubated in a water bath at 55°C (just below the optimum temperature of proteinase K that is 65°C), on a shaking platform. This was done overnight so that the reaction occurs for as long as possible.

Day 2:

5 ml buffered phenol and 5 ml chloroform with IAA (24:1 ratio of chloroform to Isoamyl alcohol) was mixed with the cell lysate. The phenol-chloroform mixture was used to dissolve the lipids and precipitate proteins. This leaves the nucleic acids in an aqueous layer with white coagulated protein left at the interphase, after centrifuging.

The mixture is centrifuged at 3000 rpm for 15 minutes (room temperature). This allows for separation of the mixture with the organic solvent (phenol-chloroform) layer at the bottom of the centrifuge tube (containing lipids), the protein interphase in the middle and the aqueous layer on the top (containing nucleic acid). The aqueous layer was transferred into a freshly labelled 50 ml centrifuge tube for DNA precipitation and the organic solvents were disposed of safely (UCL health and safety disposal of hazardous solvents protocol).

As a note, it is notoriously difficult to extract the aqueous layer right up to the interphase without disturbing the layer itself. Therefore, in addition to the phenol-chloroform, 1 g PVPP mixed with 5 ml TE was used. Once centrifuged with the organic mix, the protein interphase is held down along with the organic solvents. This allows for the maximum removal of the aqueous layer without contamination. From this aqueous layer, DNA was precipitated by adding 1500 µl of 3 M sodium acetate and 30 ml of absolute ethanol. The tube was then inverted gently until the DNA precipitated out of solution into a condensed white clump.

The DNA clump was removed with a sterile glass or plastic rod and then washed in 70% ethanol to remove as much sodium acetate as possible (in order not to interfere with any polmerase chain reactions). The DNA clump was then transferred into a labelled 1.5 ml microscrew tube containing 500 µl TE. Finally, DNA samples were stored away from light at room temperature for about a week. This allows the DNA to dissolve into solution before quantification.

2.4 DNA QUANTIFICATION

All extracted DNA samples were quantified with Picogreen (Molecular Probes) by use of fluorimetry (FLA-3000 scanner; Fuji). This was done to standardise DNA concentrations so that known amounts of DNA can be used in genotyping and sequencing.

The DNA samples were quantified by transferring 2 µl of each sample into a freshly labelled 2 ml eppendorf tube containing 78 µl TE. This was thoroughly mixed and after votexing, 10 µl of each sample was transferred into a flat bottomed well plate containing 90 µl of TE. The locations of each sample within the 96 well plates were recorded. A series of wells containing known concentrations of DNA, derived from calf thymus, were additionally transferred into the 96 well plates to use as comparison. The dilution series included DNA samples at 0, 1, 10, 20, 50, 75 and 100 ng/µl. This was done in order to produce a standard curve from which the concentration of the samples could be derived from.

The DNA samples were mixed with 100 µl Picogreen (150 µl Picogreen dissolved in 30 ml TE), which is a fluorescent dye that specifically binds to double stranded DNA. Once the picogreen is mixed, the plate containing the sample is scanned by fluorimetry. The fluorescence is directly proportional to the quantity of DNA present. Using the fluorimetry results, each DNA sample is diluted to 25 ng/µl in two 1.5 ml microscrew tubes. The rest of the DNA is labelled up as stock for future use. All DNA were stored below 5°C and in the dark to maintain quality.

2.5 SELECTING GENETIC MARKERS FOR ASSOCIATION STUDIES

In this genetic association study, fine mapping of the chromosome 11q22-24 was done so as to detect linkage disequilibrium between DNA markers and schizophrenia. Microsatellites were initially used as they are useful for identifying distant LD. However, they are not very common in the genome. Therefore, further fine mapping of the region was carried out using SNPs as these markers tend to be more common (1 in every 300-1000 bp).

2.5.1 IDENTIFYING MICROSATELLITE MARKERS

Random microsatellite markers on chromosome 11q22-24, close to the maximum LOD score position in the UCL schizophrenia family linkage study, were initially selected for fine mapping (Gurling et al. 2001). These markers are established

microsatellites obtained from the GDB Human Genome Database and the University of California Santa Cruz (UCSC) Genome Browser database. Other microsatellite markers genotyped were also chosen from these bioinformatic databases. They include unformatted (non-established) microsatellites, which were identified from simple repeats found in the UCSC Genome Browser database, May 2004 assembly. These new repeats were formatted by us for genotyping. The forward and reverse primers were created using the Primer3 design software (Rozen & Skaletsky 2000).

An M13 tail (5'-CACGACGTTGTAAAACGAC-3') was added to the 5' end of a single primer for each microsatellite marker (Section 2.6). These M13 tails are used to detect the PCR product through fluorescence after the addition of a second M13 primer that is labelled with a particular wavelength dye. This complementary M13 primer incorporates itself into the product. The M13 wavelengths used for genotyping were M13F 700 (700 nm) and M13F 800 (800 nm), which allowed for more than one marker of similar size to be genotyped at the same time on different wavelengths. Both pair of primer sequences were then tested for the possible formation of secondary structures e.g. hairpin loop, which would impede amplification. This was examined using the NetPrimer (Premier Biosoft international) program. Here, a score of under 70 is undesirable. Therefore, the primer with the best score for M13 tail attachment was chosen. All primers were ordered from MWG biotech (Ebersberg, Germany). The primers were received in lyophilised form and were rehydrated in TE to produce a concentration of 100 pmol/μl.

2.5.2 IDENTIFYING SNP MARKERS

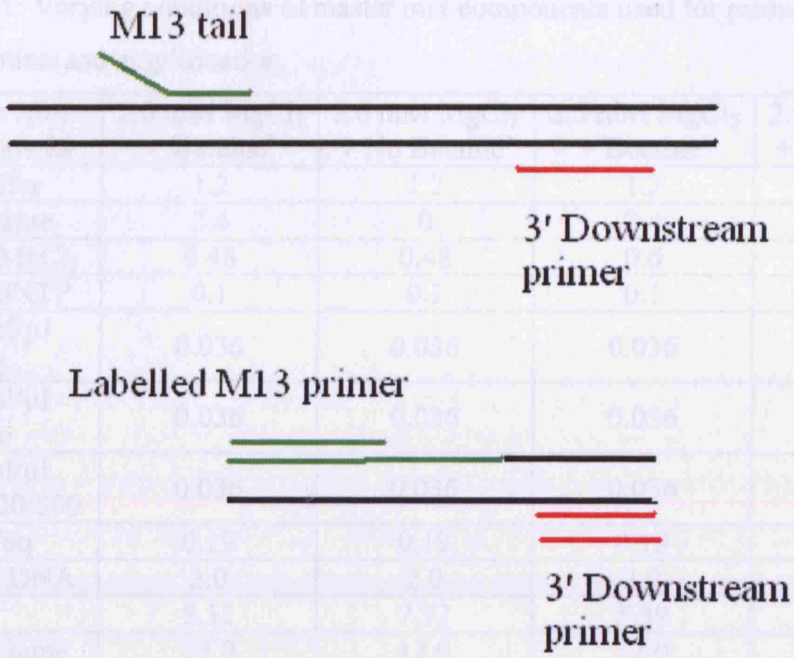
SNPs were initially chosen from Ensembl before the availability of data from the International HapMap Project (HapMap 2003). A number of programs enable the selection of tag SNPs. These are SNPs that are likely to carry identical genetic information to that of nearby SNPs in complete LD. Thus, enabling the selection of SNPs that represent other SNPs in the region without duplicating information. The data generated by HapMap can be interpreted using the Haploview program (Barrett

et al. 2005). This software helps visualise LD from a chosen region and helps to judge which SNPs are useful to type, based on their haplotypic structure. The software also contains a Tagger function that can be used as an automated SNP selection tool because it deduces all tagged SNPs found within HapMap. Both D' (Section 3.31) and r^2 (Section 3.3.2) values are shown within the Haploview graphics. When both LD values are equal to 1, the SNPs are said to be genetically identical and share the same genotypic information.

2.6 PCR AMPLIFICATION

PCR amplification of the microsatellite markers was performed using the M13 tailed primer and a second non-tailed primer. The methodology involves the dual dye infrared fluorescence technology with use of LI-COR® 4200 IR DNA sequencers (LI-COR®, Biosciences, NE, US), as described by Otting et al. (1995). A third universal M13 sequence primer, labelled with infrared dye IRD700 or IRD800, was used to hybridise against the M13 tailed locus specific primer (Figure 2.1).

Figure 2.1: M13 and PCR reaction multiplication process.



The M13 sequencing primer, with identical sequence to the tail of the forward primer, is combined with the IRD fluorescent molecule of either the 700nm or 800nm wavelength. During the initial cycles of the PCR reaction, the forward and reverse primers are incorporated into the DNA and amplify the sequence. In further cycles, the M13 primer replaces the forward primer to create an amplicon labelled with IRD fluorescent molecules.

2.6.1 PCR MASTER MIX

Newly designed primers need to be optimised to find the optimal master mix concentration and PCR condition, which produce the best result for successful amplification of the targeted region. Optimisation is carried out on DNA samples that have been excluded from the study and PCR products are visualised on agarose and polyacrylimide gels. When optimising primers, four common conditions were used to amplify the target region (Table 2.1). 12 µl PCR reactions were prepared on ice for each of the varying conditions, which differed in the concentration of Betaine and $MgCl_2$ in each reaction. Betaine helps to reduce the melting point and is useful for amplifying regions of high GC content. The variation in magnesium ion concentration influences the specificity of the PCR reaction and also acts as a co-enzyme to the Taq polymerase.

Table 2.1: Varying conditions of master mix components used for primer optimisation and amplification.

Master Mix Components	2.0 mM $MgCl_2$ + Betaine[a]	2.0 mM $MgCl_2$ + No Betaine[a]	2.5 mM $MgCl_2$ + Betaine[a]	2.5 mM $MgCl_2$ + No Betaine[a]
10X Buffer	1.2	1.2	1.2	1.2
5 M Betaine	2.4	0	2.4	0
50 mM $MgCl_2$	0.48	0.48	0.6	0.6
25 mM dNTP	0.1	0.1	0.1	0.1
100 pmol/µl F primer	0.036	0.036	0.036	0.036
100 pmol/µl R primer	0.036	0.036	0.036	0.036
100 pmol/µl M13F 700/800	0.036	0.036	0.036	0.036
1 U/µl Taq	0.19	0.19	0.19	0.19
25 ng/µl DNA	2.0	2.0	2.0	2.0
Water	5.52	7.92	5.39	7.79
Total Volume	12.0	12.0	12.0	12.0

[a] Volumes in µl.

When amplifying and genotyping a microsatellite marker, the relevant M13F 700 or 800 primers are added into the master mix. This produces a PCR product at the respective frequency.

2.6.2 PCR CYCLES

Once all master mix components were added to the genomic template, the PCR can begin. All PCRs were run on an MWG-HT Primus 96 thermal cycler (MWG Biotech, DE). The DNA template was first denatured for an extended period of time to separate the double stranded helix and to allow the template to become linear.

The PCR was run in a three step cycle:

1. Denaturation: Typically occurs at 93-95°C.

2. Reannealing: Primers bind specifically to their complementary sequence. The annealing temperature usually depends on the melting temperature (T_m) of the expected duplex. This tends to be approximately 5°C below the expected T_m.

3. DNA synthesis: Extension of the complementary strand is initiated by the annealed primer. This occurs at 70-75°C.

PCRs were run for approximately 25-35 cycles and finished with an extended hold at 72°C. This was done to complete all synthesis and extensions of the targeted region. Three standard PCR programs were required to sufficiently amplify all amplicons successfully (Table 2.2). These programs were tweaked by adjusting the annealing temperature and by adding/removing a few cycles to obtain a clean amplification of the target region.

Table 2.2: PCR cycling conditions.

Touch Down	Standard 55°C	Standard 60°C
Lid Heated to 105°C Products are denatured at 94°C for 5 minutes 3 cycles of: 94°C - 30 seconds 63°C - 30 seconds 72°C - 30 seconds 3 cycles of: 94°C - 30 seconds 60°C - 30 seconds 72°C - 30 seconds 3 cycles of: 94°C - 30 seconds 57°C - 30 seconds 72°C - 30 seconds 3 cycles of: 94°C - 30 seconds 54°C - 30 seconds 72°C - 30 seconds 3 cycles of: 94°C - 30 seconds 51°C - 30 seconds 72°C - 30 seconds 20 cycles of: 94°C - 30 seconds 48°C - 30 seconds 72°C - 30 seconds Hold at 72°C - 10 minutes Store at 4°C	Lid Heated to 105°C Products are denatured at 94°C for 5 minutes 35 cycles of: 94°C - 30 seconds 55°C - 30 seconds 72°C - 30 seconds Hold at 72°C - 10 minutes Store at 4°C	Lid Heated to 105°C Products are denatured at 94°C for 5 minutes 35 cycles of: 94°C - 30 seconds 60°C - 30 seconds 72°C - 30 seconds Hold at 72°C - 10 minutes Store at 4°C

The 12 μL PCR reactions were dehydrated to 5 μL and an equivalent volume of Fuschin loading buffer was added to each amplified product. Samples were then denatured by heating the plate to 95°C for three minutes prior to loading on a

polyacrylamide gel. This ensured that the oligonucleotides were single stranded. Microsatellite marker fragment sizes were then separated and visualised with either of the two infrared dyes on LiCor 4200L sequencers (Section 2.6.3). The best result from the variable master mix (Table 2.1) and PCR (Table 2.2) conditions were then used to amplify the entire case-control sample.

PCR reactions for each microsatellite marker were carried out separately in a 96 well PCR plate. The layout of the plate was arranged with selected samples removed to incorporate size markers into the gel. The samples that were taken out were transferred onto another plate and were subsequently loaded on a different gel.

2.6.3 POLYACRYLAMIDE GEL ELECTROPHORESIS

The Licor sequencers contain a focusing fluorescent microscope with a solid state silicon avalanche photodiode. The image output is displayed on a computer screen as a serious of bands revealed in a real time autoradiogram-type format (Middendorf et al. 1992).

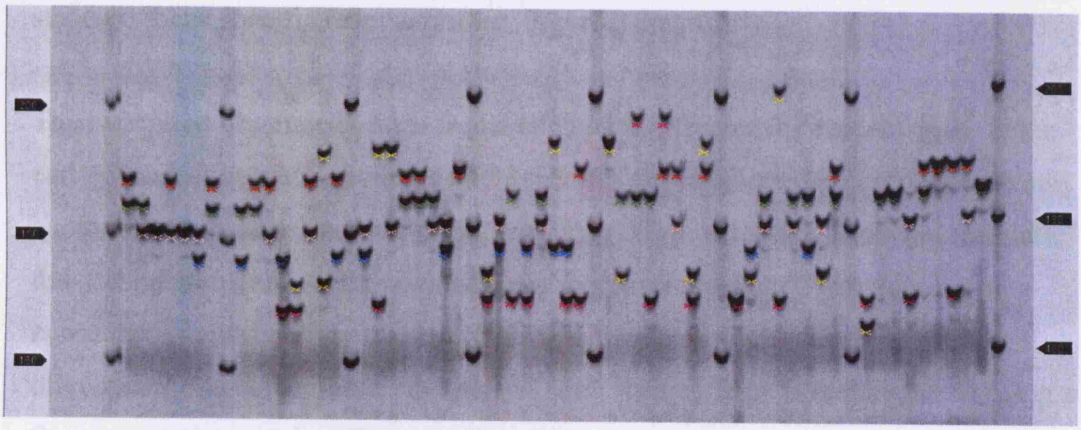
Gels were created (Section 2.2.5) by adding ammonium persulphate (APS) to polymerise the acrylamide (800 µl of 10% APS to every 100 ml of buffered Sequa gel solution). The gel apparatus consists of two 25 cm glass plates that were separated using 2 mm thick spacers. Acrylamide with added APS was then poured into the gap. A single well was prepared at the top of the gel. Once the gel had set (approximately 2 hours), a sharktooth comb was inserted into the single well to form multiple wells in which the samples were loaded. Running conditions were 1200 V, 25 mA, 30 W, signal filter 3 and scan speed 4. The sequencer maintains a constant temperature of 45 ± 1°C. Before loading samples, the gel is preheated to 45°C to help prevent the formation of secondary structure. This allows the linear DNA to run in proportion to their length.

2.7 GENOTYPING METHODS

2.7.1 SAGA-GT

Genotyping of microsatellite markers was accomplished using SAGA-GT™, LI-COR® version 3.2 (Biosciences, NE, USA) genotyping software (Figure 2.2). The genotypes were checked by eye, with all allele calls verified by a second independent person blind to initial calling. Any discrepant genotypes between the two assessors were repeated from the PCR amplification stage. All data were automatically stored in the database.

Figure 2.2: Example of a Licor gel image analysed using the SAGA-GT software.



The Licor gel image is of a tetranucleotide repeat marker that has been analysed. The allele calls for all genotyped samples are marked with coloured "x" symbols that correspond to a fragment size, which are based on ladder samples loaded at specific lanes (1, 9, 18, 27, 36, 45, 54 and 64).

When the entire case-control sample had been genotyped, randomly chosen samples from both control and schizophrenia plates were reamplified. These samples were used as cross checks to compare with original genotypes. This was done in order to assess the consistency of the process. Allelic association analyses with schizophrenia were then performed using the software CLUMP (Section 3.5.2). The CLUMP program employs an empirical Monte Carlo test of significance that does not require further correction for multiple alleles (Sham & Curtis 1995).

2.7.2 KBIOSCIENCE

Kbioscience (KBioscience, Hoddesdon, UK) were employed to undertake the genotyping for the majority of SNPs in the case-control sample. For each SNP, 50 bp either side of the polymorphism was selected from the dbSNP webpage http://www.ncbi.nlm.nih.gov/projects/SNP, within the National Centre for Biotechnology Information (NCBI) Database. This information, along with SNP name, was provided to KBioscience for SNP assay development.

KBiosciences employs a modified version of the Amplifluor (Myakishev et al. 2001) genotyping procedure (Millipore, MA, US). This method involves the allele specific amplification of SNP alleles using two tailed locus specific oligonucleotides and a standard locus specific reverse primer (Figure 2.3 and 2.4).

The two tailed oligonucleotides are labelled with different fluorescent dyes. In the early cycles of amplification, the allele specific oligonucleotides compete with one another for template. The best fitting/matching oligonucleotide binds the template, due to higher affinity, and thereby creates more of this allele's template for subsequent amplification. In the later cycles, the appropriate tailed fluorescent oligonucleotide is incorporated. The end point fluorescence is then detected using a fluorescent plate reader. There are four different possible outcomes, each with different fluorescent intensities:

1. Low intensity for both fluorescent dyes, indicating that no amplification has occurred.

2. High intensity for one fluorescent dye and low intensity for the other, indicating a homozygote template for the appropriate SNP allele.

3. The opposite scenario to (2).

4. Moderate intensities for both fluorescent dyes, indicating a heterozygous template for the SNP.

Figure 2.3: KASPar genotyping method for a two allele SNP, using allele-specific primers coupled with two Univeral Amplifour primers (Bengra et al. 2002).

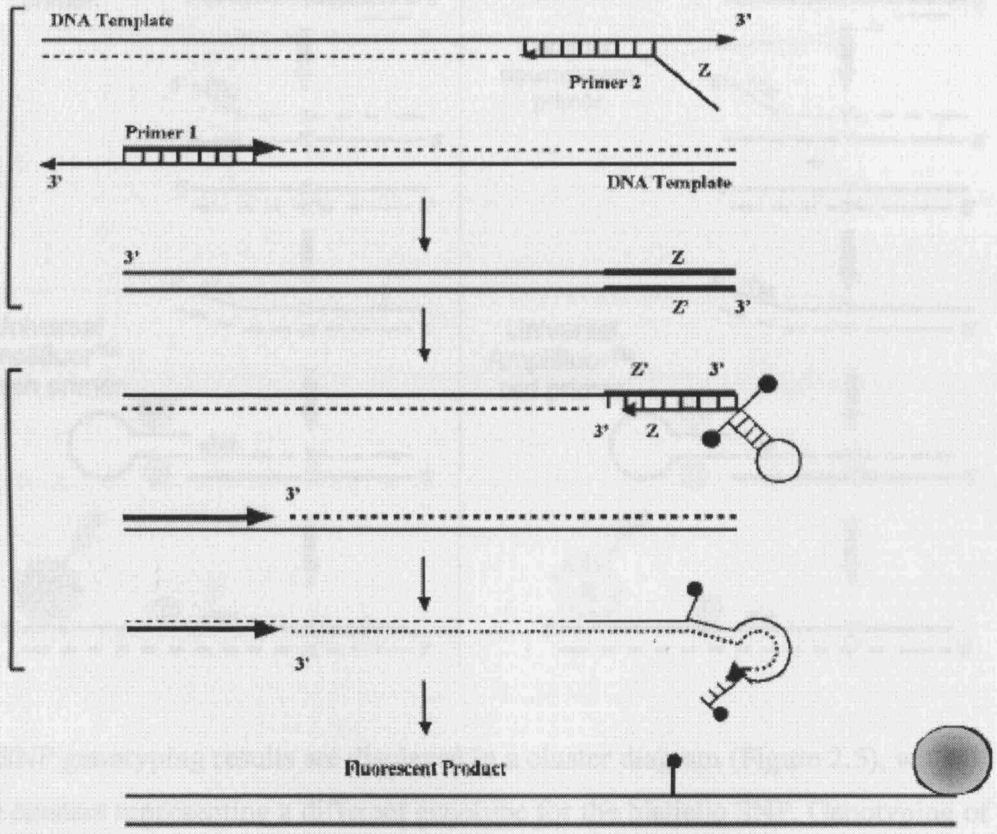
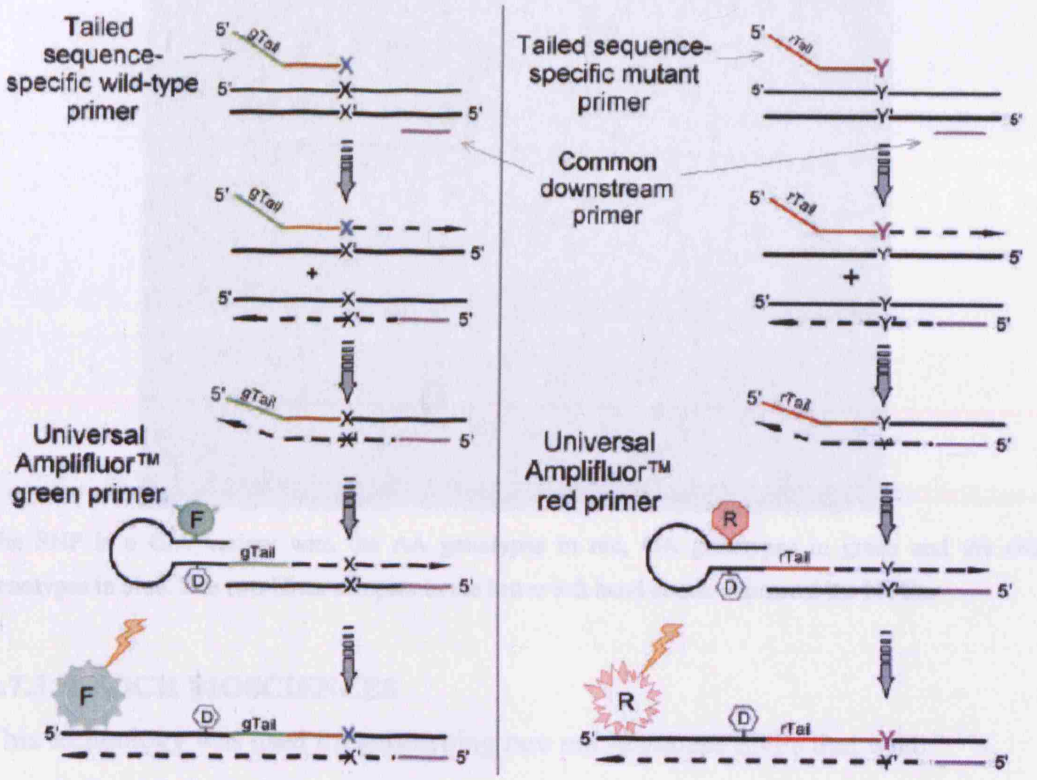
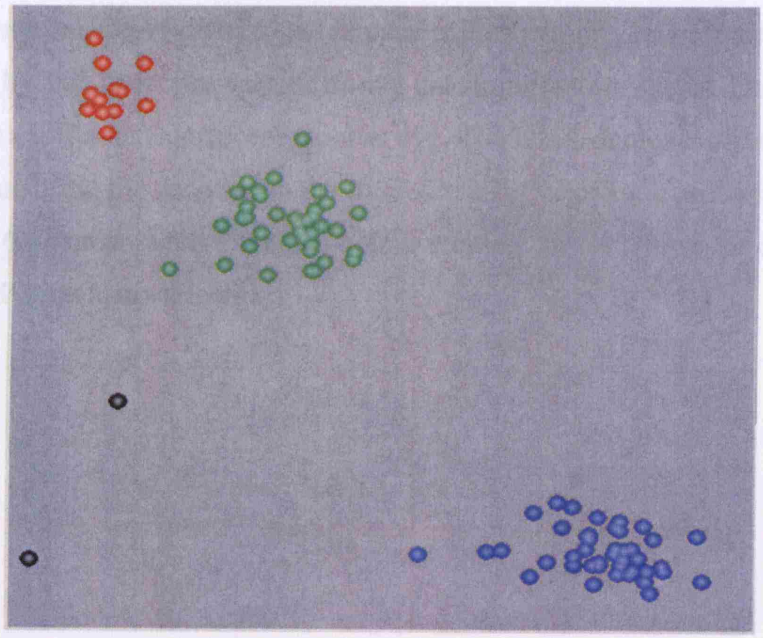


Figure 2.4: An alternative diagrammatic representation of two allele SNP detection, using allele-specific primers coupled with two Univeral Amplifour primers (Bengra et al. 2002).



The SNP genotyping results are displayed in a cluster diagram (Figure 2.5), with all three clusters representing a different genotype for the biallelic SNP. Genotyping of samples on a 96 well plate included two non-template controls (NTCs) and the results for each plate were displayed in the cluster diagram. Of the samples, 17% from each microtitre plate were reduplicated in order to detect possible errors and to confirm the reproducibility of genotypes. The data were then analysed to confirm Hardy-Weinberg equilibrium (Section 3.4).

Figure 2.5: SNP genotypes of a 96 well plate.



The SNP is a G/A variant with the AA genotypes in red, GA genotypes in green and the GG genotypes in blue. The two black samples in the lower left hand corner represent the NTCs.

2.7.3 EPOCH BIOSCIENCES

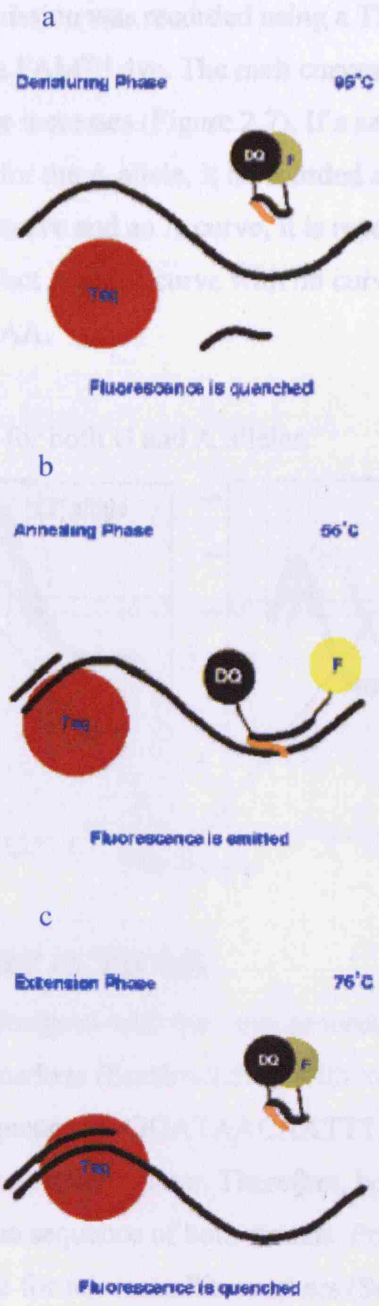
This technology was used for genotyping rare non-database SNPs that were potentially aetiological, in our entire case-control sample.

Epoch Biosciences (Epoch Biosciences merged into Nanogen, Inc., WA, US) were employed to design a genotyping assay for the G/A non-database SNP (position 117252877, UCSC Genome Browser database, March 2006 assembly) in the case-control sample. 50 bp either side of the polymorphism were selected using the UCSC Genome Browser database. This sequence was sent to Epoch Biosciences along with genomic DNA from two known heterozygotes and two known major allele homozygotes of the SNP. The received probes contained a fluorophore moiety at the 5' end and a quencher at the 3' end.

Epoch Biosciences use a probe containing a minor groove binding (MGB) sequence attached to a quencher at the 5' end and a fluorophore at the 3' end. This method

does not need the Taq DNA polymerase to cleave the probe. It simply emits fluorescence once hybridised to the sequence. Furthermore, it is protected from 5' digestion by Taq DNA polymerase during the amplification stages. Each MGB probe has a different fluorescent reporter dye. The DNA duplexes are denatured over a time course and the decrease in the fluorescent signal of each probe is measured (Figure 2.6). Comparison of the two-colour melting curves allows for differentiation of the SNP sequence variants.

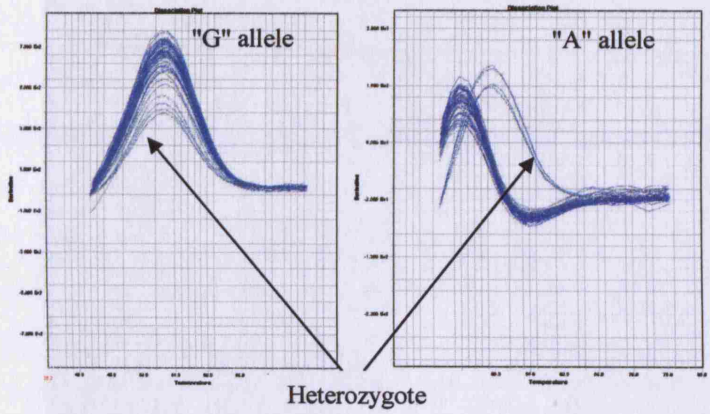
Figure 2.6: Epoch Biosciences SNP genotyping method.



The assay exploits the 5'→3' exonuclease activity of Taq DNA polymerase to cleave the probe and release the fluorophore from the quencher. (a) Denaturing phase: The probe is unattached and no fluorescence is emitted. (b) Annealing phase: The probe is attached in the minor groove and fluorescence is emitted. (c) Extension phase: The probe is unattached and no fluorescence is emitted. Figure courtesy of Epoch Biosciences, http://www.apczech.cz/pdf/DF_SNP_and_allelic.pdf.

The G allele melt curve emission was recorded using a TET™ dye and the A allele curve was recorded using a FAM™ dye. The melt curves of each sample were recorded as the temperature increases (Figure 2.7). If a sample produces a perfect G allele curve with no curve for the A allele, it is recorded as a homozygote G. If a sample produces both a G curve and an A curve, it is recorded as heterozygote GA. If a sample produces a perfect A allele curve with no curve for the G allele, it is recorded as a homozygote AA.

Figure 2.7: Melting curves for both G and A alleles.

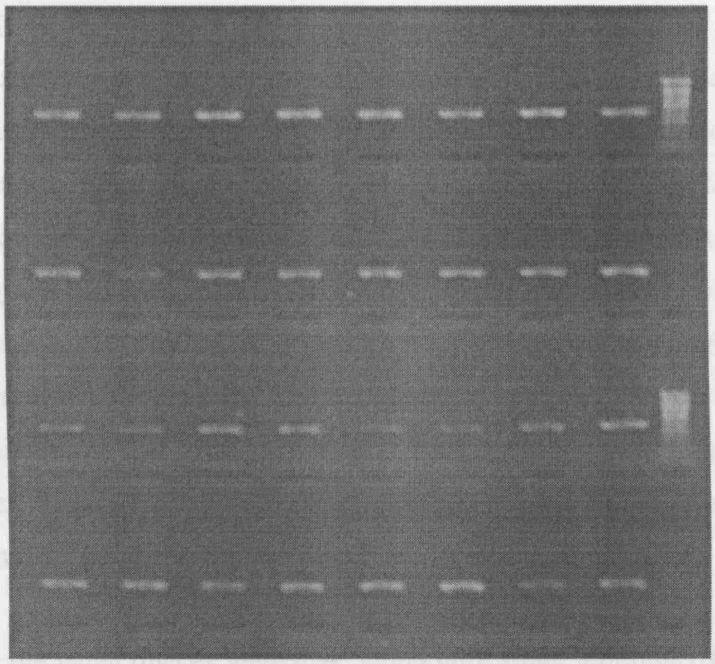


2.8 SEQUENCING REACTIONS

Sequencing primers were designed with the same procedure as those used for amplifying microsatellite markers (Section 2.5.1), with the exception that an additional reverse M13 sequence (5'-GGATAACAATTTCACACAGG-3') was added to the 5' end of the non-tailed primer. Therefore, both M13F 700 and M13R 800 were used to capture the sequence of both strands. Primers were optimised using the same procedure as those for microsatellite markers (Section 2.6.1 and 2.6.2), without the addition of the IRD 700/800 dye. The template was then cleaned (Section 2.8.1) and set up for the final PCR phase, the Sanger-Coulson chain termination sequencing method (Section 2.8.2). Both IRD 700 and 800 dyes are incorporated at this stage of the sequencing process.

DNA samples for sequencing were selected on the basis that cases of schizophrenia had inherited alleles and haplotypes that had previously shown association to schizophrenia. Once the primers were amplified, a proportion of the PCR products were analysed on an agarose gel. This was to check that all individuals have been amplified, are the correct amplicon size, contain no non-specific amplification (multiple bands) and have a low intensity of primer dimmer or none at all.

Figure 2.8: An agarose gel illustrating optimisation of sequencing primers.



1% agarose gels were made using 1X TBE (0.4 M Tris, 0.2 M sodium acetate, 20 mM EDTA, pH 8.3). Ethidium bromide was added to a final concentration of 0.5 µg/ml. Loading buffer (1X TBE, 50% v/v glycerol, 0.025% bromophenol blue) was added to DNA samples prior to loading. 4 µl of the 25 µl PCR reaction was mixed with 1 µl of loading buffer and this was subsequently loaded into the agarose gel.

2.8.1 CLEAN UP

Clean up removes primers, excess dNTPs and any primer dimer produced by the PCR reaction. The amplicon for each individual reaction was added to an equal volume of Microclean (Microzone Ltd, West Sussex, UK) in an eppendorf, vortexed and left at room temperature for 5 minutes. All eppendorfs were then centrifuged at

14000 g for 7 minutes so as to pellet the template. The supernatant was disposed of and the tubes were briefly spun again to pull down any remaining "dregs", which were then removed. Tubes were left open for 10 minutes at room temperature. The pellet was then resuspended in 10-15 µl of ultra pure water. DNA was vortexed again and centrifuged briefly. The cleaned up DNA template was then used for the Sanger-Coulson chain termination sequencing method (Section 2.8.2).

2.8.2 SANGER-COULSON CHAIN TERMINATION SEQUENCING METHOD

Once the targeted region was amplified and cleaned, the DNA template can then be sequenced. The amplicon contains M13F/R sequence that was integrated by the primers used in the initial amplification. An M13F/R primer can therefore be used for specific sequencing and simultaneous incorporation of the dye for visualisation on LICOR sequencers. During the PCR reaction, the M13F/R primer binds to the specific template and allows extension and synthesis of the complementary strand by the SequiTherm DNA polymerase. The advantage of using M13F 700 and M13R 800 is that both strands can be sequenced simultaneously and used as a cross check for one another, thus assessing whether novel bands are artefacts or real polymorphisms.

In sequencing the usual dNTPs were used along with small amounts of dideoxynucleotide (like ddATP), which is incorporated just as efficiently into a growing polynucleotide strand but prevents further synthesis. This is because the dideoxynucleotide lacks a hydroxyl group at the 3' position of the sugar component and this prevents further incorporation into the polynucleotide. Therefore, chain termination occurs wherever a dideoxynucleotide is incorporated. As a result, the reaction can run with each individual ddNTP and will terminate at every base of the sequence, thus displaying exactly where each base belongs to in a sequence.

For each sample, 3 µl of the cleaned DNA template was mixed with 6.8 µl of the sequencing master mix (Table 2.3). 2 µl of this is then mixed with 2 µl of each

ddNTP (Figure 2.9). The sequencing reaction was then PCR amplified by the Sequatherm PCR program (Table 2.4). This was carried out on a 96 well plate and enabled the sequencing of 16 individuals at the same time on a single polyacrylamide gel.

Table 2.3: Master mix reagents and volumes required to sequence 16 individuals.

Sequencing reagent for master mix (MM)	Volume (µl)
3.5X Buffer	75
10 pmol/µl M13F 700	1.33
10 pmol/µl M13R 800	1.33
SequiTherm Polymerase	11
Water	27

Figure 2.9: Sequencing reaction layout revealing how master mix and DNA were mixed with each ddNTP separately.

