Analysis of oligodendrocyte and myelination related genes in schizophrenia.

A thesis submitted in fulfilment of the
requirements for the Degree of Doctor of
Philosophy at Cardiff University
by

Timothy Rowan Peirce

2006

Department of Psychological Medicine
School of Medicine
Cardiff University
Heath Park
Cardiff CF14 4XN

Summary.

Schizophrenia is a severe psychiatric disorder with a lifetime risk of about 1%. It affects almost all domains of mental function, including perception, emotion, cognition and social performance. Evidence from family, twin and adoption studies has shown the disorder to have a high genetic component.

The aetiology of schizophrenia remains obscure but over the last decade there has been increasing evidence that abnormal oligodendrocyte function and/or myelination might play a part. The main strands of evidence that provide support for this hypothesis have been compiled from neuropathological studies, brain imaging, symptom overlap with demyelinating diseases, age related changes in myelin and evidence from micro-array studies. In this thesis, I investigated several genes of importance to oligodendrocyte function and/or myelination as candidate genes for influencing susceptibility to schizophrenia. Genes were primarily selected for study based upon their reports of altered expression in micro-array studies of post-mortem schizophrenic brain. In many cases, their potential involvement was supported by positional evidence from linkage studies and a putative functional role in the neurobiology of schizophrenia.

In total thirteen genes were studied: CNP, NOGO, NgR, OMG, NGFR, GFAP, MOG, SOX 10 and APOL 1-5. During my studies I detected, modest evidence for association with three genes (CNP, NOGO and MOG). I also showed that the expression of these genes is under the influence of cis-acting polymorphisms. Whilst, the results for these genes are modest and require independent replication, they do provide some support for the general hypothesis that oligodendrocyte and/or myelination abnormalities play a role in schizophrenia aetiology.

Declaration and statements.

DECLARATION

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed ....(candidate)

Date........21/ 12/ 2006...

STATEMENT 1

This thesis is the result of my own investigations, except where otherwise stated. Other sources are acknowledged by explicit references. A bibliography is appended.

Signed ....(candidate)

Date........21/ 12/ 2006...

STATEMENT 2

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed ....(candidate)

Date........21/ 12/ 2006...

Acknowledgements.

I owe a debt of gratitude to many people. I wish to thank most sincerely my supervisor, Professor Michael O'Donovan for providing the opportunity to undertake this research and for his invaluable advice and support throughout the course of my PhD. Without his guidance and endless patience the completion of this thesis would have been a much more arduous and less enjoyable experience.

I would also like to thank Dr Nick Bray for all of his help and advice, particularly with my allelic expression assays and during the first year of my studies, together with the members of the Cardiff schizophrenia research team, Prof Michael Owen, Dr Nigel Williams, Dr Nadine Norton, Dr Hywel Williams, Dr George Kirov, Dr Valentina Moskvina, Liam Carroll and Sarah Dwyer. Thanks must also be given to all my friends and colleagues in the laboratory for making time spent in the department such an entertaining and enjoyable experience, especially Elaine, Darko, Natalie, Dobril, Didi, Ivan, Lucy, Denise, Richard, Becky, Tom, Fran and Kate.

Finally I must thank my family and friends for their continued support, particularly my parents for their constant encouragement and belief in me throughout my years of study and research at Cardiff University.

Abbreviations.

A	Adenine
aa	Amino acid
ACN	Acetonitrile
bp	Base pair
C	Cytosine
cDNA	Complementary deoxyribonucleic acid
CI	Confidence interval
cm	Centimeter
cM	Centimorgan
CNS	Central nervous system
CT	Computerised tomography
D	Aspartic acid
dATP	Deoxyadenosine triphosphate
dGTP	Deoxyguanosine triphosphate
dCTP	Deoxycytidine triphosphate
dTTP	Deoxythymidine triphosphate
ddATP	Dideoxyadenosine triphosphate
ddGTP	Dideoxyguanosine triphosphate
ddCTP	Dideoxycytidine triphosphate
ddTTP	Dideoxythymidine triphosphate
dNTP	Deoxynucleotide triphosphate
ddNTP	Dideoxynucleotide triphosphate
Del	Deletion
df	Degrees of freedom
dH_2O	Distilled water
DHPLC	Denaturing high performance liquid chromatography
DLPFC	Dorsolateral prefrontal cortex
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
ds	Double stranded
DSM	Diagnostic and statistical manual

DTI	Diffusion tensor imaging
DZ	Dizygotic
Exo I	Exonuclease I
F	Forward
fMRI	Functional magnetic resonance imaging
FP	Fluorescence polarisation
G	Glycine
g	Grams
G	Guanine
GRR	Genotype relative risk
H	Histidine
HRR	Haplotype relative risk
HW	Hardy Weinberg
IBD	Identity by descent
ICD	International classification of diseases
Ins	Insertion
kb	Kilobase
kDa	Kilodaltons
L	Leucine
LD	Linkage disequilibrium
LOD	Logarithm of odds
M	Molar
MAF	Minor allele frequency
MALDI-TOF	Matrix assisted laser desorption ionisation time of flight
Mb	Megabase
mg	Milligram
mins	Minutes
ml	Millilitre
MLD	Metachromatic leukodystrophy
mM	Millimolar
MRC	Medical research council
MRI	Magnetic resonance imaging
mRNA	Messenger ribonucleic acid
MS	Multiple sclerosis

MTI	Magnetic transfer imaging
MZ	Monozygotic
ng	Nanogram
nm	Nanometres
nmol	Nanomolar
°C	Degrees centigrade
OMR	Oligodendrocyte and myelin related
OR	Odds ratio
P	Proline
PCR	Polymerase chain reaction
PET	Positron emission tomography
pmol	Picomolar
PNS	Peripheral nervous system
R	Arginine
R	Reverse
RNA	Ribonucleic acid
rpm	Rotations per minute
RT-PCR	Reverse transcription polymerase chain reaction
S	Serine
SAP	Shrimp alkaline phosphatise
SD	Standard deviation
secs	Seconds
SNP	Single nucleotide polymorphism
ss	Single stranded
T	Thymine
TDT	Transmission disequilibrium test
TEAA	Triethylammonium acetate
u	Units
UTR	Untranslated region
UV	Ultraviolet
V	Valine
V	Volts
VCFS	Velo-cardio facial syndrome
VNTR	Variable number tandem repeat

W	Tryptophan
Y	Tyrosine
θ	Recombination fraction
λ_s	Relative risk (siblings)
χ^2	Chi-square test for statistical significance
µg	Microgram
µl	Microlitre
µm	Micrometres

Table of contents.

Chapter 1: General Introduction

Chapter 2: Materials and Methods

Chapter 3.1: Analysis of Oligodendrocyte Related Candidate Genes Differentially Expressed in Schizophrenic Brain

Chapter 3.2: 2', 3'-Cyclic Nucleotide 3'-Phosphodiesterase (CNP) a Candidate Gene for Schizophrenia

Chapter 4: The NOGO Receptor Complex and Schizophrenia

Chapter 5: Myelin Candidate Genes and Schizophrenia

Chapter 6: The APO L Gene Family and Schizophrenia

Chapter 7: Final Discussion

Chapter 1. General Introduction:

1.1 Schizophrenia – General Description:

The concept of 'schizophrenia' and the term itself emerged in the early 1900s, where it was first used by Eugen Bleuler in his book *Dementia Praecox or the Group of Schizophrenias* (Bleuler 1911). His book followed the earlier work of Emil Kraepelin who had previously established a distinction between manic depressive insanity and what he called 'dementia praecox' (Kraepelin 1896). Bleuler, although agreeing with most of Kraepelins core symptoms (e.g hallucinations and delusions), instead saw the disorder as comprising a range of different conditions. The acceptance of this refined definition can be seen by the decline of the term 'dementia praecox' and the rise of the word 'schizophrenia'.

Schizophrenia is a common mental illness with a lifetime prevalence of about 1% (Jablensky 1997; McGrath, Saha et al. 2004). The typical age of disease onset is in adolescence or early adulthood (between 15 and 30 years of age), although in some cases, onset can occur in childhood or old age. The risk is the same for both sexes although the average age of onset for males is 3-4 years earlier than for females (Murray and Van Os 1998). The incidence of disease has been shown to be similar across a range of different countries and cultures in both the developing and developed world (Jablensky, Sartorius et al. 1992). There are however pockets of populations where the risk of schizophrenia is higher, for example the Afro-Caribbean population in the UK (Sharpley, Hutchinson et al. 2001; Cantor-Graae and Selten 2005), the risk being greater for second generation rather than first generation migrants or the general UK population. This argues against a selective migration of people early in the course of disease and indicates a possible gene-environment interaction (Sharpley, Hutchinson et al. 2001; Cantor-Graae and Selten 2005). There is also a higher incidence in urban compared to rural areas, with the risk increasing depending on the duration of urban dwelling as a child (Mortensen, Pedersen et al. 1999). There have also been other environmental factors associated with schizophrenia including: winter and spring birth, pregnancy and birth complications, delayed childhood development and cannabis use (Murray 2002).

Within the spectrum of psychiatric disorders schizophrenia is the most disabling and requires the highest level of support and intervention by the mental health services. For instance schizophrenia has been estimated to account for around 50% of admissions to psychiatric hospitals and schizophrenic patients occupy roughly 25% of psychiatric hospital beds (Mueser and McGurk 2004). The outcome for schizophrenia sufferers is variable even with treatment; some patients do well and appear to make a full recovery, but for most, the outcome is not so good and is typically one of repeated relapses requiring medication for the remainder of their lives and leading to lasting social and occupational impairment (Andreasen 1995).

Schizophrenia is characterised by a diverse group of symptoms affecting almost all domains of mental function, including disturbances of emotion, cognition and social performance. These symptoms vary between patients, creating very different patient profiles. The symptoms can include delusional beliefs, auditory hallucinations, catatonia, disorganised thought and speech patterns, apathy, blunted affect, cognitive impairment and social withdrawal. These symptoms generally lead to impaired functioning in work, school, self-sufficiency, interpersonal relationships and social interaction.

Because the signs and symptoms of schizophrenia are so heterogeneous, efforts have been made to subdivide the disorder into broad categories. The most commonly used approach is to divide the symptoms into groups classed as 'positive' and 'negative'. Positive symptoms present as a distortion or excess of 'normal' function (e.g delusions and hallucinations). Negative symptoms are conceptually a deficit or loss of normal functions (e.g apathy, flattening of affect, poverty of speech and social withdrawal). In general the presence and severity of positive symptoms tend to be episodic over time, whereas negative symptoms are more pervasive fluctuating less over time and are therefore strongly associated with poor psychosocial functioning (Hafner 2003).

The lack of a known pathogenesis or biological marker of disease means that the diagnosis of schizophrenia is based entirely on clinical presentation. This has led to researchers questioning whether the disease could be diagnosed with enough reliability to be useful for genetic research. However, the use of structured and semi-structured interviews for the identification of symptoms and operational diagnostic systems does allow reliable diagnosis. The two major diagnostic systems currently in use are the *Tenth Revision of the International Classification of Diseases* (ICD-10) (WHO 1992), and the *Diagnostic and Statistical Manual, Fourth edition* (DSM-IV) (ASP 1994). Although both systems objectively define symptoms and characteristic impairments in similar ways, there are some important differences. The DSM-IV system requires evidence of social or occupational dysfunction (not included in ICD-10) and a 6-month duration of symptoms (versus a 1 month duration in ICD-10), thereby resulting in a narrower definition of the disease in the DSM-IV system (Andreasen 1995).

1.2 Neurodevelopmental hypothesis of Schizophrenia:

The results of the various and diverse areas of research in schizophrenia, have led to the formulation of numerous theories about the biological origin of schizophrenia. However, over the past two decades schizophrenia has increasingly been viewed as a neurodevelopmental disorder. The 'neurodevelopmental hypothesis' proposes that schizophrenia is the result of abnormalities that occur in the brain during development; that the abnormalities occur before the onset of schizophrenic symptoms and are caused by a combination of genetic and environmental factors (Weinberger 1987).

Support for the hypothesis has been drawn from the imaging studies that have been performed in schizophrenia. The best replicated anatomic abnormality in schizophrenia revealed by different forms of imaging is that of lateral ventricular enlargement. Shenton and colleagues (Shenton, Dickey et al. 2001) reviewed the results of 193 MRI studies from 1988 to 2000 and reported that over 80% of the studies showed significant lateral ventricular enlargement in schizophrenia. The degree of enlargement in schizophrenics has been calculated to be around 30% on average with a slightly greater enlargement in male patients compared to females (Lawrie and Abukmeil 1998). Studies of other ventricle systems in the brain support a more general ventricular enlargement in schizophrenia. Shenton and colleagues (Shenton, Dickey et al. 2001) reported that 73% of the MRI studies they reviewed, that had studied the third ventricle showed enlargement of this region in schizophrenics. It should be noted that the range of ventricular size in schizophrenics, overlaps with the range in the normal population and thus the ventricular size of a particular schizophrenic patient may still fall within this normal range (Pearlson and Marsh 1999).

In addition to ventricular abnormalities, MRI studies have also reported limbic and temporal lobe abnormalities in schizophrenia in the amygdala, hippocampus, parahippocampal gyrus and the superior temporal gyrus (Liddle 2003). There is some evidence of a general volume reduction in the temporal lobe in schizophrenia (Liddle 2003). However, there is also evidence for more pronounced effects on specific regions within the temporal lobe. For instance of the MRI studies reviewed (Shenton, Dickey et al. 2001), 74% showed volume reductions in the medial temporal lobe structures, including the amygdala-hippocampal complex and parahippocampal gyrus.

Some studies have tried to assess the role of genetic and environmental factors in causing the brain abnormalities by studying discordant mono and dizygotic twins. These studies have shown reduced whole brain, parahippocampal and hippocampal volumes in affected twins regardless of the zygosity e.g (Suddath, Christison et al. 1990; Baare, van Oel et al. 2001), indicating the presence of an environmental contribution to these features. However, several studies have reported larger lateral and third ventricles and smaller temporal lobe volumes in the unaffected relatives of schizophrenic patients, compared to healthy matched control subjects from families without a history of schizophrenia (Lawrie, Whalley et al. 1999; Baare, van Oel et al. 2001). This would indicate that the brain abnormalities observed in schizophrenia appear to be determined at least in part by familial, possibly genetic factors.

The findings of ventricular enlargement and temporal lobe reductions may be explained by either a developmental or degenerative hypothesis. However, these brain abnormalities are observed in first episode patients before drug treatment, and before symptoms have been present for a significant length of time, arguing against the brain abnormalities being the result of illness progression or drug treatment and indicating they are in some way related to the disease aetiology itself (Harrison 1999; Wong and Van Tol 2003) rather than being secondary effects of psychosis. Additionally, studies including longitudinal follow ups have shown that there is no correlation between ventricular enlargement and the progression of illness, a finding which would be expected if the pathological process responsible for the enlarged ventricles was degenerative and advanced during the course of the illness (Raz and Raz 1990; Jaskiw, Juliano et al. 1994), providing support for the neurodevelopmental hypothesis of schizophrenia. Although, it should be noted that two recent studies have demonstrated that ventricular enlargement is progressive during the very early stages of schizophrenia (Cahn, Hulshoff Pol et al. 2002; Ho, Andreasen et al. 2003).

A second source of support for the developmental hypothesis comes from neuropathological data. The search for the neuropathology of schizophrenia has been a major but elusive goal of biological psychiatry for over 100 years. However, within the last twenty years a number of more robust structural differences have been observed in the brains of schizophrenics. Post-mortem studies of schizophrenic brain have reported ventricular enlargement, brain weight and volume reductions (Brown, Colter et al. 1986), and decreased temporal lobe volume (Bogerts, Meertz et al. 1985; Bogerts, Falkai et al. 1990), findings which have since been corroborated by more recent imaging studies. The changes have been shown to be present in first episode patients, meaning that the pathology cannot be explained as an artefact of chronic illness or drug treatment (Harrison 1999). The presence of reliable structural brain abnormalities in schizophrenia has encouraged further neuropathological research.

Some of the most reliable and important histological findings have been negative, and have been interpreted to indicate that the pathology of schizophrenia is not one of brain degeneration. First there has been no evidence of lesions, tangles or plaques detected in schizophrenic brain, and secondly there has been no reliable evidence of gliosis being present in schizophrenia, all of which would have indicated the presence of a degenerative process (Harrison 1999). Consequently, in the absence of evidence for a degenerative process, to account for reduction in the size of brain structures, the pathology of schizophrenia has been assumed to be a neurodevelopmental process resulting from alterations in the normal maturation of the brain (Weinberger 1987). Additionally, the lack of gliosis has been used to predict that the developmental deficits in schizophrenia occur in the middle stages of gestation. This is because defects in development during early gestation (neurogenesis) would be expected to lead to major abnormalities of the cerebral cortex (a symptom not present in schizophrenia), whilst the lack of gliosis is interpreted to mean that the abnormalities in development must have occurred before the third trimester of pregnancy (Harrison 1999).

With the pathology of schizophrenia likely to be one of neurodevelopment, recent neuropathological research in schizophrenia has examined the cortical cytoarchitecture of the brain, focusing mainly on the limbic system especially the hippocampus and dorsolateral prefrontal cortex (DLPFC) (Harrison 1999). The findings have included neuronal disarray in the hippocampus of schizophrenics (Kovelman and Scheibel 1984; Conrad, Abebe et al. 1991), irregular neuronal clustering in the entorhinal cortex (Jakob and Beckmann 1986; Arnold, Hyman et al. 1991; Arnold, Ruscheinsky et al. 1997), a loss of neuronal density in the hippocampus and DLPFC of schizophrenics, as well as evidence of disorganised positioning of interstitial neurons in the frontal and temporal cortex of schizophrenic patients (Akbarian, Bunney et al. 1993; Akbarian, Vinuela et al. 1993). All of these findings would implicate an abnormality that occurred in the predicted period of gestation and involved a disruption in neuronal maturation and development. However, the support for the hypothesis is made much weaker by the lack of reliable replications of these findings. All of the findings mentioned above have failed to be replicated in at least one study, and in the case of the irregular neuronal clustering present in schizophrenics at least four studies have failed to replicate this finding (Akil and Lewis 1997; Krimer, Herman et al. 1997).

Evidence for a neurodevelopmental process also comes from the many studies which have looked at the relationship between obstetric complications and schizophrenia. Several studies have reported that schizophrenic patients have an increased frequency of obstetric complications. A recent meta-analysis (Cannon, Jones et al. 2002) reported that three groups of complications were significantly associated with schizophrenia. These were: 1) complications of pregnancy (bleeding, diabetes, pre-eclampsia), 2) abnormal fetal growth and development (low birth weight, reduced head circumference, congenital malformations), and 3) complications of delivery (asphyxia, emergency caesarean section). Based on meta-analysis, the effect of obstetric complications on the risk for schizophrenia seems genuine but small, with the odds ratio of the effect of exposure to obstetric complications estimated to be about 2 (Geddes, Verdoux et al. 1999; Cannon, Jones et al. 2002). However, it can not be assumed that obstetric complications have a causative effect on schizophrenia as they might be caused themselves by some other underlying causal process. For example low birth weight and head size are markers of fetal growth retardation, but any factors which adversely affect the fetal development are likely to retard its growth and could therefore account for the lack of growth (Rapoport, Addington et al. 2005).

There have also been studies reporting a variety of other prenatal factors with increased risk for schizophrenia. These include schizophrenics having about 5% higher frequencies of winter and spring births than controls (Torrey, Miller et al. 1997), leading some to propose that the higher likelihood of maternal infections in winter and spring might affect development. There is also evidence of a higher incidence of schizophrenia with urban birth (Mortensen, Pedersen et al. 1999). This was originally thought to be due to people with schizophrenia moving into more deprived inner city areas as a consequence of their illness, but recent work suggests that urban residence during upbringing independent of place of birth is associated with a higher risk of developing schizophrenia (Pedersen and Mortensen 2001). There has also been some evidence that severe maternal malnutrition during pregnancy is associated with higher rates of schizophrenia (Susser, Neugebauer et al. 1996; Wahlbeck, Forsen et al. 2001; St Clair, Xu et al. 2005).

If the basis of schizophrenia is neurodevelopmental then one might expect organs other than the brain to show defects or for there to be other abnormalities of the nervous system to be present. There is some evidence of an increased incidence of physical abnormalities in schizophrenics, particularly of the craniofacial variety. Lane and colleagues (Lane, Kinsella et al. 1997) reported that by using 12 variables of cranial morphology they were able to correctly classify 90% of patients with schizophrenia and 80% of controls. There is also evidence of developmental delays in children who later go on to develop schizophrenia, including motor, cognitive, social and educational defects (Marenco and Weinberger 2000). For instance, Jones and colleagues (Jones, Rodgers et al. 1994) who studied a cohort of children all born in the same week in 1946, reported that the individuals who went on to develop schizophrenia achieved motor development goals, particularly walking, later than expected. In the same study the children who developed schizophrenia were rated by teachers to be more likely to play alone, be less socially confident and more anxious in social situations. Done and colleagues have replicated this finding (Done, Crow et al. 1994), reporting that the children in their cohort who later developed schizophrenia were more likely to exhibit inappropriate social behaviour. There is also some evidence of lower IQ in people who later develop schizophrenia compared to controls, with the lower the IQ the greater the risk of developing schizophrenia. Evidence for this finding is provided by two studies of army conscripts (David, Malmberg et al. 1997; Davidson, Reichenberg et al. 1999), which both showed that the IQ of conscripts who later developed schizophrenia was lower than that of conscripts who did not. This would indicate that intellectual performance was impaired before the onset of schizophrenic symptoms and could be due to earlier developmental defects.

The different strands of evidence for a neurodevelopmental hypothesis of schizophrenia as discussed above are strong and while much of the evidence can still be seen as circumstantial, controversial or inconclusive, when taken together, does suggest that neurodevelopmental defects play at least a part in the aetiology of schizophrenia. Many modified versions of the hypothesis have been proposed with abnormalities in additional processes such as cell adhesion, synaptic pruning, oligodendrocyte function and myelination also being proposed to play an important role in the aetiology of schizophrenia. The recent interest in the myelination process and a postulated role of oligodendrocytes in schizophrenia forms the basis of this thesis and will be discussed in more detail in section 1.5 of this chapter.

1.2.1 Neurochemical hypotheses of schizophrenia.

In addition to the neurodevelopmental hypothesis of schizophrenia there have also been several neurochemical hypotheses proposed. These originated following the development of antipsychotic medications for schizophrenia, and led to the hypothesis that the brain dysfunction associated with schizophrenia was at least in part caused by a disturbance in brain biochemistry and neurochemical imbalances. These hypotheses were also supported by the observation that psychosis could be induced through the administration of certain drugs, such as amphetamines and LSD. It was assumed that the neurotransmitters influenced by these drugs would reflect the biological abnormalities observed in schizophrenia. This led to extensive study of the various neurotransmitter systems and to the formation of many hypotheses for their involvement in schizophrenia.

One of the first and still one of the most widely held hypotheses is the dopamine hypothesis of schizophrenia, which proposed that the symptoms of schizophrenia were due to dopamine over-activity. This hypothesis was based initially on pharmacological evidence showing that dopamine agonists such as amphetamines, which are known to increase dopamine release, induce or exacerbate psychosis with schizophrenic like symptoms, as well as the fact that the potency of all of the typical antipsychotic drugs varies with their affinity for the D_2 subtype of dopamine receptor (van Rossum 1966; Carlsson 1978).

Another neurotransmitter hypothesised to be implicated in the aetiology of schizophrenia is glutamate. Interest was initially created by reports of reduced glutamate levels in schizophrenic patients (Kim, Kornhuber et al. 1980). The hypoglutamatergic hypothesis of schizophrenia was further supported by the discovery that the administration of antiglutamatergic drugs such as phencyclidine (PCP) and ketamine, which act by blocking the NMDA subtype of glutamate receptors, induced psychotic symptoms characteristic of schizophrenia (Moghaddam 2003). Moreover, the administration of PCP induced not only psychotic positive symptoms like amphetamine but also negative symptoms and cognitive dysfunction in healthy controls, and it therefore may provide a better model of schizophrenia (Reynolds 2002).

A further neurotransmitter that has been implicated with the aetiology of schizophrenia is serotonin. Interest in serotonin was initially generated by the discovery that atypical antipsychotics block certain subtypes of the serotonin receptor. The involvement of serotonin is further supported by the psychotomimetic properties of drugs such as LSD which act upon the serotonin receptor. Subsequent post-mortem studies have failed to find consistent changes in the serotonin system or its receptors in schizophrenia (Tamminga and Holcomb 2005). However, as serotonin is known to interact with other neurotransmitter systems including the dopaminergic, the involvement of serotonin in the aetiology of schizophrenia may not be simple and may involve the interaction of several neurotransmitter systems (Tamminga and Holcomb 2005).

Detailed discussion of all the proposed neurochemical hypotheses are beyond the remit of this thesis, but proponents exist for a contribution from all major neurotransmitter systems including those mentioned above, GABA, Acetyl choline and noradrenaline.

1.3 Genetic Epidemiology of Schizophrenia:

Genetic epidemiology studies of schizophrenia provide strong evidence for the involvement of genetic factors in the diseases aetiology. The numerous family studies which have been performed over the last century clearly show that the risk of schizophrenia is increased in the relatives of schizophrenic patients (Shih, Belmonte et al. 2004). In comparison to the 1% risk of schizophrenia in the general population, there is a modest increase to about a 2-3% risk observed in second and third degree relatives of schizophrenics, whilst the risk increases further to about 10% in first degree relatives (McGuffin, Owen et al. 1995). However, family studies do not allow us to differentiate the genetic vs environmental components of schizophrenia. In order to try and distinguish these factors investigators have used further twin and adoption study designs.

The basis of twin studies is to compare concordance rates for the disease between members of monozygotic (MZ) twin pairs and dizygotic (DZ) twin pairs. As MZ twins share 100% of their alleles, whereas DZ twins share 50%, under the assumption that each type of twin pair share the environment to a similar extent, a higher concordance in MZ than DZ twins indicates the presence of a genetic aetiology (McGuffin, Owen et al. 1995). In a recent study (Cardno and Gottesman 2000) examined the five most recent twin studies. The authors reported a monozygotic concordance estimate of 41-65%, compared with a dizygotic concordance estimate of 0-28%. They then used the data to estimate a broad genetic heritability for schizophrenia of 85%, with the remaining 15% being attributed to environmental factors.

Adoption studies also provide evidence that the increased risk of schizophrenia in the relatives of affected patients is not due to shared environmental factors. Individuals adopted into families containing an affected individual do not show an increased risk of schizophrenia, whereas an adoptee from a family with an affected individual, who is adopted by a family with no history of the disease, does carry an increased risk of schizophrenia (McGuffin, Owen et al. 1995; Shih, Belmonte et al. 2004).

Although the epidemiological studies of schizophrenia prove that genes play an important role in the disease, environmental factors are still important. This is shown by the fact that the concordance of schizophrenia in MZ twins is around 50% and the estimated heritability of schizophrenia is less than 100%. What is inherited is not schizophrenia, but rather an increased susceptibility to developing the disease.

1.4 Molecular Genetic approaches to Schizophrenia:

It is clear from the genetic epidemiology studies described above that there is a significant genetic contribution to schizophrenia. However, studies looking at different family relatives of schizophrenic individuals have shown that the relative risk of developing schizophrenia decreases too rapidly with increasing genetic distance from the affected individual for schizophrenia to be a single gene disorder or a collection of single gene disorders. It appears more likely that the mode of transmission of schizophrenia is complex, involving the action of several, perhaps many genes which each account for a small increase or decrease in susceptibility for the disease. In fact (Risch 1990; Risch 1990) calculated that the data is incompatible with the existence of any single locus conferring a relative risk in siblings (λ_s) of more than 3, and that models with several loci of $\lambda_s \leq 2$ were more likely. It is however still possible that there are some rare genes of larger effect operating in certain schizophrenic pedigree families or subpopulations. To date there have been three main genetic approaches used to try and identify schizophrenia disease genes in human populations: Linkage studies, association studies and the identification of chromosomal abnormalities in schizophrenic patients.

1.4.1. Linkage Studies:

The object of linkage studies is to try and identify regions of the genome containing genes that predispose to disease. This is done by studying related affected individuals and looking for areas of the genome that are co-transmitted along with disease in these individuals. The main assumption behind this approach is that a gene involved with a trait of interest is located physically close to the marker locus being studied whose location is known in the genome. The nearer two loci are together on a chromosome, the less likely it is that a recombination event will separate them during meiosis, and hence the more likely it is that they will co-segregate. The probability of a recombination event occurring between two markers during meiosis can be calculated to give the recombination fraction (θ). If two genetic loci are on separate chromosomes or are far apart on the same chromosome the recombination fraction (θ) will be 0.5, indicating that there is an even chance of the two markers being separated during meiosis (independent assortment). Two loci are 'linked' if during meiosis, recombination occurs between them with a probability (recombination fraction) of less than 50% ($\theta = <0.5$), a departure from independent assortment. If there is non-independent assortment between a disease and a genetic marker, this indicates that a gene responsible for causing the disease is located in the region of the genome containing that particular marker. Using the linkage method it is possible to screen the whole of the genome for co-transmitted regions using a relatively small number of evenly spaced genetic markers (around 200-300 micro-satellites). Linkage can potentially identify the chromosomal position of disease related regions/genes without any prior knowledge of the disease aetiology (Owen, Williams et al. 2004). In order to use parametric statistical methods that have been successful for simple genetic disorders a genetic model describing the mode of inheritance of the disease is required and ideally there is a strong correlation between genotype and phenotype. Unfortunately as I have mentioned above schizophrenia appears to be either polygenic or oligogenic involving multiple genes of moderate to small effect and the power of linkage to detect these genes is limited as a consequence by the reduced correlation between any single locus and the phenotype.

Although it does not seem that simple genetic forms are the rule many early linkage studies of schizophrenia were performed in the hope that at least some multiply affected schizophrenic families would be caused by inheritance of alleles of large effect which once identified could then further inform researchers of the possible gene pathways and systems involved in the disease. Unfortunately, the negative results of these early studies suggested that individual largely monogenic forms of schizophrenia are at best extremely rare. However, recently with the introduction of international collaborative studies and the availability of larger sample sizes, linkage studies have started to provide evidence of loci for schizophrenia, which have then been replicated in other samples. The linkage studies have been performed in samples consisting of large multiply affected families and also samples of affected sib-pairs, with both types of analysis providing positive evidence of loci for schizophrenia. For instance the putative susceptibility gene NRG1 was detected following the evidence of linkage to 8p12-21 detected in large multiply affected families (Stefansson, Sigurdsson et al. 2002), whilst the putative susceptibility gene DTNBP1 was detected following evidence of linkage to 6p24-22 in samples of affected sib-pairs (Schwab, Albus et al. 1995; Straub, MacLean et al. 1995).

Some of the best-supported regions of the genome include: 6p24-22 (Schwab, Albus et al. 1995; Straub, MacLean et al. 1995), 10q25-26 (Williams, Norton et al. 2003), 1q21-22 (Brzustowicz, Hodgkinson et al. 2000), 13q32-34 (Blouin, Dombroski et al. 1998; Brzustowicz, Honer et al. 1999), 8p21-22 (Blouin, Dombroski et al. 1998; Gurling, Kalsi et al. 2001) and 6q21-25 (Martinez, Goldin et al. 1999; Lindholm, Ekholm et al. 2001), with further promising findings for regions like 22q11-12 (Gill, Vallada et al. 1996; DeLisi, Shaw et al. 2002; Williams, Norton et al. 2003), 5q21-q33 (Schwab, Eckstein et al. 1997; Straub, MacLean et al. 1997; Gurling, Kalsi et al. 2001), 10p15-p11 (Schwab, Hallmayer et al. 1998; Straub, MacLean et al. 1998), 1q42 (Blackwood, Fordyce et al. 2001), and 17p11-25 (Williams, Norton et al. 2003). It is likely that some of these results represent false positives. However, it is difficult to distinguish the true from false positives on the basis of replication because in the context of the expected small effect sizes of the genes being sought, the sample sizes are not large enough to be able to reliably replicate any single finding. A method that has been proposed to circumvent the problems with replication of linkage studies is the use of meta-analysis, which merges the results of many separate studies together, thus providing an increased power to detect the genes that influence schizophrenia susceptibility. Recently, two research groups have performed meta-analysis of the schizophrenia linkage studies. Both of these studies have used different methods and reported contrasting but overlapping findings. Lewis and colleagues (Lewis, Levinson et al. 2003) looked at 20 linkage studies and reported genome-wide significant evidence for a susceptibility locus on chromosome 2q. The authors also reported evidence for the presence of susceptibility loci on chromosomes 5q, 3p, 11q, 6p, 1q, 22q, 8p, 20q and 14p, with lesser evidence for several other chromosomal regions including 16q, 18q, 10p, 15q, 6q and 17q. In the second study, Badner and Gershon (Badner and Gershon 2002) reported evidence for the existence of susceptibility loci on chromosomes 8p, 13q and 22q. The only regions receiving support from both meta-analyses were 8p and 22q. The study by Lewis and colleagues (Lewis, Levinson et al. 2003) was based upon a larger and more complete data set and it is probably reasonable to assume that some of the additional significant regions identified in the study might represent genuine susceptibility loci that were not detected in the other meta-analysis.

1.4.2. Cytogenetic Abnormalities:

Another method which has been used to try and identify the location of schizophrenia susceptibility genes has been the identification of chromosomal abnormalities in affected patients. The hope was that cytogenetic abnormalities occurring in a specific region of the genome such as translocations, inversions, deletions or duplications would either exist in greater frequency in schizophrenics than in the general population, would disrupt a region previously implicated by linkage analysis, or would be inherited along with schizophrenia in affected families. Any abnormalities detected could be pathogenic through various mechanisms such as the direct disruption of a gene or genes, the formation of an aberrant gene product by fusing together two previously separate genes, by indirectly disrupting the function of a neighbouring gene by modifying its expression, or by altering the expression of the gene in the case of deletions or duplications.

There have been many studies reporting the association of chromosomal abnormalities and schizophrenia. However, only two findings have provided convincing evidence for the possible location of a susceptibility gene. Velo-Cardio Facial Syndrome (VCFS), which is associated with small chromosome 22q11 deletions, have a high risk of schizophrenia (Bassett, Hodgkinson et al. 1998), with later studies (Murphy, Jones et al.

1999; Murphy 2002) estimating the relative risk to be around 25-30%. The prevalence of VCFS in the general population is too low for the disease to account for more than about 1% of schizophrenia cases (Owen, Williams et al. 2004), and therefore from the schizophrenia genetics point of view, the most interesting question to try and answer is whether the high rate of schizophrenia in VCFS patients can provide a clue to the location of a susceptibility gene or genes involved in schizophrenic patients without chromosome 22 deletions. There is some support for this theory from the results of linkage studies which have previously reported evidence for the location of a susceptibility locus for schizophrenia within a broad region of chromosome 22 (Gill, Vallada et al. 1996; DeLisi, Shaw et al. 2002; Williams, Norton et al. 2003). Indeed in the meta-analysis (Lewis, Levinson et al. 2003) the 22p11-q21 region was one of the most significant results. Several candidate genes in the region (COMT, ARVCF and PRODH) have been studied for association with schizophrenia, but so far none have been convincingly implicated with disease (Williams, Glaser et al. 2005).

Another result of interest was the discovery of a (1:11) (q42;q14.3) balanced reciprocal translocation which was found to co-segregate with schizophrenia and major mood disorders in a large Scottish family (St Clair, Blackwood et al. 1990; Blackwood, Fordyce et al. 2001). This translocation has been reported to disrupt two genes on chromosome 1, called disrupted in schizophrenia 1 and 2 (Disc 1 and Disc 2) (Millar, Wilson-Annan et al. 2000; Blackwood, Fordyce et al. 2001). There have subsequently been several association studies published reporting that variation within Disc1 is associated with schizophrenia and bipolar disorder in the general population (Hennah, Varilo et al. 2003; Hodgkinson, Goldman et al. 2004; Callicott, Straub et al. 2005; Thomson, Wray et al. 2005).

1.4.3. Association Studies:

Association studies aim to detect a relationship between two variables. In the case of a genetic association study it is usually a relationship between a genetic polymorphism(s) and a specific trait, such as a particular phenotype or a disease that is being tested. A polymorphism is said to be associated with a trait if a genotype or allele occurs more often than would be expected by chance in people with that particular trait. For instance an allele that is significantly over represented in a disease population when compared to the frequency in the general population would be said to be significantly associated with the disease.

The genetic loci that are thought to be involved with complex diseases such as schizophrenia are likely to be of moderate to small effect size (Risch 1990). Compared with linkage, association studies provide an alternative and more powerful means of identifying moderately common alleles of modest to small effect in realistically attainable sample sizes compared to linkage studies (Risch and Merikangas 1996). As an example Risch and colleagues (Risch and Merikangas 1996) calculated that to detect a gene with a genotype relative risk (GRR) of 2.0 with a disease allele frequency of 0.5 using a linkage study design would require around 2500 families. In contrast, they calculated that to detect the same allele using an association study design would require only around 200 families.

Unfortunately, studies based upon association do have a major disadvantage compared to linkage. It has been estimated that association studies in normal populations (no significant inbreeding) would require at least several hundreds of thousands of markers to be genotyped in order to screen the whole genome by association (Gabriel, Schaffner et al. 2002; Carlson, Eberle et al. 2004; Hirschhorn and Daly 2005). This large number of markers initially made the prospect of whole genome association studies economically impracticable, even for the most highly funded research laboratories. However, recently with the advent of new technology allowing genotyping on an extremely large scale, the prospect of whole genome association studies is now feasible.

The basic and most commonly used form of association study is a case control study involving comparing the frequency of alleles in a sample of individuals from a disease population (cases) against the frequency in a sample of unaffected individuals from some form of control population. By doing this it is possible to test for a statistically significant difference between the two populations and hence determine whether a particular genotype or allele is associated with an increased risk for a disease.

1.4.3.1. Previous association studies in schizophrenia:

Given that the numbers of markers required prohibited genome wide studies until recently, association studies of schizophrenia have been required to select specific candidate genes or relatively small regions for investigation by association. The selection of these genes has been based upon evidence that the gene is a functional candidate on the basis of its putative involvement with schizophrenic neurobiology, on the basis that the gene is located in a region of the genome implicated by schizophrenic linkage studies, or alternatively based upon a combination of the two criteria. Despite the restrictions on the scope of association studies they have still been successfully used to identify genes of minor effect in various complex genetic disorders, including Alzheimer's (Corder, Saunders et al. 1993; Strittmatter, Saunders et al. 1993), dyslexia (Francks, Paracchini et al. 2004; Cope, Harold et al. 2005), schizophrenia (Chumakov, Blumenfeld et al. 2002;

Stefansson, Sigurdsson et al. 2002; Straub, Jiang et al. 2002), type 1 diabetes (Vella, Cooper et al. 2005) and type 2 diabetes (Grant, Thorleifsson et al. 2006).

Hundreds, possibly thousands, of association studies on candidate genes have been published in schizophrenia. The results and quality of the data produced by these studies has been variable and the results inconsistent and unconvincing, due possibly to the small sample sizes and power of the studies to enable detection or replication of associations with genetic variants of moderate to small effect size (Owen, Williams et al. 2004), and incomplete marker coverage of the genes. The positive associations that have been reported have often failed to replicate, associations have also been reported with different alleles of the same variant, or associations have been reported with different variants in the same genes.

However, recently there have been a few genes reported to be associated with schizophrenia which have subsequently received strong evidence from replication studies, indicating the genes may be putative susceptibility loci for schizophrenia. Three of the most compelling are Neuregulin 1 (NRG1), Dysbindin (DTNBP1) and D-amino acid oxidase activator (DAOA; also known as G72). All three of these genes were initially selected for study on the basis of their location within regions of the genome that has been previously linked to schizophrenia.

Neuregulin (NRG1) was first reported to be associated with schizophrenia by Stefansson and colleagues (Stefansson, Sigurdsson et al. 2002), who following up their earlier linkage findings on chromosome 8p12-21, detected an association with schizophrenia for a single marker and also a more significant association with a seven marker haplotype at the 5' end of NRG1. This initial result was replicated in an independent Scottish case control sample, with association being reported individually for three of the markers in the original Stefansson (2002) seven-marker haplotype and again a more significant association with the haplotype itself (Stefansson, Sarginson et al. 2003). A further study (Williams, Preece et al. 2003) also supports NRG1 as a putative schizophrenia locus by reporting association with the haplotype, although at a moderate level of significance. Further positive results have been reported in various populations including Irish (Corvin, Morris et al. 2004), Chinese (Yang, Si et al. 2003; Tang, Chen et al. 2004; Zhao, Shi et al. 2004), South African (Hall, Gogos et al. 2004) and Portuguese (Petryshen, Middleton et al. 2005), with different haplotypes. Recently two meta-analyses of the NRG1 association data have provided further support for the genes involvement in schizophrenia (Li, Collier et al. 2006; Munafo, Thiselton et al. 2006).

Dysbindin (DTNBP1) was initially implicated in schizophrenia by Straub and colleagues (Straub, Jiang et al. 2002) who analysed their previously reported linkage region on chromosome 6p22, and detected association to schizophrenia with several individual markers in DTNBP1 and also significant association with several three-marker haplotypes across the gene. The initial result was supported by the findings of Schwab (Schwab, Knapp et al. 2003) who reported association between DTNBP1 and schizophrenia in two separate samples. They reported association with a single marker in both their samples and also increased association with various two and three marker haplotypes. Additional support has been provided by the results of several large studies from the UK (Williams, Preece et al. 2004), Bulgaria (Kirov, Ivanov et al. 2004), America (Funke, Finn et al. 2004), China (Tang, Zhou et al. 2003) and Japan (Numakawa, Yagasaki et al. 2004). All studies have reported associations with different haplotypes in DTNBP1 and schizophrenia. Interesting complementary evidence has also been reported by Bray and colleagues (Bray, Preece et al. 2005) who demonstrated that the associated risk haplotypes in DTNBP1 tag one or more cis-acting variants responsible for a relative reduction in the expression levels of the dysbindin gene.

D-amino acid oxidase activator (DAOA; also known as G72) was first shown to be associated with schizophrenia by Chumakov and colleagues (Chumakov, Blumenfeld et al. 2002), after following up on their earlier linkage findings on chromosome 13q22-34. Replication studies have subsequently reported positive association between DAOA and schizophrenia in the German (Schumacher, Jamra et al. 2004), Chinese (Wang, He et al. 2004; Zou, Li et al. 2005), Palestinian (Korostishevsky, Kaganovich et al. 2004; Korostishevsky, Kremer et al. 2006) and American (Addington, Gornick et al. 2004) populations. It is of note that positive association has also been reported between DAOA and Bipolar disorder (Hattori, Liu et al. 2003; Chen, Akula et al. 2004b; Schumacher, Jamra et al. 2004; Williams, Green et al. 2006). Interestingly when Williams and colleagues (2006) analysed their schizophrenic and bipolar data across the traditional diagnostic boundaries they detected a significant association between variation within DAOA and patients suffering from major mood episodes, but not in those schizophrenics with no history of mood disorder.

It is important to remember that the detection of a positive association between a marker and a disease does not always mean that you have detected the specific variant which is responsible for causing increased risk for the disease. Positive association between a genetic polymorphism and a trait can be detected in a particular population for several reasons:

1.4.3.2. Direct association:

Association analysis based on known functional SNPs or complete mutation discovery are termed 'direct' studies. Here the aim is to test variants directly responsible for modifying gene function and hence influencing disease risk. Unfortunately, it is not thought that all the causal variants responsible for complex diseases will be located within coding sequence, since variation located within non-coding sequence could also alter a genes function by affecting the expression, regulation or splicing of the gene. The unknown functional boundaries of genes makes complete mutation discovery impossible and it is therefore difficult to specify a subset of markers for direct studies.

1.4.3.3. Spurious association:

Association can also be caused by confounding factors such as population stratification which can lead to the detection of a false positive association. Population stratification occurs when a population is made up of several genetically distinct subgroups (e.g different ethnic backgrounds). If the ascertained cases and controls in a study have a different proportion of individuals from each genetic background any alleles that differ in frequency simply by genetic background can result in a false positive association. This type of association is termed 'spurious' association as there would not be a true positive association between the two variables detected within any of the individual genetic subgroups (Sham 2003). These same confounding factors can also result in 'false negative' association findings by masking the effect of a true association signal.

The phenomenon of 'spurious association' is obviously of major concern and molecular geneticists employ several steps in an attempt to avoid the problem. The first step is to ensure that association studies are performed in samples selected from well mixed, out-breeding populations, and that the case and control samples are carefully 'matched' for all factors except their disease status. This includes matching the samples for traits such as their ethnic origin, age, sex and geographical region.

The use of family based studies has also been strongly advocated to avoid the problem of stratification. This is preferably achieved through the use of affected-parent trio samples or, where it is not possible to collect parental samples (e.g diseases with late onset); samples consisting of affected and non-affected sibling pairs can be used. In order to test for potential association, the alleles of the parents that have been transmitted and non-transmitted to the affected probands are determined at the locus to be analysed and statistical tests are used to determine whether a particular allele is transmitted more frequently to affected individuals than expected by chance. The transmission frequency analysis is usually performed using either the haplotype based, haplotype relative risk (HRR) method (Terwilliger and Ott 1992), or the transmission disequilibrium test (TDT) (Spielman, McGinnis et al. 1993). However, family based association studies do have some disadvantages compared to more traditional case-control studies. Firstly they are more expensive as they require more genotyping due to the trio design, and secondly there are often problems with sample collection due to the availability of the parents. For example in diseases with a late age of onset, or diseases with strong social stigma and high rates of family disruption (e.g schizophrenia) the collection of family based samples is very difficult.

Additionally it has also been proposed that the use of genomic controls can help to avoid the problem of stratification. This method proposes that by genotyping a large number of unselected genetic markers across the genome the extent to which allelic variance is influenced by population stratification can be estimated and then the threshold required for statistical significance subsequently adjusted in order to reduce the level of false positive associations detected.

A further cause of false positive association findings is type I error (statistical chance) as a result of sampling variation. If the level of statistical significance adopted in a study is p=0.05, then for every twenty markers tested for association one will appear to be statistically significant by chance. A high rate of type I error in published data may result from the inadequate correction of association findings for the total number of markers tested at a locus. One method proposed for dealing with type I error is the use of the Bonferroni correction, where the corrected threshold for significance is set at 0.05 / the total number of independent association tests performed. However, since association studies of complex traits are likely to require the testing of large numbers of markers, correcting p-values using the Bonferroni correction is likely to require an initial level of statistical significance several orders of magnitude greater in order to remain significant after correction. Such levels of significance are extremely difficult to achieve, and therefore by applying such a conservative correction it is possible that true association findings may be missed. A further problem with the use of the Bonferroni correction is the existence of LD between markers. The presence of LD leads to non-independence of markers and thus results in the threshold of significance being over corrected. The answer to the problems discussed above is not simple. The best answer to the problem of type I error is the robust independent replication of positive findings in suitably powerful association samples. In order to allow for the problems associated with the non-independence of markers, the use of permutation testing provides a means to take into account the presence of LD.

1.4.3.4. Indirect association:

The final reason for a positive association being detected in a population is termed 'indirect' association. This occurs when a polymorphism being tested for association with a disease is actually a surrogate marker for the true causal variant. This situation occurs due to the existence of a phenomenon known as 'linkage disequilibrium' (LD). In this case the polymorphism detected to be showing association with a disease has no direct effect on susceptibility, but is instead in linkage disequilibrium (LD) with another polymorphism that is the actual true causal variant, and the association is therefore termed 'indirect'.

LD refers to the non-random population association of alleles. This arises when an allele at one locus occurs more frequently with a specific allele at a second locus than would be expected if the two loci were segregating randomly during recombination (Ardlie, Kruglyak et al. 2002). Normally after a mutation event the rearrangement of alleles during recombination events rapidly destroys the relationship between alleles at the mutated site and those on the chromosome on which the mutation occurred. However, if two alleles are very tightly linked the chance of them being separated during meiosis is low, and the level of LD between the markers will decay slowly. This may result in the persistence of relatively strong levels of LD between the markers for many generations. That strong LD can exist between two or more tightly linked markers allows 'indirect' association to be detected with a marker that is not itself the causal susceptibility variant for disease, but is instead displaying association due to strong LD with the true causal variant. The fact that association to a disease can be detected with markers other than the causal variant by virtue of LD is crucial to association study design. It provides geneticists with the ability to screen genes or genomic regions for association by only testing a selection of the genetic variation within that region, a technique which is termed 'linkage disequilibrium mapping'.

1.4.4. Linkage Disequilibrium (LD) mapping:

1.4.4.1. Measurement of linkage disequilibrium:

In order for geneticists to be able to use the phenomenon of LD to their advantage they need to be able to measure the strength of the LD which exists between two loci. There have been many methods developed over the years to assess the level of LD between markers but the two most commonly used statistics are D' and r^2 (Ardlie, Kruglyak et al. 2002). Both are based upon one of the earliest measures of disequilibrium to be developed, termed the population genetic parameter (symbolised by D). The statistic D quantifies disequilibrium as the difference between the observed frequencies of two marker haplotypes and the frequencies of the haplotypes that would be expected to occur if the alleles at the two loci were segregating randomly. However, although the statistic D represents the concept of LD, its numerical value is not very useful for comparing and determining the strength of LD between markers, due to its reliance on allele frequencies. This limitation led to the development of alternative measures, the two most common of which are D' and r^2. These two statistics both range from 0 (no disequilibrium) to 1 (complete disequilibrium), although the interpretation of the statistics is slightly different.

The statistic of D' is derived by dividing the population genetic parameter D by its maximum possible value, thereby adjusting for the allele frequencies at the two markers being tested. The way that D' is calculated means that D'=1 only if there has not been any recombination present between the tested markers during the history of the population. This is often called 'complete LD' (note that formally, D' can have values ranging from -1 to 1 but the sign is usually dropped in modern practice). Where D'=1, at most, three of the possible four two locus haplotypes are present in the sample. A value of D'=<1 indicates that the markers have been disrupted by recombination and represents the presence of all four of the possible two locus haplotypes in the population. While a D' value of 1 or 0 is relatively easy to understand (either complete or no disequilibrium), intermediate values are more difficult to interpret. This difficulty is caused in part by the finding that values of D' can be inflated in small sample sizes, particularly when the markers being analysed have rare allele frequencies. It has been suggested that intermediate D' values should not be used to compare the strength of LD between samples (Ardlie, Kruglyak et al. 2002).

The second commonly used statistic of LD is r^2 and represents the square of correlation (r) of the alleles at the two loci being tested. The statistic is complementary to D' and has recently become for many the LD measure of choice for comparing the strength of LD between markers and samples. r^2 is derived by dividing D^2 by the product of the four allele frequencies at the two loci being tested. A value of $r^2=1$, which is termed 'perfect LD' only occurs if the markers being tested have not been separated by recombination and also have the same allele frequency. In this case only two of the four possible two-locus haplotypes are present in the sample. When two markers are in perfect LD, genotype observations that are made at one marker provide the complete set of information about the other marker, therefore making genotyping of the second marker redundant. Compared to the D' statistic, intermediate values of r^2 are easier to interpret as the value of r^2 is directly related to the amount of information provided by one marker about another, as well as the sample size required to detect an association (Wall and Pritchard 2003). For example consider two markers, one of which (A) is involved with susceptibility to disease, and a second nearby marker (B) in LD with the susceptibility locus. It has been calculated that in order to provide the same power to detect an association that would be provided by genotyping the susceptibility locus (A) itself, marker (B) would need to be genotyped in a case control sample whose size has been increased by a factor of $1/ r^2$ (Wall and Pritchard 2003). This has been shown to be an oversimplification (Terwilliger and Hiekkalinna 2006), though on average it is correct. Similar to the D' statistic, the value of r^2 takes into account the differences in allele frequencies at the markers being tested, but additionally has the significant advantage over D' that the value of r^2 shows less inflation in small sample sizes (Weiss and Clark 2002). Due to the r^2 statistics greater stability in smaller sample sizes and the ability to interpret intermediate values, the statistic has been proposed as the LD measure of choice for use in association studies (Pritchard and Przeworski 2001).

1.4.5. Extent of LD in the human population:

The extent of linkage disequilibrium within the human genome is not purely determined by the genetic distance between the markers being tested. As a consequence, markers that are tens or even hundreds of kilobases apart may be in strong linkage disequilibrium, whereas markers that are close together on the genome might be in weak LD (Wall and Pritchard 2003). Large differences in the extent of LD exist for different regions of the genome (Abecasis, Noguchi et al. 2001; Reich, Cargill et al. 2001). This variable extent of LD observed across the genome is explained by several factors.

1.4.5.1. Factors influencing LD diversity:

Mutation age/rate and recombination probably have the most obvious impact on the extent of LD across the genome, but there are other factors which are also important.

Genetic drift: This results in changes in gene and haplotype frequencies caused by random sampling of gametes. This is of particular importance in small populations where it can result in the loss of some haplotypes from the population and therefore lead to increased levels of LD (Ardlie, Kruglyak et al. 2002).

Population growth: Rapid population growth decreases the extent of LD by reducing genetic drift (Ardlie, Kruglyak et al. 2002).

Population bottlenecks: During a population bottleneck the size of a population is significantly reduced leading to a loss of haplotypic diversity within the population and hence resulting in an increase in the extent of LD within the population (Pritchard and Przeworski 2001).

Migration (gene flow): This occurs when two genetically distinct populations migrate or admix together and create a 'hybrid' population, which can result in the generation of LD. Recently admixed populations can have very strong levels of LD across large genetic distances (Pritchard and Przeworski 2001). The rate at which the level of LD breaks down in these populations is determined by the size of the hybrid population and whether the population continues to mix with the original populations (slow rate of LD decay) or whether the population becomes isolated (fast rate of LD decay) (Pfaff, Parra et al. 2001).

Natural selection: Selection can affect the extent of LD through two mechanisms. The first is where a haplotype which flanks a (selected) favourable mutation becomes highly frequent in a population, thus inflating LD (Ardlie, Kruglyak et al. 2002). The second selection mechanism which can affect LD is through the epistatic selection of combinations of alleles at two or more loci on a chromosome thereby increasing the level of LD (Ardlie, Kruglyak et al. 2002).

Gene conversion: This is a phenomenon where a stretch of one copy of a chromosome is transferred to the other copy during meiosis. This is very similar to two very close recombination events and can result in the breakdown of LD in a manner similar to recombination and mutation (Ardlie, Kruglyak et al. 2002).

Variable recombination and mutation rates: Recombination and mutation rates are known to vary across the genome. As the breakdown of LD is thought to mostly depend on recombination and mutation rates the extent of LD is thought to be related to these rates. It has even been suggested that recombination might be restricted to certain regions of the genome (hotspots) resulting in the breakdown of LD in these regions, whilst recombination might be almost missing in other areas of the genome resulting in high levels of LD (Ardlie, Kruglyak et al. 2002).

1.4.6. Studies of the extent of LD across the genome:

It is generally accepted that LD declines with increasing genetic distance, but as I discussed there are several factors that can also influence LD and result in LD variation across the genome, between genes and also between different populations. This variation has led to an increasing number of studies examining the patterns of LD across the genome and between different populations:

One of the earliest of these reports (Kruglyak 1999), was based on a simulation study and predicted that useful levels of LD (suitable for association studies) were unlikely to span distances of more than 3kb in a general outbreed population. The authors then suggested that taking this into consideration roughly 500,000 markers would be required to be genotyped in order to perform whole genome association studies. However, the prediction that useful LD only extends over regions of 3kb or less has been rejected by several empirical studies (Abecasis, Noguchi et al. 2001; Reich, Cargill et al. 2001; Stephens, Schneider et al. 2001).