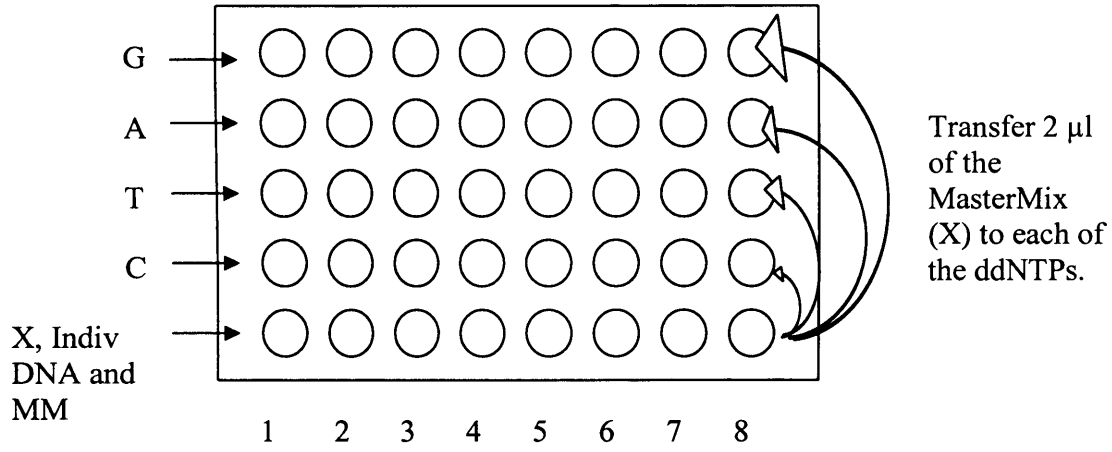


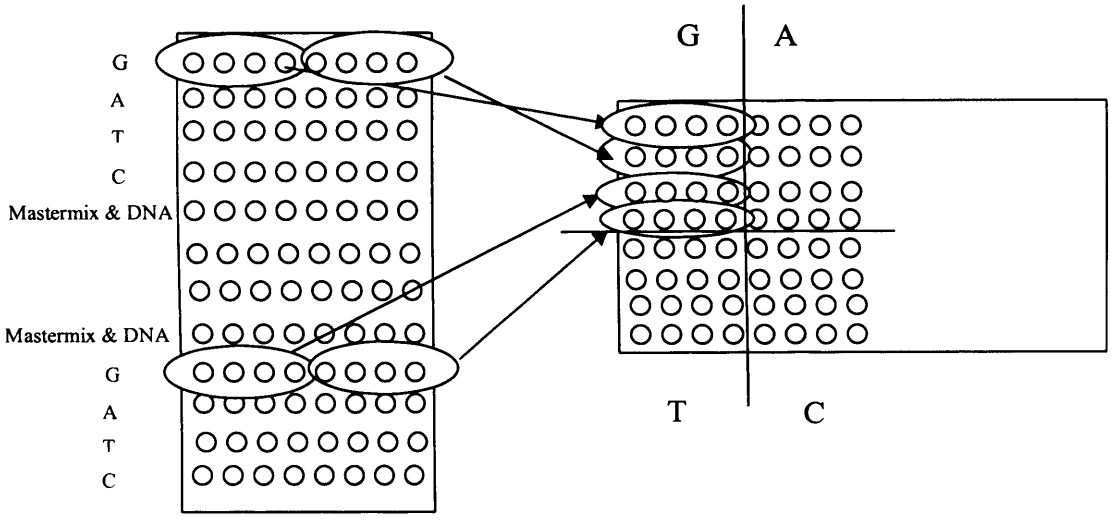
Table 2.4: PCR cycling conditions for sequencing.

Sequitherm
Lid Heated to 105°C Products are denatured at 94°C for 5 minutes 30 cycles of: 92°C - 30 seconds 60°C - 30 seconds 70°C - 1 minute Store at 4°C

2.8.3 POLYACRYLAMIDE GEL ELECTROPHORESIS

3.5 µl of the Sequitherm reaction was added to an equivalent volume of Fuschin loading buffer. 3.5 µl of this Fuschin/Sequitherm reaction mix was then transferred onto another 96 well PCR plate using the format shown in Figure 2.10. This facilitates visual scoring of potential mutations and allows recognition of polymorphisms in their heterozygote state when the products are separated on the polyacrylamide gel (Figure 2.11).

Figure 2.10: Method for transferring bases into a loading plate.

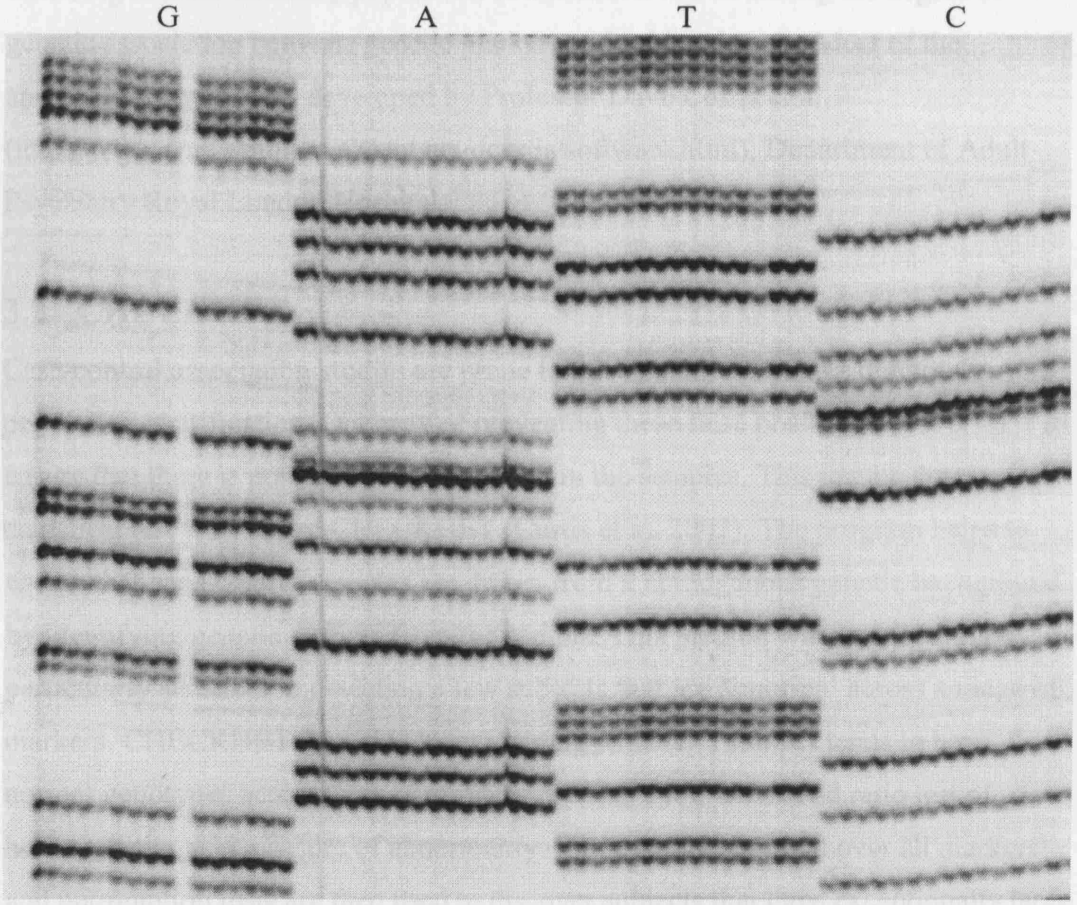


The amplified product is transferred using the specific plate format. This enables loading with a multichannel pipette contiguously, which keeps all the same nucleotides aligned.

The sequencing product was loaded on a 44 cm agarose gel and analysed by the LICOR scanner on the preset sequencing parameters (2000 V, 35 mA and 45 W). The resulting sequencing image was read and checked against the database sequence to make sure the correct region was sequenced and to also check for any abnormalities. Validity of potential mutations was assessed by comparison with the gel image produced from simultaneous sequencing reactions primed from the opposite strand. Any abnormalities were checked in random controls to see if they appeared in the same frequency as the cases. In total 32 of the most significantly

associated individuals were sequenced. The presence of any potential aetiological base pair changes were then investigated by sequencing with an equal number of random controls.

Figure 2.11: Example of a Licor DNA sequencer gel image.



The gel was created using the Sanger-Coulson chain sequencing method and illustrates the loading layout. Here, the bases of all 16 individuals appear together to help identify any base pair changes.

3.0 STATISTICAL ANALYSIS

3.1 INTRODUCTION

A range of statistical analysis was used to determine the reliability of our findings. This ranged from checking population stratification to calculating the degree of genetic association between genetic markers and schizophrenia. Most of the analytical software was developed by Professor Dave Curtis et al. (http://www.smd.qmul.ac.uk/statgen/dcurtis/software.html), Department of Adult Psychiatry Royal London Hospital.

3.2 CHECKHET

Case-control association studies are prone to false positive findings due to population stratification. One way of preventing these false positive discoveries is to ensure that there is no overall heterogeneity in the samples. This can be done via a statistical test with use of CHECKHET (Curtis et al. 2002). The program helps to ensure that case-control samples are drawn from a homogenous genetic background by identifying genetic outliers from the sample. This method was intended to be particularly sensitive in detecting a few subjects that are abnormal across a range of markers. CHECKHET does this by examining whether a subject tends to have normal genotypes across a range of markers using a log likelihood ratio test of heterogeneity as a measure of abnormality. The tests are summed over all markers and permutation tests are then used to discover subjects that have exceptionally large scores. These samples are then excluded from the analysis. The program depends on genotypes rather than allelic frequencies because it does not make the assumption that all markers are in Hardy-Weinberg equilibrium.

Fifteen reference microsatellite markers at chromosomal loci not thought to be involved in schizophrenia were genotyped in a subset of samples (200 cases and 300 controls). This was arranged so as to detect genetic heterogeneity between the cases and controls. It was also done to confirm that the samples were genetically well

matched i.e. that the ancestry selection method was valid. CHECKHET was then employed to identify subjects with an atypical genetic background (Curtis et al. 2002). No evidence of genetic heterogeneity between cases and controls was found with use of the reference markers. The CHECKHET test detected two schizophrenic subjects with abnormal genotypes. These were excluded from the sample before any chromosome 11 markers were genotyped.

3.3 LINKAGE DISEQUILIBRIUM (LD)

LD is often termed "allelic association." When alleles at two distinctive loci occur in gametes more frequently than expected (rather from other causes such as population stratification or epistasis), given the known allele frequencies and recombination fraction between the two loci, the alleles are said to be in linkage disequilibrium. Evidence for linkage disequilibrium can be helpful in mapping disease genes since it suggests that the disease causing change and genetic marker may be very close to one another. LD is also commonly used to describe the relationship between genetic markers. However, it can be disrupted by the recombination rate and the presence of new mutations.

3.3.1 LEWONTIN'S D' MEASUREMENT OF LD

Lewontin's D' is derived from D (Lewontin 1964a; Lewontin 1964b), which is a LD coefficient derived from haplotypic frequencies. Using a biallelic system e.g. SNP, the haplotypic frequencies can be simply estimated between two markers.

D measures the deviation of haplotype frequencies from the equilibrium state. LD occurs when D is significantly greater than zero. Considering two linked SNPs with alleles (A,a) and (B,b), that result in four possible haplotypes (AB, Ab, aB and ab), D can thus be calculated.

		SNP1		
		A	a	
SNP2	B	fAB	faB	fB
	b	fAb	Fab	fb
		fA	Fa	

f = frequency

As a result, the LD coefficient D could be worked out as follows:

$$D = f(AB) - f(A) \times f(B)$$

The problem with the D coefficient is that it is dependent on the marginal allele frequencies in the contingency table. This disqualifies D as a measure of LD because the data is dependent and cannot be compared for different SNPs and different populations (Devlin & Risch 1995). However, D can be normalised to D'. This allows the measurement to be comparable across SNPs and populations.

D' is defined as the absolute ratio of D compared with its maximum value, D_{max} (when $D \geq 0$), or compared with its minimal value, D_{min} (when $D < 0$). When $D' = 1$, it denotes complete LD. Historical recombination result in the decay of D' towards a value of zero.

When $D' = 1$, alleles at adjacent marker loci are in LD with each other as much as possible, given that they have different population allele frequencies. D' values less than 1 imply that the maximum extent of LD between two marker alleles has not been observed e.g. $D' = 0.87$ represents a strong LD between SNP alleles, while $D' = 0.12$ represents weaker LD between SNP alleles. D' has the same range of values (Lewontin 1988) and its sign (positive or negative) depends on the arbitrary choice of the alleles paired at the two loci.

The normalised measure of D (D') is worked out as follows:

$$D' = \frac{D}{\sqrt{p(1-p)q(1-q)}}$$

Where p and q are major and minor allele frequencies respectively, as described in the Hardy-Weinberg equation (Section 3.4).

Where $|D'| = 1$, denotes complete LD.

Where $D' = 0$, reveals no LD.

Where $|D'| < 1$, shows various degree of LD between the markers. The higher the D' value the stronger the LD.

The limitations of using D' include its insensitivity to distinguish between different degrees of LD. A $|D'| = 1$ can reflect 2 to 3 haplotypes present. There are also upward bias for small to moderate sample sizes and it can take extreme values when at least one allele frequency is small (Devlin & Risch 1995).

3.3.2 r^2 MEASUREMENT OF LD

r^2 is another measurement of LD (Pritchard & Przeworski 2001). It is thought to overcome the problems of D', as r^2 is able to detect LD in small sample sizes and is also far more reliable for detecting LD for low allele frequencies compared to D'.

The equation for r^2 is as follows:

$$r^2 = \frac{(f(AB) - f(A)f(B))^2}{fA.fa.fB.fb} \text{ or } \frac{D^2}{fA.fa.fB.fb}$$

Where f is the frequency of the allele.

When $r^2 = 1$, the markers are said to be in perfect LD. Observations of one marker provides complete information about the other marker, thus making them both redundant. Therefore, they are genetically identical and can be used as a tag SNP for the other in association studies.

$r^2 = 1$ when no recombination has occurred between the markers and the allele frequencies are identical. It also corresponds to a situation where two haplotypes are present out of a possible four.

In general r^2 is more useful for dividing closely located SNPs into blocks, when the purpose is to identify tagging SNPs. However r^2 represents statistical association at the population level as well as incorporating LD and has no direct relationship with recombination like D'. r^2 depends on marker allele frequencies and can be difficult to interpret when comparing multiple markers in a region (Hedrick 1987). D' is directly related to recombination fraction and its generalisation to more than two loci is the only measure of LD not sensitive to allele frequencies.

When judging LD, both D' and r^2 values are used to describe the LD relationship between markers. They were both used to help select the most appropriate SNPs for genotyping, using data from the HAPMAP.

3.4 HARDY–WEINBERG EQUILIBRIUM (HWE)

Microsatellites and SNP genotyping data were all checked to make sure that they follow Hardy-Weinberg equilibrium before any further statistical analysis was considered. This was done as an indicator to whether the alleles had been genotyped correctly.

The equation for HWE is as follows:

$$p^2 + 2pq + q^2 = 1$$

Where p is the major allele and q is the minor allele.

The concept of HWE is used to understand the genetic characteristics of populations (Hardy 1908; Weinberg 1908). The equation denotes that if p is the frequency of one allele (A) for a biallelic locus, then the HWE-expected frequency will be p^2 for the AA genotype, $2pq$ for the Aa genotype and q^2 for the aa genotype. The three genotypic proportions should sum to 1, as should the allele frequencies.

The derivation of the HWE is appreciated by use of a Punnett's square, as shown in Figure 3.1.

Figure 3.1: Using a Punnett's square to derive the HWE equation.

		Females	
		A (p)	a (q)
Males	A (p)	AA (p^2)	Aa (pq)
	a (q)	Aa (pq)	aa (q^2)

The final three possible genotypic frequencies in the offspring are:

$$f(AA) = p^2,\ f(Aa) = 2pq \text{ and } f(aa) = q^2$$

Testing for HWE is used for quality control of large scale genotyping and is an important method to identify systematic genotyping errors in unrelated individuals (Gomes et al. 1999). Genotyping errors are known specifically to affect certain genetic measurements such as LD, upon which association studies depend.

Tests for deviation of HWE were carried out with a simple χ^2 test. This was done by using the observed genotyped frequencies for the data and the expected genotype

frequencies from the equation. Once the data was found to be in Hardy-Weinberg equilibrium, statistical analysis was then used to test for association between the marker alleles and schizophrenia (Section 3.5). However, any markers lacking HWE were repeated by Kbioscience using an alternative Taqman method.

3.5 STATISTICAL TESTS FOR ASSOCIATION BETWEEN MARKER ALLELES AND DISEASE

Allelic association analysis with schizophrenia was performed using a simple χ^2 test for biallelic SNPs (Section 3.5.1) and by using CLUMP for microsatellite markers (Section 3.5.2).

3.5.1 χ^2 TEST FOR ASSOCIATION OF SNPS

SNP genotyping data was first checked for accuracy by comparing the 17% of samples that were duplicated. The data were then analysed to confirm HWE. Once these criteria were satisfied, the genotypic information was checked for association. Allelic association was tested by forming a table with the allele frequencies for both cases and controls. A χ^2 test was then performed, with one degree of freedom (df). Genotypic association was also tested using a similar table representing each genotype frequency and performing a χ^2 test, with 2 df.

3.5.2 CLUMP TEST FOR ASSOCIATION OF MICROSATELLITES

Allelic tests of association for microsatellite markers were undertaken using a program called CLUMP (Sham & Curtis 1995). Genotyping data was first tested for accuracy by cross checking the genotypes with duplicates samples (minimum of 17%). The data was then tabulated with allele frequencies for the microsatellite in both cases and controls. This was fed in to CLUMP, which performed a number of χ^2 tests for association.

CLUMP is a program designed to assess the significance of the departure of observed values, within a 2 x N contingency table, from the expected values conditional on the marginal totals. The significance is assessed using a Monte Carlo

approach, by performing repeated simulations on randomly generated tables with the same marginal values using simulated data and counting the number of times the χ^2 value is associated with the table under consideration by chance. This means that the significance levels assigned is unbiased because the accuracy depends on the number of simulations. It also means that no special considerations need to be made for small expected values (Sham & Curtis 1995).

CLUMP was used to calculate maximal χ^2 values by clumping together columns into a new two by two table in a way that is designed to maximise the χ^2 value. The method produces an "inflated" χ^2 value but does not cause any problems in interpretation because its significance is assessed using a Monte Carlo method, which takes into account the increased chance of a false positive result.

The output of CLUMP generates four χ^2 results and their significance is evaluated by testing how many times the value produced is exceeded by chance with the Monte Carlo simulation method. Subtests of the CLUMP program include:

T1:

Pearson's χ^2 statistic of the 'raw' 2 x N contingency table. The χ^2 value is calculated from the original table supplied. However, if there are small expected values (< 5) in some cells, then these values might not follow the expected distribution of a chi-squared statistic, with N-1 df. In spite of this, the significance can be reliably accessed using the Monte Carlo simulations.

T2:

The χ^2 statistic of a table with rare alleles grouped together to prevent small expected cell counts. If any cell has a value less than 5, then its column is clumped together with the column of the next smallest value. This process is repeated until all columns have values of 5 or more. The significance is assessed by Monte Carlo simulations and the df is N-1, where N is the number of resulting columns after they were clumped.

T3:

The largest of the χ^2 statistics of 2 x 2 tables, each of which compares one allele against the rest grouped together. The 2 x 2 table is obtained by comparing one column of the original table against the total of all other columns. This tests the hypothesis that there is one particular column with a number of cells deviating from the expected values. For each column in turn, all the other columns are combined to assess the χ^2 value. The columns that produce the maximal χ^2 value is then used and the results are outputted. Columns containing values less than 5 are not considered and these columns are clumped together with the next column of lowest value. In this case, the T2 table would be used. The significance of the result is assessed with Monte Carlo simulations.

T4:

The largest of the χ^2 statistics of all possible 2 x 2 tables, which compares any combination of alleles against the rest to produce the maximum χ^2 value possible. The 2 x 2 table is obtained by clumping the columns of the original table, so as to maximise the chi-squared value. Its significance is assessed using Monte Carlo simulations.

To select regions of interest for further investigation, the most significant output of the four CLUMP subtest statistics were used to show the degree of allelic association for each microsatellite marker. Likewise, the pointwise P values for markers was considered individually rather than attempting to derive an overall P value for the region on the basis of combining results from multiple markers. CLUMP employs an empirical Monte Carlo test of significance that does not require further correction for multiple alleles.

In order to run the CLUMP, the program requires the number of columns (alleles) in the table, the number of permutations to perform and a random number that starts the calculation in order to produce the pseudo-random tables for Monte Carlo

simulations. The number of permutation should be significantly large to produce a satisfactorily accurate estimate of the true significance that is achieved. 9,999 or more permutations were generally used in this study.

3.6 GENECOUNTING

Multilocus genotypes were analysed for haplotypic association with schizophrenia using GENECOUNTING (Zhao et al. 2000; Zhao et al. 2002; Curtis et al. 2006). The software contains several programs:

3.6.1 LDPAIRS

Pairwise LD between all markers was calculated with LDPAIRS. This program determines LD between any set of selected markers and calculates a P value for the statistical significance of the test for LD between the markers. Using estimated haplotype frequencies, it can also measure Cramer's V, which is equivalent to the square root of r^2 (Section 3.3.2). In addition, the absolute value of the D' statistic (Section 3.3.1) between the commonest alleles at each locus are revealed (Bishop et al. 1975).

3.6.2 SCANGROUP

SCANGROUP compares haplotype frequencies in cases and controls of haplotypes built from subsets of markers, consisting either of groups of contiguous markers or of all possible combinations of a specified number of markers. The program is useful for detecting any genotyped plates that may have been swapped over, which would produce aberrant results.

3.6.3 RUNGC

RUNGC estimates haplotype frequencies between unrelated subjects and has the ability to deal with multiallelic markers and markers with missing genotypes. The program computes maximum-likelihood estimates of haplotype frequencies from phase-unknown case-control data. The significance of any overall haplotypic association with schizophrenia is then obtained with a permutation test.

Tests for association are based on likelihood ratio tests comparing estimated haplotype frequencies in the control samples with the estimated frequencies in the case sample. The null hypothesis is that haplotypes have the same frequencies in cases and controls. RUNGC examines haplotypic association when taking a global (overall) analysis of haplotype frequency differences. Permutation testing is then undertaken to generate an empirical significance value for the strength of the haplotypic association.

This program displays the frequencies of haplotypes and groups them side by side for comparison. By adding a second output parameter (.hap), the RUNGC program is able to show grouped individuals containing their haplotypes. This enables the selection of individuals with the most significant haplotypes for specific sequencing of those "most at risk individuals".

4.0 A GENETIC ASSOCIATION STUDY TO FINE MAP A CHROMOSOME 11q22-24 SCHIZOPHRENIA SUSCEPTIBILITY GENE

4.1 INTRODUCTION

Previous linkage analyses of multiply affected schizophrenic families have confirmed the involvement of the chromosome 11q22-24 region, with LOD scores of 3.4 and 3.1 (Maziade et al. 1995; Gurling et al. 2001). As a result, a genetic association study was performed to fine map a chromosome 11q22-24 schizophrenia susceptibility gene using the UCL case-control sample.

4.2 METHODS

Five microsatellite markers in the chromosome 11q22-24 region were randomly selected for fine mapping. The markers, D11S1998, D11S925, D11S964, D11S836 and D11S934, are all located close to the maximum LOD score position in the UCL schizophrenia family linkage study (Gurling et al. 2001). Further microsatellite markers, D11S29, AAT13, D11S4127, TTTC20.2, D11S939, D11S976, CA22, and D11S4195, were then chosen non-randomly based on the findings of the original five microsatellites genotyped. Unformatted microsatellite markers AAT13, TTTC20.2 and CA22 were identified from simple repeats found in the UCSC Genome Browser database, May 2004 assembly. These new repeats were formatted by us for genotyping. The forward and reverse primers were created using the Primer3 design software (Rozen & Skaletsky 2000). Primer sequences for D11S29, D11S4127, D11S1998, D11S939, D11S976, D11S4195, D11S925, D11S964, D11S836 and D11S934 were all obtained from the GDB Human Genome Database and Ensembl. The location and primer sequences for all genotyped microsatellite markers are shown in Table 4.1.

Table 4.1: Microsatellite marker positions from UCSC Genome Browser database, March 2006 assembly, used for fine mapping of schizophrenia on chromosome 11q22-24.

Chromosome 11 Marker	Marker Location (bp)	Polymorphism	Forward Primer	Reverse Primer
D11S29	116,993,514	$(AC)_{24}$	TCTAGCTCCACCATCCTGTG	ACAACACACTGCCACAAGAC
AAT13	117,138,789	$(AAT)_{13}$	CCATGAGGTGGAGGTTGC	TTTAAGCAACGTGGGGATCT
D11S4127	117,149,328	$(CA)_{23}$	ATGAGAAGTGCCATCCAGC	ACTATGCCCAGTGTGTGTGC
D11S1998	117,202,941	$(GATA)_{24}$	AGCCATCAACTAGCTTTCCC	GGGAGGCACCAACAGATG
TTTC20.2	117,234,550	$(TTTC)_{20}$	CGGAGAGTCGATTCTGCTTC	ATCGCACCACTGCACACTC
D11S939	117,311,513	$(AC)_{30}$	TCAGTTTCCTCATCCACCAA	AAAAGCACATATTTTGCTCAGAGTC
D11S976	117,421,197	$(CA)_{15}$	TCAGTTCAAGTTGCAGAAGAA	CAACTTGGTGACAGCCTTTC
CA22	117,466,730	$(CA)_{22}$	AGGCTGCAGTGAGCTGAAAT	TGCTAGAAGCTGAGGCACAA
D11S4195	117,501,743	$(AC)_{23}$	GTGGCCCAGGCTGTTC	GCTGCTAAATGTCACACTGAGA
D11S925	120,333,421	$(TG)_{17}$	AGAACCAAGGTCGTAAGTCCTG	TTAGACCATTATGGGGGCAA
D11S964	123,087,535	$(TTTC)_{51}$	ACTTCAGCCTCGGTGACAG	TGTTCTGCCTCTGTTGTTAC
D11S836	123,169,313	$(TG)_{19}$	GCCTCTGAAGTGGCTAAATA	CCCCTCACCACATCACTTG
D11S934	125,585,522	$(AC)_{15}$	GCTGTCCCTGACAACTACATGC	TTCCATCAGAACTGGGAATGAG

Genotyping of microsatellite markers was done in an initial UCL case-control sample of 443 controls and 480 schizophrenic patients. PCR amplification of the microsatellite markers was performed using an M13 tailed primer and a second non-tailed primer. A third universal M13 sequence primer, labelled with infrared dye IRD700 or IRD800, was used to hybridise against the M13 tailed locus specific primer. Microsatellite marker fragment sizes were then separated and visualised with either of the two infrared dyes on LiCor 4200L sequencers. Genotyping was accomplished using SAGA-GT genotyping software. The genotypes were then checked by eye, with all allele calls verified by a second independent person blind to initial calling. Any discrepant genotypes between the two assessors were repeated from the PCR amplification stage.

In order to follow association found with microsatellite markers, SNPs rs11999, rs869789, rs529623, rs479991, rs2019655, rs10790212, rs3087563, rs1815774, rs876798, rs876797, rs4938445, rs4938446, rs11216598 and rs631898 were all chosen from Ensembl. Further SNPs rs4579962, rs678776, rs11216567, rs10892181, rs3168238, rs564989, rs12363888, rs11608153, rs512481, rs476130, rs7121573, rs873713, rs10790218, rs11605223, rs3809043, rs3809042 and rs497768 were then selected from the International HapMap Project (CEU population) based on the results of the Tagger function within the Haploview version 3.32 software program (Barrett et al. 2005; de Bakker et al. 2005). The defaults settings employed were aggressive tagging use of two and three marker haplotypes, r^2 threshold of 0.8 and a LOD threshold for multimarker tests of 3.0.

All SNP markers were genotyped by Kbioscience using the KASPar SNP genotyping procedure (Section 2.7.2). These SNPs were genotyped in an increased UCL case-control sample, now consisting of 496 unrelated cases and 488 ancestrally matched controls. Of the samples, 17% from each microtitre plate were reduplicated in order to detect possible errors and to confirm the reproducibility of genotypes. The data were then analysed to confirm Hardy-Weinberg equilibrium (HWE). Any markers lacking HWE were repeated using an alternative Taqman method. Before

any association analysis, the genotype data were assessed using the SCANGROUP program from the GENECOUNTING software (Zhao et al. 2000; Zhao et al. 2002; Curtis et al. 2006). This program highlights any potential genotyping errors by identifying significant differences in haplotypic frequencies between any single 96 well microtitre plate and haplotypic frequencies in all other plates combined. It is argued that genotyping errors can materialise from rare haplotypes occurring on just a single plate, whereas true haplotypic associations with the disease will appear as differences in haplotype frequencies between cases and controls spread across a number of different plates.

Allelic association analysis with schizophrenia was performed using a simple χ^2 test for biallelic SNPs (Section 3.5.1) and the software CLUMP for microsatellite markers (Section 3.5.2). Multilocus genotypes were then analysed for haplotypic association with schizophrenia using the RUNGC program within the GENECOUNTING software. This computes for maximum likelihood estimates of haplotype frequencies from phase unknown case-control data. The significance of any overall haplotypic association with schizophrenia was then obtained with a permutation test (Zhao et al. 2000; Zhao et al. 2002; Curtis et al. 2006). GENECOUNTING was also used to calculate pairwise LD between all markers with use of the LDPAIRS program. This was visualised using the LocusView 2.0 software obtained from the Broad Institute website: http://www.broad.mit.edu/mpg/locusview (Petryshen et al. 2003).

4.3 RESULTS

4.3.1 MICROSATELLITE GENOTYPING RESULTS

One of the five randomly selected microsatellites that was initially genotyped in chromosome 11q22-24 region, D11S1998, showed allelic association with schizophrenia (P = 0.021; CLUMP T2). After this, other microsatellite markers to be genotyped were selected non-randomly by virtue of being close to marker D11S1998. From these non-random microsatellites, TTTC20.2 was also found to be

associated with schizophrenia (P = 0.048; CLUMP T4). All other microsatellite markers genotyped, did not show association with schizophrenia; P values are for CLUMP T1 statistic with all CLUMP results displayed in Table 4.2: D11S29 (P = 0.899), AAT13 (P = 0.423), D11S4127 (P = 0.521), D11S939 (P = 0.460), D11S976 (P = 0.393), CA22 (P = 0.597), D11S4195 (P = 0.287), D11S925 (P = 0 .379), D11S964 (P = 0.132), D11S836 (P = 0.423) and D11S934 (P = 0.561).

Table 4.2: Allelic association tests of microsatellite markers with schizophrenia at the chromosome 11q22-24 region.

C11 Marker and Sample	Marker Location (bp)	Prior Marker Distance (bp)	Number of Observed Alleles at each Respective Fragment Size by Population																							P from CLUMP Monte Carlo Subtest			
																										T1	T2	T3	T4
D11S29	116,993,514		165	167	169	171	173	175	177	179	181	183	185	187	193											0.899	0.669	0.336	0.454
Control			160	4	1	4	15	196	54	63	69	17	9	1	1														
Case			115	4	1	5	15	149	40	50	76	14	10	1	0														
AAT13	117,138,789	145,275	235	241	244	247	250	253	256	259	262	265	268													0.423	0.446	0.436	0.418
Control			39	632	4	57	24	0	7	22	44	4	1																
Case			33	481	3	44	12	4	5	28	36	3	1																
D11S4127	117,149,328	10,539	99	105	107	109	111	113	115	117	119	121	123	125												0.521	0.426	0.333	0.863
Control			0	2	4	19	180	130	120	290	90	5	1	1															
Case			1	5	6	29	152	123	94	268	81	5	0	0															
D11S1998	117,202,941	53,613	150	158	162	166	170	174	178	182	186															0.030	0.021	0.094	0.058
Control			60	1	2	38	405	275	75	7	1																		
Case			86	0	0	27	428	239	60	6	0																		
TTTC20.2	117,234,550	31,609	232	236	240	242	244	246	248	250	252	254	256	258	260	262	264	266	268	270	274	278	282			0.057	0.056	0.144	0.048
Control			8	3	22	3	130	11	97	54	113	65	75	60	31	60	12	46	3	17	9	2	1						
Case			4	14	16	5	130	10	137	47	104	77	69	62	51	57	14	32	4	23	5	1	0						
D11S939	117,311,513	76,963	255	257	259	261	263	265	267																	0.460	0.278	0.480	0.587
Control			0	3	195	71	4	216	49																				
Case			1	5	178	60	7	164	49																				
D11S976	117,421,197	109,684	135	137	139	141	143	145	147	149	151	153	155	157	159											0.393	0.508	0.563	0.692
Control			2	73	78	83	8	20	22	201	118	18	2	9	2														
Case			0	80	79	82	16	23	15	181	101	13	6	4	2														
CA22	117,466,730	45,533	244	246	248	250	252																			0.597	0.515	0.465	0.692
Control			9	68	559	4	0																						
Case			8	52	533	6	1																						
D11S4195	117,501,743	35,013	286	290	294	296	298	300	302	304	306	308	310	312												0.287	0.345	0.273	0.309
Control			1	0	3	3	293	8	20	48	142	13	35	6															
Case			0	1	4	7	222	7	15	46	87	14	35	2															
D11S925	120,333,421	2,831,678	186	194	196	198	202	204	206	208	210	212	213	214	216	217	218	220	222	224	226					0.379	0.227	0.937	0.161
Control			1	189	21	0	2	0	2	19	48	80	2	60	74	0	53	9	1	1	0								
Case			0	175	18	1	4	1	1	12	48	81	4	53	59	5	40	13	4	1	2								
D11S964	123,087,535	2,754,114	282	286	290	294	298	302	304	306	308	310	312	314	316	318	320	322	324	326	328	330				0.132	0.218	0.311	0.342
Control			3	128	34	50	223	174	3	96	1	64	3	22	4	9	2	4	1	0	1	2							
Case			1	95	36	54	211	153	2	117	5	65	1	22	2	2	3	2	0	3	0	0							
D11S836	123,169,313	81,778	235	241	244	247	250	253	256	259	262	265	268													0.423	0.446	0.436	0.418
Control			39	632	4	57	24	0	7	22	44	4	1																
Case			33	481	3	44	12	4	5	28	36	3	1																
D11S934	125,585,522	2,416,209	183	195	197	199	201	203	205	207	209	211	213	215	217	219	221	223	225	227	229					0.561	0.347	0.119	0.292
Control			12	3	5	107	40	94	142	39	18	3	2	8	14	39	21	7	8	3	1								
Case			5	5	4	123	28	104	136	18	14	2	2	6	12	34	18	7	6	2	0								

4.3.2 SNP GENOTYPING RESULTS

After microsatellite analysis, SNP markers localised near positively associated markers D11S1998 and TTTC20.2 were all considered for genotyping. The first set of SNP markers genotyped were chosen from Ensembl and included SNPs rs11999, rs869789, rs529623, rs479991, rs2019655, rs3087563, rs876797, rs4938446 and rs631898. From this, SNP rs4938446 (P = 0.025) was found to be associated with schizophrenia with two further SNPs, rs869789 (P = 0.097) and rs876797 (P = 0.086), showing a trend toward association (Table 4.3). The second set of SNP markers genotyped were also selected from Ensembl and consisted of SNPs rs876798, rs4938445 and rs11216598. They were specifically chosen by virtue of having a high LD to SNP rs4938446, based on the D' value. This set of markers also included SNPs rs10790212 and rs1815774, whose assay failed to work in the first set of SNP markers and were now redesigned using the reverse strand. From this set of markers, SNPs rs1815774 (P = 0.049) and rs4938445 (P = 0.010) were both found to be associated with schizophrenia.

Finally, SNP markers selected from the International HapMap Project using Haploview's Tagger function were genotyped (Barrett et al. 2005; de Bakker et al. 2005). Two SNPs, rs3168238 (P = 0.009) and rs497768 (P = 0.023), were found to be associated with schizophrenia. A further four SNPs, rs11216567 (P = 0.072), rs7121573 (P = 0.076), rs873713 (P = 0.091) and rs10790218 (P = 0.099), all showed a trend toward association. None of the other SNPs genotyped were found to be associated with schizophrenia.

Table 4.3: Allelic association tests of SNP markers with schizophrenia at the chromosome 11q23.3 locus.