1.4.7. The block structure of LD:

Recent studies have suggested that the pattern of LD across the genome can be structured into blocks of high and low LD, termed 'haplotype blocks'. Two early studies (Daly, Rioux et al. 2001; Jeffreys, Kauppi et al. 2001) reported that LD in the genome was highly structured and broken up into 'blocks' of high LD with low haplotypic diversity (recombination coldspots) interspersed by genomic regions with high levels of recombination (recombination hotspots). These initial studies were performed on relatively small regions of the genome of about 500kb, but a more recent study has been performed on a larger 13Mb region of the genome and demonstrated that a region of this size could also be defined into distinct blocks of high LD separated by recombination hotspots (Gabriel, Schaffner et al. 2002). There have since been several other LD studies performed which have examined large chromosomal regions and again reported that the genome can be divided into 'haplotype block' structures of high LD separated by hotspots of recombination (Dawson, Abecasis et al. 2002; Phillips, Lawrence et al. 2003).

1.4.8. The use of haplotype blocks in LD mapping:

The existence of often high levels of LD throughout the genome has led to the prospect of using 'tag' SNPs to reduce the number of markers that are required to be genotyped in order to perform a whole genome association study (Johnson, Esposito et al. 2001; Cardon and Abecasis 2003). The term 'tag' SNP refers to a SNP that when genotyped individually or as part of a multi-marker haplotype, captures a high proportion of the common haplotypic information surrounding a gene or genomic region. In order to determine which tag SNPs capture the genetic information of a region, initially a highly dense set of SNPs needs to be genotyped in a representative panel of individuals (Kruglyak 2005). Using the data the minimum number of SNPs required to be genotyped in order to capture all or a user defined proportion of the genetic variation within that region by LD (either individually or through multi-marker haplotypes) can be calculated. This technique therefore reduces the amount of genotyping required to cover a region, whilst maintaining the power to detect at least the common known variants present.

The possibility of using tag SNPs to screen a genomic region for susceptibility variants for common disease inspired the foundation of the international HapMap consortium. This project has obtained genotype information for more than six million SNPs (roughly one SNP every 1kb) across the genome in 269 individuals from four different ethnic populations (Consortium. 2005). The project therefore allows genetic researchers to determine the tag SNPs within a region without the need to perform an initial dense SNP genotyping study in their own sample. However, the use of the HapMap does rely on the assumption that the LD structure in the population of interest is the same as in the populations studied by the HapMap consortium, an assumption which is less reliable in populations of African origin, as these populations generally contain more genetic variation and less LD (Hirschhorn and Daly 2005).

Support for the HapMap project is provided by the results of a large genotyping study (Hinds, Stuve et al. 2005). The authors presented genotyping data for more than 1500000 SNPs across the genome (~1 SNP every 1.8kb) in 71 individuals from three different populations of European-American, African-American and Asian-American ancestry. Using the genotype data the authors calculated the LD structure across the genome and then used this to determine blocks (bins) of SNPs where each block had at least one tag SNP with an r^2 of at least 0.8 with all of the other SNPs in that block. Using this pairwise r^2 method to define blocks and then select tag SNPs, the authors showed that in order to capture the common genetic information would require 30% of the overall SNPs to be genotyped in the European-American population, 28% of the SNPs to be genotyped for the Asian-American population and 50% of the SNPs to be genotyped in the African –American population (Hinds, Stuve et al. 2005). Although the results of this study indicate that by using tag SNPs the quantity of genotyping required to perform a whole genome association study can be significantly reduced, the authors concluded that at best a comprehensive whole genome study will still require genotyping of at least several hundreds of thousands of SNPs for each individual in the association sample (Hinds, Stuve et al. 2005). Indeed in a recent study (Barrett and Cardon 2006) assessing the coverage of the phase II HapMap data it was reported that nearly all common variation in the European-American and Asian-American populations can be captured with $r^2 \geq 0.8$ using approximately 500000 carefully selected SNPs. However, to capture nearly all the common variation in the African-American population would require more than twice as many SNPs. The authors also reported that in all cases the gain in coverage achieved with increasing numbers of markers showed a diminishing level of return. For example, although 500000 SNPs are required to capture the common variation in the European-American population, a set of 250000 SNPs would capture 85% of the variation present (Barrett and Cardon 2006). A consequence of this effect is that a large proportion of the tag markers are capturing 'singleton' SNPs that are not in strong LD with any other SNPs.

In an effort to reduce the number of SNPs required to be genotyped further methods of tagging the genetic information across the genome have been proposed. The most common of these alternatives is the use of multi-marker haplotypes in order to tag the genetic variation across the genome (Stram 2004; de Bakker, Yelensky et al. 2005). The basis of this method is that a specific multimarker combination (haplotype) can serve as a proxy for several individual tag SNPs, and thus make the genotyping of these tags individually redundant. Using this method each individually typed tag as well as each specific haplotype is tested for association.

Recently de Bakker and colleagues (de Bakker, Yelensky et al. 2005) using the available HapMap genotyping data have calculated that by using a pairwise r^2 method to select tag SNPs reduced the number of genotypes required by 65% in the European population compared with testing all of the common SNPs directly. The subset of tag SNPs maintained the same level of relative power as would be achieved by testing all of the available genetic variation. The authors then went on to calculate whether the use of multimarker haplotypes would improve the genotyping efficiency any further. They reported that by using a haplotype tagging technique the number of tag SNPs required to be genotyped was reduced by a further 30% compared to the number required using the pairwise r^2 method, whilst still maintaining the same relative power (de Bakker, Yelensky et al. 2005).

1.5 Oligodendrocytes, Myelination and Schizophrenia:

Over the last decade there has been increasing evidence suggesting that abnormal oligodendrocyte function and/or myelination plays a role in schizophrenia aetiology. The main strands of evidence that provide support for this come from studies of neuropathology, imaging analysis, symptom overlap with demyelinating diseases, age related changes in myelin and evidence from micro-array studies.

Oligodendrocytes are the myelinating glial cells in the central nervous system and are essential for the development of the neuronal system and for maintaining the function of mature neurons. Oligodendrocytes are the predominant cell type in the white matter of the brain where their main known function is the myelination of axons. Satellite oligodendrocytes also exist which are involved in regulating the microenvironment around neurons (Baumann and Pham-Dinh 2001).

During myelination oligodendrocytes create highly branched membrane processes that target and ensheath the axons of neurons in multiple layers of myelin to form the myelin sheath (Baumann and Pham-Dinh 2001). The myelin sheath is a highly organised multi-lamellar membrane structure, consisting of a specialised set of lipids and proteins. The presence of the myelin sheath around axons is essential for the rapid conduction of action potentials along axons by the process of saltatory conduction (Sherman and Brophy 2005). Proper myelination of the axons of the central nervous system by oligodendrocytes is therefore critical to normal neuronal function, a fact demonstrated by the involvement of abnormal myelination in diseases such as leukodystrophies and multiple sclerosis (MS) in the central nervous system (Baumann and Pham-Dinh 2001).

Oligodendrocytes and myelin have also been proposed to play other functions in the CNS. Oligodendrocytes influence neuronal maturation. For example they trigger the clustering of sodium channels around the nodes of Ranvier on axons (Sherman and Brophy 2005). It has also been suggested that oligodendrocytes are able to promote the radial growth of axons, an important process, as conduction velocity has been shown to be correlated with axonal diameter (Baumann and Pham-Dinh 2001). Additionally there is also evidence that oligodendrocytes secrete factors that promote axonal maintenance and survival although the molecular mechanism is still to be determined (Baumann and Pham-Dinh 2001; Sherman and Brophy 2005). Finally there is evidence that the inhibition of axonal growth and regeneration in the CNS is in part caused by myelin proteins on the surface of the myelin sheath (see Chapter 4, section 4.1).

Abnormalities in the myelination process and/or factors that affect myelination and the deposition of myelin such as the function of oligodendrocytes could all therefore have a profound effect on brain function.

1.5.1. Oligodendrocyte and myelin related gene expression in schizophrenia.

Some of the strongest evidence for the existence of an oligodendrocyte and/or myelination abnormality in schizophrenia comes from micro-array studies of oligodendrocyte and myelin related (OMR) gene expression. The study of global mRNA expression levels offers a complementary approach to the traditional strategy of selecting candidate genes and/or biological pathways for association study in the absence of any prior functional or positional evidence linking those genes to the disease aetiology (Williams, O'Donovan et al. 2002). Although there are several potential methodological as well as experimental and sampling hurdles to overcome with the use of complex post mortem tissue, the results of recent studies have suggested that reliable and robust changes in gene expression can be detected in human post mortem brain tissue. For example one of the earliest micro-array studies to be performed (Hakak, Walker et al. 2001) reported the downregulation of ErbB3 in schizophrenia, a result that has since been independently replicated (Tkachev, Mimmack et al. 2003; Aston, Jiang et al. 2004). The altered expression of ErbB3 becomes more aetiologically relevant in light of the evidence strongly implicating variation in the gene encoding one of its ligands, Neuregulin 1 (NRG1), with susceptibility to schizophrenia (Stefansson, Sigurdsson et al. 2002; Stefansson, Sarginson et al. 2003; Owen, Craddock et al. 2005; Li, Collier et al. 2006). Other less compelling examples include the identification of the G protein receptor kinase 3 gene as a candidate gene for bipolar disorder (Niculescu, Segal et al. 2000) and the regulator of G protein signalling 4 (RGS4) as a candidate gene for schizophrenia (Mirnics, Middleton et al. 2001) findings that have received supportive evidence from follow up genetic studies (Chowdari, Mirnics et al. 2002; Barrett, Hauger et al. 2003). In the case of RGS4 and GRK3, each gene was selected for genetic analysis because as well as displaying altered expression, each mapped to a putative region of linkage for the relevant disorder. Therefore each candidate was supported by data from two non-hypothesis based methods, an approach that has been termed ‘convergent functional genomics’ (Niculescu, Segal et al. 2000).

One of the earliest micro-array studies of schizophrenia suggested that oligodendrocyte and myelin related (OMR) dysfunction might play a role in schizophrenia (Hakak, Walker et al. 2001). The authors assayed over 6000 gene transcripts and detected 89 genes that displayed significant altered expression levels in post-mortem dorsolateral prefrontal-cortex brain samples of schizophrenic patients (n=12) compared to matched controls (n=12). All of the gene expression was predominantly up-regulated, except for a series of six genes related to oligodendrocyte and myelin function that were down-regulated in schizophrenic brain. The genes that were down-regulated were: myelin and lymphocyte protein (MAL), gelsolin (GSN), myelin associated glycoprotein (MAG), 2'-3'-cyclic nucleotide 3'-phosphodiesterase (CNP), transferrin (TF), and ErbB3. The altered expression of several of these genes has subsequently been replicated using independent post-mortem brain samples and different global expression assaying techniques, see table 1.1 for a summary of the OMR gene expression results.

The first of these replication studies (Tkachev, Mimmack et al. 2003) reported down-regulation of genes encoding major oligodendrocyte and myelination proteins, along with various oligodendrocyte transcription factors in post-mortem pre-frontal cortex brain of schizophrenic patients compared to controls. The authors validated their micro-array findings by quantitative PCR experiments to confirm the reductions in expression observed for the genes in schizophrenic brain. The genes reported to be down-regulated included the oligodendrocyte related genes; myelin associated glycoprotein (MAG), transferrin (TF), ErbB3, myelin basic protein (MBP), proteolipid protein 1 (PLP1), myelin associated oligodendrocyte basic protein (MOBP), claudin 11 (CLDN11) and myelin oligodendrocyte glycoprotein (MOG) as well as the oligodendrocyte lineage transcription factors Olig1, Olig2, and Sox 10.

A further micro-array study (Aston, Jiang et al. 2004) reported significant decreases in the expression levels of the myelin related genes MAG, plasmolipin (PLLP), PLP1 and ErbB3 in the post mortem temporal cortex of schizophrenic subjects compared to controls. The authors also observed reduced expression of three further myelin related genes (MAL, TF and CNP) that had been previously reported to be under-expressed in schizophrenics (Hakak, Walker et al. 2001).

The altered expression of OMR genes is additionally supported by the results of a cDNA micro-array study (Sugai, Kawamura et al. 2004) where the genes MOBP, MBP and MAL were all reported to be significantly down regulated in post mortem schizophrenic brain. The authors also reported that genes involved in oligodendrocyte and myelin processes had the highest frequency of expression alterations in schizophrenic patients compared to controls in their study. Further support is provided by the results of another micro-array and qPCR study (Iwamoto, Bundo et al. 2005a) reporting the decreased expression of OMR genes including; SOX10, Olig2, MAG, PLP1, MOBP and MOG in the prefrontal cortex of schizophrenic brain. Finally a recent micro-array analysis studying expression patterns of 15 different regions of post mortem schizophrenic brain compared to controls, with expression differences being subsequently confirmed by follow-up qPCR analysis in independent samples (Katsel, Davis et al. 2005b) reported the down regulation of a large number of OMR genes in schizophrenia across multiple brain regions. The study reported down regulated expression results for the genes QKI, TF, MAG, CLDN5, PMP22, Sox 10, CNP, GSN, ErbB3, MAL, MOG, CLDN11 and Olig2, in multiple regions of schizophrenic brain (Katsel, Davis et al. 2005b).

Recently two large studies by the same group (Katsel, Davis et al. 2005a; Katsel, Davis et al. 2005b) have reported that the largest expression differences and the greatest number of dysregulated OMR genes were observed in the cingulate cortex, the superior temporal gyrus and the hippocampus of schizophrenic brain. Most recently from the same group (Dracheva, Davis et al. 2006) has also reported region specific expression patterns of OMR genes in schizophrenic brain. The authors reported that the expression of MAG, CNP, SOX10, CLDN11, and PMP22 was reduced in the hippocampus and anterior cingulate cortex but not in the putamen of the post mortem schizophrenic brains that they studied. It is interesting to note that the altered expression of OMR genes is not exclusive to schizophrenia but has also been observed in post-mortem brain from other psychiatric diseases including bipolar disorder (Tkachev, Mimmack et al. 2003; Iwamoto, Bundo et al. 2005a) and unipolar depression (Aston, Jiang et al. 2005).

It should be noted that not all of the global expression assays performed have reported the altered expression of OMR genes in schizophrenia. Some of the early studies (Mirnics, Middleton et al. 2000; Middleton, Mirnics et al. 2002; Vawter, Crook et al. 2002) did not detect expression differences in schizophrenic post-mortem brain which have subsequently been reported in multiple later studies. It has been pointed out that the variable reproducibility of the data may be attributable to methodological artefact, as the different arrays available in these studies may not have included probes for identical genes, or may have included probes for different gene transcripts (Katsel, Davis et al. 2005c). Indeed it has been shown that different micro-arrays may produce different results even when the same samples are compared across platforms (Hollingshead, Lewis et al. 2005).

Table 1.1 Summary of expression results for OMR genes in schizophrenia.

Table summarising the reported altered expression of oligodendrocyte and myelin related genes in schizophrenia. Showing the study, number of patients and controls studied, the brain regions studied and the OMR genes displaying altered expression.

Author	Patients / Controls	Brain region studied	OMR genes displaying altered expression
Hakak et al, 2001	12 / 12	Dorso-lateral pre-frontal cortex	MAL, CNP, MAG, TF, GSN, ErbB3 (down-regulated)
Tkachev et al, 2003	15 / 15	Pre-frontal cortex	MAG, TF, ErbB3, MBP, MOBP, CLDN11, MOG, PLP1, Olig1, Olig2, Sox 10 (down-regulated) MPZL1, GFAP (up-regulated)
Aston et al, 2004	12 / 14	Temporal gyrus	MAG, PLLP, PLP1, ErbB3, MAL, TF, CNP (down-regulated)
Sugai et al, 2004	6 / 6	Dorso-lateral pre-frontal cortex	MOBP, MBP, MAL (down-regulated)
Iwamoto et al, 2005	13 / 15	Pre-frontal cortex	SOX10, Olig2, MAG, PLP1, MOBP, MOG (down-regulated)
Katsel et al, 2005	30 / 25	Superior temporal cortex, Cingulate cortex, Hippocampus	QKI, TF, MAG, CLDN5, PMP22, CNP, Sox 10, GSN, ErbB3, MAL, MOG, CLDN11, Olig2 (down-regulated
Dracheva et al, 2006	13 / 13	Anterior cingulate cortex, Hippocampus	MAG, CNP, SOX10, CLDN11, PMP22 (down-regulated)

1.5.2. Neuroproteomic analysis:

There have currently not been many studies which have examined oligodendrocyte or myelin related genes at the protein level in schizophrenic brain. However, Flynn and colleagues (Flynn, Lang et al. 2003) reported that CNP protein levels were significantly reduced in the prefrontal cortex of schizophrenic brain. The reduction of CNP protein in schizophrenic brain has been subsequently replicated by two further independent groups in the prefrontal cortex (Prabakaran, Swatton et al. 2004) and also the anterior cingulate cortex (Dracheva, Davis et al. 2006).

There is also further evidence for the potential involvement of OMR genes in schizophrenia pathology provided by the results of animal knockout studies. Lappe-Siefke and colleagues (Lappe-Siefke, Goebbels et al. 2003) produced a CNP knockout mouse which displayed both oligodendrocyte and axonal defects. The authors additionally reported that these CNP null mice displayed features of central nervous system pathology (reduced brain volume, enlarged ventricles and corpus callosum atrophy) consistent with some pathological features observed in schizophrenia (Harrison 1999).

1.5.3. Imaging techniques provide evidence of white matter abnormalities:

Following the development of imaging techniques like magnetic resonance imaging (MRI), there is now a strong basis of evidence for white matter abnormalities in schizophrenia.

A meta-analysis of the MRI studies up till 1998 suggests reduction in the overall white matter volume in schizophrenic brain (Wright, Rabe-Hesketh et al. 2000). Reduction of white matter in the prefrontal cortex of schizophrenics has been reported (Buchanan, Vladar et al. 1998; Sanfilipo, Lafargue et al. 2000; Sigmundsson, Suckling et al. 2001), and this has additionally been shown to be associated with the negative symptoms of the disease. The use of MRI to study white matter changes has recently been complemented by the advent of other imaging techniques, including magnetic transfer imaging (MTI) and diffusion tensor imaging (DTI) which provides greater levels of resolution and detail on the white matter changes present.

MTI allows the indirect visualisation of protons bound to macromolecular structures, such as myelin and cell membranes in white matter (Foong, Symms et al.

2001). These structures are essentially invisible to conventional MRI because of their very short relaxation times (Kubicki, Park et al. 2005). The exchange of magnetization between bound protons and free water is represented by the magnetization transfer ratio (MTR), which provides a quantitative measure of both the amount and structural integrity of the macromolecules in the tissue (Kubicki, Park et al. 2005). Thus MTI does not visualise myelin directly, but is instead interpreted as providing an index of the integrity of myelin. One of the earliest MTI studies of schizophrenia demonstrated a decrease in the myelin or axonal membrane integrity in the temporal regions of schizophrenic brain, as well as other widespread cortical regions (Foong, Symms et al. 2001). A more recent MTI study looking at first episode schizophrenic patients found reduced myelin integrity in the prefrontal cortex and other regions of the frontal lobe (Bagary, Symms et al. 2003). The existence of white matter integrity deficits at the onset of schizophrenia suggests that the abnormalities involving myelin may be of aetiological relevance rather than a consequence of disease progression or medication.

DTI measures the sum of the vectors of water diffusion. This sum represents the anisotropy of the particular tissue, which in turn is thought to represent the coherence of structures within that particular tissue or region. Tissue that is uniform and aligned along a similar axis will have a high anisotropy, whereas tissue that is disorganised and less uniform will have a lower anisotropy. DTI data are generally thought to reflect both the coherence of white matter bundles and also the degree to which myelin is wrapped around axons, and therefore normal axons and myelin which are aligned in the same plane have a higher anisotropy than disorganised structures which are aligned in multiple planes. The observation of lower anisotropy within white matter is thought to be due either to a loss of coherence in white matter fibre tracts, to changes in the number/density of interconnecting fibre tracts or to changes in myelination.

Recently there have been around 18 DTI studies on schizophrenia with most reporting white matter abnormalities (Kubicki, McCarley et al. 2005). Decreased anisotropy along white matter tracts have been found in numerous regions of schizophrenic brain including the prefrontal cortex, temporal lobes and the various regions connecting these regions such as the anterior cingulum (Lim, Hedehus et al. 1999; Agartz, Andersson et al. 2001; Wang, Sun et al. 2004). Although not providing conclusive proof, the decreases in anisotropy seen in many regions of the schizophrenic brain are consistent with the theory that a myelin and/or oligodendrocyte abnormality plays a part in schizophrenia aetiology, probably leading to a functional or structural deficit in connectivity.

Kubicki and colleagues (Kubicki, Park et al. 2005) performed both DTI and MTI analysis on the same brain regions of schizophrenic patients to try and define better what sort of abnormalities were present in schizophrenia. DTI displayed decreased anisotropy of white matter in widespread brain regions, with MTI analysis also displaying myelination alterations in many of the same brain regions. However, the authors also reported regions that were altered only in DTI analysis or MTI analysis. This was taken to be an indication that the white matter abnormalities observed in schizophrenia are likely to be caused by a combination of abnormal coherence, disorganisation of white matter fibre tracts and also myelin disruption, rather than just one type of abnormality. Whilst, DTI and MTI studies of schizophrenic brain are consistent with the presence of myelin abnormalities in schizophrenia, it should be noted that the findings do not definitively prove the existence of such abnormalities. The reduced anisotropy observed in schizophrenic brain can also be explained by other processes such as changes in the number/density of interconnecting fibre tracts.

The majority of DTI studies have looked at chronic schizophrenic patients making it difficult to exclude the possibility that reduced anisotropy might result from ongoing medication and illness progress. However, recently several studies have tried to resolve this problem. Szeszko and colleagues (Szeszko, Ardekani et al. 2005) looking at first episode schizophrenic patients reported reduced white matter anisotropy in the frontal and temporal gyrus of schizophrenic brain, suggesting that the white matter abnormalities observed in schizophrenics are present at the onset of illness and are not the result of any secondary factors. A further study (Kumra, Ashtari et al. 2004) looking at early onset schizophrenia in adolescents reported lower anisotropy in frontal lobe white matter. This would suggest that the white matter abnormalities are present in schizophrenics at a young age and are not due to age or illness progression.

There are a few further interesting findings from DTI studies that have been reported, such as the correlations between white matter anisotropy and both negative disease symptoms and the level of performance when undertaking executive function and memory tasks (Kubicki, McCarley et al. 2005). Finally there have been decreases in anisotropy reported in several demyelinating diseases such as multiple sclerosis and various leukodystrophies, suggesting that DTI does in fact provide a true indication of myelin integrity and further supporting the involvement of a myelination deficit in schizophrenia (Davis, Stewart et al. 2003).

1.5.4. Symptom overlaps with demyelinating disease:

If the evidence of white matter abnormalities observed in both schizophrenia and demyelinating diseases are due to similar defects in myelination one might expect to see some evidence of shared symptoms between the various diseases.

Metachromatic leukodystrophy (MLD) is a disease caused by an accumulation of sulfatides in the nervous system, which leads to widespread demyelination usually beginning in the frontal lobes of the brain. The rare late-onset form of MLD which develops during adolescence to early adulthood often presents with acute symptoms of schizophrenia, before leading to progressive mental deterioration (Davis, Stewart et al. 2003; Stewart and Davis 2004). Psychotic symptoms have been shown to occur in about 53% of late-onset cases of MLD, and possibly more importantly psychotic symptoms are only reported in cases presenting in adolescence and early adulthood which is the developmental period during which schizophrenia usually presents (Hyde, Ziegler et al. 1992). The psychotic symptoms in MLD patients are only present during the early stages of the disease when the demyelination is restricted to the frontal lobes. As the demyelination spreads through the brain the psychotic symptoms disappear and are then replaced by neurological deficits and dementia (Hyde, Ziegler et al. 1992).

Multiple sclerosis (MS) is characterised by patches of demyelination throughout the central nervous system. The diverse locations lead to a wide variety of disease presentations and symptoms. Psychiatric symptoms are frequently reported in MS, although psychotic symptoms are rarely observed. The most interesting finding with regard to schizophrenia is the relationship between the location of demyelination in MS brain and the associated symptoms. For instance an MRI study (Honer, Hurwitz et al. 1987) showed that although MS patients with psychiatric symptoms had the same number of overall lesions as MS patients with no psychiatric symptoms, their lesions were more abundant in the temporal lobe of the brain. A further study (Reischies, Baum et al. 1988) demonstrated that the severity of the psychotic symptoms observed in MS patients was correlated with the amount of demyelination in the frontal lobe rather than any other brain regions.

The differences between MLD and MS symptoms are of relevance to schizophrenia. The timing of the myelination defect appears to be important. In MLD if the disease presents in early adulthood or adolescence, the same age of onset period as schizophrenia, then the MLD patients often present with schizophrenic symptoms, indeed they are often misdiagnosed as schizophrenic (Hyde, Ziegler et al. 1992). However, MS has a later age of onset than schizophrenia and is rarely seen to present with psychotic symptoms (Davis, Stewart et al. 2003). Secondly, the location of the defect in myelination appears to be important. In MLD cases displaying psychotic symptoms the demyelination begins in the frontal lobes, however, when the demyelination expands to the rest of the brain there is an accompanied loss of psychosis (Hyde, Ziegler et al. 1992). MS on the other hand rarely presents with psychotic symptoms, but on the rare occasions that it does the majority of the demyelination lesions are found in the frontal and temporal lobes of the brain (Honer, Hurwitz et al. 1987; Reischies, Baum et al. 1988). These observations would suggest that defective connectivity in the frontal lobes of the brain caused by abnormalities in myelination is of relevance to the development of psychotic symptoms and possibly schizophrenia.

1.5.5. Age related white matter changes:

The involvement of a myelination abnormality in schizophrenia aetiology would be additionally supported if the brain areas that have been previously implicated in schizophrenia such as the frontal lobe underwent developmental myelination events at the same time as the onset of schizophrenia. The expansion of white matter and the myelination process actively proceeds from childhood until the early twenties (Giedd, Blumenthal et al. 1999). Connections between the cortex and hippocampus have been shown to be actively myelinating during childhood, adolescence and adulthood (Benes, Turtle et al. 1994), whilst the frontal and temporal lobes are the last cortical regions to complete the myelination process (Goldman-Rakic and Selemon 1997). It is interesting that myelination processes occurring during late adolescence to early adulthood, the normal age of onset in schizophrenia, are concentrated in the frontal and temporal lobes of the brain, both regions which have been implicated in the aetiology of schizophrenia. This observation supports the possible involvement of these brain regions along with a white matter abnormality in the aetiology of schizophrenia. However, the association between the timing of myelination in the brain regions associated with schizophrenia and the typical age of onset in schizophrenia can only be seen as speculative.

1.5.6. Evidence of myelin dysfunction and abnormalities in schizophrenia:

Direct examination of myelin sheaths in the brains of schizophrenia patients has demonstrated several ultrastructural abnormalities, including alterations to the myelinated fibres between the prefrontal cortex and caudate nucleus, along with a reduction in the compactness of the myelin sheath lamellae (Stewart and Davis 2004).

An electron microscopy study (Uranova, Orlovskaya et al. 2001) looking at post-mortem tissue from the prefrontal cortex and caudate nucleus of schizophrenic brain has demonstrated evidence of apoptosis and necrosis of oligodendroglial cells along with damaged myelin sheath lamellae, leading to the formation of concentric lamellar bodies (indicators of damaged myelinated fibres). The authors also reported that these dramatic changes in myelin ultrastructure could not be accounted for by the effects of any confounding factors such as post-mortem delay or neuroleptic medication.

Oligodendrocyte clustering and density have been reported to be significantly reduced in the white matter of the superior frontal gyrus of schizophrenic brain compared to controls (Hof, Haroutunian et al. 2002; Hof, Haroutunian et al. 2003). The authors of these two studies reported that oligodendrocyte clustering in schizophrenic subjects was 30% lower than in control brain, and proposed that this reduced oligodendrocyte organisation might contribute to the reduced white matter tract coherence observed in schizophrenia. The authors additionally reported that the density of oligodendrocytes was 28% lower in schizophrenics compared to controls. Finally the total number of oligodendrocytes present in the cortical layer III and white matter of Brodmans area 9 (superior frontal gyrus) was also shown to be significantly reduced in schizophrenic brain (average reduction of 27%) compared to matched control subjects (Hof, Haroutunian et al. 2003).

There have also been further reports of reduced oligodendrocyte density in other regions of the brain. Uranova and colleagues (Uranova, Vostrikov et al. 2004) demonstrated a significant reduction (average reduction of 25%) of oligodendrocyte density in the prefrontal cortex of schizophrenic brain samples compared to controls. Stark and colleagues (Stark, Uylings et al. 2004) also reported a significant reduction

(33%) in the number of glial cells observed in the anterior cingulate cortex of schizophrenic brain compared to healthy controls, although it should be noted that not all of the glial cells are oligodendrocytes.

Finally in the anterior thalamic nucleus of schizophrenic brain, Byne and colleagues (Byne, Kidkardnee et al. 2006) reported an average reduction of 17% in the volume of both neurons and oligodendrocytes. The authors also reported a significant 40% reduction in the total number of oligodendrocytes along with a significantly reduced ratio of oligodendrocytes to neurons in the post-mortem schizophrenic brain studied.

1.5.7. Association analysis of oligodendrocyte and myelin related genes:

As described above there is now evidence consistent with the presence of oligodendrocyte and/or myelination abnormalities in schizophrenic patients. However, the molecular mechanisms responsible for these defects are still uncertain, and it is not known whether the abnormalities are a cause or effect of schizophrenia. Whilst, the neuropathological and imaging studies support the presence of oligodendrocyte and/or myelin abnormalities in schizophrenia, the neuropathological evidence is still controversial and the imaging findings may be explained by several other processes. Evidence from micro-array studies does indicate the presence of altered OMR gene expression in schizophrenia, but again does not reveal whether these are primary or secondary effects. Genetic association studies of OMR genes are required to try and resolve whether the observed oligodendrocyte/myelination abnormalities in schizophrenia play a primary aetiological role in the disorder or whether they are secondary effects as a consequence of the disease. Recently researchers have started to perform association analyses targeting OMR genes to determine if they are significantly associated with susceptibility for the disease.

One gene which has received strong support for involvement in schizophrenia aetiology is neuregulin 1 (NRG1) (Stefansson, Sigurdsson et al. 2002; Stefansson, Sarginson et al. 2003; Williams, Preece et al. 2003; Li, Collier et al. 2006). Neuregulin has been shown to have numerous functions, including influences on nervous system and oligodendrocyte development, synapse formation, neuron-glial signalling, synaptic transmission and myelination (Calaora, Rogister et al. 2001; Corfas, Roy et al. 2004; Harrison and Law 2006). In vitro experiments studies have shown that NRG1 induces the differentiation of oligodendrocyte precursors (Vartanian, Fischbach et al. 1999), as well as inducing the proliferation and survival of oligodendrocyte precursors (Canoll, Musacchio et al. 1996; Fernandez, Tang et al. 2000). Recently it has been reported that one of the associated SNPs in NRG1 is predicted to lead to a gain/loss of a myelin transcription factor binding site (Law, Lipska et al. 2006).

The OMR gene which has received the most attention after NRG1 is reticulon 4 (RTN4, also referred to as NOGO). RTN4 is a myelin associated protein that has been shown to inhibit the growth of neurons and nerve terminals. Interest in RTN4 was first generated by Novak and colleagues (Novak, Kim et al. 2002) who reported that RTN4 mRNA was up-regulated in schizophrenic brain, and additionally reported a significant association between a genetic variant in the 3'UTR of the gene and schizophrenia susceptibility (this gene is discussed more thoroughly in chapter 4).

Another OMR gene where there has been more than a single report of association to schizophrenia susceptibility is myelin associated glycoprotein (MAG). MAG has been shown to be play important roles in myelination such as maintaining myelin-axonal contacts and the structure of myelin sheaths (Gravel, Peterson et al. 1996; Hakak, Walker et al. 2001) and has been reported to be down-regulated in schizophrenia (see section 1.5.1). Association between MAG and schizophrenia has been reported in family and case-control based studies at both the individual and haplotype level (Wan, Yang et al. 2005; Yang, Qin et al. 2005). However, it should be noted that although both studies were performed in Han Chinese populations there are inconsistencies in the patterns of the association findings, and until the results are confirmed by further replication studies the association should be viewed with caution.

A further gene that has recently been reported to be associated with susceptibility to schizophrenia is the quaking homolog KH domain RNA-binding gene (QKI) (Aberg, Saetre et al. 2006a; Aberg, Saetre et al. 2006b). The authors also demonstrated that two of the QKI mRNA splice variants are down-regulated in schizophrenic patients a finding which supports the findings of a previous micro-array study which reported that QKI was the most robustly down-regulated gene in their schizophrenic brain samples (Katsel, Davis et al. 2005b).

There has also been association reported between other OMR genes and schizophrenia such as: oligodendrocyte lineage transcription factor 2 (Olig2) (Georgieva, Moskvina et al. 2006), myelin oligodendrocyte glycoprotein (MOG) (Liu, Qin et al. 2005), proteolipid protein (PLP1) (Qin, Gao et al. 2005), claudin 5 (Sun, Wei et al. 2004) and gelsolin (GSN) (Xi, Qin et al. 2004). However the results have yet to be replicated in independent samples.

1.5.8. Possible role of oligodendrocyte and myelin genes in schizophrenia:

Although the evidence for association between altered OMR function and schizophrenia is growing, the molecular basis for the putative altered function and how this might contribute to schizophrenia is still unknown. It is not yet clear whether the observed oligodendrocyte and myelin abnormalities in schizophrenia are a primary or secondary deficit of the disease.

One of the hypotheses of schizophrenia pathology is that the disease is caused by a disruption in connectivity between neurons, cortical areas and sub-areas in the central nervous system (Friston 1998). However, the basis of this potentially abnormal connectivity remains to be discovered. It has been proposed that the altered connectivity could be due to disruptions in either the 'micro-connectivity' (connections between adjacent cells) and/or disruptions in the 'macro-connectivity' (connections between different distinct brain regions) of schizophrenic brain. The presence of a white matter abnormality in schizophrenic brain would complement the hypothesis that schizophrenia is a disease caused by disrupted connectivity, as alterations to the myelination process or the oligodendrocyte structure of the brain could have profound effects on the connectivity of the brain at either the cellular or cortical level.

Thus I decided to focus this series of studies on genes involved with the oligodendrocyte and myelination processes and their potential involvement with increased susceptibility for schizophrenia.

1.6. Strategy of the study:

The OMR genes studied were primarily selected based upon their reported altered expression in micro-array studies of schizophrenic brain, supported by positional evidence from linkage studies and/or the genes putative functional role in the neurobiology of schizophrenia.

At the time this project was initiated two micro-array studies had been published reporting the altered expression of genes in schizophrenic brain (Hakak, Walker et al. 2001; Mimmack, Ryan et al. 2002). The first of these studies (Hakak, Walker et al. 2001) reported the down-regulation of a group of six genes related to oligodendrocyte and/or myelin function in the dorso-lateral prefrontal cortex of schizophrenic brain; myelin and lymphocyte protein (MAL), 2′, 3′-cyclic nucleotide 3′- phosphodiesterase (CNP), myelin associated glycoprotein (MAG), transferrin (TF), gelsolin (GSN) and neuregulin receptor tyrosine kinase 3 (ErbB3).

The second study (Mimmack, Ryan et al. 2002) reported up-regulation of several members of the apolipoprotein L (APOL) gene family in the prefrontal cortex of schizophrenic patients compared to matched controls, a consistent up-regulation of the APOL 1, 2 and 4 genes was observed. In a subsequent real-time PCR study the authors displayed a 2.6 fold up-regulation of APOL 1, a 2.4 fold up-regulation of APOL 2 and a 2.7 fold up-regulation of APOL 4 relative to control samples.

I therefore decided to begin this project by focusing my studies upon the genes displaying altered expression in these two micro-array studies. Altered gene expression can be attributable to either *cis-* or *trans*-acting effects. For the purposes of genetic association studies it is the former *cis*-acting sequence variants that are of interest, and I therefore, first sought to confirm whether the genes expression was influenced by *cis*-acting polymorphisms. Where the presence of *cis*-acting variants was confirmed, the genes were then subsequently screened for evidence of association with schizophrenia susceptibility.

All of the OMR genes subsequently studied as part of this project (RTN4, RTN4R, OMG, NGFR, GFAP, MOG and SOX 10) were also selected for investigation based primarily upon reports of their altered expression in schizophrenic brain, supported by evidence from positional, functional or association studies. The experimental approaches used to study these genes varied and is described in more detail in each of the results chapter.

Chapter 2. Materials and Methods:

2.1. DNA Samples:

Genomic DNA used in this study was derived from schizophrenia patients and blood donor control individuals. The case-control association sample consisted of 709 Caucasian subjects with schizophrenia from the United Kingdom and Ireland (483 males and 226 females) and 710 controls (477 males and 233 females). Control subjects were group matched to the cases for age, sex and ethnicity and were recruited from the National blood transfusion service (Wales). All patients had a consensus diagnosis of schizophrenia according to DSM-IV criteria made by two independent raters following a semi-structured interview by trained psychiatrists or psychologists using the Schedules for Clinical Assessment in Neuropsychiatry interview (Wing, Babor et al. 1990) and review of case records. High levels of reliability (kappa >0.8) were achieved between raters for diagnoses. All cases were screened to exclude substance-induced psychotic disorder or psychosis due to a general medical condition. The mean age at first psychiatric contact was 23.6 (SD 7.7) and the mean age at ascertainment was 41.8 (SD 13.5). The blood donor controls were not specifically screened for psychiatric illness but individuals were not taking regular prescribed medications. In neither country are blood donors remunerated even for expenses and as a result, blood donors are not enriched for substance abusers or socially disadvantaged people who may have relatively high rates of psychosis. Ethics committee approval was obtained in all regions where patients were recruited and informed written consent was obtained from all participants. Blood samples were taken and DNA extracted by other members of the Cardiff schizophrenia research group.

2.1.2. Mutation screening sample:

A screening sample consisting of fourteen unrelated Caucasian schizophrenia patients (chromosomes =28) was used when screening for polymorphisms. All subjects were from the UK, met DSM-IV criteria for schizophrenia and also had at least one affected sibling. For the DHPLC mutation screening set, power was calculated using the equation: $1-(1-f)^n$, where f = allele frequency and n = number of chromosomes examined. Thus, a screening set of 28 chromosomes provided a power of 80% to detect polymorphisms with a minor allele frequency of 0.05, and a power of 95% to detect polymorphisms with a minor allele frequency of 0.1.

2.1.3. Pooled DNA samples:

The case control sample was made up of four separate stages the first three each containing 184 cases and 184 controls and the fourth containing 157 cases and 158 controls giving a total number of 709 cases and 710 controls. The four stages of the sample were also available in a single pool. Equal volumes of quantified DNA for each of the samples within the four case and control stages were mixed together to create case and control pools for each stage. The total association sample provides a power of 94% to detect a variant with a minor allele frequency of 0.1, conferring an odds ratio (OR) of 1.5, at the 0.05 significance level. The pools were constructed by Nadine Norton as described (Norton, Williams et al. 2002). Briefly, the concentration of the DNA samples being used to construct the pools was measured using the PicoGreen Quantification Reagent (Invitrogen, Paisley) in a Fluoroscan Ascent fluorimeter (Labsystems, Cambridge) and then diluted to 4ng/ul (+/- 0.5ng/ul). Pools were then constructed by combining equal volumes of each sample. All pools were then validated by genotyping several SNPs both in the pools and individually in the corresponding cases and controls that made up the pools.

2.1.4. DNA Extraction:

Genomic DNA was extracted from either blood, mouthwash or from sections of post-mortem brain tissue using standard laboratory procedures. The DNA concentration of each sample was assayed by taking spectrophotometer readings of DNA aliquots. Samples were subsequently stored at -20°C until required. I was not directly involved in the diagnosis or collection of patients or the extraction of DNA samples.

2.2. Characterisation of gene sequences:

With the emergence of the human genome sequence, there is now a multitude of information resources available for any given gene. This information is available on the Internet through various sources such as the UCSC Genome browser (http://genome.ucsc.edu/), Ensembl (http://www.ensembl.org/) and the National Centre for Biotechnology Information (NCBI) databases (http://www.ncbi.nlm.nih.gov/). Therefore these sites were the first point of reference for all candidate genes studied in this project.

The cDNA sequence for each gene was initially obtained from either the NCBI website (http://www.ncbi.nlm.nih.gov), or the UCSC Genome browser (http://genome.ucsc.edu/). Intron/Exon boundaries were determined by alignment between cDNA sequence and genomic clones in the Genbank databases using the BLAST search program (http://www.ncbi.nlm.nih.gov/BLAST/Blast.cgi), or the BLAT sequence alignment software (http://genome.ucsc.edu/cgi-bin/hgBlat).

2.3. Polymerase Chain Reaction (PCR):

The polymerase chain reaction (PCR) is a technique used to rapidly amplify large quantities of a specific target section of genomic DNA, and is the initial step in all of the methods used in this project. The process requires two oligonucleotide primers which are specific to the target sequence being amplified. This means they must be complementary to the sequence directly flanking the target region and are generally about 15-25 nucleotides in length.

2.3.1. Primer design:

PCR fragments for this project were designed to span exons, untranslated regions (UTR's), 5' and 3' flanking regions and putative promoters. All of the primers used in the project were designed using the Internet program Primer 3.0 (www.genome.wi.mit/edu/cgi-bin/primer/primer3_www.cgi). This package checks the primers for primer pair compatibility, G-C ratio, primer dimers and hairpin formation. All of these things can affect the PCR efficiency and the quality of the end product produced. Primer 3.0 also provides a suggested optimum annealing temperature for each of the primer pairs. The designed primers were then ordered from Sigma-Genosys (http://orders.sigma-genosys.eu.com) (Haverhill, UK).

2.3.2. PCR conditions:

All PCR's were performed on MJ thermocyclers (MJ Research) under standard conditions of 1x PCR buffer (Qiagen), 1.5 mM $MgCl_2$, 0.25mM of dNTPs, 0.8pmol of each primer, 0.5 units of Hotstart Taq polymerase (Qiagen) and 16ng of genomic DNA. The following PCR mixture for a 12μl reaction was used as a starting point for the optimisation of all the PCR reactions:

Reagent	Volume	Final Concentration / Amount
Genomic DNA template (4ng/μl)	4.0μl	16ng
Forward Primer (10pmol/μl)	1.0μl	0.83pmol
Reverse Primer (10pmol/μl)	1.0μl	0.83pmol
2.5mM dNTP mix	1.2μl	0.25mM
10x PCR Buffer (15 mM $MgCl_2$.)	1.2μl	1x Buffer: 1.5 mM $MgCl_2$
Hot Start Taq Polymerase (5u/μl)	0.1μl	0.5 units
Distilled deionised PCR water	Variable	
DMSO (Optional)	0.6-1.2μl	5-10% of total rxn volume
Final Volume	12μl	

Each of the different PCR reactions was optimised to find the most effective temperatures and conditions for unique amplification of the various fragments. Initially a genomic control sample was amplified at a gradient (range) of different annealing temperatures and the conditions generating the best product were selected for use in further experiments. Cycling conditions were one cycle at 95°C for 15 mins followed by 35 cycles of 95°C 30 secs, appropriate annealing temperature 30 secs, 72°C 45 secs, and a final extension step at 72°C for 10 mins. A negative control containing water instead of genomic template was run with each PCR reaction to ensure reagents were not contaminated with genomic template. The PCR products were then separated by electrophoresis on agarose gels stained with ethidium bromide to check the product size and absence of amplification in negative controls.

2.4. Agarose Gel Electrophoresis:

Following PCR amplification, the products were checked to ensure that amplification had been successful. This was done by running the products on a 1.5-3.0% Agarose gel. The percentage of the gels varied depending on the fragment size and the required resolution of the product bands. (As the fragment size decreases, the gel percentage should increase to maintain optimal resolution.) This is true because at a higher percentage the small fragments will run more slowly through the gel and hence stay in a tighter band and give better resolution. For a 1.5% gel, 1.5g of agarose (standard grade, Sigma-Aldrich) was added to 0.5x TBE (50mM Tris pH 7.4, 50mM Boric acid ph 7.4, 10mM EDTA pH 8.0) buffer (ultra pure electrophoretic sequencing grade, National diagnostics) to give a final volume of 100ml. The mixture was heated until all the agarose was dissolved and then 1.5μl ethidium bromide solution (10mg/ml) was added. The gel was poured into a gel former and the appropriate well formers were placed in the gel. The gel was then allowed to set at room temperature and the well combs removed carefully. The gel was then placed in 0.5 x TBE buffer. A 5-8μl aliquot of the PCR product, together with 3-5μl of a loading buffer (6x loading buffer: 15% Ficoll, 0.25% Bromophenol blue, 0.25% Xylene cyanol in water) was then placed in the wells and the gel run for about 30 minutes at 100 V in a gel tank. The DNA fragments were always run alongside a 1kb DNA ladder (Invitrogen, Paisley UK) to confirm that the products were of the correct size. The gel was then visualised on a UV transilluminator to check product quantity, quality and specificity. Photographs were taken using a Kodak Electrophoresis Gel documentation and analysis system.

2.5. Post-PCR Purification:

Several of the techniques used in this project require the PCR products used to be purified of any unincorporated oligonucleotide primers and dNTP's. To achieve this, an enzyme digestion reaction was used. One unit of shrimp alkaline phosphatase and exonuclease I (Amersham Life sciences) was added to 12μl of PCR product, and then incubated on a MJ thermocycler at 37°C for 1 hour followed by 20 minutes at 85°C in order to inactivate the two enzymes. The shrimp alkaline phosphatase dephosphorylates any excess dNTP's rendering them inert whilst the exonuclease I digests remaining single stranded primers.

2.6. Denaturing High Performance Liquid Chromatography:

The identification of DNA sequence variation is one of the key steps in genomic analysis. There are numerous techniques available to do this. One of the most widely used methods available at the time I started my thesis for comparative DNA analysis was called Denaturing High Performance Liquid Chromatography (DHPLC). This is a highly sensitive and specific technique for detecting new or unknown sequence variation. The method depends upon the detection of heteroduplexes in PCR products by ion-pair reverse-phase high-performance liquid chromatography, under partially denaturing conditions which can be modified by the temperature at which the HPLC column is maintained (O'Donovan, Oefner et al. 1998). The sensitivity of the method is highly dependent on the temperature at which the analysis is undertaken. The optimal temperature for DHPLC was predicted using DHPLC Melt (Jones, Austin et al. 1999). The reaction is temperature sensitive because different DNA sequences will have unique compositions of bases and hence a varying amount of hydrogen bonds. These bonds require different levels of energy to break them and hence the sequences will have unique melting temperatures.

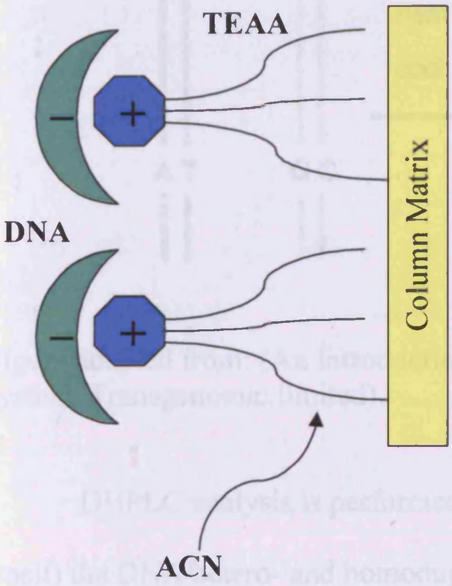
DHPLC offers significant advantages for high throughput screening of single nucleotide polymorphisms (SNPs) and short tandem repeats (STRs) in animal, bacterial and plant genomes. The procedure is almost totally automated which eliminates time and labour in gel preparation, running and analysis. A further advantage is that the PCR products are loaded directly without the need for purification. This again saves time and labour, but also saves money that would have been spent on purification. The technique is fast, cost effective and reliably detects single base changes, as well as small insertions and deletions (Jones, Austin et al. 1999). The method has proven to be reliable in several studies and the protocols used by me have been shown to have high specificity and sensitivity (O'Donovan, Oefner et al. 1998; Jones, Austin et al. 1999).

2.6.1. Chemistry of DHPLC:

DNA molecules are negatively charged. However the coated polystyrene beads comprising the stationary phase of the chromatography (DNAsep) column is neutral and hydrophobic, so therefore DNA fragments cannot directly bind to the column. A bridging molecule is required to help the binding of the DNA to the stationary phase. This molecule is triethylammonium acetate (TEAA). Once the DNA is bound to the column it is only removed when the TEAA bridge is broken. This is achieved by the addition of an organic solvent Acetonitrile (ACN). The column is washed in a solution containing an increasing ACN gradient until the DNA's bond to the column is broken and the DNA is eluted. Therefore the buffers containing the TEAA and ACN are important factors in the DHPLC method. See figure 2.1 for an explanation of how TEAA binds the DNA fragments to the column.

Figure 2.1 Binding of DNA with the help of TEAA.

This figure shows how TEAA enables the DNA fragments to bind to the DHPLC column.



- The column matrix is electrically neutral and hydrophobic.
- TEAA has a positively charged head and a hydrophobic tail. This tail can therefore bind to the hydrophobic column.
- The positive head of the TEAA attracts the negative DNA fragments and binds to them.
- The bond between the TEAA and the column is only broken when ACN is added to the column.
- The ACN removes the DNA from the column allowing it to be eluted.

The other principle behind DHPLC is the formation of heteroduplexes. A sample which is heterozygous for a single nucleotide polymorphism (SNP) has a 1:1 ratio of each allele. Heating the PCR product to 95°C and then slowly cooling it hybridises the PCR products and forms a mixture of hetero- and homoduplexes. See figure 2.2 for an explanation of the production of heteroduplexes.

Figure 2.2 Heteroduplex formation.

This figure shows how the formation of heteroduplexes occurs. If an individual is heterozygous for a SNP then the PCR products derived from that individual will contain roughly equal numbers of molecules carrying each base at the polymorphic site of each allele, (in this case A and G). When the PCR products are denatured and then cooled slowly half the DNA forms homoduplexes and half will form heteroduplexes. This is where the DNA forms together with a corresponding strand from the other allele, (i.e A to C instead of A to T). Heteroduplexes are not as stable to heat denaturation as the homoduplexes. The significance of this will be explained later.

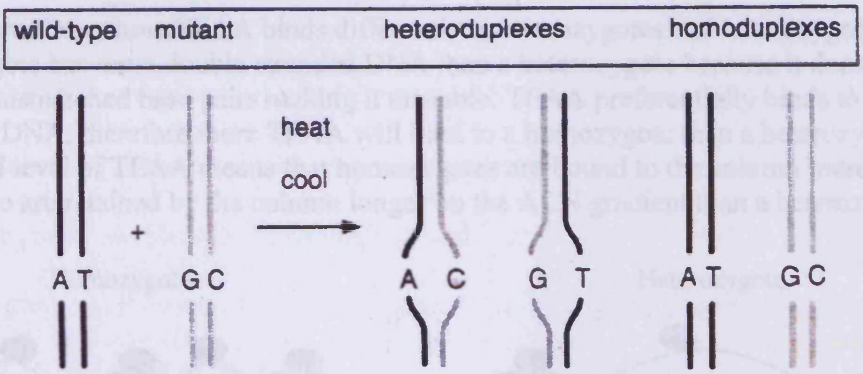
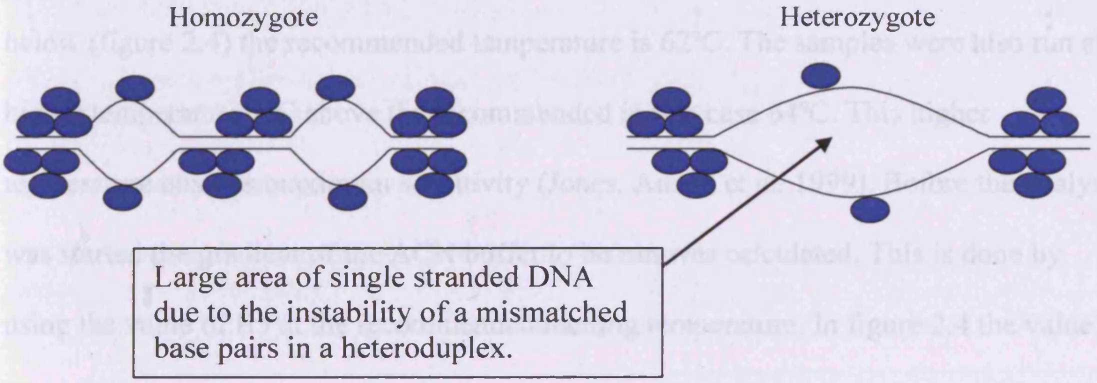


Figure adapted from: (An introduction to the Transgenomic WAVE DNA analysis system, Transgenomic limited).

DHPLC analysis is performed at a temperature sufficient to partially denature (melt) the DNA hetero- and homoduplexes. The hetero- and homoduplexes can now be resolved from each other by ion-pair reversed-phase liquid chromatography. Heteroduplexes denature more completely at a lower temperature than a corresponding homoduplex (see figure 2.3). TEAA preferentially binds to double stranded DNA because of the reduced negative charge in a single stranded molecule. Hence more TEAA molecules bind under partially denaturing conditions to a homoduplex than a heteroduplex (See figure 2.3). As there is a reduced amount of TEAA bound to the heteroduplexes they are bound less strongly to the DNAsep column than homoduplexes. This in turn means that they are eluted from the column at an earlier point in the ACN gradient because a lower ACN concentration is sufficient to disrupt the weaker TEAA-DNA bond. As the DNA is eluted from the column it is detected by a UV (260nm) detector and displayed as a chromatogram (see figure 2.5 for an example chromatogram).

Figure 2.3 Denaturation and TEAA binding.

This figure shows how TEAA binds differently to homozygotes and heterozygotes. A homozygote has more double stranded DNA than a heterozygote because it does not have a set of mismatched base pairs making it unstable. TEAA preferentially binds to double stranded DNA, therefore more TEAA will bind to a homozygote than a heterozygote. The increased level of TEAA means that homozygotes are bound to the column more securely and hence are retained by the column longer on the ACN gradient than a heterozygote.



Large area of single stranded DNA due to the instability of a mismatched base pairs in a heteroduplex.

2.6.2. DHPLC Method:

Polymorphism detection was performed by using DHPLC on a Wave™ DNA Fragment Analysis System (Transgenomic). The first step in the process was to amplify the desired fragment of genomic sequence using a PCR reaction. Before analysis the specific melting temperature for the fragment had to be determined. This was done by using the DHPLC melt program: (http://insertion.stanford.edu/melt.html). The genomic sequence for the fragment was pasted into the program which calculated the optimum melt temperatures for analysis as well the recommended buffer gradient. Experience in the laboratory suggests that a simple correction has to be performed on the suggested buffer gradient in order to determine the optimal gradient for use when analysing the fragment. The gradient of the buffers is important because it determines when the DNA fragments are washed off the DHPLC column. The higher the percentage of buffer B the more likely the DNA will be washed off.

DHPLC melt gives the recommended melting temperature for the fragment and the corresponding value for B3 at that temperature (see figure 2.4). The value of B3 is required in order to calculate the gradient of the ACN buffer to be run. In the example below (figure 2.4) the recommended temperature is 62°C. The samples were also run at a higher temperature 2°C above the recommended in this case 64°C. This higher temperature ensures maximum sensitivity (Jones, Austin et al. 1999). Before the analysis was started the gradient of the ACN buffer to be run was calculated. This is done by using the value of B3 at the recommended melting temperature. In figure 2.4 the value of B3 at 62°C is 63.8. In order to calculate the ACN buffer gradient I subtracted 2-3 off the B3 value to determine the percentage of the ACN buffer to use at the start of the gradient and then added 8 to this value in order to determine the percentage of the buffer that should be run at the end of the gradient. This gives you the increasing range of the ACN buffer concentrations required be run on the DNAsep column in order to elute the DNA fragment (in figure 2.4 the gradient is 60.8%-68.8%). When the reaction was started the percentage of ACN buffer being injected into the column is at the lowest value, and then as the reaction progresses the concentration increases until the upper value is reached at the end of the gradient. Normally the DNA fragment is eluted from the DNAsep column at the mid-point of the gradient and washed off to be detected by the UV detector.