Chromosome 11 Marker and Sample	Marker Location (bp)	Prior Marker Distance (bp)	Number of Observed Alleles for each Respective Allelic Base		χ^2	P[a]
rs11999	117196128		A	C	0.91	0.340
Control			542	300		
Case			521	261		
rs869789	117196596	468	A	G	2.75	0.097
Control			107	753		
Case			138	772		
rs4579962	117197544	948	A	G	0.03	0.860
Control			274	688		
Case			266	680		
rs529623	117198465	921	C	T	0.60	0.440
Control			473	395		
Case			475	427		
rs678776	117201954	3489	A	G	0.00	0.975
Control			285	671		
Case			282	666		
rs479991	117202168	214	A	G	1.17	0.280
Control			665	197		
Case			715	187		
D11S1998	117202941	773	See Table 4.2		12.60	0.021[b]
Control						
Case						
rs11216567	117203693	752	A	G	3.25	0.072
Control			880	70		
Case			829	89		
rs10892181	117204796	1103	G	T	0.49	0.484
Control			351	609		
Case			320	594		
rs2019655	117205088	292	A	G	0.60	0.437
Control			208	664		
Case			230	674		
rs10790212	117207900	2812	C	T	0.47	0.493
Control			640	240		
Case			666	232		
rs3087563	117213147	5247	C	T	0.09	0.763
Control			415	451		
Case			422	472		
rs3168238	117214855	1708	G	T	6.80	0.009
Control			76	874		
Case			47	887		
rs564989	117214964	109	C	T	0.30	0.584
Control			222	742		
Case			226	712		
rs12363888	117218397	3433		T	0.04	0.843
Control			907	55		
Case			909	53		
rs11608153	117221873	3476	A	G	0.48	0.489
Control			821	129		
Case			805	115		

rs512481 Control Case	117224090	2217	A 452 425	G 494 507	0.90	0.344
rs476130 Control Case	117233093	9003	C 890 853	T 70 77	0.64	0.423
TTTC20.2 Control Case	117234550	1457	See Table 4.2		17.54	0.048[c]
rs7121573 Control Case	117235040	490	A 382 331	G 564 579	3.15	0.076
rs1815774 Control Case	117236649	1609	C 342 311	G 528 583	3.87	0.049
rs876798 Control Case	117242650	6001	C 125 123	T 739 797	0.45	0.503
rs876797 Control Case	117242957	307	A 207 186	C 651 712	2.94	0.086
rs873713 Control Case	117245126	2169	A 230 197	T 710 733	2.86	0.091
rs10790218 Control Case	117247205	2079	A 374 334	G 576 602	2.73	0.099
rs4938445 Control Case	117250213	3008	A 299 250	G 579 632	6.68	0.010
rs4938446 Control Case	117250259	46	A 296 271	T 570 655	5.00	0.025
rs11216598 Control Case	117253662	3403	A 727 771	G 149 153	0.07	0.798
rs631898 Control Case	117253688	26	A 419 466	G 451 434	2.31	0.128
rs11605223 Control Case	117254054	366	C 490 462	G 470 494	1.41	0.235
rs3809043 Control Case	117254325	271	C 929 898	G 29 34	0.57	0.452
rs3809042 Control Case	117254330	5	C 27 23	G 925 899	0.21	0.646
rs497768 Control Case	117255950	1620	C 416 371	G 500 553	5.21	0.023

[a] Two-tailed significance (P) from 2 x 2 χ^2, with 1 df.

[b] P from CLUMP Monte Carlo subtest T2.

[c] P from CLUMP Monte Carlo subtest T4.

4.3.3 HAPLOTYPIC ASSOCIATION RESULTS

Tests of three marker haplotypic association using SNPs rs11216598, rs11605223 and rs497768 (Hap-F7) provided significant evidence for association to schizophrenia, with a global empirical permutation significance of P = 0.0013 (Table 4.4). Here, the haplotype A-G-G is estimated to be present in 11.4% of controls and an elevated value of 16.8% in cases. Several other two, three and four marker tests of haplotypic association were also found to be positively associated.

Table 4.4: Haplotypic association tests with schizophrenia at the chromosome 11q23.3 locus.

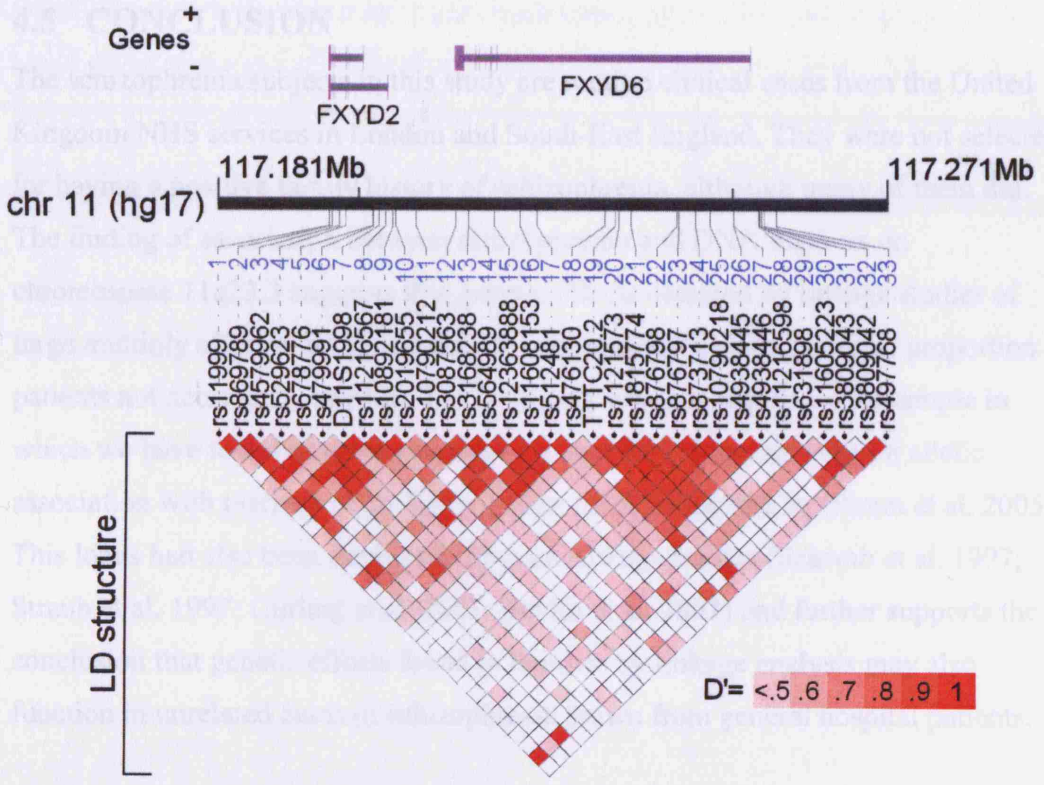
Number of Markers and Haplotype Reference	Haplotype	Global Empirical P[a]	Alleles Increasing in Cases	Estimated Haplotype Frequency (%)	
				Controls	Cases
2:					
HAP-F1	rs3168238-rs497768	0.0028	T-G	47.9	56.0
3:					
HAP-F2	rs564989-rs476130-rs876797	0.0043	C-C-C	11.4	14.9
HAP-F3	rs564989-rs476130-rs497768	0.0048	T-C-G	41.6	48.7
HAP-F4	rs564989-rs876797-rs873713	0.0010	C-C-T	16.9	23.8
HAP-F5	rs476130-rs11605223-rs497768	0.0027	C-G-G	11.3	17.1
HAP-F6	rs1815774-rs4938445-rs4938446	0.0066	G-G-T	58.1	64.0
HAP-F7	rs11216598-rs11605223-rs497768	0.0013	A-G-G	11.4	16.8
HAP-F8	rs11605223-rs3809043-rs497768	0.0021	G-C-G	11.5	17.3
4:					
HAP-F9	rs3168239-rs564989-rs876797-rs873713	0.0013	T-C-C-T	17.9	23.1
HAP-F10	rs3168238-rs12363888-rs11605223-rs497768	0.0057	T-C-G-G	10.8	17.2
HAP-F11	rs3168238-rs476130-rs11605223-rs497768	0.0026	T-C-G-G	10.7	16.9
HAP-F12	rs3168238-rs11605223-rs3809043-rs497768	0.0016	T-G-C-G	10.8	17.3
HAP-F13	rs3168238-rs11605223-rs3809042-rs497768	0.0035	T-G-G-G	10.9	17.2
HAP-F14	rs1815774-rs10790218-rs4938445-rs11605223	0.0055	G-G-G-G	40.5	46.7
HAP-F15	rs1815774-rs4938445-rs4938446-rs11605223	0.0159	G-G-T-G	40.6	47.2
HAP-F16	rs10790218-rs4938445-rs4938446-rs11605223	0.0066	G-G-T-G	40.8	46.3
HAP-F17	rs11216598-rs11605223-rs3809043-rs497768	0.0019	A-G-C-G	11.4	16.8

[a] Haplotype permutation test empirical P, based on 9,999 permutations.

4.3.4 LINKAGE DISEQUILIBRIUM RESULTS

LD between all markers genotyped in the chromosome 11q23.3 locus was examined using LDPAIRS, an accessory programme of the GENECOUNTING software. The result of pairwise LD statistics between all markers genotyped and their respective locations is illustrated in Figure 4.1, with the use of LocusView 2.0 (Petryshen et al. 2003). A higher D′ value is displayed when there is greater red intensity.

Figure 4.1: Pairwise LD between markers genotyped in the chromosome 11q23.3 locus.



LD structure (D') between marker pairs is indicated by the coloured matrices. Marker locations are provided in Table 4.3 and their locations are shown relative to both the FXYD genes displayed above the chromosome bar. The gene structures contain vertical lines representing exons. Genomic positions are in accordance with the National Centre for Biotechnology Information build 35 hg 17 of the human genome assembly. The figure was generated using LocusView 2.0 (Petryshen et al. 2003).

4.4 DISCUSSION

Nearly all the genotyped markers that show significant association with schizophrenia or a positive trend toward association are localised within the FXYD domain containing ion transport regulator 6 (FXYD6) gene. The exceptions to these are markers rs869789, D11S1998, and rs11216567, situated in FXYD2. These three markers show significant LD to several of the markers from FXYD6. Therefore, the evidence from the genetic association analysis points toward FXYD6 as the most likely gene for susceptibility to schizophrenia.

4.5 CONCLUSION

The schizophrenia subjects in this study are routine clinical cases from the United Kingdom NHS services in London and South-East England. They were not selected for having a positive family history of schizophrenia, although many of them did. The finding of association between schizophrenia and DNA markers on chromosome 11q23.3 suggests that genetic effects detected by linkage studies of large multiply affected pedigrees may also be operative in a significant proportion of patients not necessarily having a family history of schizophrenia. The sample in which we have found positive association with FXYD6 has also shown allelic association with markers at the Epsin 4 gene on chromosome 5q (Pimm et al. 2005). This locus had also been implicated by prior linkage studies (Schwab et al. 1997; Straub et al. 1997; Gurling et al. 2001; Devlin et al. 2002) and further supports the conclusion that genetic effects found in families by linkage analysis may also function in unrelated cases of schizophrenia drawn from general hospital patients.

To verify the significance of the positive allelic and haplotypic associations with schizophrenia found in the UCL sample, replication was attempted in a Scottish case-control sample provided to us by the University of Aberdeen. Confirmation of genetic association with schizophrenia requires attention to the statistical power of case-control replication studies. The critical issue is the proven presence of locus heterogeneity for schizophrenia. The linkage studies of schizophrenia have shown strong evidence of heterogeneity, with different subgroups of families containing

different susceptibility loci. Therefore, any one susceptibility gene will increase susceptibility to schizophrenia in only a minority of cases. The UCL schizophrenia sample has the power to detect allelic association of 0.99 at $P < 0.05$ and of 0.91 at $P < 0.001$, with the assumption of complete LD between markers and disease alleles when the minor marker allele frequency is < 10% and with a difference of 5% in allele frequencies between cases and controls. If the allele frequency is as high as 50%, the power is 0.90 at $P = 0.05$, with a 10% allele frequency difference. However, the power of the sample is only 0.41 at $P = 0.05$ for a 5% allele frequency difference. Further attempts at replication should aim to have 0.8 power for a significance of 0.05. This may require 1,300 cases and 1,300 controls, even with the assumption of complete LD between markers and disease mutations. All power values were calculated using a Genetic Power Calculator obtained from the website: http://pngu.mgh.harvard.edu/~purcell/gpc (Purcell et al. 2003).

5.0 CONFIRMATION OF THE GENETIC ASSOCIATION BETWEEN THE FXYD6 GENE AND SCHIZOPHRENIA

5.1 INTRODUCTION

In our genetic association study of chromosome 11q22-24 (Section 4.0), seven microsatellite or SNP markers revealed empirically significant allelic association with schizophrenia in the UCL case-control sample (Choudhury et al. 2007). These markers, D11S1998, rs3168238, TTTC20.2, rs1815774, rs4938445, rs4938446 and rs497768, are all localised within or near the FXYD6 gene. A further six SNPs, rs869789, rs11216567, rs7121573, rs876797, rs873713 and rs10790218, displayed a trend toward association. Several haplotypes were also found to be associated with schizophrenia e.g. haplotype Hap-F21 comprising of markers rs10790212-rs4938445-rs497768, produced a global empirical permutation significance of P = 0.002. To confirm the importance of the positive associations discovered in the UCL sample, several SNP markers were selected for additional genotyping in an independent Scottish case-control sample provided by the University of Aberdeen.

5.2 BRIEF METHODS AND MATERIALS

All methods employed were as described in Sections 2.0 and 3.0. The Aberdeen case-control sample consisted of 858 schizophrenia subjects and 591 "supernormal" controls. This gives a total of 2,433 individuals available (1,354 with schizophrenia and 1,079 of controls), when combined with the UCL sample. The cases were recruited through Scottish psychiatric hospitals and met DSM-III-R or DSM-IV criteria of schizophrenia, with use of an operational criteria checklist (OPCRIT). Diagnosis of schizophrenia was confirmed through agreement by two independent senior psychiatrists. This was based on a combination of psychiatric case note inspection and in the majority of cases (n=717) via the structured clinical interviews for DSM-III-R/DSM-IV (SCID). All controls were drawn from the same region of Scotland and were ethnically matched with the case population. They were enlisted

as volunteers though general practices and screened for absence of psychiatric illness. The samples used for controls were also sex matched for research purposes. Informed consent was obtained from all schizophrenic patients and control individuals. A proportion of the Aberdeen case-control sample has been employed previously in a genetic association test between NRG1 and schizophrenia (Stefansson et al. 2003).

To validate the findings of significant allelic and haplotypic association between genetic markers at FXYD6 and schizophrenia, replication in an independent case-control sample from the University of Aberdeen had been performed. This was done using SNPs rs869789, rs11216567, rs10790212, rs3168238, rs876797, rs4938445 and rs497768. The genotyping of SNP markers in the Aberdeen case-control sample were determined and analysed using the same procedure as described for the UCL sample (Section 4.2).

5.3 RESULTS

With a one-tailed test of significance, SNPs rs4938445 (P = 0.044) and rs497768 (P = 0.037) were both found to show allelic association with schizophrenia in the Aberdeen case-control sample (Table 5.1). The allele frequency increase for SNP rs3168238 (P = 0.074), although not significantly associated with schizophrenia, displayed a trend in the opposite direction to that found in the UCL sample. This may indicate that aetiological base pair changes are present on different haplotypic backgrounds within the two samples or that the disease haplotype frequencies are different between the North and the South of the UK.

When the data from the UCL and the Aberdeen samples were combined and analysed together, SNPs rs869789 (P = 0.023), rs4938445 (P = 0.002) and rs497768 (P = 0.005) all showed significant allelic association with schizophrenia using a two-tailed test of significance (Table 5.1). For these three SNPs, evidence for association was stronger in the combined samples then in the individual UCL and Aberdeen

samples. Two further SNPs, rs11216567 (P = 0.058) and rs876797 (P = 0.063), also displayed a trend toward association when samples were combined (Table 5.1).

Table 5.1: Allelic association tests with schizophrenia at the FXYD6 locus in the UCL, Aberdeen and combined samples.

Chromosome 11 Marker and Sample	Marker Location (bp)	Prior Marker Distance (bp)	UCL[a] No. (%) of Observed Alleles for Allelic Base		χ^2	P[b]	Aberdeen No. (%) of Observed Alleles for Allelic Base		χ^2	P[c]	Combined (UCL and Aberdeen) No. (%) of Observed Alleles for Allelic Base		χ^2	P[b]
rs869789	117196596		A	G	2.75	0.097	A	G	2.00	0.079	A	G	5.17	0.023
Control			107 (0.12)	753 (0.88)			173 (0.15)	985 (0.85)			280 (0.14)	1738 (0.86)		
Case			138 (0.15)	772 (0.85)			282 (0.17)	1384 (0.83)			420 (0.16)	2156 (0.84)		
rs11216567	117203693	7097	A	G	3.25	0.072	A	G	2.48	0.058	A	G	3.59	0.058
Control			880 (0.93)	70 (0.07)			1040 (0.92)	108 (0.08)			1920 (0.92)	178 (0.08)		
Case			829 (0.90)	89 (0.10)			1528 (0.92)	128 (0.08)			2357 (0.92)	217 (0.08)		
rs10790212	117207900	4207	C	T	0.47	0.493	C	T	0.04	0.420	C	T	0.43	0.510
Control			640 (0.73)	240 (0.27)			860 (0.74)	296 (0.26)			1500 (0.74)	536 (0.26)		
Case			666 (0.74)	232 (0.26)			1242 (0.75)	420 (0.25)			1908 (0.75)	652 (0.25)		
rs3168238	117214855	6955	G	T	6.80	0.009	G	T	3.18	0.074[b]	G	T	0.501	0.479
Control			76 (0.08)	874 (0.92)			47 (0.04)	1103 (0.96)			123 (0.06)	1977 (0.94)		
Case			47 (0.05)	887 (0.95)			92 (0.06)	1558 (0.94)			139 (0.05)	2445 (0.95)		
rs876797	117242957	28102	A	C	2.94	0.086	A	C	0.82	0.182	A	C	3.45	0.063
Control			207 (0.24)	651 (0.76)			250 (0.22)	882 (0.78)			457 (0.23)	1533 (0.77)		
Case			186 (0.21)	712 (0.79)			337 (0.21)	1295 (0.79)			523 (0.21)	2007 (0.79)		
rs4938445	117250213	7256	A	G	6.68	0.010	A	G	2.94	0.043	A	G	9.61	0.002
Control			299 (0.34)	579 (0.66)			376 (0.33)	770 (0.67)			675 (0.33)	1349 (0.67)		
Case			250 (0.28)	632 (0.72)			485 (0.30)	1145 (0.70)			735 (0.29)	1777 (0.71)		
rs497768	117255950	5737	C	G	5.21	0.023	C	G	3.21	0.037	C	G	7.87	0.005
Control			416 (0.45)	500 (0.55)			521 (0.45)	633 (0.55)			937 (0.45)	1133 (0.55)		
Case			371 (0.40)	553 (0.60)			676 (0.42)	944 (0.58)			1047 (0.41)	1497 (0.59)		

[a] Original data as shown in Table 4.3 (Choudhury et al. 2007).

[b] Two-tailed significance (P) from 2 x 2 χ^2, with 1 df.

[c] One-tailed significance (P) from 2 x 2 χ^2, with 1 df.

The three marker haplotype rs10790212-rs4938445-rs497768 (Hap-F21), which produced a global empirical association in the UCL sample (empirical P = 0.0024), was also found to be replicated in the Aberdeen case-control sample (empirical P = 0.0127). When data from both samples were combined, haplotype Hap-F21 exhibited strengthened empirical significance for haplotypic association with schizophrenia (empirical P = 0.0005). Other significantly associated haplotypes in the UCL sample were also found to be replicated in the Aberdeen sample and when both case-control samples were combined (Table 5.2).

Table 5.2: Tests of haplotypic association with schizophrenia at the FXYD6 locus for comparison of UCL, Aberdeen and combined samples.

Number of Markers and Haplotype Reference	Haplotype and Sample	Global Empirical P[a]	Alleles Increasing in Cases	Estimated Haplotype Frequencies (%)	
				Controls	Cases
2:					
HAP-F18	rs10790212-rs4938445				
	UCL	0.0183	C-G	44.7	48.9
	Aberdeen	0.0266	C-G	44.6	49.7
	Combined	0.0093	C-G	44.7	49.4
HAP-F1	rs3168238-rs497768				
	UCL	0.0028	T-G	47.9	56.0
	Aberdeen	0.1173	T-G	52.0	54.2
	Combined	0.0157	T-G	50.1	54.9
3:					
HAP-F19	rs11216567-rs10790212-rs4938445				
	UCL	0.0038	A-C-G	40.2	44.4
	Aberdeen	0.0453	A-C-G	38.6	44.2
	Combined	0.0687	A-C-G	39.1	44.1
HAP-F20	rs11216567-rs10790212-rs497768				
	UCL	0.0442	A-C-G	35.2	40.3
	Aberdeen	0.0104	A-C-G	39.3	40.0
	Combined	0.0348	A-C-G	37.5	40.0
HAP-F21	rs10790212-rs4938445-rs497768				
	UCL	0.0024	C-G-G	23.4	27.9
	Aberdeen	0.0127	C-G-G	25.5	27.6
	Combined	0.0005	C-G-G	24.7	27.7

[a] Haplotype permutation test empirical P, based on 9,999 permutations.

5.4 DISCUSSION

The observation of allelic and haplotypic association in the UCL sample, with use of nominal empirically derived P values, and the fact that the same alleles and haplotypes were associated in the second Aberdeen sample, suggest that the association between FXYD6 and schizophrenia is unlikely to be a false positive result. The strongest positive associations were found with markers within FXYD6 and not in neighbouring genes, which leads us to believe that this gene was the most likely to be involved in the aetiology of schizophrenia. One microsatellite marker in FXYD2 was found to be associated with schizophrenia, but none of the SNPs within FXYD2 were significantly associated in either the UCL or the Aberdeen case-control samples. We found consistency as to which alleles and haplotypes were associated with schizophrenia in the UCL and the Aberdeen samples. Even though an allele at marker rs3168238 showed a trend to be associated with schizophrenia with an allele different from the one that showed significant association in the UCL sample, it was nevertheless found that the same allele was in the haplotype showing significant association with schizophrenia in the Aberdeen sample.

5.5 CONCLUSION

Association between FXYD6 markers and schizophrenia in a second case-control sample has been replicated and thus confirms FXYD6 as a gene increasing susceptibility to schizophrenia. There was increased statistical significance when both the UCL and the Aberdeen samples were combined. Further independent replications of association are now required from other populations.

The FXYD6 gene has been sequenced in schizophrenic subjects from our UCL sample, who have inherited the alleles and haplotypes contributing to the positive association found. This was done in order to identify aetiological base pair changes or mutations affecting the expression or function of FXYD6 that are involved in the molecular pathology of schizophrenia. These may be "private" mutations not yet identified or genetic variants that are already in databases.

6.0 FUNCTIONAL SCREENING OF THE FXYD6 GENE FOR AETIOLOGICAL BASE PAIR CHANGES OR MUTATIONS THAT INCREASE THE GENETIC SUSCEPTIBILITY TO SCHIZOPHRENIA

6.1 INTRODUCTION TO RESEQUENCING THE FXYD6 GENE

Using the UCL case-control sample, fine mapping of the chromosome 11q22-24 region has revealed the FXYD6 gene to be associated with schizophrenia (Section 4.0). The genetic association study was replicated in an independent case-control sample from a Scottish population provided to us by the University of Aberdeen (Section 5.0). This replication confirmed that DNA markers within the FXYD6 gene are significantly associated with schizophrenia and that FXYD6 is a putative susceptibility gene in causing the genetic disease (Choudhury et al. 2007).

The next step in substantiating FXYD6 as a schizophrenia susceptibility gene was to uncover DNA changes that may cause abnormal gene structure or expression. The genetic markers that were found to be associated with schizophrenia are common non-coding polymorphisms and thus not considered to be functional SNPs or aetiological base pair changes. SNP rs497768 is situated over 3.3 kb 5' of the FXYD6 gene, whereas SNP rs3168238 is located in the middle of the 9th intron and is therefore unlikely to be involved in affecting the splicing of the exon. The other three SNPs associated with schizophrenia, rs1815774, rs4938445 and rs4938446, are all localised to the first intron, which is over 18.3 kb long. Hence, it was necessary to resequence the FXYD6 gene as far as possible in an attempt to identify any base sequence abnormalities.

6.1.1 MUTATION MECHANISMS

To begin resequencing of the FXYD6 gene, it was first necessary to consider where any aetiological mutations were most likely to be found. Botstein and Risch (2003) reviewed the past successes of linkage and association studies to outline the different types of mutations that have been discovered in human diseases and the relative frequencies of these mutations (Table 6.1).

Table 6.1: The different mutations that lead to human diseases and their relative frequencies (Botstein & Risch 2003).

Change	Number	% of total
Deletion	6,085	21.8
Insertion/Duplication	1,911	6.8
Complex rearrangement	512	1.8
Repeat variations	38	0.1
Missense/Nonsense	16,441	58.9
Splicing	2,727	9.8
Regulatory	213	0.8
Total	27,027	100.0

The data indicates that most Mendelian clinical phenotypes are associated with alterations in normal coding sequence proteins, while very few (0.8%) are associated with regulatory changes. Linkage studies have proved very successful in the identification of rare, high-risk, disease associated mutations. This is because they have clear inheritance patterns in single large kindreds.

It is clearly important to consider where the mutations causing schizophrenia or other complex disorders are most likely to be found. Botstein & Risch (2003) favoured the coding regions of genes due to the fact that so far, moderate to low risk polymorphisms have mainly appeared to be found in this vicinity. Outside the coding regions, the mechanisms by which diseases arise are complex and varied. Eukaryotic genes encoding for proteins are usually organised into multiple exons, which must be spliced to produce the mRNA that is then translated into the protein. A 5' promoter element, contiguous with the transcription start site, is required to assemble the protein complex necessary for RNA synthesis (Levine & Tjian 2003).

For many genes, the region immediately upstream of the minimal promoter contains sufficient transcription factor binding sites to direct correct expression of the gene. These are regulatory promoters. Genes may also require multiple *cis*-acting distant genomic elements for correct gene expression to occur. These enhancers can be located upstream, within introns, or downstream of the transcription unit (Kleinjan et al. 2001). The transcription unit is made up of the transcribed exons and introns from the promoter region to the polyadenylation site. The genomic regions harbouring regulatory elements can stretch up to 1 Mb in either direction of the transcription unit (Pfeifer et al. 1999; Kimura-Yoshida et al. 2004). Some or all of these elements may exist within the introns of neighbouring genes, which often have no function related to the gene that they are localised within (Kleinjan et al. 2001; Lettice et al. 2002).

Genes can be classified into three groups based on the properties of their regulation:

1) Housekeeping genes are required for the functioning of most or all cells. In general, they are expressed ubiquitously and have promoters that are active in all cells without the need for enhancers.

2) Tissue-specific genes play a specific role in the particular function of the differentiated cell type. These genes are regulated through one or few specific enhancers.

3) Developmental regulator genes function in specific tissues at a distinct time period of development, sometimes at critically defined levels. They are strictly inactive in all other tissues and time points. These genes require multiple enhancer elements that need to be fitted altogether into the *cis* region surrounding the gene.

So far, most genes where disturbance of long-range control has been observed are key developmental regulators (Kleinjan & van Heyningen 2005). The effect of a regulatory element located far from its gene is clearly illustrated by the callipyge mutation found in sheep. The callipyge phenotype (CLPG) is characterised by hindquarter muscle overgrowth that only affects heterozygotes with paternal inheritance of the CLPG mutation. Using DNA from mutants and the mosaic founder individual, the CLPG mutation was identified as a single base substitution. The mutation was discovered in a region of strong sequence conservation 33 kb upstream of one of the genes involved (Freking et al. 2002). Undoubtedly, finding such mutations and their effects is evidently difficult. A further example of a small deletion or mutation in a *cis*-element is provided by the preaxial polydactyl mutation. Here, a single nucleotide substitution located 1 Mb from the causative gene produces a severe genetic defect (Lettice et al. 2003).

Multiple studies have recently revealed an abundance of copy number variation (CNV) within submicroscopic DNA segments (Redon et al. 2006). These CNVs range from kilobases to megabases in size and can take form as deletions, insertions, duplications or complex multisite variants (Iafrate et al. 2004; Sharp et al. 2005). They have been found in all humans and other mammals that have been examined (Freeman et al. 2006). CNVs influence gene expression, phenotypic variation and adaptation by disrupting genes and altering gene dosage. They can cause diseases e.g. microdeletion or microduplication disorders, or confer risk to complex disease traits such as HIV-1 and glomerulonephritis. The CNVs often represent an appreciable minority of causative alleles in genes where other forms of mutation are more strongly associated with the disease e.g. Parkinson's and Alzheimer's disease (Singleton et al. 2003; Rovelet-Lecrux et al. 2006). The Database of Genomic Variants reveals no evidence of CNVs at the FXYD6 locus in the general population (Iafrate et al. 2004; Zhang et al. 2006).

With regard to resequencing the FXYD6 gene, the decision was made to focus attention upon the exons, including their surrounding splice sites, and the area 5′ to the gene, which incorporates the promoter region.

6.2 METHODS AND SAMPLE SELECTION FOR SEQUENCING THE FXYD6 GENE

To discover the aetiological mutations within the FXYD6 gene requires resequencing the affected cases and control individuals. Schizophrenic samples were selected for sequencing following analysis of the haplotypic association results (Table 4.4). It can be seen that certain haplotypes had a marked increase of frequency in cases compared to the controls. These were selected for use in identifying samples likely to contain the aetiological base pair changes. Haplotype Hap-F7 (rs11216598, rs11605223 and rs497768) was chosen because it exhibited a high global permutation significance (empirical P = 0.0013) and contained a haplotype elevated in cases. Here, the A-G-G haplotype was present with estimated frequencies of 11.4% in controls and 16.8% in cases. Samples from the schizophrenic patients were chosen if they contained the A-G-G haplotype in either the homozygous or the heterozygous state. From this, 32 samples were selected for PCR amplification and the amplified PCR products were bi-directionally sequenced to screen for aetiological base pair changes. When mutations were found, they were screened in a further 32 control samples selected at random. This was done so as to determine whether the mutation were more frequent in the schizophrenic cases than in control individuals.

Sequencing of the exons included approximately 100 bp of the intronic region to take into account any abnormal intron/exon splice site variations. Primers were designed to amplify 250-600 bp, with any large sequencing regions covered by overlapping amplicons. The list of primer sequences used for sequencing the FXYD6 gene is listed in Table 6.2. Those primers that failed to amplify were redesigned with different annealing positions. Sequencing commenced using the Sanger-Coulson Chain termination method (Section 2.8.2).

Table 6.2: Forward and reverse primers used to sequence the FXYD6 exons and 5' upstream (promoter) region.

Location[a]	M13F Primer		M13R Primer	
	Name	Sequence[b]	Name	Sequence[b]
5' Upstream	FXYD6_Pr13_M13F	CACGACGTTGTAAAACGACtctggaagatgcgtgttca(	FXYD6_Pr13_M13R	GGATAACAATTTCACACAGGacaagcctaactgatgtacgtg
5' Upstream	FXYD6_Pr12_M13F	CACGACGTTGTAAAACGActcaggatgaaagcctgtcc	FXYD6_Pr12_M13R	GGATAACAATTTCACACAGGccctagccaacttcggttct
5' Upstream	FXYD6_Pr11_M13F	CACGACGTTGTAAAACGACaaggactctgcctggtcttg	FXYD6_Pr11_M13R	GGATAACAATTTCACACAGGtgcaaaggagtctgagaaaatg
5' Upstream	FXYD6_Pr10_M13F	CACGACGTTGTAAAACGACgctggctggagatcaaaatg	FXYD6_Pr10_M13R	GGATAACAATTTCACACAGGtaaatgtcgtgtcccagacc
5' Upstream	FXYD6_Pr09_M13F	CACGACGTTGTAAAACGACtcaggcagcgtatgtggata	FXYD6_Pr09_M13R	GGATAACAATTTCACACAGgtaatcccagggagtcagca
5' Upstream	FXYD6_Pr08_M13F	CACGACGTTGTAAAACGActagtggggtgccacttcat	FXYD6_Pr08_M13R	GGATAACAATTTCACACAGGaagccgtgagaggatggata
5' Upstream	FXYD6_Pr07_M13F	CACGACGTTGTAAAACGACtccccactaaagaatcctgct	FXYD6_Pr07_M13R	GGATAACAATTTCACACAGGgtgggtatgcatggacacaa
5' Upstream	FXYD6_Pr06_M13F	CACGACGTTGTAAAACGACaggaatggtttccacccttt	FXYD6_Pr06_M13R	GGATAACAATTTCACACAGGtggagaggaaagatgtggtc
5' Upstream	FXYD6_Pr05_M13F	CACGACGTTGTAAAACGACtcagggctctgaactgtgttt	FXYD6_Pr05_M13R	GGATAACAATTTCACACAGGtgcaagcctatttgggaaag
5' Upstream	FXYD6_Pr04_M13F	CACGACGTTGTAAAACGACgcattaggtaggggggacctg	FXYD6_Pr04_M13R	GGATAACAATTTCACACAggaaccagtaggggaaccat
5' Upstream	FXYD6_Pr03_M13F	CACGACGTTGTAAAACGAcagagtggcatccattcagc	FXYD6_Pr03_M13R	GGATAACAATTTCACACAGGtgggaggcttagctcatagtg
5' Upstream	FXYD6_Pr02_M13F	CACGACGTTGTAAAACGACgacagcctccatctcagcat	FXYD6_Pr02_M13R	GGATAACAATTTCACACAGGtaacccagcaccagacacag
5' Upstream & CpG	FXYD6_Pr01_M13F	CACGACGTTGTAAAACGACcacgacacagccattcat	FXYD6_Pr01_M13R	GGATAACAATTTCACACAGGtgaggcttgagctgctttc
CpG & Exon 1 (5' UTR)	FXYD6_Ex01_M13F	CACGACGTTGTAAAACGAcagccagcccgactccta	FXYD6_Ex01_M13R	GGATAACAATTTCACACAGGcacacccactgaacacacaa
Exon 2 (5' UTR)	FXYD6_Ex02_M13F	CACGACGTTGTAAAACGACttacccctgtttccaagctg	FXYD6_Ex02_M13R	GGATAACAATTTCACACAGGctggagaggctggatgttct
Exon 3 (5' UTR)	FXYD6_Ex03_M13F	CACGACGTTGTAAAACGACgctgctggtttatgggtgtt	FXYD6_Ex03_M13R	GGATAACAATTTCACACAGgcacccagggatagatttca
Exon 4 (5' UTR & CDS)	FXYD6_Ex04_M13F	CACGACGTTGTAAAACGACggcccctgactagacatgaa	FXYD6_Ex04_M13R	GGATAACAATTTCACACAGGtcatgagtgcagagcaaagc
Exon 5 & 6 (both CDS)	FXYD6_Ex05n06_M13F	CACGACGTTGTAAAACGACttcccctaacccccctaactg	FXYD6_Ex05n06_M13R	GGATAACAATTTCACACAGGctctcacactgtccccacct
Exon 7 (CDS)	FXYD6_Ex07_M13F	CACGACGTTGTAAAACGActttgccccagaagaagcta	FXYD6_Ex07_M13R	GGATAACAATTTCACACAGGtgggctgaccttcaatgttt
Exon 8 (CDS)	FXYD6_Ex08_M13F	CACGACGTTGTAAAACGACagcctgtttggtggagtgag	FXYD6_Ex08_M13R	GGATAACAATTTCACACAGGtgacccttcagtccagacct
Exon 9 (CDS & 3' UTR)	FXYD6_Ex09_M13F	CACGACGTTGTAAAACGAcagactcctgaaggctcctc	FXYD6_Ex09_M13R	GGATAACAATTTCACACAGGaccgagggactgagagaggt
Exon 10 (3' UTR)	FXYD6_Ex10a_M13F	CACGACGTTGTAAAACGACtcaaggaagggagagaagca	FXYD6_Ex10all_M13R	GGATAACAATTTCACACAggaaagcgagtccacagttc
Exon 10 (3' UTR)	FXYD6_Ex10bII_M13F	CACGACGTTGTAAAACGACttgtgtttgttagtgaactgtgga	FXYD6_Ex10b_M13R	GGATAACAATTTCACACAGGcagatggggactcctcgtg
Exon 10 (3' UTR)	FXYD6_Ex10c_M13F	CACGACGTTGTAAAACGACtatcccacctgatcccagtc	FXYD6_Ex10c_M13R	GGATAACAATTTCACACAGGcactgacaggcagagcaaag
Exon 10 (3' UTR)	FXYD6_Ex10d_M13F	CACGACGTTGTAAAACGACgaaggactcctgttccgttg	FXYD6_Ex10d_M13R	GGATAACAATTTCACACAGGccttcctcttccctctgacc

[a] CpG = CpG island; UTR = Untranslated region; CDS = Coding sequence.

[b] M13F/R tails are indicated in the uppercase of the primer sequences.

6.3 RESULTS FOR THE RESEQUENCING OF THE FXYD6 EXONS AND PROMOTER REGION

All ten exons and 5 kb of the promoter region were successfully sequenced in the FXYD6 gene. The mutations found through screening are listed in Table 6.3 including their location and the frequency of the mutation found in both the 32 schizophrenic and 32 control samples.

Table 6.3: Results of resequencing the FXYD6 gene including the location and frequency of all mutations found through screening.