See figure 2.4 below for a sample DHPLC melt output and calculation:

Figure 2.4: Example Melt output for a DNA fragment:

This figure shows an example melt output that was calculated for a DNA fragment. The fragment was subsequently analysed at the recommended temperature of 62°C and also at 64°C. The temperature of +2°C was run to maximise sensitivity (Jones, Austin et al. 1999).

Melting temperature of entire strand: 61.75°C

Temperature = 58° Celsius: 496-497
Temperature = 59° Celsius: 494-495,
Temperature = 60° Celsius: 490-493,
Temperature = 61° Celsius: 1-24,
Temperature = 62° Celsius: 25-489,

This section shows the different melting temperatures for the DNA fragment and which section of the fragment (bp) are melted at that temperature.

This sequence should be run at all of the following temperatures to ensure detection of all polymorphisms.

Recommended temperature: 62°

The following concentrations are rough recommendations only. Some adjustment may be necessary. Run from B1 to B2 in 0.5 minutes, then B2 to B3 in 3.0 minutes.

Temperature (°C)	**B1**	**B2**	**B3**
62	**50.0**	**58.4**	**63.8**
61	**50.0**	**59.8**	**65.2**
60	**50.0**	**61.2**	**66.6**
59	**50.0**	**62.5**	**67.9**

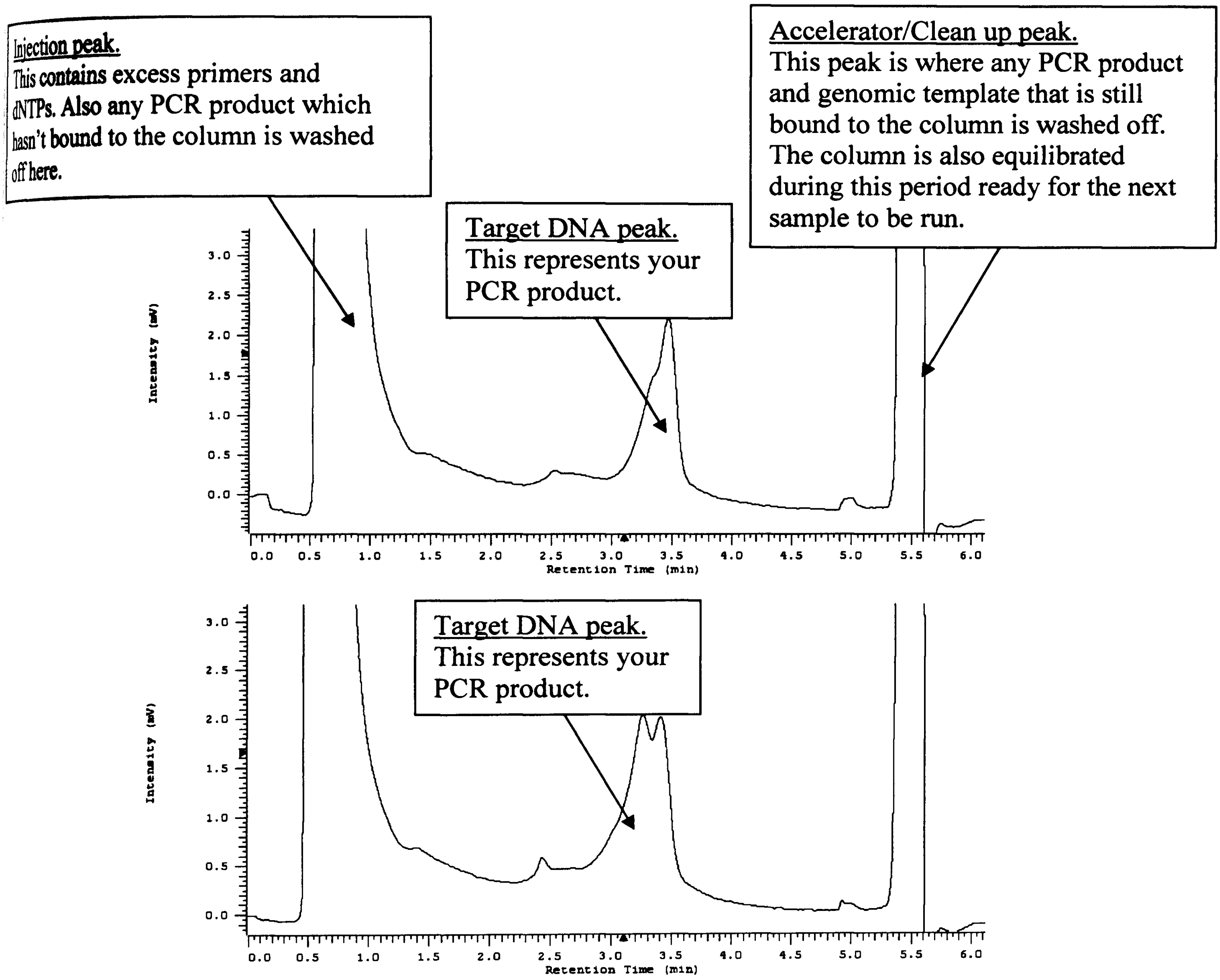
where B=0.1M TEAA, pH 7, 25% acetonitrile

This section provides the corresponding value of B3 for the fragments various melting temperatures enabling the calculation of the ACN buffer gradient required to elute the DNA from the DNAsep column.

Before the PCR samples are analysed by DHPLC the denaturation step was carried out, by heating the sample to 95°C for 5 minutes and then cooling to 45°C at a rate of 1°C/minute, to promote the formation of heteroduplexes. 5-10µl of the PCR product was then injected onto the DNAsep column using a Transgenomic WAVE DNA analysis system (Transgenomic Limited). The product was then eluted using the calculated acetonitrile gradient in 0.1M (TEAA) buffer pH7.0, at a constant flow rate of 0.9ml/min. The gradient was calculated as described above. The gradient is produced by mixing buffer A (0.1 M TEAA, 0.1 mM EDTA) and B (25% Acetonitrile, 0.1M TEAA). The eluted DNA fragments were detected by a UV detector set at an absorbance of 260nm, which then produced a chromatogram (see figure 2.5). Analysis took between 5 and 8 mins/sample including column equilibration and cleaning. Once the analysis is complete the chromatograms are interpreted visually in order to determine the presence of heteroduplexes.

Figure 2.5: Example of DHPLC chromatogram outputs.

This figure shows an example of two of the chromatograms produced when fragments were analysed. In the second chromatogram you can see a distinct change in the pattern of the peak which indicates heteroduplex formation.



Injection peak.
This contains excess primers and dNTPs. Also any PCR product which hasn't bound to the column is washed off here.

Target DNA peak.
This represents your PCR product.

Accelerator/Clean up peak.
This peak is where any PCR product and genomic template that is still bound to the column is washed off. The column is also equilibrated during this period ready for the next sample to be run.

Target DNA peak.
This represents your PCR product.

If a possible heteroduplex was found, the PCR products representing a heteroduplex and a homoduplex (normally at least two of each type) were sequenced in order to determine the sequence variation present in the fragment.

2.7. DNA Sequencing:

Automated sequencing was carried out using an ABI 3100 automated sequencer, and the Big Dye^{TM} Terminator Cycle Sequencing Ready reaction Kit (PE Biosystems.) This is based upon Sanger dideoxynucleotide sequencing, with the four ddNTPs being labelled with four different fluorophors. Each fluorophor has a different emission wavelength when excited by the laser in the ABI 3100, which allows all four termination reactions to be performed in a single tube. The relative size of each terminated sequencing product is then determined by electrophoresis in a capillary.

2.7.1. Pre-Sequencing Purification:

PCR reactions were cleaned prior to sequencing using one of two techniques. Either the QIAquick PCR purification kit (Qiagen), and a microcentrifuge, or a multiscreen PCR purification plate (Millipore) and a plate centrifuge. The processes remove residual dNTPs, primers and buffer from the PCR products by chromatography.

2.7.2. Sequencing Reaction:

Sequencing reactions were performed using the Big Dye Terminator cycle sequencing kit (Applied Biosystems) in a total reaction volume of 10μl. Two reactions were performed for each PCR fragment, one using the forward PCR primer and one the reverse primer. The reaction mix and thermocycling conditions used are detailed below:

Sequencing Reaction Mix:

Big Dye Terminator Mix	2μl
Big Dye Sequencing Buffer	2μl
Primer (F or R) (5 pmol)	1μl
Purified PCR product	5μl

Thermocycling conditions:

96°C	2 mins	
96°C	10 secs	25 cycles
50°C	5 secs	
60°C	4 mins	

2.7.3. Post-Sequencing Purification:

In order to generate high quality DNA sequence data it is important to remove unincorporated dye terminators, salt and any other impurities from the sequencing product prior to electrophoresis. This was achieved by gel filtration, using Sephadex (Sigma) loaded Multiscreen HV filtration plates (Millipore). Dry Sephadex G-50 resin was added to the 96 wells of the filtration plate using a column loader (Millipore). 300μl of ddH_2O was then added to each well to hydrate the resin and the plate left to stand at room temperature for 3 hours. Before use the filtration plate was then placed on top of a 96 well collection plate and spun in a centrifuge at 2500 rpm (900xg) for 5 min to remove any excess water and compact the columns. The individual sequencing products were diluted with 10μl of ddH_2O and added to the centre of the individual columns. The filtration plate was then placed on top of a fresh 96 well plate, secured in position and spun at 2500 rpm (900xg) for 5 min. The eluted sequencing products were then dried down using a speed vacuum before being prepared to run on the capillary sequencer.

2.7.4. Sequencing Sample Electrophoresis:

Purified sequencing products were prepared for electrophoresis by being resuspended in 10µl Hi-Di Formamide (Applied Biosystems). They were then subsequently run on an ABI PRISM® 3100 Genetic Analyser using a 36cm capillary array and Pop-6™ polymer. The resultant sequencing data was analysed using Sequence Analysis software (PE Applied Biosystems) before being imported into Sequencher (Gene Codes Corporation) for identification of polymorphisms.

2.8. Pooled Genotyping:

Where possible all the SNPs studied in this project were typed by pooled genotyping in order to assay their allele frequencies for association with schizophrenia. Pooled genotyping was performed by primer extension using the ABI PRISM® SNaPshot™ Multiplex kit (Applied Biosystems).

2.8.1. SNaPshot Genotyping:

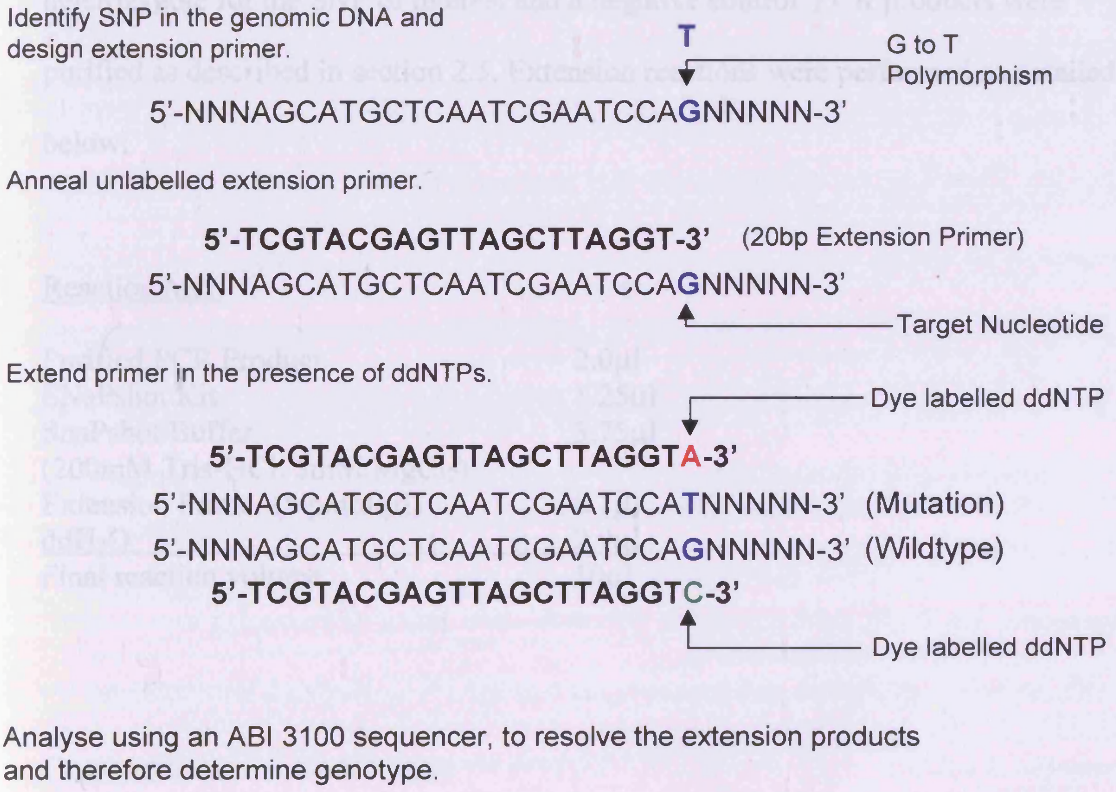
The SNaPshot procedure is based on the dideoxy single base extension of an unlabeled oligonucleotide primer. The primer binds to a complementary template in the presence of fluorescently labelled ddNTPs and AmpliTaq DNA Polymerase. The polymerase extends the unlabelled primer by one nucleotide, adding a single ddNTP to the 3′ end. The ddNTPs are each labelled with a unique fluorescent dye so that they can be distinguished after electrophoresis on an ABI 3100 sequencer (see table 2.1). Depending upon which base is present at the polymorphic site, a different ddNTP is incorporated at the end of the primer, therefore allowing the base at the polymorphism to be identified (See figure 2.6).

Table 2.1: Fluorescent ddNTPs used in SNaPshot genotyping:

ddNTP	Dye Label	Colour in Analysed Data
A	dR6G	Green
C	dTAMRA	Black
G	dR110	Blue
T (U)	dROX	Red

Figure 2.6: Diagram of SNaPshot extension reaction:

The 3' end of the extension primer should terminate directly adjacent to the polymorphic base as shown below. When you carry out the extension reaction the primer will anneal to the sequence and the ddNTPs will extend the primer by one base. Depending upon which of the alleles is represented in the sample you will get different coloured peaks when the data are analysed. In the case of a heterozygote you will get two different coloured peaks the same size but they will be roughly half the height of the normal homozygote peaks.


Identify SNP in the genomic DNA and design extension primer.

5'-NNNAGCATGCTCAATCGAATCCAGNNNNN-3'

Anneal unlabelled extension primer.

5'-TCGTACGAGTTAGCTTAGGT-3'

5'-NNNAGCATGCTCAATCGAATCCAGNNNNN-3'

Extend primer in the presence of ddNTPs.

5'-TCGTACGAGTTAGCTTAGGTA-3'

5'-NNNAGCATGCTCAATCGAATCCATNNNNN-3' (Mutation)

5'-NNNAGCATGCTCAATCGAATCCAGNNNNN-3' (Wildtype)

5'-TCGTACGAGTTAGCTTAGGTC-3'

Analyse using an ABI 3100 sequencer, to resolve the extension products and therefore determine genotype.

2.8.1.1 Extension Primer Design:

Single base primer extension reactions do not allow flexibility with respect to the location of the 3' end of the primer, with the exception of using either the sense or anti-sense strand of the sequence. It is therefore important to assess the sequence flanking the SNP in order to enable an optimal primer to be designed. All the extension primers used in this project were designed by me using software developed in house by Dobril Ivanov (http://m034.pc.uwcm.ac.uk/FP_Primer.html). This software assess the secondary structure of all the possible extension primers, determines the optimal length of the primer and whether to use the sense or anti-sense DNA strand.

2.8.1.2 SNaPshot Reaction:

Each DNA pool sample was PCR amplified in triplicate along with a heterozygote for the SNP of interest and a negative control. PCR products were purified as described in section 2.5. Extension reactions were performed as detailed below:

Reaction Mix:

Purified PCR Product	2.0μl
SNaPshot Kit	1.25μl
SnaPshot Buffer (200mM Tris-HCl. 5mM $MgCl_2$)	3.75μl
Extension Primer (5 pmol/μl)	0.1μl
ddH_2O	2.9μl
Final reaction volume	10μl

Thermocycling conditions:

94°C for 2 min
94°C for 5 sec
43°C for 5 sec } 25 Cycles
60°C for 5 sec
15°C forever

Following the extension reaction, samples were treated with 1 U of shrimp alkaline phosphatase at 37°C for 1 hour followed by 20 min at 85°C. Left untreated, the unincorporated fluorescent ddNTPs co-migrate with the labelled primers of interest. Removal of the 5' phosphoryl groups by phosphatase treatment alters the migration of the unincorporated fluorescent ddNTPs and thus prevents interference. In order to prepare the samples for electrophoresis 1µl of treated extension product was added to 9µl of HiDiformamide (Applied Biosystems).

The samples were run on the ABI PRISM® 3100 Genetic Analyser using a 36cm capillary and Pop4 polymer (Applied Biosystems). The ABI 3100 separates fragments by capillary electrophoresis. The capillaries contain the GeneScan Polymer (POP4). Samples are loaded by electrokinetic injection; the capillary and electrode are placed into the sample, voltage is applied, and the negatively charged DNA enters the capillary and migrates toward the positive electrode at the other end of the capillary. As the DNA fragments in the capillaries pass the detection cell they are illuminated by a laser. This light excites the attached dye labels causing them to fluoresce. The fluorescence that is emitted is separated by wavelength (each dye has a unique spectral profile/colour) and recorded by a charge-coupled device (CCD) camera. The fluorescence information is then converted into electronic data, which are processed by the Genescan Analysis v3.7 (Applied Biosystems) data collection software. The output were then analysed using the program Genotyper® v2.5 (Applied Biosystems).

The only element in the SNaPshot primer extension reaction that is altered for different polymorphisms is the amount of extension primer added to the reaction. This is of particular importance when genotyping DNA pools where, for accurate estimation of allele frequencies, peak heights representing the product corresponding to each allele (which are a direct count of the fluorescence emitted) should be above 1000 and below 6000 relative fluorescence intensity units. Genescan analysis software is unable to accurately quantify fluorescent signals outside this range. Analysis of the peak heights obtained in optimisation reactions determine the appropriate concentration of extension primer to use for further experiments. If the peak heights were above 1000 but below 6000, the extension primer concentration would remain at 5 pmol/µl for genotyping the pools. However if the peak heights were above 6000 or below 1000, the concentration was adjusted to correct the heights accordingly. The concentration (c) of the extension primer for the required peak height is calculated as $c = Y'/(Y/X)$ where, Y' = required peak height, Y = initial peak height, and X = initial primer concentration (Norton, Williams et al. 2002). The estimated allele frequency in each pool was determined as the mean of three primer extension assays after correction for the degree of unequal allelic representation detected in a heterozygote described as 'Correction for unequal allelic detection'.

2.8.1.3 Correction for Unequal Allelic Detection:

For a bi-allelic marker, the primer extension products for each allele are not equally represented. Possible explanations include differential PCR amplification of alleles and differential efficiencies of incorporation of the ddNTPs for each allele specific reaction (Norton, Williams et al. 2002). A further explanation is the unequal emission energies of the different fluorescent dyes. In order to allow for unequal representation of alleles the estimated allele frequencies from pools were corrected using the mean of the ratios obtained from the analyses of four heterozygotes using the equation:

Frequency of allele A $f(a) = A/(A+kB)$ where A and B are the peak heights of the primer extension products representing alleles A and B in pools, and k = is the mean of the four replicates of A/B ratios observed in a heterozygote (Hoogendoorn, Norton et al. 2000).

Based upon the number of subjects in each pool, the estimated allele frequencies can be used to estimate the number of each allele in the pool. The allele numbers in the cases and controls are then used to test for allelic association using a Chi-squared test under 1 degree of freedom. Any SNP showing an allelic association in the pools at the $p=0.05$ level was typed individually in the case control sample.

2.9. Allelic Expression Assay:

Genetic influences on gene expression may be either *cis*-acting or *trans*-acting. *Cis*-acting factors usually operate through effects on RNA transcription, RNA stability or translational efficiency and are usually located within the affected gene locus itself, although they can also be located some distance from the coding sequence. *Trans*-acting factors usually operate remotely on both copies of a gene, such as effects on transcription factors or hormones, and are not physically located at the influenced gene locus. Recently, it has been proposed that the expression of an identified disease gene can be used as an endophenotype for quantitative linkage analysis (expression quantitative trait loci; eQTL), and may therefore provide a method for mapping both *cis*- and *trans*-acting polymorphic loci which affect gene expression and modulate disease susceptibility. However, in order to assay the genes studied in this project for potential *cis*-acting influences on gene expression I used a highly quantitative allele specific expression assay developed in the host department (Bray, Buckland et al. 2003a; Bray, Buckland et al. 2003b).

The principle of the assay is that in heterozygous carriers of a cis-acting polymorphism that affects the transcription or stability of a species of mRNA, the quantity of mRNA originating from each gene copy will be unequal. The simplest method for distinguishing between mRNA molecules originating from each copy of a pair of autosomal genes is to use an exonic polymorphism as a copy specific tag (Bray, Buckland et al. 2003a). It is then possible to apply quantitative methods of allele discrimination to mRNA samples originating from individuals who are known to be heterozygous for the marker polymorphism in order to measure relative copy expression (Bray, Buckland et al. 2003a). Factors that can influence the measured amount of total expression of a specific gene (e.g. tissue preparation, mRNA quality, drug exposure, pre-agonal state, hormones, trans-effects secondary to regulatory polymorphisms in other genes) will, in the absence of a cis-acting allele-specific interaction, influence the amount of mRNA originating from each chromosome equally, and are therefore controlled for by this assay. While not allowing the measurement of total RNA abundance, the allelic expression assay does have the major advantage of allowing the detection of small polymorphic cis-acting influences even in the face of large trans-acting influences.

Post-mortem brain tissue was obtained from frontal, parietal or temporal cortex of 60 white European unrelated anonymized human adults of whom 50 were from the UK (the MRC London Neurodegenerative Diseases Brain Bank) and 10 were from Sweden (Department of Clinical Neuroscience, Karolinska Institute, Stockholm). All were free from a diagnosis of psychiatric or neurological disorder at the time of death. Genomic DNA and RNA were extracted from each individual tissue sample, with subsequent DNAse treatment of RNA. All of the DNA/RNA extraction and subsequent reverse transcription was kindly performed by Nick Bray (Bray, Buckland et al. 2003a).

Heterozygotes for the transcribed marker SNP were identified by genotyping genomic DNA from all subjects. Allelic expression was estimated as described by (Bray, Buckland et al. 2003a; Bray, Buckland et al. 2003b). Briefly, DNAse-treated RNA samples were subject to RT-PCR, primer extension with allele-specific dye terminator incorporation using the proprietary SNaPshot kit (PE Applied Biosystems), and the relative levels of the products representing each transcribed allele measured on a capillary sequencer. Samples were assayed using primers based on single exonic sequence, capable of amplifying either genomic DNA or cDNA. The cDNA samples were assayed alongside the corresponding heterozygote genomic DNA, which represents the assay output in the presence of a perfect 1:1 ratio of the two alleles. The ratio obtained from genomic DNA thus provided a correction factor for any inequalities in the efficiency of allelic representation specific to each assay (Hoogendoorn, Norton et al. 2000). Absence of genomic DNA in the RNA extracts was confirmed by including RNA samples that had not been reverse transcribed. Analysis of heterozygous samples was performed as two separate experiments. In each experiment, two cDNA samples were assayed in duplicate for each heterozygous individual (as two separate RT reactions), alongside the corresponding genomic DNA sample.

When assessing the results of the expression assays an arbitrary threshold of a 20% relative expression difference was used to indicate the presence of a *cis*-acting variant affecting gene expression as described by (Yan, Yuan et al. 2002; Bray, Buckland et al. 2003a). On the result graphs this threshold is represented by a ratio greater than 1:1.2 or lower than 1:0.83.

Calculation of power for allelic expression assays is based upon the binomial distribution, Hardy-Weinberg equilibrium at the regulatory SNP, and no LD with the marker SNP. The probability of an individual being homozygous at a putative regulatory locus with alleles in Hardy-Weinberg equilibrium is $p^2 + q^2$, where p and q are the two allele frequencies. The probability of n individuals all being homozygous (and therefore not detected in the allelic expression assay) for a regulatory polymorphism is then $(p^2 + q^2)^n$. This also applies for n individuals selected for heterozygosity at the marker locus if there is no relationship (ie LD) between the genotypes at each locus. The power to detect one heterozygote is then $1-(p^2 + q^2)^n$. If the marker and the regulatory variant are in LD, the power will be increased as a higher proportion of people selected for heterozygosity at the marker locus will also be heterozygous for the regulatory variant.

2.10 Individual Genotyping:

A variety of techniques were used to perform this including SNaPshot, Acycloprime™ Fluorescence Polarisation (FP), Amplifluor (Fluorescence Intensity) and Sequenom MALDI-TOF genotyping. These are described in greater detail below.

2.10.1 SNaPshot Genotyping:

In order to genotype individual DNA samples using the SNaPshot kit (Applied Biosystems), samples were individually amplified in a standard 12µl PCR reaction, and then processed in the same way as explained for pooled DNA samples (see section 2.8.1 above). When the data are imported into Genotyper for analysis the peaks are manually labelled in order to distinguish between the homozygotes of both alleles and heterozygotes. The results are then converted into a table and exported into an Excel workbook containing all of the sample information. The cases and controls are then separated according to their sample IDs, and the genotypes counted and formed into contingency tables for statistical analysis.

2.10.1.1 Multiplex SNaPshot genotyping:

Due to the high cost of genotyping samples individually using SNaPshot the genotyping assays are usually run as multiplexes. To type multiple SNPs simultaneously the assays need to be designed to ensure that none of the reaction products conflict with each other on the SNaPshot electrophorogram. To achieve this, the extension primers are designed to be different lengths ranging from 15-35bp long with a minimum gap between each SNP of around 10bp. It is also possible to assay two SNPs using extension primers that are the same length if the SNPs do not have any incorporated ddNTPs in common. It is also possible to use a mixture of forward and reverse extension primers in the assay allowing a further degree of flexibility. By using a combination of forward and reverse primers it is usually possible to design multiplexes containing 6-8 SNPs depending on the specific composition of the alleles present.

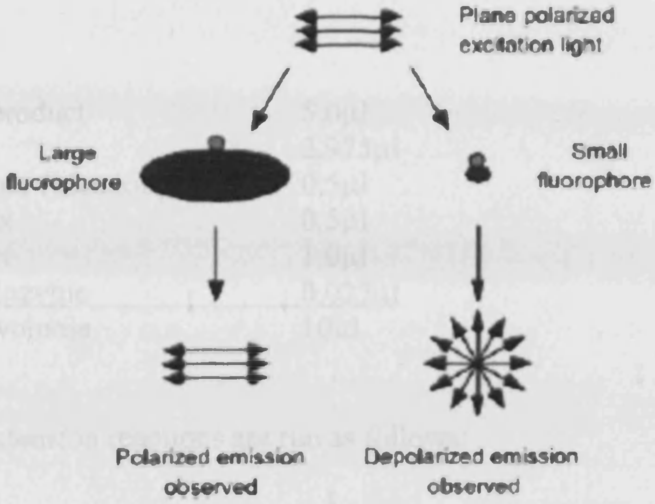
Sequences containing SNPs of interest are either amplified by PCR as separate reactions and then pooled together or are amplified as part of a multiplex PCR reaction. The PCR product is then cleaned as described in section 2.5. The SNaPshot reaction is performed as described in section 2.8.1, except that multiple extension primers are added to the reaction. The assay is then run and analysed on the ABI 3100 genetic analyser along with the internal size standard (LIZ 120) (Applied biosystems). The data are then exported into Genotyper 2.5 (Applied biosystems) and the genotypes determined. It is important to optimise the SNaPshot reaction to ensure that the allele peak heights are within the acceptable range and that all of the alleles are distinguishable from each other on the electrophorogram as the output is not determined by size alone but also in part by conformation.

2.10.2 Acycloprime™ Fluorescence Polarisation (FP):

The Acycloprime-FP process is also based upon single base primer extension using one of two fluorescently labelled ddNTP terminators. The identity of this base is determined by the increased fluorescence polarization of its attached dye (Hsu, Chen et al. 2001; Kwok 2002). The principle is that variation in fluorescence polarization occurs between fluorescently labelled molecules of different molecular weight. A small molecule containing a fluorophore rotates and tumbles rapidly, whereas a large molecule rotates and spins more slowly. If a polarising light is used to excite the fluorphore, small molecules will emit light in all directions and the degree of polarisation will be lost. However, large molecules rotating slowly tend to emit more light in the plane of polarisation therefore maintaining high levels of fluorescence polarisation (see figure 2.7) (Kwok 2002).

Figure 2.7: Fluorescence polarisation.

Adapted from the Perkin Elmer website: (www.perkinelmer.com/lifesciences).



A fluorophore linked to a large molecule has slow motion and the emitted light remains polarised

A fluorophore linked to a small molecule has fast motion and the emitted light is depolarised

To perform the assay 4µl of each DNA sample was aliquoted into 96 well PCR plates (ABgene) and a standard 8µl PCR cocktail aliquoted into the PCR plates containing the DNA samples. The samples were then amplified using standard PCR cycling conditions (as described in section 2.3) and then cleaned to remove any unincorporated dNTPs and unused PCR primers using the standard SAP and EXO I PCR purification method (as described in section 2.5). Following clean up of the PCR, 5µl of the product was aliquoted into 96 well black PCR plates (ABgene). Each AcycloPrime FP reaction only uses a mix of two ddNTPs (Acyclo Terminator mix), it is therefore important that the correct ddNTP mix, corresponding to the two alleles of the SNP being genotyped is used. The ddNTP combinations used are G/A, G/C, G/T, C/A, C/T and A/T and are all labelled with R110/Tamra fluorescent dyes. The extension primer is diluted to 10pmol/µl and the number of reaction cycles is variable, with the majority of assays working at 10-25 cycles. For each individual reaction the following reaction mix is added to the PCR:

Reaction Mix:

Purified PCR product	5.0µl
ddH_2O	2.975µl
Extension Primer (10pmol/µl)	0.5µl
Terminator Mix	0.5µl
Reaction Buffer	1.0µl
Acycloprime Enzyme	0.025µl
Total reaction volume	10µl

The samples extension reactions are run as follows:

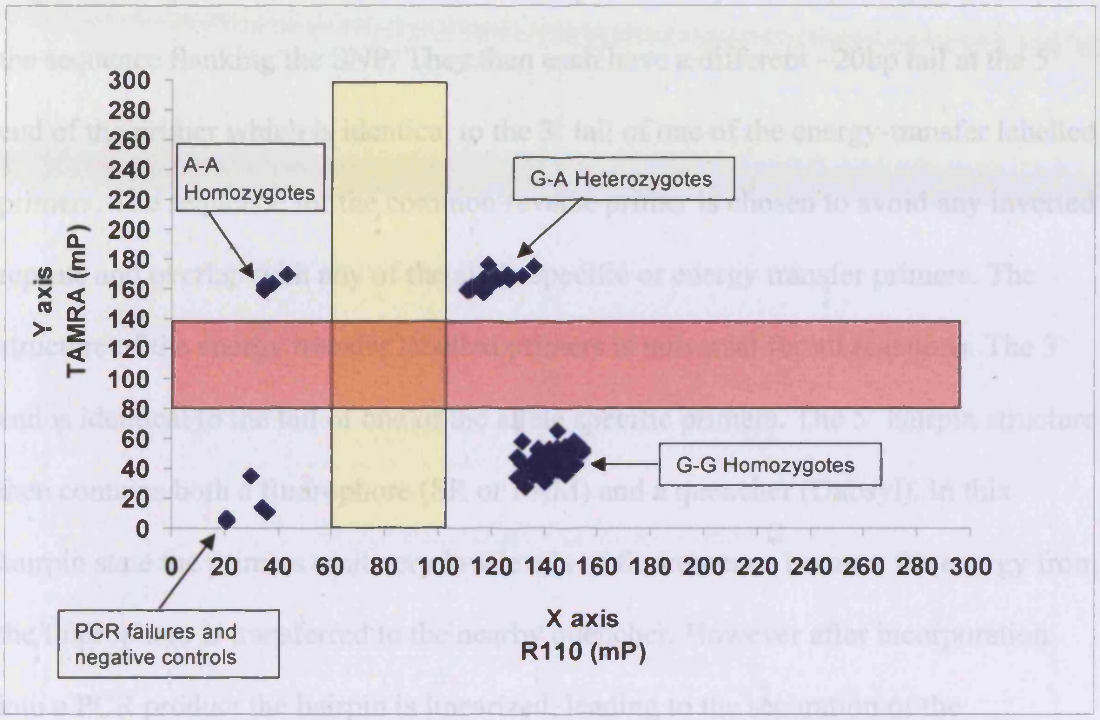
Thermocycling conditions:

95°C	2 min	
95°C	15 sec	5 – 30 cycles
55°C	30 sec	

Once the primer extension reaction has been completed the samples were chilled for about 10 minutes before being analysed. This increases viscosity and exaggerates the effect of molecular size on fluorescence polarization. The samples are then analysed using an Analyst HTS Assay Detection platform (LJL Biosystems) which is a fluorescence polarization (FP) capable fluorometer. After analysis the data is imported into an Excel macro (available from Perkin Elmer) and the genotypes scored by cluster analysis. See example of output graph below:

Figure 2.9: Example of an FP genotyping output graph.

The graph shows one of the FP genotyping output graphs generated during the individual genotyping of the CNP rs12952915 polymorphism. The clusters represent the different types of genotype, with the small cluster near zero representing the reaction dropouts and negative controls.



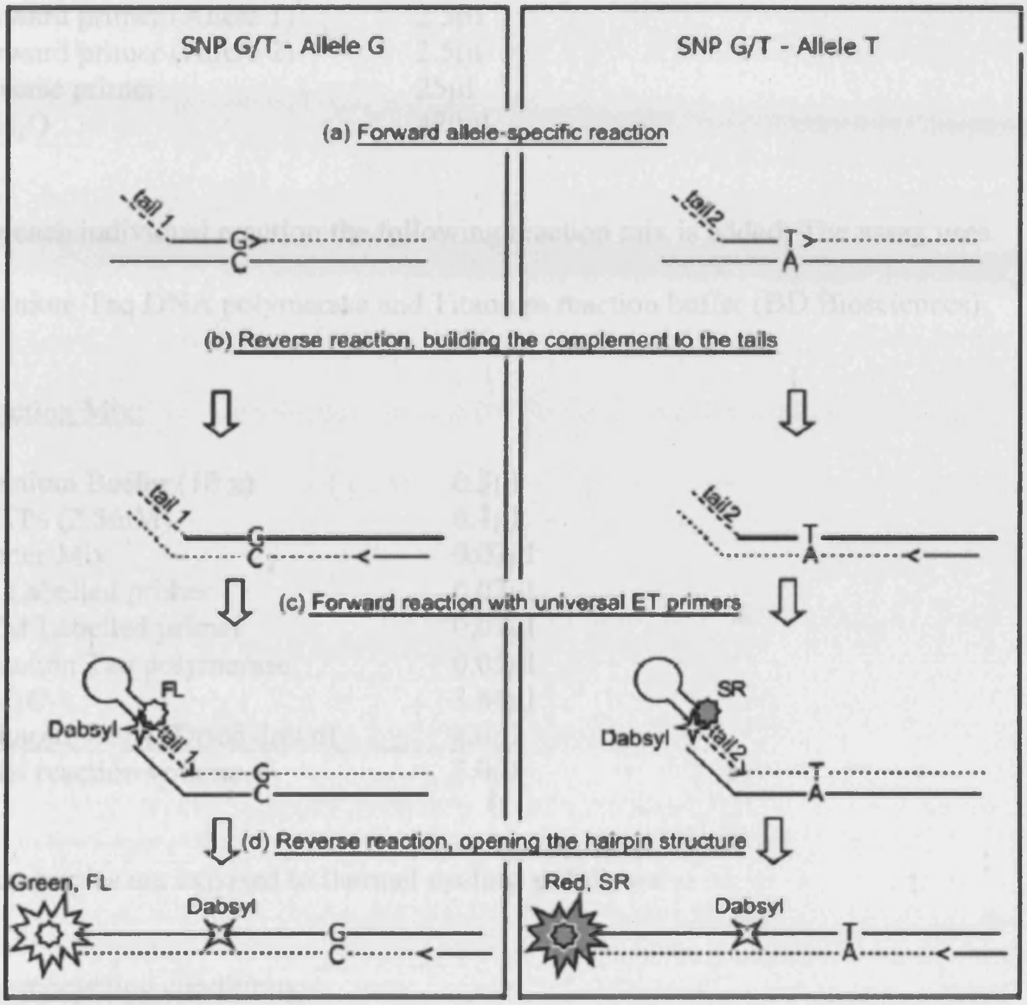
2.10.3. Amplifluor:

This technique is based on PCR amplification of genomic DNA using two, allele-specific tailed primers that introduce priming sites for universal energy-transfer labelled primers. The reaction requires a single step and no post-PCR handling (Myakishev, Khripin et al. 2001). The assays are designed with the use of the Amplifluor™ AssayArchitect™ software available at the amplifluor website (www.serologicals.com).

Genomic DNA is PCR amplified in the presence of five primers: Two forward allele-specific tailed primers, a single reverse primer, and two 'universal' energy-transfer labelled primers each containing a sequence identical to one of the allele specific tailed primers and labelled with either a green dye (fluorescein, FAM) or a red dye (sulforhodamine, SR). The aim of the assay is to have one homozygote yield green fluorescence only, the other homozygote red only, and the heterozygote a mix of red and green fluorescence. The two different allele specific primers contain a single allele specific nucleotide at the 3' end preceded by 15-20bp complementary to the sequence flanking the SNP. They then each have a different ~20bp tail at the 5' end of the primer which is identical to the 3' tail of one of the energy-transfer labelled primers. The sequence for the common reverse primer is chosen to avoid any inverted repeats and overlap with any of the allele specific or energy transfer primers. The structure of the energy transfer labelled primers is universal for all reactions. The 3' end is identical to the tail of one of the allele specific primers. The 5' hairpin structure then contains both a fluorophore (SR or FAM) and a quencher (Dabsyl). In this hairpin state the primers emit very low levels of fluorescence because the energy from the fluorophore is transferred to the nearby quencher. However after incorporation into a PCR product the hairpin is linearized, leading to the separation of the fluorophore and quencher and an enhancement in the level of fluorescence. The step by step reaction is shown below in figure 2.10.

Figure 2.10: Amplifluor™ genotyping of a G/T polymorphism.

Adapted from the Chemicon website: (www.chemicon.com).



a) Allele specific primers containing a 5' tail unique to one of the energy transfer primers anneal to the target DNA and are extended by Taq polymerase. This leads to the production of a fragment with either tail 1 or tail 2 at the 5' end, (or both in the case of a heterozygote).

b) The common reverse primer then binds to this amplified product and is elongated by Taq polymerase resulting in synthesis of the complement to the tail sequence.

c) Energy-transfer labelled primers anneal specifically to the tail sequence of the reverse reaction and are elongated by Taq polymerase.

d) The complementary strands are then synthesized using the product of the previous reaction as template. This leads to the opening of the hairpin structure at the tail of the energy-transfer labelled primers as they are incorporated into the double stranded product, separating the fluorophore and quencher which in turn gives the desired increase in green or red fluorescence.

The PCR primer mix used in the amplifluor reaction is as follows: (All primer stocks are kept at a concentration of 100pmol/μl).

Forward primer (Allele 1)	2.5μl
Forward primer (Allele 2)	2.5μl
Reverse primer	25μl
ddH$_2$O	470μl

For each individual reaction the following reaction mix is added: The assay uses Titanium Taq DNA polymerase and Titanium reaction buffer (BD Biosciences).

Reaction Mix:

Titanium Buffer (10 x)	0.5μl
dNTPs (2.5mM)	0.4μl
Primer Mix	0.07μl
SR Labelled primer	0.07μl
FAM Labelled primer	0.07μl
Titanium Taq polymerase	0.05μl
ddH$_2$O	3.84μl
Genomic DNA (Dried down)	4.0μl
Total reaction volume	5.0μl

The samples are exposed to thermal cycling as follows:

Thermocycling conditions:

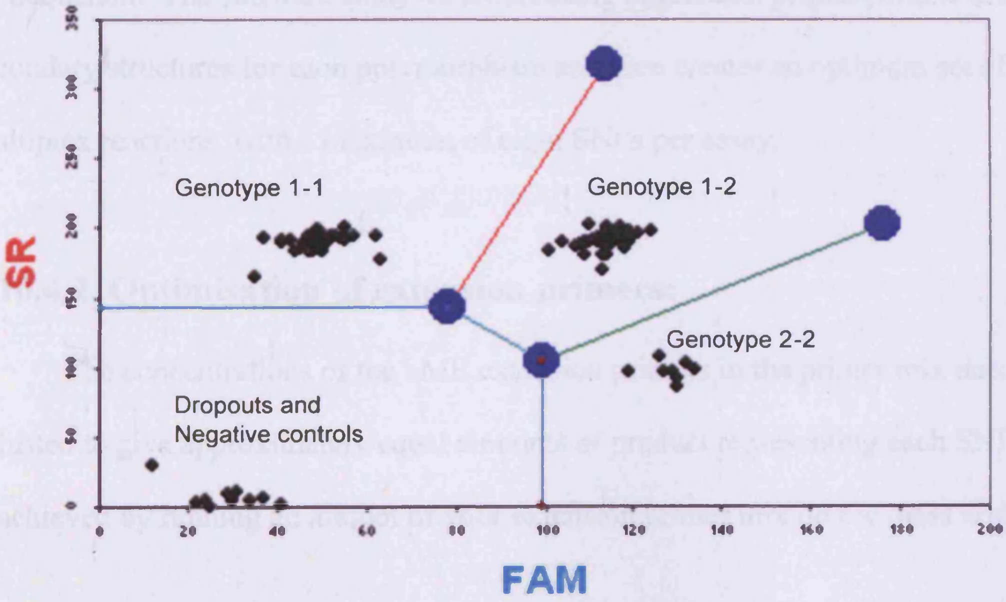
96°C	4 min	
96°C	10 sec	20 cycles
58/60°C	5 sec	
72°C	10 sec	
96°C	10 sec	22 cycles
55°C	20 sec	
72°C	40 sec	
72°C	3 min	

The reaction cycling conditions are in two stages. The first stage at 58-60°C involves the annealing and extension of the allele specific primers on the DNA template. These extended primers are then used with the reverse primer to create the complimentary tails to the labelled universal primers. The second stage involves the annealing of the universal labelled primers to their complimentary sequence and then a subsequent extension of these labelled primers to unfold their hairpin structures and allow a fluorescent signal to be generated (see figure 2.10 above).

The samples are analysed using an Analyst HTS Assay Detection Platform (LJL Biosystems). The results in the form of fluorescent signal intensity readings for the two dyes (FAM and SR) are then imported into an Excel worksheet (available from www.serologicals.com) and plotted on a graph. The clusters of points on the graph represent the three genotype classes and the worksheet then creates a file containing the genotypes in the form of 11, 12, 22, where 1 corresponds to the allele specific to the SR dye and 2 corresponds to the allele specific to the FAM dye (see figure 2.11 below).

Figure 2.11: Example of an Amplifluor genotyping graph:

This figure shows an example of an amplifluor genotyping graph. The fluorescence intensity values cluster into four groups. The upper left cluster on the plot is homozygotes with high fluorescence intensity for SR, the upper right cluster is heterozygotes with high fluorescence intensity for both the SR and FAM dyes, and the lower right cluster is homozygotes with high fluorescence intensity for the FAM dye. The bottom left cluster on the graph represents the failed PCR reactions and negative controls with low fluorescence intensity for both dyes.



2.10.4. Sequenom MALDI-TOF:

The Sequenom MassArray™ Homogenous MassExtend™ (hME) uses matrix assisted laser desorption ionisation time-of-flight (MALDI-TOF) mass spectrometry to analyse primer extension products. The method like several of the other assays is based on PCR followed by an allele specific primer extension step. However, unlike other assays Sequenom extension primers can be designed to anneal either directly or closely adjacent to the polymorphic site, meaning that primer extension is not always by just a single base. This variable extension is made possible by using a specific mixture of terminator nucleotides during the extension reaction including both appropriate deoxynucleotides and dideoxynucleotides to terminate the reaction. The mass of the primer extension product generated is then analysed and used to determine the nucleotides present at the polymorphism.

2.10.4.1. Multiplex assay design:

In order to make the method as high throughput and cost effective as possible the PCR and hME primer extension reactions are carried out in multiplex. In order to achieve this, multiplexes were designed using the Assay design 2.0 software provided by Sequenom. The software analyses the flanking sequences, primer parameters and secondary structures for each polymorphism and then creates an optimum set of multiplex reactions, with a maximum of eight SNPs per assay.

2.10.4.2. Optimisation of extension primers:

The concentrations of the hME extension primers in the primer mix need to be adjusted to give approximately equal amounts of product representing each SNP. This is achieved by running an aliquot of your extension primer mix on the mass array to determine the initial peak heights, followed by an adjustment of the primer concentrations in the mix if required to ensure equal peak heights for all of the extension primers.

2.10.4.3. Multiplex PCR optimisation:

This is achieved by amplifying 16 control DNA samples and also 8 negative controls (water). 2µl of each sample is aliquoted into a 384 well PCR plate and then dried down by incubating at 65°C for 15mins. The forward and reverse PCR primers for all of the SNPs are combined into a primer mix with a final concentration of 1pmol/µl for each of the separate primers. The following mastermix is then made and 5µl added to each of the samples:

Multiplex PCR Reaction Mix:

Genomic DNA (dried)	2µl
ddH_2O	3.39µl
PCR Buffer 10x (Qiagen)	0.625µl
$MgCl_2$ (25mM)	0.325µl
dNTPs (25mM)	0.1µl
Hotstart Taq (Qiagen)	0.06µl
F and R PCR primers (1pmol/µl)	0.5µl
Total reaction volume	5µl

The reaction is performed on a thermocycler using the following PCR program:

Sequenom PCR program:

95°C	15mins	
95°C	20 secs	45 cycles
56°C	30 secs	
72°C	1 mins	
72°C	3 mins	
4°C	hold	

The PCR products are examined by electrophoresis on an agarose gel to check that the correct sized products have been generated and there is no contamination present. Due to the large amount of primers being added to the reaction it is important to recognise any primer dimers that are present in either the DNA samples or the negative controls and not mistake them for either PCR product or contamination. Once the multiplex PCR has been optimised successfully the association sample can be aliquoted into 384 well PCR plates and then amplified as described above using the same multiplex PCR mix and PCR program.

2.10.4.4. PCR purification:

Before the hME primer extension reaction, the PCR product must be cleaned in order to remove any unincorporated dNTPs from the reaction. A 2μl mix containing 0.3μl of shrimp alkaline phosphatase (SAP), 0.17μl of SAP buffer and 1.53μl of ddH_20 is added to each PCR reaction and then incubated at 37°C for 30 mins followed by an enzyme denaturation step of 85°C for 15 mins.

2.10.4.5. hME primer extension reaction:

Once the multiplex PCR has been SAP treated the hME extension reaction can then be performed. It is important to use the right mixture of terminator nucleotides (hME extension mix) to perform the reaction. This is important to ensure that the appropriate dideoxynucleotides and specific deoxynucleotide are present to extend the primer through the polymorphic site and then to terminate the extension reaction at the appropriate point in the sequence. This ensures that the allele-specific extension products generated each have a unique mass, which can then be used to identify the genotypes present. The following reaction mix is prepared and then 2µl of the mix added to each cleaned PCR product:

<u>hME Extension Reaction Mix:</u>

ddH_2O	0.76µl
Appropriate hME Extend Mix	0.2µl
Adjusted Extension Primer Mix	1.0µl
Thermosequenase	0.04µl

The reaction proceeds on a thermo cycler using the following program:

<u>hME Extension reaction program:</u>

94°C	2 mins	
94°C	5 secs	cycle 100 times.
52°C	5 secs	
72°C	5 secs	
4°C	forever	

2.10.4.6. Desalting of extension product:

Prior to mass spectroscopy any residual salts remaining in the reaction products must be removed as these interfere with the MALDI-TOF analysis. This is achieved by the addition of Sequenom clean resin, which forms complexes with the salt ions in the reaction. 6mg of resin is added to each reaction followed by a 30 minute mixing step on a plate rotator. Following mixing, the reactions are centrifuged at 3000rpm for 10 mins to prepare the samples to be spotted onto the mass array chip.

2.10.4.7. Running and analysing spectra:

The desalted samples are spotted onto a mass array Spectrochip (Sequenom) using a Sequenom nanodispenser robot. The Spectrochip is then loaded into the MALDI-TOF mass spectrometer ready for each of the samples to be analysed. The

MALDI-TOF reaction is performed by firing a laser at the spotted multiplex extension product on the spectrochip resulting in the ionisation of the product. The length of time for the resultant ions to be extracted into the mass analyser (time of flight) is recorded and the masses and corresponding genotypes of the extension product is determined in real time by the mass array software (Sequenom). The software automatically calls the genotypes for each of the SNPs in the multiplex assay according to the masses present, and then produces a results output file which is saved to an attached computer.

2.10.5 Genescan Fluorescent PCR:

This method is particularly useful for genotyping insertion/deletions and repeat elements. A FAM fluorophore fluorescently labelled PCR primer (either the F or R primer), is used in a standard PCR reaction and then the PCR products analysed on an ABI 3100 genetic analyser (Applied biosystems) to determine the size of the PCR products and therefore the corresponding alleles present.

The PCR is performed effectively as described in section 2.3. The PCR product is then diluted (usually 1:50) in order to produce suitable electrophorogram peak heights, and then subjected to electrophoresis on the ABI 3100 genetic analyser (Applied biosystems) along with a suitable size standard (usually ROX 400/500) (Applied biosystems). During electrophoresis the products are excited by a laser light resulting in fluorescence which is then detected by a photo multiplier tube (PMT). The samples are subsequently analysed with the Genescan software (Applied biosystems), to determine the product sizes and therefore the corresponding alleles present.

2.11. Tag SNP identification by Entropy:

During some of the projects in this thesis, a low redundancy set of SNPs for haplotype analysis was selected using an Entropy method developed in house by V Moskvina, and described in (Peirce, Bray et al. 2006). Entropy as a measure of haplotypic diversity was calculated as:

$$E = -\sum_{i=1}^{k} f_i \log_2 f_i$$

where f_i is the frequency of the i-th haplotype, and k is the number of distinct haplotypes in the sample. The haplotype frequencies f_i were obtained using the EM-algorithm realized in the "Eh-plus" software (Xie 1993; Zhao, Curtis et al. 2000).

The basic principle of defining tag-markers is that the larger the value of entropy, the more informative is the set of markers forming the haplotype. It should be noted that the informativity of a set of tag markers is relative to the quantity of genetic information input into the analysis. For m biallelic markers (e.g. SNPs), the number of possible haplotypes is 2^m. The full amount of information for these m markers is:

$$E_{full} = -\sum_{i=1}^{2^m} f_i \log_2 f_i ,$$

where f_i are the m-markers haplotype frequencies. Calculating entropies for each subset of these m markers and dividing them by E_{full}, we obtain the relative amount of information (values between 0 and 1) provided by a particular subset of markers, compared with the full set of markers. By this method, haplotype-tagging SNPs can be selected even in the absence of a clear haplotypic block structure. A similar approach is described in (Ackerman, Usen et al. 2003).

2.12. Statistical Analysis:

During this thesis contingency tables were used to test genotypes and alleles for association with schizophrenia, and to calculate odds ratios. All analyses are 2-tailed. For the DHPLC mutation screening set, power was calculated using the equation: $1-(1-f)^n$, where f = allele frequency and n = number of chromosomes examined. Thus a screening set of 28 chromosomes (14 individuals) has a 95% power to detect an allele with a frequency of 10%. Power estimations and sample size requirements for case-control samples were calculated using the web based program provided by the UCLA department of statistics (http://calculators.stat.ucla.edu/). The D' and r^2 estimates of linkage disequilibrium (LD) were calculated using various versions of Haploview (http://www.broad.mit.edu/personal/jcbarret/haploview/index.php). Haplotype frequency analysis was performed using EH plus (Xie 1993) with a permutation test (Zhao, Curtis et al. 2000). All polymorphisms typed individually were tested for deviation from Hardy-Weinberg equilibrium using a goodness of fit test implemented by the HW program (in house software) written by Peter McGuffin and Peter Holmans. Group comparisons of allelic expression assay data were analyzed using *t*-tests (two tailed), and correlation analyses were performed using the Pearsons correlation coefficient.

Chapter 3.1. Analysis of oligodendrocyte related candidate genes differentially expressed in schizophrenic brain.

3.1.1 Introduction:

In this section, I describe my analysis of the oligodendrocyte and/or myelination related (OMR) genes that were identified as being differentially expressed in the brains of schizophrenics in the study which provided much of the impetus for this thesis (Hakak, Walker et al. 2001) (described in chapter 1, section 1.51). The OMR genes reported to be down-regulated in that study are myelin and lymphocyte protein (MAL), 2′, 3′-cyclic nucleotide 3′- phosphodiesterase (CNP), myelin associated glycoprotein (MAG), transferrin (TF), gelsolin (GSN) and avian erythroblastic leukemia viral oncogene homolog 3 (ErbB3).

MAL is a proteolipid expressed by oligodendrocytes and Schwann cells. Expression occurs during myelin formation and the protein is then predominantly found in compact myelin (Frank 2000). MAL has been proposed to have several functions including vesicular transport of myelin proteins and lipids to the myelin membrane during myelination and as a structural component in compact myelin (Frank 2000).

CNP is an oligodendroglial transmembrane protein synthesized by oligodendrocytes early in myelination, is associated with non-compacted myelin and is widely used as a marker protein of myelin-forming glial cells (Braun 2004). Although the proteins precise functions are unknown, it has been shown to interact with mitochondria and cytoskeletal proteins and has been proposed to promote microtubule assembly or act as a membrane anchor for tubulin (Bifulco, Laezza et al.

2002; Lee, Gravel et al. 2005). These and other putative functions for CNP are discussed more extensively in section 3.2.

MAG is a minor but important component of myelin. It is expressed only by myelin forming cells when they come into contact with axons and is thought to be involved in the initiation of myelination in the CNS (Stewart and Davis 2004). MAG has also been shown to be important in maintaining myelin-axonal contacts and the structure of myelin sheaths (Gravel, Peterson et al. 1996; Hakak, Walker et al. 2001)

TF is an iron transport glycoprotein usually synthesized in hepatocytes, but which is also synthesized in other cell types including oligodendrocytes. There is evidence that the function of TF in the brain is not restricted to iron transport but my also involve a role in myelin production and the maintenance of myelin stability (Sow, Lamant et al. 2006). It has been shown that in the brains of mice expressing reduced transferrin levels, the myelin sheath is 30% thinner than normal (Saleh, Espinosa de los Monteros et al. 2003).

GSN is a cytosolic protein which is expressed in high concentrations in developing oligodendrocytes and is also enriched in neuronal growth cones (Stewart and Davis 2004). GSN has been shown to have anti-apoptotic properties; a process which is essential for nervous system development (Xi, Qin et al. 2004).

ErbB3 is a receptor-protein tyrosine kinase that binds neuregulin proteins and is believed to be involved in cell survival, migration and differentiation (Stewart and Davis 2004). One of the ErbB3 ligands is Neuregulin 1 (NRG1), which is in turn involved in numerous cellular functions including the establishment of the oligodendroglial lineage, the maintenance and growth of neuronal cells, and the survival of oligodendroglial cells (Corfas, Roy et al. 2004). Additionally NRG1 has now been found to be associated with schizophrenia in several populations and samples (Stefansson, Sigurdsson et al. 2002; Stefansson, Sarginson et al. 2003; Williams, Preece et al. 2003; Owen, Craddock et al. 2005; Li, Collier et al. 2006).

Altered gene expression can result from either *cis* or *trans*-acting factors. If a genes altered expression plays a primary aetiological role in disease rather than being either involved in the pathophysiological pathway or a secondary effect of the disease, then it is the former *cis*-acting genetic effects that are of relevance. Therefore for the purposes of genetic association studies the presence of *cis*-acting variants is of interest, with the hypothesis being that if altered expression is of aetiological relevance, then the mechanisms causing this phenomenon can be expected to be associated with disease. Therefore the initial focus of this study was to screen the genes reported to be down-regulated in schizophrenia (Hakak, Walker et al. 2001) for the presence of *cis*-acting sequence variants.

3.1.2 Results:

The first experiments were allelic expression assays to ascertain whether the reported expression differences (Hakak, Walker et al. 2001) could potentially be attributable to *cis*-acting genetic factors. At the time of this study only four of the genes of interest had a marker SNP in public databases suitable for use in the allelic expression assay. In order to be suitable for use in the assay the marker has to be located within the exonic sequence of the gene, and have a common enough minor allele frequency to provide the assay with the power to detect common variation affecting the expression of the gene (see table 3.1). The four genes with suitable markers were *CNP, MAG, TF* and *ErbB3*. The results of the allelic expression assays are summarised below in table 3.1. The allelic expression analysis depicting all individual data points are shown below in figures 3.1, 3.2, 3.3 and 3.4.

Table 3.1: Summary of the allelic expression assay results;

Summary of the allelic expression assay results showing the gene name, marker SNP used in the assay, the number of informative heterozygotes, the power of the assay to detect a regulatory variant at a frequency of 0.1 and 0.05, the number of individuals displaying an expression difference greater than 20% and the maximum expression difference observed between alleles.

GENE	Marker SNP assayed	Number of assayed heterozygotes	Power to detect a regulatory variant at a frequency of (0.1/0.05)	No of individuals displaying variation ≥ 20%	Maximum expression difference
CNP	rs2070106	25	99% (0.1) 92% (0.05)	21 (21-67%)	67%
MAG	rs1126770	22	99% (0.1) 89% (0.05)	0	11%
TF	rs8649	18	97% (0.1) 83% (0.05)	0	13%
ERBB3	rs2271189	24	99% (0.1) 91% (0.05)	0	20%

Only one of the four genes (CNP) displayed evidence for the presence of a cis-acting sequence variant influencing the expression of the gene (see figures 3.1-3.4). Of the 60 subjects for whom brain was available, 25 individuals were heterozygous for CNP SNP rs2070106 and therefore informative for analysis. Allelic expression analysis of these individuals revealed that in each case, the A allele was expressed at a lower level relative to the G allele (figure 3.1). The data points from each individual cDNA sample showed good reproducibility, with a mean co-efficient of variation of 0.07. When data from all the individuals were combined and then analysed using a Mann-Whitney test, the A allele was significantly under-expressed ($p < 0.0001$). The analysis was repeated and gave similar results ($p < 0.0001$). Across the two experiments, the mean relative reduction of the A allele was 24%. Since the same allele is under expressed in every sample, the data suggest that the altered expression is caused by either a direct effect of the marker SNP itself or the effect of an untested variant which is in strong LD with the marker SNP.

No evidence for the presence of any common *cis*-acting sequence variants affecting the expression of MAG, TF or ErbB3 was detected. The failure to detect any common *cis*-acting polymorphisms cannot be attributed to a lack of power, as all three genes had a good power to detect regulatory variants at a frequency of 0.1 and 0.05 (see table 1.1). It should also be noted that the power calculations are conservative, in that they assume that there is no LD present between the marker variant and the functional variant. If there is LD present between the marker and regulatory variant, the power will be increased as a higher proportion of people selected for heterozygosity at the marker locus will also be heterozygous for the regulatory variant.

Figure 3.1: CNP (rs2070106) allelic expression assay:

The graph shows the corrected genomic and cDNA ratios for CNP SNP rs2070106 in 25 heterozygotes. Data are expressed as the mean of the ratio of A:G for two measurements for each genomic DNA sample and of four measurements of each cDNA sample.

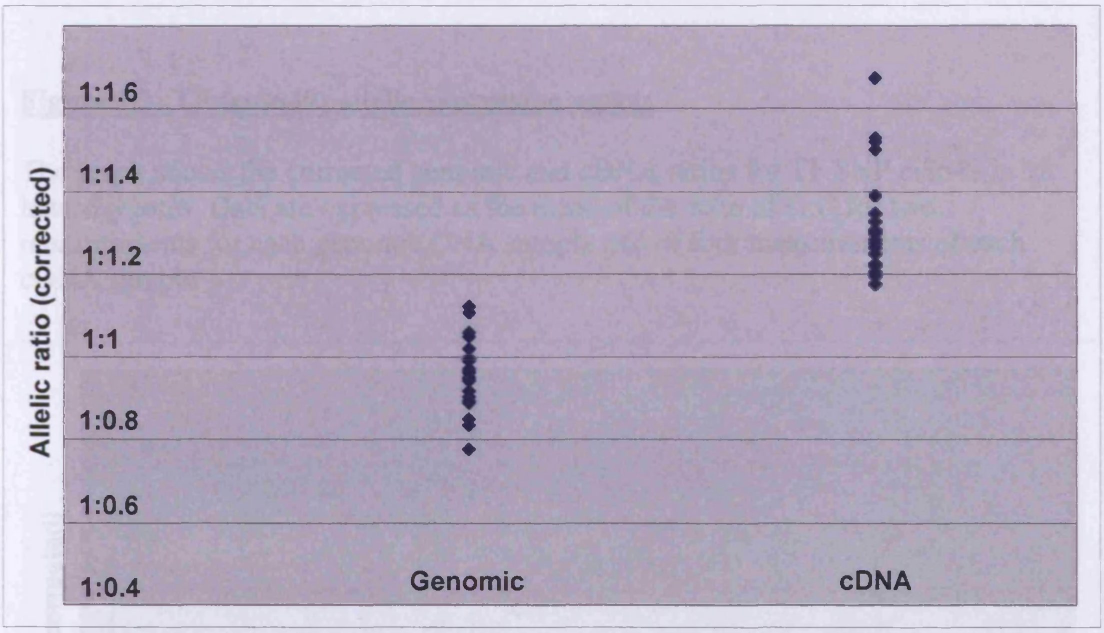


Figure 3.2: MAG (rs1126770) allelic expression assay:

The graph shows the corrected genomic and cDNA ratios for MAG SNP rs1126770 in 22 heterozygotes. Data are expressed as the mean of the ratio of C:T for two measurements for each genomic DNA sample and of four measurements of each cDNA sample.

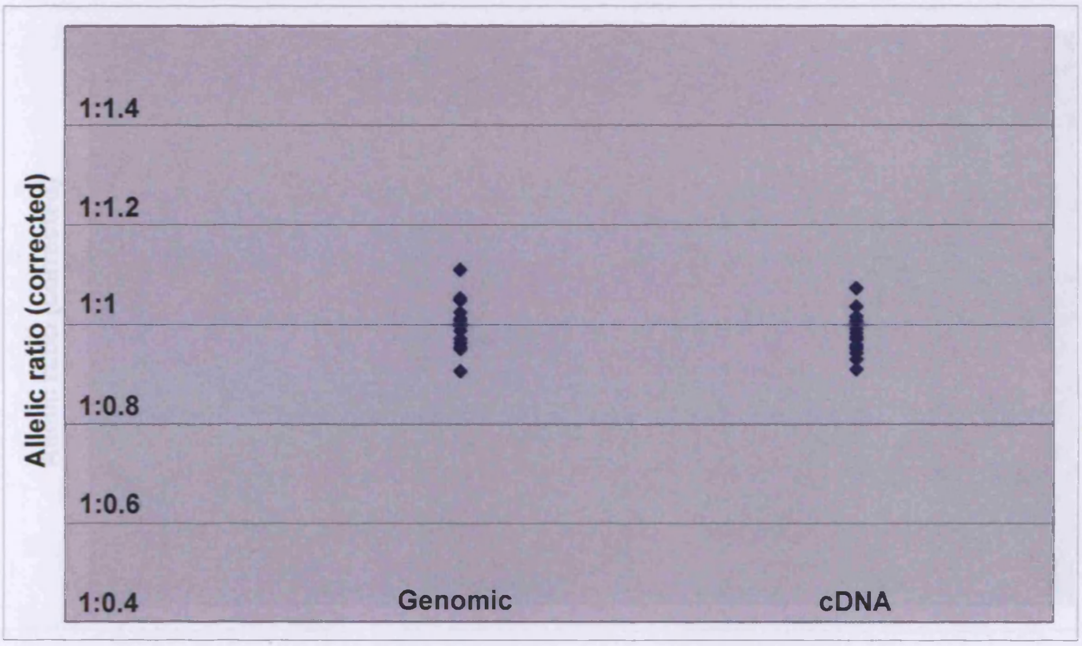


Figure 3.3: TF (rs8649) allelic expression assay:

The graph shows the corrected genomic and cDNA ratios for TF SNP rs8649 in 18 heterozygotes. Data are expressed as the mean of the ratio of G:C for two measurements for each genomic DNA sample and of four measurements of each cDNA sample.

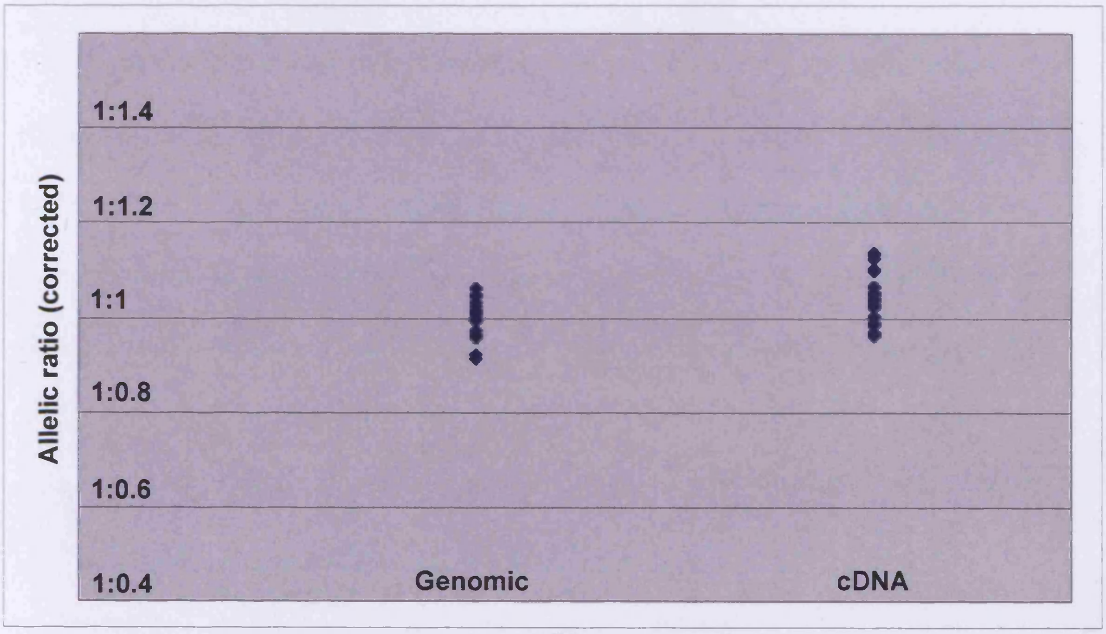
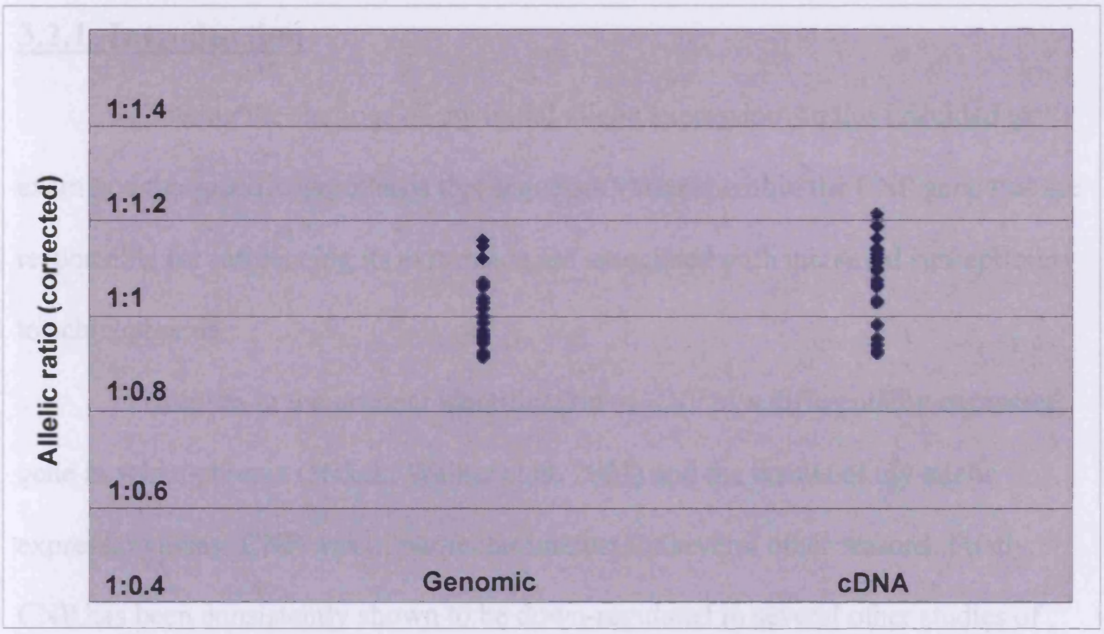


Figure 3.4: ErbB3 (rs2271189) allelic expression assay:

The graph shows the corrected genomic and cDNA ratios for ErbB3 SNP rs2271189 in 24 heterozygotes. Data are expressed as the mean of the ratio of G:A for two measurements for each genomic DNA sample and of four measurements of each cDNA sample.



Chapter 3.2. 2′, 3′-Cyclic Nucleotide 3′- Phosphodiesterase (CNP) a candidate gene for Schizophrenia:

3.2.1. Introduction

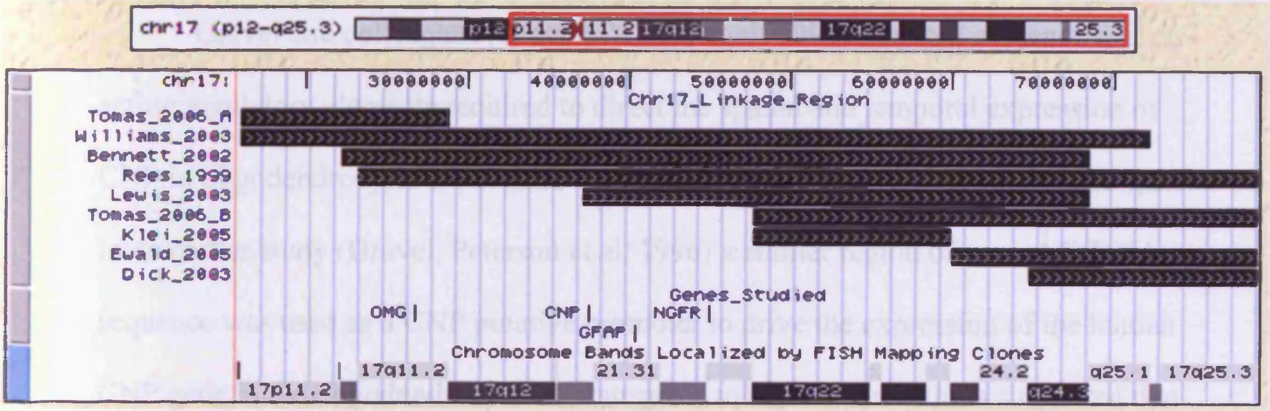
Following the findings of my initial allelic expression studies I decided to examined the specific hypothesis that sequence variants within the CNP gene that are responsible for influencing its expression are associated with increased susceptibility to schizophrenia.

In addition to the original identification of CNP as a differentially expressed gene in schizophrenia (Hakak, Walker et al. 2001) and the results of my allelic expression assay, CNP was of particular interest for several other reasons. Firstly, CNP has been consistently shown to be down-regulated in several other studies of schizophrenia (Aston, Jiang et al. 2004; Katsel, Davis et al. 2005b; Dracheva, Davis et al. 2006) (see chapter 1, table 1.1). The mRNA findings are also supported by reports of a significant reduction at the protein level in both the anterior frontal cortex (Flynn, Lang et al. 2003) and the hippocampus (Dracheva, Davis et al. 2006) of schizophrenic brains compared to controls.

Secondly, CNP maps to chromosome 17q21.2, spans 11 kb of genomic sequence and contains four exons. This is a region of the genome in which my host department had at the time I started, recently observed genome wide significant evidence for linkage (LOD 8.32, genome-wide empirical p value ≤ 0.02) to schizophrenia in a single multiply affected pedigree (Williams, Norton et al. 2003). The involvement of this region in schizophrenia has been supported by a recent meta-analysis (Lewis, Levinson et al. 2003) and additional linkage studies (Rees, Fenton et al. 1999; Klei, Bacanu et al. 2005) (see figure 3.5). Additionally there is also significant (Dick, Foroud et al. 2003; Tomas, Canellas et al. 2006) and suggestive (Bennett, Segurado et al. 2002; Ewald, Wikman et al. 2005) evidence of linkage to this region in bipolar disorder (see figure 3.5). Given the reports of linkage to the region of chromosome 17 where CNP is located and the evidence for altered expression of CNP in schizophrenia, the gene is a plausible candidate gene for study as to its involvement in susceptibility to schizophrenia.

Figure 3.5. Chromosome 17 linkage region.

Summary of the linkage findings reported on chromosome 17 for schizophrenia and bipolar disorder. The figure shows the linkage studies and the corresponding region of chromosome 17 they reported to be linked to disease, along with the location of the four genes located on chromosome 17 which were studied as part of this project.



CNP transcripts encode for two prenylated myelin protein isoforms, CNP1 (46 kDa, 400 amino acids) and CNP2 (48 kDa, 420 amino acids) that differ by the presence of alternative translation start sites in the first two exons of the gene (O'Neill, Minuk et al. 1997). Both isoforms are identical except that the CNP2 isoform has a 20 amino acid extension at the N-terminus (Lee, O'Neill et al. 2006).

There are also two promoter regions of the gene encoding for the two mRNA transcripts, one of which (CNP2) can generate both CNP protein isoforms.

CNP accounts for about 4% of the protein in central nervous system myelin, but only accounts for about 0.4% of the protein in peripheral nervous system myelin (Rasband, Tayler et al. 2005). Numerous studies have shown an association between CNP protein expression and myelinating cells (Braun 2004). However, immuno-electron microscopy of myelin has shown that CNP is associated only with non-compacted myelin regions, such as the inner mesaxon and paranodal loops, and that it is absent from compacted myelin (Lappe-Siefke, Goebbels et al. 2003). Analysis of the amino acid sequence of CNP shows a high homology between the different mammalian species. For instance the homology between the mouse and human CNP sequence is 85% (Braun 2004).

Gravel and colleagues (Gravel, Di Polo et al. 1998) reported evidence that *cis*-acting regulatory elements required to direct the spatial and temporal expression of CNP in oligodendrocytes are located within 4kb of the 5'end of the mouse CNP gene. In an earlier study (Gravel, Peterson et al. 1996) a similar region of mouse 5' DNA sequence was used as a CNP putative promoter to drive the expression of the human CNP gene in the oligodendrocytes of transgenic mice, and the authors suggested that the human and mouse CNP promoter sequences are therefore similar.

CNP is found in cytoplasm, and its expression appears at an early stage of development in oligodendrocyte precursors, preceding the onset of myelination. Its expression is then maintained at a high level as myelination proceeds (Lee, Gravel et al. 2005). Although high levels of CNP are seen in differentiating oligodendrocytes in parallel with the major structural proteins such as myelin basic protein (MBP) or proteolipid protein (PLP), the expression of CNP has been shown to occur before either of these proteins are expressed, indicating a possible role for CNP before or during the initiation of myelination (Braun 2004). CNP protein continues to be expressed throughout adulthood, when it shows a high turnover compared with the other major myelin related proteins, suggesting a possible ongoing role for CNP in the CNS after neuronal development and myelination (Flynn, Lang et al. 2003).

The precise function of the CNP gene in oligodendrocytes is not known, although the expression pattern of the protein would suggest an important role in oligodendrocyte development, myelination and possibly maintenance of the myelin sheath (Lee, Gravel et al. 2005). These hypotheses are supported by various animal studies. Several studies have demonstrated that premyelinating oligodendrocytes from transgenic CNP over-expressing mice produce aberrant and redundant membranous extensions from their oligodendrocyte processes and myelin internodes (Gravel, Peterson et al. 1996; Yin, Peterson et al. 1997). However, in contrast to the over-expressing CNP mice, a study by (Lappe-Siefke, Goebbels et al. 2003) reported that CNP-null transgenic mice appeared to have normal myelination although they did display progressive neurodegeneration, axonal loss and swelling, which led to premature death. The mice also displayed ataxia and hindlimb impairments at about four months of age, as well as progressive reduction in overall brain size and progressive enlargement of ventricles and corpus callosum atrophy (Lappe-Siefke, Goebbels et al. 2003). Enlarged ventricles and corpus callosum atropy are consistent with some pathological features observed in schizophrenia (Harrison 1999; Cahn, Hulshoff Pol et al. 2002; Ho, Andreasen et al. 2003). More recent analysis of the CNP-null mice has revealed that the deficiency of CNP leads to disrupted axoglial interactions in the CNS and major abnormalities in the structure of the paranodal loops as early as three months of age, and before the onset of any axonal or neuronal degeneration (Rasband, Tayler et al. 2005). The authors suggest that CNP deficiencies result in disrupted axoglial signalling and underlie the mechanism responsible for the progressive axonal death observed later in the CNP-null mice.

Interestingly CNP gene expression has been shown to be increased by lithium treatment (Wang and Young 1996). Lithium is often used as a highly effective treatment for psychiatric diseases, although its role in schizophrenia is not well established. If CNP is in fact involved with the control of the myelination process then the therapeutic effects of lithium treatment might be due to its regulating one or more of these myelination processes.

A final piece of circumstantial evidence concerning a possible relationship between CNP and schizophrenia comes from its potential interaction with the protein quaking homolog, KH domain RNA binding (mouse) (QKI). QKI is an RNA binding protein in myelin producing cells and is required for the cell-cycle arrest of oligodendrocyte progenitor cells and the subsequent differentiation of oligodendrocytes (Larocque, Galarneau et al. 2005). Zhang and colleagues (Zhang and Feng 2001) showed that CNP binds to QKI transcripts weakly and that the expression of CNP mRNA is nearly normal in the brain of QKI null transgenic mice. However, the authors also report that CNP protein levels were significantly reduced in the brains of the same QKI null mice. The authors propose that this reduction in CNP is unlikely to result from a failure in CNP translation but instead is probably the result of accelerated protein degradation possibly as a consequence of a lack of QKI. This finding is of interest as QKI has recently been shown to be the most significantly down-regulated gene in a micro-array study of schizophrenic brain (Katsel, Davis et al. 2005b). The gene has also been reported to be associated with schizophrenia (Aberg, Saetre et al. 2006a; Aberg, Saetre et al. 2006b) and my host department has recently observed a significant association (p=0.002) between several genetic variants in the gene and susceptibility to schizophrenia (Georgieva, O'Donovan, unpublished data). If the theory Zhang and colleagues (Zhang and Feng 2001) proposed, that a lack of QKI protein leads to the increased degradation of CNP protein is correct, then the reports of reduced QKI expression in schizophrenic brain (Katsel, Davis et al. 2005b), would tie in with the results demonstrating a reduction of CNP at the protein level in schizophrenic brain (Flynn, Lang et al. 2003; Dracheva, Davis et al. 2006) and may indicate the presence of an interaction between the roles these two genes play in schizophrenia aetiology.

So in summary based on the reports of linkage to the broad region of chromosome 17 where CNP is located, evidence for altered expression of CNP in schizophrenic brain at both the mRNA and protein level, the genes proposed function in oligodendrocyte function and the myelination process, and the results of the CNP knockout mouse studies reporting central nervous system pathological characteristics reminiscent of some of the pathological features observed in schizophrenic patients, CNP is a highly plausible candidate gene for schizophrenia, albeit one of many highly plausible genes.

3.2.2 Materials and Methods:

In this section I describe methodological issues specific to this section of the thesis. A more detailed analysis of the methods and techniques used can be found in chapter 2.

3.2.2.1 Patient Samples:

The case/control sample used in this chapter is described in chapter 2. The pedigree showing evidence for linkage to 17q consists of a single generation of 6 siblings with DSM IV schizophrenia (diagnosed as described in chapter 2) from which DNA was available. DNA was also available from one parent of unknown diagnostic status.

3.2.2.2 Polymerase Chain Reaction PCR:

Primers to amplify the exons, promoters, flanking sequences and introns of the CNP genomic sequence based upon alignment of mRNA sequence (NM_033133) and the corresponding genomic sequence (NT_010755.14). The primers used to generate the various amplimers are listed in tables 3.3 and 3.4 respectively. PCR amplification was performed under standard conditions.

Table 3.3: Primer information for the amplimers used during the initial CNP screening experiments:

This table shows the details of the primers used during the initial mutation screening of the CNP gene in 14 cases, including the forward and reverse PCR primer sequences, the size of the fragments and the optimised reaction conditions and annealing temperatures. HS indicates the use of hot start taq polymerase and GC indicates the use of a GC rich PCR protocol.

Fragment Name	Genomic region covered	Forward Primer (5' to 3')	Reverse Primer (5' to 3')	Size (bp)	Annealing Temperature (°C)
CNP_1	Exon 1	CTTGGACAGGAAACTCAAAAAT	CTGCGTGTTGAGGGGTAACT	441bp	HS54
CNP_2	Exon 1	GACCCTGCCTCACTCACTTC	GTTTGCAAGACGCTTCCTC	472bp	GC60
CNP_3	Exon 1	CCCCCGGGCTATGTAAAG	GTAAAGGGTCCCCTCCTCTG	566bp	GC60
CNP_4	Intron 1	CAGTCTCATCACGCGGAAC	GGCTGAGTGCACAGAGAAGG	458bp	HS54
CNP_5	Intron 1	AGGAATGTGTGGCTTTTTCG	AGGAATGTGTGGCTTTTTCG	450bp	HS54
CNP_6	Exon 2	CATTTCAAAGAGCTGCTCCA	CATTTCAAAGAGCTGCTCCA	446bp	HS54
CNP_7	Exon 2	AGGACCTGGCTGCCTACTG	ATTGTCGCAGCTCCTTCTTG	363bp	HS54
CNP_8	Exon 3	TGAGCTGTAGTGGGTTGCTG	TGGCACCGGTACCTATCTTC	360bp	HS60
CNP_9	Exon 4	TGTTAAAGAAATCTTACTCCAAGG	CACTCATATGATGGTGCAGGA	452bp	HS54
CNP_10	Exon 4	ACTACGGGAAAGGCAAACCT	GGGGCTCTGAGGTGTCTCC	460bp	HS54
CNP_11	Exon 4	GCCCAGAGCTGGGTGAGT	GCGGCTGCTTAACTGTTCTG	445bp	HS54
CNP_12	Exon 4	CCTCCCCTTCACACCATAAC	AGGATTAACATGATTTATTTCATTATC	519bp	HS50

Table 3.4: Primer information for the amplimers used during the additional CNP screening experiments:

This table shows the details of the primers used during the additional screening of the complete genomic sequence of the CNP gene, including the forward and reverse PCR primer sequences, the size of the fragments and the optimised reaction conditions and annealing temperatures.

Fragment Name	Genomic region covered	Forward Primer (5' to 3')	Reverse Primer (5' to 3')	Size (bp)	Annealing Temperature (°C)
CNP_b_1	5' Flanking/Put prom seq	CAGGTCCCTTGACTGCAAAC	GAGAAAAGCGGGAAGAAAGA	597bp	HS 60
CNP_b_2	5' Flanking/Put prom seq	CGGCCTGACATGAACAAACT	TCTGGTTCCTGTCTGCTGAA	425bp	HS 60
CNP_b_3	5' Flanking/Put prom seq	GGCCTTCGGGAGAATTAGAG	TTCTTCCACGATGCTCTTCC	409bp	HS 60
CNP_b_4	5' Flanking/Put prom seq	TCACTCTTGCAAACCCTGAG	CCTGTCCAAGTGCTGTGCTA	504bp	HS 60
CNP_b_5	5' Flanking/Put prom seq	TCTGGGTGGCATAATCAGGT	CTGCGTGTTGAGGGGTAACT	549bp	HS 60
CNP_b_6	Intronic Sequence	GGACTTCCTGCCGCTCTACT	CCAAGACAAGGGACAGTGGT	427bp	HS 60
CNP_b_7	Intronic Sequence	CTTTCCTCCCACCACACCTA	GTAAACCCCACATCCCTCCT	423bp	HS 60
CNP_b_8	Intronic Sequence	ACTCATGGGTTTGTGGCATC	TGGTGCACACCTGTAGTCCT	444bp	HS 60
CNP_b_9	Intronic Sequence	CAACCTCTGCCTCCCAGA	TCACAGGTACAGGGCTTGCT	423bp	HS 60
CNP_b_10	Intronic Sequence	GGGAATTTGTGCCTGTATGG	CTTGATTCGTCCCATCTTGG	438bp	HS 60
CNP_b_11	Intronic Sequence	GCTGGTCGGTAGCACAAAA	CTTGCTGGCCCATTACATTT	436bp	HS 60
CNP_b_12	Intronic Sequence	TCTGCCCAGGATTAGAGCAG	CTCAGCTCTGCCATGTACCA	470bp	HS 60
CNP_b_13	Intronic Sequence	GGGAACACTGGGGTATATGG	CCAGGAGGCAGAGGATACAG	414bp	HS 60
CNP_b_14	Intronic Sequence	CAATGTAAAGCCAGAGAATTTTAACT	CAGCAACCCACTACAGCTCA	406bp	HS 60
CNP_b_15	Intronic Sequence	GCACCTGGCCTGAGAATTT	GGCACCGGTACCTATCTTCA	430bp	HS 60
CNP_b_16	Intronic Sequence	AGTCTTCCCCAGGGACACAT	TGTGGATATGGGTGACATGC	547bp	HS 60
CNP_b_17	Intronic Sequence	CTTACACCCACAGCCACTGA	CGCACACAACAGCTGCTACT	544bp	HS 60
CNP_b_18	Intronic Sequence	AGTCTGACCTCACCCTGTCC	CGTGAAGGCCTTGGAGTAAG	527bp	HS 60
CNP_b_19	Intronic Sequence	GGGTCAGCTGTTCCTCAAGA	AGGTTTGCCTTTCCCGTAGT	457bp	HS 60
CNP_b_20	3' Flanking sequence	TCGGGTTGTCACTCAACAAA	TCTTCCCTGGTGCTTGAAGT	518bp	HS 60
CNP_b_21	3' Flanking sequence	TGGTATGGGTTTCTCCCAGT	GGGTTCGAGGGAGATGGTAG	534bp	HS 60

3.2.2.3 Pooled Genotyping:

All of the polymorphisms studied in the CNP gene were genotyped in the pooled association sample by primer extension using the ABI SNaPshot reaction method. The SNPs, their flanking sequences and the extension primers used for the pooled genotyping assays are shown in table 3.5. For each SNP the pools were amplified in triplicate along with an individual heterozygous for each SNP.

3.2.2.4 Individual Genotyping techniques:

During the completion of the experiments presented in this chapter several individual genotyping techniques were utilised. These included Acycloprime FP, Amplifluor, Sequenom MALDI-TOF, and multiplex SNaPshot genotyping. All of these techniques are described in section 2.10. All genotypes were scored blind to affection status, and each plate of samples contained a mixture of equal numbers of cases and controls.

Table 3.5: Polymorphism and primer extension reaction information for the CNP gene.

This table shows the nineteen polymorphisms detected within the CNP gene, their rs numbers if available, the sequence flanking each of the polymorphisms, the extension primer used to assay the polymorphism by primer extension and the expected alleles in the primer extension reaction.

SNP number.	rs Number	SNP Flanking sequence	Extension primer	Expected Alleles
CNP -4359 G/A	rs4796750	GCCGACCCCCATTC[A/G]TGCTTCCCGCAGGC	TAAGCGCAGTAGCCGACCCCCATTC	G (Blue) & A (Green)
CNP -2086 C/T	rs8080978	GGCGCCAACCACCA[C/T]GCCTGGCTAATTTT	GGGACTACAGGCGCCAACCACCA	C (Black) & T (Red)
CNP -2084 C/G	rs8080979	CGCCAACCACCACG[C/G]CTGGCTAATTTTTG	CTGTACTAAAAATACAAAAATTAGCCAG	G (Blue) & C (Black)
CNP -1148 AG Ins/Del	rs10540926	GGTGCCAAGAAGAA[AG]GGTTTCAGGAGGGT	GAAATGAGAAAGGTGCCAAGAAGAA	A (Green) & G (Blue)
CNP -351 G/A	rs12952915	TCCTGCTAATCCAC[G/A]GGACCGGGGAGGGT	GCCTCACTCACTTCCTGCTAATCCAC	G (Blue) & A (Green)
CNP -96 T/C	rs4432296	GGCCGGGCTCGGGT[T/C]GTGCCACCGCTGGA	CGGGAGTCCAGCGGTGGCAC	A (Green) & G (Blue)
CNP 1868 T/G	rs8078650	ATGTAGGTGGCAGG[T/G]TGGGGCCTTATAAG	GCGACAATGTAGGTGGCAGG	T (Red) & G (Blue)
CNP 1944 A/T		*TACTCAAAGGATGG[A/T]GCACGAAGCAGCAG*	*CTCTCTGCCTCCTGCTGCTTCGTGC*	*T (Red) & A (Green)*
CNP 2371 A/G	rs4258677	TCAGGCCTCGGACT[A/G]CTCAATTTACTCAT	GCCACAAACCCATGAGTAAATTGAG	T (Red) & C (Black)
CNP 3235 A/G		TTTATTTATGAGAC[G/A]GAGTCTTGCTCTGT	TCAAAGTTGCTTTTAATATTTATTTATGAGAC	G (Blue) & A (Green)
CNP 3582 C/T	*rs12952456*	*GCAGTGGGGCAATT[C/T]GGCTTACTGCAACC*	*CGGGAGATGGAGGTTGCAGTAAGCC*	G (Blue) & A (Green)
CNP 3625 T/C	rs12950675	TCAAGCAATTCTCC[T/C]GCCTCAGCCACCCA	ATCTCCCGGGTTCAAGCAATTCTCC	T (Red) & C (Black)
CNP 4926 A/G	rs12602950	TCCGAGTGTTTTGC[G/A]CTGGGCCTCGGTGG	GATGCTTAGTGTCCGAGTGTTTTGC	G (Blue) & A (Green)
CNP 6961 G/A	rs2070106	GCCATCTTCACGGG[G/A]TACTACGGGAAAGG	GTCAGGGCCATCTTCACGGG	G (Blue) & A (Green)
CNP 7230 G/A	rs11079028	CCTCTAATGCTCAC[G/A]CTCCCAACACAAGG	CGCCCTCTTCCCCTCTAATGCTCAC	G (Blue) & A (Green)
CNP 7284 G/A		CCTGGACCAAAGCT[G/A]ACGAGGCTGGGCCA	CCATTCAGGAACCTGGACCAAAGCT	G (Blue) & A (Green)
CNP 8157 T/C	rs11296	GCCCTCCACTGCCC[C/T]CTGTTCAGTAACAG	GTTCCTGGGGTGCCCTCCACTGCCC	C (Black) & T (Red)
CNP 8634 C/T	rs4796751	GGAAGCTCCTGGTA[T/C]GGGTTTCTCCCAGT	CCCTAGAGAACTGGGAGAAACCC	A (Green) & G (Blue)
CNP 8962 C/T		CACCTTGAACACCA[C/T]GCTTGACAGGACAG	GGCTACTCACCTTGAACACCA	C (Black) & T (Red)

3.2.3 Results.

Given the results of the allelic expression assays I decided to undertake additional genetic studies of CNP. Initially this involved testing the specific hypothesis that the lower expression A allele of rs2070106 would be associated with schizophrenia. Individual genotyping of marker rs2070106 in the case control sample was consistent with this hypothesis at a level that meets conventional levels of significance (p = 0.04, OR = 1.2, 95% CI (1.0-1.40)). Genotypic analysis of the data revealed no stronger evidence of association (see table 3.6).

Table 3.6: Individual genotyping of rs2070106 in the schizophrenia sample:

Individual genotyping results for marker rs2070106 in the schizophrenia association sample. Genotypic and allelic counts as well as allele frequencies in the cases and controls, with the corresponding chi-square and p-values are given.

SNP	**Genotypes**			**Alleles**	
rs2070106 G>A	**G/G**	**G/A**	**A/A**	**G**	**A**
Cases	288	319	74	895 (0.66)	467 (0.34)
Controls	325	303	59	953 (0.7)	421 (0.3)
	χ^2 = 4.31 (2df)	P = 0.11		χ^2 = 4.15 (1df)	P = **0.04**

Given the positive association detected for rs2070106, which supports the hypothesis that lower expression would be associated with disease, I decided to expand my study of the CNP gene further in order to look for non-synonymous SNPs and other variants that might be responsible for the genes altered expression and also any markers in LD with rs2070106. This was performed initially by de novo mutation discovery of the CNP exons and flanking intronic sequence using 14 unrelated schizophrenic individuals. This revealed two novel sequence variants (rs12952915, 7284 G>A) and two variants already lodged in the Chip bioinformatics database (rs4432296 and rs2070106). SNP labelling is based on genomic sequence (NT

010755.14) and follows the labelling protocol as suggested (den Dunnen and Antonarakis 2001). Of these polymorphisms, two were located in the coding sequence (rs2070106 and 7284 G>A), neither encoded for an amino acid change, although the rs2070106 polymorphism is a synonymous mutation in the final exon coding for glycine (G376G). The polymorphisms were then genotyped through the pooled case control sample, which revealed a trend for association between the lower expression A allele of rs2070106 and schizophrenia ($p = 0.06$) consistent with the individual genotyping data, but no trend for the other three polymorphisms (see table 3.9).

To continue the search for a functional locus in LD with rs2070106, located within either intronic, 5' or 3' sequence, I performed a more detailed analysis of the CNP locus. This was performed by genotyping in pools markers from the public databases at a minimum density of one SNP per 5kb across the gene (see table 3.9, markers highlighted in grey). No further evidence for association to schizophrenia was detected.

However, the optimal approach to LD mapping remains a subject of controversy and I have therefore used two approaches. The first was as described above to genotype all the above markers in pools on the premise that reduced degrees of freedom means that single locus LD analysis based upon a fairly dense set of markers maybe more powerful than multi-locus analysis if genotypes at one of the tested markers correlate strongly with those at the undetected locus (Kruglyak 2005). This type of technique has been used successfully in my laboratory to detect a putative susceptibility locus for dyslexia (Cope, Harold et al. 2005) and Alzheimer's disease (Grupe, Li et al. 2006). The second approach I used was based upon haplotype analysis, where the aim was to type a subset of the SNPs that could retrieve >95% of the haplotype diversity across the gene that would have been captured by individually genotyping all of the detected SNPs. This was achieved by using an entropy method to determine the redundant markers (see materials and methods section). In order to perform this analysis, all of the SNPs I had detected along with the markers contained within the 5kb grid were initially genotyped in 96 unrelated DNA control samples to allow the LD structure to be elucidated and to allow estimation of the haplotype frequencies. The subsequent entropy analysis revealed that by typing three further markers individually (rs12952915, rs11296 and rs4796751) in addition to rs2070106, I could extract 98% of the haplotypic diversity that would have been provided by genotyping all the markers individually.

These three additional markers were genotyped individually in the association sample but displayed no further evidence for association at the individual level (see table 3.7). Haplotype analysis was performed on all of the possible haplotype combinations, but none of the 2, 3 or 4 marker global haplotypes were significant (see table 3.8) and there was no more significant evidence for association for any of the single haplotypes. All of the genotypes were in Hardy-Weinberg equilibrium.

Table 3.7: Individual genotyping for CNP markers rs12952915, rs11296 and rs4796751 in the schizophrenia sample.:

Individual genotyping results for CNP markers rs12952915, rs11296 and rs4796751. Genotypic and Allelic counts, the allele frequencies and the corresponding chi-square and p-values for the markers are given.

SNP	Genotypes			Alleles	
rs12952915 G>A	**G/G**	**G/A**	**A/A**	**G**	**A**
Cases	445	212	19	1102 (0.82)	250 (0.18)
Controls	445	205	28	1095 (0.81)	261 (0.19)
	$\chi^2 = 1.84$ (2df)	P = 0.39		$\chi^2 = 0.25$ (1df)	P = 0.62
rs11296 T>C	**T/T**	**T/C**	**C/C**	**T**	**C**
Cases	582	73	1	1237 (0.94)	75 (0.06)
Controls	597	83	4	1277 (0.93)	91 (0.07)
	$\chi^2 = 2.05$ (2df)	P = 0.36		$\chi^2 = 1.01$ (1df)	P = 0.32
rs4796751 G>A	**G/G**	**G/A**	**A/A**	**G**	**A**
Cases	435	165	19	1035 (0.84)	203 (0.16)
Controls	483	176	19	1142 (0.84)	214 (0.16)
	$\chi^2 = 0.18$ (2df)	P = 0.91		$\chi^2 = 0.18$ (1df)	P = 0.67

Table 3.8 Initial CNP haplotype analysis.

Results of the initial CNP haplotype analysis. All 2, 3 and 4 marker haplotypes are shown, with their corresponding global p-values and also the most significant single haplotype p-values. The markers 1-4 refer to SNP's rs12952915, rs2070106, rs11296, and rs4796751 respectively.

Haplotype	Most significant single haplotype	Global p-value
1-2	0.07	0.09
1-3	0.62	0.86
1-4	0.65	0.86
2-3	0.06	0.16
2-4	0.12	0.43
3-4	0.67	0.84
1-2-3	0.1	0.17
1-2-4	0.09	0.54
1-3-4	0.44	0.89
2-3-4	0.12	0.52
1-2-3-4	0.09	0.58

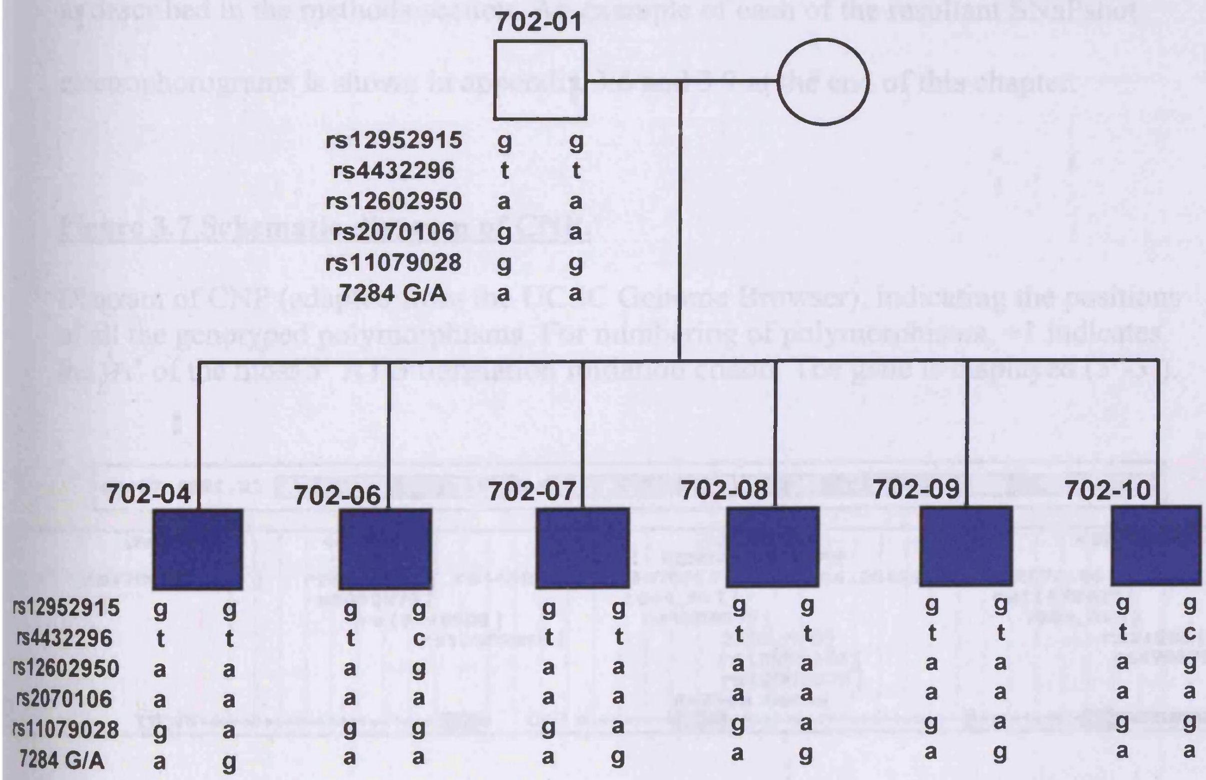
3.2.3.1 Family 702 pedigree analysis:

As mentioned before, my host department previously reported genome wide significant evidence for linkage (LOD 8.32, genome-wide empirical p value ≤ 0.02) to schizophrenia in a single pedigree (family 702) (Williams, Norton et al. 2003). In order to determine if this linkage finding might be the result of variation in CNP I sequenced the exons of CNP in all 7 of the available family members of the family 702 pedigree. Two additional sequence variants present in the family were discovered (4926 A>G, 7230 G>A) and genotyped through the pooled association sample. The results are presented in table 3.9. Neither was associated with schizophrenia or changed the predicted amino-acid sequence of the encoded CNP.

Based upon the linkage data (Williams, Norton et al. 2003) it was expected that all of the affected siblings would show identity by descent for both maternal and paternal chromosomes, and therefore with respect to any one CNP marker, the affected individuals would all have the same genotype. However, for some of the markers (rs4432296, rs12602950, rs11079028 and 7284 G>A) homozygous and heterozygous genotypes were observed indicating that among the siblings, a minimum of three parental chromosomes were represented with respect to the CNP locus (see figure 3.6). The availability of parental DNA allowed me to determine that each affected subject received the same paternal copy of CNP and therefore that for the CNP locus, both maternal chromosomes were represented in the affected individuals. Interestingly, despite this, all the affected individuals in the family are homozygous for the minor lower expression A allele at marker rs2070106 (see figure 3.6). Under the null hypothesis the prior probability that the single transmitted paternal chromosome transmitted would carry this allele is equal to the control minor allele frequency (0.3) while that of the mother being homozygous for this allele is the square of the allele frequency (0.09). Therefore given the known pattern of chromosomal transmissions in this family, the prior probability that by chance, both parents would have genotypes that would allow all affected individuals to be homozygous for the lower expression allele = 0.027 (0.3*0.09). If the presence of linkage is not allowed for, given the known parental genotypes (father heterozygous, mother uninformative homozygous), the corresponding probability of all the siblings being homozygous for the A allele is $0.5^6 = 0.015$ (the probability of the heterozygous father transmitting the A allele to all six of the siblings).

Figure 3.6 Family 702 pedigree.

Family 702 pedigree showing the six affected male siblings diagnosed with DSM-IV schizophrenia. The genotypes present at six CNP markers are given for the father and the six siblings, indicating the presence of a minimum of three parental chromosomes represented within the siblings.



3.2.3.2 Analysis of the complete CNP genomic sequence.

After performing the analysis on the exonic sequence of the family 702 pedigree I decided that in order to study the gene more comprehensively the entire 11 Kb of CNP genomic sequence should be screened for sequence variation in the 14 unrelated schizophrenic patients and also the family 702 pedigree. The screening of this additional sequence resulted in an additional 9 polymorphisms being detected (rs8080978, rs8080979, rs10540926, rs8078650, 1944 A>T, 3235 A>G, rs12952456, rs12950675 and 8962 C>T). All of these markers were then genotyped through the pooled association sample, but none showed any evidence of association with schizophrenia (see table 3.9). All of the additional SNPs were also genotyped in 96 unrelated control subjects in order to calculate the D' and r^2 values (see figure 3.8 and table 3.10). This genotyping was achieved by using two multiplex SNaPshot reactions as described in the methods section. An example of each of the resultant SNaPshot electrophorograms is shown in appendix 3.6 and 3.7 at the end of this chapter.

Figure 3.7 Schematic diagram of CNP.

Diagram of CNP (adapted from the UCSC Genome Browser), indicating the positions of all the genotyped polymorphisms. For numbering of polymorphisms, +1 indicates the 'A' of the most 5' ATG translation initiation codon. The gene is displayed (5'-3').

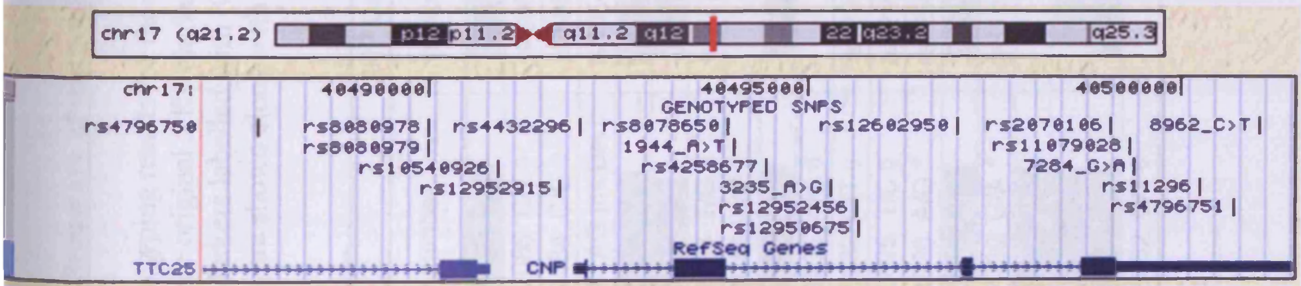


Table 3.9: Summary of the pooled genotyping data for all of the CNP polymorphisms studied in this project:

Pooled genotyping results for the CNP SNPs. Markers highlighted in grey were detected during the initial gene screening experiment or were studied in the original 5 Kb marker grid. Markers labelled with (*) were detected during the sequencing of the exonic sequence in the family 702 pedigree. Markers labelled with (◊) were detected during the sequencing of the remaining genomic sequence of CNP. The markers minor allele frequencies are shown along with the corresponding chi-square and p-values.

SNP Data			**Pooled Genotyping data**				
SNP number.	rs Number	UCSC position (HG 16, July 2003)	Minor Allele Cases (freq)	Minor Allele Controls (freq)	Difference	Chi-square	P-value
CNP -4359 G/A	rs4796750	40,487,709	A= (0.3)	A= (0.31)	0.01	0.3	0.58
CNP -2086 C/T ◊	rs 8080978	40,489,982	T= (0.2)	T= (0.22)	0.02	1.7	0.20
CNP -2084 C/G ◊	rs 8080979	40,489,984	G= (0.05)	G= (0.05)	0	0.2	0.63
CNP -1148 AG Ins/Del ◊	rs10540926	40,490,919	AG Del= (0.2)	AG Del= (0.22)	0.02	0.6	0.45
CNP -351 G/A	rs12952915	40,491,717	A= (0.22)	A= (0.22)	0	0.005	0.94
CNP -96 T/C	rs4432296	40,491,972	C= (0.33)	C= (0.32)	0.01	0.02	0.89
CNP 1868 T/G ◊	rs 8078650	40,493,936	G= (0.2)	G= (0.21)	0.01	0.53	0.47
CNP 1944 A/T ◊		40,494,011	T= (0.03)	T= (0.03)	0	0.006	0.94
CNP 2371 A/G	rs4258677	40,494,439	G= (0.12)	G= (0.11)	0.01	0.04	0.84
CNP 3235 A/G ◊		40,495,303	G= (0.28)	G= (0.3)	0.02	0.9	0.34
CNP 3582 C/T ◊	rs12952456	40,495,650	T= (0.19)	T= (0.2)	0.01	0.18	0.67
CNP 3625 T/C ◊	rs12950675	40,495,693	C= (0.15)	C= (0.17)	0.02	2.1	0.15
CNP 4926 A/G *	rs12602950	40,496,994	G= (0.29)	G= (0.31)	0.02	1.1	0.30
CNP 6961 G/A	rs2070106	40,499,029	A= (0.28)	A= (0.22)	0.06	3.5	**0.06**
CNP 7230 G/A *	rs11079028	40,499,298	A= (0.22)	A= (0.2)	0.02	1.2	0.27
CNP 7284 G/A		40,499,352	A= (0.13)	A= (0.13)	0	0.2	0.67
CNP 8157 T/C	rs11296	40,500,225	C= (0.1)	C= (0.11)	0.01	0.3	0.58
CNP 8634 C/T	rs4796751	40,500,702	T= (0.18)	T= (0.19)	0.01	0.3	0.58
CNP 8962 C/T ◊		40,501,030	T= (0.05)	T= (0.04)	0.01	1.5	0.22

Figure 3.8: LD structure of the CNP genomic region.

Haploview LD plot of the CNP genomic region displaying the D' values between the CNP markers and the LD block present as defined by Gabriel and colleagues (2002).

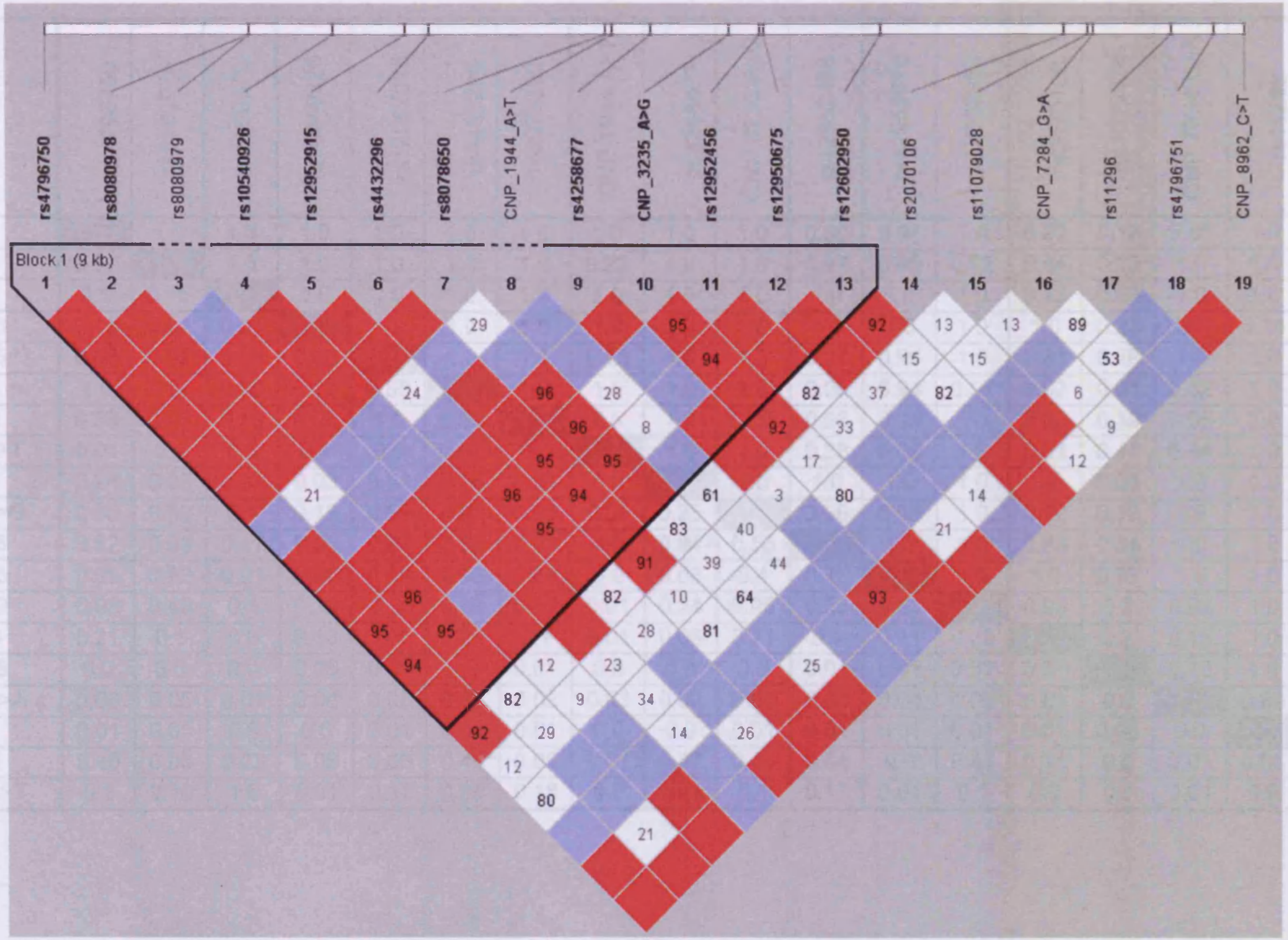


Table 3.10: Summary of the LD between CNP markers:

Summary of the LD analysis results for the CNP polymorphisms in 96 controls. The table shows the marker names and the D' and r^2 values. The results shown above the diagonal line of grey squares are D' values and the results shown below the diagonal line are r^2 values.