Primer	Mutation	Location	Position (bp)[a]	Allelic Frequency[b]	
				Controls	Cases
FXYD6_Ex01	rs12790073 (G/A)	Intron 1	117252487	4/64	3/64
FXYD6_Ex05n06	rs473057 (C/T)	Intron 4	117218097	38/64	24/64
	rs17596129 (C/G)	Intron 4	117218059	5/64	2/64
FXYD6_Ex08	rs496164 (T/C)	Intron 7	117216442	38/62	27/64
FXYD6_Ex10a	rs516655 (C/T)	Intron 9	117214238	53/64	41/64
	ndSNP (G/A)	Exon 10 (3' UTR)	117214013	5/64	1/64
	Deletion (-TG)	Exon 10 (3' UTR)	117213967-8	14/64	17/64
	Insertion (+TG)	Exon 10 (3' UTR)	117213967-8	1/64	1/64
FXYD6_Ex10b	rs520333 (C/T)	Exon 10 (3' UTR)	117213835	30/64	22/64
	rs3087563 (A/G)	Exon 10 (3' UTR)	117213147	–	–
FXYD6_Pr01	ndSNP (G/A)	CpG Island	117252877	0/58	3/58
FXYD6_Pr03	rs11216598 (T/C)	5' Upstream	117253662	–	–
	rs631898 (A/G)	5' Upstream	117253688	–	–
FXYD6_Pr04	rs11605223 (G/C)	5' Upstream	117254054	–	–
	rs3809043 (C/G)	5' Upstream	117254325	–	–
	rs3809042 (G/C)	5' Upstream	117254330	–	–
FXYD6_Pr08	rs497768 (G/C)	5' Upstream	117255950	–	–

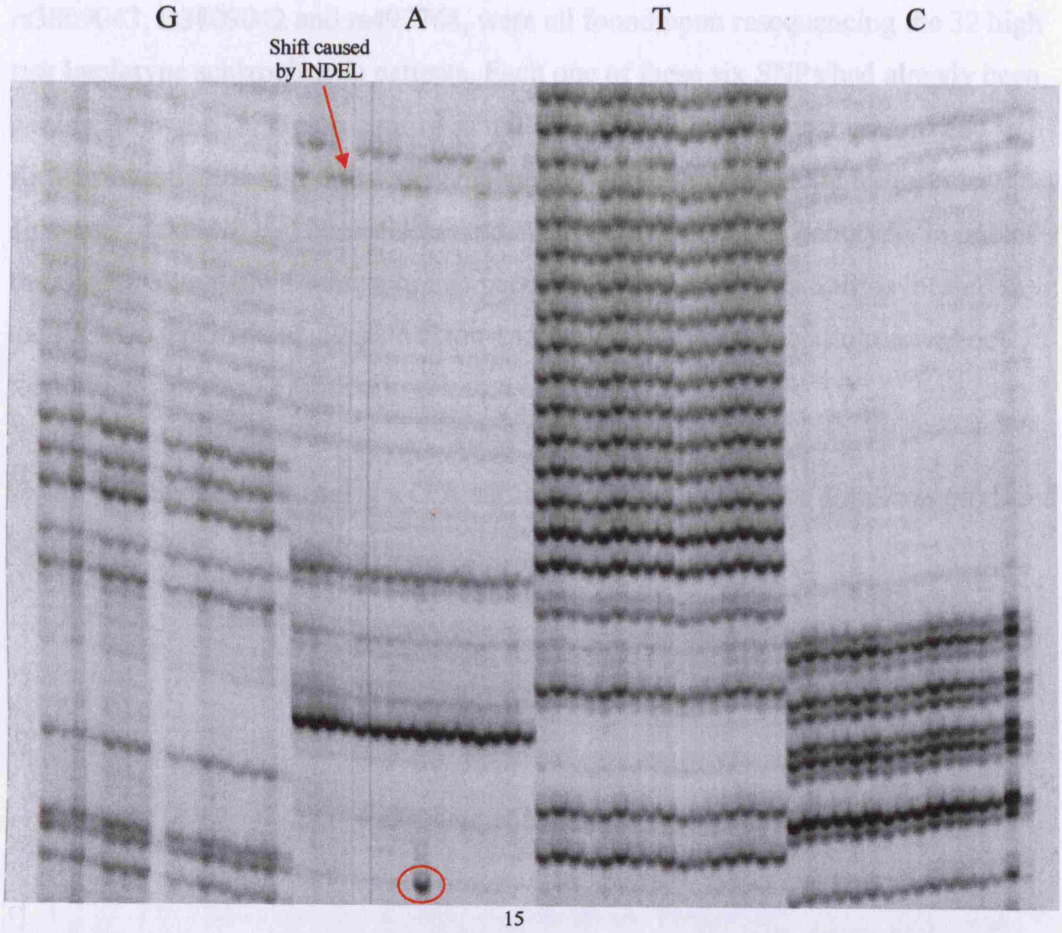
[a] Positions taken from UCSC Genome Browser database, March 2006 assembly.

[b] Frequencies for SNPs already genotyped in the UCL case-control sample are not listed and can be found in Table 4.3.

6.3.1 EXONIC / INTRONIC REGION

Seven database SNPs, rs12790073, rs473057, rs17596129, rs496164, rs516655, rs520333 and rs3087563, were all found in the exonic/intronic regions of the FXYD6 gene. The frequency of each SNP was greater in the controls than those observed in the 32 high risk haplotype schizophrenic patients. A non-database SNP (ndSNP) of G/A substitution and a TG insertion/deletion (INDEL) was also discovered. The ndSNP was found to appear in one schizophrenic patient (heterozygote GA genotype) and 3 control individuals (3 heterozygote GA and 1 homozygous AA genotype). The INDEL is part of a polymorphic sequence of TG. Both mutations are located in exon 10 (3'UTR) and are displayed in Figure 6.1.

Figure 6.1: Sequence showing the TG INDEL and a G/A substitution SNP within the amplicon produced from the FXYD6_Ex10a primers.

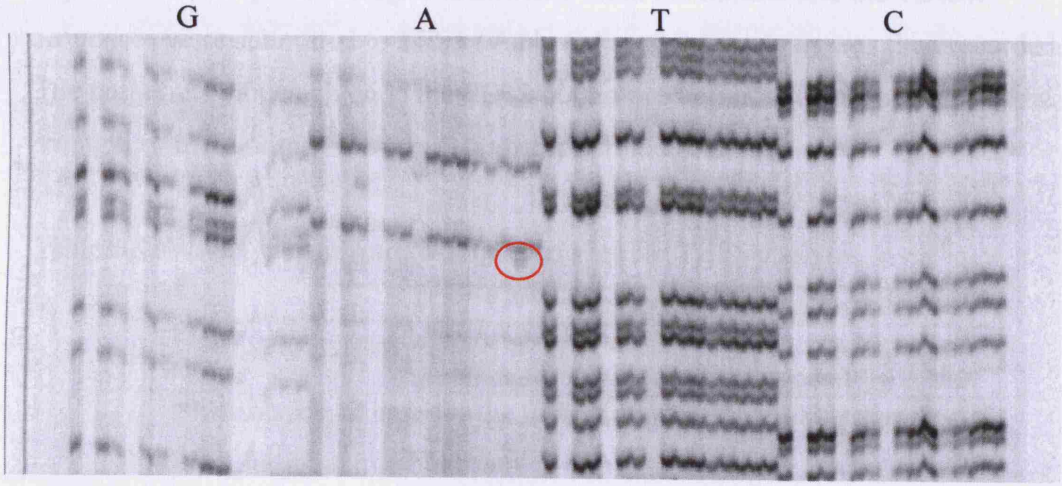


The repeated sequence of TG ends with a TG deletion in some of the samples. This caused the sequence to be shifted down by two bases just before the ACT nucleotide sequence (see arrow). One of the schizophrenic patients (sample 15) has a TG insertion in addition to the TG deletion (-TG,+TG heterozygote genotype). The INDEL (position 117213967-8, UCSC Genome Browser database, March 2006 assembly) was found to be in approximately equal number of cases and controls. A ndSNP of G/A substitution (position 117214013, UCSC Genome Browser database, March 2006 assembly), was discovered as a heterozygote genotype in one of the schizophrenic patient (highlighted circle). However, the frequency of this SNP was greater in the controls (5/64) than in the cases (1/64). Both mutations are located in exon 10 (3' UTR) of the FXYD6 gene.

6.3.2 PROMOTER REGION

In the FXYD6 promoter region, database SNPs rs11216598, rs631898, rs11605223, rs3809043, rs3809042 and rs497768, were all found upon resequencing the 32 high risk haplotype schizophrenic patients. Each one of these six SNPs had already been genotyped in the UCL case-control sample (Section 4.0) and was therefore not further investigated in the control samples for comparison of allelic frequencies. However, a ndSNP had been discovered, as a heterozygote GA genotype, in one of the chosen 32 screened schizophrenic patients (Figure 6.2). This SNP is located in the FXYD6 CpG Island, close to Exon 1 (5' UTR). The G/A substitution was not found in any of the 32 randomly selected controls.

Figure 6.2: Sequence showing a G/A substitution SNP within the amplicon produced from the FXYD6_Pr01 primers.



A rare SNP of G/A substitution (position 117252877, UCSC Genome Browser database, March 2006 assembly), was revealed as a heterozygous GA genotype in one of the 32 chosen schizophrenic patients (highlighted circle). This ndSNP was not found in any of the 32 randomly selected control individuals screened. The mutation is located as the 51st base in the CpG Island of the FXYD6 gene, close to the 1st exon (5' UTR).

6.4 FURTHER ANALYSIS OF THE FXYD6 CpG ISLAND G/A ndSNP

Following the screening results of the FXYD6 CpG Island ndSNP, the next stage involved looking at the effects of the mutation on the binding of transcription factors. The ndSNP also needed to be genotyped in the whole UCL case-control sample.

6.4.1 TRANSCRIPTION ELEMENT SEARCH SYSTEM (TESS)

The novel mutation identified by resequencing the FXYD6 promoter region was examined in an attempt to identify any alterations to possible transcription factor binding sites. 25 bases either side of the ndSNP were selected and analysed using TESS, a transcription factor binding site prediction tool obtained from the website: http://www.cbil.upenn.edu/cgi-bin/tess/tess. Both the normal and the variant sequences were submitted to TESS (Table 6.4) and the analysis was then recorded. The numerical values given in the TESS output predict, in a logarithmic scale, the strength of the binding for various transcription factors.

Table 6.4: Wild-type and variant sequence used for TESS analysis.

	Sequence[a]
Wild-type	ggcggcagtctcctgttgcttctcg**G**agcttgactgccctttggaaaaatc
Variant	ggcggcagtctcctgttgcttctcg**A**agcttgactgccctttggaaaaatc

[a] Substitution of base for the SNP is shown in capital letters and is highlighted in a bold font.

The output from TESS (not shown) indicated no significant alterations in the predicted binding efficiencies for any transcription factors. It was however decided to continue the genetic association test of the ndSNP in the UCL case-control sample.

6.4.2 A GENETIC ASSOCIATION STUDY OF THE FXYD6 CpG ISLAND G/A ndSNP

Genotyping of the rare ndSNP in the UCL case-control samples was done using EPOCH probes and a qPCR machine. The methods for this are described in Section 2.7.3 and the allelic association result is shown in Table 6.5.

Table 6.5: Allelic association test of the FXYD6 CpG Island ndSNP (G/A) with schizophrenia in the UCL sample.

Sample	G	A	χ^2	P[a]
Control	920	14	0.04	0.851
Case	919	13		

[a] Two-tailed significance (P) from 2 x 2 χ^2, with 1 df.

The χ^2 analysis of the genetic association test revealed no significant association to schizophrenia in the UCL case-control sample (P = 0.851).

6.5 DISCUSSION

Resequencing of the exons and the promoter region of the FXYD6 gene has led to the discovery of several mutations including ndSNPs. Only two of these mutations were found at a higher frequency in the 32 selected high risk haplotype schizophrenic patients when compared to the 32 randomly chosen control individuals (Table 6.3). One of these SNPs, rs497768, had already been genotyped in the whole UCL case-control sample and was found to be significantly associated to schizophrenia (P = 0.023). This marker is a common polymorphism and is thus unlikely to be a functional SNP or an aetiological base pair change that causes the genetic disorder. The other mutation with an increased frequency in the 32 schizophrenia samples is a ndSNP of G/A substitution (position 117252877, UCSC Genome Browser database, March 2006 assembly), located in the CpG Island of the FXYD6 gene. Bioinformatical screening with the TESS prediction program estimated no significant disruption by the variant to any possible transcription factor binding sites. A genetic association test of this SNP was then found to be negatively associated to schizophrenia after genotyping in the whole UCL case-control sample.

Because of time constraints, other non-database mutations found were not genotyped in the whole UCL case-control sample.

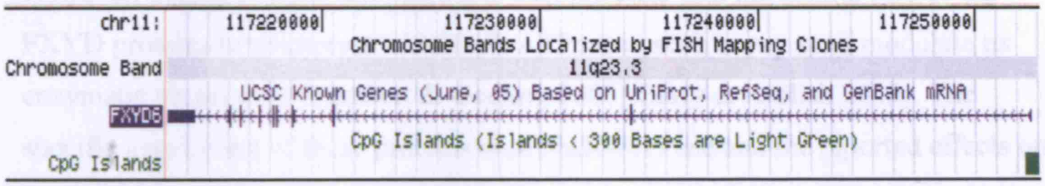
Further complexities that could have been explored in the search for aetiological mutations affecting the expression or function of FXYD6 include the methylation state of the promoter region, particularly the CpG Island. A larger area of the 5' and 3' region could also be sequenced due to possible distant mutations that may affect the gene, as previously discussed in Section 6.1.1. In addition, the introns should be fully sequenced because they are known to be important in gene regulation e.g. mutations in introns can cause incorrect splicing resulting in "exon skipping" or premature termination of the protein. It is interesting to note that four of the markers genotyped within intron 1, TTTC20.2, rs1815774, rs4938445 and rs4938446, have all shown significant association to schizophrenia in the UCL sample, with four other SNP markers, rs7121573, rs876797, rs873719 and rs10790218, all showing a trend toward association (Table 4.2 and 4.3). However, even though the majority of positively associated markers and those showing a trend toward association are placed in the same intron, they are still separated by over 15.7 kb. Additionally, markers in the opposite end of the gene have displayed positive association to schizophrenia. Therefore, it is possible that the site of the aetiological base pair change or mutation is located somewhere towards the 3' end of FXYD6 gene. Full length cDNA clone sequences containing the whole of FXYD2 with FXYD6 have been observed in the UCSC Genome Browser database, March 2006 assembly. Therefore, both genes need to be screened by sequencing.

7.0 THE FXYD DOMAIN – CONTAINING ION TRANSPORT REGULATOR 6 (FXYD6) GENE

7.1 BASIC GENE CHARACTERISTICS

The human FXYD (pronounced fix-id) Domain-Containing Ion Transport Regulator 6 (FXYD6) gene is located on chromosome 11q23.3, genomic position 117,212,903-117,252,577 (UCSC Genome Browser database, March 2006 assembly). The 40 kb gene consists of 10 exons and contains a CpG island situated in the 5' end overlapping exon 1 (Figure 7.1).

Figure 7.1: The location and features of the FXYD6 gene.



Transcription of the gene occurs on the negative strand to produce a 1.8 kb transcript. The sequence of four preceding and one nucleotide following the ATG start codon (cgccATGg) conforms to the Kozak consensus sequence for the initiation of translation (Kozak 1987). The start codon is located in exon 4 and the DNA transcription ends with the TGA (opal or umber) stop codon in exon 9. Therefore, the coding sequence spans 6 exons. The transcript is translated into a protein of 95 amino acid residues (Figure 7.2), with an estimated molecular weight of 10.5 kDa (Yamaguchi et al. 2001).

Figure 7.2: The protein sequence of the FXYD6 gene (*Homo sapiens*).

MELVLVFLCSLLAPMVLASAAEKEKEMDPFHYDYQTLRIGGLVFAVVLFSVGILLI
LSRRCKCSFNQKPRAPGDEEAQVENLITANATEPQKAEN

7.2 THE FXYD FAMILY OF GENES

FXYD6 encodes for the protein phosphohippolin and is part of a family which includes seven known mammalian FXYD genes (Yamaguchi et al. 2001; Kadowaki et al. 2004). Two other family members exist in *Danio rerio*, which are not obvious orthologs of any identified mammalian FXYD protein (Sweadner & Rael 2000). The complete FXYD gene family was established recently by expressed sequence tags database analysis, in a detailed genetic study that has provided complete cDNA sequences, protein signature sequences and expression patterns for the genes (Sweadner & Rael 2000). All FXYD proteins share homology for a single common transmembrane domain and are named after the invariant extracellular motif. Early studies indicate a variety of unrelated functions for these proteins (Table 7.1). However, during the last few years it is become clear that the central role of the FXYD proteins is to interact with Na,K-ATPase in such a way as to modulate its enzymatic properties (Crambert & Geering 2003; Garty & Karlish 2005). The specific association of these proteins with Na,K-ATPase and the reported effects on the enzyme function are listed in Table 7.1.

Table 7.1: Functional roles reported for each FXYD family member.

Gene	Protein	Function Reported in	
		Early Studies	Association with Na,K-ATPase
FXYD1	PLM (phospholemman)	Major membrane phosphoprotein in heart and muscle.[a] Evokes Cl^- conductance in *Xenopus laevis* oocytes.[b] Conducts other ions and osmolytes.[c]	Accessory protein of Na,K-ATPase in cerebellum and choroid plexus.[j]
FXYD2	γ (subunit of Na,K-ATPase)	Proteolipid that is copurified with renal Na,K-ATPase.[d]	Specific component of Na,K-ATPase in mammalian cells.[k] Modulates Na^+ and K^+ affinity of renal Na,K-ATPase.[l]
FXYD3	Mat-8 (mammary tumor protein of 8 kDa)	*neu*- and *ras*-induced gene.[e] Evokes Cl^- conductance in *Xenopus laevis* oocytes.[f]	Regulator of Na,K-ATPase.[m]
FXYD4	CHIF (corticosteroid hormone-induced factor)	Aldosterone-induced gene evokes K^+ conductance in *Xenopus laevis* oocytes.[g]	Regulator of Na,K-ATPase.[n]
FXYD5	RIC (related to ion channel) / Dysadherin	Cancer associated E2a-Pbx1 induce gene that downregulates E-cadherin.[h]	Modulates Na,K-ATPase properties.[o]
FXYD6	Php (phosphohippolin)	EST clones homologous to other FXYD proteins.[i]	Regulator of Na,K-ATPase expressed in the inner ear.[p]
FXYD7	FXYD7	EST clones homologous to other FXYD proteins.[i]	Brain-specific regulator of Na,K-ATPase α 1-β isozymes.[q]

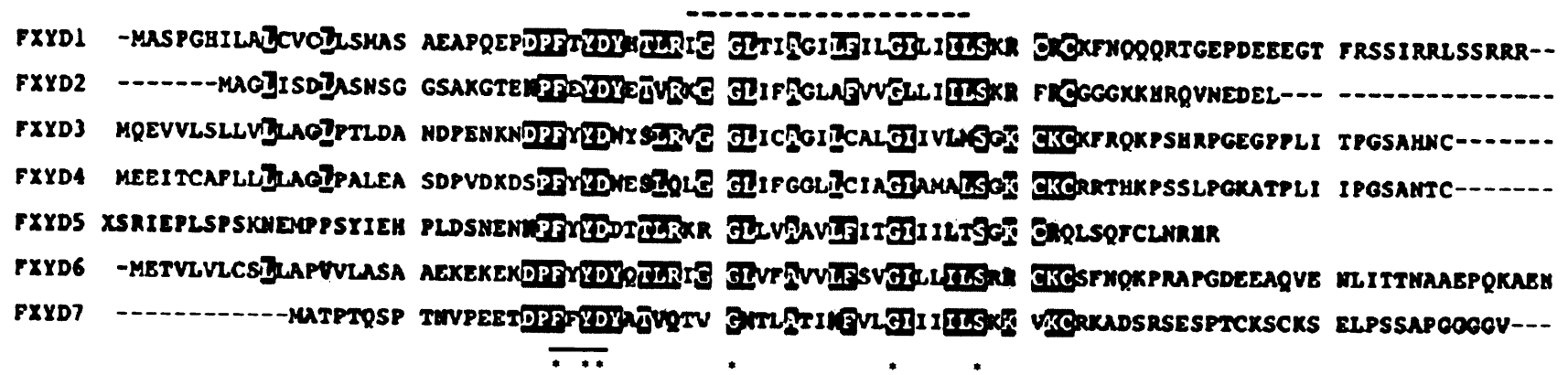
[a] Palmer et al. 1991; [b] Moorman et al. 1992; [c] Kowdley et al. 1997, Moorman & Jones 1998; [d] Forbush et al. 1978, Mercer et al. 1993; [e] Morrison & Leder 1994; [f] Morrison et al. 1995; [g] Attali et al. 1995; [h] Fu & Kamps 1997, Ino et al. 2002; [i] Sweadner & Rael 2000; [j] Crambert et al. 2002, Feschenko et al. 2003; [k] Beguin et al. 1997, Therien et al. 1999, Pu et al. 2001; [l] Therien et al. 1997, Arystarkhova et al. 1999; [m] Crambert et al. 2005; [n] Beguin et al. 2001, Garty et al. 2002, Lindzen et al. 2003; [o] Lubarski et al. 2005; [p] Delprat et al. 2007b; [q] Beguin et al. 2002.

A small acidic proteolipid (γ) belonging to what is now recognised as the FXYD family was first discovered in association with renal Na,K-ATPase nearly 30 years ago (Forbush et al. 1978). FXYD6 is the final mammalian member of the FXYD gene family for which a functional interaction with the Na,K-ATPase has been demonstrated (Delprat et al. 2007b). Due to the small amount of research on this novel gene, a greater understanding of FXYD6 can be obtained from existing data of the structural and functional interactions between Na,K-ATPase and other FXYD proteins. This includes discussing the putative physiological roles of FXYD proteins as well as possible interactions with other ion transport systems.

7.2.1 PRIMARY STRUCTURE, SEQUENCE HOMOLOGY AND MEMBRANE TOPOLOGY OF FXYD PROTEINS

FXYD proteins are short polypeptides (66-178 amino acids) with a single transmembrane segment. Homology among family members is observed in a stretch of 35 amino acids in and around the transmembrane domain, but is not observed outside this region (Figure 7.3). Only six residues are fully conserved in all family members and species. These are the extracellular phenylalanine, tyrosine and aspartic acid residues that define the FXYD motif, two transmembrane glycines and a serine residue at the membrane-cytoplasm interface. It is likely that these residues are involved in an essential function common to all FXYD proteins e.g. structural interaction with the Na,K-ATPase. The requirement of the FXYD motif has already been shown to be important for structural interaction in both FXYD2 (γ) and FXYD4 (CHIF) proteins (Beguin et al. 2001), but not for FXYD7 (Crambert et al. 2004). In addition, the two invariant glycines within the transmembrane span signify that they are helix packing sites (Sweadner & Rael 2000).

Figure 7.3: Multiple sequence alignment of the seven mammalian FXYD family proteins (*Mus musculus*).

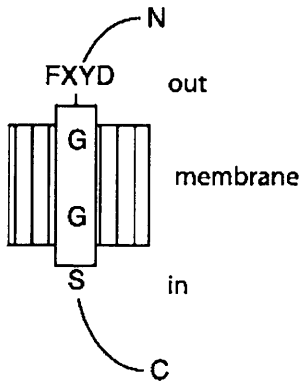


Each FXYD family member is numbered according to the order in which they were initially sequenced and share a common 35 amino acid signature sequence encompassing a single transmembrane span and the adjacent regions (Sweadner & Rael 2000). The transmembrane domain is specified with a dotted bar directly above the sequences and the six invariant amino acid residues are marked by asterisks below the sequences. Conserved amino acid residues amongst the mammalian FXYD family members are indicated as shaded residues. The signature motif in which the FXYD family of genes are named after is underlined below the FXYD sequence, where X is an unspecified variant amino acid residue (usually Y but can also be E, T or F). The FXYD motif is preceded by proline residue (PFXYD) in all FXYD genes and species (Sweadner & Rael 2000), with the exception of lysine (LFXYD) in *Xenopus laevis* FXYD2 (γ) proteins (Beguin et al. 1997; Crambert & Geering 2003). The basic residues following the invariant serine residue and then cysteines at one or both of two positions with another basic residue in between, forms a typical membrane protein stop-transfer signal that is important for biosynthesis in the rough endoplasmic reticulum (Sweadner & Rael 2000). Figure modified from Garty and Karlish (2006).

The genomic organisation of the seven mammalian FXYD members is unusual in that each transcript is made up of at least six exons and that the core 35 amino acid conserved region is formed by three separate exons (Sweadner et al. 2003). The large number of exons in the genes of the relatively small FXYD proteins may reflect modular gene assembly (Sweadner & Rael 2000; Sweadner et al. 2000). Evidence from intron-exon junctions suggest that the proteins in this family were assembled from discrete structural modules via exon shuffling. The multiple exon organisation of the FXYD genes could serve to enable high levels of structural and functional diversity among the family members.

Analysis of the FXYD membrane topology established that they are hydrophobic type I single-span membrane proteins, with an acidic extracellular NH_2 terminus and a basic cytoplasmic COOH terminus (Figure 7.4). NH_2 terminal signal peptide cleavage has been observed in mature protein, except for γ and FXYD7 (Palmer et al. 1991; Beguin et al. 1997; Therien et al. 1997; Beguin et al. 2001; Beguin et al. 2002; Cornelius & Mahmmoud 2003; Lubarski et al. 2005). Helical wheel projection of the transmembrane segment demonstrates that most of the residues conserved among the FXYD proteins are preferentially located on one face of the helix (Sweadner & Rael 2000).

Figure 7.4: Membrane topology of FXYD proteins.



The positions of conserved amino acid residues (Figure 7.3), including the FXYD motif, are indicated. The type I single-span membrane protein has an extracellular NH_2 terminus, containing the FXYD family signature, and a cytoplasmic COOH terminus. Figure taken from Geering (2005).

7.2.2 TISSUE DISTRIBUTION OF FXYD PROTEINS IN MAMMALS

FXYD proteins are widely distributed with prominent expression in tissues that perform fluid and solute transport or that are electrically excitable (Sweadner & Rael 2000). Most FXYD family members exhibit a high degree of tissue specificity e.g. γ is detected only in the kidney (Pu et al. 2001; Wetzel & Sweadner 2001), CHIF is detected only in kidney and colon (Shi et al. 2001; Garty et al. 2002), PLM is detected mainly in heart and skeletal muscles (Feschenko et al. 2003; Wetzel & Sweadner 2003; Zhang et al. 2003), whereas FXYD7 is brain specific (Beguin et al. 2002). There is limited data on cell and tissue specificity of the other FXYD proteins. The available information from literature on organ/tissue distribution of FXYD mRNA and protein is summarised in Table 7.2.

Table 7.2: The expression of FXYD mRNA and proteins in major organs and tissues.

Gene	Protein	Reported Expression of	
		mRNA	Protein
FXYD1	PLM (phospholemman)	Heart, liver and skeletal muscle.[a]	Heart (myocytes), brain (cerebellum and choroids plexus) and kidney (extraglomerular mesangium).[b]
FXYD2	γ (subunit of Na,K-ATPase)	Kidney, heart and stomach.[c]	Kidney (thick ascending limb of the loop of Henle, distal convoluted tubule, inner medullary collecting duct and macula densa).[d]
FXYD3	Mat-8 (mammary tumor protein of 8 kDa)	Colon, stomach and uterus.[e]	Stomach (mucous cells).[f]
FXYD4	CHIF (corticosteroid hormone-induced factor)	Kidney collecting duct and distal colon.[g]	Kidney collecting duct and distal colon.[h]
FXYD5	RIC (related to ion channel) / Dysadherin	Heart, brain, spleen, lung, liver, skeletal muscle, kidney and testis.[i]	Kidney, intestine, lung, heart and spleen.[j]
FXYD6	Php (phosphohippolin)	Brain and kidney.[k]	CNS and inner ear.[l]
FXYD7	FXYD7	Brain (cerebellum, cerebrum, hippocampus and stem).[m]	Brain (cerebellum, cerebrum, hippocampus and stem).[m]

[a] Morrison et al. 1995, Bogaev et al. 2001; [b] Feschenko et al. 2003, Wetzel & Sweadner 2003, Zhang et al. 2003; [c] Beguin et al. 1997; [d] Pu et al. 2001, Wetzel & Sweadner 2001, Pihakaski-Maunsbach et al. 2003, Pihakaski-Maunsbach et al. 2005; [e] Morrison et al. 1995; [f] Crambert et al. 2005; [g] Attali et al. 1995, Capurro et al. 1996; [h] Shi et al. 2001, Garty et al. 2002, Pihakaski-Maunsbach et al. 2003; [i] Fu & Kamps 1997; [j] (Morrison et al. 1995; [k] Yamaguchi et al. 2001; [l] Kadowaki et al. 2004, Delprat et al. 2007b; [m] Beguin et al. 2002.

Since FXYD proteins have different tissue distribution and functional effects, it is hypothesised that each FXYD protein is expressed in a tissue-specific manner and functions by modulating Na,K-ATPase (Crambert & Geering 2003). This is by altering the kinetic behaviour of the enzyme to the specific requirements of a given tissue, cell type or physiological state (Crambert & Geering 2003; Garty & Karlish 2005). Therefore, FXYD proteins behave as regulatory subunits that adjust the ion pump properties of Na,K-ATPase in response to external signals.

7.2.3 FUNCTIONAL INTERACTIONS OF FXYD PROTEINS AND Na,K-ATPase

Various expression systems have shown that the association of FXYD proteins is not integral for Na,K-ATPase activity (Therien & Blostein 2000). However, there is compelling evidence for specific association with the α/β complex of Na,K-ATPase as demonstrated by co-immunoprecipitation. Significant functional affects of Na,K-ATPase by all seven FXYD proteins have been published (Therien et al. 2001; Crambert & Geering 2003; Cornelius et al. 2005; Crambert et al. 2005; Garty & Karlish 2005; Delprat et al. 2007b). The adaptation in the kinetic properties of Na,K-ATPase (apparent Na^+ and/or K^+ affinity) involves changes in the active Na^+ and K^+ transport that is based on the particular needs of different cells (Garty & Karlish 2006). The modulatory fold-effects of FXYD proteins on kinetic parameters such as $K_{1/2}$ Na^+ (cytoplasmic/intracellular Na^+), $K_{1/2}$ K^+ (extracellular K^+), V_{max} (maximal rate of Na,K-ATPase) and K'_{ATP} are usually approximately two-fold or less (Therien et al. 1997; Arystarkhova et al. 1999; Therien et al. 1999; Beguin et al. 2001; Pu et al. 2001; Beguin et al. 2002; Crambert et al. 2002; Garty et al. 2002; Feschenko et al. 2003; Mahmmoud et al. 2003; Zouzoulas et al. 2003; Fuller et al. 2004; Crambert et al. 2005; Jia et al. 2005; Jones et al. 2005; Silverman et al. 2005; Zouzoulas et al. 2005). However, these moderate effects are likely to have long-term physiological importance in maintaining cation homeostasis (Garty & Karlish 2006).

Although the functional effects of FXYD proteins on Na,K-ATPase activity is well documented, there is a lack of information detailing the molecular mechanism

involved. This would have provided better understanding on the functional role of FXYD proteins as well as complementing studies on the relationship between structure and function. It is thought that the effect of FXYD proteins on Na,K-ATPase activity causes alterations in the rate limiting steps i.e. binding of the transported cations, particularly cytoplasmic Na^+ ions (or competing K^+ ions). This limits the rate of active Na^+ pumping *in vivo*. However, direct observations of purified FXYD/Na,K-ATPase complexes (α/β/FXYD) expressed in *P. pastoris* have yet to be conclusive (Cohen et al. 2005; Lifshitz et al. 2006).

7.2.4 STRUCTURAL INTERACTIONS OF FXYD PROTEINS AND Na,K-ATPase

The functional effect of FXYD proteins is accounted for by its interaction with the α/β complex of Na,K-ATPase. Evidence suggest that there are multiple sites of interaction between FXYD and the α/β subunits (Therien et al. 1997; Therien et al. 1999; Beguin et al. 2001; Pu et al. 2001). Helical wheel projection reveal that the three CHIF residues identified as important for the stability of the CHIF-α/β complex (Gly-45, Met-55 and Ala-56) face two different planes of the membrane (Lindzen et al. 2003). Therefore, they are likely to interact with at least two different helixes on the α and/or β subunit of Na,K-ATPase. Peptides corresponding to the transmembrane segment of γ have been shown to reduce Na^+ affinity of the α/β complex in HeLa cell membranes (Zouzoulas et al. 2003). This confirms the functional role of the transmembrane segments.

To determine how many FXYD proteins can interact with each α/β complex, information on possible stoichiometries of α/β to γ subunits in renal Na,K-ATPase, and HeLa cells expressing γ, was obtained using co-immunoprecipitation (Mercer et al. 1993; Kuster et al. 2000). The experiment was done in conditions known to stabilise the protein and to optimise co-immunoprecipitation efficiency. The purified renal Na,K-ATPase was shown to contain α:β:γ (both γa and γb splice variants) at a ratio close to 1:1:1. This suggests that there is an even combination of α/β/γa and

α/β/γb complexes. However, the data is also compatible with an equal proportion of mixed α/β/γa/γb complexes and α/β complexes without γ.

As there is little known structural data on the Na,K-ATPase enzyme itself, direct structural information for the Na,K-ATPase-FXYD interaction is consequently small. The molecular structure of Na,K-ATPase with FXYD proteins may eventually become elucidated using FXYD protein structures that are currently being determined through nuclear magnetic resonance (NMR) spectroscopy (Marassi & Crowell 2003).

7.2.5 POSSIBLE PHYSIOLOGICAL ROLES OF FXYD PROTEINS

The physiological roles of FXYD proteins can be proposed on the basis of their observed effects *in vitro*, from phenotypic analysis of knockout (KO) mice models and in genetic disorders associated by mutations found within FXYD genes. Initial analyses of γ and CHIF KO mice models have been reported (Aizman et al. 2002; Goldschimdt et al. 2004; Jones et al. 2005). A single nucleotide mutation in human γ has also been linked to familial dominant renal hypomagnesemia (Meij et al. 2000).

The opposite functional effects of γ and CHIF in renal Na,K-ATPase pump activity matches their dissimilar patterns of expression along the nephron (Table 7.2). This may relate to the different physiological needs of the cells in which they are being expressed in (Geering 2006). The thick ascending limb of the loop of Henle contains prominent levels of γ and active Na,K-ATPase (α/β/γ complexes). It is characterised by a very high rate of Na^+ pumping through transepithelial Na^+ reabsorption. This generates the renal salt gradients. The effect of γ in lowering cellular Na^+ affinity, deduced from γ KO mice experiments and transfected cells, implies a major homeostatic function in the physiological role of the γ subunit (Arystarkhova et al. 1999; Beguin et al. 2001; Pu et al. 2001; Jones et al. 2005). In contrast to γ, CHIF is expressed in the kidney's cortical collecting duct (CCD) and is induced by corticosteroid plus K^+ loading (Attali et al. 1995; Capurro et al. 1996; Wald et al. 1996; Wald et al. 1997; Shi et al. 2001; Pihakaski-Maunsbach et al. 2003).

Therefore, it functions in the crucial role of K^+ homeostasis in CCD cells through hormonal regulation (Palmer 1999; Feraille & Doucet 2001; Giebisch 2001). CHIF increases Na^+ affinity to adapt the pumping rate to a lower intracellular Na^+ level. This is by responding effectively to increases in cytosolic Na^+ associated with corticoid-induced Na^+ permeability. In conclusion, γ and CHIF modulate opposite Na,K-ATPase apparent affinity for Na^+ based on their physiological relevance and segment specific localisation (Cornelius & Mahmmoud 2003; Geering 2005). The adaptation reflects on the functional role of Na,K-ATPase in Na^+ reabsorption for different parts of the renal tubule.

The single nucleotide mutation in human γ linked to familial dominant renal hypomagnesemia causes a G41R replacement (Gly-41 with Arg) in the transmembrane segment (Meij et al. 2000). This abolishes the interaction of γ with Na,K-ATPase. As a result, the γ subunit fails to be incorporated into the plasma membrane and instead accumulates in the Golgi apparatus. The routing of Na,K-ATPase α/β subunits are not affected throughout this process. It is thought that the absence of γ may indirectly lead to a loss of Mg^{2+} in the distal convoluted tubule (DCT) i.e. reduced Na,K-ATPase activity might elevate cell Na^+ and thus trigger off secondary effects of Mg^{2+} entry reduction into the DCT (Meij et al. 2000). The DCT plays an important role in determining the final urinary excretion of Mg^{2+} even though it reabsorbs only approximately 10% of the filtered Mg^{2+} (Dai et al. 2001).

Further observations suggest the possibility of additional γ and CHIF functions through unknown mechanisms. The two γ splice variants, γa and γb, structurally vary in their most NH_2-terminal amino acids (Kuster et al. 2000; Sweadner et al. 2000). They have no noticeable functional differences (Beguin et al. 2001; Pu et al. 2001), yet some of their tissue distributions are dissimilar (Pihakaski-Maunsbach et al. 2003). This implies that both variant γ proteins may have separate roles depending on the specific cellular location where they are expressed. γ and CHIF proteins have also been recently detected within the same cells of the inner medullary collecting duct (Pihakaski-Maunsbach et al. 2003). Since γ and CHIF

have opposite effects on Na^+ affinity and are not found in mixed α/β/γ/CHIF complexes, it is likely that a regulatory mechanism exists by which some post-translational modification determines the interaction of α with either γ or CHIF (Garty et al. 2002). Finally, the possibility of FXYD proteins modulating ion transport proteins other than Na,K-ATPase can not be excluded (Section 7.2.6).