SNP ID	rs4796750	rs8080978	rs8080979	rs10540926	rs12952915	rs4432296	rs8078650	CNP 1944 A>T	rs4258677	CNP 3235 A>G	rs12952456	rs12950675	rs12602950	rs2070106	rs11079028	CNP 7284 G>A	rs11296	rs4796751	CNP 8962 C>T
rs4796750		1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.96	0.94	1.0	0.92	0.12	0.81	1.0	1.0	1.0
rs8080978	0.57		1.0	1.0	1.0	1.0	1.0	0.22	1.0	1.0	0.97	0.96	0.92	0.84	0.09	1.0	1.0	0.2	1.0
rs8080979	0.1	0.16		1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.13	0.09	1.0	1.0	1.0	1.0
rs10540926	0.13	0.08	0.01		1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.23	0.34	0.14	1.0	1.0
rs12952915	0.58	0.97	0.17	0.08		1.0	1.0	0.25	1.0	1.0	0.97	0.96	1.0	0.83	0.28	1.0	1.0	0.27	1.0
rs4432296	1.0	0.57	0.09	0.13	0.58		1.0	1.0	1.0	1.0	0.96	0.94	0.97	0.92	0.01	0.82	1.0	1.0	1.0
rs8078650	0.59	0.93	0.18	0.08	1.0	0.59		0.3	1.0	0.96	0.96	0.95	1.0	0.83	0.39	0.65	1.0	0.25	1.0
CNP 1944 A>T	0.01	0.0	0.0	0.1	0.0	0.01	0.0		1.0	1.0	0.28	0.09	1.0	0.61	0.41	0.44	1.0	1.0	1.0
rs4258677	0.24	0.03	0.01	0.03	0.04	0.24	0.03	0.0		1.0	1.0	1.0	1.0	1.0	0.03	1.0	1.0	0.93	1.0
CNP 3235 A>G	0.98	0.59	0.1	0.13	0.59	0.97	0.58	0.01	0.25		0.96	0.95	1.0	0.92	0.17	0.8	1.0	1.0	1.0
rs12952456	0.52	0.93	0.17	0.08	0.93	0.52	0.93	0.0	0.04	0.53		1.0	1.0	0.83	0.34	1.0	1.0	0.22	1.0
rs12950675	0.39	0.72	0.01	0.06	0.69	0.38	0.64	0.0	0.03	0.4	0.75		1.0	1.0	0.37	1.0	1.0	0.14	1.0
rs12602950	0.98	0.49	0.1	0.13	0.56	0.93	0.57	0.01	0.24	0.95	0.55	0.41		0.93	0.2	0.82	1.0	1.0	1.0
rs2070106	0.21	0.1	0.0	0.13	0.1	0.21	0.1	0.01	0.06	0.21	0.1	0.11	0.23		0.2	0.15	1.0	1.0	0.13
rs11079028	0.0	0.0	0.0	0.05	0.01	0.0	0.01	0.0	0.0	0.0	0.01	0.01	0.01	0.01		0.13	1.0	0.05	0.09
CNP 7284 G>A	0.06	0.05	0.01	0.08	0.05	0.06	0.02	0.03	0.02	0.06	0.05	0.04	0.06	0.01	0.0		0.89	0.53	1.0
rs11296	0.01	0.01	0.0	0.0	0.01	0.01	0.01	0.0	0.0	0.01	0.01	0.01	0.01	0.01	0.08	0.0		1.0	1.0
rs4796751	0.46	0.03	0.23	0.06	0.06	0.44	0.05	0.01	0.47	0.45	0.04	0.0	0.44	0.11	0.0	0.01	0.01		1.0
CNP 8962 C>T	0.1	0.16	1.0	0.01	0.17	0.09	0.18	0.0	0.01	0.1	0.17	0.01	0.1	0.0	0.0	0.01	0.0	0.23	

3.2.3.3 Additional entropy analysis.

Following the comprehensive mutation detection that was carried out on the CNP locus in both the family 702 pedigree and the mutation screening set I performed an additional entropy analysis on the 19 CNP polymorphisms. By calculating the haplotype frequencies of all the markers, it was possible to utilise the entropy analysis technique, to determine the subset of markers that would extract >95% of the genetic information contained within the gene that would have been captured by individually genotyping all of the nineteen discovered SNPs. This entropy analysis revealed that by individually typing a further six additional markers (rs4796750, rs10540926, 3235 A>G, rs12602950, rs11079028 and 7284 G>A) in addition to the four markers that had already been genotyped individually, it would be possible to extract >98% of the haplotypic diversity that would have been provided by genotyping all of the markers (see table 3.11).

Table 3.11. Results of Entropy analysis.

Entropy analysis results for the 19 CNP polymorphisms. The number of polymorphisms required to be genotyped and the maximum percentage of haplotypic diversity that could be extracted by genotyping those polymorphisms is shown.

Number of markers to be genotyped	Maximum percentage of haplotypic diversity extracted
1	0.24
2	0.44
3	0.62
4	0.76
5	0.88
6	0.96
7	0.98
8	0.99
9	1.0
10	1.0

Genotyping of the additional six markers was performed using a Sequenom MALDI-TOF multiplex assay, as described in the materials and methods chapter. No evidence for association with schizophrenia at either the genotypic or allelic level was found (see Table 3.12).

Table 3.12: Additional individual genotyping results for CNP markers in the schizophrenia case control sample:

Individual genotyping data produced when the CNP markers (rs4796750, rs10540926, 3235 A>G, rs12602950, rs11079028, and 7284 G>A) were typed in the schizophrenia case control sample. Genotypic and allelic counts, the allele frequencies and the corresponding chi-square and p-values for the markers are given.

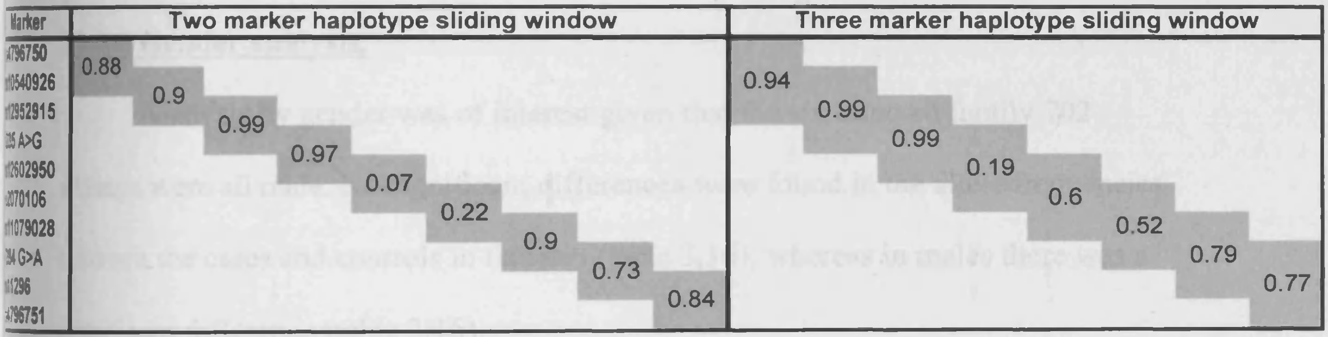
SNP Information	Summary of Genotypic data			Summary of Allelic data	
rs4796750 G>A	G/G	G/A	A/A	G	A
Cases	330	240	46	900 (0.73)	332 (0.27)
Controls	347	254	55	948 (0.72)	364 (0.28)
	$\chi^2 = 0.37$ (2df)	P = 0.83		$\chi^2 = 0.20$ (1df)	P = 0.65
rs10540926 AG Ins/Del	Ins/Ins	Ins/Del	Del/Del	AG Ins	AG Del
Cases	414	194	23	1022 (0.81)	240 (0.19)
Controls	435	203	28	1073 (0.81)	259 (0.19)
	$\chi^2 = 0.27$ (2df)	P = 0.87		$\chi^2 = 0.08$ (1df)	P = 0.78
3235 A>G	A/A	A/G	G/G	A	G
Cases	331	238	56	900 (0.72)	350 (0.28)
Controls	344	265	55	953 (0.72)	375 (0.28)
	$\chi^2 = 0.53$ (2df)	P = 0.77		$\chi^2 = 0.02$ (1df)	P = 0.89
rs12602950 A>G	A/A	A/G	G/G	A	G
Cases	347	259	61	953 (0.71)	381 (0.29)
Controls	365	283	62	1013 (0.71)	407 (0.29)
	$\chi^2 = 0.18$ (2df)	P = 0.91		$\chi^2 = 0.004$ (1df)	P = 0.95
rs11079028 G>A	G/G	G/A	A/A	G	A
Cases	334	173	25	841 (0.79)	223 (0.21)
Controls	381	177	28	939 (0.80)	233 (0.2)
	$\chi^2 = 0.69$ (2df)	P = 0.71		$\chi^2 = 0.39$ (1df)	P = 0.53
7284 G>A	G/G	G/A	A/A	G	A
Cases	475	144	9	1094 (0.87)	162 (0.13)
Controls	506	148	10	1160 (0.87)	168 (0.13)
	$\chi^2 = 0.08$ (2df)	P = 0.96		$\chi^2 = 0.04$ (1df)	P = 0.85

3.2.3.4 Haplotype analysis.

All possible haplotype combinations were analysed, up to and including the ten marker haplotype (data not shown). The performance of exhaustive haplotype analysis was necessary as the percentage of haplotypic diversity predicted to be captured by the entropy method is based upon the subsequent analysis of all possible haplotypes. None of the global p-values were more significant than the result detected individually for marker rs2070106 (see table 3.13), (table 3.13 only shows 2 and 3 marker sliding windows for ease of presentation).

Table 3.13. CNP Haplotype analysis.

CNP haplotype analyses results. The two and three marker haplotype sliding windows are shown, with their corresponding global p-values.



Marker	Two marker haplotype sliding window	Three marker haplotype sliding window
4796750	0.88	0.94
10540926	0.9	0.99
12952915	0.99	0.99
25 A>G	0.97	0.19
12602950	0.07	0.6
2070106	0.22	0.52
11079028	0.9	0.79
34 G>A	0.73	0.77
1296	0.84	
4796751		

3.2.3.5 Familial schizophrenia analysis:

Given the previous evidence reporting linkage of the region of chromosome 17 containing CNP to schizophrenia in a multiply affected family (Williams, Norton et al. 2003), I decided to analyse the rs2070106 data in the subset of our association sample which had a positive family history of schizophrenia. In the association sample a patient is classified as having a positive family history if they have at least one affected first degree relative. When this subset of the cases was tested against the control sample there was no evidence for association between marker rs2070106 and susceptibility to familial schizophrenia at either the genotypic (p=0.29) or allelic (p=0.12) level (see table 3.14).

Table 3.14: Individual genotyping of rs2070106 in the familial schizophrenia sample:

Individual genotyping results for marker rs2070106 in the familial subset of the schizophrenia association sample. Genotypic and allelic counts as well as allele frequencies are given, with the corresponding chi-square and p-values.

SNP	Genotypes			Alleles	
rs2070106 G>A	G/G	G/A	A/A	G	A
Cases	58	69	16	185 (0.65)	101 (0.35)
Controls	325	303	59	953 (0.7)	421 (0.3)
	$\chi^2 = 2.51$ (2df)	P = 0.29		$\chi^2 = 2.40$ (1df)	P = 0.12

3.2.3.6 Gender analysis.

Analysis by gender was of interest given that the six affected family 702 siblings were all male. No significant differences were found in the allele frequencies between the cases and controls in females (table 3.16), whereas in males there was a significant difference (table 3.15).

Table 3.15: Individual genotyping data for marker rs2070106 in the male sample:

Summary of the rs2070106 genotyping data in males. Genotypic and allelic counts, the allele frequencies and the corresponding chi-square and p-values for the data are given.

SNP	Genotypes			Alleles	
rs2070106 G>A	G/G	G/A	A/A	G	A
Cases	190	216	58	596 (0.64)	332 (0.36)
Controls	214	202	41	630 (0.69)	284 (0.31)
	$\chi^2 = 4.76$ (2df)	P = 0.09		$\chi^2 = 4.58$ (1df)	P = **0.03**

Table 3.16: Individual genotyping data for marker rs2070106 in the female sample:

Summary of the rs2070106 genotyping data in females. Genotypic and allelic counts, the allele frequencies and the corresponding chi-square and p-values for the data are given.

SNP	Genotypes			Alleles	
rs2070106 G>A	G/G	G/A	A/A	G	A
Cases	98	103	13	299 (0.69)	135 (0.31)
Controls	108	98	18	314 (0.70)	134 (0.30)
	χ^2 = 1.19 (2df)	P = 0.55		χ^2 = 0.15 (1df)	P = 0.70

If the observed male gender effect is a genuine finding and not the result of chance or small female sample size, then a significant difference between the male and female cases might be expected to exist. When the analysis was performed a significant difference between the genotypes and also a trend towards significance between the allele frequencies present in the male and female cases was revealed (table 3.17). The analysis was also repeated in the control sample to determine whether the observed gender difference was present in the general healthy population. This analysis revealed no significant difference between either the genotypic or allelic frequencies in the male and female controls (table 3.18).

Table 3.17: Genotyping data for rs2070106 in male and female cases:

Genotyping data for rs2070106 in the male and female schizophrenic patients. Genotypic and allelic counts, the allele frequencies and the corresponding chi-square and p-values for the data are given.

SNP	Genotypes			Alleles	
rs2070106 G>A	G/G	G/A	A/A	G	A
Male Cases	190 (0.41)	216 (0.46)	58 (0.13)	596 (0.64)	332 (0.36)
Female Cases	98 (0.46)	103 (0.48)	13 (0.06)	299 (0.69)	135 (0.31)
	χ^2 = 6.66 (2df)	P = **0.03**		χ^2 = 2.86 (1df)	P = 0.09

Table 3.18: genotyping data for rs2070106 in male and female controls:

Genotyping data for rs2070106 in the male and female controls. Genotypic and allelic counts, the allele frequencies and the corresponding chi-square and p-values for the data are given.

SNP	Genotypes			Alleles	
rs2070106 G>A	G/G	G/A	A/A	G	A
Male Controls	214 (0.47)	202 (0.44)	41 (0.09)	630 (0.69)	284 (0.31)
Female Controls	108 (0.48)	98 (0.44)	18 (0.08)	314 (0.70)	134 (0.30)
	$\chi^2 = 0.22$ (2df)	P = 0.89		$\chi^2 = 0.19$ (1df)	P = 0.66

3.3 Discussion:

I screened four genes (CNP, MAG, TF and ErbB3) reported, in the study (Hakak, Walker et al. 2001) which formed the inspiration for this thesis, to be down-regulated in schizophrenia for the presence of *cis*-acting sequence variants affecting the genes expression in human brain. Altered gene expression can be a consequence of either *cis*- or *trans*-acting effects. If altered expression plays a primary aetiological role in disease, rather than being either involved in the pathophysiological pathway or being a secondary effect of the disease, it is the former *cis*-acting genetic effects that are of relevance. I therefore first sought to confirm whether expression of any of the genes was in fact influenced by cis-acting polymorphisms. The results of the allelic expression analyses clearly confirmed the presence of such *cis*-acting variants only in the case of CNP. The observation that in every sample, the A allele of CNP marker rs2070106 is under expressed compared to the G allele suggests that either the marker SNP itself is responsible for the altered expression, or that the actual functional variant is in strong linkage disequilibrium (LD) with the marker SNP.

If it is postulated that the under expression of CNP in schizophrenia is of aetiological relevance then the mechanisms responsible for this phenomenon can be expected to be associated with the disorder. This leads to the following specific hypothesis. If CNP under expression is aetiologically significant then the lower expression A allele of marker rs2070106 will be associated with disease. When the marker was genotyped individually in the case-control sample the results were consistent with this hypothesis, although they only just meet the conventional criteria for significance (p = 0.04, OR = 1.2, 95% CI (1.0-1.40)). Additional analysis of the data revealed a slightly stronger association in males (p=0.03), along with a significant difference between the male and female cases (see table 3.17).

Following these initial results I performed further direct and indirect genetic analysis of CNP, in order to look for non-synonymous SNPs, functional variants in LD with rs2070106 and any further association signals. This analysis involved the use of a dense set of markers as well as ultimately sequencing the complete genomic sequence of CNP and its flanking sequence. A total of nineteen sequence variants were detected and subsequently analysed for association (see figure 3.7 and table 3.9). Analysis of these polymorphisms revealed no further evidence for association to schizophrenia, either at the single marker or haplotype level. This would suggest that if the original finding reflects a true association, marker rs2070106 is likely to be the schizophrenia susceptibility variant *per se*, or is in LD with a marker beyond the region that I sequenced.

I have also tried to determine whether variation at the CNP locus might be responsible for the previous report of linkage to 17q in a single multiply affected family. In order to do this I sequenced the exons of CNP in all the family members of the 702 pedigree. This analysis did not reveal any common polymorphisms or mutations in the exonic sequence that might account for the linkage. Interestingly, the results of this analysis did reveal that the 17q region of the genome where CNP resides had a lower level of IBD sharing than had been previously predicted by the linkage study. However, all of the affected siblings within the pedigree were homozygous for the lower expression A allele of marker rs2070106. This observation of transmissions in the siblings would be expected to occur by chance around 3 times in 100 (0.03), and provides independent support for CNP playing a role in susceptibility to schizophrenia. It should be noted, that as the putative susceptibility allele has a homozygosity rate of 10% in the general population, the allele is unlikely to be a highly penetrant recessive genotype, and therefore unless there are loci that interact with rs2070106 fixed in the family cannot account for the reported linkage finding in the multiply affected family on its own.

Whilst it has been possible during this project to confirm the presence of *cis*-acting sequence variants that affect the expression of the CNP gene, the expression assay does not allow the exact location or function of these variants to be determined. The location of *cis*-acting variants may be in transcribed sequence itself, promoters, intronic sequence or in distant regulatory elements (Levine and Tjian 2003), making it a very difficult proposition to locate the variants. Although mutation analysis of CNP did not reveal polymorphisms predicted to change the amino acid sequence of the gene, rs2070106 showing modest association to schizophrenia and associated with lower expression is located in the genes coding sequence close to the 3'UTR, where the occurrence of SNPs may affect the genes expression. At present, bioinformatics analysis (http://pupasnp.bioinfo.cnio.es/) does not predict any functional role for rs2070106. However, there have been examples reported where synonymous SNPs have been shown to affect gene expression (Duan, Wainwright et al. 2003; Capon, Allen et al. 2004; Pagani, Raponi et al. 2005), so it is possible this marker itself is responsible for the altered expression. However, it should be noted that an alternative explanation would be that the association and altered expression observed at this marker are as a result of strong linkage disequilibrium with the true functional variant. If this is true, then it is unlikely that the functional variant is located within the genes exonic sequence or introns as these have been comprehensively screened as part of this project. As mentioned earlier these could be located either distal or proximal to the gene in regulatory elements. Analysis of the available HapMap data revealed the presence of strong LD extending up to 300kb from the CNP locus.

It is also of note that rare sequence variants with small minor allele frequencies may not have been detected during this project. The screening set used in the project to perform mutation detection provided a power of 80% to detect polymorphisms with a minor allele frequency of 0.05, and a 95% power to detect polymorphisms with a minor allele frequency of 0.1. This means it is likely that polymorphisms with minor allele frequencies of around 0.05 or lower will not have been detected and tested for association to schizophrenia. Recent advances in mutation screening technology and the advent of high throughput genotyping techniques mean that if the project was being repeated today it would now be feasible to perform mutation screening on large numbers (hundreds) of cases and controls and thus allow rare polymorphisms with low minor allele frequencies to be detected and tested for association. For example a screening set of one hundred cases and one hundred controls would provide a power of 98% to detect polymorphisms with a minor allele frequency of 0.01 and a power of 87% to detect polymorphisms with a minor allele frequency of 0.005.

A further point is the design of the project itself. At the time the project was conceived, the genetic databases and genome information available on the internet were limited compared to what is available now. If the project was being designed at the present time I would perform it differently. The rapid expansion in the genetic information available combined with the release of the latest version of the HapMap project allows for the selection of specific 'tag' markers in order to screen genes using an LD mapping approach, as well as allowing for the use of haplotype analysis. The method allows the selection of the minimal set of markers that are required to be genotyped in order to obtain the genetic information that would be extracted by typing all of the genetic markers contained within a gene either individually or through the use of haplotype analysis. Even at the present time there are only five CNP markers with genotype information for the European-American population contained within the HapMap database, including rs2070106. Analysis of the data revealed these markers to all be singleton tag markers. Using my earlier entropy analysis it was possible to calculate that if the five markers were genotyped they would extract about 80% of the haplotypic diversity that would have been obtained by typing all of the 19 detected SNPs in CNP. Although the HapMap database currently does not contain all of the genetic diversity within CNP, it does provide reasonable coverage of the genetic information contained within the gene and would still be a good starting point for the genes investigation.

Recently it has been shown that CNP interacts with mitochondria and cytoskeletal proteins and may act to promote microtubule assembly or act as a membrane anchor for tubulin and as such CNP has been proposed as a potential candidate responsible for coordinating the reorganisation of the cytoskeleton during the process outgrowth of oligodendrocytes. This theory is supported by the results of several studies. Firstly, Bifulco and colleagues (Bifulco, Laezza et al. 2002) demonstrated that CNP was a membrane bound microtubule associated protein that was able to link tubulin to membranes. The authors proposed that CNP may regulate cytoplasmic microtubule distribution, by acting as a membrane anchor for microtubules. Secondly, Lee and colleagues (Lee, Gravel et al. 2005) demonstrated that CNP interacts with tubulin heterodimers and promotes microtubule assembly. The authors additionally demonstrated that CNP was able to induce microtubule and F-actin reorganisation, which are both necessary procedures for process outgrowth and arborisation in oligodendrocytes. During myelinogenesis oligodendrocytes create highly branched processes that target and wrap around axons to form the myelin sheath. The presence of the myelin sheath is essential for the successful and rapid delivery of action potentials along the axons. This important role of oligodendrocytes is itself critically dependant upon the establishment of an adequate oligodendrocyte support structure. This is provided by the cytoskeleton, which is constructed from microtubules and microfilaments (Lee, Gravel et al. 2005). Disruption to the structure of this cytoskeleton or its component parts might have a critical impact on the functionality of oligodendrocytes and subsequently upon the formation of the myelin sheath and the function of axons themselves.

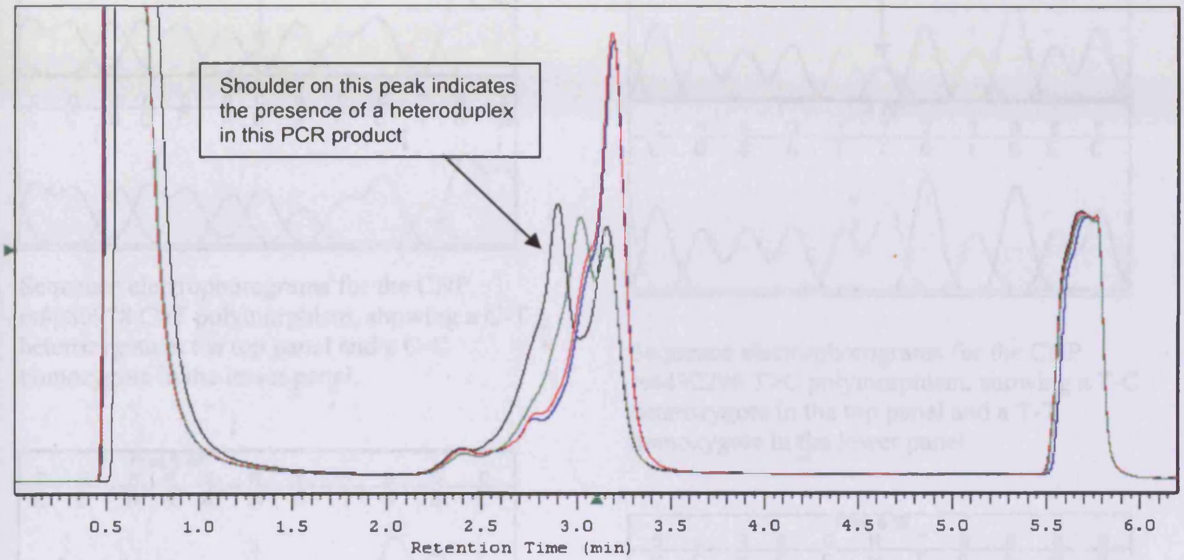
Also, recent analysis of CNP-null mice has revealed that deficiency of CNP leads to disrupted axoglial interactions in the CNS and also major abnormalities in the paranodes and paranodal loops (Rasband, Tayler et al. 2005). The paranodal loops are the principal site of axoglial contacts and have been shown to influence the regulation of several important axonal functions including the clustering of ion channels and cell adhesion molecules, and also the control of axonal survival (Rasband, Tayler et al. 2005). It has therefore been suggested that CNP is required for maintaining the integrity of paranodes and that disrupted axoglial signalling caused by a deficiency in CNP underlies the progressive axonal death observed in the CNS of CNP-null mice (Rasband, Tayler et al. 2005). Given the reduced expression of CNP observed in schizophrenic brain and the proposed role of CNP in maintaining both the structure and function of the cytoskeleton and also the integrity of paranodes and axoglial signalling. The gene may play an important role in causing the oligodendrocyte and myelination abnormalities observed in schizophrenic brain.

In conclusion I have shown that CNP expression is under the influence of cis-acting polymorphism, and that consistent with previous micro-array findings of reduced CNP expression in schizophrenic brain, the allele associated with lower CNP expression is also associated with schizophrenia. The findings, while supported by a strong directional prior hypothesis and by further observations in a family showing linkage to the region, are weak and certainly require independent replication in large suitably powered association samples. This might be problematic given the power required to detect markers with small effect size. For example, in the case of rs2070106 with a minor allele frequency (controls) of 0.3, an odds ratio of 1.2 and a disease prevalence of 0.01, a sample of 1100 cases and 1100 controls is required to provide an 80% power to detect association at p=0.05, and a sample of 1500 cases and 1500 controls is required to provide a 90% power to detect association at p=0.05. Nevertheless, the results provide support for the specific hypothesis that reduced CNP expression in schizophrenic brain may be aetiologically relevant to schizophrenia. Moreover, the data also provides support for the more general hypothesis that altered myelination and/or oligodendrocyte function may play a role in schizophrenia aetiology. These findings therefore provide further inspiration to study other OMR genes in schizophrenia.

3.4 Appendix: Samples of Experimental data.

Appendix 3.1: DHPLC chromatograph.

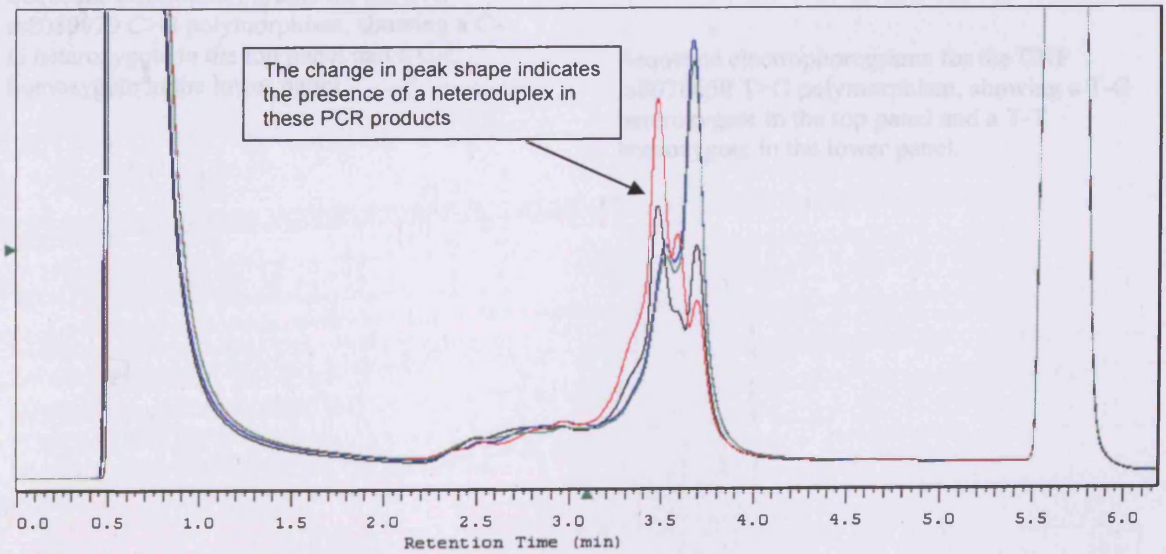
Example of DHPLC traces produced for CNP amplimer b1 analysed at 64˚C, for screening set samples 6, 12, 13 and 14. The graph shows the peaks representing these samples PCR products, and changes in the peak shape indicates the presence of a heteroduplex.



Shoulder on this peak indicates the presence of a heteroduplex in this PCR product

Appendix 3.2: DHPLC chromatograph.

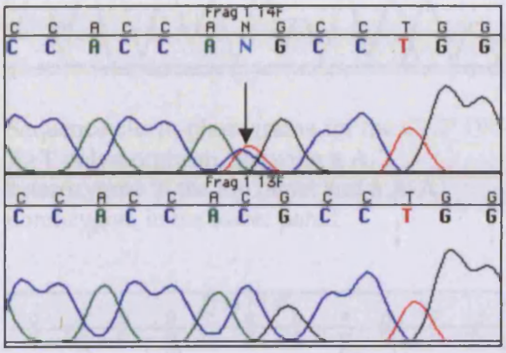
Example of DHPLC traces produced for CNP amplimer b20 analysed at 63˚C, for screening set sample 12 and three members of the family 702 pedigree. The graph shows the peaks representing these samples PCR products, and changes in the peak shape indicates the presence of a heteroduplex.



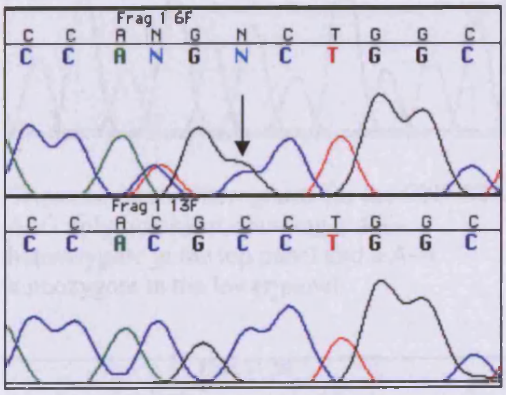
The change in peak shape indicates the presence of a heteroduplex in these PCR products

Appendix 3.3: Sequencing electrophorograms for CNP.

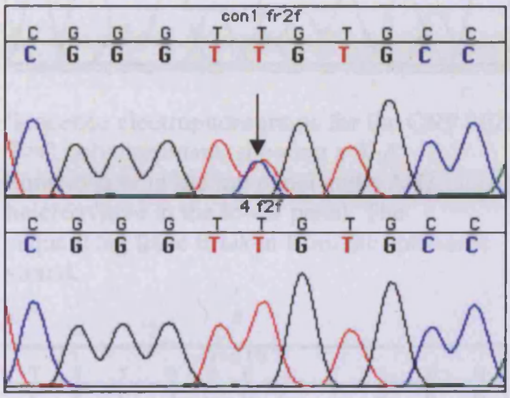
This appendix shows examples of the different sequencing electrophorograms that were produced during the classification of each of the polymorphisms detected as a result of screening the entire CNP genomic sequence. The polymorphism is indicated with an arrow.



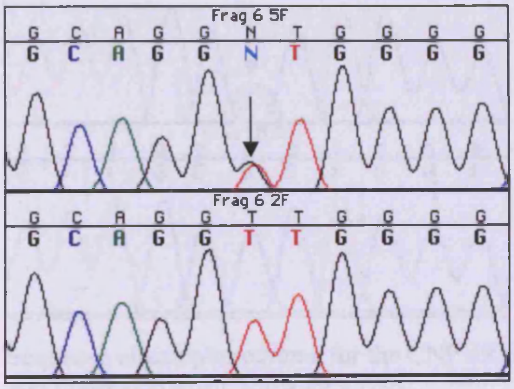
Sequence electrophorograms for the CNP rs8080978 C>T polymorphism, showing a C-T heterozygote in the top panel and a C-C homozygote in the lower panel.



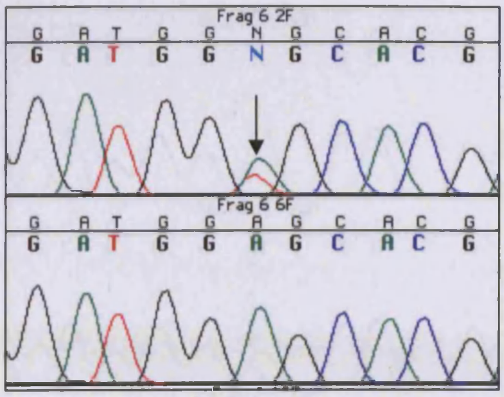
Sequence electrophorograms for the CNP rs8080979 C>G polymorphism, showing a C-G heterozygote in the top panel and a C-C homozygote in the lower panel.



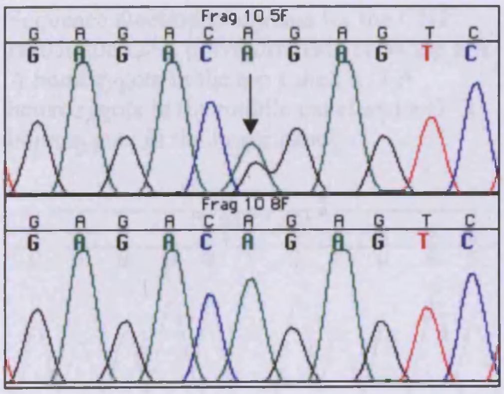
Sequence electrophorograms for the CNP rs4432296 T>C polymorphism, showing a T-C heterozygote in the top panel and a T-T homozygote in the lower panel.



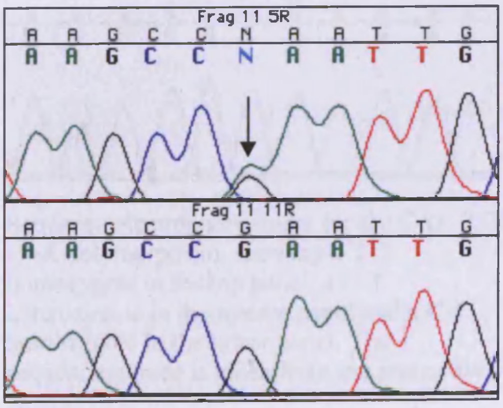
Sequence electrophorograms for the CNP rs8078650 T>G polymorphism, showing a T-G heterozygote in the top panel and a T-T homozygote in the lower panel.



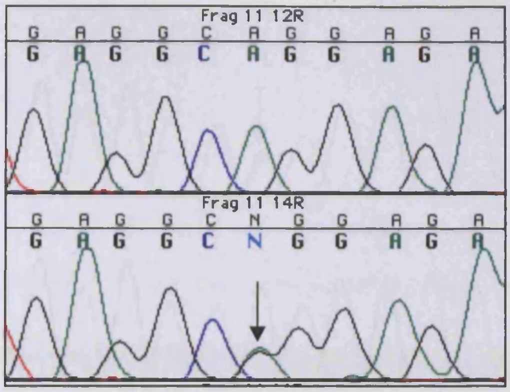
Sequence electrophorograms for the CNP 1944 A>T polymorphism, showing a A-T heterozygote in the top panel and a A-A homozygote in the lower panel.



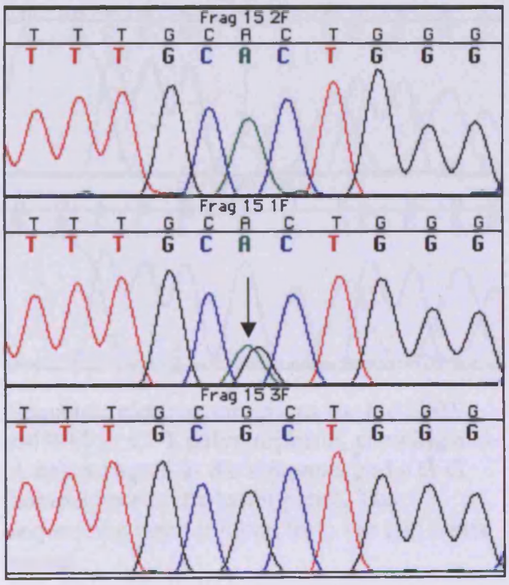
Sequence electrophorograms for the CNP 3235 A>G polymorphism, showing a A-G heterozygote in the top panel and a A-A homozygote in the lower panel.



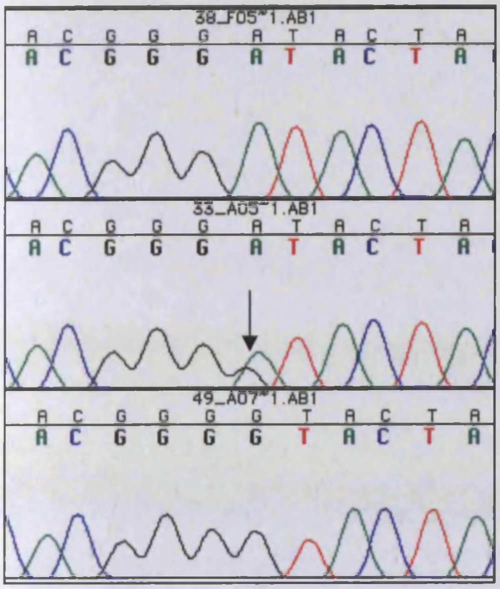
Sequence electrophorograms for the CNP 3582 C>T polymorphism, showing a G-A heterozygote in the top panel and a G-G homozygote in the lower panel. The sequencing trace is taken from the anti-sense strand.



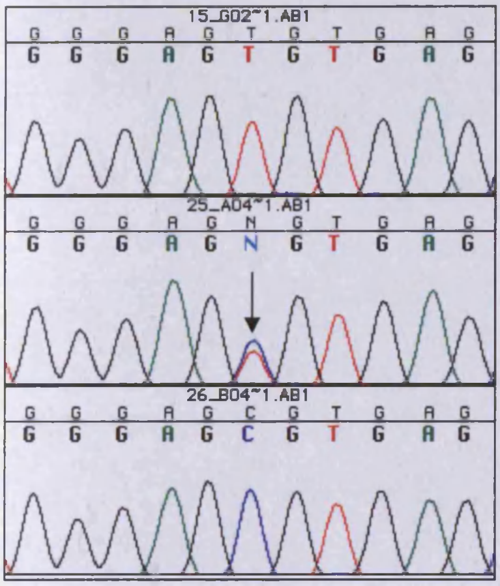
Sequence electrophorograms for the CNP 3625 T>C polymorphism, showing a A-A homozygote in the top panel and a A-G heterozygote in the lower panel. The sequencing trace is taken from the anti-sense strand.



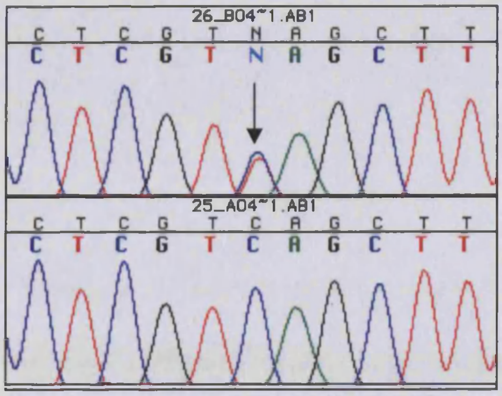
Sequence electrophorograms for the CNP 4926 A>G polymorphism, showing a A-A homozygote in the top panel, a A-G heterozygote in the middle panel and a G-G homozygote in the lower panel.



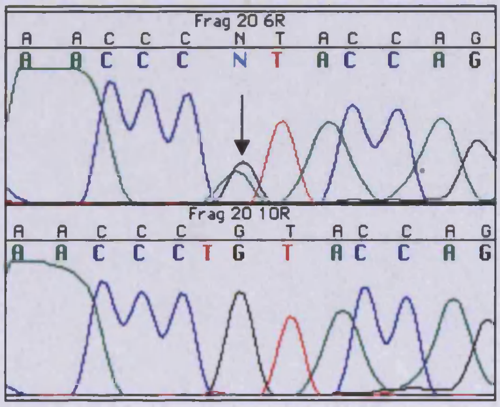
Sequence electrophorograms for the CNP rs2070106 G>A polymorphism, showing a A-A homozygote in the top panel, a G-A heterozygote in the middle panel and a G-G homozygote in the lower panel.



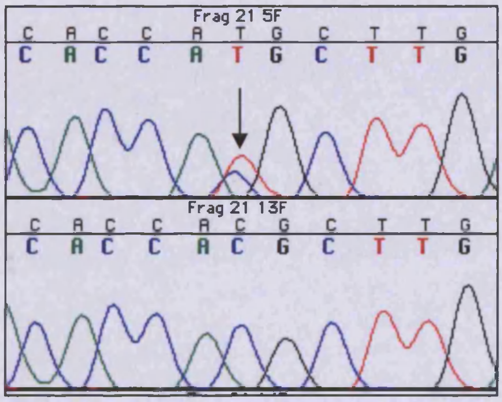
Sequence electrophorograms for the CNP 7230 G>A polymorphism, showing a T-T homozygote in the top panel, a C-T heterozygote in the middle panel and a C-C homozygote in the lower panel. The sequencing trace is taken from the anti-sense strand.



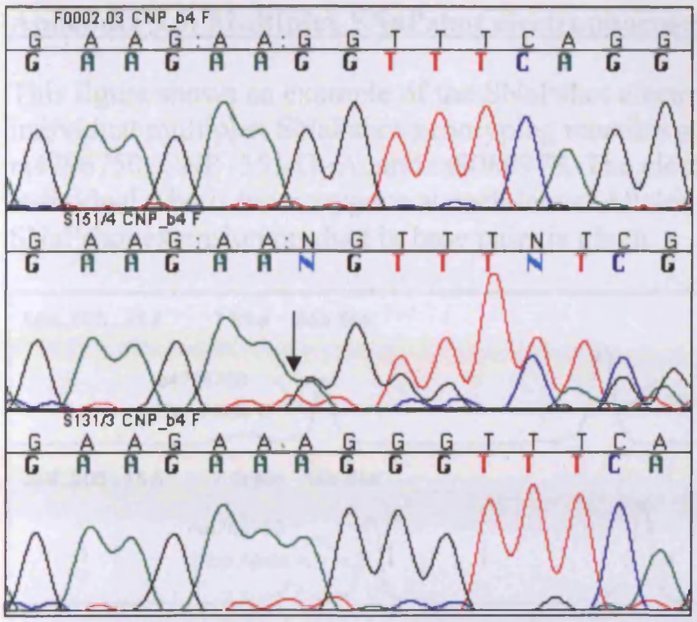
Sequence electrophorograms for the CNP 7284 G>A polymorphism, showing a C-T heterozygote in the top panel and a C-C homozygote in the lower panel. The sequencing trace is taken from the anti-sense strand.



Sequence electrophorograms for the CNP rs4796751 C>T polymorphism, showing a G-A heterozygote in the top panel and a G-G homozygote in the lower panel. The sequencing trace is taken from the anti-sense strand.



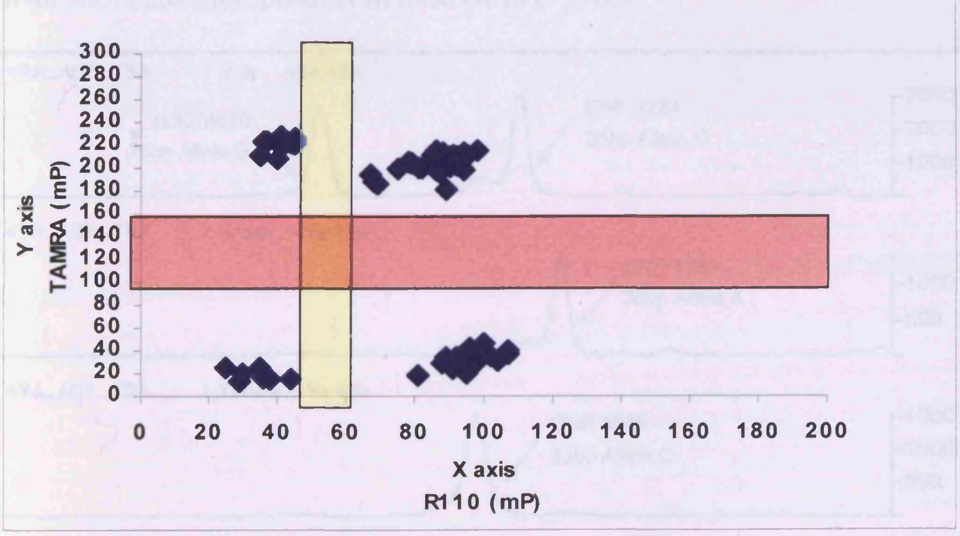
Sequence electrophorograms for the CNP 8962 C>T polymorphism, showing a C-T heterozygote in the top panel and a C-C homozygote in the lower panel.



Sequence electrophorograms for the CNP -1148 AG Ins/Del polymorphism, showing a AG-Del homozygote in the top panel, a AG-Ins heterozygote in the middle panel and an AG-Ins homozygote in the lower panel.

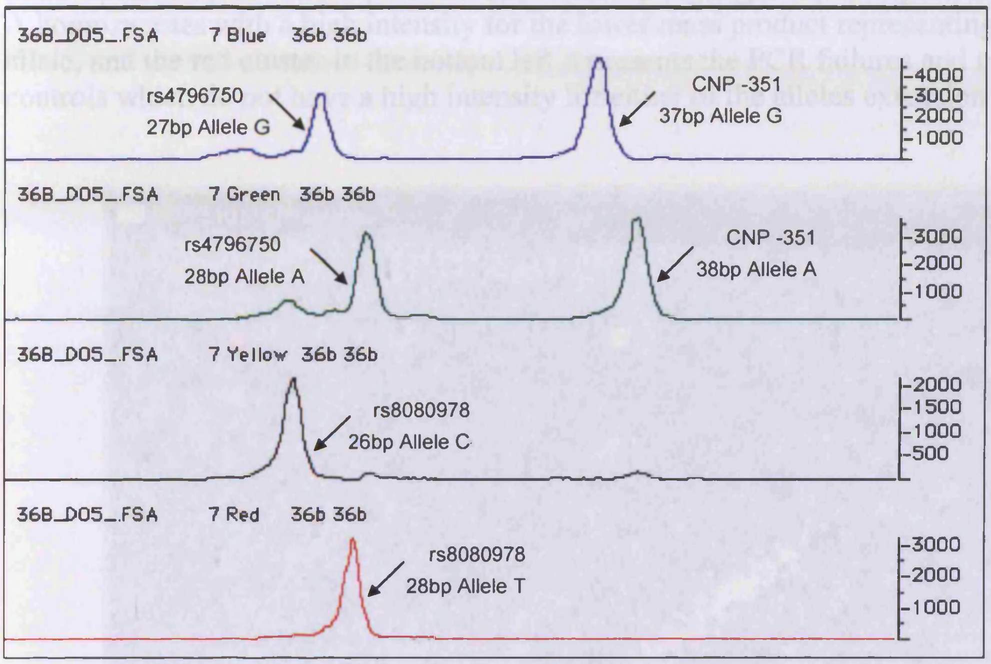
Appendix 3.4: FP genotyping graph for rs2070106.

This figure shows an example of an FP genotyping graph produced during individual genotyping of CNP rs2070106. The FP values cluster into four groups. The upper left corner of the plot represents the homozygote samples for allele A with a high FP for TAMRA, the right upper corner of the plot shows the heterozygote samples with a high FP for TAMRA and ROX, the lower right corner of the plot shows the homozygote samples for allele G with a high FP for ROX, and the bottom left corner of the graph shows the failed PCR reactions and negative controls with low FP values for both TAMRA and ROX.



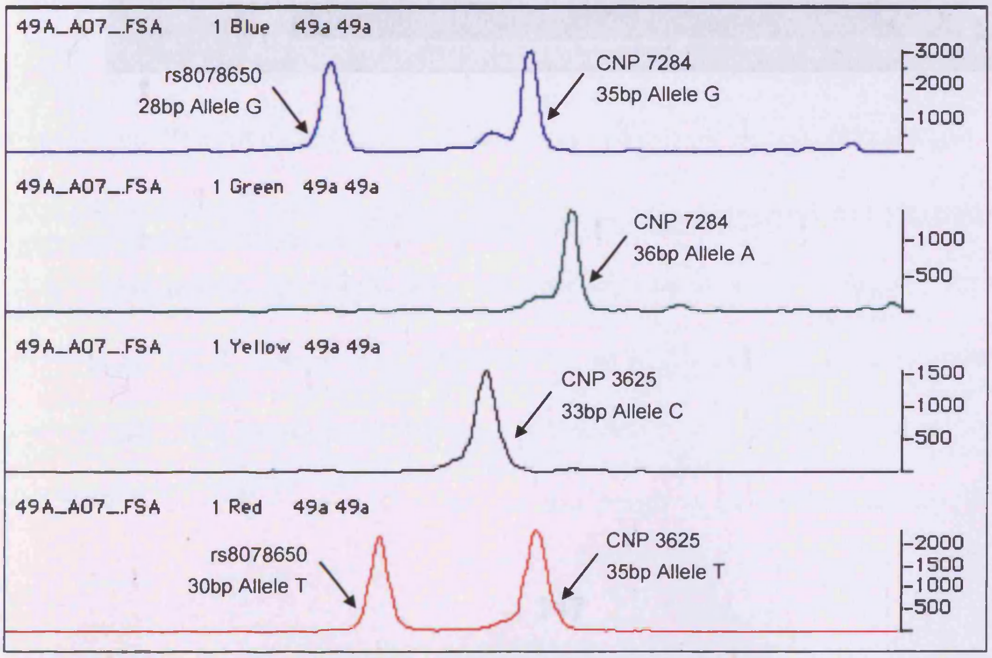
Appendix 3.5: Multiplex SNaPshot electrophorogram for CNP:

This figure shows an example of the SNaPshot electrophorograms produced by the individual multiplex SNaPshot genotyping reaction performed on the CNP markers rs4796750, CNP -351 G>A, and rs8080978. The electrophorogram shows an individual who is heterozygous at each locus. Alleles are labelled and the size of each SNaPshot extension product in base pairs is given.



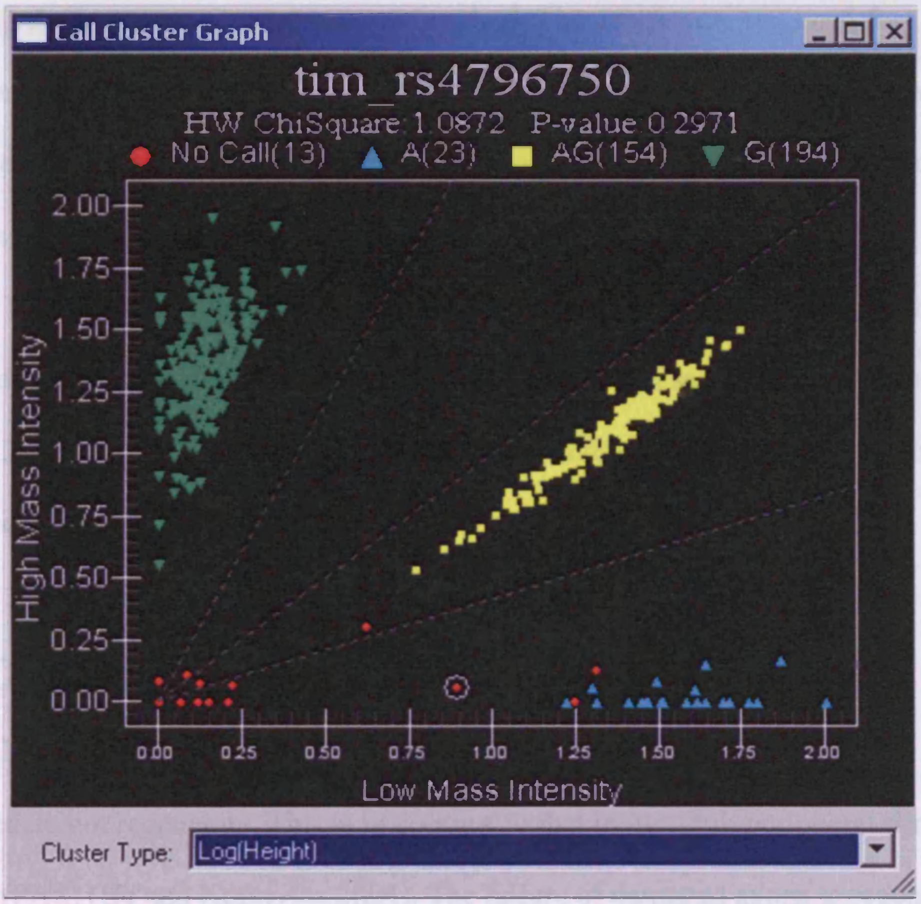
Appendix 3.6: Multiplex SNaPshot electrophorogram for CNP:

This figure shows an example of the SNaPshot electrophorograms produced by the individual multiplex SNaPshot genotyping reaction performed on the CNP markers rs8078650, CNP 7284 G>A, and CNP 3625 C>T. The electrophorogram shows an individual who is heterozygous at each locus. Alleles are labelled and the size of each SNaPshot extension product in base pairs is given.



Appendix 3.7: Sequenom MALDI-TOF genotyping output graph:

This figure shows an example of a Sequenom genotyping output graph for the CNP marker rs4796750. The graph displays a plot of the genotype intensities. The green cluster in the top left represents the G-G homozygotes with a high intensity for the higher mass extension product representing the G allele, the yellow cluster in the mid right represents the G-A heterozygotes with a high intensity for both the high and low mass extension products, the blue cluster in the bottom right corner represents the A-A homozygotes with a high intensity for the lower mass product representing the A allele, and the red cluster in the bottom left represents the PCR failures and negative controls which do not have a high intensity for either of the alleles extension products.



Cluster Type: Log(Height)

Chapter 4. The NOGO receptor complex and schizophrenia:

4.1.1. Introduction:

In this chapter I report analysis of four genes involved in the NOGO receptor complex: Neurite outgrowth inhibitor (NOGO), NOGO receptor (NGR), Oligodendrocyte myelin glycoprotein (OMG) and Nerve growth factor receptor (NGFR). A direct genetic approach was applied. The genes were screened for polymorphisms and the identified SNPs tested for association with schizophrenia in DNA pools. The genes were also assayed for the presence of *cis*-acting sequence variants affecting their expression in brain.

The NOGO complex genes were selected for study primarily on the basis of reported over-expression of NOGO mRNA in post mortem schizophrenic brain (Novak, Kim et al. 2002), supported by positional and/or functional evidence for their involvement in schizophrenia.

4.1.2. The NOGO Receptor Complex:

In the adult central nervous system (CNS), axons which are lesioned or damaged do not regenerate. This is in contrast to that in the adult peripheral nervous system (PNS) (He and Koprivica 2004). The failure of damaged axons to regenerate in the adult CNS is thought to be the result of inhibitory factors (Woolf and Bloechlinger 2002). Much of this inhibition has been attributed to myelin (Filbin 2003). Myelination appears to coincide with the loss of neuronal regenerative ability (Novak, Kim et al. 2002). Oligodendrocytes, as well as making myelin, produce growth inhibitory proteins that become embedded within the myelin sheath (Woolf and Bloechlinger 2002). An important breakthrough in the understanding of axon regeneration was the identification of several of these inhibitory molecules and their receptors. At least three proteins contribute to the inhibitory action of CNS myelin on axonal growth and development: NOGO (Chen, Huber et al. 2000; GrandPre, Nakamura et al. 2000; Prinjha, Moore et al. 2000), OMG (Wang, Koprivica et al. 2002) and MAG (McKerracher, David et al. 1994). Fournier and colleagues (Fournier, GrandPre et al. 2001) identified a binding partner for NOGO that they called the NOGO receptor (NgR). NgR was shown to be expressed on the surface of certain neurons and in combination with NOGO lead to axonal growth cone collapse. More recently it has been shown that all three of the major myelin neuronal inhibitors interact and bind to the same receptor (NgR) with high affinity (Liu, Fournier et al. 2002; Wang, Koprivica et al. 2002). It has been suggested that there is functional redundancy within the myelin inhibitors, as the presence of any one is sufficient to activate an inhibitory signal via the receptor complex (Filbin 2003). Although NgR is essential for NOGO, MAG and OMG to exert their inhibitory effect, it is unable to transduce the inhibition signal across membranes as it contains no transmembrane domains. The inhibitory signal is instead transmitted via a transmembrane co-receptor, the nerve growth factor receptor (NGFR)/p75 neurotrophin receptor (Wang, Kim et al. 2002; Wong, Henley et al. 2002). It has also been shown that RhoA a member of the Rho GTPase family is activated by this inhibitory complex and is also involved in inhibiting neurite outgrowth (Hasegawa, Fujitani et al. 2004). RhoA in its GTP-bound form induces rigidity in the actin cytoskeleton and it is thought this prevents axon elongation and results in growth cone collapse (Hasegawa, Fujitani et al. 2004).

The myelin associated neurite inhibitors have been proposed to have other functions including involvement in myelination, maintaining functional and structural relationships between glial cells and axons and in maintaining and regulating the structural plasticity of the nervous system (He and Koprivica 2004).

Three independent mouse knockout models have been made for NOGO. The first (Kim, Li et al. 2003) showed improved axon regeneration in young Nogo knockout mice after spinal cord injury. However a second study (Zheng, Ho et al. 2003) displayed no evidence of increased axon regeneration. The third (Simonen, Pedersen et al. 2003) reported an improvement in axonal growth but only a small number of axons regenerated. These results demonstrate the redundancy of the inhibitory molecules in the regulation of axon regeneration.

The myelin inhibitory proteins and their receptors are expressed in both the CNS and the PNS during development suggesting they may in part control the growth of axons during this period (Huber, Kolodkin et al. 2003). As neuronal networks develop, the growth of axons is progressively halted. It has been postulated that the lack of axonal regeneration in the adult CNS may result from reactivation of the mechanisms responsible for controlling growth of the complex neuronal system during development (Woolf 2003). During development the course of an individual axon is controlled by many extracellular guidance prompts as well as axonal signalling (He and Koprivica 2004). The extension of axons from their cell body begins with the formation of small growths whose tips form a specialised structure called the axonal growth cone. The growth cone interacts with its surrounding environment to determine the direction and extent of axon elongation. When the growth cone comes into contact with CNS myelin, its cytoskeletal structure is altered causing it to collapse preventing further axonal growth (Woolf and Bloechlinger 2002). The myelin inhibitory proteins are thought to play a role in this process. Recently, it has also been suggested that the myelin inhibitory proteins might function as a repulsive contact-guidance mechanism guiding axons along myelinated tracts as well as being responsible for terminating growth by collapsing the axonal growth cone (Raisman 2004). If the NOGO complex does have a function in the growth and guidance of axons it is possible that any alteration in the expression or functional ability of the genes could allow misregulation of development, which could disrupt connectivity of the CNS. In fact it has been suggested that the main role of the NOGO complex may be more related to the long term preservation of the existing wiring and connectivity present in the adult CNS, and that the suppression of any neuronal regeneration or growth is a necessary consequence of this function (Woolf and Bloechlinger 2002).

4.1.3 The NOGO receptor genes:

4.1.3.1 Neurite outgrowth inhibitor (NOGO):

The neurite outgrowth inhibitor (NOGO), also known as Reticulon 4 (RTN4), gene is located on chromosome 2p13-p14, spans 78kb of genomic sequence and contains nine exons. Several genome scan studies (Coon, Myles-Worsley et al. 1998; DeLisi, Shaw et al. 2002; Straub, MacLean et al. 2002) and a meta-analysis (Lewis, Levinson et al. 2003) have reported evidence that this region is linked to schizophrenia.

NOGO is a member of the reticulon family of proteins, which are usually associated with the endoplasmic reticulum. The NOGO gene encodes for three transcripts (NOGO-A, -B and –C) by alternative splicing and promoter use. The longest transcript is NOGO-A and contains a unique amino-terminal region (He and Koprivica 2004). NOGO-A is highly expressed in oligodendrocytes but not in Schwann cells, a finding consistent with the notion that NOGO-A acts as a CNS myelin associated inhibitor. NOGO-A is also expressed in certain central and peripheral neurons (He and Koprivica 2004). The NOGO-A isoform that is observed in oligodendrocytes mostly associates with the endoplasmic reticulum, however, it has also been observed on the oligodendrocyte surface and on the innermost loop of the myelin membrane (Filbin 2003). NOGO-A has at least two inhibitory domains. One of these is termed NOGO-66 as it is coded for by a 66-amino acid sequence and is common to all isoforms of NOGO. The second domain termed amino-NOGO is unique to NOGO-A and is located in the extended N-terminus of the protein. The most common structure of NOGO-A observed in the myelin membrane places the NOGO-66 domain outside the cell and the N- and C-terminals including the amino-NOGO domain within the cytoplasmic space (Filbin 2003).

The NOGO gene is expressed during neural development. Halabiah and colleagues (Al Halabiah, Delezoide et al. 2005) showed that NOGO and its receptor are expressed during human fetal development, with the strongest expression being observed in the cortical and ventricular zones of the brain. The authors suggested that the early expression might indicate a more diverse function for the gene than just its known neuronal growth inhibition activity and that it may be involved with shaping the neuronal circuitry during development (Gregorio, Mury et al. 2005). Mingorance and colleagues (Mingorance, Fontana et al. 2004) have suggested that the NOGO receptor complex may play an important role during the development of the hippocampal connections in the brain and that they might subsequently be involved in maintaining neuronal plasticity in the adult brain. The hippocampus is a brain region that has been implicated in the aetiology of schizophrenia (Harrison 1999).

Novak and colleagues (Novak, Kim et al. 2002) reported that NOGO mRNA is over expressed in schizophrenic post mortem cerebral cortex. Quantitative RT-

PCR, showed an average 60% over expression of NOGO in the schizophrenic samples. In principle this might result in increased NOGO receptor stimulation and consequently prematurely limit neural development, cause disturbed or abnormal neural connectivity, as well as interfering with any neural regeneration or remodelling in the mature nervous system. All of the above would fit in with the hypothesis that schizophrenia is a disorder of altered connectivity (Friston 1998; Benes 2000).

The NOGO gene has also been implicated in multiple sclerosis (MS). Karnezis and colleagues (Karnezis, Mandemakers et al. 2004) looked at NOGO deficient mice affected with experimental autoimmune encephalomyelitis (EAE), an animal model of MS. They reported that the EAE mice lacking NOGO displayed less signs of inflammatory demyelination and axonal damage than NOGO expressing control mice with EAE, and subsequently developed less severe phenotypes. They concluded that NOGO is likely to play a role in the development of EAE and as a result may be involved in the aetiology of MS. Satoh and colleagues (Satoh, Onoue et al. 2005) reported that NOGO expression was up-regulated in surviving oligodendrocytes at the edge of active demyelinating lesions in MS brains compared to controls, whilst the expression of the NOGO receptor was also increased in astrocytes and microglia/macrophages within these MS lesions compared to their expression levels in the brains of controls. The authors suggested that this increased expression of NOGO and its receptor may play a role in the active demyelinating lesions observed in MS.

Several molecular genetic studies of schizophrenia have targeted NOGO. The results have been varied but there has been some evidence for an association to schizophrenia reported (see table 4.1). Initially Novak and colleagues (Novak, Kim et al. 2002) detected a CAA insertion polymorphism within the 3'UTR of the gene that they reported to be associated with schizophrenia. The frequency of individuals homozygous for the insertion was significantly higher in schizophrenics compared to controls (p=0.002). The authors also detected a second polymorphism in the 3'UTR, a TATC deletion upstream of the CAA insertion which they did not test for association with schizophrenia but was reported to be almost always co-inherited with the associated insertion (Novak, Kim et al. 2002). More recently (Tan, Chong et al. 2005) examined NOGO in a Chinese case control study. They did not observe a significant association for either polymorphism in their overall sample, but when they analysed females alone they detected significant association to schizophrenia for both the CAA insertion (p=0.03) and the TATC deletion (p=0.01). They also examined haplotypes in females and again detected a significant association to schizophrenia with the haplotype containing the CAA insertion and TATC deletion polymorphisms. Interestingly, a second Chinese case control study (Chen, Gu et al. 2004a) reported no evidence of association with the polymorphisms individually, but examination of the supplementary data from that paper reveals a significant association to the same haplotype (p=0.03) in females. It should be noted that there are some haplotype frequency differences between the two Chinese studies raising the concern of an error in allele labelling.

Some studies have failed to provide evidence of association with schizophrenia. Two studies (Covault, Lee et al. 2004; Gregorio, Mury et al. 2005) typed the CAA insertion polymorphism in case control samples. Xiong and colleagues (Xiong, Rouleau et al. 2005) typed the polymorphisms in case control and trio samples, but all failed to find any association with schizophrenia for either polymorphism individually or at the haplotypic level. It should be noted that the sample sizes of the negative replication studies are small and therefore the power to detect an association might be low. It is also of note that the sample studied by Xiong and colleagues (2005) was ascertained from a mixture of ethnic backgrounds. One interesting observation from the replication studies is that the allele frequencies of the two polymorphisms differ quite significantly between the different samples and ethnic groups tested, making it hard to compare the results of the studies together.

Table 4.1 Summary of previous RTN4 association studies in schizophrenia.

This table shows the sample sizes, polymorphisms tested, the corresponding allele frequencies and a summary of the results for all of the previous RTN4 association studies.

Author	Cases/Controls (Ethinicity)	Polymorphism	Allele Freq (Cases)	Allele Freq (Controls)	Results
Novak et al, 2002	81 / 61 (North American Caucasians)	CAA insertion in the 3'UTR	CAA Ins (0.41)	CAA Ins (0.22)	CAA insertion significantly associated with schizophrenia (p=0.002).
Covault et al, 2004	57 / 243 (European Americans)	CAA insertion in the 3'UTR	CAA Ins (0.43)	CAA Ins (0.43)	No significant difference between cases and controls.
Gregorio et al, 2005	181 / 427 (Brazilian)	CAA insertion in the 3'UTR	CAA Ins (0.36)	CAA Ins (0.38)	No significant difference between cases and controls.
Xiong et al, 2005	462 / 153 (96 French-Canadian, 83 Tunisian, 283 Other Europeans)	CAA insertion and TATC deletion in the 3'UTR	CAA Ins (0.43) TATC Del (0.50)	CAA Ins (0.42) TATC Del (0.49)	No significant difference between cases and controls.
Tan et al, 2005	363 /253 (Females 93/129) (Ethnic Chinese)	CAA insertion and TATC deletion in the 3'UTR	CAA Ins (0.35) (Female 0.36) TATC Del (0.39) (Female 0.41)	CAA Ins (0.30) (Female 0.27) TATC Del (0.34) (Female 0.30)	Significant allelic and haplotypic association with CAA insertion (p=0.03) and TATC deletion (p=0.01) in female schizophrenics.
Chen et al, 2004	403 / 347 (Han Chinese)	CAA insertion and TATC deletion in the 3'UTR	CAA Ins (0.69) TATC Del (0.65)	CAA Ins (0.66) TATC Del (0.61)	Significant haplotypic association in female schizophrenics (p=0.03).

4.1.3.2 NOGO receptor (NgR):

The NOGO receptor (NgR)/Reticulon 4 receptor (RTN4R) gene is located on chromosome 22q11, spans 26kb of genomic sequence and contains two exons. The gene maps on the edge of the VCFS critical region. This region is deleted in most cases of VCFS, a syndrome associated with a marked increased risk of developing schizophrenia (Murphy, Jones et al. 1999; Murphy 2002). A wider region of chromosome 22q has also been implicated in several linkage studies (Gill, Vallada et al. 1996; DeLisi, Shaw et al. 2002; Williams, Norton et al. 2003), and a recent meta-analysis of the linkage data has provided further support for the involvement of this region (Lewis, Levinson et al. 2003).

Sinibaldi and colleagues (Sinibaldi, De Luca et al. 2004) recently looked at the NgR gene as a potential candidate gene for schizophrenia. They used DHPLC to screen 120 schizophrenic patients and detected three rare coding mutations, two non-synonymous (R119W and R196H) and one synonymous (L18L). The two non-synonymous SNPs were each detected in a single schizophrenic patient but none of 300 controls when they were subsequently genotyped for these polymorphisms. However since the design was based only on screening cases, we cannot conclude that the rare non-synonymous SNPs are associated with schizophrenia, because by design any unique (or very rare variants) will only to be detected in cases. Thus at present there is no reliable evidence implicating NgR in schizophrenia.

4.1.3.3 Oligodendrocyte myelin glycoprotein (OMG):

Oligodendrocyte myelin glycoprotein (OMG) is located on chromosome 17q11.2 approximately 10 Mb centromeric of the CNP locus, where it is embedded within intron 1 of the neurofibromatosis type 1 (NF1) gene. It spans 2.5kb of genomic sequence and contains two exons, with the first exon being untranslated. The OMG and NF1 genes are transcribed in different directions. The 17q11 region has previously been implicated with schizophrenia in a single pedigree (Williams, Norton et al. 2003). The involvement of this region in schizophrenia has been supported by a recent meta-analysis (Lewis, Levinson et al. 2003), and an additional linkage study (Klei, Bacanu et al. 2005). Additionally there is also significant (Dick, Foroud et al. 2003) and suggestive (Bennett, Segurado et al. 2002; Ewald, Wikman et al. 2005) evidence of linkage to this region in Bipolar disorder.

OMG is a glycosylphosphatidylinositol (GPI)-linked protein expressed in both oligodendrocytes and neurons. OMG is a relatively minor component of CNS and PNS myelin accounting for less than 1% of the structure (Filbin 2003). The timing of OMG expression in oligodendrocytes correlates with the onset of myelination; the protein can be detected on the surface of oligodendrocytes and on myelin layers adjacent to axons (He and Koprivica 2004). Vourc'h and colleagues (Vourc'h, Dessay et al. 2003) have shown an increase in OMG expression during the post-natal period of CNS development, with the peak expression level occurring during the later stages of myelination. This would indicate that OMG may be involved in CNS myelogenesis, possibly playing a role in processes such as the arrest of oligodendrocyte proliferation, the arrest of myelination or the compaction of myelin (Vourc'h, Dessay et al. 2003). Finally, in a recent study Vourc'h and colleagues (Vourc'h, Martin et al. 2003) reported an association between a G21D G>A non-synonymous SNP in the OMG gene and autism. Although there is no strong case for a link between the two diseases, both autism and schizophrenia are neurodevelopmental disorders, making it feasible that genes involved in one disease may also be involved with the other.

4.1.3.4 Nerve growth factor receptor (NGFR/p75):

The Nerve growth factor receptor (NGFR)/p75 Neurotrophin receptor is located on chromosome 17q21-q22, spans 19kb of genomic sequence and contains six exons. As I mentioned earlier in chapter 3 this region of chromosome 17 has been putatively linked to schizophrenia in a single pedigree (see section 3.2.1, figure 3.5).

NGFR/p75 is highly expressed in both glial cells and neurons during development. Expression then declines to a low level in adulthood (He and Koprivica 2004). As well as binding to nerve growth factor, NGFR has also been shown to bind to all known neurotrophins. Neurotrophins are key mediators of peripheral nervous system myelination and play an important role in the survival and development of neurons and the establishment and maintenance of synapses during development and adulthood (Rajakumar, Leung et al. 2004). Cosgaya and colleagues (Cosgaya, Chan et al. 2002) have shown that myelin formation is dependant on the presence of functional p75 receptors. In a further study Chan and colleagues (Chan, Cosgaya et al. 2001) demonstrated that the brain derived neurotrophic factor (BDNF) acts through the p75 receptor to enhance the formation of myelin, and that removal of endogenous BDNF inhibited myelination. Several independent studies have previously reported associations between polymorphisms in the BDNF gene and psychiatric disorders such as schizophrenia (Neves-Pereira, Cheung et al. 2005; Chen, Chen et al. 2006), bipolar disorder (Sklar, Gabriel et al. 2002; Lohoff, Sander et al. 2005; Green, Raybould et al. 2006) and major depression (Schumacher, Jamra et al. 2005). A more recent study (Rajakumar, Leung et al. 2004), demonstrated that rats with compromised p75 receptor function during cortex development were associated with dopamine hyper-responsivity and impaired prepulse inhibition, both features associated with schizophrenia (Marenco and Weinberger 2000; Parwani, Duncan et al. 2000).

4.2 Materials and Methods:

Briefly described here are the specific methods used to study the NOGO receptor complex genes. More detailed explanations of the methods and samples used are described in Chapter 2.

4.2.1 Polymerase Chain reaction (PCR):

PCR fragments to span exons, UTRs, putative promoter and limited amounts of 5' and 3' flanking regions were designed using the web based program Primer 3.0 (http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi). The primers used to generate amplimers for the purpose of DHPLC for each gene studied are listed below in tables 4.2, 4.3, 4.4 and 4.5 respectively.

It was necessary to attempt a long range PCR due to regions of chromosome 22 having high homology with the NgR exon 1 and flanking intronic sequence. It was hoped that by using primers that anneal specifically to perform the long range PCR, the product could then be used to perform a subsequent nested PCR reaction amplifying only the region of interest in NgR. The long range PCR was performed using the Expand long template PCR system (Roche Diagnostic, Germany). The reaction mix was performed as per manufacturer's instructions and is shown below:

Expand Long Template PCR System, Reaction mix:

Reagent	Volume	Final Concentration / Amount
Genomic DNA template (40ng/μl)	6.0μl	240ng
Forward Primer (10pmol/μl)	3.0μl	300nM
Reverse Primer (10pmol/μl)	3.0μl	300nM
10 mM dNTP mix	2.5μl	500uM
Buffer (22.5 mM $MgCl_2$.)	5.0μl	1x Buffer: 2.25 mM $MgCl_2$
Expand long template enzyme (5u/μl)	0.75μl	3.75 units
Distilled deionised water	29.75μl	
Total reaction volume	50.0μl	

Expand Long Template PCR System, Cycling Conditions:

Temperature	Time	Cycles
94°C	2 mins	1
94°C 55-65°C 68°C	10 secs 30 secs 15 min	10
94°C 55-65°C 68°C	15 secs 30 secs 15 mins + 20 secs per cycle	25
68°C	15 mins	1
4°C	hold	

A variety of different conditions, including a selection of amplification primers and reaction mixes and protocols were attempted but unfortunately amplification of the specific genomic region of interest was unable to be achieved.

Table 4.2. PCR primers and conditions used for RTN4/NOGO assays.

This table shows details of the RTN4 PCR assays for the purpose of mutation detection, including PCR primers, size of fragment, the reaction conditions and annealing temperature. (HS indicates that hot start taq polymerase was used).

Fragment Name	Spanning	Forward Primer (5' to 3')	Reverse Primer (5' to 3')	Size (bp)	Annealing Temp (°C)
RTN 4_1	Putative Prom	CCCATGGCAATCAAACTTTC	TCACTGCATGCCTACTCTGG	412bp	HS 60
RTN 4_2	Putative Prom	TTGAGCAACTTGAAAGCAGGT	CGAAATTGGGAGGAAAACAG	486bp	HS 60
RTN 4_3	Putative Prom	AAGGCCCCAAACTCTCTTTG	CCACTTTGGGTAACCGTGAG	471bp	HS 64
RTN 4_4	Exon 1	ACTCTCAGTTCCCCCAGGAT	ACTGAGCCGAGGGACCTACT	522bp	HS 64
RTN 4_5	Exon 1	TGGCCTCACTCCTAGCTCAT	CGAAGTCATTTCCGAAGTCC	498bp	HS 64
RTN 4_6	Exon 1	AGGACGAAGACCTGGAGGAG	ACCAGCCCAAAGCATCTG	505bp	HS 58
RTN 4_7	Exon 1/Intron 1	CCCCTCCCTCTTTTGTGTG	CACCTTTTCCAAAGGAGCAA	412bp	HS 60
RTN 4_8	Exon 2	TAGGCACCCACCCGTAGTAG	CTGGGCCAAGTCCACACTAA	548bp	HS 64
RTN 4_9	Exon 3	CTGAGTGAAAGCTTGGGATG	GACAAAACATTCTCGGAACCA	510bp	HS 64
RTN 4_10	Exon 3/Intron 3	GAGCCTGTGATACGCTCCTC	GGCAATTACATGCTGTGCAA	493bp	HS 64
RTN 4_11	Intron 3/ Exon 4	ATTTACCCAGTGCCCCTTCT	GAAAGAAGCGGCTGAGAGAG	485bp	HS 64
RTN 4_12	Exon 4	TGTCCTGCTTGAAACTGCTG	CCCATACTCGCTCAAATGGT	506bp	HS 64
RTN 4_13	Exon 4	GAGTTGCAGTGGAAGCTCCT	TCTCAGAATCCTGTGCTGCT	462bp	HS 64
RTN 4_14	Exon 4	TAACCCAGCAGCAACTGAGA	GGTGATGAGCTGGGCTGTAT	513bp	HS 64
RTN 4_15	Exon 4	TGCAGTTCCTAGTGCTGGTG	AGGTGGCAAAGCACTGAGTT	514bp	HS 64
RTN 4_16	Exon 4	CCTGACGTTCCACAAAAACA	AAAGGTCATGGGGCAATTCT	499bp	HS 60
RTN 4_17	Exon 4/Intron 4	CTAATGCCCCGGATGGAG	GGTGAATGCTGAGTGCAAGA	498bp	HS 64
RTN 4_18	Exon 5	TGCAGCTCCATTTGGTATGA	AAACCGAATGCCAGTGAAAC	544bp	HS 60
RTN 4_19	Exon 6	GATGGAGCAAATCCAAATTGA	CTCCTCCATCATGAGCCTTT	468bp	HS 60
RTN 4_20	Exon 7	GTGGCAGAGCATCAGATTCA	AACACAAGTTTACAGCCATGTGA	524bp	HS 64
RTN 4_21	Exon 8	GGGTAAGATGCACCCTTGGT	AAGCTCCTGGCACACAATCT	407bp	HS 64

RTN 4_22	Exon 9/ Exon 10	GTGGCTCCTCAACAGCAGAT	TTCTCACAATCCAGCACACC	526bp	HS 64
RTN 4_23	Exon 10/ Exon 11	ACCTCTCTTTCAGGCACAGA	TTCAATCCAGGGATTTTTGC	463bp	HS 60
RTN 4_24	Exon 11	AGCAGCAGCAAACAAAACTG	CGGCAAGACTATCTGCAACA	402bp	HS 64
RTN 4_25	Exon 11	CCTGTCTTGACTGCCATGTG	GGTGGCTGGAAGGTAAAGGT	434bp	HS 60
RTN 4_26	Exon 11	GCAGGAATGACAAAGCTTGC	TTCACAATGCATTCCACAGA	473bp	HS 60
RTN 4_27	Exon 11	TCCAAAAATGTTGTTTGTTTGC	AGCAATAGTGTCCGTGTGTCA	527bp	HS 60
RTN 4_28	Exon 11	TGCACCAATGTTTTAACTCACA	CTGGCTACTTGCATTTGTCA	488bp	HS 60

Table 4.3: PCR primers and conditions used for RTN4-R/NOGO-R assays.

This table shows details of the RTN4-R PCR assays for the purpose of mutation detection, including PCR primers, size of fragment, the reaction conditions and annealing temperature. (HS indicates that hot start taq polymerase was used).

Fragment Name	Spanning	Forward Primer (5' to 3')	Reverse Primer (5' to 3')	Size (bp)	Annealing Temp (°C)
NOGO-R Long PCR	Dup Exon Var A	CAATGCATACGTCCCATTGAC	CATGAACTAACGTGTCTCTCAGGAT	11,436bp	Long Range PCR 60-68 (Gel Extract)
NOGO-R Prom 1a A	Putative Prom	GGTCCTCACTTCCTCTGCAC	GCAGGTGTCCCTTCTCAATC	426bp	HS 60
NOGO-R Prom 1a B	Putative Prom	CTGACTCCCAGCACTTCTCC	CTCACGAAACCCATGAATCC	487bp	HS 60
NOGO-R Ex 1a A	Exon 1	CAGCAAGGGGCTTCTGTG	CGGGGACTGAAAGTCGTTT	535bp	Couldn't Opt
NOGO-R Ex 1a B	Exon 1	AATGCGCCGCCTTTGTCT	GAAGCGGCAACTTTTCCTG	577bp	Couldn't Opt
NOGO-R Prom 1b A	Putative Prom	AGAAAGAGAGCAGAGAGGGTGT	CAGCTGGCTAAGAGCAGGAT	459bp	HS 66
NOGO-R Prom 1b B	Putative Prom	CTGTGGTCCACTCCTTCCTC	ATAGCCTCCCAGTGGGTTCT	430bp	HS 66
NOGO-R Ex 1b	Exon 1	CATGTAGGGCCTTGGTGAAG	CAGACTCCCCATACCACACC	599bp	HS 66
NOGO-R Ex 2A	Exon 2	AGGTCCTCCTGAGACCCAAT	GAGCTGTGCATTATCGCTGA	439bp	HS 66
NOGO-R Ex 2B	Exon 2	GCAACCTCACCATCCTGTG	GGCAGCGCTGATAGATTGTT	484bp	HS 66
NOGO-R Ex 2C	Exon 2	AGCCTCGACCGTCTCCTACT	GCTCCAGTACTGAGGCCTTG	439bp	HS 66
NOGO-R Ex 2D	Exon 2	CCTTACCATCCCATCTGGAC	TGTACACACACCTGGCTGCT	526bp	HS 61
NOGO-R Ex 2E	Exon 2	AGAAGGCTCAGGTGCCCTA	GCACCTGACATTGGAGCTG	545bp	HS 66

Table 4.4: PCR primers and conditions used for OMG assays.

This table shows details of the OMG PCR assays for the purpose of mutation detection, including PCR primers, size of fragment, the reaction conditions and annealing temperature. (HS indicates that hot start taq polymerase was used).

Fragment Name	Spanning	Forward Primer (5' to 3')	Reverse Primer (5' to 3')	Size (bp)	Annealing Temperature (°C)
OMG_1	Putative Prom	ACAGCCTAAACCCAAATTCA	ACTCTAGGAGCCACCTTTTGTAT	467bp	HS 60
OMG_2	Putative Prom	GTGGCAATAAAGCACCCAAT	TTGCACATTTCATCTTTCTGG	512bp	HS 60
OMG_3	Putative Prom	CCACTGAAGAGTGGAGCAAA	TCCCCCAACATGGTGATTAT	531bp	HS 60
OMG_4	Exon 1	CCTAAAAGGGAAGACGGAACA	AAGCTCTGCAACCACCTGAT	534bp	HS 60
OMG_5	Exon 1	TGTGCACCTCTTGCCATAAA	GCACTTGAAAAATCCAGACCA	501bp	HS 60
OMG_6	Exon 1	TGCTGCAGTGTCAGTTGAGG	CCTTCAATTGGGCTATGTGG	521bp	HS 60
OMG_7	Exon 1	GGGGGAGGGAGAGAGAAAAT	CCAGGGTCCTCAGATTGGTA	517bp	HS 60
OMG_8	Exon 1	TATGCACAGAGAGGCACAGG	GGTCAAAAGATTGGTCTGGAA	546bp	HS 60
OMG_9	Exon 1	TGCCCTCCAAACTACATATCG	GCACCAAACGTTGTTTCCTT	454bp	HS 60
OMG_10	Exon 1	TGCAGACAGTGGACACCATT	TCAGGGAGGAGTGCTTTCAT	518bp	HS 60
OMG_11	Exon 1	GCTCAATGTTGTGGTCATGC	TGGAAAGCCAAATGACACAG	413bp	HS 60
OMG_12	Exon 1	AAAAATGAAATGTTTTAAGGCTTC	GCTTTTTGAAGAGCTTTTGTCC	409bp	HS 60
OMG_13	Exon 1	CATCCCATGAATGGTTATGCT	CTGTCCCATGCCTAGCCTTA	535bp	HS 60

Table 4.5: PCR primers and conditions used for NGFR assays.

This table shows details of the NGFR PCR assays for the purpose of mutation detection, including PCR primers, size of fragment, the reaction conditions and annealing temperature. (HS indicates that hot start taq polymerase was used).

Fragment Name	Spanning	Forward Primer (5' to 3')	Reverse Primer (5' to 3')	Size (bp)	Annealing Temperature (°C)
NGFR_1	Putative Prom	CGCTCTTGATCTCCCTTTTG	TGCACACACACACTGACACC	506bp	HS 60
NGFR_2	Putative Prom	AAGCCCATGGAGACACTCAC	GGAGTCCATTCCGACATCAT	406bp	HS 64
NGFR_3	Putative Prom	GAGTGCATGAGAACGGTGTG	TTAGAGCCTCTCACCCATCC	529bp	HS 60
NGFR_4 Redesign	Exon 1	GAACAGGAACCGCAGTGG	GCTAACACTCACCCCCAGAA	494bp	HS 62
NGFR_5	Exon 1	CGGACCGAGCTGGAAGTC	TTCTCTCGCCTTGGTTCTGT	518bp	HS 64
NGFR_6	Exon 2	CCCCTGTCAGCTTTGATGTT	CTCCTCCCTCTGCACACTTC	523bp	HS 64
NGFR_7	Exon 3	CCAACTTTCACAGGCTAGGG	CACACGGTGTTCTGCTTGTC	402bp	HS 64
NGFR_8 Redesign	Exon 3	GCAATAGGGGAGGGAGAGAC	CAGAGAGGGGTTCTTTGCTG	664bp	HS 62
NGFR_9	Exon 4	CAGAGTCACCCAGCAAGTCA	AGACAGGGATGAGGTTGTCG	428bp	HS 64
NGFR_10	Exon 4	GCCTTTTCTCCCAGAGATCC	GGGATGAGGACACAAGCAGT	561bp	HS 64
NGFR_11	Exon 5	CGGCAGAGCACAGATACTCA	AAAGCCTCCATTTCCCAAGT	441bp	HS 64
NGFR_12	Exon 5/Intron 5	AGGCTGAGGAAACAGAAGCA	TGAACTCCTTTCCCCAACAG	582bp	HS 60
NGFR_13	Exon 6	GCTGGAGTGACAGGAGGAAG	CTGCACAGACTCTCCACGAG	475bp	HS 64
NGFR_14	Exon 6	CCCGAGCACATAGACTCCTT	GTTCCTCGATTCCTCCCTCT	588bp	HS 64
NGFR_15	Exon 6	CCTTCTCCCACACTGCTAGG	CTTCAAGCCATTTCCCAGAA	466bp	HS 60
NGFR_16	Exon 6	GCACCGCCTTCTCTAAATGA	ATTCCCTCCGATGCTTCTCT	496bp	HS 60
NGFR_17	Exon 6	CCTCACCTTGCAACACACAG	TAGAGGGGAGCCACACTCAC	514bp	HS 64
NGFR_18	Exon 6	GACTGTGACCTGTGGGGATT	AAGGAGGGGAGGTGATAGGA	542bp	HS 60
NGFR_19	Exon 6	CCCTTGAGCTGAGATGGAAC	CTCTGACTTGCCTGTGGTCA	483bp	HS 60
NGFR_20	Exon 6	AGGAAGCTAGCCAGCTGTGA	AGCCTCCCTTTGTCTTCCTC	451bp	HS 60
NGFR_21	Exon 6	GGTTGTTTGGCCAAGTCCT	TTTCACAGGACAGCATCTGG	506bp	HS 60

4.2.2 Genotyping:

Every SNP identified was typed in pools using the ABI SNaPshot™ kit. PCR primers used are listed in tables 4.2 to 4.5 and extension primers are listed in tables 4.6, 4.20, 4.22 and 4.24.

Any SNP showing an association in the pools at the p=0.05 level was typed individually. Where a polymorphism could not be assayed with a primer extension assay, like the VNTR in fragment RTN4 24, it was typed individually in the UK case control sample. This was performed by assaying the polymorphism by size using a fluorescently labelled PCR primer. The forward PCR primer was labelled with the fluorophore FAM. The PCR products were analysed on the ABI 3100 Genetic Analyser. An example of an electrophorogram is shown in the appendix at the end of this chapter (section 4.5, appendix 4.4).

4.3 Results:

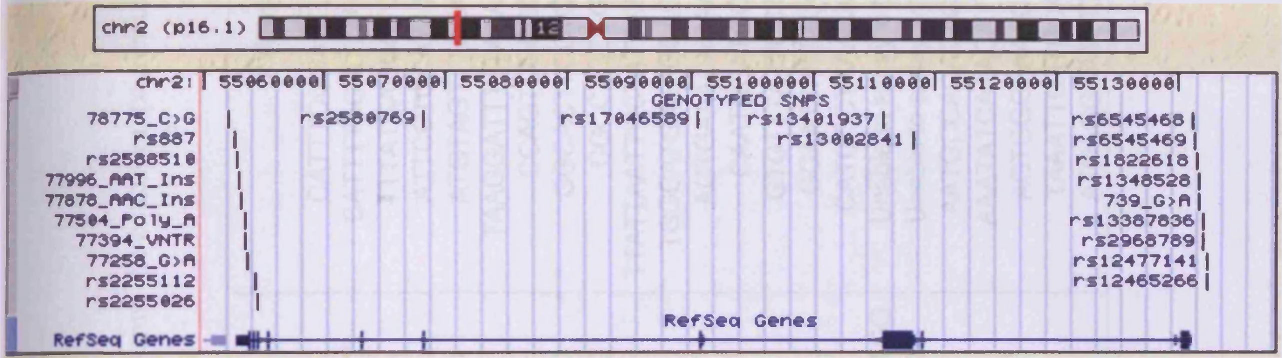
4.3.1 NOGO / RTN4:

Determination of NOGO genomic structure was based on alignment of the cDNA sequence NM_020532 and the genomic sequence available from the UCSC Genome Bioinformatics website (http://genome.ucsc.edu).

I identified nineteen SNPs in NOGO along with two insertion/deletions a VNTR and a variable length poly A sequence (See appendix 4.3 for sequencing results). There were no coding SNPs, synonymous or non-synonymous. The variants were located in 5'-UTR (rs6545468), 3'-UTR (77258 G>A, 77394 VNTR, rs11464312, 77878 AAC-Ins and 77996 AAT-Ins), 5' flanking region (rs12465266, rs12477141, rs2968789, rs13387836, -739 A>G, rs1348528, rs1822618 and rs6545469), 3' flanking region (rs2588510, rs887 and 78775 C>G) and introns (rs13002841, rs13401937, rs17046589, rs2580769, rs2255026 and rs2255112) (see figure 4.1).

Figure 4.1. Schematic diagram of NOGO.

Diagram of NOGO (adapted from the UCSC Genome Browser), indicating the positions of the identified polymorphisms. For numbering of polymorphisms, +1 indicates the 'A' of the ATG translation initiation codon. The gene is displayed (3'-5').



The 77878 AAC-Ins and the 77394 VNTR are the same polymorphisms reported to be associated with schizophrenia in previous studies (Novak, Kim et al. 2002; Chen, Gu et al. 2004a; Tan, Chong et al. 2005). Seventeen of the detected polymorphisms were successfully genotyped in the pooled association sample. The extension primers used for pooled genotyping are listed in table 4.6. The minor allele frequency of SNPs -739 A>G, 77258 G>A, 78775 C>G and the 77996 AAT-Ins were undetectable using pooled genotyping. The 77394 VNTR and 77504 poly A polymorphisms were incompatible with primer extension genotyping assays. The calculated allele counts and frequencies for all of the genotyped polymorphisms are shown in table 4.7. There were no significant differences observed between the allele frequencies of the cases and controls for any of the assayed polymorphisms.

Table 4.6: Polymorphism and primer extension reaction information.

This table shows the sequence flanking each of the NOGO/RTN4 polymorphisms, the extension primer used to assay the polymorphism and the expected alleles in the primer extension reaction.

SNP number.	rs Number	SNP Flanking sequence	Extension primer	Expected Alleles
-1447 G>A	rs12465266	CTAATTCTTGAA[G/A]TCCAGTGCGCAT	CATTTCGGAGGTACTAATTCTTGAA	G (Blue) & A (Green)
-1409 C>T	rs12477141	CACATCTAATTT[C/T]AGCCTAGCTACA	CATTTTACACTTACAGCACATCTAATTT	C (Black) & T (Red)
-1196 T>C	rs2968789	TCTGCAACTACT[T/C]TCTTTAAACTCT	TTTATGAGGCTGGAGAGTTTAAAGA	A (Green) & G (Blue)
-1123 A>G	rs13387836	AAACCAATGACT[A/G]TGGGGCTTCCTG	ATTCCTGCATATCAAACCAATGACT	A (Green) & G (Blue)
-739 A>G	N/A	CTTTGAGTGAAC[A/G]AATAATYTTTTA	ATGTAGTTTGGAACTTTGAGTGAAC	A (Green) & G (Blue)
-732 C>T	rs1348528	TGAACRAATAAT[C/T]TTTTATTTCAAA	TAAGGATTTTTAAAGATTTTTGAAATAAAA	G (Blue) & A (Green)
-646 C>G	rs1822618	TTCTTCCTTTCC[C/G]GGGAGGAAAGTT	CCAGTCGGGTTTCTTCCTTTCC	C (Black) & G (Green)
-472 C>T	rs6545469	TGAACGCAATCC[C/T]AGCCGGTAAGAG	GGCACTTCCTCTCTCTTACCGGCT	G (Blue) & A (Green)
-205 G>C	rs6545468	CGGCGGCGGCAA[G/C]TGGGGACAGGGC	CGCCACCCGCCCTGTCCCCA	C (Black) & G (Blue)
22715 A>G	rs13002841	ATTTTAGTAAAC[A/G]GTAAGTCTTTAC	TTATTAATTGCTATTAACATTTCATTTTAGTAAAC	A (Green) & G (Blue)
25222 C>A	rs13401937	TTCAGGTAATCC[C/A]ATCTATATGCAG	TGGCATGAAAGCACAAAACTGCATATAGAT	G (Blue) & T (Red)
40387 A>C	rs17046589	GTTGTATACCAG[A/C]AAATGTCATGCT	AGTGAAACAGAACAGCATGACATTT	T (Red) & G (Blue)
62818 T>G	rs2580769	TCAGGTGAGATG[T/G]CTGGAAAACAAG	CAATGCATGCCTTGTTTTCCAG	A (Green) & C (Black)
76398 G>A	rs2255026	CCAAATCATTCC[G/A]TTCTTTTTGAAA	GTGTTTATTCCTCCCAAATCATTCC	G (Blue) & A (Green)
76636 A>C	rs2255112	GTGGTTAATAAG[A/C]TTCTTTAGCAAC	GGAATTATTACCGTTGCTAAAGAA	T (Red) & G (Blue)
77258 G>A	N/A	GTTTCACAGATC[G/A]TTGTTAGATCTT	CATGGCTAAAAATAAAGATCTAACAA	C (Black) & T (Red)
77394 VNTR	N/A	TTTCCTATCT[ATCT/AATCTATCT]GAGGCACTGG	Unable to assay with an extension primer.	
77504 Poly A	rs11464312	TAGAAAAAAAAAAAAAAA[A/-]GCCCTTTTCA	Unable to assay with an extension primer.	
77878 AAC-Ins	N/A	TAGAACTCCAAC[AAC/-]ATCAATTTCATT	AATGCCACACACATAGAACTCCAACA	T (Red) & A (Green)
77996 AAT-Ins	N/A	TATGCAAGAAAT[AAT/-]TATTAATTACAA	AAATATCAAACATTGTTATGCAAGAAAT	T (Red) & A (Green)
78188 G>A	rs2588510	GCTTCAACCAGC[G/A]AATGGTAGCTTG	AGTGGGTAAACTGCTTCAACCAGC	G (Blue) & A (Green)
78368 T>C	rs887	CAATGTTTTAAC[T/C]CACATATATCAT	TAAATTCATGCACCAATGTTTTAAC	T (Red) & C (Black)
78775 C>G	N/A	GTAAATGTTAGT[C/G]TATTCTACAATT	ATGAGGCATTTTCAATTGTAGAATA	G (Blue) & C (Black)

Table 4.7: Pooled genotyping data for NOGO/RTN4.

This table shows the estimated minor allele frequencies and allele counts in pooled samples of 535 cases and 552 controls for all of the polymorphisms in the NOGO/RTN4 gene, along with their corresponding p-values.

SNP Data		Pooled Genotyping data				
SNP number.	rs Number	Minor Allele Cases (freq)	Minor Allele Controls (freq)	Difference	Chi-square	P-value
-1447 G>A	rs12465266	G= 136 (0.13)	G= 151 (0.14)	0.01	0.44	0.51
-1409 C>T	rs12477141	T= 116 (0.11)	T= 133 (0.12)	0.01	0.78	0.38
-1196 T>C	rs2968789	C= 98 (0.09)	C= 117 (0.11)	0.02	1.26	0.26
-1123 A>G	rs13387836	G= 352 (0.33)	G= 358 (0.32)	0.01	0.05	0.82
-739 A>G		Minor allele frequency too rare to assay.				
-732 C>T	rs1348528	T= 128 (0.12)	T= 136 (0.12)	0	0.07	0.79
-646 C>G	rs1822618	G= 121 (0.11)	G= 144 (0.13)	0.02	1.53	0.22
-472 C>T	rs6545469	T= 94 (0.09)	T= 75 (0.07)	0.02	3.00	0.08
-205 G>C	rs6545468	C= 460 (0.43)	C= 479 (0.43)	0	0.03	0.85
22715 A>G	rs13002841	G= 102 (0.1)	G= 123 (0.11)	0.01	1.52	0.22
25222 C>A	rs13401937	C= 131 (0.12)	C= 111 (0.1)	0.02	2.63	0.11
40387 A>C	rs17046589	C= 160 (0.15)	C= 196 (0.18)	0.03	3.11	0.07
62818 T>G	rs2580769	T= 233 (0.22)	T= 245 (0.22)	0	0.05	0.82
76398 G>A	rs2255026	G= 166 (0.16)	G= 192 (0.17)	0.01	1.39	0.24
76636 A>C	rs2255112	C= 456 (0.43)	C= 490 (0.44)	0.01	0.69	0.41
77258 G>A		Minor allele frequency too rare to assay.				
77394 VNTR		Incompatible with a primer extension assay.				
77504 Poly A	rs11464312	Incompatible with a primer extension assay.				
77878 AAC-Ins		AAC Ins= 475 (0.44)	AAC Ins= 491 (0.45)	0.01	0.001	0.97
77996 AAT-Ins		Too rare to assay.				
78188 G>A	rs2588510	A= 445 (0.42)	A= 460 (0.42)	0	0.001	0.97
78368 T>C	rs887	C= 344 (0.32)	C= 377 (0.34)	0	0.98	0.32
78775 C>G		Minor allele frequency too rare to assay.				

As the 77394 VNTR was incompatible with primer extension genotyping in the pools, and was of high a priori interest because it is reported to be in strong LD with the 77878 AAC-Ins polymorphism previously associated with schizophrenia, I genotyped the polymorphism individually in the case-control sample. This was achieved by using a Genescan fluorescent PCR assay (as described in section 2.10.5). The results of the individual genotyping are shown in tables 4.8 and 4.9 below. The VNTR was found to be significantly associated with schizophrenia at both the genotypic (p=0.02) and allelic level (p=0.02). Homozygote genotypes containing the smaller repeat length (allele 1) are over represented in cases compared to the controls, and this is reflected in the allelic association.

Table 4.8: Genotype counts for the NOGO 77394 VNTR.

Genotype counts for the 77394 VNTR polymorphism in the case-control sample, and corresponding p-value after analysis using CLUMP. Allele 2 is the ATCT Insertion and allele 3 is the AATCTATCT Insertion.

77394 VNTR	**Cases (n = 658)**		**Controls (n = 712)**	
Genotype	**Count**	**Frequency**	**Count**	**Frequency**
1-1	165	0.25	145	0.20
1-2	303	0.46	355	0.50
1-3	32	0.05	20	0.03
2-2	135	0.21	171	0.24
2-3	22	0.03	17	0.02
3-3	1	0.001	4	0.005
			χ^2 = 12.73	**P= 0.02**

Table 4.9: Allele counts for the NOGO 77394 VNTR.

Allele counts for the 77394 VNTR polymorphism in the association sample and corresponding chi-square and p-values. Allele 2 is the ATCT Insertion and allele 3 is the AATCTATCT Insertion.

SNP	**Alleles**		
77394 VNTR	**1**	**2**	**3**
Cases	665 (0.51)	595 (0.45)	56 (0.04)
Controls	665 (0.47)	714 (0.50)	45 (0.03)
		χ^2 = 7.77	**p = 0.02**

Given previous reports of the 77878 AAC-Ins polymorphisms association with schizophrenia, I also genotyped this polymorphism individually even though it did not show an association with schizophrenia when genotyped in the pooled sample, see table 4.7. The results are shown in table 4.10. No evidence for association with schizophrenia was found at either the genotypic or allelic level. The allele frequencies are similar to the results of the pooled genotyping, except for a small difference (0.03) in the frequencies of the controls. This is possibly the result of the enlarged control sample that was individually genotyped or may be attributable to experimental variation in the pooled genotyping experiment.

Table 4.10: Individual genotyping results for 77878 AAC in/del polymorphism.

Genotype and allele counts and frequencies for the 77878 AAC in/del polymorphism in the schizophrenia association sample. Both genotypic and allelic chi-square and p-values are shown.

SNP	Genotypes			Alleles	
77878 AAC Ins	**Del/Del**	**Del/Ins**	**Ins/Ins**	**Del**	**Ins**
Cases (n=666)	200 (0.30)	343 (0.52)	123 (0.18)	743 (0.56)	589 (0.44)
Controls (n=714)	228 (0.32)	374 (0.52)	112 (0.16)	830 (0.58)	598 (0.42)
	$\chi^2 = 2.02$	P = 0.36		$\chi^2 = 1.54$	P = 0.21

4.3.1.1 Gender analysis.

Following the results of Tan and colleagues (Tan, Chong et al. 2005), which reported an association between the two NOGO 3'UTR polymorphisms and schizophrenia in females, I analysed my data according to gender.

The results of the 77394 VNTR analyses are shown in tables 4.11, 4.12, 4.13 and 4.14 below. The VNTR was associated with schizophrenia at the genotypic level (p=0.02) in females (table 4.11). At the allelic level the association with schizophrenia was also stronger in females (p=0.004), with the smaller allele again being over represented in the cases compared to the controls (table 4.12). In males the polymorphism displayed a trend towards significance at the genotypic (p=0.08) level (table 4.13) but did not show any evidence for association with schizophrenia at the allelic (p=0.422) level (table 4.14).

Table 4.11: Genotyping of the NOGO 77394 VNTR in females.

Genotype counts for the 77394 VNTR polymorphism in the female association sample and corresponding p-value after analysis using CLUMP. Allele 2 is the ATCT Insertion and allele 3 is the AATCTATCT Insertion.

77394 VNTR	Cases (n = 211)		Controls (n = 232)	
Genotype	**Count**	**Frequency**	**Count**	**Frequency**
1-1	48	0.23	41	0.18
1-2	107	0.51	120	0.52
1-3	9	0.04	5	0.02
2-2	37	0.18	63	0.27
2-3	9	0.04	3	0.01
3-3	1	0.004	0	0
			χ^2 = 12.23	**P= 0.022**

Table 4.12: Allele counts for the NOGO 77394 VNTR in females.

Allele counts for the 77394 VNTR polymorphism in the female association sample and the corresponding chi-square and p-values. Allele 2 is the ATCT Ins and allele 3 is the AATCTATCT Ins.

SNP	Alleles		
77394 VNTR	**1**	**2**	**3**
Cases (n=211)	212 (0.50)	190 (0.45)	20 (0.05)
Controls (n=232)	207 (0.45)	249 (0.54)	8 (0.01)
		χ^2 = 11.17	**p = 0.004**

Table 4.13: Genotyping of the NOGO 77394 VNTR in males.

Genotype counts for the 77394 VNTR polymorphism in the male association sample and the corresponding p-value after analysis using CLUMP. Allele 2 is the ATCT Insertion and allele 3 is the AATCTATCT Insertion.

77394 VNTR	**Cases (n = 431)**		**Controls (n = 479)**	
Genotype	**Count**	**Frequency**	**Count**	**Frequency**
1-1	113	0.26	104	0.22
1-2	187	0.44	234	0.49
1-3	23	0.05	15	0.03
2-2	96	0.22	108	0.23
2-3	12	0.03	14	0.03
3-3	0	0.0	4	0.008
			χ^2 = 9.66	**P= 0.08**

Table 4.14: Allele counts for the NOGO 77394 VNTR in males.

Allele counts for the 77394 VNTR polymorphism in the male association sample and the corresponding chi-square and p-values. Allele 2 is the ATCT Ins and allele 3 is the AATCTATCT Ins.

SNP	**Alleles**		
77394 VNTR	**1**	**2**	**3**
Cases (n=431)	436 (0.51)	391 (0.45)	35 (0.04)
Controls (n=479)	457 (0.48)	464 (0.48)	37 (0.04)
		χ^2 = 1.72	**p = 0.42**

Given the gender effect observed with the 77394 VNTR I also performed an analysis comparing the female cases against the male cases. No evidence of a significant difference was observed between the genders of the cases at either the genotypic (p=0.19) or allelic (p=0.88) level, suggesting that the increased association observed in females is likely to be attributable to differences in the control sample. The female controls were subsequently analysed against the males, revealing no evidence of a significant difference at the genotypic level (p=0.19), but a weakly significant difference was observed between the genders at the allelic level (p=0.04).

Although the 77878 AAC-Ins was not associated with schizophrenia in the whole case-control sample given the increased association observed in females for the 77394 VNTR, I re-analysed the AAC insertion polymorphism to look for a gender effect. The results are shown in table 4.15 and 4.16 below. No association with schizophrenia was observed, although there was a trend similar to that reported (Tan, Chong et al. 2005) with the homozygote insertion frequency being more frequent in the female cases than controls and the insertion allele also being more frequent in the cases (table 4.15). In males, there was no evidence of association observed at either the genotypic (p=0.82) or allelic (p=0.61) level (table 4.16).

Table 4.15: Individual genotyping results for 77878 AAC Ins in females.

Genotype and allele counts and frequencies for the 77878 AAC Insertion polymorphism, in the female association sample. Both genotypic and allelic chi-square and p-values are shown.

SNP	Genotypes			Alleles	
77878 AAC Ins	**Del/Del**	**Del/Ins**	**Ins/Ins**	**Del**	**Ins**
Cases (n=211)	61 (0.29)	113 (0.54)	37 (0.17)	235 (0.56)	187 (0.44)
Controls (n=234)	77 (0.33)	128 (0.55)	29 (0.12)	282 (0.60)	186 (0.40)
	$\chi^2 = 2.58$	P = 0.28		$\chi^2 = 1.90$	P = 0.17

Table 4.16: Individual genotyping results for 77878 AAC Ins in males.

Genotype and allele counts and frequencies for the 77878 AAC Insertion polymorphism in the male association sample. Both genotypic and allelic chi-square and p-values are shown.

SNP	Genotypes			Alleles	
77878 AAC Ins	**Del/Del**	**Del/Ins**	**Ins/Ins**	**Del**	**Ins**
Cases (n=439)	135 (0.31)	221 (0.50)	83 (0.19)	491 (0.56)	387 (0.44)
Controls (n=479)	151 (0.32)	245 (0.51)	83 (0.17)	547 (0.57)	411 (0.43)
	$\chi^2 = 0.39$	P = 0.82		$\chi^2 = 0.26$	P = 0.61

4.3.1.2 Haplotype analysis.

As the previous studies (Novak, Kim et al. 2002; Tan, Chong et al. 2005) reported that the two polymorphisms were in high linkage disequilibrium with each other I analysed my data to determine the extent of LD in my sample. This revealed that the two polymorphisms are in strong LD with each other (D'= 0.96). As the VNTR is not a bi-allelic marker I was unable to directly calculate an r^2 value between the two markers. However, by collapsing the 77394 VNTR genotyping data (grouping the two larger alleles 2 and 3 together) I was able to calculate both a D' and r^2 value. This revealed that the two polymorphisms are in strong LD with each other (D'=0.96, r^2 =0.74).

Haplotype analysis for the two markers revealed no stronger evidence for association (global p-value = 0.07, 5df) and none of the individual haplotypes were more significantly associated than the individual association finding (table 4.17).

Table 4.17. Haplotype analysis of NOGO.

Haplotype counts and frequencies for the NOGO data, the corresponding global p-value and the individual haplotype chi-square and p-values.

Haplotype	Cases (freq)	Controls (freq)	Chi-square	P-value
AAC Del - ATCT Del	93 (0.07)	81 (0.06)	2.32	0.13
AAC Del - ATCT Ins	583 (0.45)	701 (0.49)	5.91	0.02
AAC Del - AATCTATCT Ins	54 (0.04)	45 (0.03)	2.13	0.14
AAC Ins - ATCT Del	565 (0.43)	581 (0.41)	1.64	0.2
AAC Ins - ATCT Ins	7 (0.005)	13 (0.009)	1.10	0.29
AAC Ins - AATCTATCT Ins	2 (0.001)	1 (0.0002)	0.85	0.36
Global p-value	**0.07 (5 df)**			

I also performed haplotype analysis on the data according to gender as a result of the gender effect observed for the VNTR polymorphism. The haplotype analysis of the female data showed evidence of association (global p-value = 0.01, 5df) (table

4.18) but was less significant than the association result for the individual VNTR genotyping data in females. Although none of the individual female haplotype combinations were more significant than the original VNTR association, the frequency of the putative risk haplotype (AAC Ins - ATCT Del) (Tan, Chong et al. 2005) is over represented (increase of 5%) in the cases compared to the controls, whilst the putative protective haplotype (AAC Del - ATCT Ins) is over represented (increase of 9%) in the controls (table 4.18) a finding consistent with the observations made in previous studies (Chen, Gu et al. 2004a; Tan, Chong et al. 2005). The haplotype analysis of the male data revealed no evidence of association (global p-value = 0.24, 5df) with no individual haplotype combinations displaying any association (table 4.19).

Table 4.18. Haplotype analysis of NOGO in females.

Haplotype counts and frequencies for the female NOGO data, the corresponding global p-value and the individual haplotype chi-square and p-values.