7.2.6 REGULATION OF ION TRANSPORT PROTEINS OTHER THAN Na,K-ATPase BY FXYD PROTEINS

Initial publications of most FXYD proteins reported findings that were unrelated to Na,K-ATPase regulation (Table 7.1). Thus, PLM and Mat-8 were shown to evoke Cl^- channel activity in *Xenopus laevis* oocytes (Moorman et al. 1992; Morrison et al. 1995). Other studies suggested PLM mediated permeation of various other ions as well as large molecules (Moorman et al. 1995; Kowdley et al. 1997; Moorman & Jones 1998; Chen et al. 1999; Morales-Mulia et al. 2000; Moran et al. 2001; Davis et al. 2004). γ and CHIF were also described as manifesting ionic channel activity in expression systems (Attali et al. 1995; Minor et al. 1998; Sha et al. 2001; Jespersen et al. 2006). γ evokes a large non-specific conductance that allows permeation of the plant carbohydrate inulin, whereas CHIF-induced conductance is K^+-specific and may reflect activation of KCNQ1 ion channels. RIC is involved in downregulating E-cadherin and also promotes metastasis (Ino et al. 2002). Finally, recent findings suggest PLM directly regulates the Na^+/Ca^{2+} exchanger NCX1 in cardiac myocytes (Song et al. 2002; Zhang et al. 2003; Mirza et al. 2004; Ahlers et al. 2005; Song et al. 2005). These observations raise the possibility that FXYD proteins perform other cellular functions in addition to altering pump kinetics and may modulate ion transport proteins other than Na,K-ATPase. However, the physiological relevance of these putative functions remains unclear.

7.2.7 CONCLUSIONS

From data obtained from other FXYD family members, it can be concluded that FXYD6 is expressed in a tissue-specific fashion and functions by modulating the ion pump properties of Na,K-ATPase (Section 7.2.2). This involves changes in the

active Na^+ and K^+ transports that are based on the specific requirements of the different cells in which they are being expressed. Its interaction with Na,K-ATPase causes alteration in the enzyme's apparent affinities for Na^+, K^+, ATP and V_{max} (Section 7.2.3). These protein fold-effects on kinetic parameters such as $K_{1/2}$ Na^+, $K_{1/2}$ K^+, V_{max} and K'_{ATP} are usually around two-fold. Although modest, the kinetic outcomes are likely to have profound long-term physiological importance in maintaining cation homeostasis.

Mechanistic details for the functional effect of phosphohippolin are lacking. However, based on the evidence from other FXYD proteins, it is thought that FXYD6 proteins structurally interact with the α/β complex of Na,K-ATPase (Section 7.2.4). This may cause changes in the rate limiting steps of Na,K-ATPase activity, which could subsequently alter the rate of active Na^+ pumping *in vivo*. Helical wheel projection of CHIF residues reveal that the FXYD/Na,K-ATPase interaction is likely to occur with at least two different helixes on the α and/or β Na,K-ATPase subunit. This involves some of the six amino acid residues that are conserved in all FXYD family members and species (Section 7.2.1). Therefore, the FXYD motif and the transmembrane segment of phosphohippolin are essential for structural interaction with Na,K-ATPase.

Phenotypic analyses of γ and CHIF KO mice models provide important information on the possible physiological role of phosphohippolin (Section 7.2.5). The distinct and opposite functional effects of γ and CHIF on the Na,K-pump are consistent with their different patterns of expression along the nephron. This suggests that the alteration of cellular Na^+ affinity by phosphohippolin is tissue-specific because it modifies Na,K-ATPase transport properties based on the physiological condition of the tissue. Splice variants of FXYD6 may also be expressed in dissimilar tissue distributions (Section 7.2.5). Therefore, Na,K-ATPase activity may be additionally regulated in an isoform-specific manner where alternative splice variants of phosphohippolin have different functional effects.

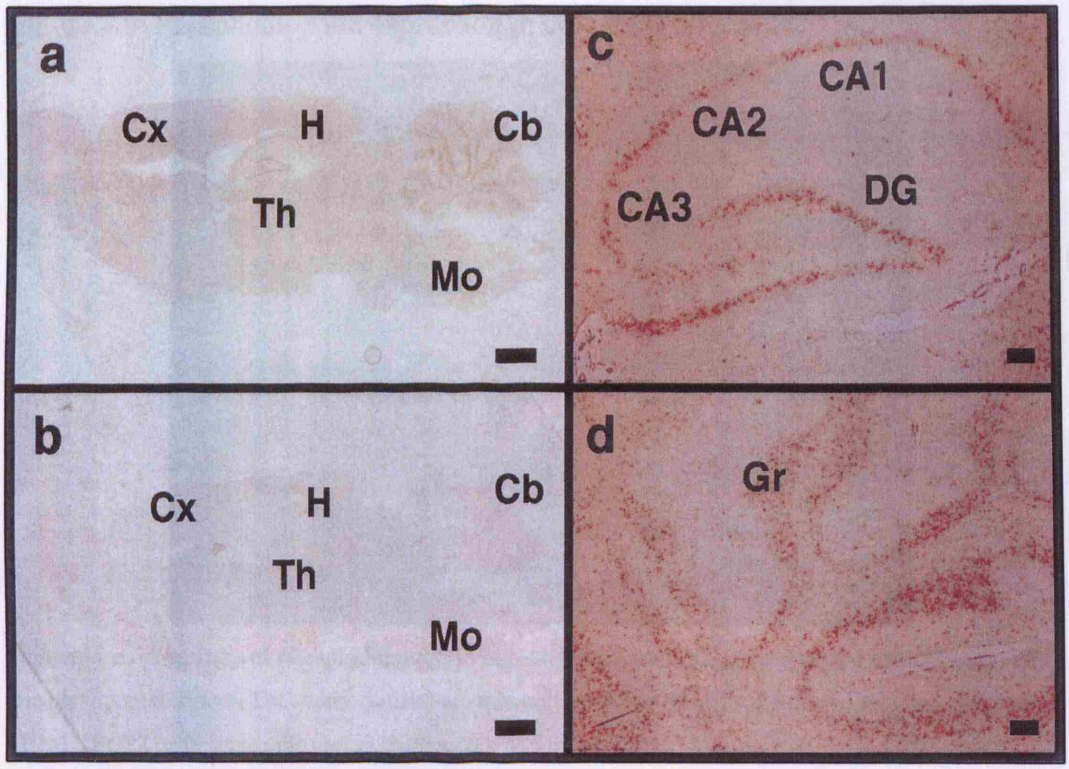
Finally, the discovery of a single nucleotide mutation in human γ, which causes dominant renal hypomagnesemia, provides an insight for a potential mechanism where mutations in FXYD6 can cause diseases (Section 7.2.5). The replacement of the conserved glycine residue into an arginine residue (G41R) indicates that the dysregulation of Na,K-ATPase by FXYD proteins can be implicated in pathophysiological states. We should also not exclude the possibility of phosphohippolin controlling activity of other transporters in addition to modulating the Na,K-ATPase (Section 7.2.6). This is based on the accumulating data of FXYD protein functions other then their involvement in Na,K-ATPase regulation. However, no evidence of this has yet to be reported for FXYD6.

7.3 EXPRESSION OF FXYD6 IN MAMMALIAN TISSUE

Yamaguchi et al. (2001) identified and cloned FXYD6 from the cDNA library of the rat hippocampus. The deduced sequenced comprised of 93 amino acid residues with a core FXYD motif and a single transmembrane domain. Tissue distribution of phosphohippolin mRNA was studied by Northern blot analysis (Yamaguchi et al. 2001). Abundant expression of the 1.5 kb phosphohippolin transcript was observed in the brain and kidney, whilst also being detected in lower intensities of various other tissues i.e. heart, testis, liver and lung.

The localisation of phosphohippolin mRNA in rat brain was analysed by *in situ* hybridisation (Yamaguchi et al. 2001). Although the scattered expression of phosphohippolin was noticed in the entire brain (Figure 7.5a), strong expression was seen in the hippocampus, including the dentate gyrus (Figure 7.5c) and the CA1-3 layers (Figure 7.5c), plus the granular cell layer of cerebellum (Figure 7.5d). The sense probe detected no signal (Figure 7.5b), indicating that the positive signal was specific to the phosphohippolin mRNA. Magee (1998) reported that the dendritic hyperpolarisation-activated currents modify the integrative properties of hippocampal CA1 pyramidal neurons. The high levels of phosphohippolin expressed in the hippocampus thus suggest that it may be involved in the integration of synaptic inputs (Section 7.5).

Figure 7.5: Localisation of phosphohippolin mRNA in the brain by *in situ* hybridisation (*Rattus norvegicus*).

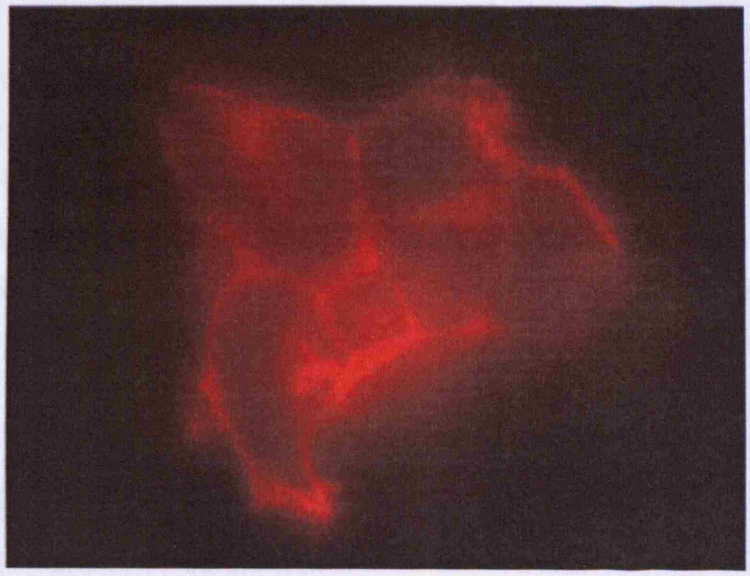


Sagittal sections of the adult rat brain were hybridised with a digoxigenin labelled RNA probe. The signal was visualised with 5-bromo-4-chloro-3-indolyl-phosphate and nitro blue tetrazolium chloride. Scale bar: 1 mm for a/b and 100 µm for c/d. (a) Strong expression of phosphohippolin mRNA in the hippocampus (H) and cerebellum (Cb) was detected [Key for other brain regions: cortex (Cx), thalamus (Th) and medulla pons (MO)]. (b) No signal was observed with the sense probe (control). (c) Magnification of the hippocampus region: CA1-3 and dentate gyrus (DG) were strongly stained. (d) Magnification of the cerebellum: The granular cell layer (Gr) was strongly stained. Figure taken from Yamaguchi et al. (2001).

In the follow-up study, Kadowaki et al. (2004) analysed the tissue distribution and developmental expression of the FXYD6 protein by raising and purifying the polyclonal antibodies against the carboxyl-terminal sequence of phosphohippolin (c-Php). Using confocal microscopy, fluorescence immunostaining in non-differentiated pheochromocytoma (PC12) cells revealed phosphohippolin

distribution in neuronal cells to be localised predominately with the plasma membrane (Figure 7.6).

Figure 7.6: Phosphohippolin expression in cultured PC12 cells.



Immunocytochemistry of phosphohippolin in non-differentiated PC12 cells reveal an affiliation with the plasma membrane. The methods used to stain cells were as described in Racoosin and Swanson (1994). PC12 cells were cultured in Dulbecco's Modified Eagle Medium with 10% fetal bovine serum. They were fixed with 4% paraformaldehyde for 30 min, permeabilised with 1% saponin and blocked with 0.5% bovine serum albumin plus 1% goat serum for 20 min. The cells were then incubated with primary c-Php antibody in phosphate buffered saline for 1 h at room temperature. Following wash steps, the cells were incubated with fluorescent isothiocyanate-conjugated anti-rabbit IgG for 45 min at room temperature. The glass slides were viewed using laser scanning confocal microscope. Figure taken from Kadowaki et al. (2004).

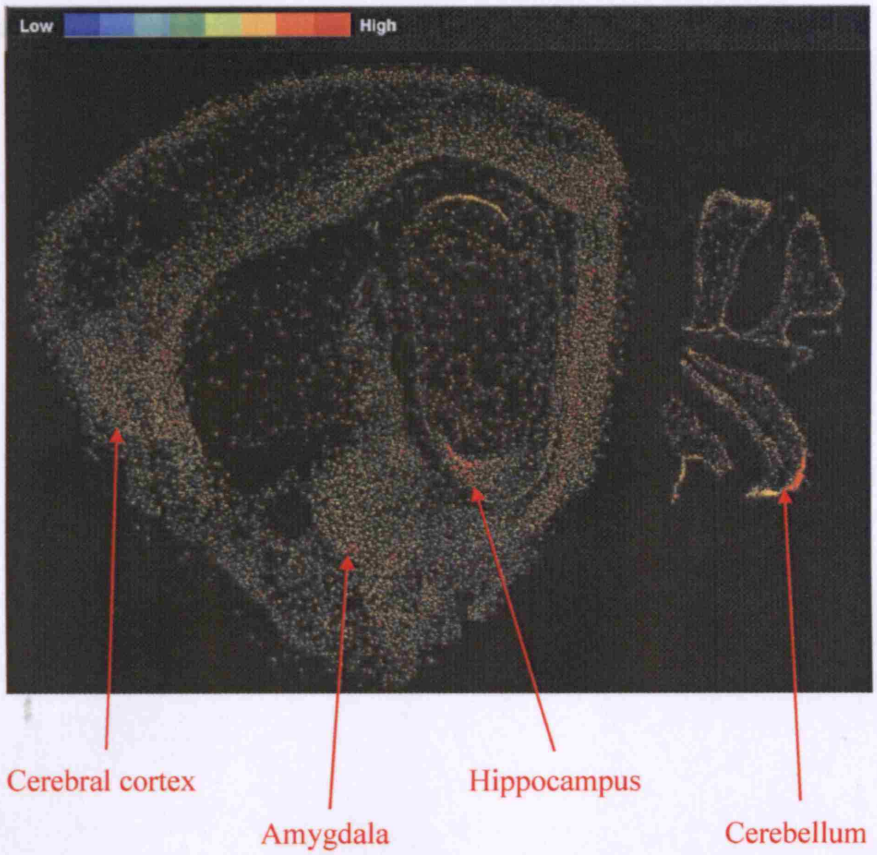
For distribution in various tissues of postnatal 10 week old rats, Western blot analysis disclosed abundant phosphohippolin expression in the rat brain with weak (lung, colon and testis) or no (liver, kidney, pancreas, small intestine, muscle and thymus) expression observed in other tissues (Kadowaki et al. 2004). In addition, the expression of phosphohippolin in four regions of the rat brain (forebrain, hippocampus, cerebellum and brain stem) significantly changed during postnatal development. The greatest amount of phosphohippolin was recorded in postnatal 3 week old rat brain with substantial capacity of phosphohippolin still existing in the

forebrain, hippocampus and cerebellum of postnatal 10 week old rats (Kadowaki et al. 2004). These findings suggest that phosphohippolin may play an important role in neuronal excitability of the central nervous system during postnatal development and in the adult brain.

The immunohistochemistry results displayed spatial expression of phosphohippolin within the neuronal fibres of the lateral habenula nucleus, hypothalamus, amygdaloid body, hippocampus, thalamus, stria terminalis, cingulum, olfactory bulb, cerebral cortex and cerebellum (Kadowaki et al. 2004). The distribution of phosphohippolin was identified as unique in the cerebellum, with a predominant expression pattern in the granule layer of the posterior lobe.

Additional gene expression data obtained from *in situ* hybridisation of mouse brain, as observed in the Allen Brain Atlas Database (Lein et al. 2006), shows FXYD6 to be present in high concentrations of the amygdala, hippocampus, cerebral cortex and cerebellum (Figure 7.7). These results match those found by Kadowaki et al. (2004). The highest level of expression in the hippocampus was detected in the pyramidal layer of the CA1 field, as also revealed by Yamaguchi et al. (2001).

Figure 7.7: mRNA expression of FXYD6 in the brain (*Mus musculus*).

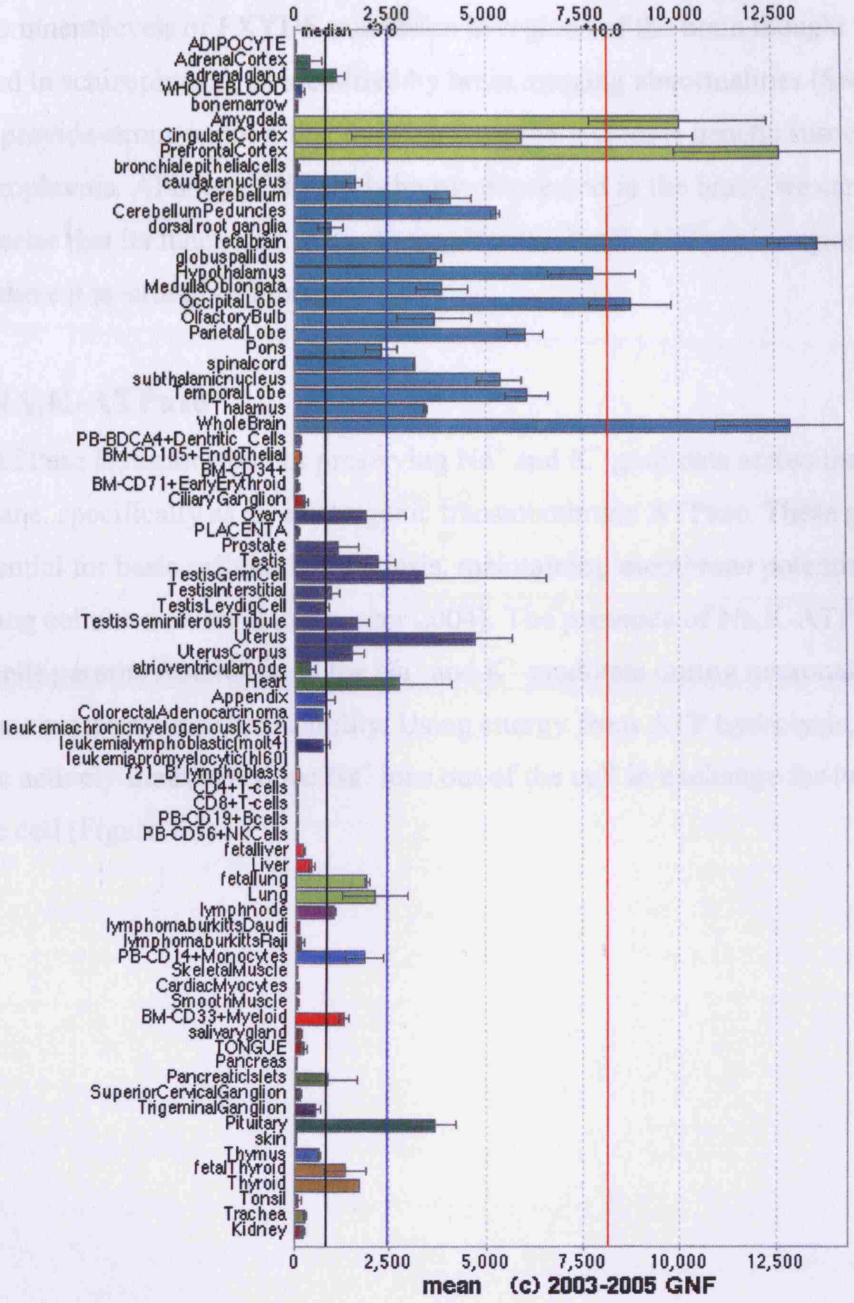


In situ hybridisation (ISH) of FXYD6 mRNA using digoxigenin labelled riboprobes. Level of gene expression is indicated by the intensity of coloured staining. Controls were performed in order to establish reproducibility and to demonstrate both qualitative and quantitative concordance to conventional radioactive ISH methodologies. The ISH Image Pane is a sagittal section of the mouse brain (pane Fxyd6_16 of Fxyd6-sagittal-b05-0107 image set) displaying gene expression of FXYD6. Figure taken from the Allen Brain Atlas (Lein et al. 2006), a genomewide image database of gene expression in the mouse brain.

Finally, the Novartis Gene Expression Atlas Database (GNF SymAtlas) shows that the mRNA expression of human FXYD6 is primarily in the brain (Su et al. 2002), as reported in rat samples by both Yamaguchi et al. (2001) and Kadowaki et al. (2004). The highest level of expression was found in the prefrontal cortex, amygdala, occipital lobe, hypothalamus and temporal lobe regions of the brain (Figure 7.8). The relatively huge capacity of mRNA expression in the fetal brain correlates with

the fact that all FXYD genes are expressed early in fetal life and display a great amount of tissue specificity (Sweadner & Rael 2000).

Figure 7.8: mRNA expression of FXYD6 in various tissues (*Homo sapiens*).



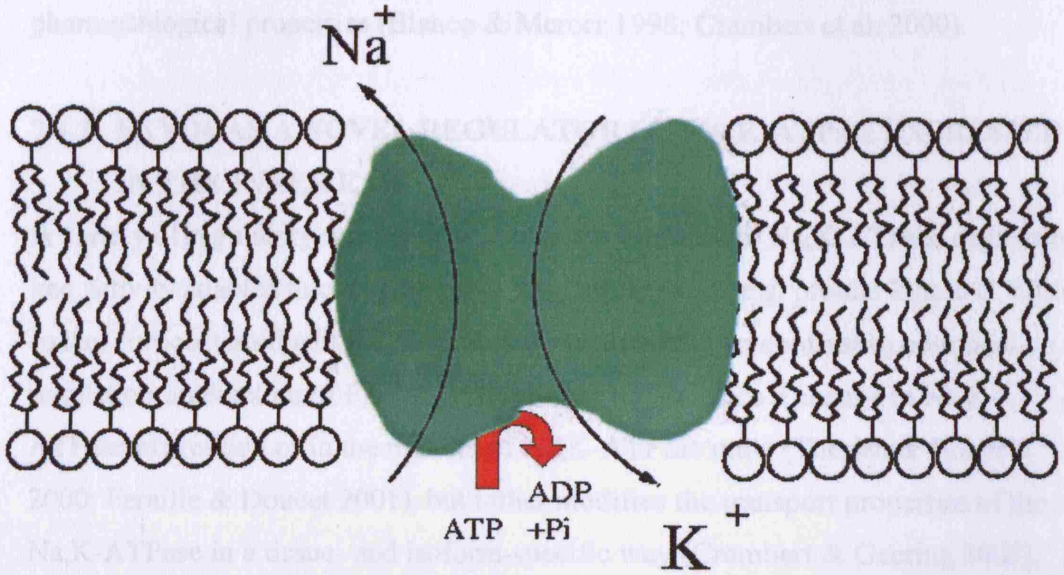
Human tissue samples were labelled to U95A high density oligonucleotide arrays as described in Lockhart et al. (1996) and Sandberg et al. (2000). Primary image analysis of the arrays was performed using Affymetrix GeneChip 3.2. The image is scaled to an average hybridisation intensity of 200. Median value is calculated with $^{+}3$ and $^{+}10$ as multiples of the median. The black, blue and red lines correspond to the median, 3 x median and 10 x median value respectively. Error bars were determined from duplicate samples/hybridisations. Figure taken from GNF SymAtlas (Su et al. 2002).

The prominent levels of FXYD6 expression in regions of the brain thought to be involved in schizophrenia, as identified by brain imaging abnormalities (Section 1.5.1), provide strong support that the FXYD6 gene increases genetic susceptibility to schizophrenia. Although FXYD6 is highly expressed in the brain, we can only hypothesise that its functional role is to regulate the Na,K-ATPase in regions of the brain where it is strongly expressed.

7.4 NA,K-ATPase

Na,K-ATPase is responsible for preserving Na^+ and K^+ gradients across the plasma membrane, specifically as an electrogenic transmembrane ATPase. These gradients are essential for basic cellular homeostasis, maintaining membrane potential and regulating cellular volume (Horisberger 2004). The presence of Na,K-ATPase in nerve cells permits restoration of the Na^+ and K^+ gradients during neuronal activity and thus ensures neuronal excitability. Using energy from ATP hydrolysis, Na,K-ATPase actively transports three Na^+ ions out of the cell in exchange for two K^+ ions into the cell (Figure 7.9).

Figure 7.9: The mechanism of the Na,K-ATPase pump activity.



ATP bound Na,K-ATPase binds with three intracellular Na^+ ions. The ATP is then hydrolysed, leading to the transfer of the γ-phosphate into a highly conserved aspartate residue of the Na,K-ATPase pump. ADP is subsequent released following Na,K-ATPase phosphorylation. A conformational change in the Na,K-ATPase pump exposes the Na^+ ions to the outside, where they are released due to their low affinity with phosphorylated Na,K-ATPase. The Na,K-ATPase pump then binds with two extracellular K^+ ions. This causes the dephosphorylation of Na,K-ATPase, which reverts back to its previous conformational state. As a result, the K^+ ions are transported into the cell. The unphosphorylated form of the Na,K-ATPase pump has a higher affinity for Na^+ ions than K^+ ions, so the two bound K^+ ions are released. ATP binds to the Na,K-ATPase and the whole process starts again.

7.4.1 Na,K-ATPase SUBUNITS

The Na,K-ATPase protein is P-type ATPase heterodimer composed of a catalytic α subunit and a regulatory β subunit (Crambert & Geering 2003). The α subunit, which has 10 transmembrane segments, hydrolyses ATP, becomes phosphorylated during the catalytic cycle and transports the cations (Kuhlbrandt 2004). The type II β subunit acts as a molecular chaperone necessary for the correct and stable membrane insertion of the α subunit (Geering 2001). It also modulates the transport properties of the enzyme (Jaisser et al. 1994; Crambert et al. 2000). Four α and three β isoforms have been identified, which show a tissue-specific expression and can potentially

form 12 different oligomeric Na,K-ATPase isozymes with distinct transport and pharmacological properties (Blanco & Mercer 1998; Crambert et al. 2000).

7.4.2 FXYD6 AS A NOVEL REGULATOR OF NA,K-ATPase EXPRESSED IN THE INNER EAR

A variety of regulatory mechanisms guarantee appropriate Na,K-ATPase expression and activity adapted to changing physiological demands e.g. protein kinase C, nitric oxide, corticosteroid and FXYD proteins (Section 7.2). In contrast to hormonal regulation, interaction of FXYD proteins does not produce a change in Na,K-ATPase expression or in the number of Na,K-ATPase units (Therien & Blostein 2000; Feraille & Doucet 2001), but rather modifies the transport properties of the Na,K-ATPase in a tissue- and isoform-specific way (Crambert & Geering 2003).

The delicate precision of the cochlea, which mediates the transduction of sound waves into nerve impulses, depends on the endolymph ionic composition and the endocochlear potential. This is maintained by Na,K-ATPase activity (Kuijpers & Bonting 1970a; Kuijpers & Bonting 1970b). From Western blot analysis and immunohistochemistry, Delprat et al. (2007b) detected phosphohippolin in various rat epithelial cells in the auditory neurones. Phosphohippolin was shown to co-localise with Na,K-ATPase in the stria vascularis and modulate the Na,K-ATPase transport properties in a way that is compatible with a role in the endolymph production and/or endococlear potential generation. Expression studies in *Xenopus laevis* oocytes revealed that phosphohippolin associates with Na,K-ATPase $\alpha1/\beta1$ and $\alpha1/\beta2$ isozymes, which are preferentially expressed in different regions of the inner ear, but not with $\alpha2/\beta1$ and $\alpha3/\beta1$ isozymes (Delprat et al. 2007b). The apparent K^+ and Na^+ affinities of $\alpha1/\beta1$ and $\alpha1/\beta2$ isozymes were found to be different i.e. association of FXYD6 with Na,K-ATPase $\alpha1/\beta1$ isozymes decreases their apparent K^+ and Na^+ affinity (two-fold), whereas association with $\alpha1/\beta2$ isozymes slightly increases their apparent K^+ and Na^+ affinity.

In the follow-up study, reverse transcriptase-polymerase chain reaction (RT-PCR) analysis by Delprat et al. (2007a) demonstrated that FXYD6 is present as two splice variants in the inner ear. Sequencing of the PCR products revealed one sequence corresponding to that of the database entries (FXYD6) and one identical sequence lacking alanine at position 19 (FXYD6delA). Both variants were found to co-immunoprecipitate with Na,K-ATPase after expression in *Xenopus laevis* oocytes. In addition, the FXYD6delA variant was able to associate with Na,K-ATPase α1/β1 and α1/β2 isozymes just like the FXYD6 variant (Delprat et al. 2007a; Delprat et al. 2007b). The physiological significance of the posttranslational modified splice variant in unknown since both FXYD6 species are expressed in the inner ear, but it is not possible to distinguish between them in the tissue.

7.5 FXYD6, NA,K-ATPASE AND SCHIZOPHRENIA

The current theory for the involvement of Na,K-ATPase centres around the glutamatergic hypothesis of schizophrenia. Interest in the neurotransmitter glutamate and the reduced function of the N-methyl-D-Aspartate (NMDA) glutamate receptor in schizophrenia is based on the abnormally low levels of glutamate receptors found in postmortem brains of people previously diagnosed with schizophrenia (Konradi & Heckers 2003). Glutamate blocking drugs (antagonist) such as phencyclidine (PCP) and ketamine have been discovered to mimic the symptoms (both positive and negative) and cognitive problems associated with the disease (Royston & Simpson 1991; Lahti et al. 2001). Post mortem studies of schizophrenic brains have suggested that NMDA and kainate receptors are decreased in temporal lobe structures, such as the hippocampus and the entorhinal cortex, but increase in the frontal cortex (Royston & Simpson 1991). The fact that reduced glutamate function is linked to poor performance on tests requiring frontal lobe and hippocampal function plus the fact that glutamate can affect dopamine function, all of which have been implicated in schizophrenia, suggest a possibly causal role of glutamate pathways in schizophrenia (Deutsch et al. 1989; Coyle 1996; Coyle et al. 2003).

Na,K-ATPase is an associated molecule of the glutamate transporters. It helps to regulate the function of glutamate synapses by generating Na^+ and K^+ transmembrane gradients (Balcar 2002; Shigeri et al. 2004). Proposals that Na,K-ATPase could be altered in schizophrenia have been described in the literature for a long time (Parker & Hoffman 1964; Viukari 1972; Rybakowski & Lehmann 1994) and have been mentioned recently in relation to possible inhibition of Na,K-ATPase by endogenous inhibitors (Kurup & Kurup 2003). The first Na,K-ATPase study in schizophrenia was reported by Seeman and O'Brien (1963). Here, the Na,K-ATPase activity in schizophrenic cases was found to be over twice the mean value of normal subjects. The suggestion was made that schizophrenia might be a disease of cell excitability. Using evidence that chlorpromazine inhibits Na,K-ATPase (Akera & Brody 1968), Viukari (1972) stated that Na,K-ATPase might act as a binding site for several drugs and that schizophrenia may be linked to the Na,K-ATPase system. A relative lack of intracellular sodium and a prevailing effect of potassium in schizophrenia were proposed. Any involvement of Na,K-ATPase in schizophrenia, particularly through a deficient glutamatergic neurotransmission, is likely to occur through Na,K-ATPase subunits that are linked to glutamate transport (Cholet et al. 2002).

Both Yamaguchi et al. (2001) and the Allen Brain Atlas Database (Lein et al. 2006) revealed high levels of FXYD6 expressed in the CA1 layer of the hippocampus (Section 7.3). After the prefrontal and the superior temporal cortices, the hippocampus is the structure most consistently implicated in acute psychosis and schizophrenia (Schmajuk 2001; Harrison 2004; Heckers 2004). A number of studies describe reductions of hippocampal volume in schizophrenic patients (Nelson et al. 1998; Wright et al. 2000). The hippocampal CA1 pyramidal neurones have been reported to be modified by dendritic hyperpolarisation-activated currents (Magee 1998). Therefore, the high levels of phosphohippolin expressed in the hippocampus suggests that it may be involved in the integration of synaptic inputs by inducing slow activating voltage-dependent ion current. Through modulation of Na,K-ATPase activity, the induced transmembrane ionic flow could respond to neuronal resting

membrane potential changes that play an important role in cellular signal transduction (Pape 1996; Luthi & McCormick 1998).

Dizocilpine, a member of the PCP family, is a potent, non-competitive NMDA receptor antagonist (Wong et al. 1986), which was found to induce psychosis and to enhance the symptoms in schizophrenics (Javitt & Zukin 1991; Lahti et al. 1995). Gamma rhythm generating circuits are prominent in the hippocampal formation (Leung 1998; Gloveli et al. 2005a; Gloveli et al. 2005b; Mann et al. 2005) and it has been suggested that functional alterations in gamma frequency underlie some of the core symptoms of schizophrenia (Lee et al. 2003). Kehrer et al. (2007) investigated dizocilpine induced hippocampal gamma activity in schizophrenia. They proposed that cellular alterations in Na^+/K^+ pump activity and increases in phasic inhibition of dizocilpine cells to be the underlying mechanisms responsible for the more depolarised resting membrane potentials. This shift in the resting membrane potential of the CA1 pyramids, to a more depolarised level, may reflect one of the basic pathophysiological mechanisms in schizophrenia and could possibly involve phosphohippolin through Na,K-ATPase maintenance of Na^+/K^+ gradients.

8.0 THESIS DISCUSSION

8.1 PROJECT SUMMARY

The chromosome 11q22-24 region has been shown to be one of the most well established linkages to schizophrenia by a meta-analysis of twenty genome scans (Section 1.9.1 and 1.10). This provided considerable confidence that a schizophrenia susceptibility locus was likely to exist within this vicinity.

In this thesis I report fine mapping of a schizophrenia susceptibility gene in the chromosome 11q22-24 region, determined on the basis of a UCL sample of 496 cases and 488 supernormal controls. Seven microsatellite or SNP markers localised within or near the FXYD6 gene showed empirically significant allelic associations with schizophrenia (Section 4.3.1 and 4.3.2). Several haplotypes were also found to be associated (Section 4.3.3). Confirmation was then performed by genotyping selected SNPs in the Aberdeen sample consisting of 858 cases and 591 controls. Two of these markers were found to be associated with schizophrenia in the Scottish sample (Section 5.3).

Sequencing of the exons and the promoter region of the FXYD6 gene led to the discovery of a rare non-database SNP, which was found at a significantly higher frequency in the 32 selected high risk haplotype schizophrenic patients when compared to the 32 randomly chosen control individuals (Section 6.3.2). Genotyping this SNP in the whole UCL case-control sample produced no significant association to schizophrenia (Section 6.4.2).

The FXYD6 gene encodes for a protein called phosphohippolin, which is highly expressed in regions of the brain thought to be involved in schizophrenia (Section 1.5.1). It is part of a family that includes seven known mammalian FXYD genes. These FXYD proteins function by modulating the kinetic properties of Na,K-ATPase to the specific physiological requirements of the tissue (Section 7.2.3).

Phosphohippolin is predominantly expressed in the brain, with high levels of expression in the hippocampus and the cerebellum (Section 7.3).

Aetiological base pair changes in FXYD6 or in associated promoter/control regions are likely to cause abnormal function or expression of phosphohippolin and increase genetic susceptibility to schizophrenia.

8.2 FUTURE WORK

8.2.1 FURTHER REPLICATIONS

Given the positive findings from two different samples, further replication studies from independent research groups are required to supplement and promote the results. Finding FXYD6 association with schizophrenia in other populations will help it become a more credible contender as schizophrenia susceptibility candidate gene. However, attention should be paid to the methodology before any replications are attempted. This is because replications of association studies have become a universal concern in the field of genetics. Several journals have proposed guidelines for genetic association studies to provide assistance with study plan and manuscript preparation (Saito et al. 2006). These guidelines help to remove false positive results that arise through chance or study bias and also provide tips on how to display results so that can be easily interpreted.

False positive results can arise from a number of different reasons (Section 1.8.4). An example of this is population stratification, which can occur do to hidden allele frequency differences detected from ethnic/ancestral divergence amongst the cases and the controls. Here, when cases and controls are drawn from two or more ethnic groups, the disease prevalence varies by race along the genetic variant frequency. A recent meta-analysis concluded that population heterogeneity was in evidence for association studies (Ioannidis et al. 2001). Despite this, over 50% of allelic associations in case-control samples have been replicated and confirmed (Lohmueller et al., (2003). Therefore, any future investigations of FXYD6 in

schizophrenia must involve attempts at replicating the positive associations reported here. Such replications should aim to use the same genetic markers of the initial UCL sample study (Section 4.0) with sufficient numbers of cases and controls used to obtain adequate power.

8.2.2 RESEQUENCING

In the resequencing of the FXYD6 gene, or any genes that are thought to be involved in susceptibility to schizophrenia, it is important to discover aetiological base pair changes or to confirm that no such abnormalities have been found. Like any other procedure, the bi-directional method used for resequencing is not without faults.

A larger region of the 5' and 3' UTR should also be sequenced to detect long range control regions or enhancer elements that could affect the expression efficiency of the gene. These locus control regions could be further than 14 kb away and situated within another gene (Ho et al. 2006). Alternatively, the FXYD6 gene could contain a long distance control region or an alternative gene that is responsible for the genetic association. Therefore, it is important to complete sequencing of the entire genomic extent of the gene. It is thus not unreasonable to suggest that any future work should involve resequencing the entire intronic region of the FXYD6 gene. Indeed several diseases have been found to have a mutation within the introns e.g. Myotonic dystrophy type 2 (DM2), which is caused by a CCTG repeat expansion (mean ~5000) located in intron 1 of the zinc finger protein 9 (ZNF9) gene (Liquori et al. 2001). Also, closely linked or even causative abnormalities can be found in introns e.g. the three SNP haplotype in intron 1 of OCA2 (Oculocutaneous albinism type 2), which is thought to influence eye colour variation (Duffy et al. 2007). Finding an aetiological mutation could have implications for the pathogenesis, diagnosis and treatment of schizophrenic patients in the future.

8.2.3 GENOTYPING

In any future work, the ndSNPs discovered during resequencing should be genotyped in the whole UCL case-control sample. In addition, more database SNPs including those found in resequencing and further SNPs chosen from the Hapmap should be genotyped across the region so that the role of FXYD6 as a schizophrenia susceptibility gene can be further investigated. It should also be noted that HapMap does not contain all known SNPs covering a gene and that typing tagged SNPs within the FXYD6 gene does not necessarily provide complete information. Therefore, fully sequencing the gene should help find additional SNPs that require genotyping in the whole UCL case-control sample.

8.2.4 GENOME SCANS

Whole genome scan association studies, using large numbers of SNPs, have been carried out in several other diseases including myocardial infarction, breast cancer and osteoarthritis (Abel et al. 2006). Recently, Mah et al (2006) carried out a genomewide scan using over 25,000 SNPs located within approximately 14,000 genes. They repeated the scan within several populations and found a marker on chromosome 1q32 within PLXNA2, a novel candidate gene for schizophrenia (Section 1.9.2.1).

Future work will include high density genomewide association studies aiming to reinforce current findings and/or identify further candidate genes. The identification of genes by markers in LD with specific loci may not be the final route by which complex disorders are completely understood. The loci may simply be regions for interaction or control. Understanding these interactions and control mechanisms will possibly be the main focus of attention for genetic scientists over the next decade.

8.2.5 KNOCKOUT (KO) MICE MODELS

KO mice models can be used to study how the presence/absence of FXYD6 affects the activity of the sodium pump. Thus, observing the characteristics of FXYD6 KO mice may provide information to better understand how the gene causes or

contributes to schizophrenia in humans. However, it is important to be aware of the fact that deletion of FXYD6 proteins may have complex physiological consequences e.g. turning on compensatory homeostatic mechanisms, which obscure phenotypic changes in the absence of additional physiological stresses such as osmotic stress or salt loading. An example of this is γ KO mice that are viable, without observable pathology. Here, urinary secretion of Na^+, K^+ and $Mg2^+$ was found to be unaffected (59), despite the association of G41R mutation with renal hypomagnesemia (64). In addition, γ is reported to be expressed in the embryo and required for the cavitation process during the blastocyst stage (76). Yet, no delay in blastocoel formation or embryonic development was detected in the KO mouse (59). Another aspect of the physiological complexity of FXYD6 protein function is that maintenance of ionic balance *in vivo* must involve both short-term and long-term regulatory mechanisms. Finally, we can not exclude the possibility that FXYD6 proteins modulate ion transport proteins other than Na,K-ATPase.

8.2.6 METHYLATION STATUS OF THE FXYD6 CpG ISLAND

DNA methylation involves the addition of a methyl group to the number 5 carbon of the cytosine pyrimidine ring. Methylation of the CpG Island suppresses gene expression of the upstream gene. This can silence the gene resulting in loss of gene function. The activity is mainly important for developmental roles and specific tissue expression. We have found a rare ndSNP in the CpG Island (Section 6.3.2). It would be interesting to investigate how this mutation could affect the methylation status in the promoter region.

Measuring DNA methylation can be carried out by sodium bisulfite treatment, which creates sequence differences by converting unmethylated cytosines to uracils, but leaving methylated cytosine unchanged. The differences can then be detected quantitatively by several techniques e.g. sequencing subclones or PCR products, restriction digest or pyrosequencing (Clark et al. 1994; Xiong & Laird 1997; Colella et al. 2003).

8.2.7 MRI SCANS OF FXYD6 ASSOCIATED CASES

Magnetic Resonance Imaging (MRI) is a powerful scanning device, which is able to build up detailed pictures of internal tissue organs using a safe form of radio waves. It is primarily used in medical imaging to demonstrate pathological or other physiological alterations of living tissues. If multiple studies confirm that FXYD6 is a schizophrenia susceptibility gene, investigation of how abnormally expressed or abnormally functioning FXYD6 effects brain morphology of schizophrenic patients would be a possible future step. This could be carried out by scanning selected patients who have inherited the alleles and haplotypes associated with schizophrenia. It would be interesting to see which areas of the brain are affected and then contrast this to other schizophrenics and normal controls.

8.3 CONCLUSION

FXYD6 has been implicated in an English UCL case-control sample and then replicated in a larger Scottish Aberdeen case-control sample. It is an attractive schizophrenia susceptibility candidate gene due to the fact that is it is widely expressed in regions of the brain thought to be involved in the disease. The greatest amount of phosphohippolin expression is found in the hippocampus and the cerebellum. Thus, suggesting that it may be involved in the integration of synaptic inputs and play an important role in neuronal excitability. Further association studies in independent populations are required to fully confirm FXYD6 as a susceptibility gene for schizophrenia.

9.0 ELECTRONIC REFERENCES

Allen Brain Atlas, http://www.brain-map.org

Ensembl, http://www.ensembl.org

Database of Genomic Variants, http://projects.tcag.ca/variation

GDB Human Genome Database, http://www.gdb.org

Genetic Power Calculator, http://pngu.mgh.harvard.edu/~purcell/gpc

GNF SymAtlas, http://symatlas.gnf.org/SymAtlas

Haploview, http://www.broad.mit.edu/mpg/haploview

International Haplotype Project, http://www.hapmap.org

KBioscience, http://www.Kbioscience.co.uk

LocusView2.0, http://www.broad.mit.edu/mpg/locusview

MWG, http://www.mwg-biotech.com/html/all/index.php

National Center for Biotechnology Information (NCBI), http://www.ncbi.nih.gov

NetPrimer, http://www.premierbiosoft.com/netprimer/netprlaunch/netprlaunch.html

Primer3, http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi

Sigma, http://www.sigmaaldrich.com

TESS, http://www.cbil.upenn.edu/cgi-bin/tess/tess

UCSC Genome Browser, http://genome.ucsc.edu

10.0 REFERENCES

Abel, K., Reneland, R., Kammerer, S., Mah, S., Hoyal, C., Cantor, C. R., Nelson, M. R., et al. (2006). Genome-wide SNP association: identification of susceptibility alleles for osteoarthritis. Autoimmun Rev 5(4): 258-63.

Ahlers, B. A., Zhang, X. Q., Moorman, J. R., Rothblum, L. I., Carl, L. L., Song, J., Wang, J., et al. (2005). Identification of an endogenous inhibitor of the cardiac Na+/Ca2+ exchanger, phospholemman. J Biol Chem 280(20): 19875-82.

Aizman, R., Asher, C., Fuzesi, M., Latter, H., Lonai, P., Karlish, S. J. & Garty, H. (2002). Generation and phenotypic analysis of CHIF knockout mice. Am J Physiol Renal Physiol 283(3): F569-77.

Akbarian, S., Kim, J. J., Potkin, S. G., Hetrick, W. P., Bunney, W. E. & Jr.Jones, E. G. (1996). Maldistribution of interstitial neurons in prefrontal white matter of the brains of schizophrenic patients. Arch Gen Psychiatry 53(5): 425-36.

Akera, T. & Brody, T. M. (1968). Inhibition of brain sodium- and potassium-stimulated adenosine triphosphatase activity by chlorpromazine free radical. Mol Pharmacol 4(6): 600-12.

Andreasen, N. C. (1995). Symptoms, signs, and diagnosis of schizophrenia. Lancet 346(8973): 477-81.

APA (1994). APA 1994 Diagnostic and statistical manual of mental disorders. Washington, DC, American Psychiatric Association.

Arystarkhova, E., Wetzel, R. K., Asinovski, N. K. & Sweadner, K. J. (1999). The gamma subunit modulates Na(+) and K(+) affinity of the renal Na,K-ATPase. J Biol Chem 274(47): 33183-5.

Attali, B., Latter, H., Rachamim, N. & Garty, H. (1995). A corticosteroid-induced gene expressing an "IsK-like" K+ channel activity in Xenopus oocytes. Proc Natl Acad Sci U S A 92(13): 6092-6.

Badner, J. A. & Gershon, E. S. (2002). Meta-analysis of whole-genome linkage scans of bipolar disorder and schizophrenia. Molecular Psychiatry 7: 405-411.

Bakker, S. C., Hoogendoorn, M. L., Selten, J. P., Verduijn, W., Pearson, P. L., Sinke, R. J. & Kahn, R. S. (2004). Neuregulin 1: genetic support for schizophrenia subtypes. Mol Psychiatry 9(12): 1061-3.

Balcar, V. J. (2002). Molecular pharmacology of the Na+-dependent transport of acidic amino acids in the mammalian central nervous system. Biol Pharm Bull 25(3): 291-301.

Baron, M. (1986). Genetics of schizophrenia: I. Familial patterns and mode of inheritance. Biol Psychiatry 21(11): 1051-66.

Baron, M. (2001). Genetics of schizophrenia and the new millennium: progress and pitfalls. Am J Hum Genet 68(2): 299-312.

Baron, M., Gruen, R., Asnis, L. & Kane, J. (1982). Schizoaffective illness, schizophrenia and affective disorders: morbidity risk and genetic transmission. Acta Psychiatr Scand 65(4): 253-62.

Barrett, J. C., Fry, B., Maller, J. & Daly, M. J. (2005). Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21(2): 263-5.

Beguin, P., Crambert, G., Guennoun, S., Garty, H., Horisberger, J. D. & Geering, K. (2001). CHIF, a member of the FXYD protein family, is a regulator of Na,K-ATPase distinct from the gamma-subunit. Embo J 20(15): 3993-4002.

Beguin, P., Crambert, G., Monnet-Tschudi, F., Uldry, M., Horisberger, J. D., Garty, H. & Geering, K. (2002). FXYD7 is a brain-specific regulator of Na,K-ATPase alpha 1-beta isozymes. Embo J 21(13): 3264-73.

Beguin, P., Wang, X., Firsov, D., Puoti, A., Claeys, D., Horisberger, J. D. & Geering, K. (1997). The gamma subunit is a specific component of the Na,K-ATPase and modulates its transport function. Embo J 16(14): 4250-60.

Bengra, C., Mifflin, T. E., Khripin, Y., Manunta, P., Williams, S. M., Jose, P. A. & Felder, R. A. (2002). Genotyping of essential hypertension single-nucleotide polymorphisms by a homogeneous PCR method with universal energy transfer primers. Clin Chem 48(12): 2131-40.

Bishop, Y., Fienberg, S. E. & Holland, P. W. (1975). Discreate Multivariate Analysis: Theory and Practice, Cambridge, Mass., MIT Press.

Blackwood, D., Fordyce, A., Walker, M., St. Clair, D., Porteous, D. & Muir, W. (1998). Thirty year follow up of a family showing association of schizophrenia with a balanced translocation t(1:11)(q42.1,q14.3). American Journal of Medical Genetics 81: 532.

Blackwood, D. H., Fordyce, A., Walker, M. T., St Clair, D. M., Porteous, D. J. & Muir, W. J. (2001). Schizophrenia and affective disorders--cosegregation with a translocation at chromosome 1q42 that directly disrupts brain-expressed genes: clinical and P300 findings in a family. Am J Hum Genet 69(2): 428-33.

Blanco, G. & Mercer, R. W. (1998). Isozymes of the Na-K-ATPase: heterogeneity in structure, diversity in function. Am J Physiol 275(5 Pt 2): F633-50.

Bogaev, R. C., Jia, L. G., Kobayashi, Y. M., Palmer, C. J., Mounsey, J. P., Moorman, J. R., Jones, L. R., et al. (2001). Gene structure and expression of phospholemman in mouse. Gene 271(1): 69-79.

Botstein, D. & Risch, N. (2003). Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. Nat Genet 33 Suppl: 228-37.

Brzustowicz, L. M., Simone, J., Mohseni, P., Hayter, J. E., Hodgkinson, K. A., Chow, E. W. & Bassett, A. S. (2004). Linkage disequilibrium mapping of schizophrenia susceptibility to the CAPON region of chromosome 1q22. Am J Hum Genet 74(5): 1057-63.

Capurro, C., Coutry, N., Bonvalet, J. P., Escoubet, B., Garty, H. & Farman, N. (1996). Cellular localization and regulation of CHIF in kidney and colon. Am J Physiol 271(3 Pt 1): C753-62.

Cardno, A. G. & Gottesman, II (2000). Twin studies of schizophrenia: from bow-and-arrow concordances to star wars Mx and functional genomics. Am J Med Genet 97(1): 12-7.

Castle, D., Wessely, S., Der, G. & Murray, R. M. (1991). The incidence of operationally defined schizophrenia in Camberwell, 1965-84. Br J Psychiatry 159: 790-4.

Chen, X., Dunham, C., Kendler, S., Wang, X., O'Neill, F. A., Walsh, D. & Kendler, K. S. (2004). Regulator of G-protein signaling 4 (RGS4) gene is associated with schizophrenia in Irish high density families. Am J Med Genet B Neuropsychiatr Genet 129(1): 23-6.

Chen, Z. H., Jones, L. R. & Moorman, J. R. (1999). Ion currents through mutant phospholemman channel molecules. Receptors Channels 6(6): 435-47.

Cholet, N., Pellerin, L., Magistretti, P. J. & Hamel, E. (2002). Similar perisynaptic glial localization for the Na+,K+-ATPase alpha 2 subunit and the glutamate transporters GLAST and GLT-1 in the rat somatosensory cortex. Cereb Cortex 12(5): 515-25.

Choudhury, K., McQuillin, A., Puri, V., Pimm, J., Datta, S., Thirumalai, S., Krasucki, R., et al. (2007). A genetic association study of chromosome 11q22-24 in two different samples implicates the FXYD6 gene, encoding phosphohippolin, in susceptibility to schizophrenia. Am J Hum Genet 80(4): 664-72.

Chowdari, K. V., Mirnics, K., Semwal, P., Wood, J., Lawrence, E., Bhatia, T., Deshpande, S. N., et al. (2002). Association and linkage analyses of RGS4 polymorphisms in schizophrenia. Hum Mol Genet 11(12): 1373-80.

Chumakov, I., Blumenfeld, M., Guerassimenko, O., Cavarec, L., Palicio, M., Abderrahim, H., Bougueleret, L., et al. (2002). Genetic and physiological data implicating the new human gene G72 and the gene for D-amino acid oxidase in schizophrenia. Proc Natl Acad Sci U S A 99(21): 13675-80.

Clark, S. J., Harrison, J., Paul, C. L. & Frommer, M. (1994). High sensitivity mapping of methylated cytosines. Nucleic Acids Res 22(15): 2990-7.

Cohen, E., Goldshleger, R., Shainskaya, A., Tal, D. M., Ebel, C., le Maire, M. & Karlish, S. J. (2005). Purification of Na+,K+-ATPase expressed in Pichia pastoris reveals an essential role of phospholipid-protein interactions. J Biol Chem 280(17): 16610-8.

Colella, S., Shen, L., Baggerly, K. A., Issa, J. P. & Krahe, R. (2003). Sensitive and quantitative universal Pyrosequencing methylation analysis of CpG sites. Biotechniques 35(1): 146-50.

Collins, F. S. (1992). Positional cloning: let's not call it reverse anymore. Nat Genet 1(1): 3-6.

Collins, F. S., Guyer, M. S. & Charkravarti, A. (1997). Variations on a theme: cataloging human DNA sequence variation. Science 278(5343): 1580-1.

Cordell, H. J. (2002). Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans. Hum Mol Genet 11(20): 2463-8.

Cornelius, F. & Mahmmoud, Y. A. (2003). Functional modulation of the sodium pump: the regulatory proteins "Fixit". News Physiol Sci 18: 119-24.

Cornelius, F., Mahmmoud, Y. A., Meischke, L. & Cramb, G. (2005). Functional significance of the shark Na,K-ATPase N-terminal domain. Is the structurally variable N-Terminus involved in tissue-specific regulation by FXYD proteins? Biochemistry 44(39): 13051-62.

Corvin, A. P., Morris, D. W., McGhee, K., Schwaiger, S., Scully, P., Quinn, J., Meagher, D., et al. (2004). Confirmation and refinement of an 'at-risk' haplotype for schizophrenia suggests the EST cluster, Hs.97362, as a potential susceptibility gene at the Neuregulin-1 locus. Mol Psychiatry 9(2): 208-13.

Coyle, J. T. (1996). The glutamatergic dysfunction hypothesis for schizophrenia. Harv Rev Psychiatry 3(5): 241-53.

Coyle, J. T., Tsai, G. & Goff, D. (2003). Converging evidence of NMDA receptor hypofunction in the pathophysiology of schizophrenia. Ann N Y Acad Sci 1003: 318-27.

Crambert, G., Fuzesi, M., Garty, H., Karlish, S. & Geering, K. (2002). Phospholemman (FXYD1) associates with Na,K-ATPase and regulates its transport properties. Proc Natl Acad Sci U S A 99(17): 11476-81.

Crambert, G. & Geering, K. (2003). FXYD proteins: new tissue-specific regulators of the ubiquitous Na,K-ATPase. Sci STKE 2003(166): RE1.

Crambert, G., Hasler, U., Beggah, A. T., Yu, C., Modyanov, N. N., Horisberger, J. D., Lelievre, L., et al. (2000). Transport and pharmacological properties of nine different human Na, K-ATPase isozymes. J Biol Chem 275(3): 1976-86.

Crambert, G., Li, C., Claeys, D. & Geering, K. (2005). FXYD3 (Mat-8), a new regulator of Na,K-ATPase. Mol Biol Cell 16(5): 2363-71.

Crambert, G., Li, C., Swee, L. K. & Geering, K. (2004). FXYD7, mapping of functional sites involved in endoplasmic reticulum export, association with and regulation of Na,K-ATPase. J Biol Chem 279(29): 30888-95.

Curtis, D. (1996). Genetic dissection of complex traits. Nat Genet 12(4): 356-8.

Curtis, D., Knight, J. & Sham, P. C. (2006). Program report: GENECOUNTING support programs. Ann Hum Genet 70(Pt 2): 277-9.

Curtis, D., North, B. V., Gurling, H. M., Blaveri, E. & Sham, P. C. (2002). A quick and simple method for detecting subjects with abnormal genetic background in case-control samples. Ann Hum Genet 66(Pt3): 235-44.

Curtis, D. & Sham, P. C. (1995). Model-free linkage analysis using likelihoods. Am J Hum Genet 57(3): 703-16.

Dai, L. J., Ritchie, G., Kerstan, D., Kang, H. S., Cole, D. E. & Quamme, G. A. (2001). Magnesium transport in the renal distal convoluted tubule. Physiol Rev 81(1): 51-84.

Datta, S., McQuillin, A., Rizig, M. A., Blaveri, E., Thirumalai, S., Kalsi, G., Lawrence, J., et al. (2008). A threonine to isoleucine mis-sense mutation in the pericentriolar material 1 (PCM1) gene is strongly associated with schizophrenia. In Press.

Davis, C. E., Patel, M. K., Miller, J. R., John, J. E., 3rd, Jones, L. R., Tucker, A. L., Mounsey, J. P., et al. (2004). Effects of phospholemman expression on swelling-activated ion currents and volume regulation in embryonic kidney cells. Neurochem Res 29(1): 177-87.

de Bakker, P. I., Yelensky, R., Pe'er, I., Gabriel, S. B., Daly, M. J. & Altshuler, D. (2005). Efficiency and power in genetic association studies. Nat Genet 37(11): 1217-23.

Delprat, B., Puel, J. L. & Geering, K. (2007a). Dynamic expression of FXYD6 in the inner ear suggests a role of the protein in endolymph homeostasis and neuronal activity. Dev Dyn 236(9): 2534-40.

Delprat, B., Schaer, D., Roy, S., Wang, J., Puel, J. L. & Geering, K. (2007b). FXYD6 is a novel regulator of Na,K-ATPase expressed in the inner ear. J Biol Chem 282(10): 7450-6.

Deutsch, S. I., Mastropaolo, J., Schwartz, B. L., Rosse, R. B. & Morihisa, J. M. (1989). A "glutamatergic hypothesis" of schizophrenia. Rationale for pharmacotherapy with glycine. Clin Neuropharmacol 12(1): 1-13.

Devlin, B., Bacanu, S. A., Roeder, K., Reimherr, F., Wender, P., Galke, B., Novasad, D., et al. (2002). Genome-wide multipoint linkage analyses of multiplex schizophrenia pedigrees from the oceanic nation of Palau. Mol Psychiatry 7(7): 689-94.

Devlin, B. & Risch, N. (1995). A comparison of linkage disequilibrium measures for fine-scale mapping. Genomics 29(2): 311-22.

Done, D. J., Johnstone, E. C., Frith, C. D., Golding, J., Shepherd, P. M. & Crow, T. J. (1991). Complications of pregnancy and delivery in relation to psychosis in adult life: data from the British perinatal mortality survey sample. BMJ 302(6792): 1576-80.

Duffy, D. L., Montgomery, G. W., Chen, W., Zhao, Z. Z., Le, L., James, M. R., Hayward, N. K., et al. (2007). A Three-Single-Nucleotide Polymorphism Haplotype in Intron 1 of OCA2 Explains Most Human Eye-Color Variation. Am J Hum Genet 80.

Elston, R. C. & Campbell, M. A. (1970). Schizophrenia: evidence for the major gene hypothesis. Behav Genet 1(1): 3-10.

Evans, K., McGrath, J. & Milns, R. (2003). Searching for schizophrenia in ancient Greek and Roman literature: a systematic review. Acta Psychiatr Scand 107(5): 323-30.

Falconer, D. S. (1965). Maternal effects and selection response. Genetics today 3: 763-764.

Faraone, S. V., Matise, T., Svrakic, D., Pepple, J., Malaspina, D., Suarez, B., Hampe, C., et al. (1998). Genome scan of European-American schizophrenia pedigrees: results of the NIMH Genetics Initiative and Millennium Consortium. Am J Med Genet 81(4): 290-5.

Farmer, A. E., McGuffin, P. & Gottesman, II (1984). Searching for the split in schizophrenia: a twin study perspective. Psychiatry Res 13(2): 109-18.

Feraille, E. & Doucet, A. (2001). Sodium-potassium-adenosinetriphosphatase-dependent sodium transport in the kidney: hormonal control. Physiol Rev 81(1): 345-418.

Feschenko, M. S., Donnet, C., Wetzel, R. K., Asinovski, N. K., Jones, L. R. & Sweadner, K. J. (2003). Phospholemman, a single-span membrane protein, is an accessory protein of Na,K-ATPase in cerebellum and choroid plexus. J Neurosci 23(6): 2161-9.

Forbush, B., 3rd, Kaplan, J. H. & Hoffman, J. F. (1978). Characterization of a new photoaffinity derivative of ouabain: labeling of the large polypeptide and of a proteolipid component of the Na, K-ATPase. Biochemistry 17(17): 3667-76.

Frankle, W. G., Lerma, J. & Laruelle, M. (2003). The synaptic hypothesis of schizophrenia. Neuron 39(2): 205-16.

Freeman, J. L., Perry, G. H., Feuk, L., Redon, R., McCarroll, S. A., Altshuler, D. M., Aburatani, H., et al. (2006). Copy number variation: new insights in genome diversity. Genome Res 16(8): 949-61.

Freking, B. A., Murphy, S. K., Wylie, A. A., Rhodes, S. J., Keele, J. W., Leymaster, K. A., Jirtle, R. L., et al. (2002). Identification of the single base change causing the callipyge muscle hypertrophy phenotype, the only known example of polar overdominance in mammals. Genome Res 12(10): 1496-506.

Fu, X. & Kamps, M. P. (1997). E2a-Pbx1 induces aberrant expression of tissue-specific and developmentally regulated genes when expressed in NIH 3T3 fibroblasts. Mol Cell Biol 17(3): 1503-12.

Fuller, W., Eaton, P., Bell, J. R. & Shattock, M. J. (2004). Ischemia-induced phosphorylation of phospholemman directly activates rat cardiac Na/K-ATPase. Faseb J 18(1): 197-9.

Garty, H. & Karlish, S. J. (2005). FXYD proteins: tissue-specific regulators of the Na,K-ATPase. Semin Nephrol 25(5): 304-11.

Garty, H. & Karlish, S. J. (2006). Role of FXYD proteins in ion transport. Annu Rev Physiol 68: 431-59.

Garty, H., Lindzen, M., Scanzano, R., Aizman, R., Fuzesi, M., Goldshleger, R., Farman, N., et al. (2002). A functional interaction between CHIF and Na-K-ATPase: implication for regulation by FXYD proteins. Am J Physiol Renal Physiol 283(4): F607-15.

Geering, K. (2001). The functional role of beta subunits in oligomeric P-type ATPases. J Bioenerg Biomembr 33(5): 425-38.

Geering, K. (2005). Function of FXYD proteins, regulators of Na, K-ATPase. J Bioenerg Biomembr 37(6): 387-92.

Geering, K. (2006). FXYD proteins: new regulators of Na-K-ATPase. Am J Physiol Renal Physiol 290(2): F241-50.

Gejman, P. V., Ram, A., Gelernter, J., Friedman, E., Cao, Q., Pickar, D., Blum, K., et al. (1994). No structural mutation in the dopamine D2 receptor gene in alcoholism or schizophrenia. Analysis using denaturing gradient gel electrophoresis. Jama 271(3): 204-8.

Gerber, D. J., Hall, D., Miyakawa, T., Demars, S., Gogos, J. A., Karayiorgou, M. & Tonegawa, S. (2003). Evidence for association of schizophrenia with genetic variation in the 8p21.3 gene, PPP3CC, encoding the calcineurin gamma subunit. Proc Natl Acad Sci U S A 100(15): 8993-8.

Gershon, E. S., Martinez, M., Goldin, L. R. & Gejman, P. V. (1990). Genetic mapping of common diseases: the challenges of manic-depressive illness and schizophrenia. Trends Genet 6(9): 282-7.

Giebisch, G. (2001). Renal potassium channels: function, regulation, and structure. Kidney Int 60(2): 436-45.

Gill, M., McGuffin, P., Parfitt, E., Mant, R., Asherson, P., Collier, D., Vallada, H., et al. (1993). A linkage study of schizophrenia with DNA markers from the long arm of chromosome 11. Psychol Med 23(1): 27-44.

Glantz, L. A. & Lewis, D. A. (1997). Reduction of synaptophysin immunoreactivity in the prefrontal cortex of subjects with schizophrenia. Regional and diagnostic specificity. Arch Gen Psychiatry 54(10): 943-52.

Gloveli, T., Dugladze, T., Rotstein, H. G., Traub, R. D., Monyer, H., Heinemann, U., Whittington, M. A., et al. (2005a). Orthogonal arrangement of rhythm-generating microcircuits in the hippocampus. Proc Natl Acad Sci U S A 102(37): 13295-300.

Gloveli, T., Dugladze, T., Saha, S., Monyer, H., Heinemann, U., Traub, R. D., Whittington, M. A., et al. (2005b). Differential involvement of oriens/pyramidale interneurones in hippocampal network oscillations in vitro. J Physiol 562(Pt 1): 131-47.

Goldschimdt, I., Grahammer, F., Warth, R., Schulz-Baldes, A., Garty, H., Greger, R. & Bleich, M. (2004). Kidney and colon electrolyte transport in CHIF knockout mice. Cell Physiol Biochem 14(1-2): 113-20.

Gomes, I., Collins, A., Lonjou, C., Thomas, N. S., Wilkinson, J., Watson, M. & Morton, N. (1999). Hardy-Weinberg quality control. Ann Hum Genet 63(Pt 6): 535-8.

Gottesman, II & Shields, J. (1976). A critical review of recent adoption, twin, and family studies of schizophrenia: behavioral genetics perspectives. Schizophr Bull 2(3): 360-401.

Gottesman, I. I. (1991). Schiziophrenia Genesis. New York, W.H. Freeman.

Gottesman, I. I. & Shields, J. (1967). A polygenic theory of schizophrenia. Proc Natl Acad Sci U S A 58(1): 199-205.

Gottesman, I. I. & Shields, J. (1972). A Twin Study Vantage Point. London, Academic Press.

Green, M. F. (2006). Cognitive impairment and functional outcome in schizophrenia and bipolar disorder. J Clin Psychiatry 67(10): e12.

Gurling, H. M., Critchley, H., Datta, S. R., McQuillin, A., Blaveri, E., Thirumalai, S., Pimm, J., et al. (2006). Genetic association and brain morphology studies and the chromosome 8p22 pericentriolar material 1 (PCM1) gene in susceptibility to schizophrenia. Arch Gen Psychiatry 63(8): 844-54.

Gurling, H. M., Kalsi, G., Brynjolfson, J., Sigmundsson, T., Sherrington, R., Mankoo, B. S., Read, T., et al. (2001). Genomewide genetic linkage analysis confirms the presence of susceptibility loci for schizophrenia, on chromosomes 1q32.2, 5q33.2, and 8p21-22 and provides support for linkage to schizophrenia, on chromosomes 11q23.3-24 and 20q12.1-11.23. Am J Hum Genet 68(3): 661-73.

Gurling, H. M., Sherrington, R. P., Brynjolfsson, J., Read, T., Curtis, D., Mankoo, B. J., Potter, M., et al. (1989). Recent and future molecular genetic research into schizophrenia. Schizophr Bull 15(3): 373-82.

Gusella, J. F., Wexler, N. S., Conneally, P. M., Naylor, S. L., Anderson, M. A., Tanzi, R. E., Watkins, P. C., et al. (1983). A polymorphic DNA marker genetically linked to Huntington's disease. Nature 306(5940): 234-8.

Haas, S. J., Hill, R., Krum, H., Liew, D., Tonkin, A., Demos, L., Stephan, K., et al. (2007). Clozapine-associated myocarditis: a review of 116 cases of suspected myocarditis associated with the use of clozapine in Australia during 1993-2003. Drug Saf 30(1): 47-57.

Hacia, J. G., Brody, L. C., Chee, M. S., Fodor, S. P. & Collins, F. S. (1996). Detection of heterozygous mutations in BRCA1 using high density oligonucleotide arrays and two-colour fluorescence analysis. Nat Genet 14(4): 441-7.

Hafner, H. & an der Heiden, W. (1997). Epidemiology of schizophrenia. Can J Psychiatry 42(2): 139-51.

Hannerz, H., Borga, P. & Borritz, M. (2001). Life expectancies for individuals with psychiatric diagnoses. Public Health 115(5): 328-37.

Hansen, R. A. & Atchison, B. (2000). Conditions in occupational therapy: effect on occupational performance. Baltimore, Lippincott, Williams & Williams: 54–74.

HapMap (2003). The International HapMap Project. Nature 426(6968): 789-96.

Hardy, G. H. (1908). Mendelian proportions in a mixed population. Science 18: 49-50.

Harrison, P. J. (2004). The hippocampus in schizophrenia: a review of the neuropathological evidence and its pathophysiological implications. Psychopharmacology (Berl) 174(1): 151-62.

Harrison, P. J. & Weinberger, D. R. (2005). Schizophrenia genes, gene expression, and neuropathology: on the matter of their convergence. Mol Psychiatry 10(1): 40-68.

Hashimoto, R., Suzuki, T., Iwata, N., Yamanouchi, Y., Kitajima, T., Kosuga, A., Tatsumi, M., et al. (2005). Association study of the frizzled-3 (FZD3) gene with schizophrenia and mood disorders. J Neural Transm 112(2): 303-7.

Hauser, E. R., Haines, J., Goldgar, D.E., (2006). Nonparametric Linkage Analysis. Genetic Analysis of Complex Disease. J. L. Haines, Pericak-Vance, M.A., Wiley: 283-328.

Heckers, S. (2004). The hippocampus in schizophrenia. Am J Psychiatry 161(11): 2138-9.

Hedrick, P. W. (1987). Gametic disequilibrium measures: proceed with caution. Genetics 117(2): 331-41.

Heston, L. L. (1966). Psychiatric disorders in foster home reared children of schizophrenic mothers. Br J Psychiatry 112(489): 819-25.

Ho, Y., Elefant, F., Liebhaber, S. A. & Cooke, N. E. (2006). Locus control region transcription plays an active role in long-range gene activation. Mol Cell 23(3): 365-75.

Horisberger, J. D. (2004). Recent insights into the structure and mechanism of the sodium pump. Physiology (Bethesda) 19: 377-87.

Hovatta, I., Terwilliger, J. D., Lichtermann, D., Makikyro, T., Suvisaari, J., Peltonen, L. & Lonnqvist, J. (1997). Schizophrenia in the genetic isolate of Finland. Am J Med Genet 74(4): 353-60.

Iafrate, A. J., Feuk, L., Rivera, M. N., Listewnik, M. L., Donahoe, P. K., Qi, Y., Scherer, S. W., et al. (2004). Detection of large-scale variation in the human genome. Nat Genet 36(9): 949-51.

Ino, Y., Gotoh, M., Sakamoto, M., Tsukagoshi, K. & Hirohashi, S. (2002). Dysadherin, a cancer-associated cell membrane glycoprotein, down-regulates E-cadherin and promotes metastasis. Proc Natl Acad Sci U S A 99(1): 365-70.

Ioannidis, J. P., Ntzani, E. E., Trikalinos, T. A. & Contopoulos-Ioannidis, D. G. (2001). Replication validity of genetic association studies. Nat Genet 29(3): 306-9.

Iwata, N., Suzuki, T., Ikeda, M., Kitajima, T., Yamanouchi, Y., Inada, T. & Ozaki, N. (2004). No association with the neuregulin 1 haplotype to Japanese schizophrenia. Mol Psychiatry 9(2): 126-7.

Jaisser, F., Jaunin, P., Geering, K., Rossier, B. C. & Horisberger, J. D. (1994). Modulation of the Na,K-pump function by beta subunit isoforms. J Gen Physiol 103(4): 605-23.

Jakob, H. & Beckmann, H. (1986). Prenatal developmental disturbances in the limbic allocortex in schizophrenics. J Neural Transm 65(3-4): 303-26.

Javitt, D. C. & Zukin, S. R. (1991). Recent advances in the phencyclidine model of schizophrenia. Am J Psychiatry 148(10): 1301-8.

Jeffreys, A. J., Wilson, V. & Thein, S.L., (1985). Hypervariable "minisatellite" regions in human DNA. Nature 314: 67-73.

Jespersen, T., Grunnet, M., Rasmussen, H. B., Jorgensen, N. B., Jensen, H. S., Angelo, K., Olesen, S. P., et al. (2006). The corticosteroid hormone induced factor: a new modulator of KCNQ1 channels? Biochem Biophys Res Commun 341(4): 979-88.

Jia, L. G., Donnet, C., Bogaev, R. C., Blatt, R. J., McKinney, C. E., Day, K. H., Berr, S. S., et al. (2005). Hypertrophy, increased ejection fraction, and reduced Na-K-ATPase activity in phospholemman-deficient mice. Am J Physiol Heart Circ Physiol 288(4): H1982-8.

Job, D. E., Whalley, H. C., Johnstone, E. C. & Lawrie, S. M. (2005). Grey matter changes over time in high risk subjects developing schizophrenia. Neuroimage 25(4): 1023-30.

Johnstone, E. C., Crow, T. J., Frith, C. D., Husband, J. & Kreel, L. (1976). Cerebral ventricular size and cognitive impairment in chronic schizophrenia. Lancet 2(7992): 924-6.

Johnstone, E. C., Macmillan, J. F. & Crow, T. J. (1987). The occurrence of organic disease of possible or probable aetiological significance in a population of 268 cases of first episode schizophrenia. Psychol Med 17(2): 371-9.

Jones, D. H., Li, T. Y., Arystarkhova, E., Barr, K. J., Wetzel, R. K., Peng, J., Markham, K., et al. (2005). Na,K-ATPase from mice lacking the gamma subunit (FXYD2) exhibits altered Na+ affinity and decreased thermal stability. J Biol Chem 280(19): 19003-11.

Kadowaki, K., Sugimoto, K., Yamaguchi, F., Song, T., Watanabe, Y., Singh, K. & Tokuda, M. (2004). Phosphohippolin expression in the rat central nervous system. Brain Res Mol Brain Res 125(1-2): 105-12.

Kallman, F. J. (1953). Heredity and eugenics. Am J Psychiatry 109(7): 491-3.

Kalsi, G., Mankoo, B. S., Curtis, D., Brynjolfsson, J., Read, T., Sharma, T., Murphy, P., et al. (1995). Exclusion of linkage of schizophrenia to the gene for the dopamine D2 receptor (DRD2) and chromosome 11q translocation sites. Psychol Med 25(3): 531-7.

Kan, Y. W. & Dozy, A.M. (1978). Polymorphism of DNA sequence adjacent to human beta-globin structural gene: relationship to sickle mutation. Proc. Natl. Acad. Sci. 75(11): 5631-5.

Kaufmann, C. A., Suarez, B., Malaspina, D., Pepple, J., Svrakic, D., Markel, P. D., Meyer, J., et al. (1998). NIMH Genetics Initiative Millenium Schizophrenia Consortium: linkage analysis of African-American pedigrees. Am J Med Genet 81(4): 282-9.

Kehrer, C., Dugladze, T., Maziashvili, N., Wojtowicz, A., Schmitz, D., Heinemann, U. & Gloveli, T. (2007). Increased inhibitory input to CA1 pyramidal cells alters hippocampal gamma frequency oscillations in the MK-801 model of acute psychosis. Neurobiol Dis 25(3): 545-52.

Kendler, K. S., McGuire, M., Gruenberg, A. M., O'Hare, A., Spellman, M. & Walsh, D. (1993). The Roscommon Family Study. IV. Affective illness, anxiety disorders, and alcoholism in relatives. Arch Gen Psychiatry 50(12): 952-60.

Kendler, K. S. & Zerbin-Rudin, E. (1996). Abstract and review of "Studien Uber Vererbung und Entstehung Geistiger Storungen. I. Zur Vererbung und Neuentstehung der Dementia praecox." (Studies on the inheritance and origin of mental illness: I. To the problem of the inheritance and primary origin of dementia praecox). 1916. Am J Med Genet 67(4): 338-42.