Haplotype	Cases (freq)	Controls (freq)	Chi-square	P-value
AAC Del - ATCT Del	27 (0.06)	27 (0.06)	0.14	0.71
AAC Del - ATCT Ins	187 (0.44)	245 (0.53)	5.95	0.01
AAC Del - AATCTATCT Ins	20 (0.05)	7 (0.02)	7.75	0.005
AAC Ins - ATCT Del	184 (0.44)	180 (0.39)	2.29	0.13
AAC Ins - ATCT Ins	2 (0.005)	4 (0.01)	0.75	0.39
AAC Ins - AATCTATCT Ins	1 (0.002)	1 (0.002)	0.89	0.34
Global p-value	**0.01 (5 df)**			

Table 4.19. Haplotype analysis of NOGO in males.

Haplotype counts and frequencies for the male NOGO data, the corresponding global p-value and the individual haplotype chi-square and p-values.

Haplotype	Cases (freq)	Controls (freq)	Chi-square	P-value
AAC Del - ATCT Del	62 (0.07)	53 (0.06)	2.23	0.14
AAC Del - ATCT Ins	384 (0.45)	456 (0.48)	1.39	0.24
AAC Del - AATCTATCT Ins	33 (0.04)	37 (0.04)	0.001	0.97
AAC Ins - ATCT Del	368 (0.43)	401 (0.42)	0.25	0.62
AAC Ins - ATCT Ins	3 (0.004)	7 (0.008)	1.02	0.31
AAC Ins - AATCTATCT Ins	2 (0.002)	1 (0.001)	2.71	0.1
Global p-value	**0.24 (5 df)**			

4.3.1.3 Allelic expression assay of NOGO.

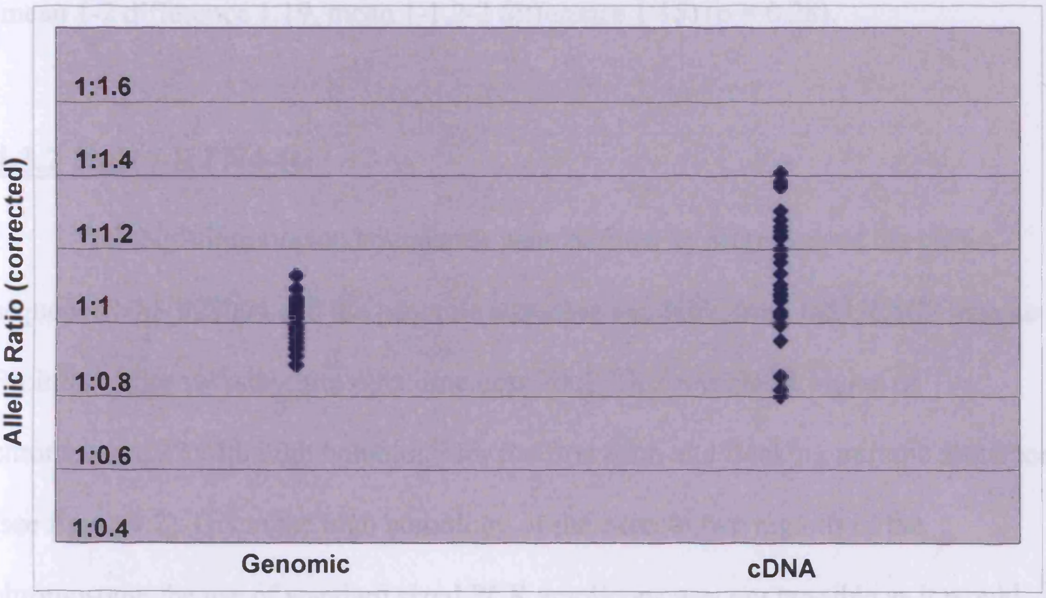
Given that the main rational for studying NOGO was based upon the observed over expression of NOGO in schizophrenia (Novak, Kim et al. 2002), I assayed the NOGO/RTN4 gene for evidence of *cis*-acting sequence variation affecting the gene expression using the allele specific expression assay (Bray, Buckland et al. 2003a), described in section 2.8.

The G>C SNP rs6545468, located in the 5'UTR was chosen as a tag SNP for the assay as it has a high minor allele frequency (0.4). Of the 60 brain samples that were genotyped for SNP rs6545468, 35 were heterozygous and informative for the assay. Of the 35 heterozygotes assayed 17 displayed an expression difference greater than the arbitrary 20% threshold when compared to the corresponding average genomic ratio (see Graph 4.1). On the graph this threshold is represented by a ratio greater than 1:1.2 or lower than 1:0.83. The data points from each individual cDNA sample showed reasonable reproducibility, with an average SD of the mean of 0.11. The individuals who demonstrated a relative expression difference which exceeded the 20% threshold showed expression differences ranging from a 27% increase of the

G allele relative to the C allele (1:0.79), to a 40% decrease of the G allele relative to the C allele (1:1.40).

Graph 4.1: RTN4 Allelic Expression assay: rs6545468.

Corrected genomic and cDNA ratios for rs6545468. Comparison between the observed corrected genomic ratios and the corrected cDNA ratios assayed for NOGO/RTN4 (n=35). Data are expressed as the mean of the ratio of G:C for two measurements of each genomic DNA sample and two measurements of each cDNA sample.



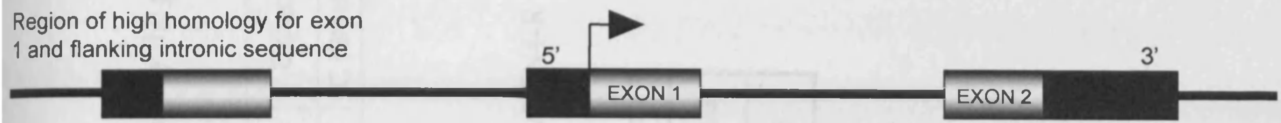
I also analysed the allelic expression data to determine whether the 77394 VNTR genotype correlated with the relative expression difference observed in the allelic expression assay by comparing the mean relative expression difference of the samples heterozygous (1-2) for the VNTR against the mean relative expression difference for the homozygous samples (1-1, 2-2). The heterozygotes were compared against both homozygotes as allelic expression differences are assumed to reflect heterozygosity at functional loci. The analysis revealed no evidence of a correlation between a samples VNTR genotype and the relative allelic expression difference observed in the rs6545468 assay (mean 1-2 difference 1.17, mean 1-1,2-2 difference 1.09) (p = 0.14). Analysis of the markers revealed low levels of LD (D'=0.27, r^2=0.07), and all four haplotypes were observed. Given the low LD, the phase is difficult to establish and as such the allelic expression differences would not be expected in one direction. Therefore the analysis was also repeated after converting all of the relative expression differences to be positive and again revealed no evidence of a correlation between VNTR genotype and the observed relative expression difference (mean 1-2 difference 1.19, mean 1-1,2-2 difference 1.15) (p = 0.28).

4.3.2 NgR / RTN4-R:

The NgR intron/exon boundaries were defined by alignment of the cDNA sequence NM_023004 and the genomic sequence available from the UCSC Genome Bioinformatics website (http://genome.ucsc.edu). This revealed a region of chromosome 22 with high homology for the first exon and flanking intronic sequence (see figure 4.2). Given the high homology of the exon to two regions of the chromosome the use of standard sized PCR amplicons was not possible as it would have amplified two genomic stretches simultaneously. To circumvent this problem a long range PCR followed by a subsequent nested PCR reaction was attempted. However, after several attempts at performing a long range PCR reaction to amplify the specific region of the genome containing the exon of interest, it was only possible to obtain a PCR product which contained multiple products of varying molecular sizes. Further efforts were made to screen the exon by gel extracting the PCR product band of the correct size and then attempting to perform nested PCR using this product as a template. However, no specific NgR products were obtained.

Figure 4.2: Schematic diagram of NgR on chromosome 22.

Diagram of the NgR gene (not to scale), representing the genomic structure on chromosome 22. The genes two exons are displayed with the start of translation indicated by an arrow. The region of high homology for exon one and its flanking intronic sequence is displayed.



Mutation detection revealed only one SNP in NgR which was located in the 5' flanking region (-1807 C>T). The polymorphism identified was genotyped in the pooled association sample. No evidence of association with schizophrenia was found (table 4.21). The extension primer used for pooled genotyping is listed in table 4.20. It should be noted that the minor allele frequency of the polymorphism is vary rare (0.01-0.02), which increases the likelihood of technical variation occurring whilst performing pooled genotyping.

4.3.2.1 Allelic expression assay of NgR.

I was unable to assay the NgR gene for *cis*-acting variants affecting the expression of the gene, due to the lack of an exonic polymorphism within my study or the public databases.

Table 4.20: Polymorphism and primer extension reaction information.

This table shows the sequence flanking each of the NgR/RTN4-R polymorphisms, the extension primer used to assay the polymorphism and the expected alleles in the primer extension reaction.

SNP number.	rs Number	SNP Flanking sequence	Extension primer	Expected Alleles
-1807 C>T	N/A	CTTCCTCCTGCC[C/T]GGCTCCAACCCC	TGGTCCACTCCTTCCTCCTGCC	C (Black) & T (Red)

Table 4.21: Pooled genotyping data for NgR/RTN4-R.

This table shows the estimated allele frequencies and allele counts for all of the polymorphisms in the NOGO-R/RTN4-R gene, along with their corresponding p-values.

SNP Data		Pooled Genotyping data				
SNP number.	rs Numbers	Alleles Cases (freq)	Alleles Controls (freq)	Difference	Chi-square	P-value
-1807 C>T	N/A	C= 1082 (0.98) T= 22 (0.02)	C= 1090 (0.99) T= 14 (0.01)	0.01	1.81	0.18

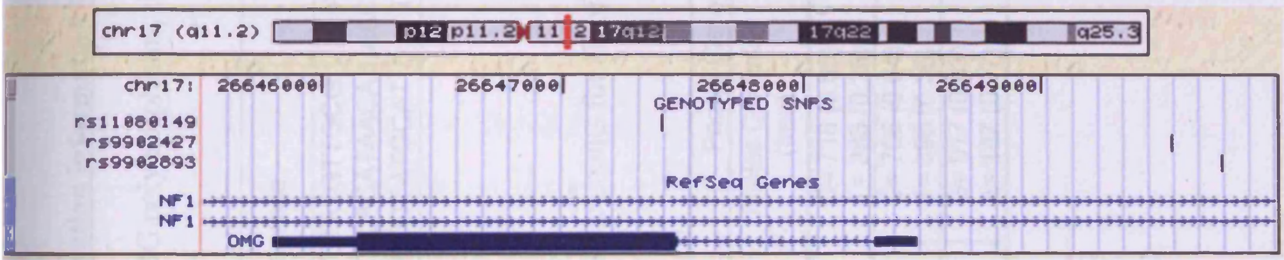
4.3.3 OMG:

Determination of OMG intron/exon boundaries was based on the alignment of the cDNA sequence NM_002544 and the genomic sequence available from the UCSC Genome Bioinformatics website (http://genome.ucsc.edu).

Mutation detection assays using DHPLC and subsequent sequencing analysis identified three SNPs in the OMG gene (See appendix 4.3 for sequencing traces), one non-synonymous SNP G21D (rs11080149) and two 5' flanking region SNPs (rs9902893 and rs9902427) (see figure 4.3).

Figure 4.3 Schematic diagram of OMG.

Diagram of the OMG gene (adapted from the UCSC Genome Browser), indicating the positions of the identified polymorphisms. Gene is shown (3'-5').



All three polymorphisms were genotyped in the pooled association sample. The extension primers used for pooled genotyping are listed in table 4.22. The calculated allele counts and frequencies for all of the genotyped polymorphisms are shown in table 4.23. No significant differences were observed between the allele frequencies in the cases and controls for any of the polymorphisms. The non-synonymous G21D polymorphism has been previously reported to be associated with autism (Vourc'h, Martin et al. 2003).

Table 4.22: Polymorphism and primer extension reaction information for OMG.

This table shows the sequence flanking each of the OMG polymorphisms, the extension primer used to assay the polymorphism and the expected alleles in the primer extension reaction.

SNP number.	rs Number	SNP Flanking sequence	Extension primer	Expected Alleles
-2289 C>T	rs9902893	ATCAACAATAAA[C/T]TTCTATGTTGCG	AATAAAGCACCCAATCAACAATAAA	C (Black) & T (Red)
-2081 C>T	rs9902427	AAAGAGACTTTC[C/T]TTACACATAACA	AACGAGTAATATTTAGCATGTTATGTGTAA	G (Blue) & A (Green)
61 G>A	rs11080149	TTCTCACACCTG[G/A]TATTTTATGCAT	TTGGAGAGGACAAATGCATAAAATA	C (Black) & T (Red)

Table 4.23: Pooled genotyping data for OMG.

This table shows the estimated allele frequencies and allele counts for all of the polymorphisms in the OMG gene, along with their corresponding p-values.

SNP Data		Pooled Genotyping data				
SNP number.	rs Numbers	Alleles Cases (freq)	Alleles Controls (freq)	Difference	Chi-square	P-value
-2289 C>T	rs9902893	C= 718 (0.67) T= 352 (0.33)	C= 718 (0.65) T= 386 (0.35)	0.02	1.04	0.31
-2081 C>T	rs9902427	C= 715 (0.67) T= 355 (0.33)	C= 706 (0.64) T= 396 (0.36)	0.03	1.98	0.16
61 G>A	rs11080149	G= 912 (0.85) A= 158 (0.15)	G= 917 (0.83) A= 187 (0.17)	0.02	1.92	0.17

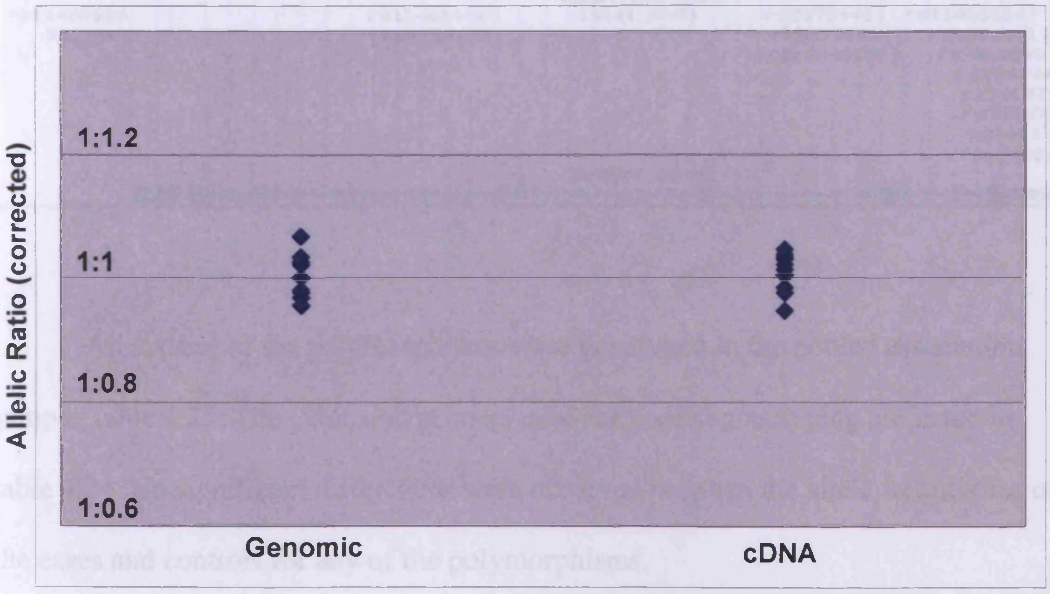
4.3.3.1 Allelic expression of OMG.

Given the two pronged screening strategy employed in this project (described in chapter 1), I assayed the OMG gene for *cis*-acting sequence variation affecting the gene expression using the allele specific expression assay (Bray, Buckland et al. 2003a) described in section 2.8.

Marker rs11080149 was selected for analysis given its fairly informative minor allele frequency (0.17). Of 60 subjects for whom brain was available, 12 individuals were heterozygous for SNP rs11080149. Of the 12 OMG heterozygous individuals, none showed a relative expression difference beyond the 20% limit of measurement error (cDNA ratios varied from 1:1.05 to 1:0.95) (Graph 4.2). The data points from each individual cDNA sample showed good reproducibility, with an average SD of the mean of 0.04. The assay provides no evidence for the presence of a *cis*-acting variant affecting the allelic expression of OMG.

Graph 4.2: OMG Allelic Expression assay: rs11080149.

Corrected genomic and cDNA ratios for rs11080149. Comparison between the observed corrected genomic ratios and the corrected cDNA ratios assayed for OMG (n=12). Data are expressed as the mean of the ratio of G:A for two measurements of each genomic DNA sample and two measurements of each cDNA sample.

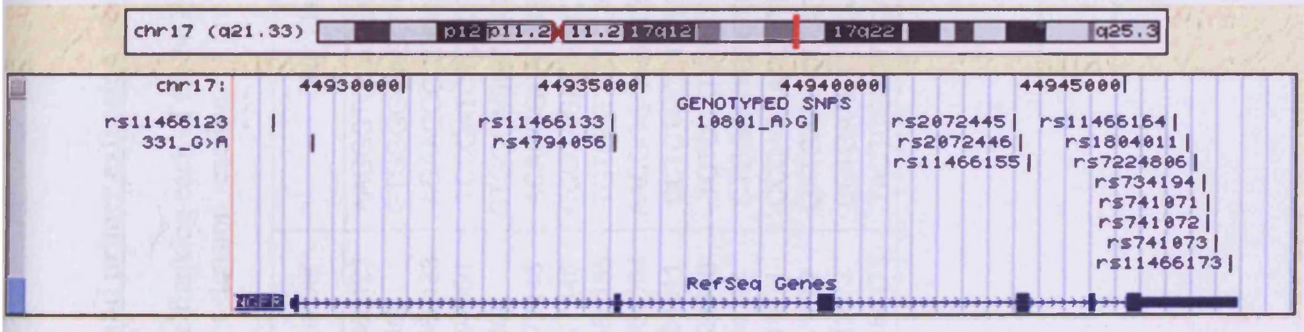


4.3.4 NGFR:

Determination of the NGFR intron/exon boundaries was based on the alignment of the cDNA sequence NM_002507 and the genomic sequence available from the UCSC Genome Bioinformatics website (http://genome.ucsc.edu). Mutation detection assays identified fifteen SNPs along with an insertion/deletion (See appendix 4.3 for sequencing traces). There was one non-synonymous SNP S205L (rs2072446), one coding synonymous SNP G265G (rs11466155), eight 3'-UTR SNPs (rs11466164, rs1804011, rs7224806, rs734194, rs741071, rs741072, rs741073 and rs11466173), one 5' flanking region SNP (rs11466123) and five intronic polymorphisms (331 G>A, rs11466133, rs4794056, 10801 A>G and rs2072445) (see figure 4.4).

Figure 4.4: Schematic diagram of NGFR.

Diagram of NGFR (adapted from the UCSC Genome Browser), indicating the positions of the identified polymorphisms. For numbering of polymorphisms, +1 indicates the 'A' of the ATG translation initiation codon. Gene is shown (5'-3').



All sixteen of the polymorphisms were genotyped in the pooled association sample, table 4.25. The extension primers used for pooled genotyping are listed in table 4.24. No significant differences were observed between the allele frequencies of the cases and controls for any of the polymorphisms.

Table 4.24: Polymorphism and primer extension reaction information.

This table shows the sequence flanking each of the NGFR polymorphisms, the extension primer used to assay the polymorphism and the expected alleles in the primer extension reaction.

SNP number.	rs Number	SNP Flanking sequence	Extension primer	Expected Alleles
-486 A>C	rs11466123	AAGGCTCCAGGG[A/C]GAAGGTGAAGCC	CCATCTTCCTCCGCCTCTGGCTTCACCTTC	T (Red) & G (Blue)
331 G>A	N/A	GTGGGGGAGCTG[G/A]GAGGGGTCTTTC	CATGCCCCCTCTTGAAAGACCCCTC	C (Black) & T (Red)
6583 C-Ins	rs11466133	AGAACCCCCCCC[C/-]AACCCACCCCAG	GGTGTTTGATTCCCCGGAAGAACCCCCCCC	C (Black) & A (Green)
6637 C>T	rs4794056	TCCGATCTCCCT[C/T]CATCCAGGTGTC	GCCAGTCTGACCCTCCGATCTCCCT	C (Black) & T (Red)
10801 A>G	N/A	CTGGCAGGCAAT[A/G]GGGGAGGGAGAG	TCAGCGTCCTGGCAGGCAAT	A (Green) & G (Blue)
14932 G>T	rs2072445	ACACGGCAGTGG[G/T]TTAGAGCTAAAA	CCCCACTCCTCCCTTTTAGCTCTAA	C (Black) & A (Green)
15039 C>T	rs2072446	CCCCAGAGGGCT[C/T]GGACAGCACAGC	ATTACACGGTCCACACCCCCAGAGGGCT	C (Black) & T (Red)
15220 C>T	rs11466155	TGTGGTTGTGGG[C/T]CTTGTGGCCTAC	CATCCTGGCTGCTGTGGTTGTGGG	C (Black) & T (Red)
18255 A>G	rs11466164	AACGGGGAGGCC[A/G]AGTGCAGGCTGG	CCCAAAGCCAGACAGCAACGGGGAGGCC	A (Green) & G (Blue)
18490 C>A	rs1804011	CCTCTGCCTGTC[C/A]CTCTCAGGCATG	GGTGCTGCTTCCCTCTGCCTGTC	C (Black) & A (Green)
18684 T>C	rs7224806	TCTCCATGAGTT[T/C]TTTCTCTTGGGC	AGAAATCTCACTTTTCTCCATGAGTT	T (Red) & C (Black)
18829 T>G	rs734194	CTGGCGTCTGTC[T/G]TCAAGGGCTTAC	CATTCCTCCACGTGTAAGCCCTTGA	A (Green) & C (Black)
18910 T>C	rs741071	CCCAGAAGGTTG[T/C]GATGAAGAAAAG	GGGTTGGCTGGGCCCAGAAGGTTG	T (Red) & C (Black)
18923 T>C	rs741072	GATGAAGAAAAG[T/C]GGGCCAGTGTGG	CCGCATTCCCACACTGGCCC	A (Green) & G (Blue)
19106 G>A	rs741073	CCTGAAGTTGGA[G/A]TGAGTGTGGCTC	CTGTTCTGTTTTGCCTGAAGTTGGA	G (Blue) & A (Green)
19421 G>C	rs11466173	TACTCCCCTTGA[G/C]CTGAGATGGAAC	GGCCCCTTTTTACTCCCCTTGA	G (Blue) & C (Black)

Table 4.25: Pooled genotyping data for NGFR.

Estimated minor allele frequencies and minor allele counts for all of the polymorphisms detected in NGFR, along with their corresponding chi-square and p-values.

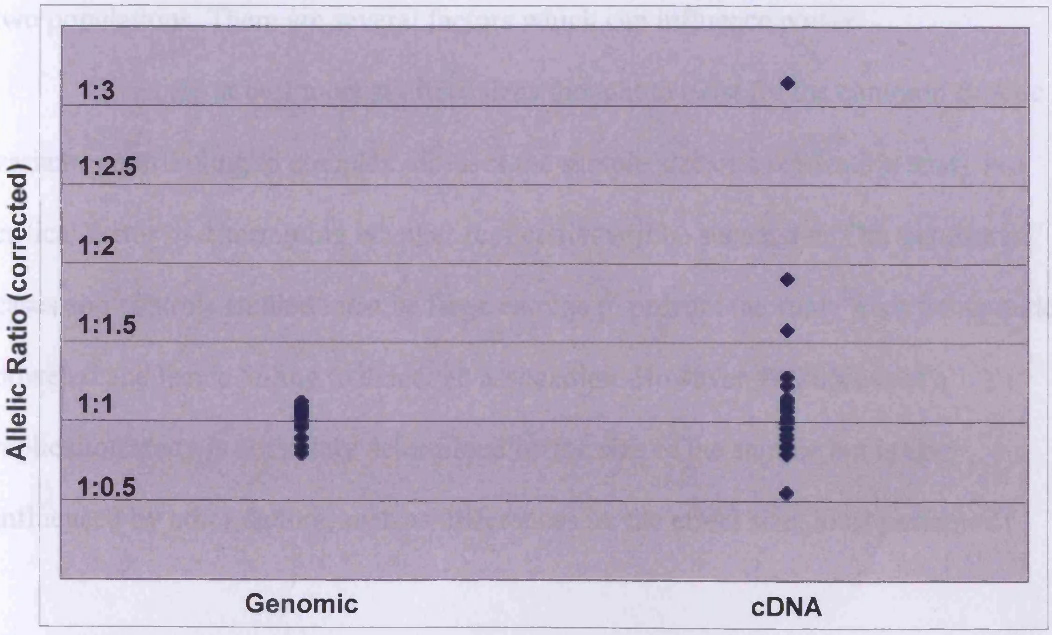
SNP Data		Pooled Genotyping data				
SNP number.	rs Numbers	Minor Allele Cases (freq)	Minor Allele Controls (freq)	Difference	Chi-square	P-value
-486 A>C	rs11466123	C= 24 (0.02)	C= 27 (0.02)	0	0.1	0.76
331 G>A		A= 53 (0.05)	A= 48 (0.04)	0.01	0.45	0.50
6583 C-Ins	rs11466133	C Ins= 125 (0.12)	C Ins= 136 (0.12)	0	0.21	0.65
6637 C>T	rs4794056	T= 180 (0.17)	T= 168 (0.15)	0.02	1.04	0.31
10801 A>G		G= 173 (0.16)	G= 176 (0.16)	0	0.02	0.89
14932 G>T	rs2072445	T= 96 (0.09)	T= 102 (0.09)	0	0.05	0.83
15039 C>T	rs2072446	T= 54 (0.05)	T= 63 (0.06)	0.01	0.47	0.49
15220 C>T	rs11466155	T= 371 (0.35)	T= 382 (0.35)	0	0.001	0.97
18255 A>G	rs11466164	G= 157 (0.15)	G= 162 (0.15)	0	0.0001	0.99
18490 C>A	rs1804011	A= 307 (0.29)	A= 328 (0.3)	0	0.27	0.60
18684 T>C	rs7224806	C= 138 (0.11)	C= 152 (0.11)	0.01	0.36	0.55
18829 T>G	rs734194	G= 137 (0.13)	G= 156 (0.14)	0.01	0.82	0.37
18910 T>C	rs741071	C= 467 (0.44)	C= 483 (0.44)	0	0.002	0.96
18923 T>C	rs741072	C= 496 (0.46)	C= 508 (0.46)	0	0.03	0.87
19106 G>A	rs741073	A= 319 (0.3)	A= 332 (0.3)	0	0.02	0.89
19421 G>C	rs11466173	C= 80 (0.08)	C= 71 (0.06)	0.02	0.92	0.34

4.3.4.1 Allelic expression assay of NGFR.

Marker rs1804011 was selected for allelic expression analysis given its high informativity (minor allele frequency 0.3). Of the 60 subjects for whom brain was available, 26 were heterozygous for SNP rs1804011. Of the 26 NGFR heterozygous individuals, seven showed a relative expression difference of ≥ 20% compared to the corresponding average genomic ratios (see Graph 4.3). The data points from each individual cDNA sample showed good reproducibility, with an average SD of the mean of 0.09. The C:A ratio in cDNA departed from the genomic 1:1 ratio in both directions with the expression difference ranging from a 215% increase in the A allele relative to the C allele (1:3.15) to a 82% decrease of the A allele relative to the C allele (1:0.55). That the expression difference goes in both directions on the graph might reflect the effect of a *cis*-acting variant that is in low LD with the tag SNP that was used or the presence of more than one variant affecting the genes expression.

Graph 4.3: NGFR Allelic Expression assay: rs1804011.

Corrected genomic and cDNA ratios for rs1804011. Comparison between the observed corrected genomic ratios and the corrected cDNA ratios assayed for NGFR (n=26). Data are expressed as the mean of the ratio of C:A for two measurements of each genomic DNA sample and two measurements of each cDNA sample.



4.4 Discussion:

In this chapter I have sought to test the hypothesis that genetic variants within the NOGO gene and also within its associated receptor complex might be associated with susceptibility to schizophrenia. There have been previous studies (Novak, Kim et al. 2002; Chen, Gu et al. 2004a; Tan, Chong et al. 2005) reporting association between various polymorphisms in the NOGO gene and schizophrenia susceptibility, although it should be noted that there have also been several negative association studies published (Covault, Lee et al. 2004; Gregorio, Mury et al. 2005; Xiong, Rouleau et al. 2005).

There are three main reasons for the failure of an association study to replicate: a) the original association finding is a false positive caused by statistical chance or other problems with the study design (e.g genotyping error) and the lack of replication is correct, b) the original association finding is true and the lack of replication is a false negative result caused by a lack of power in the second study, c) the original association finding is true in that particular population but there is no association present in a second population because of differences that exist in the phenotypes between the samples or genetic and environmental backgrounds of the two populations. There are several factors which can influence power.

Given the at best modest effect sizes thought to exist for the common genetic variants contributing to complex diseases the sample size of a replication study is a critical factor in determining whether replication will be successful. The samples of cases and controls studied must be large enough to prevent the study from being under powered and hence failing to detect an association. However the success of a replication study is not solely determined by the size of the sample but is also influenced by other factors, such as differences in; the effect size, local patterns of

LD, the allele frequency of the disease allele, and differences in the gene-gene and gene-environment interactions between the populations studied.

The ability to replicate an association finding is influenced by the effect size of the associated variant. Researchers reporting a negative replication often state that the power of their study was sufficient to detect an associated variant with a certain magnitude of genetic effect. However, it is important to recognise that the true effect size of the genetic variant may have been over estimated in the initial association study. In a meta-analysis study of 370 genetic studies addressing 36 reported associations (Ioannidis, Ntzani et al. 2001), the authors demonstrated that the initial effect size of the associated variant was over estimated in 26 of the 36 reported associations studied. Hence a replication study might report a negative result due to unknowingly being under powered to detect the associated variant with its lower genetic effect size.

The local LD structure around the associated variant can also influence the results of a replication study. This is of particular importance if the variant reported to be associated with disease is not the true causal variant, but is instead associated with disease by virtue of linkage disequilibrium with the causal variant. If the extent of historical recombination is different between the associated and causal variants in different populations, then the strength of the association will also differ between the populations and may account for the lack of positive replication. This problem could be avoided by testing the causal variant itself, or a marker which is in high LD with the causal variant in both populations, however this requires the identification of the true causal variant, something which is often not possible in molecular genetic studies of complex diseases.

The allele frequency of the associated disease allele is a significant factor in determining the success of a replication study. The frequency of an allele can differ depending on the ethnicity of the population being studied. Two distinct populations from different ethnic backgrounds are likely to have different genetic histories, rates of disease and also different allele frequencies at some loci. Thus if a replication study is performed in a population that is not identical to the original association population allele frequency differences for the reported disease allele may account for the lack of a positive association finding. Indeed, the original association may be true in the specific population tested, whilst not being associated in the second genetically distinct population. This factor is likely to be of particular importance if an association is reported with a rare genetic variant. Rare variants are thought to have a higher likelihood of being population specific, due to originating fairly recently in history and hence still being specific to a founder population or slightly less recently and being found in one major ethnic group (Risch 2000). This would tend to make any association reported with rare variants less replicable in distinct populations. Common variants are thought to be more likely to be found globally in most populations, making an association more reproducible in many ethnically diverse populations (Risch 2000).

The allele frequency of a disease allele can also be influenced by ascertainment bias relating to the phenotype. The phenotype of a person normally determines their selection for study (e.g schizophrenic or healthy control). However the method of selection can also influence the phenotype being selected. For instance, a sample of schizophrenic patients ascertained from a hospital inpatient clinic, may be enriched for specific phenotype components, for example gender, disease severity, or violence, compared to a sample of outpatient schizophrenics living in the community.

If phenotypic variation results from genetic variation, certain allele frequencies in the inpatient sample may differ from those of an independent sample of schizophrenics collected using a different ascertainment method. Thus association reported in the inpatient association sample may not replicate due to the associated variant being a specific risk factor for the particular phenotype selected in your sample.

A further factor relating to the replication of association studies and the disease allele frequency is the reliability of disease diagnosis. In schizophrenia and other psychiatric disorders there is no definitive diagnostic test for the presence of disease and hence a diagnosis is made by clinicians using symptom rating scales. This provides the potential for diagnoses to be made slightly differently depending on a patients clinician, the symptom rating scale being used and their interpretation of the patients symptoms. This could lead to different patients being diagnosed with slightly different phenotypes (schizophrenia/schizoaffective /bipolar disorder etc) even though both display identical symptoms. These differences in diagnosis could again lead to allele frequency differences between association study samples and therefore influence the outcome of replication studies.

Finally it is likely that common complex diseases are the result of both genetic and environmental factors and their interactions with each other. If an associated variants influence on disease susceptibility is only present on a certain genetic and/or environmental background, then the different gene-gene and gene-environment interactions that are present in distinct populations might influence the replication of an association study by changing the effect size at any given locus.

When undertaking molecular genetic studies researchers often ignore the possibility that there may be genotype-environment interactions present. However it is likely that distinct populations will have been exposed to different environment factors and hence in a similar way to gene-gene interactions; it is possible that different combinations of gene-environment interactions may be responsible for disease susceptibility in different populations. For instance it has been shown that a variant within the promoter of the hepatic ligase gene has a different influence on high density lipoprotein metabolism depending on the presence of a high-fat or low-fat diet (Ordovas, Corella et al. 2002). Thus an association study of this genetic variant in populations with different dietary fat intakes might give inconsistent results. The influence of gene-environment interactions on replication studies may be impossible to avoid as collecting two independent samples exposed to exactly the same environment is virtually impossible.

The only way to confirm/exclude the report of an association is by gaining confidence through the findings of many large sufficiently powered independent replication studies reporting association with the same allele, in the same phenotype, with the same direction of effect and with a p-value low enough to survive stringent correction for multiple testing.

Four genes involved with the NOGO receptor complex (NOGO, NgR, OMG and NGFR) have been examined in this study. A total of forty three polymorphisms were detected by DHPLC in the four genes. Of these forty three variants, thirty seven were genotyped in the pooled association sample. Of the seven variants that were not genotyped, the minor allele frequency of four was too rare to allow reliable allele frequency estimates by pooled genotyping and two of the polymorphisms were not compatible with genotyping by using a primer extension assay. Of the thirty seven polymorphisms successfully typed in the pooled samples none showed significant differences between the schizophrenia case and control pools.

However, when individual genotyping was performed on the NOGO 77394 VNTR polymorphism, a significant (uncorrected) association with schizophrenia was observed at both the genotypic (p=0.02) and allelic level (p=0.02), with the shorter repeat VNTR being over represented in the schizophrenics compared to control subjects. As one of the previous studies (Tan, Chong et al. 2005) reported the presence of a gender effect in females, my data was also analysed according to gender, revealing that the association was stronger in females, with the smaller repeat of the VNTR being strongly associated with schizophrenia (p=0.004). In light of the previous reports of association between the 77394 AAC insertion and schizophrenia, and the fact the VNTR and insertion had been reported to be in strong LD with each other, I decided to also genotype the AAC insertion individually. No evidence of association between the 77394 AAC insertion and schizophrenia in either my whole sample or the female subgroup was observed, although the AAC insertion frequency was increased in schizophrenics in both of the tested samples. Haplotypic analysis based on the two NOGO variants was significant but less so than the VNTR polymorphism alone. Although the two polymorphisms did not both display evidence of association with schizophrenia in my sample they do however both appear to be highly related to each other (D'=0.96, r^2 =0.74).

In addition to the genetic analysis of the four NOGO receptor complex genes three were analysed for *cis*-acting variants affecting the genes expression in the brain. The NgR gene was unable to be assayed for altered expression due to the lack of a suitable marker SNP to perform the assay. Both the NOGO and NGFR genes displayed evidence for the presence of genetic variants affecting their expression in the brain. The relative cDNA allelic expression differences observed for both NOGO and NGFR deviated from the genomic 1:1 ratio in both directions, which might indicate either the effect of a *cis*-acting variant that is only in weak LD with the tag SNP used (NOGO rs6545468, NGFR rs1804011) or possibly the presence of more than one genetic variant affecting the genes expression. The evidence of a *cis*-acting variant causing an expression difference in the NOGO gene is consistent with the exhibition of a cis-acting allele. If the findings of Novak and colleagues (Novak, Kim et al. 2002) who reported a 60% over expression of NOGO cDNA in schizophrenic brain compared to controls are confirmed in other samples, this would suggest it is at least possible that it may reflect a primary change resulting from a regulatory variant rather than a secondary effect. However, as yet the causal variant is unknown.

As NOGO has been proposed to have an important role in the control of synaptic plasticity and neuronal migration, altered expression of this gene in schizophrenia may contribute to the abnormal neuronal connectivity which is postulated as a mechanism underlying schizophrenia (Friston 1998; Benes 2000). Unfortunately during this project it was not possible to determine the exact location of the genetic variants causing the altered expression in either the NOGO or NGFR genes. However, the 3'UTR of genes are known to play an important role in the regulation of mRNA stability, translation initiation and mRNA localisation (Kuersten and Goodwin 2003) and contain various elements responsible for regulating gene expression, including the poly A tail which enhances translation (Munroe and Jacobson 1990), AU rich sequences responsible for inhibiting translation (Kruys, Marinx et al. 1989), exonic splicing enhancer (ESE) sequences (Schaal and Maniatis 1999) and various other sequences such as seleno-cysteine insertions which may have a profound effect on mRNA stability or translational efficiency (Low and Berry 1996). It is possible therefore that polymorphisms contained within the 3'UTR of a gene may have a profound effect on the genes expression and as a result it is possible that the polymorphisms contained within the NOGO genes 3'UTR contribute to the altered expression of the gene seen in schizophrenia (Novak, Kim et al. 2002). However, it should be noted that when I analysed my data for the 77394 VNTR against the allelic expression data for NOGO, I detected no evidence of a significant correlation between the VNTR genotype present and the extent of the relative allelic expression difference observed in my sample, suggesting this variant is not functional or at least not the predominant functional variant.

The results of my study are not entirely consistent with the findings from other association studies (see table 4.1). Although the evidence for association in my sample is modest, the associated allele of the 77394 VNTR is the same as that reported by Tan and colleagues (Tan, Chong et al. 2005), and the association is also stronger in the female subgroup of my sample ($p=0.004$). However, Novak and colleagues (Novak, Kim et al. 2002) reported association between the 77878 AAC insertion and schizophrenia, but there was no evidence for association of this polymorphism detected in my sample. LD analysis of the two polymorphisms revealed that they are related ($D'=0.96$, $r^2=0.74$). However, as the polymorphisms are not in perfect LD it is feasible that association with schizophrenia may not be detected for both polymorphisms.

There is no obvious explanation for the differences observed in the results of the various association studies, but it might be feasible that different LD structures or alleles are responsible for association to schizophrenia in different populations. This would be supported by the marked observed differences in the two polymorphisms allele frequencies across the various previous association studies (table 4.1). The studies were performed in several ethnically distinct populations including samples consisting of North American and European Caucasians, ethnic and Han Chinese and

Brazilians. Similar allele frequencies to those detected in my study for the 77878 AAC-insertion were reported in two studies (Covault, Lee et al. 2004; Xiong, Rouleau et al. 2005) which were both composed of European Caucasians. However, the allele frequency of the polymorphism reported in the other studies was inconsistent. Surprisingly even the two large studies performed in samples of Chinese descent (Chen, Gu et al. 2004a; Tan, Chong et al. 2005) reported markedly different allele frequencies for both of the polymorphisms studied, indicating the possible existence of allele frequency differences within the Chinese population or alternatively allele labelling error.

It is also possible that the sample sizes used in the various studies were not sufficiently large enough to enable the detection of genetic associations with small effect sizes. Two of the negative replication studies (Covault, Lee et al. 2004; Gregorio, Mury et al. 2005) were performed in small samples of cases (table 4.1) whilst the other negative study (Xiong, Rouleau et al. 2005) was performed in a moderately large case sample (462) but the small control sample (153) may have compromised the power of the study to detect an association.

In order to try and clarify a possible role for NOGO in susceptibility to schizophrenia further the use of independent replication studies performed in sufficiently large and powerful association samples is required. These studies should ideally test all of the non-redundant genetic variation contained within the gene for association to schizophrenia. Additionally given the increasing evidence that some genetic susceptibility loci may be common to diseases crossing the whole spectrum of psychiatric disorders it has been suggested that in order to comprehensively study genes for association to psychiatric diseases a more detailed approach to the analysis of sub-phenotypes should be employed (Craddock, O'Donovan et al. 2006; Craddock, Owen et al. 2006). Such studies would be facilitated by extending the work presented here by identifying the variant driving the allele specific differences in expression.

The results of my studies on NgR, OMG and NGFR provided no evidence that sequence variations within these genes are associated with susceptibility to schizophrenia. The allelic expression assay of NGFR provided evidence of at least one *cis*-acting sequence variant affecting the genes expression in the brain, whilst no evidence of altered expression was detected for the other two genes. Again as for NOGO, further studies to identify the variant responsible would help power future studies.

However, on the basis of my results alone I cannot exclude these genes from playing a role in the aetiology of schizophrenia. In order to do this with any confidence would at least require the genotyping of all the non-redundant genetic variation contained within the gene in a large, sufficiently powered association sample, which was representative of both the general disease and healthy populations. As most of the genotyping performed in this project was performed in pooled samples the extent of the linkage disequilibrium structure (LD) between the polymorphisms was not determined for any of the genes and hence the level of redundancy was not ascertained. Retrospective analysis of the data contained within the HapMap (version 2.0) database revealed only three genetic variants within OMG, of which one was detected in my study (rs11080149) and very low levels of LD across the gene. Analysis of NgR again revealed a low level of genetic variation within the gene (6 markers) and moderately high levels of LD across the gene. NGFR did contain a high level of genetic variation, including seven markers detected in my study, but very little evidence of LD was detected spanning the gene. On the basis of the HapMap data it appears that not all of the non-redundant genetic variation contained within these genes was genotyped in this project. It is however not surprising that all of the non-redundant variation was not screened as a) the genes intronic sequence was not screened for genetic variation, b) SNPs incompatible with pooled genotyping or that had a minor allele frequency which was too rare were not genotyped in the pooled samples, c) the screening set used for mutation detection might not have been powerful enough to detect some rare variants (power of 80% to detect a variant with a MAF of 0.05). It is also a possibility that the association sample employed in this project was not sufficiently large enough to provide the power required to detect association present with rare genetic variants or variants with small genetic effect sizes. As an example, the association sample used in this project provides a power of 94% to detect a variant with a MAF of 0.1, conferring an odds ratio of 1.5, at the 0.05 significance level, but only a power of 32% to detect a variant with a MAF of 0.1, conferring an odds ratio of 1.2, at the 0.05 significance level.

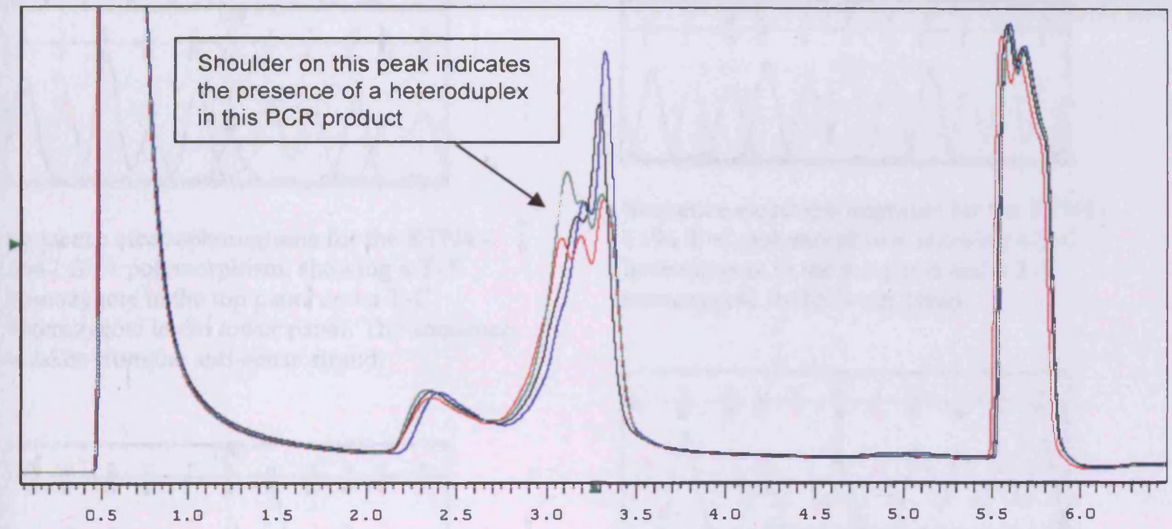
In summary the results of my study provide at least weak support for the hypothesis that the NOGO receptor complex may play a role in the aetiology of schizophrenia. Additionally my results provide support for sequence variation within the gene playing a role in altering the expression of the gene in the human brain and possibly schizophrenia. However, the results of this study need independent replication in other large sufficiently powered association samples, before any more definitive conclusions can be drawn from the results.

4.5 Appendix:

Samples of Experimental data.

Appendix 4.1.

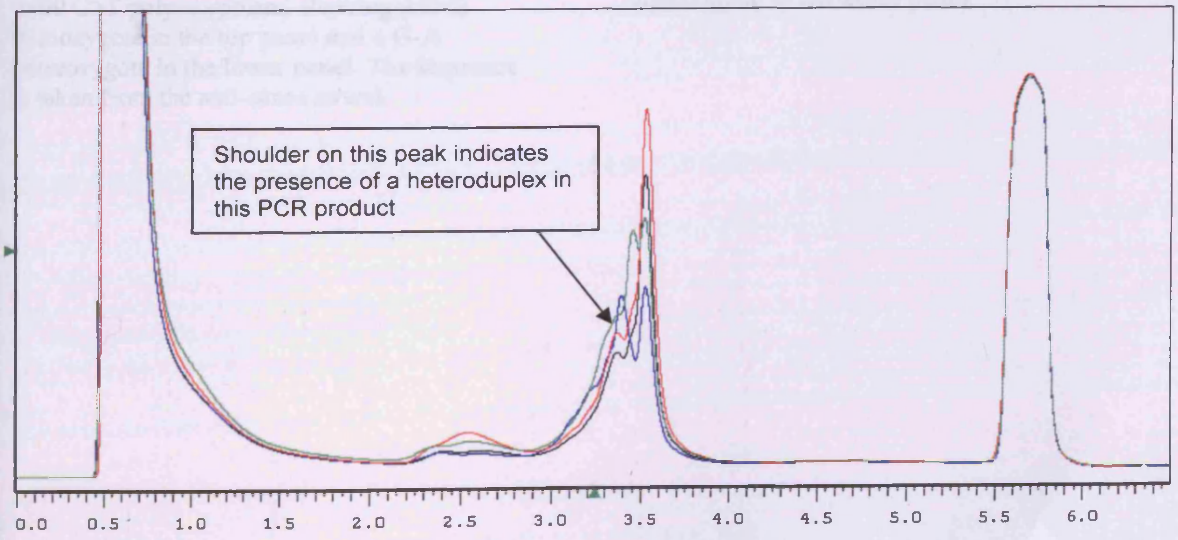
Example of DHPLC chromatographs produced for RTN4 amplimer 22 analysed at 56˚C, for screening set samples 1, 2, 3 and 4. The graph shows the peaks representing these samples PCR products, and changes in the peak shape indicates the presence of a heteroduplex.



Shoulder on this peak indicates the presence of a heteroduplex in this PCR product

Appendix 4.2.

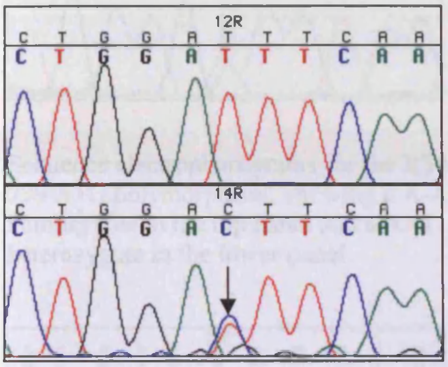
Example of DHPLC chromatographs produced for NGFR amplimer 6 analysed at 62˚C, for screening set samples 2, 3, 4 and 8. The graph shows the peaks representing these samples PCR products, and changes in the peak shape indicates the presence of a heteroduplex.



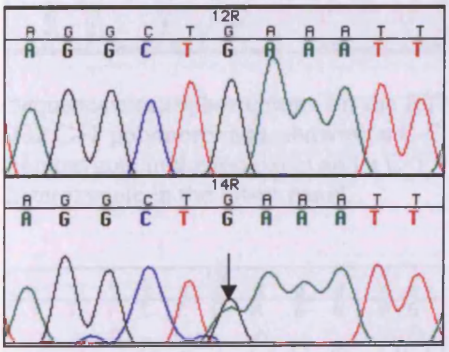
Shoulder on this peak indicates the presence of a heteroduplex in this PCR product

Appendix 4.3:

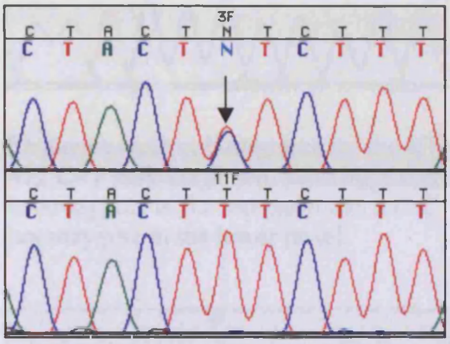
Examples of the different sequencing electrophorograms produced for each of the polymorphisms that were detected as a result of screening the four NOGO receptor complex genes in this project. The SNP is indicated with an arrow.



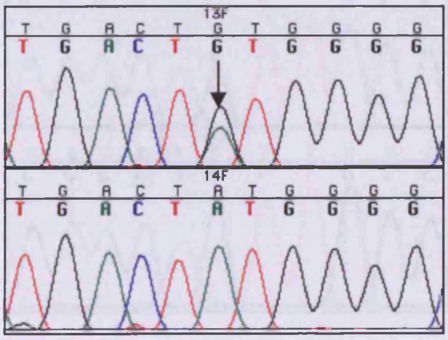
Sequence electrophorograms for the RTN4 - 1447 G>A polymorphism, showing a T-T homozygote in the top panel and a T-C heterozygote in the lower panel. The sequence is taken from the anti-sense strand.



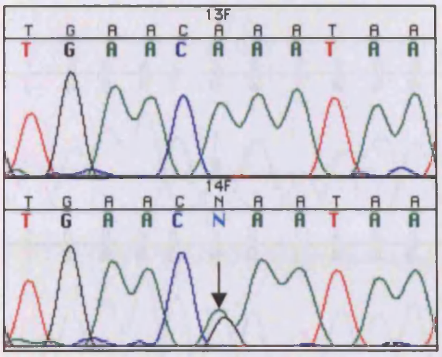
Sequence electrophorograms for the RTN4 - 1409 C>T polymorphism, showing a G-G homozygote in the top panel and a G-A heterozygote in the lower panel. The sequence is taken from the anti-sense strand.



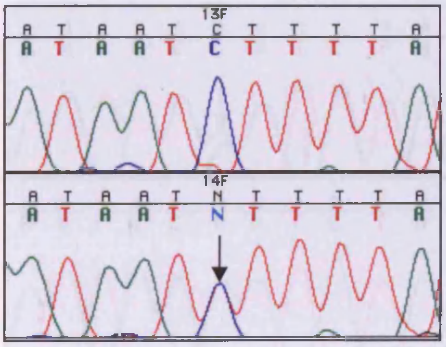
Sequence electrophorograms for the RTN4 - 1196 T>C polymorphism, showing a T-C heterozygote in the top panel and a T-T homozygote in the lower panel.



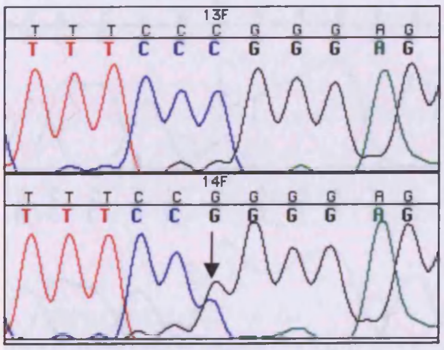
Sequence electrophorograms for the RTN4 - 1123 A>G polymorphism, showing a A-G heterozygote in the top panel and a A-A homozygote in the lower panel.



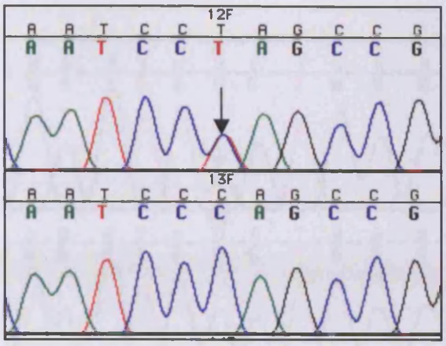
Sequence electrophorograms for the RTN4 - 739 A>G polymorphism, showing a A-A homozygote in the top panel and a A-G heterozygote in the lower panel.



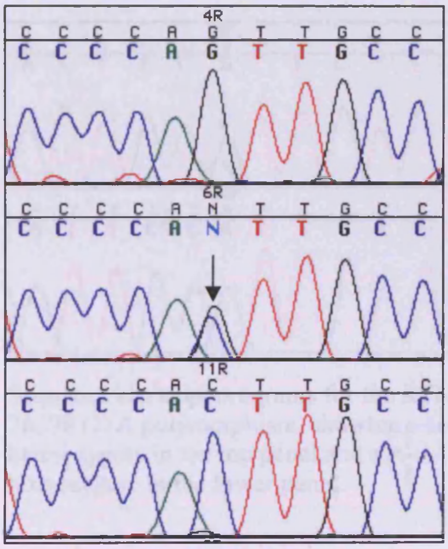
Sequence electrophorograms for the RTN4 - 732 C>T polymorphism, showing a C-C homozygote in the top panel and a C-T heterozygote in the lower panel.



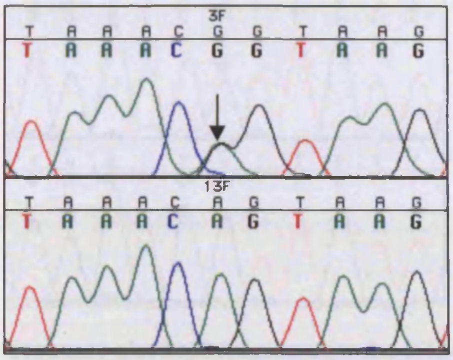
Sequence electrophorograms for the RTN4 - 646 C>G polymorphism, showing a C-C homozygote in the top panel and a C-G heterozygote in the lower panel.



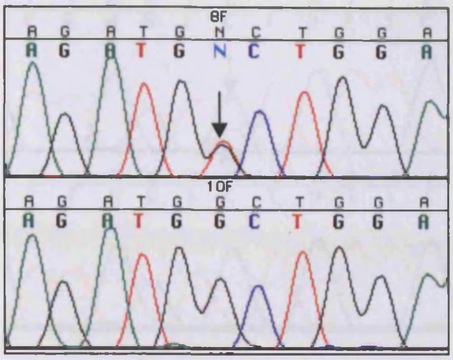
Sequence electrophorograms for the RTN4 - 422 C>T polymorphism, showing a C-T heterozygote in the top panel and a C-C homozygote in the lower panel.



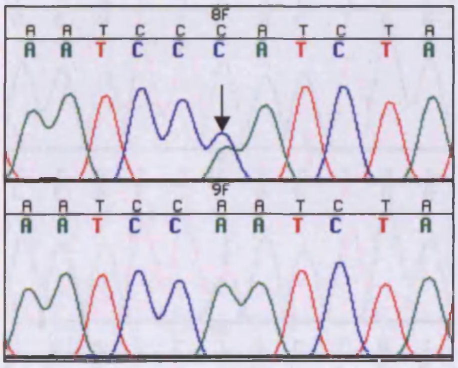
Sequence electrophorograms for the RTN4 - 205 G>C polymorphism, showing a G-G homozygote in the top panel, a G-C heterozygote in the middle panel and a C-C homozygote in the lower panel.