Kety, S. S., Rosenthal, D., Wender, P. H. & Schulsinger, F. (1971). Mental illness in the biological and adoptive families of adpoted schizophrenics. Am J Psychiatry 128(3): 302-6.

Kety, S. S., Rosenthal, D., Wender, P. H., Schulsinger, F. & Jacobsen, B. (1976). Mental illness in the biological and adoptive families of adopted individuals who have become schizophrenic. Behav Genet 6(3): 219-25.

Kety, S. S., Wender, P. H., Jacobsen, B., Ingraham, L. J., Jansson, L., Faber, B. & Kinney, D. K. (1994). Mental illness in the biological and adoptive relatives of schizophrenic adoptees. Replication of the Copenhagen Study in the rest of Denmark. Arch Gen Psychiatry 51(6): 442-55.

Kidd, K. K. & Cavalli-Sforza, L.L. (1973). An Analysis of the genetics of schizophrenia. Soc. Biol. 20: 254-265.

Kimura-Yoshida, C., Kitajima, K., Oda-Ishii, I., Tian, E., Suzuki, M., Yamamoto, M., Suzuki, T., et al. (2004). Characterization of the pufferfish Otx2 cis-regulators reveals evolutionarily conserved genetic mechanisms for vertebrate head specification. Development 131(1): 57-71.

Kinoshita, Y., Suzuki, T., Ikeda, M., Kitajima, T., Yamanouchi, Y., Inada, T., Yoneda, H., et al. (2005). No association with the calcineurin A gamma subunit gene (PPP3CC) haplotype to Japanese schizophrenia. J Neural Transm 112(9): 1255-62.

Kirov, G., Ivanov, D., Williams, N. M., Preece, A., Nikolov, I., Milev, R., Koleva, S., et al. (2004). Strong evidence for association between the dystrobrevin binding protein 1 gene (DTNBP1) and schizophrenia in 488 parent-offspring trios from Bulgaria. Biol Psychiatry 55(10): 971-5.

Kleinjan, D. A., Seawright, A., Schedl, A., Quinlan, R. A., Danes, S. & van Heyningen, V. (2001). Aniridia-associated translocations, DNase hypersensitivity, sequence comparison and transgenic analysis redefine the functional domain of PAX6. Hum Mol Genet 10(19): 2049-59.

Kleinjan, D. A. & van Heyningen, V. (2005). Long-range control of gene expression: emerging mechanisms and disruption in disease. Am J Hum Genet 76(1): 8-32.

Konradi, C. & Heckers, S. (2003). Molecular aspects of glutamate dysregulation: implications for schizophrenia and its treatment. Pharmacol Ther 97(2): 153-79.

Korostishevsky, M., Kaganovich, M., Cholostoy, A., Ashkenazi, M., Ratner, Y., Dahary, D., Bernstein, J., et al. (2004). Is the G72/G30 locus associated with schizophrenia? single nucleotide polymorphisms, haplotypes, and gene expression analysis. Biol Psychiatry 56(3): 169-76.

Kovelman, J. A. & Scheibel, A. B. (1984). A neurohistological correlate of schizophrenia. Biol Psychiatry 19(12): 1601-21.

Kowdley, G. C., Ackerman, S. J., Chen, Z., Szabo, G., Jones, L. R. & Moorman, J. R. (1997). Anion, cation, and zwitterion selectivity of phospholemman channel molecules. Biophys J 72(1): 141-5.

Kozak, M. (1987). An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Res 15(20): 8125-48.

Kraepelin, E. (1907). Text book of psychiatry, 7th edition. Translated by A. R. Diefendorf. London, Macmillan.

Krebs, M. O., Guillin, O., Bourdell, M. C., Schwartz, J. C., Olie, J. P., Poirier, M. F. & Sokoloff, P. (2000). Brain derived neurotrophic factor (BDNF) gene variants association with age at onset and therapeutic response in schizophrenia. Mol Psychiatry 5(5): 558-62.

Kringlen, E. (2000). Twin studies in schizophrenia with special emphasis on concordance figures. Am J Med Genet 97(1): 4-11.

Kuhlbrandt, W. (2004). Biology, structure and mechanism of P-type ATPases. Nat Rev Mol Cell Biol 5(4): 282-95.

Kuijpers, W. & Bonting, S. L. (1970a). The cochlear potentials. I. The effect of ouabain on the cochlear potentials of the guinea pig. Pflugers Arch 320(4): 348-58.

Kuijpers, W. & Bonting, S. L. (1970b). The cochlear potentials. II. The nature of the cochlear endolymphatic resting potential. Pflugers Arch 320(4): 359-72.

Kurup, R. K. & Kurup, P. A. (2003). Schizoid neurochemical pathology-induced membrane Na(+)-K+ ATPase inhibition in relation to neurological disorders. Int J Neurosci 113(12): 1705-17.

Kuster, B., Shainskaya, A., Pu, H. X., Goldshleger, R., Blostein, R., Mann, M. & Karlish, S. J. (2000). A new variant of the gamma subunit of renal Na,K-ATPase. Identification by mass spectrometry, antibody binding, and expression in cultured cells. J Biol Chem 275(24): 18441-6.

Lahti, A. C., Koffel, B., LaPorte, D. & Tamminga, C. A. (1995). Subanesthetic doses of ketamine stimulate psychosis in schizophrenia. Neuropsychopharmacology 13(1): 9-19.

Lahti, A. C., Weiler, M. A., Tamara Michaelidis, B. A., Parwani, A. & Tamminga, C. A. (2001). Effects of ketamine in normal and schizophrenic volunteers. Neuropsychopharmacology 25(4): 455-67.

Lalouel, J. M., Rao, D. C., Morton, N. E. & Elston, R. C. (1983). A unified model for complex segregation analysis. Am J Hum Genet 35(5): 816-26.

Lander, E. & Kruglyak, L. (1995). Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. Nat Genet 11(3): 241-7.

Lander, E. S. & Schork, N. J. (1994). Genetic dissection of complex traits. Science 265(5181): 2037-48.

Law, A. J., Lipska, B. K., Weickert, C. S., Hyde, T. M., Straub, R. E., Hashimoto, R., Harrison, P. J., et al. (2006). Neuregulin 1 transcripts are differentially expressed in schizophrenia and regulated by 5' SNPs associated with the disease. Proc Natl Acad Sci U S A 103(17): 6747-52.

Lawrie, S. M., Whalley, H. C., Abukmeil, S. S., Kestelman, J. N., Donnelly, L., Miller, P., Best, J. J., et al. (2001). Brain structure, genetic liability, and psychotic symptoms in subjects at high risk of developing schizophrenia. Biol Psychiatry 49(10): 811-23.

Lee, K. H., Williams, L. M., Breakspear, M. & Gordon, E. (2003). Synchronous gamma activity: a review and contribution to an integrative neuroscience model of schizophrenia. Brain Res Brain Res Rev 41(1): 57-78.

Lein, E. S.Hawrylycz, M. J.Ao, N.Ayres, M.Bensinger, A.Bernard, A.Boe, A. F., et al. (2006). Genome-wide atlas of gene expression in the adult mouse brain. Nature 445(7124): 168-76.

Lettice, L. A., Heaney, S. J., Purdie, L. A., Li, L., de Beer, P., Oostra, B. A., Goode, D., et al. (2003). A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. Hum Mol Genet 12(14): 1725-35.

Lettice, L. A., Horikoshi, T., Heaney, S. J., van Baren, M. J., van der Linde, H. C., Breedveld, G. J., Joosse, M., et al. (2002). Disruption of a long-range cis-acting regulator for Shh causes preaxial polydactyly. Proc Natl Acad Sci U S A 99(11): 7548-53.

Leung, L. S. (1998). Generation of theta and gamma rhythms in the hippocampus. Neurosci Biobehav Rev 22(2): 275-90.

Levine, M. & Tjian, R. (2003). Transcription regulation and animal diversity. Nature 424(6945): 147-51.

Levinson, D. F., Levinson, M. D., Segurado, R. & Lewis, C. M. (2003). Genome scan meta-analysis of schizophrenia and bipolar disorder, part I: Methods and power analysis. Am J Hum Genet 73(1): 17-33.

Lewis, C. M., Levinson, D. F., Wise, L. H., DeLisi, L. E., Straub, R. E., Hovatta, I., Williams, N. M., et al. (2003). Genome scan meta-analysis of schizophrenia and bipolar disorder, part II: Schizophrenia. Am J Hum Genet 73(1): 34-48.

Lewontin, R. C. (1964a). The Interaction of Selection and Linkage. I. General Considerations; Heterotic Models. Genetics 49: 49-67.

Lewontin, R. C. (1964b). The Interaction of Selection and Linkage. Ii. Optimum Models. Genetics 50: 757-82.

Lewontin, R. C. (1988). On measures of gametic disequilibrium. Genetics 120(3): 849-52.

Li, T., Ma, X., Sham, P. C., Sun, X., Hu, X., Wang, Q., Meng, H., et al. (2004a). Evidence for association between novel polymorphisms in the PRODH gene and schizophrenia in a Chinese population. Am J Med Genet B Neuropsychiatr Genet 129(1): 13-5.

Li, T., Stefansson, H., Gudfinnsson, E., Cai, G., Liu, X., Murray, R. M., Steinthorsdottir, V., et al. (2004b). Identification of a novel neuregulin 1 at-risk haplotype in Han schizophrenia Chinese patients, but no association with the Icelandic/Scottish risk haplotype. Mol Psychiatry 9(7): 698-704.

Liddell, H. G. & Scott, R. (1980). Greek-English Lexicon. Oxford, Oxford University Press.

Liddle, P. F. (1987). The symptoms of chronic schizophrenia. A re-examination of the positive-negative dichotomy. Br J Psychiatry 151: 145-51.

Liddle, P. F., Friston, K. J., Frith, C. D., Hirsch, S. R., Jones, T. & Frackowiak, R. S. (1992). Patterns of cerebral blood flow in schizophrenia. Br J Psychiatry 160: 179-86.

Lieberman, J. A., Stroup, T. S., McEvoy, J. P., Swartz, M. S., Rosenheck, R. A., Perkins, D. O., Keefe, R. S., et al. (2005). Effectiveness of antipsychotic drugs in patients with chronic schizophrenia. N Engl J Med 353(12): 1209-23.

Lifshitz, Y., Lindzen, M., Garty, H. & Karlish, S. J. (2006). Functional interactions of phospholemman (PLM) (FXYD1) with Na+,K+-ATPase. Purification of alpha1/beta1/PLM complexes expressed in Pichia pastoris. J Biol Chem 281(23): 15790-9.

Lindzen, M., Aizman, R., Lifshitz, Y., Lubarski, I., Karlish, S. J. & Garty, H. (2003). Structure-function relations of interactions between Na,K-ATPase, the gamma subunit, and corticosteroid hormone-induced factor. J Biol Chem 278(21): 18738-43.

Liou, Y. J., Lai, I. C., Wang, Y. C., Bai, Y. M., Lin, C. C., Lin, C. Y., Chen, T. T., et al. (2006). Genetic analysis of the human ENTH (Epsin 4) gene and schizophrenia. Schizophr Res 84(2-3): 236-243.

Liquori, C. L., Ricker, K., Moseley, M. L., Jacobsen, J. F., Kress, W., Naylor, S. L., Day, J. W., et al. (2001). Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of ZNF9. Science 293(5531): 864-7.

Litt, M. & Luty, J. A. (1989). A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. Am J Hum Genet 44(3): 397-401.

Liu, H., Heath, S. C., Sobin, C., Roos, J. L., Galke, B. L., Blundell, M. L., Lenane, M., et al. (2002). Genetic variation at the 22q11 PRODH2/DGCR6 locus presents an unusual pattern and increases susceptibility to schizophrenia. Proc Natl Acad Sci U S A 99(6): 3717-22.

Lockhart, D. J., Dong, H., Byrne, M. C., Follettie, M. T., Gallo, M. V., Chee, M. S., Mittmann, M., et al. (1996). Expression monitoring by hybridization to high-density oligonucleotide arrays. Nat Biotechnol 14(13): 1675-80.

Lohmueller, K. E., Pearce, C. L., Pike, M., Lander, E. S. & Hirschhorn, J. N. (2003). Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. Nat Genet 33(2): 177-82.

Lubarski, I., Pihakaski-Maunsbach, K., Karlish, S. J., Maunsbach, A. B. & Garty, H. (2005). Interaction with the Na,K-ATPase and tissue distribution of FXYD5 (related to ion channel). J Biol Chem 280(45): 37717-24.

Luthi, A. & McCormick, D. A. (1998). H-current: properties of a neuronal and network pacemaker. Neuron 21(1): 9-12.

Magee, J. C. (1998). Dendritic hyperpolarization-activated currents modify the integrative properties of hippocampal CA1 pyramidal neurons. J Neurosci 18(19): 7613-24.

Mah, S., Nelson, M. R., Delisi, L. E., Reneland, R., Markward, N., James, M. R., Nyholt, D. R., et al. (2006). Identification of the semaphorin receptor PLXNA2 as a candidate for susceptibility to schizophrenia. Mol Psychiatry 11(5): 471-8.

Mahmmoud, Y. A., Cramb, G., Maunsbach, A. B., Cutler, C. P., Meischke, L. & Cornelius, F. (2003). Regulation of Na,K-ATPase by PLMS, the phospholemman-like protein from shark: molecular cloning, sequence, expression, cellular distribution, and functional effects of PLMS. J Biol Chem 278(39): 37427-38.

Mann, E. O., Radcliffe, C. A. & Paulsen, O. (2005). Hippocampal gamma-frequency oscillations: from interneurones to pyramidal cells, and back. J Physiol 562(Pt 1): 55-63.

Marassi, F. M. & Crowell, K. J. (2003). Hydration-optimized oriented phospholipid bilayer samples for solid-state NMR structural studies of membrane proteins. J Magn Reson 161(1): 64-9.

Matthysse, S., Lange, K. & Wagener, D. K. (1979). Continuous variation caused by genes with graduated effects. Proc Natl Acad Sci U S A 76(6): 2862-5.

Matthysse, S. W. & Kidd, K. K. (1976). Estimating the genetic contribution to schizophrenia. Am J Psychiatry 133(2): 185-91.

Maziade, M., Raymond, V., Cliche, D., Fournier, J. P., Caron, C., Garneau, Y., Nicole, L., et al. (1995). Linkage results on 11Q21-22 in Eastern Quebec pedigrees densely affected by schizophrenia. Am J Med Genet 60(6): 522-8.

McGlashan, T. H. & Hoffman, R. E. (2000). Schizophrenia as a disorder of developmentally reduced synaptic connectivity. Arch Gen Psychiatry 57(7): 637-48.

McGue, M., Gottesman, II & Rao, D. C. (1985). Resolving genetic models for the transmission of schizophrenia. Genet Epidemiol 2(1): 99-110.

McGuffin, P., Farmer, A. E., Gottesman, II, Murray, R. M. & Reveley, A. M. (1984). Twin concordance for operationally defined schizophrenia. Confirmation of familiality and heritability. Arch Gen Psychiatry 41(6): 541-5.

Meij, I. C., Koenderink, J. B., van Bokhoven, H., Assink, K. F., Groenestege, W. T., de Pont, J. J., Bindels, R. J., et al. (2000). Dominant isolated renal magnesium loss is caused by misrouting of the Na(+),K(+)-ATPase gamma-subunit. Nat Genet 26(3): 265-6.

Meltzer, H. Y. (1997). Treatment-resistant schizophrenia--the role of clozapine. Curr Med Res Opin 14(1): 1-20.

Mercer, R. W., Biemesderfer, D., Bliss, D. P., Jr., Collins, J. H. & Forbush, B., 3rd (1993). Molecular cloning and immunological characterization of the gamma polypeptide, a small protein associated with the Na,K-ATPase. J Cell Biol 121(3): 579-86.

Middendorf, L. R., Bruce, J. C., Bruce, R. C., Eckles, R. D., Grone, D. L., Roemer, S. C., Sloniker, G. D., et al. (1992). Continuous, on-line DNA sequencing using a versatile infrared laser scanner/electrophoresis apparatus. Electrophoresis 13(1): 487-494.

Millar, J. K., Wilson-Annan, J. C., Anderson, S., Christie, S., Taylor, M. S., Semple, C. A., Devon, R. S., et al. (2000). Disruption of two novel genes by a translocation co-segregating with schizophrenia. Hum Mol Genet 9: 1415-1423.

Minor, N. T., Sha, Q., Nichols, C. G. & Mercer, R. W. (1998). The gamma subunit of the Na,K-ATPase induces cation channel activity. Proc Natl Acad Sci U S A 95(11): 6521-5.

Mirnics, K., Middleton, F. A., Stanwood, G. D., Lewis, D. A. & Levitt, P. (2001). Disease-specific changes in regulator of G-protein signaling 4 (RGS4) expression in schizophrenia. Mol Psychiatry 6(3): 293-301.

Mirza, M. A., Zhang, X. Q., Ahlers, B. A., Qureshi, A., Carl, L. L., Song, J., Tucker, A. L., et al. (2004). Effects of phospholemman downregulation on contractility and [Ca(2+)]i transients in adult rat cardiac myocytes. Am J Physiol Heart Circ Physiol 286(4): H1322-30.

Moises, H. W., Yang, L., Kristbjarnarson, H., Wiese, C., Byerley, W., Macciardi, F., Arolt, V., et al. (1995). An international two-stage genome-wide search for schizophrenia susceptibility genes. Nat Genet 11(3): 321-4.

Moorman, J. R., Ackerman, S. J., Kowdley, G. C., Griffin, M. P., Mounsey, J. P., Chen, Z., Cala, S. E., et al. (1995). Unitary anion currents through phospholemman channel molecules. Nature 377(6551): 737-40.

Moorman, J. R. & Jones, L. R. (1998). Phospholemman: a cardiac taurine channel involved in regulation of cell volume. Adv Exp Med Biol 442: 219-28.

Moorman, J. R., Palmer, C. J., John, J. E., 3rd, Durieux, M. E. & Jones, L. R. (1992). Phospholemman expression induces a hyperpolarization-activated chloride current in Xenopus oocytes. J Biol Chem 267(21): 14551-4.

Morales-Mulia, M., Pasantes-Morales, H. & Moran, J. (2000). Volume sensitive efflux of taurine in HEK293 cells overexpressing phospholemman. Biochim Biophys Acta 1496(2-3): 252-60.

Moran, J., Morales-Mulia, M. & Pasantes-Morales, H. (2001). Reduction of phospholemman expression decreases osmosensitive taurine efflux in astrocytes. Biochim Biophys Acta 1538(2-3): 313-20.

Morris, D. W., Rodgers, A., McGhee, K. A., Schwaiger, S., Scully, P., Quinn, J., Meagher, D., et al. (2004). Confirming RGS4 as a susceptibility gene for schizophrenia. Am J Med Genet 125B(1): 50-3.

Morrison, B. W. & Leder, P. (1994). neu and ras initiate murine mammary tumors that share genetic markers generally absent in c-myc and int-2-initiated tumors. Oncogene 9(12): 3417-26.

Morrison, B. W., Moorman, J. R., Kowdley, G. C., Kobayashi, Y. M., Jones, L. R. & Leder, P. (1995). Mat-8, a novel phospholemman-like protein expressed in human breast tumors, induces a chloride conductance in Xenopus oocytes. J Biol Chem 270(5): 2176-82.

Morton, N. E. (1955). Sequential tests for the detection of linkage. Am J Hum Genet 7(3): 277-318.

Mulcrone, J., Whatley, S. A., Marchbanks, R., Wildenauer, D., Altmark, D., Daoud, H., Gur, E., et al. (1995). Genetic linkage analysis of schizophrenia using chromosome 11q13-24 markers in Israeli pedigrees. Am J Med Genet 60(2): 103-8.

Munafo, M. R., Bowes, L., Clark, T. G. & Flint, J. (2005). Lack of association of the COMT (Val158/108 Met) gene and schizophrenia: a meta-analysis of case-control studies. Mol Psychiatry 10(8): 765-70.

Munafo, M. R., Thiselton, D. L., Clark, T. G. & Flint, J. (2006). Association of the NRG1 gene and schizophrenia: a meta-analysis. Mol Psychiatry 11(6): 613.

Mutsuddi, M., Morris, D.W., Waggoner, S.G., Daly, M.J., Scolnick, E.M. & Sklar, P. (2006). Analysis of high-resolution HapMap of DTNBP1 (Dysbindin) suggests no consistency between reported common variant associations to schizophrenia. Am J Hum Genet 79(5): 903-9.

Myakishev, M. V., Khripin, Y., Hu, S. & Hamer, D. H. (2001). High-throughput SNP genotyping by allele-specific PCR with universal energy-transfer-labeled primers. Genome Res 11(1): 163-9.

Nakamura, Y., Leppert, M., O'Connell, P., Wolff, R., Holm, T., Culver, M., Martin, C., et al. (1987). Variable number of tandem repeat (VNTR) markers for human gene mapping. Science 235(4796): 1616-22.

Nanko, S., Gill, M., Owen, M., Takazawa, N., Moridaira, J. & Kazamatsuri, H. (1992). Linkage study of schizophrenia with markers on chromosome 11 in two Japanese pedigrees. Jpn J Psychiatry Neurol 46(1): 155-9.

Nelson, M. D., Saykin, A. J., Flashman, L. A. & Riordan, H. J. (1998). Hippocampal volume reduction in schizophrenia as assessed by magnetic resonance imaging: a meta-analytic study. Arch Gen Psychiatry 55(5): 433-40.

Neves-Pereira, M., Cheung, J. K., Pasdar, A., Zhang, F., Breen, G., Yates, P., Sinclair, M., et al. (2005). BDNF gene is a risk factor for schizophrenia in a Scottish population. Mol Psychiatry 10(2): 208-12.

Numakawa, T., Yagasaki, Y., Ishimoto, T., Okada, T., Suzuki, T., Iwata, N., Ozaki, N., et al. (2004). Evidence of novel neuronal functions of dysbindin, a susceptibility gene for schizophrenia. Hum Mol Genet 13(21): 2699-708.

Ott, J. (1991). Analysis of Human Genetic Linkage. Baltimore, John Hopkins University Press.

Otting, N. & Bontrop, R. E. (1995). Evolution of the major histocompatibility complex DPA1 locus in primates. Hum Immunol 42(2): 184-7.

Owen, M. J., O'Donovan, M. C. & Harrison, P. J. (2005). Schizophrenia: a genetic disorder of the synapse? BMJ 330(7484): 158-9.

Palmer, C. J., Scott, B. T. & Jones, L. R. (1991). Purification and complete sequence determination of the major plasma membrane substrate for cAMP-dependent protein kinase and protein kinase C in myocardium. J Biol Chem 266(17): 11126-30.

Palmer, L. G. (1999). Potassium secretion and the regulation of distal nephron K channels. Am J Physiol 277(6 Pt 2): F821-5.

Pape, H. C. (1996). Queer current and pacemaker: the hyperpolarization-activated cation current in neurons. Annu Rev Physiol 58: 299-327.

Parker, J. C. & Hoffman, J. F. (1964). Failure to Find Increased Sodium, Potassium-Atpase in Red Cell Ghosts of Schizophrenics. Nature 201: 823.

Peralta, V. & Cuesta, M. J. (2001). How many and which are the psychopathological dimensions in schizophrenia? Issues influencing their ascertainment. Schizophr Res 49(3): 269-85.

Petryshen, T., Kirby, A. & Ainscow, M. (2003). LocusView2.0 Copyright (c) 2003, 2004 - The Broad Institute, unpublished software.

Petryshen, T. L., Middleton, F. A., Kirby, A., Aldinger, K. A., Purcell, S., Tahl, A. R., Morley, C. P., et al. (2005). Support for involvement of neuregulin 1 in schizophrenia pathophysiology. Mol Psychiatry 10(4): 366-74, 328.

Pfeifer, D., Kist, R., Dewar, K., Devon, K., Lander, E. S., Birren, B., Korniszewski, L., et al. (1999). Campomelic dysplasia translocation breakpoints are scattered over 1 Mb proximal to SOX9: evidence for an extended control region. Am J Hum Genet 65(1): 111-24.

Pierri, J. N., Volk, C. L., Auh, S., Sampson, A. & Lewis, D. A. (2001). Decreased somal size of deep layer 3 pyramidal neurons in the prefrontal cortex of subjects with schizophrenia. Arch Gen Psychiatry 58(5): 466-73.

Pihakaski-Maunsbach, K., Tokonabe, S., Vorum, H., Rivard, C. J., Capasso, J. M., Berl, T. & Maunsbach, A. B. (2005). The gamma-subunit of Na-K-ATPase is incorporated into plasma membranes of mouse IMCD3 cells in response to hypertonicity. Am J Physiol Renal Physiol 288(4): F650-7.

Pihakaski-Maunsbach, K., Vorum, H., Locke, E. M., Garty, H., Karlish, S. J. & Maunsbach, A. B. (2003). Immunocytochemical localization of Na,K-ATPase gamma subunit and CHIF in inner medulla of rat kidney. Ann N Y Acad Sci 986: 401-9.

Pimm, J., McQuillin, A., Thirumalai, S., Lawrence, J., Quested, D., Bass, N., Lamb, G., et al. (2005). The Epsin 4 gene on chromosome 5q, which encodes the clathrin-associated protein enthoprotin, is involved in the genetic susceptibility to schizophrenia. Am J Hum Genet 76(5): 902-7.

Pritchard, J. K. & Przeworski, M. (2001). Linkage disequilibrium in humans: models and data. Am J Hum Genet 69(1): 1-14.

Pu, H. X., Cluzeaud, F., Goldshleger, R., Karlish, S. J., Farman, N. & Blostein, R. (2001). Functional role and immunocytochemical localization of the gamma a and gamma b forms of the Na,K-ATPase gamma subunit. J Biol Chem 276(23): 20370-8.

Pulver, A. E., Nestadt, G., Goldberg, R., Shprintzen, R. J., Lamacz, M., Wolyniec, P. S., Morrow, B., et al. (1994). Psychotic illness in patients diagnosed with velo-cardio-facial syndrome and their relatives. J Nerv Ment Dis 182(8): 476-8.

Purcell, S., Cherny, S. S. & Sham P. C. (2003). Genetic power calculator: design of linkage and association genetic mapping studies of complex traits. Bioinformatics 19(1): 149-50.

Puri, V., McQuillin, A., Choudhury, K., Datta, S., Pimm, J., Thirumalai, S., Krasucki, R., et al. (2007). Fine mapping by genetic association implicates the chromosome 1q23.3 gene UHMK1, encoding a serine/threonine protein kinase, as a novel schizophrenia susceptibility gene. Biol Psychiatry 61(7): 873-9.

Puri, V., McQuillin, A., Thirumalai, S., Lawrence, J., Krasucki, R., Choudhury, K., Datta, S., et al. (2006). Failure to confirm allelic association between markers at the CAPON gene locus and schizophrenia in a British sample. Biol Psychiatry 59(2): 195-7.

Racoosin, E. L. & Swanson, J. A. (1994). Labeling of endocytic vesicles using fluorescent probes for fluid-phase endocytosis. Cell Biology: A Laboratory Handbook. J. E. Celis. New York, Academic Press: 375-380.

Rao, D. C., Morton, N. E., Gottesman, II. & Lew, R. (1981). Path analysis of qualitative data on pairs of relatives: application to schizophrenia. Hum Hered 31(6): 325-33.

Redon, R., Ishikawa, S., Fitch, K. R., Feuk, L., Perry, G. H., Andrews, T. D., Fiegler, H., et al. (2006). Global variation in copy number in the human genome. Nature 444(7118): 444-54.

Reich, D. E. & Lander, E. S. (2001). On the allelic spectrum of human disease. Trends Genet 17(9): 502-10.

Risch, N. (1990a). Genetic linkage and complex diseases, with special reference to psychiatric disorders. Genet Epidemiol 7(1): 3-16.

Risch, N. (1990b). Linkage strategies for genetically complex traits. I. Multilocus models. Am J Hum Genet 46(2): 222-8.

Risch, N. (1990c). Linkage strategies for genetically complex traits. II. The power of affected relative pairs. Am J Hum Genet 46(2): 229-41.

Risch, N. (1990d). Linkage strategies for genetically complex traits. III. The effect of marker polymorphism on analysis of affected relative pairs. Am J Hum Genet 46(2): 242-53.

Risch, N. & Baron, M. (1984). Segregation analysis of schizophrenia and related disorders. Am J Hum Genet 36(5): 1039-59.

Rivolta, C., Sharon, D., DeAngelis, M. M. & Dryja, T. P. (2002). Retinitis pigmentosa and allied diseases: numerous diseases, genes, and inheritance patterns. Hum Mol Genet 11(10): 1219-27.

Rizig, M. A., McQuillin, A., Puri, V., Choudhury, K., Datta, S., Thirumalai, S., Lawrence, J., et al. (2006). Failure to confirm genetic association between schizophrenia and markers on chromosome 1q23.3 in the region of the gene encoding the Regulator of G-protein Signalling 4 protein (RGS4). Am J Med Genet B Neuropsychiatr Genet 141(3): 296-300.

Rosenthal, D., Wender, P. H., Kety, S. S., Welner, J. & Schulsinger, F. (1971). The adopted-away offspring of schizophrenics. Am J Psychiatry 128(3): 307-11.

Rovelet-Lecrux, A., Hannequin, D., Raux, G., Le Meur, N., Laquerriere, A., Vital, A., Dumanchin, C., et al. (2006). APP locus duplication causes autosomal dominant early-onset Alzheimer disease with cerebral amyloid angiopathy. Nat Genet 38(1): 24-6.

Royston, M. C. & Simpson, M. D. C. (1991). Post-mortem neurochemistry of schizophrenia. Neurobiology and Psychiatry. R. Kerwin, D. Dawbarn, J. McCulloch & C. Tamminga. Cambridge, Cambridge University Press: 1-14.

Rozen, S. & Skaletsky, H. (2000). Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol 132: 365-86.

Rybakowski, J. K. & Lehmann, W. (1994). Decreased activity of erythrocyte membrane ATPases in depression and schizophrenia. Neuropsychobiology 30(1): 11-4.

Sadock, B. J. & Sadock, V. A. (2003). Kaplan & Sadock's synopsis of psychiatry: behavioral sciences/clinical psychiatry. Philadelphia, Lippincott, Williams & Wilkins.

Saito, Y. A., Talley, N. J., de Andrade, M. & Petersen, G. M. (2006). Case-control genetic association studies in gastrointestinal disease: review and recommendations. Am J Gastroenterol 101(6): 1379-89.

Sandberg, R., Yasuda, R., Pankratz, D. G., Carter, T. A., Del Rio, J. A., Wodicka, L., Mayford, M., et al. (2000). Regional and strain-specific gene expression mapping in the adult mouse brain. Proc Natl Acad Sci U S A 97(20): 11038-43.

Schindler, K. M., Pato, M. T., Dourado, A., Macedo, A., Azevedo, M. H., Kennedy, J. L. & Pato, C. N. (2002). Association and linkage disequilibrium between a functional polymorphism of the dopamine-2 receptor gene and schizophrenia in a genetically homogeneous Portuguese population. Mol Psychiatry 7(9): 1002-5.

Schmajuk, N. A. (2001). Hippocampal dysfunction in schizophrenia. Hippocampus 11(5): 599-613.

Schmidt, S. (2006). Advanced Parametric Linkage Analysis. Genetic Analysis of Complex Disease. J. L. Haines, Pericak-Vance, M.A., Wiley. 2: 255-281.

Schneider, K. (1959). Clinical Psychopathology. New York, Grune and Stratton.

Schumacher, J., Jamra, R. A., Freudenberg, J., Becker, T., Ohlraun, S., Otte, A. C., Tullius, M., et al. (2004). Examination of G72 and D-amino-acid oxidase as genetic risk factors for schizophrenia and bipolar affective disorder. Mol Psychiatry 9(2): 203-7.

Schwab, S. G., Eckstein, G. N., Hallmayer, J., Lerer, B., Albus, M., Borrmann, M., Lichtermann, D., et al. (1997). Evidence suggestive of a locus on chromosome 5q31 contributing to susceptibility for schizophrenia in German and Israeli families by multipoint affected sib-pair linkage analysis. Mol Psychiatry 2(2): 156-60.

Schwab, S. G., Knapp, M., Mondabon, S., Hallmayer, J., Borrmann-Hassenbach, M., Albus, M., Lerer, B., et al. (2003). Support for association of schizophrenia with genetic variation in the 6p22.3 gene, dysbindin, in sib-pair families with linkage and in an additional sample of triad families. Am J Hum Genet 72(1): 185-90.

Scott, M. P., Matsudaira, P., Lodish, H., Darnell, J., Zipursky, L., Kaiser, C. A., Berk, A., et al. (2004). Molecular Cell Biology. San Francisco, W. H. Freeman.

Seeman, P. M. & O'Brien, E. (1963). Sodium-Potassium-Activated Adenosine Triphosphatase in Schizophrenic Erythrocytes. Nature 200: 263-4.

Sha, Q., Lansbery, K. L., Distefano, D., Mercer, R. W. & Nichols, C. G. (2001). Heterologous expression of the Na(+),K(+)-ATPase gamma subunit in Xenopus oocytes induces an endogenous, voltage-gated large diameter pore. J Physiol 535(Pt 2): 407-17.

Sham, P. C. & Curtis, D. (1995). Monte Carlo tests for associations between disease and alleles at highly polymorphic loci. Ann Hum Genet 59(Pt 1): 97-105.

Sharp, A. J., Locke, D. P., McGrath, S. D., Cheng, Z., Bailey, J. A., Vallente, R. U., Pertz, L. M., et al. (2005). Segmental duplications and copy-number variation in the human genome. Am J Hum Genet 77(1): 78-88.

Shenton, M. E., Kikinis, R., Jolesz, F. A., Pollak, S. D., LeMay, M., Wible, C. G., Hokama, H., et al. (1992). Abnormalities of the left temporal lobe and thought disorder in schizophrenia. A quantitative magnetic resonance imaging study. N Engl J Med 327(9): 604-12.

Sherman, A. D., Davidson, A. T., Baruah, S., Hegwood, T. S. & Waziri, R. (1991). Evidence of glutamatergic deficiency in schizophrenia. Neurosci Lett 121(1-2): 77-80.

Shi, H., Levy-Holzman, R., Cluzeaud, F., Farman, N. & Garty, H. (2001). Membrane topology and immunolocalization of CHIF in kidney and intestine. Am J Physiol Renal Physiol 280(3): F505-12.

Shifman, S., Bronstein, M., Sternfeld, M., Pisante-Shalom, A., Lev-Lehman, E., Weizman, A., Reznik, I., et al. (2002). A highly significant association between a COMT haplotype and schizophrenia. Am J Hum Genet 71(6): 1296-302.

Shigeri, Y., Seal, R. P. & Shimamoto, K. (2004). Molecular pharmacology of glutamate transporters, EAATs and VGLUTs. Brain Res Brain Res Rev 45(3): 250-65.

Silverman, B. Z., Fuller, W., Eaton, P., Deng, J., Moorman, J. R., Cheung, J. Y., James, A. F., et al. (2005). Serine 68 phosphorylation of phospholemman: acute isoform-specific activation of cardiac Na/K ATPase. Cardiovasc Res 65(1): 93-103.

Sims, A. (2002). Symptoms in the mind: an introduction to descriptive psychopathology. Edinburgh, Elsevier Science Ltd.

Singleton, A. B., Farrer, M., Johnson, J., Singleton, A., Hague, S., Kachergus, J., Hulihan, M., et al. (2003). alpha-Synuclein locus triplication causes Parkinson's disease. Science 302(5646): 841.

Slater, E. & Cowie, V. (1971). The Genetics of Mental Disorders. London, Oxford University Press,.

Sobell, J. L., Richard, C., Wirshing, D. A. & Heston, L. L. (2005). Failure to confirm association between RGS4 haplotypes and schizophrenia in Caucasians. Am J Med Genet B Neuropsychiatr Genet 139(1): 23-7.

Song, J., Zhang, X. Q., Ahlers, B. A., Carl, L. L., Wang, J., Rothblum, L. I., Stahl, R. C., et al. (2005). Serine 68 of phospholemman is critical in modulation of contractility, [Ca2+]i transients, and Na+/Ca2+ exchange in adult rat cardiac myocytes. Am J Physiol Heart Circ Physiol 288(5): H2342-54.

Song, J., Zhang, X. Q., Carl, L. L., Qureshi, A., Rothblum, L. I. & Cheung, J. Y. (2002). Overexpression of phospholemman alters contractility and [Ca(2+)](i) transients in adult rat myocytes. Am J Physiol Heart Circ Physiol 283(2): H576-83.

Spitzer, R. L. & Endicott, J. (1977). The Schedule for Affective Disorders and Schizophrenia, Lifetime Version, 3rd edition. NY, New York State Psychiatric Institute.

Spitzer, R. L., Endicott, J. & Robins, E. (1978). Research Diagnostic Criteria for a selected group of functional disorders, 3rd edition. New York, New York State Psychiatric Institute.

Sribney, W. M. & Swift, M. (1992). Power of sib-pair and sib-trio linkage analysis with assortative mating and multiple disease loci. Am J Hum Genet 51(4): 773-84.

St Clair, D., Blackwood, D., Muir, W., Carothers, A., Walker, M., Spowart, G., Gosden, C., et al. (1990). Association within a family of a balanced autosomal translocation with major mental illness. Lancet 336: 13-16.

Stefansson, H., Sarginson, J., Kong, A., Yates, P., Steinthorsdottir, V., Gudfinnsson, E., Gunnarsdottir, S., et al. (2003). Association of neuregulin 1 with schizophrenia confirmed in a Scottish population. Am J Hum Genet 72(1): 83-7.

Stefansson, H., Sigurdsson, E., Steinthorsdottir, V., Bjornsdottir, S., Sigmundsson, T., Ghosh, S., Brynjolfsson, J., et al. (2002). Neuregulin 1 and susceptibility to schizophrenia. Am J Hum Genet 71(4): 877-92.

Straub, R. E., Jiang, Y., MacLean, C. J., Ma, Y., Webb, B. T., Myakishev, M. V., Harris-Kerr, C., et al. (2002). Genetic variation in the 6p22.3 gene DTNBP1, the human ortholog of the mouse dysbindin gene, is associated with schizophrenia. Am J Hum Genet 71(2): 337-48.

Straub, R. E., MacLean, C. J., O'Neill, F. A., Walsh, D. & Kendler, K. S. (1997). Support for a possible schizophrenia vulnerability locus in region 5q22-31 in Irish families. Mol Psychiatry 2(2): 148-55.

Su, A. I., Cooke, M. P., Ching, K. A., Hakak, Y., Walker, J. R., Wiltshire, T., Orth, A. P., et al. (2002). Large-scale analysis of the human and mouse transcriptomes. Proc Natl Acad Sci U S A 99(7): 4465-70.

Sullivan, P. F., Kendler, K. S. & Neale, M. C. (2003). Schizophrenia as a complex trait: evidence from a meta-analysis of twin studies. Arch Gen Psychiatry 60(12): 1187-92.

Sweadner, K. J., Arystarkhova, E., Donnet, C. & Wetzel, R. K. (2003). FXYD proteins as regulators of the Na,K-ATPase in the kidney. Ann N Y Acad Sci 986: 382-7.

Sweadner, K. J. & Rael, E. (2000). The FXYD gene family of small ion transport regulators or channels: cDNA sequence, protein signature sequence, and expression. Genomics 68(1): 41-56.

Sweadner, K. J., Wetzel, R. K. & Arystarkhova, E. (2000). Genomic organization of the human FXYD2 gene encoding the gamma subunit of the Na,K-ATPase. Biochem Biophys Res Commun 279(1): 196-201.

Talkowski, M. E., Seltman, H., Bassett, A. S., Brzustowicz, L. M., Chen, X., Chowdari, K., Collier, D. A., et al. (2006). Evaluation of a Susceptibility Gene for Schizophrenia: Genotype Based Meta-Analysis of RGS4 Polymorphisms from Thirteen Independent Samples. Biol Psychiatry 60(2): 152-62.

Tang, R. Q., Zhao, X. Z., Shi, Y. Y., Tang, W., Gu, N. F., Feng, G. Y., Xing, Y. L., et al. (2006). Family-based association study of Epsin 4 and Schizophrenia. Mol Psychiatry 11(4): 395-9.

Therien, A. G. & Blostein, R. (2000). Mechanisms of sodium pump regulation. Am J Physiol Cell Physiol 279(3): C541-66.

Therien, A. G., Goldshleger, R., Karlish, S. J. & Blostein, R. (1997). Tissue-specific distribution and modulatory role of the gamma subunit of the Na,K-ATPase. J Biol Chem 272(51): 32628-34.

Therien, A. G., Karlish, S. J. & Blostein, R. (1999). Expression and functional role of the gamma subunit of the Na, K-ATPase in mammalian cells. J Biol Chem 274(18): 12252-6.

Therien, A. G., Pu, H. X., Karlish, S. J. & Blostein, R. (2001). Molecular and functional studies of the gamma subunit of the sodium pump. J Bioenerg Biomembr 33(5): 407-14.

Thiselton, D. L., Webb, B. T., Neale, B. M., Ribble, R. C., O'Neill, F. A., Walsh, D., Riley, B. P., et al. (2004). No evidence for linkage or association of neuregulin-1 (NRG1) with disease in the Irish study of high-density schizophrenia families (ISHDSF). Mol Psychiatry 9(8): 777-83.

Thompson, P. M., Vidal, C., Giedd, J. N., Gochman, P., Blumenthal, J., Nicolson, R., Toga, A. W., et al. (2001). Mapping adolescent brain change reveals dynamic wave of accelerated gray matter loss in very early-onset schizophrenia. Proc Natl Acad Sci U S A 98(20): 11650-5.

Tienari, P., Wynne, L. C., Moring, J., Laksy, K., Nieminen, P., Sorri, A., Lahti, I., et al. (2000). Finnish adoptive family study: sample selection and adoptee DSM-III-R diagnoses. Acta Psychiatr Scand 101(6): 433-43.

Tienari, P., Wynne, L. C., Sorri, A., Lahti, I., Laksy, K., Moring, J., Naarala, M., et al. (2004). Genotype-environment interaction in schizophrenia-spectrum disorder. Long-term follow-up study of Finnish adoptees. Br J Psychiatry 184: 216-22.

Tochigi, M., Otowa, T., Suga, M., Rogers, M., Minato, T., Yamasue, H., Kasai, K., et al. (2006). No evidence for an association between the BDNF Val66Met polymorphism and schizophrenia or personality traits. Schizophr Res 87(1-3): 45-7.

Tsuang, M. T., Bucher, K. D. & Fleming, J. A. (1983). A search for 'schizophrenia spectrum disorders'. An application of a multiple threshold model to blind family study data. Br J Psychiatry 143: 572-7.

Turner, T. (2007). Chlorpromazine: unlocking psychosis. Bmj 334 Suppl 1: s7.

Vieland, V., Greenberg, D. A., Hodge, S. E. & Ott, J. (1992a). Linkage analysis of two-locus diseases under single-locus and two-locus analysis models. Cytogenet Cell Genet 59(2-3): 145-6.

Vieland, V. J., Hodge, S. E. & Greenberg, D. A. (1992b). Adequacy of single-locus approximations for linkage analyses of oligogenic traits. Genet Epidemiol 9(1): 45-59.

Vieland, V. J. & Huang, J. (2003). Two-locus heterogeneity cannot be distinguished from two-locus epistasis on the basis of affected-sib-pair data. Am J Hum Genet 73(2): 223-32.

Viukari, M. (1972). Epilepsy, schizophrenia, and sodium-potassium-A.T.P.ase. Lancet 1(7753): 749.

Wald, H., Goldstein, O., Asher, C., Yagil, Y. & Garty, H. (1996). Aldosterone induction and epithelial distribution of CHIF. Am J Physiol 271(2 Pt 2): F322-9.

Wald, H., Popovtzer, M. M. & Garty, H. (1997). Differential regulation of CHIF mRNA by potassium intake and aldosterone. Am J Physiol 272(5 Pt 2): F617-23.

Wang, D. G., Fan, J. B., Siao, C. J., Berno, A., Young, P., Sapolsky, R., Ghandour, G., et al. (1998). Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. Science 280(5366): 1077-82.

Wang, Z. W., Black, D., Andreasen, N. C. & Crowe, R. R. (1993). A linkage study of chromosome 11q in schizophrenia. Arch Gen Psychiatry 50(3): 212-6.

Watanabe, Y., Muratake, T., Kaneko, N., Nunokawa, A. & Someya, T. (2006). No association between the brain-derived neurotrophic factor gene and schizophrenia in a Japanese population. Schizophr Res 84(1): 29-35.

Weber, J. L. & May, P. E. (1989). Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am J Hum Genet 44(3): 388-96.

Wei, J. & Hemmings, G. P. (2004). Lack of a genetic association between the frizzled-3 gene and schizophrenia in a British population. Neurosci Lett 366(3): 336-8.

Weinberg, W. (1908). On the demonstration of heredity in man. Württem 64: 368-382.

Weinberger, D. R. (1995). From neuropathology to neurodevelopment. Lancet 346(8974): 552-7.

Wender, P. H., Rosenthal, D., Kety, S.S., Schulsinger, F. & Welner, J. (1974). Crossfostering. A research strategy for clarifying the role of genetic and experiential factors in the etiology of schizophrenia. Arch Gen Psychiatry 30: 121-128.

Wetzel, R. K. & Sweadner, K. J. (2001). Immunocytochemical localization of Na-K-ATPase alpha- and gamma-subunits in rat kidney. Am J Physiol Renal Physiol 281(3): F531-45.

Wetzel, R. K. & Sweadner, K. J. (2003). Phospholemman expression in extraglomerular mesangium and afferent arteriole of the juxtaglomerular apparatus. Am J Physiol Renal Physiol 285(1): F121-9.

WHO (1992). The ICD-10 classification of mental and behavioural disorders: clinical descriptions and dignostic guidelines. Geneva, World Health Organisation.

Williams, H. J., Williams, N., Spurlock, G., Norton, N., Ivanov, D., McCreadie, R. G., Preece, A., et al. (2003a). Association between PRODH and schizophrenia is not confirmed. Mol Psychiatry 8(7): 644-5.

Williams, N. M., Preece, A., Morris, D. W., Spurlock, G., Bray, N. J., Stephens, M., Norton, N., et al. (2004a). Identification in 2 independent samples of a novel schizophrenia risk haplotype of the dystrobrevin binding protein gene (DTNBP1). Arch Gen Psychiatry 61(4): 336-44.

Williams, N. M., Preece, A., Spurlock, G., Norton, N., Williams, H. J., McCreadie, R. G., Buckland, P., et al. (2004b). Support for RGS4 as a susceptibility gene for schizophrenia. Biol Psychiatry 55(2): 192-5.

Williams, N. M., Preece, A., Spurlock, G., Norton, N., Williams, H. J., Zammit, S., O'Donovan, M. C., et al. (2003b). Support for genetic variation in neuregulin 1 and susceptibility to schizophrenia. Mol Psychiatry 8(5): 485-7.

Winokur, G., Tsuang, M. T. & Crowe, R. R. (1982). The Iowa 500: affective disorder in relatives of manic and depressed patients. Am J Psychiatry 139(2): 209-12.

Witte, J. S., Elston, R. C. & Schork, N. J. (1996). Genetic dissection of complex traits. Nat Genet 12(4): 355-6.

Wong, E. H., Kemp, J. A., Priestley, T., Knight, A. R., Woodruff, G. N. & Iversen, L. L. (1986). The anticonvulsant MK-801 is a potent N-methyl-D-aspartate antagonist. Proc Natl Acad Sci U S A 83(18): 7104-8.

Wright, A., Charlesworth, B., Rudan, I., Carothers, A. & Campbell, H. (2003). A polygenic basis for late-onset disease. Trends Genet 19(2): 97-106.

Wright, I. C., Rabe-Hesketh, S., Woodruff, P. W., David, A. S., Murray, R. M. & Bullmore, E. T. (2000). Meta-analysis of regional brain volumes in schizophrenia. Am J Psychiatry 157(1): 16-25.

Xiong, Z. & Laird, P. W. (1997). COBRA: a sensitive and quantitative DNA methylation assay. Nucleic Acids Res 25(12): 2532-4.

Yamaguchi, F., Yamaguchi, K., Tai, Y., Sugimoto, K. & Tokuda, M. (2001). Molecular cloning and characterization of a novel phospholemman-like protein from rat hippocampus. Brain Res Mol Brain Res 86(1-2): 189-92.

Yang, J., Si, T., Ling, Y., Ruan, Y., Han, Y., Wang, X., Zhang, H., et al. (2003a). Association study of the human FZD3 locus with schizophrenia. Biol Psychiatry 54(11): 1298-301.

Yang, J. Z., Si, T. M., Ruan, Y., Ling, Y. S., Han, Y. H., Wang, X. L., Zhou, M., et al. (2003b). Association study of neuregulin 1 gene with schizophrenia. Mol Psychiatry 8(7): 706-9.

Zaidel, D. W., Esiri, M. M. & Harrison, P. J. (1997). Size, shape, and orientation of neurons in the left and right hippocampus: investigation of normal asymmetries and alterations in schizophrenia. Am J Psychiatry 154(6): 812-8.

Zhang, J., Feuk, L., Duggan, G. E., Khaja, R. & Scherer, S. W. (2006). Development of bioinformatics resources for display and analysis of copy number and other structural variants in the human genome. Cytogenet Genome Res 115(3-4): 205-14.

Zhang, X. Q., Qureshi, A., Song, J., Carl, L. L., Tian, Q., Stahl, R. C., Carey, D. J., et al. (2003). Phospholemman modulates Na+/Ca2+ exchange in adult rat cardiac myocytes. Am J Physiol Heart Circ Physiol 284(1): H225-33.

Zhao, J. H., Curtis, D. & Sham, P. C. (2000). Model-free analysis and permutation tests for allelic associations. Hum Hered 50(2): 133-9.

Zhao, J. H., Lissarrague, S., Essioux, L. & Sham, P. C. (2002). GENECOUNTING: haplotype analysis with missing genotypes. Bioinformatics 18(12): 1694-5.

Zhao, X., Shi, Y., Tang, J., Tang, R., Yu, L., Gu, N., Feng, G., et al. (2004). A case control and family based association study of the neuregulin1 gene and schizophrenia. J Med Genet 41(1): 31-4.

Zouzoulas, A., Dunham, P. B. & Blostein, R. (2005). The effect of the gamma modulator on Na/K pump activity of intact mammalian cells. J Membr Biol 204(1): 49-56.

Zouzoulas, A., Therien, A. G., Scanzano, R., Deber, C. M. & Blostein, R. (2003). Modulation of Na,K-ATPase by the gamma subunit: studies with transfected cells and transmembrane mimetic peptides. J Biol Chem 278(42): 40437-41.

11.0 APPENDIX

11.1 PAIRWISE LD STATISTICS BETWEEN MARKERS GENOYPED IN THE CHROMOSOME 11q23.3 LOCUS OF THE ORIGINAL UCL GENETIC ASSOCIATION STUDY (SECTION 4.0)

Figure 11.1: Cramer's V and P value pairwise statistics.

Chromosome 11 Marker	Marker Location (bp)	Prior Marker Distance (bp)		Cramer's V																																		
				1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33
rs11999	117196128		1		0.283	0.438	0.125	0.006	0.177	0.209	0.149	0.083	0.046	0.384	0.208	0.201	0.087	0.246	0.009	0.093	0.017	0.165	0.002	0.013	0.021	0.007	0.006	0.009	0.007	0.006	0.065	0.080	0.005	0.013	0.120	0.013
rs869789	117196596	468	2	0.00000		0.251	0.430	0.252	0.211	0.159	0.115	0.290	0.219	0.197	0.019	0.035	0.099	0.097	0.111	0.208	0.302	0.172	0.055	0.040	0.040	0.011	0.017	0.037	0.088	0.103	0.026	0.084	0.055	0.039	0.156	0.040
rs4579662	117197544	948	3	0.00000	0.00000		0.236	0.044	0.332	0.379	0.378	0.495	0.051	0.264	0.006	0.228	0.073	0.079	0.154	0.296	0.086	0.172	0.042	0.050	0.083	0.096	0.104	0.034	0.043	0.035	0.061	0.062	0.047	0.028	0.103	0.041
rs529623	117198465	921	4	0.00031	0.00000	0.00000		0.325	0.495	0.345	0.319	0.377	0.430	0.190	0.042	0.079	0.302	0.168	0.041	0.329	0.198	0.202	0.015	0.027	0.011	0.002	0.000	0.001	0.008	0.000	0.038	0.035	0.002	0.008	0.089	0.019
rs878776	117201954	3489	5	0.80640	0.00000	0.16809	0.00000		0.110	0.229	0.183	0.488	0.836	0.268	0.236	0.224	0.142	0.040	0.162	0.236	0.166	0.282	0.095	0.112	0.079	0.216	0.195	0.108	0.135	0.143	0.139	0.072	0.230	0.013	0.013	0.084
rs479991	117202168	214	6	0.00000	0.00000	0.00000	0.00000	0.00033		0.286	0.161	0.383	0.283	0.104	0.201	0.085	0.173	0.314	0.217	0.264	0.103	0.185	0.064	0.057	0.052	0.099	0.094	0.039	0.055	0.057	0.034	0.037	0.046	0.000	0.053	0.026
D11S1998	117202941	773	7	0.00000	0.00224	0.00000	0.00000	0.00000	0.00000		0.977	0.416	0.192	0.183	0.202	0.074	0.143	0.153	0.107	0.131	0.198	0.167	0.150	0.147	0.117	0.171	0.160	0.126	0.109	0.112	0.026	0.152	0.083	0.065	0.162	0.120
rs11216567	117203693	752	8	0.00000	0.00070	0.00000	0.00000	0.00000	0.00000	0.00000		0.395	0.146	0.113	0.181	0.042	0.093	0.069	0.118	0.055	0.083	0.169	0.029	0.030	0.113	0.088	0.087	0.025	0.023	0.024	0.006	0.120	0.013	0.040	0.032	0.071
rs10892181	117204796	1103	9	0.01549	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000		0.423	0.299	0.177	0.145	0.167	0.169	0.039	0.111	0.087	0.233	0.026	0.034	0.117	0.139	0.121	0.015	0.021	0.017	0.193	0.023	0.164	0.039	0.116	0.056
rs2019655	117205088	292	10	0.16809	0.00000	0.11828	0.00000	0.00000	0.00000	0.00000	0.00008	0.00002	0.00000		0.228	0.396	0.256	0.088	0.055	0.093	0.321	0.117	0.256	0.037	0.036	0.056	0.077	0.084	0.026	0.024	0.031	0.101	0.008	0.162	0.054	0.087	0.094
rs10790212	117207900	2812	11	0.00000	0.00000	0.00000	0.00000	0.00000	0.00180	0.00070	0.00073	0.00000	0.00000			0.431	0.146	0.084	0.358	0.220	0.172	0.123	0.324	0.124	0.134	0.046	0.094	0.093	0.115	0.183	0.182	0.105	0.059	0.179	0.024	0.081	0.076
rs3087563	117213147	5247	12	0.00000	0.55978	0.80640	0.19762	0.00000	0.00000	0.00001	0.00000	0.00000	0.00000	0.00000			0.226	0.242	0.249	0.333	0.214	0.235	0.243	0.038	0.038	0.210	0.216	0.213	0.044	0.026	0.025	0.008	0.148	0.095	0.055	0.107	0.136
rs3168238	117214855	1708	13	0.00000	0.30781	0.00000	0.01414	0.00000	0.00282	0.51095	0.10959	0.00000	0.00000	0.00000	0.00000			0.146	0.063	0.087	0.253	0.076	0.193	0.059	0.050	0.158	0.094	0.100	0.055	0.096	0.108	0.229	0.168	0.151	0.035	0.033	0.117
rs564989	117214964	109	14	0.01531	0.00705	0.02133	0.00000	0.00001	0.00000	0.00678	0.00329	0.00000	0.00872	0.04604	0.00000	0.00000			0.423	0.143	0.515	0.509	0.177	0.001	0.015	0.118	0.132	0.141	0.002	0.009	0.000	0.087	0.119	0.007	0.062	0.050	0.077
rs12363888	117218397	3433	15	0.00000	0.00054	0.00552	0.00000	0.19226	0.00000	0.00654	0.01697	0.00000	0.09082	0.00000	0.00000	0.00000	0.01514	0.00000		0.064	0.228	0.071	0.235	0.086	0.083	0.054	0.090	0.082	0.099	0.050	0.055	0.061	0.001	0.182	0.010	0.034	0.094
rs11608153	117221873	3476	16	0.77719	0.00147	0.00000	0.20589	0.00000	0.00000	0.06273	0.00001	0.22694	0.00667	0.00000	0.00000	0.00000	0.00806	0.00001	0.00021			0.363	0.054	0.298	0.239	0.244	0.129	0.382	0.406	0.245	0.313	0.308	0.067	0.166	0.102	0.044	0.002	0.075
rs512481	117224090	2217	17	0.00932	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.04564	0.07363	0.00055	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000			0.264	0.282	0.226	0.235	0.144	0.148	0.161	0.232	0.206	0.194	0.035	0.009	0.080	0.027	0.123	0.009
rs476130	117233093	9003	18	0.59684	0.00000	0.00843	0.00000	0.00000	0.00072	0.00000	0.05187	0.01055	0.00015	0.00003	0.00000	0.00188	0.00000	0.00103	0.10822	0.00000			0.233	0.178	0.173	0.080	0.134	0.137	0.179	0.140	0.154	0.117	0.047	0.165	0.096	0.034	0.213
TTTC20.2	117234550	1457	19	0.19392	0.13082	0.07451	0.02695	0.00000	0.04729	0.72935	0.22538	0.00079	0.00003	0.00000	0.00003	0.01036	0.11325	0.00017	0.00000	0.00000	0.00027		0.958	0.961	0.577	0.733	0.736	0.936	0.798	0.805	0.231	0.218	0.511	0.177	0.191	0.278
rs7121573	117235040	490	20	0.88748	0.09426	0.17971	0.59668	0.00215	0.04550	0.00534	0.33229	0.38966	0.24706	0.00018	0.24008	0.06441	0.88817	0.00419	0.00000	0.00000	0.00000	0.00000		0.961	0.477	0.861	0.671	0.951	0.779	0.788	0.041	0.042	0.477	0.060	0.011	0.158
rs1815774	117236649	1609	21	0.65476	0.20875	0.11099	0.38966	0.00045	0.07013	0.00466	0.32220	0.26932	0.24008	0.00003	0.22374	0.12289	0.59668	0.00697	0.00000	0.00000	0.00000	0.00000	0.00000		0.496	0.693	0.686	0.960	0.794	0.806	0.033	0.041	0.489	0.047	0.018	0.159
rs876798	117242650	6001	22	0.51698	0.21166	0.06440	0.70839	0.01292	0.09194	0.18848	0.00247	0.00034	0.07547	0.16587	0.00000	0.00001	0.00021	0.10687	0.00005	0.00001	0.00371	0.00000	0.00000	0.00000		0.193	0.194	0.488	0.463	0.469	0.081	0.064	0.246	0.049	0.038	0.105
rs876797	117242957	307	23	0.80640	0.70839	0.00428	1.00000	0.00000	0.00212	0.00175	0.00953	0.00006	0.02207	0.00410	0.00000	0.00315	0.00007	0.00746	0.00000	0.00002	0.00002	0.00000	0.00000	0.00000	0.00000		0.982	0.701	0.543	0.545	0.111	0.148	0.344	0.027	0.023	0.115
rs873713	117245126	2169	24	0.80640	0.58385	0.00157	1.00000	0.00000	0.00398	0.00473	0.00645	0.00030	0.01447	0.00410	0.00000	0.00114	0.00001	0.01352	0.00000	0.00000	0.00001	0.00000	0.00000	0.00000	0.00000	0.00000		0.703	0.555	0.556	0.095	0.145	0.328	0.027	0.008	0.103
rs10790218	117247205	2079	25	0.72891	0.24702	0.26547	1.00000	0.00050	0.19490	0.03802	0.40961	0.61011	0.39615	0.00034	0.17032	0.08123	1.00000	0.00129	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000		0.821	0.831	0.028	0.071	0.483	0.063	0.005	0.153
rs4938445	117250213	3008	26	0.77719	0.00806	0.16584	0.75171	0.00003	0.06928	0.13586	0.43854	0.47950	0.43100	0.00000	0.40961	0.00448	0.75201	0.11530	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000		0.992	0.002	0.018	0.411	0.029	0.025	0.149
rs4938446	117250259	46	27	0.80640	0.00184	0.24352	1.00000	0.00001	0.05848	0.11534	0.41659	0.54845	0.31251	0.00000	0.42366	0.00109	1.00000	0.08641	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000		0.002	0.017	0.428	0.041	0.025	0.150
rs11216568	117253662	3403	28	0.06137	0.42366	0.05778	0.23015	0.00004	0.27732	0.98842	0.80640	0.00000	0.00285	0.00140	0.80640	0.00000	0.01463	0.09628	0.04443	0.27737	0.00091	0.00016	0.20308	0.29424	0.00729	0.00094	0.00367	0.35937	0.88817	1.00000		0.362	0.417	0.073	0.317	0.240
rs831868	117253688	26	29	0.02108	0.05377	0.05249	0.26547	0.02084	0.23338	0.00750	0.00032	0.44638	0.77719	0.07099	0.00001	0.00000	0.00031	0.88748	0.00000	0.72917	0.13448	0.00223	0.18462	0.18465	0.05005	0.00000	0.00001	0.02448	0.54858	0.55978	0.00000		0.126	0.061	0.166	0.136
rs11605223	117254054	366	30	0.84140	0.10034	0.14349	1.00000	0.00000	0.15123	0.70526	0.65456	0.00000	0.00000	0.00000	0.00392	0.00000	0.77719	0.00000	0.00147	0.01263	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00011		0.172	0.151	0.588
rs3809043	117254325	271	31	0.65476	0.26637	0.34817	0.75171	0.65476	1.00000	0.60476	0.18670	0.20878	0.11530	0.44633	0.10687	0.21461	0.01292	0.72904	0.16366	0.36516	0.00808	0.11719	0.04443	0.13105	0.13800	0.39615	0.37713	0.04340	0.32220	0.17261	0.02504	0.04659	0.00000		0.030	0.205
rs3809042	117254330	5	32	0.00001	0.00005	0.00045	0.05187	0.65456	0.07186	0.04564	0.28565	0.00617	0.01277	0.02913	0.00035	0.24704	0.07013	0.20591	0.88748	0.00010	0.23672	0.01866	0.66924	0.52714	0.23013	0.41655	0.72904	0.88748	0.40280	0.38333	0.00000	0.00000	0.00000	0.16367		0.191
rs497768	117255950	1620	33	0.65456	0.22378	0.20034	0.54845	0.01128	0.40280	0.11393	0.03197	0.08430	0.00453	0.02084	0.00007	0.00030	0.02132	0.00279	0.02285	0.75171	0.00000	0.00000	0.00000	0.00000	0.00234	0.00057	0.00127	0.00000	0.00001	0.00000	0.00000	0.00008	0.00000	0.00000	0.00000	

P value

Figure 11.2: Absolute value of D' and LOD for LD pairwise statistics.

Absolute value of D'

Chromosome 11 Marker	Marker Location (bp)	Prior Marker Distance (bp)	LOD for LD	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33
rs11909	117196128		1		0.971	0.508	0.185	0.014	0.245	0.050	0.673	0.085	0.110	0.462	0.272	0.552	0.113	0.731	0.030	0.120	0.079	0.053	0.004	0.024	0.038	0.017	0.016	0.010	0.007	0.007	0.105	0.110	0.007	0.099	0.998	0.021
rs869780	117196596	468	2	19.5		0.999	1.000	0.966	0.999	0.291	0.947	0.968	0.957	0.814	0.045	0.053	0.137	0.989	0.715	0.552	0.414	0.018	0.175	0.131	0.249	0.051	0.078	0.120	0.326	0.378	0.030	0.159	0.138	0.083	0.379	0.086
rs4579962	117197544	948	3	36.3	15.4		0.404	0.107	1.000	0.190	0.779	0.589	0.056	0.275	0.010	0.539	0.209	0.519	0.631	0.442	0.469	0.073	0.085	0.104	0.098	0.285	0.305	0.069	0.102	0.082	0.086	0.099	0.076	0.240	0.992	0.076
rs529623	117198465	921	4	2.8	48.7	11.6		0.536	0.874	0.148	0.976	0.541	0.807	0.339	0.043	0.319	0.506	0.639	0.099	0.376	0.631	0.021	0.018	0.033	0.029	0.004	0.000	0.001	0.011	0.000	0.078	0.038	0.002	0.043	0.390	0.020
rs678776	117201954	3489	5	0.0	16.9	0.4	22.6		0.321	0.118	0.920	1.000	0.951	0.687	0.381	0.552	0.391	0.251	0.273	0.338	0.879	0.022	0.115	0.131	0.302	0.262	0.234	0.129	0.139	0.149	0.201	0.111	0.351	0.045	0.116	0.148
rs479991	117202168	214	6	5.6	11.6	33.8	60.8	2.8		0.268	0.999	0.974	0.937	0.118	0.362	0.610	0.182	0.675	0.296	0.535	0.669	0.010	0.095	0.083	0.245	0.101	0.097	0.058	0.070	0.074	0.040	0.069	0.087	0.002	0.611	0.057
D11S1998	117202941	773	7	8.8	5.2	27.7	28.0	10.3	19.0		1.000	0.039	0.078	0.144	0.046	0.120	0.091	0.286	0.061	0.138	0.630	0.210	0.161	0.168	0.112	0.202	0.171	0.137	0.114	0.110	0.015	0.010	0.060	0.052	0.221	0.094
rs11216567	117203693	752	8	4.9	2.5	29.2	23.5	7.4	8.5	158.7		0.970	0.838	0.615	0.623	0.516	0.169	0.925	0.996	0.167	0.710	0.020	0.075	0.076	0.149	0.538	0.525	0.063	0.051	0.055	0.042	0.392	0.041	0.702	0.628	0.271
rs10892181	117204796	1103	9	1.3	18.1	54.1	28.1	70.8	40.5	36.2	36.6		0.988	0.372	0.226	0.733	0.404	0.925	0.134	0.139	0.399	0.199	0.045	0.059	0.218	0.347	0.298	0.025	0.043	0.033	0.575	0.031	0.220	0.280	0.937	0.065
rs2019655	117205088	292	10	0.4	13.2	0.5	45.1	190.0	22.7	7.2	3.9	46.4		0.662	0.726	0.553	0.276	0.393	0.420	0.525	0.702	0.094	0.051	0.048	0.241	0.082	0.088	0.035	0.029	0.037	0.129	0.014	0.282	0.165	0.920	0.190
rs10790212	117207900	2812	11	25.1	9.6	13.0	7.1	15.4	2.1	5.9	2.5	17.4	11.5		0.682	0.924	0.070	0.878	0.951	0.268	0.706	0.227	0.262	0.292	0.190	0.291	0.284	0.248	0.451	0.444	0.390	0.097	0.300	0.077	0.819	0.109
rs3087563	117213147	5247	12	7.4	0.1	0.0	0.4	10.8	8.2	8.3	7.0	6.0	35.5	40.5		0.899	0.457	0.967	0.818	0.240	0.852	0.133	0.046	0.047	0.550	0.383	0.373	0.054	0.041	0.039	0.015	0.156	0.101	0.311	0.613	0.152
rs3168238	117214855	1708	13	7.7	0.2	10.5	1.3	10.6	1.9	1.6	0.6	5.3	12.2	6.8	13.1		0.997	0.980	0.849	0.895	0.993	0.038	0.177	0.144	0.241	0.666	0.697	0.161	0.244	0.279	0.388	0.635	0.567	0.724	0.748	0.513
rs564989	117214964	109	14	1.3	1.6	1.2	17.2	4.4	5.2	4.6	1.9	5.9	1.5	0.9	11.4	5.5		0.968	0.668	0.993	0.968	0.240	0.003	0.021	0.162	0.444	0.466	0.004	0.011	0.000	0.350	0.215	0.013	0.245	0.545	0.120
rs12363888	117218397	3433	15	11.4	2.6	1.7	5.8	0.4	19.8	4.6	1.2	7.1	0.6	25.8	16.1	1.3	30.2		0.994	1.000	0.997	0.031	0.278	0.261	0.089	0.198	0.183	0.315	0.138	0.153	0.112	0.006	0.743	0.213	0.854	0.448
rs11608153	117221873	3476	16	0.0	2.2	5.2	0.3	5.1	8.2	3.2	4.3	0.3	1.6	12.0	23.3	1.5	4.2	3.0		1.000	0.484	0.730	0.487	0.484	0.828	0.533	0.570	0.490	0.546	0.543	0.078	0.428	0.262	0.615	0.034	0.167
rs512481	117224090	2217	17	1.5	8.9	18.7	21.8	11.2	13.9	3.4	0.7	2.6	20.6	5.6	9.0	16.4	79.5	14.6	30.7		0.970	0.207	0.305	0.327	0.382	0.294	0.315	0.319	0.327	0.305	0.073	0.010	0.086	0.158	0.793	0.011
rs476130	117233093	9003	18	0.1	14.7	1.5	7.4	7.5	2.5	9.2	0.8	1.4	3.1	3.7	12.9	2.1	58.5	2.3	0.6	18.6		0.010	0.774	0.777	0.686	0.854	0.868	0.797	0.715	0.781	0.900	0.161	0.571	0.150	0.710	0.633
TTTC20.2	117234550	1457	19	5.5	5.9	6.5	7.4	14.1	6.9	32.3	5.3	10.0	12.1	23.3	12.1	8.1	6.0	11.0	18.0	15.3	10.7		0.962	0.963	0.927	0.949	0.974	0.935	0.898	0.894	0.118	0.184	0.256	0.188	1.000	0.125
rs7121573	117235040	490	20	0.0	0.6	0.4	0.1	2.0	0.9	4.7	0.2	0.2	0.3	3.1	0.3	0.7	0.0	1.8	12.6	10.8	8.2	301.6		0.992	0.930	0.990	0.972	0.973	0.902	0.908	0.072	0.053	0.600	0.412	0.053	0.231
rs1815774	117236649	1609	21	0.0	0.3	0.6	0.2	2.7	0.7	4.8	0.2	0.3	0.3	3.8	0.3	0.5	0.1	1.6	12.4	11.4	7.1	305.6	342.5		0.949	0.986	0.967	0.967	0.909	0.915	0.056	0.053	0.631	0.327	0.083	0.241
rs876798	117242650	6001	22	0.1	0.3	0.7	0.0	1.3	0.6	2.4	2.0	2.8	0.7	0.4	8.3	4.2	3.0	0.6	3.6	4.1	1.8	73.4	57.6	63.5		0.895	0.886	0.934	0.776	0.791	0.451	0.158	0.608	0.650	0.091	0.303
rs876797	117242957	307	23	0.0	0.0	1.8	0.0	9.0	2.0	5.4	1.5	3.5	1.1	1.8	9.2	1.9	3.4	1.6	27.3	3.9	4.0	139.0	131.4	136.0	6.4		0.993	1.000	0.675	0.686	0.133	0.275	0.637	0.270	0.260	0.248
rs873713	117245126	2169	24	0.0	0.1	2.2	0.0	7.9	1.8	4.8	1.6	2.8	1.3	1.8	8.9	2.3	4.1	1.3	33.3	5.0	4.5	140.9	133.8	131.2	6.2	282.1		0.996	0.683	0.694	0.116	0.267	0.598	0.263	0.091	0.220
rs10790218	117247205	2079	25	0.0	0.3	0.3	0.0	2.6	0.4	3.5	0.1	0.1	0.2	2.8	0.4	0.7	0.0	2.2	13.7	11.9	8.1	285.3	326.1	319.2	61.6	145.8	156.8		0.937	0.942	0.048	0.091	0.620	0.439	0.022	0.228
rs4938445	117250213	3008	26	0.0	1.5	0.4	0.0	3.8	0.7	2.7	0.1	0.1	0.1	7.4	0.1	1.8	0.0	0.5	20.4	9.3	4.7	176.7	170.6	180.9	47.2	66.5	70.7	200.4		0.995	0.003	0.026	0.607	0.235	0.220	0.256
rs4938446	117250259	46	27	0.0	2.1	0.3	0.0	4.3	0.8	2.8	0.1	0.1	0.2	7.2	0.1	2.3	0.0	0.6	19.9	8.2	6.0	185.5	179.6	190.7	50.6	68.3	71.8	210.7	349.8		0.002	0.024	0.625	0.326	0.225	0.271
rs11216598	117253662	3403	28	0.8	0.1	0.8	0.3	3.7	0.3	0.4	0.0	6.9	1.9	2.2	0.0	8.8	1.3	0.7	0.9	0.3	2.4	11.1	0.4	0.2	1.6	2.4	1.8	0.2	0.0	0.0		0.810	0.925	0.876	0.855	0.618
rs631896	117253688	26	29	1.2	0.8	0.8	0.3	1.2	0.3	4.5	2.8	0.1	0.0	0.7	4.5	6.4	2.8	0.0	5.6	0.0	0.5	9.3	0.4	0.4	0.8	4.6	4.3	1.1	0.1	0.1	31.3		0.129	0.331	0.999	0.158
rs11605223	117254054	366	30	0.0	0.6	0.5	0.0	11.3	0.4	1.2	0.0	5.4	5.4	6.1	1.8	5.6	0.0	7.2	2.2	1.4	6.5	61.3	54.9	55.3	13.1	25.7	25.2	56.4	38.5	42.9	41.4	3.2		0.931	0.906	0.687
rs3809043	117254325	271	31	0.0	0.3	0.2	0.0	0.0	0.0	1.4	0.4	0.3	0.5	0.1	0.6	0.3	1.3	0.0	0.4	0.2	1.5	6.0	0.9	0.5	0.5	0.2	0.2	0.9	0.2	0.4	1.1	0.9	7.1		0.984	0.952
rs3809042	117254330	5	32	4.3	3.6	2.7	0.8	0.0	0.7	3.4	0.2	1.6	1.3	1.0	2.8	0.3	0.7	0.3	0.0	3.3	0.3	7.7	0.0	0.1	0.3	0.1	0.0	0.0	0.2	0.2	17.7	6.1	5.0	0.4		1.000
rs497768	117255950	1620	33	0.0	0.3	0.4	0.1	1.4	0.2	2.8	1.0	0.6	1.8	1.2	3.4	2.8	1.2	1.9	1.1	0.0	10.7	15.8	5.2	5.0	2.0	2.6	2.3	4.9	4.2	5.0	11.3	3.5	86.0	10.6	8.7	