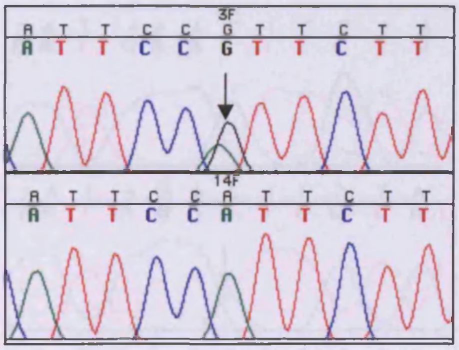
Sequence electrophorograms for the RTN4 22705 A>G polymorphism, showing a A-G heterozygote in the top panel and a A-A homozygote in the lower panel.



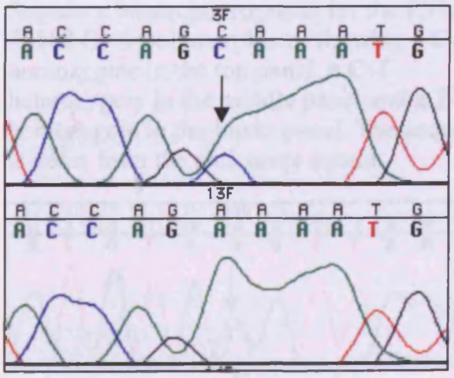
Sequence electrophorograms for the RTN4 62818 T>G polymorphism, showing a T-G heterozygote in the top panel and a G-G homozygote in the lower panel.



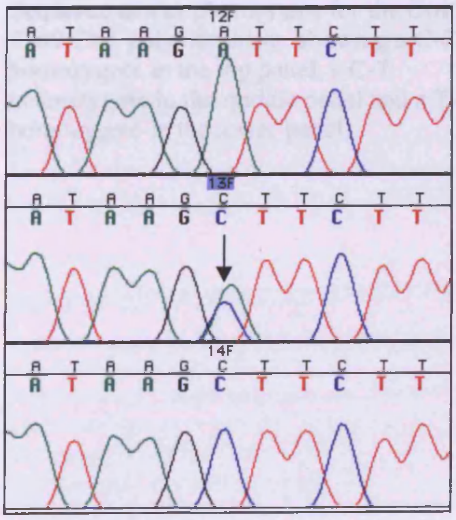
Sequence electrophorograms for the RTN4 25222 C>A polymorphism, showing a C-A heterozygote in the top panel and a A-A homozygote in the lower panel.



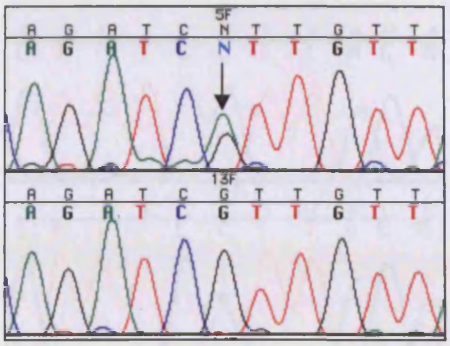
Sequence electrophorograms for the RTN4 76398 G>A polymorphism, showing a G-A heterozygote in the top panel and a A-A homozygote in the lower panel.



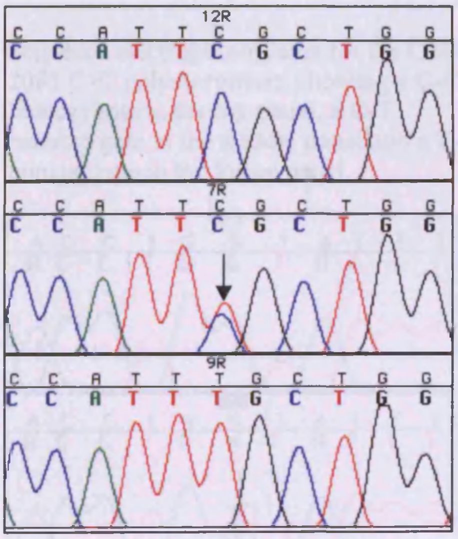
Sequence electrophorograms for the RTN4 40387 A>C polymorphism, showing a A-C heterozygote in the top panel and a A-A homozygote in the lower panel.



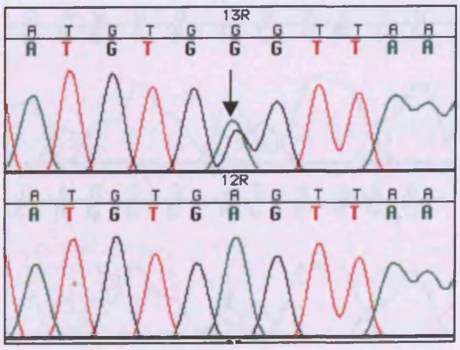
Sequence electrophorograms for the RTN4 76636 A>C polymorphism, showing a A-A homozygote in the top panel, a A-C heterozygote in the middle panel and a C-C homozygote in the lower panel.



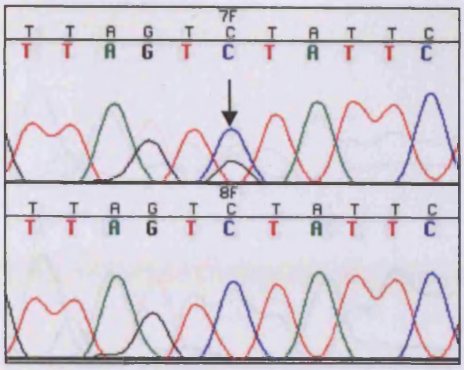
Sequence electrophorograms for the RTN4 77258 G>A polymorphism, showing a G-A heterozygote in the top panel and a G-G homozygote in the lower panel.



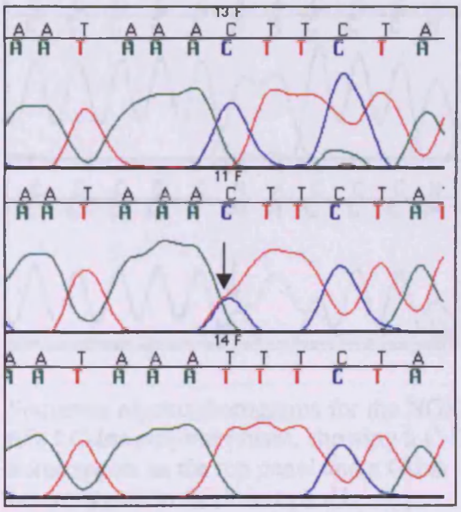
Sequence electrophorograms for the RTN4 78188 G>A polymorphism, showing a C-C homozygote in the top panel, a C-T heterozygote in the middle panel and a T-T homozygote in the lower panel. The sequence is taken from the anti-sense strand.



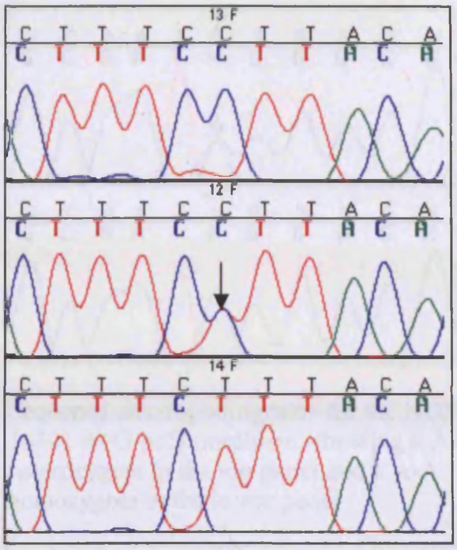
Sequence electrophorograms for the RTN4 78368 T>C polymorphism, showing a A-G heterozygote in the top panel and a A-A homozygote in the lower panel. The sequence is taken from the anti-sense strand.



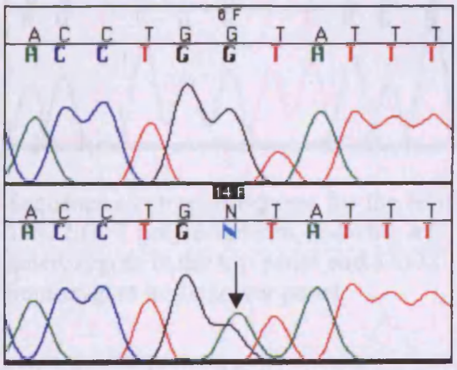
Sequence electrophorograms for the RTN4 78775 C>G polymorphism, showing a C-G heterozygote in the top panel and a C-C homozygote in the lower panel.



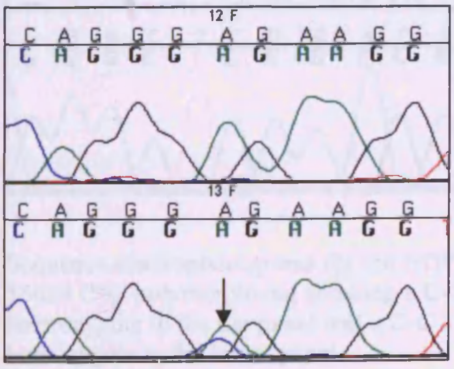
Sequence electrophorograms for the OMG -2289 C>T polymorphism, showing a C-C homozygote in the top panel, a C-T heterozygote in the middle panel and a T-T homozygote in the lower panel.



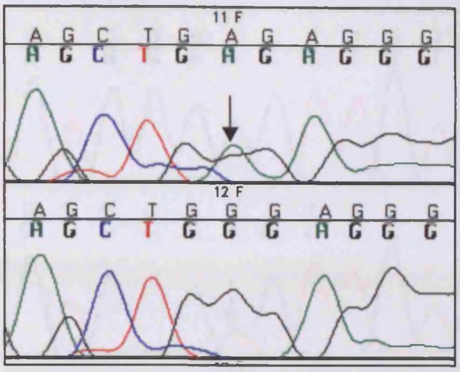
Sequence electrophorograms for the OMG -2081 C>T polymorphism, showing a C-C homozygote in the top panel, a C-T heterozygote in the middle panel and a T-T homozygote in the lower panel.



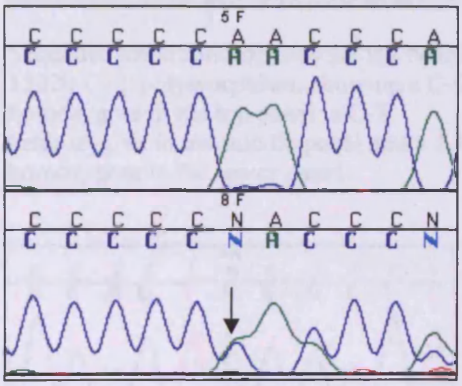
Sequence electrophorograms for the OMG 61 G>A polymorphism, showing a G-G homozygote in the top panel and a G-A heterozygote in the lower panel.



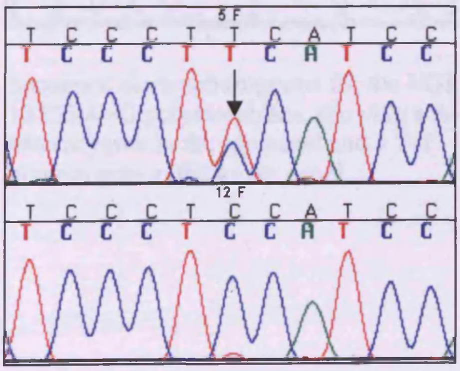
Sequence electrophorograms for the NGFR -486 A>C polymorphism, showing a A-A homozygote in the top panel and a A-C heterozygote in the lower panel.



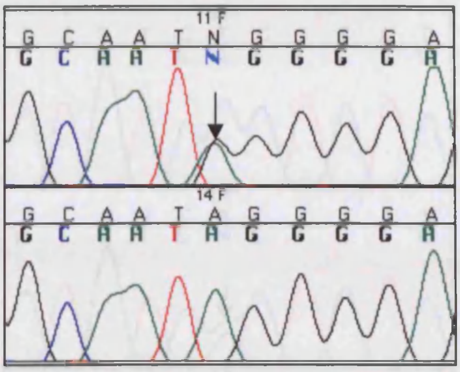
Sequence electrophorograms for the NGFR 331 G>A polymorphism, showing a G-A heterozygote in the top panel and a G-G homozygote in the lower panel.



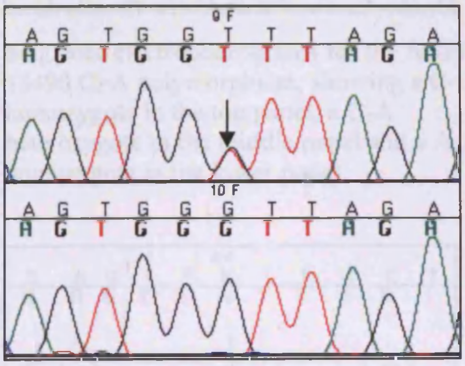
Sequence electrophorograms for the NGFR 6583 C-Ins polymorphism, showing a C-Del homozygote in the top panel and a C-Ins heterozygote in the lower panel.



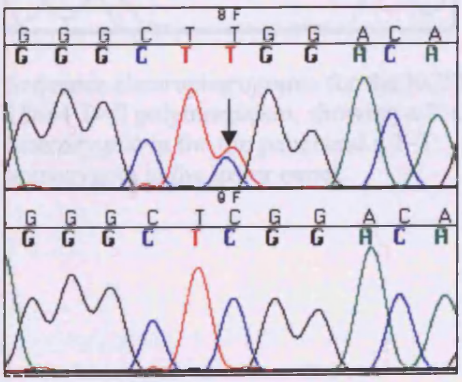
Sequence electrophorograms for the NGFR 6637 C>T polymorphism, showing a C-T heterozygote in the top panel and a C-C homozygote in the lower panel.



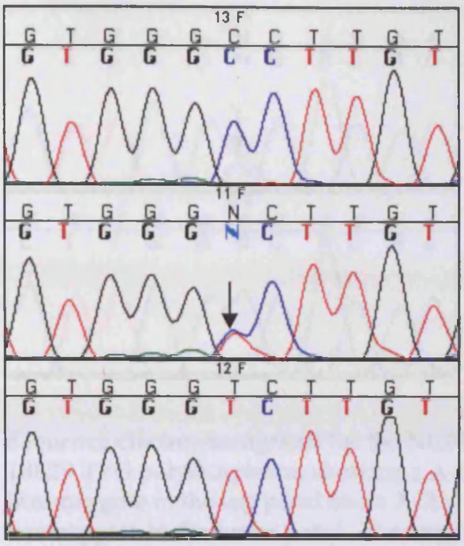
Sequence electrophorograms for the NGFR 10801 A>G polymorphism, showing a A-G heterozygote in the top panel and a A-A homozygote in the lower panel.



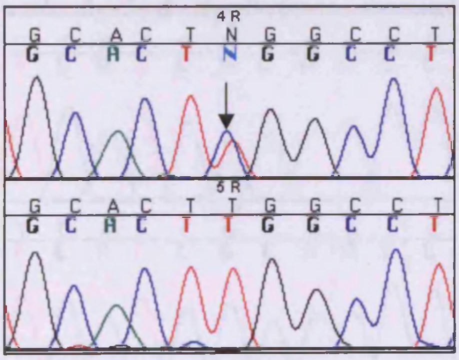
Sequence electrophorograms for the NGFR 14932 G>T polymorphism, showing a G-T heterozygote in the top panel and a G-G homozygote in the lower panel.



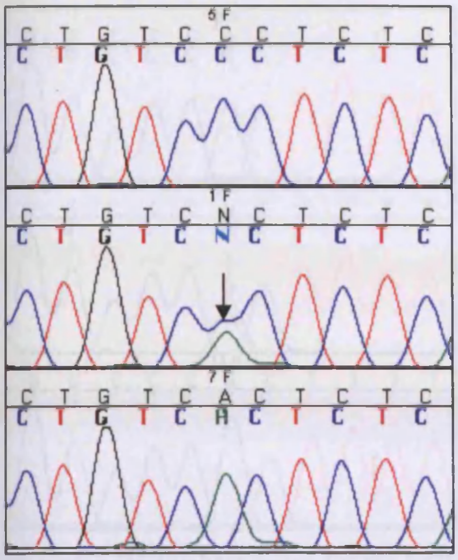
Sequence electrophorograms for the NGFR 15039 C>T polymorphism, showing a C-T heterozygote in the top panel and a C-C homozygote in the lower panel.



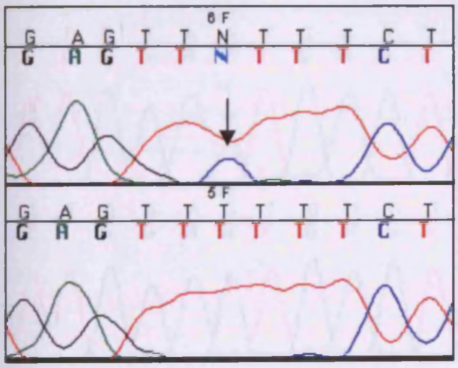
Sequence electrophorograms for the NGFR 15220 C>T polymorphism, showing a C-C homozygote in the top panel, a C-T heterozygote in the middle panel and a T-T homozygote in the lower panel.



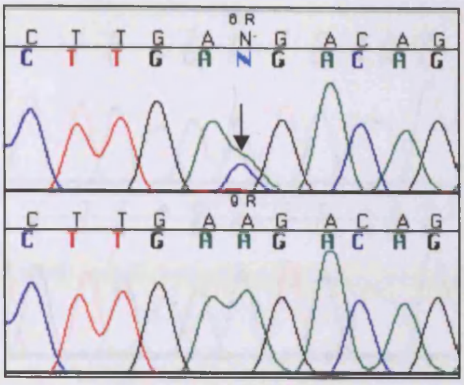
Sequence electrophorograms for the NGFR 18255 A>G polymorphism, showing a A-G heterozygote in the top panel and a T-T homozygote in the lower panel.



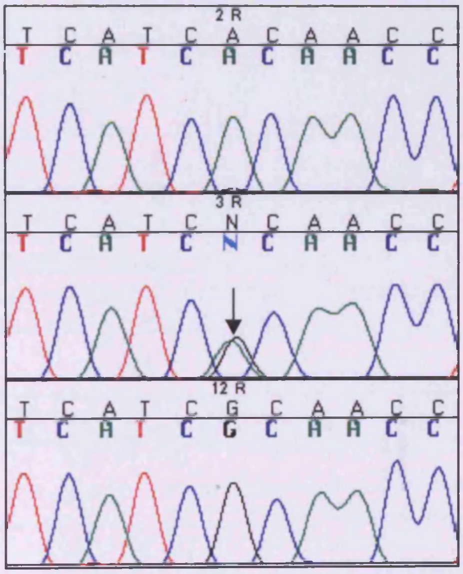
Sequence electrophorograms for the NGFR 18490 C>A polymorphism, showing a C-C homozygote in the top panel, a C-A heterozygote in the middle panel and a A-A homozygote in the lower panel.



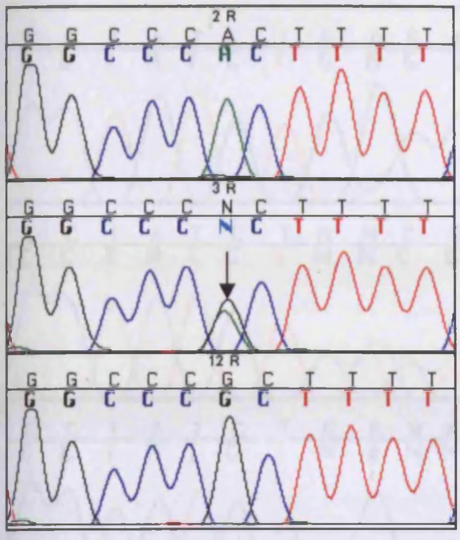
Sequence electrophorograms for the NGFR 18684 T>C polymorphism, showing a T-C heterozygote in the top panel and a T-T homozygote in the lower panel.



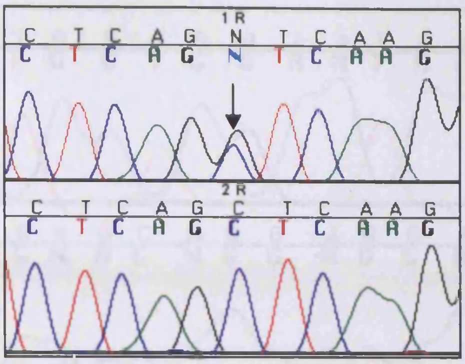
Sequence electrophorograms for the NGFR 18829 T>G polymorphism, showing a A-C heterozygote in the top panel and a A-A homozygote in the lower panel. The sequence is taken from the anti-sense strand.



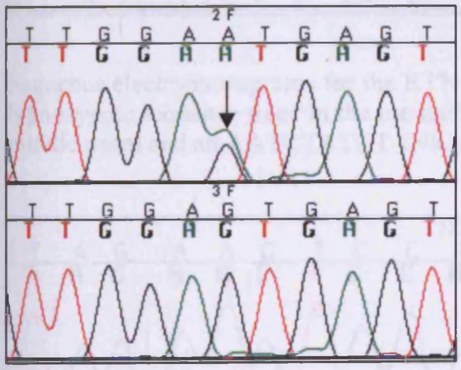
Sequence electrophorograms for the NGFR 18910 T>C polymorphism, showing a A-A homozygote in the top panel, a A-G heterozygote in the middle panel and a G-G homozygote in the lower panel. The sequence is taken from the anti-sense strand.



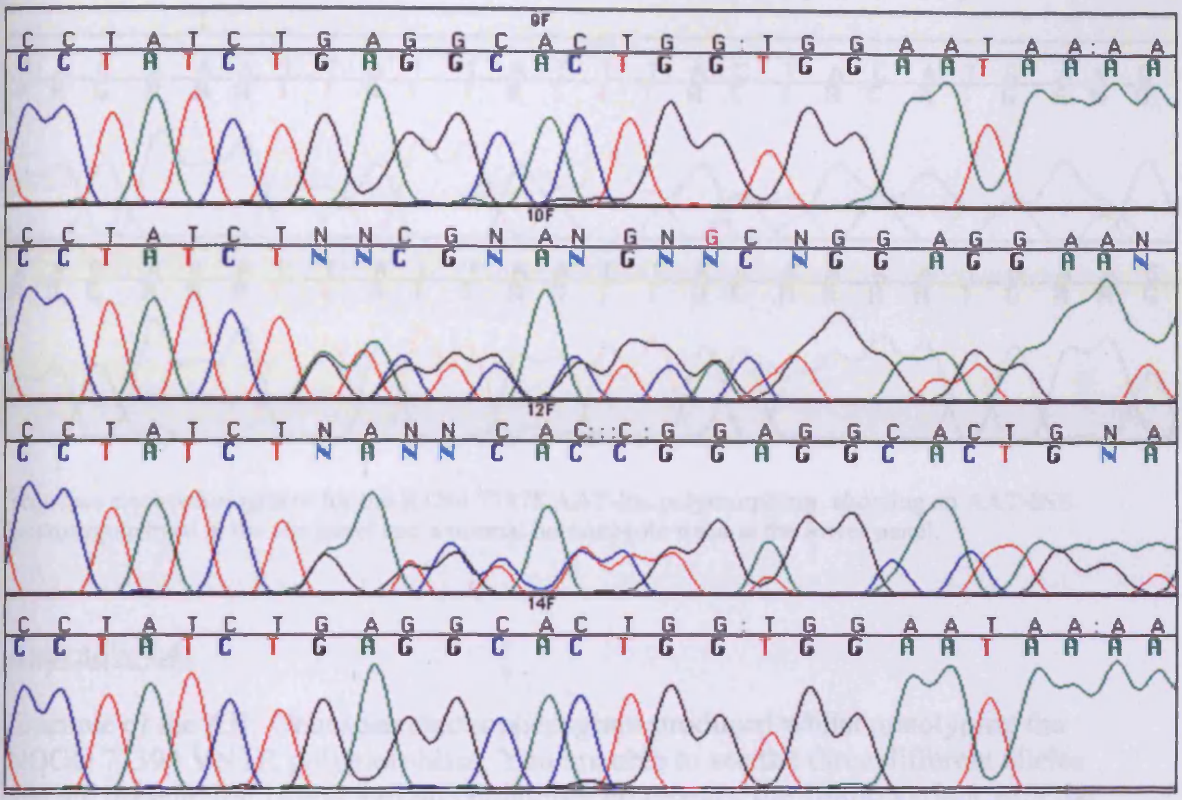
Sequence electrophorograms for the NGFR 18923 T>C polymorphism, showing a A-A homozygote in the top panel, a A-G heterozygote in the middle panel and a G-G homozygote in the lower panel. The sequence is taken from the anti-sense strand.



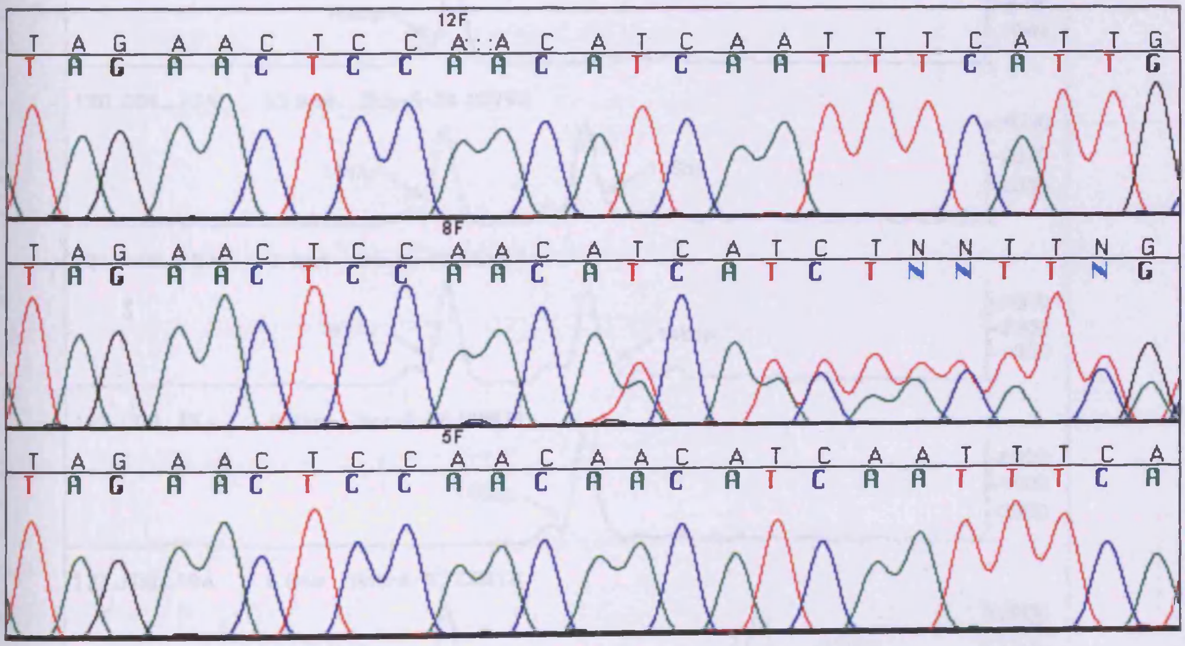
Sequence electrophorograms for the NGFR 19421 G>C polymorphism, showing a G-C heterozygote in the top panel and a C-C homozygote in the lower panel.



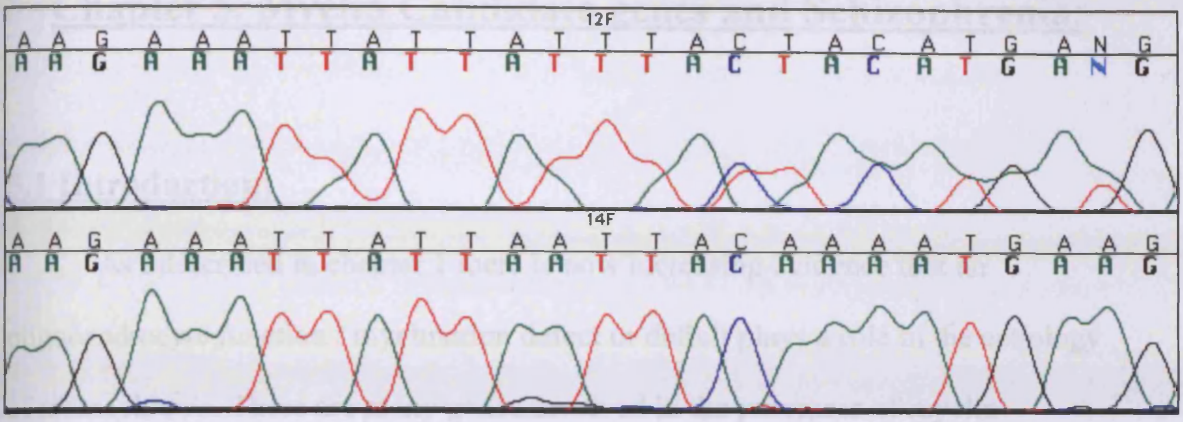
Sequence electrophorograms for the NGFR 19106 G>A polymorphism, showing a G-A heterozygote in the top panel and a G-G homozygote in the lower panel.



Sequence electrophorograms for the RTN4 77394 VNTR polymorphism, showing a normal homozygote sequence trace in the top and lower panels, an ATCT-INS heterozygote in the upper middle panel and an AATCTATCT-INS heterozygote in the lower middle panel.



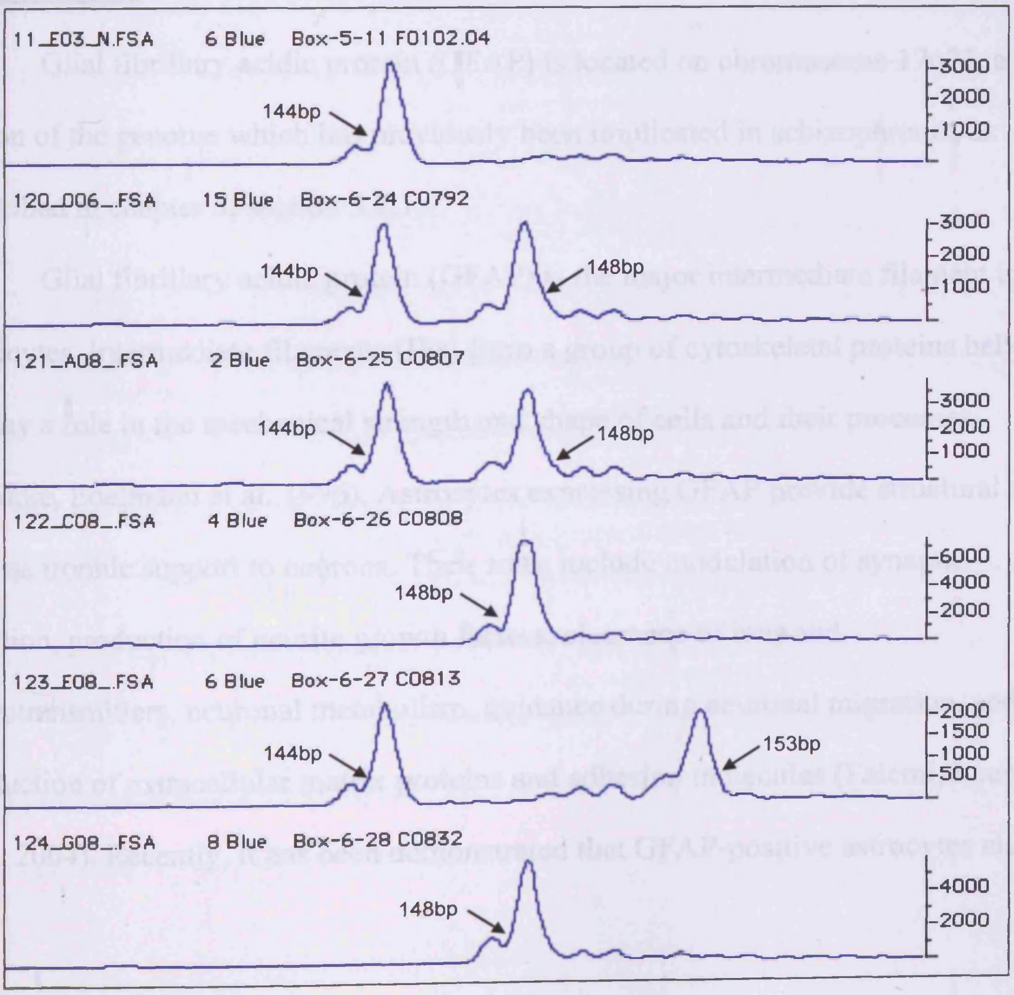
Sequence electrophorograms for the RTN4 77878 AAC-Ins polymorphism, showing a normal homozygote sequence trace in the top panel an AAC-INS heterozygote trace in the middle panel and an AAC-INS homozygote trace in the lower panel.



Sequence electrophorograms for the RTN4 77878 AAT-Ins polymorphism, showing an AAT-INS heterozygote trace in the top panel and a normal homozygote trace in the lower panel.

Appendix 4.4:

Example of the ABI Genescan electrophorogram produced whilst genotyping the NOGO 77394 VNTR polymorphism. You are able to see the three different alleles that are present, the 144bp variant containing no repeats, the 148bp variant with the ATCT insertion and the 153bp variant with the AATCTATCT insertion present.



Chapter 5. Myelin Candidate genes and Schizophrenia:

5.1 Introduction:

As I described in chapter 1 there is now increasing evidence that an oligodendrocyte function / myelination defect or deficit plays a role in the aetiology of schizophrenia. There are many genes involved in the processes of myelin formation, function and maintenance. In this chapter I examine three OMR genes; glial fibrillary acidic protein (GFAP), myelin oligodendrocyte glycoprotein (MOG) and sex-determining region Y-box 10 (SOX 10), selected for study on the basis of reported altered expression in schizophrenia and previous linkage evidence.

5.1.1 GFAP:

Glial fibrillary acidic protein (GFAP) is located on chromosome 17q21, a region of the genome which has previously been implicated in schizophrenia, as described in chapter 3, section 3.2.1.

Glial fibrillary acidic protein (GFAP) is the major intermediate filament in astrocytes. Intermediate filaments (IFs) form a group of cytoskeletal proteins believed to play a role in the mechanical strength and shape of cells and their processes (Liedtke, Edelmann et al. 1996). Astrocytes expressing GFAP provide structural as well as trophic support to neurons. Their roles include modulation of synaptic function, production of neurite growth factors, clearance of ions and neurotransmitters, neuronal metabolism, guidance during neuronal migration, and production of extracellular matrix proteins and adhesion molecules (Fatemi, Laurence et al. 2004). Recently, it has been demonstrated that GFAP-positive astrocytes also serve as primary progenitors of new oligodendrocytes as well as neurons in the adult brain (Menn, Garcia-Verdugo et al. 2006). Astrocytes are also important for the development and maintenance of myelin sheaths and the central nervous system (CNS) (Liedtke, Edelmann et al. 1996; Webster, O'Grady et al. 2005). Astrocytes ensheath the areas of neurons that are not covered by oligodendrocytes, such as the nodes of Ranvier, neuronal cell bodies and synapses (Cotter, Pariante et al. 2001). Also oligodendrocytes share intercellular junctions solely with astrocytes, such that each oligodendrocyte is isolated from other oligodendrocytes except via an astrocyte intermediary (Webster, O'Grady et al. 2005). These astrocyte intermediaries are important as they link oligodendrocytes together allowing the cells to communicate and also exchange ions and small molecules between oligodendrocytes (Webster, O'Grady et al. 2005). Therefore abnormal astrocytic numbers, density and/or function could have profound effects on the viability and activity of oligodendrocytes and ultimately impact the integrity of the neuronal systems (Webster, O'Grady et al. 2005).

Animal models have also suggested that GFAP is necessary for the long-term maintenance of CNS myelination. Liedtke and colleagues (Liedtke, Edelmann et al. 1996) showed that GFAP null mice displayed abnormal myelination. Abnormalities included the presence of actively myelinating oligodendrocytes in adults, nonmyelinated axons in optic nerve, reduced myelin thickness in spinal cord, as well as alterations in the blood-brain barrier, disorganisation of white matter architecture and vascularisation, and hydrocephalus in older mice associated with loss of white matter. The authors therefore concluded that GFAP expression is essential for normal white matter architecture and that its absence led to CNS demyelination (Liedtke, Edelmann et al. 1996).

Several studies have shown a decrease in glial number and GFAP levels in various areas of the brain from individuals with schizophrenia and mood disorder, supporting the previous micro-array results (Tkachev, Mimmack et al. 2003) who showed GFAP expression differences in both schizophrenics and bipolar patients versus controls. Johnston-Wilson and colleagues (Johnston-Wilson, Sims et al. 2000) examining GFAP in astroglia, described a reduction of GFAP protein levels in the frontal cortex of post mortem brain from schizophrenic and mood disorder patients compared to controls. Whilst, Rajkowska and colleagues (Rajkowska, Miguel-Hidalgo et al. 2002) have shown a reduction in the number of GFAP labelled astrocytes in the dorsolateral prefrontal cortex of schizophrenic patients compared to control subjects. In a further study, Fatemi and colleagues (Fatemi, Laurence et al. 2004) showed that GFAP levels were reduced by 14.5% in schizophrenic cerebellum versus controls and more recently Webster and colleagues (Webster, O'Grady et al. 2005) found decreased levels of GFAP in the white matter of post mortem anterior cingulate cortex (ACC) from schizophrenia and bipolar disorder patients compared to controls. This difference was also present in the grey matter but did not reach levels of statistical significance. The anterior cingulate cortex is an area which has been previously implicated in the pathophysiology of schizophrenia, and is involved in the regulation of emotion, attention, cognition, arousal and motor control (Webster, O'Grady et al. 2005).

Mutation in the exonic sequence of GFAP causes Alexander disease, a rare disorder of the CNS (Aoki, Haginoya et al. 2001; Li, Johnson et al. 2005). Forms of the disease exist which differ by age of onset and also in their clinical and pathological presentation. However they are all united by the defining feature of the disease which is the widespread presence of protein aggregates called Rosenthal fibres containing GFAP in astrocytes (Aoki, Haginoya et al. 2001; Li, Johnson et al. 2005). The most common form of the disorder affects infants and young children and is characterised by progressive failure of myelination and an absence of myelin in the frontal lobes, usually leading to death before adulthood. However, the juvenile and adult forms of the disease progress more slowly with myelin and mental ability remaining relatively intact (Aoki, Haginoya et al. 2001; Li, Johnson et al. 2005). Symptoms of psychosis have not been reported in Alexander disease, although the disease is usually fatal before patients reach adolescence (Aoki, Haginoya et al. 2001), the period usually associated with the development of psychosis.

Considering the various lines of evidence discussed above, including the location of GFAP in a region of linkage to schizophrenia, the evidence for altered expression of GFAP in schizophrenia, and the genes involvement in the myelination process, I decided to investigate GFAP.

5.1.2 MOG:

Myelin oligodendrocyte glycoprotein, (MOG) is located on chromosome 6p21.3 close to the human leukocyte antigen (HLA) region. It is a member of the immunoglobulin super family and is expressed on the surface of oligodendrocytes and the lamellae of myelin sheaths in the CNS. The human leukocyte antigen (HLA) system has been examined in several molecular genetic studies of schizophrenia. The results have been variable, with a number of associations reported between different antigens and schizophrenia (Sasaki, Matsushita et al. 1999; Akaho, Matsushita et al. 2000; Nunes, Borelli et al. 2005) although there have also been many studies reporting negative results within the region reviewed by (Wright, Nimgaonkar et al. 2001). There have also been several genetic studies looking for linkage to schizophrenia in this region. Some reported non-significant results, whilst others (Schwab, Albus et al. 1995; Straub, MacLean et al. 1995) showed linkage to schizophrenia both telomeric and centromeric to the HLA region. More recently (Schwab, Hallmayer et al. 2000; Straub, MacLean et al. 2002) have both reported significant linkage to schizophrenia in the HLA region (6p22.3) and have also presented evidence supporting the case for dysbindin (DTNBP1) as a schizophrenia gene in this same region (Straub, Jiang et al. 2002; Schwab, Knapp et al. 2003). These findings are supported by the results of a meta-analysis (Lewis, Levinson et al. 2003) reporting evidence of linkage between 6p21-6p22 and schizophrenia. Put together these results provide evidence for the presence of a susceptibility locus in the region of 6p21-22.

MOG has been proposed to play several roles in the CNS. These roles include; controlling and maintaining myelination, modulating cell to cell interactions, acting as a cellular adhesion molecule between bundles of nerve fibres, regulating oligodendrocyte microtubule stability, and/or mediating interactions between myelin and the immune system (Johns and Bernard 1999). However the precise function of MOG in the CNS still remains to be determined (Johns and Bernard 1999). MOG itself is only a minor component of CNS myelin representing around 0.05-0.1% of the total myelin protein (Johns and Bernard 1999). MOG is specifically expressed on the outermost lamellae of the myelin sheath as well as the cell body and processes of oligodendrocytes. Expression of MOG is developmentally late and correlates with the later stages of myelinogenesis (Clements, Reid et al. 2003). It has been suggested that MOG has a role in the completion, compaction, and/or maintenance of myelin, evidence that further suggests a role for MOG as an adhesion molecule in the CNS (Clements, Reid et al. 2003).

Sites near the HLA locus have been implicated in other neuropsychiatric disorders, such as narcolepsy and obsessive compulsive disorder (OCD) and the demyelinating disease MS (Wright, Nimgaonkar et al. 2001). One study (Zai, Bezchlibnyk et al. 2004) reported an association between OCD and MOG. The authors reported both association to an individual SNP and a slightly more significant haplotypic association. Multiple sclerosis (MS) is an inflammatory disease of the CNS characterised by localised myelin destruction and axonal degeneration (Clements, Reid et al. 2003). Antibody induced demyelination is an important component of pathology in multiple sclerosis. Antibodies are directed against surface antigens of myelin and oligodendrocytes in the CNS, in particular antibodies to MOG, which is a highly encephalitogenic glycoprotein concentrated in the outer lamella of the myelin sheath and hence exposed to the environment. Antibodies to MOG have been shown to be elevated in MS patients, and they have therefore been implicated as mediators of MS demyelination (Marta, Montano et al. 2005).

Gene expression studies have reported a decrease in MOG mRNA expression in post mortem brain from schizophrenic and bipolar patients compared to controls. Tkachev and colleagues (Tkachev, Mimmack et al. 2003) reported a 2.58 fold decrease of MOG expression in schizophrenics, a finding supported by the results of an independent study (Katsel, Davis et al. 2005b) reporting a 2.05 fold down-regulation of MOG in schizophrenics. In a further study (Aston, Jiang et al. 2005) reported a 3.2 fold decrease of expression in bipolar patients compared to controls.

Previous genetic studies of MOG and schizophrenia have yielded mixed results. Liu and colleagues (Liu, Qin et al. 2005) screened MOG using microsatellites in 532 trios and found weak association (p=0.02) between microsatellite D6S299 and schizophrenia, although the result would not withstand correction for multiple testing.

A three marker haplotype gave a slightly more significant result (p=0.01). However, (Zai, King et al. 2005) studied four polymorphisms in MOG for association with schizophrenia, using 111 schizophrenia probands and their families. They did not observe significant evidence for biased transmission of alleles at any of the polymorphisms. Haplotype analysis was also not significant. They also used a second sample of 182 cases and matched controls which again did not show significant association for any of the polymorphisms or haplotypes tested.

The MOG gene can be considered as a possible candidate for schizophrenia for several reasons a) its location in an established linkage region for schizophrenia, b) its association with white matter abnormalities, c) its altered expression in schizophrenic patients, d) its involvement in diseases with phenotypic overlaps with schizophrenia such as multiple sclerosis and e) the modest evidence from previous studies which show putative evidence of association to schizophrenia.

5.1.3 SOX 10:

Sex-determining region Y-box 10 (SOX 10) is located on chromosome 22q13.1, which is a putative region of linkage to schizophrenia (Owen, Craddock et al. 2005). There have been several studies suggesting that the 22q11-13 region on chromosome 22 is linked to schizophrenia (Gill, Vallada et al. 1996; DeLisi, Shaw et al. 2002; Williams, Norton et al. 2003), and two meta-analyses of the linkage data have provided further support for the involvement of this region (Badner and Gershon 2002; Lewis, Levinson et al. 2003). Sox 10 is located 18 Megabases downstream from the VCFS region, VCFS being a syndrome in which there is an increased risk of developing schizophrenia. It has been shown that about a third of VCFS cases suffer schizophrenia like psychosis (Murphy, Jones et al. 1999; Murphy 2002), a figure which is much higher than the 1% prevalence rate of schizophrenia in the general population.

Sox 10 belongs to the SRY box family of transcription factors, it contains a high mobility group (HMG) DNA binding domain and a trans-activation domain at its C-terminus. The HMG domain binds to and bends the minor groove of DNA, causing a conformational change, which in turn facilitates cooperative binding of transcriptional factors (Wei, Miskimins et al. 2004). Sox 10 is involved in development of neural crest derivatives and fate determination of glial cells in the neural crest lineage (Inoue, Khajavi et al. 2004). It also controls the proliferation and differentiation of Schwann cells and oligodendrocytes (Inoue, Khajavi et al. 2004). Glial cells of the peripheral nervous system (PNS) are generated by neural crest cells. These glia include satellite cells located in ganglia as well as Schwann cells which ensheath peripheral axons (Britsch, Goerich et al. 2001). Without Sox 10, no glia develop in the peripheral nervous system (Stolt, Rehberg et al. 2002). Sox 10 is also strongly expressed in the CNS in regions such as the corpus callosum and cerebral and cerebellar cortex (Stolt, Rehberg et al. 2002). The onset of SOX 10 expression in the CNS coincides with the onset of oligodendrocyte development (Stolt, Schmitt et al. 2005). Unlike other Sox proteins, Sox 10 continues to be expressed and is involved in activating myelin gene expression as part of the terminal differentiation program in oligodendrocytes (Stolt, Schmitt et al. 2005). Sox 10 is thought to stimulate the transcription of myelin related genes in both the peripheral and central nervous system. These genes include protein zero in Schwann cells of the PNS and myelin basic protein (MBP) in the CNS (Wei, Miskimins et al. 2004). Protein zero makes up more than 50% of the total myelin protein in mature Schwann cells, where it is involved in myelin compaction (Peirano, Goerich et al. 2000).

Several studies using DNA micro-arrays to study gene expression in psychiatric diseases have reported down regulation of Sox 10 in post-mortem schizophrenic brain. Tkachev and colleagues (Tkachev, Mimmack et al. 2003) reported a 1.88-1.99 fold down regulation of SOX 10 in schizophrenia, a result supported by an independent micro-array study (Katsel, Davis et al. 2005b). Aston and colleagues (Aston, Jiang et al. 2005) reported a 1.5 fold decrease in expression in bipolar patients compared to controls. Recently, following up on these results, (Iwamoto, Bundo et al. 2005b) showed that patients with schizophrenia had a higher percentage of methylated Sox 10 and lower expression levels of Sox 10 compared with control subjects. There was an inverse correlation between the percentage of methylated Sox 10 and the expression level, suggesting that downregulation of Sox 10 was possibly caused by Sox 10 methylation. The authors also showed that methylation status of Sox 10 was correlated with down regulation of other oligodendrocyte genes that were differentially expressed in schizophrenia, including MAG, OLIG 2, MOBP and CNP. They subsequently concluded that the expression in the genes is under the direct or indirect control of Sox 10.

Studies on patients with heterozygous Sox 10 mutations have identified an association between SOX 10 mutations and deficiencies in peripheral and/or central nervous system myelin (Britsch, Goerich et al. 2001). Mutations in SOX 10 have been shown to cause neurological diseases including peripheral demyelinating neuropathy, central demyelinating leukodystrophy, Waardenburg syndrome, and Hirschsprung disease (Iwamoto, Bundo et al. 2005b). Leukodystrophies are a heterogeneous group of genetic diseases that affect primarily the central nervous system myelin. They typically involve the motor systems with prominent CNS symptoms and slow mental deterioration (Pingault, Bondurand et al. 2001). Some are degenerative (demyelinating diseases), whereas others are due to a developmental defect in CNS myelin formation (dysmyelinating diseases) (Pingault, Bondurand et al. 2001).

This interpretation of the data from Sox 10 deficient humans is consistent with the findings of animal studies that have shown Sox 10 null mice have profound defects in both peripheral and central myelin (Britsch, Goerich et al. 2001).

In the peripheral nervous system, expression of erbB3 is thought to be influenced by Sox 10 (Britsch, Goerich et al. 2001), although the presence of Sox 10 binding sites within the erbB3 gene has not been reported and it is therefore not clear whether Sox 10 regulation of this gene is direct (Mollaaghababa and Pavan 2003). ErbB3 is a receptor kinase that binds neuregulins. ErbB3 like Sox 10 has a critical function in the development of neural crest derivatives including Schwann cells and oligodendrocytes (Mollaaghababa and Pavan 2003). The involvement of Sox 10 in the regulation of erbB3 is made more interesting by reports implicating one of the erbB3 ligands, Neuregulin 1 (NRG1) as a susceptibility gene for schizophrenia in several studies (Stefansson, Sigurdsson et al. 2002; Stefansson, Sarginson et al. 2003; Owen, Craddock et al. 2005; Li, Collier et al. 2006), and reports of altered erbB3 expression in schizophrenia (Hakak, Walker et al. 2001; Tkachev, Mimmack et al. 2003; Aston, Jiang et al. 2004; Katsel, Davis et al. 2005b).

5.2 Methods:

Analysis of the three genes was based upon a direct association approach as described in Chapter 1.

PCR fragments to span exonic sequence, putative promoter and limited amounts of 5' and 3' flanking regions were designed and the primers used to generate amplimers for the purpose of DHPLC for each gene studied are listed below in tables 5.1, 5.2, and 5.3 respectively.

Table 5.1: PCR primers and conditions for GFAP assays.

The table shows the fragment name, region of genomic sequence amplified, the Forward and Reverse primer sequences, the size of the fragment and the reaction conditions used to amplify all of the GFAP fragments utilised during the assaying of the gene.

Fragment Name	Spanning	Forward Primer (5' to 3')	Reverse Primer (5' to 3')	Size (bp)	Annealing Temp (°C)
GFAP_1	5' Flank / Put Prom	GGCATCTTGGGATAAAAGCA	AGGTCACCCCTTCCTTGTCT	548bp	HS 64
GFAP_2	5' Flank / Put Prom	CATGTGTAACGGCTTTGTGG	AAAGGGAAGGAAGGGCTGTA	483bp	HS 64
GFAP_3	5' Flank / Put Prom	GCCCTTGAGGGTACAGAACA	AGGCAGGTGGTACCTGGAG	461bp	HS 64
GFAP_4	Exon 1	ACCCTTTCTTGACCCACCTT	GCGAACCTTCTCGATGTAGC	569bp	HS 64
GFAP_5	Exon 1	AATGCTGGCTTCAAGGAGAC	CCAGCCATGAATGAAACACA	584bp	HS 64
GFAP_6	Exon 2 / Exon 3	TTAAGGCAAAAGCCCACATC	CCAGCGACTCAATCTTCCTC	518bp	HS 64
GFAP_7	Exon 3	TAAGGCAAGGCTGAGGAATG	GTACCAGGGCTCAGGGTACA	451bp	HS 64
GFAP_8	Exon 4	TTGGCAGGTAGAGACAGGAA	GGCAAGCCATCTCACTTCTC	485bp	HS 60
GFAP_9	Exon 5	TCATTCATTCGCTTGCTGAC	CTCTCCAGGGACTCGTTCTG	571bp	HS 60
GFAP_10 Redesign	Exon 6	CTCGAAACCTCCCCAAGTC	TGACTGCCTGCTATGTGTGA	439bp	HS 60
GFAP_11	Exon 7	TTTTTCAGCCTTGTGTGCAG	TCTGCAGGTTGGAGAAGGTC	533bp	HS 60
GFAP_12	Exon 7	AAGGAAAAGGGCGTTCAAGT	ACACTCAGAAGGGCAGTGCT	463bp	HS 60
GFAP_13	Exon 8	GCAGCAGCACTATTGGGATT	AGGGACTGAGGAAACGGAAT	448bp	HS 64
GFAP_14	Exon 8	GTGGTGAAGACCGTGGAGAT	CAGAGGGGGTGGTGAGATAA	557bp	HS 64
GFAP_15	Exon 9	AACCACAGTGGGAGCTCATC	TTGGGGGTGAGTTTCTTGTT	525bp	HS 60
GFAP_16 Redesign	Exon 9	ACCTGAGCTCCCCACCAC	CCCCACCCATCTTAGACTGA	594bp	HS 60
GFAP_17	Exon 9	CTGGGCTCAAGCAGTCTACC	CACCCACAACCCCTACTTGT	466bp	HS 64
GFAP_18	Exon 9	TGTGTACGTGGAGAGGGATG	ACTTTGCTCGTGCCTCAGTT	545bp	HS 60
GFAP_19	Exon 9	TGCTGCTTCTTAACCCCAGT	GCAGTGCCCTGAAGATTAGC	548bp	HS 60
GFAP_20	Exon 9	CTTTCCCTCCACTCCTTCCT	GGTCTGGGACCTGATCTGAG	522bp	HS 64
GFAP_21	Exon 9	GTGCCCTACTTCCATCCTGA	CCTGAAGGGTGTCAACAGGT	536bp	HS 60
GFAP_22	Exon 9	GTGCCCATGAGATGTCACAC	CCTCCATGTCTTCACCACCT	598bp	HS 62 DMSO

Table 5.2: PCR primers and conditions for the MOG assays.
The table shows the fragment name, region of genomic sequence amplified, the Forward and Reverse primer sequences, the size of the fragment and the reaction conditions used to amplify all of the MOG fragments utilised during the assaying of the gene.

Fragment Name	Spanning	Forward Primer (5' to 3')	Reverse Primer (5' to 3')	Size (bp)	Annealing Temperature (°C)
MOG_1	5' Flank / Put Prom	CAGCTCTGCCACTCCCTAGT	TTGGGGGTGGTAAGATGTGT	364bp	HS 62
MOG_2	5' Flank / Put Prom	TTTTGACAGCTTGCTTGGTG	GCAGCAAGGGGTAGAAACAC	402bp	HS 62
MOG_3	5' Flank / Put Prom	TCTAGCTCCTTACGCCCATC	GACCAGCATTGGTAGCAGGT	377bp	HS 62
MOG_4	5' Flank / Put Prom	ATTTGGTGAGGCATACTTGC	TCTACCGCGAATCCAAACTC	404bp	HS 62
MOG_5	Exon 1	TTGCCTCTCTGAGCACACC	TCCTGGAGATCCAACCTTTG	393bp	HS 62
MOG_6	Exon 1	CCCACTGATTAGCCCCTAGA	CCCCAGGAAACTAGCCTTTC	505bp	HS 62
MOG_7	Exon 2	TTTGAAGCAGCCCTTCTCAT	TTGGCCACCCAACAGTTAAT	477bp	HS 62
MOG_8	Exon 3	GTGTGTCATGCACCTTCCAC	CAATGTGAGACCCTGCCTCT	255bp	HS 62
MOG_9	Exon 4	ATGCTGAGTGTTGGGGATTC	TCCAAGAAGCCAGCTCATTT	231bp	HS 62
MOG_10	Exon 5 / Exon 6	TTTCCAGGCTGCAGAGAAAT	GAGAGGAAGGGACTCAACCA	505bp	HS 62
MOG_11	Exon 7	GGAGGCAGTACTTTTCTCATCC	TCTTTTGGCTTCCCATTCAG	353bp	HS 62
MOG_12	Exon 7 / Exon 8	GGTTGCCTTGATCATCTGCT	TCCTCCTTTCCTCTGGGAAT	486bp	HS 62
MOG_13	Exon 8	AAAAGCCAGGAAAGGGAGAC	CCCATAGCAAAGATGCCAGT	538bp	HS 62
MOG_14	Exon 8	CTTGGGGACATCTCATCCAT	GTAGGCTGCCAAGACAAGGA	542bp	HS 62
MOG_15	Exon 8	AAACCTGGAGATGGGGAGAA	GCCCACCACTATGCTCAGTT	444bp	HS 62
MOG_16 redesign	Exon 8	CTTGATGGTCAGACACAAACA	AATGGAAGCATGGAGGTGAT	684bp	HS 62
MOG_17	Exon 8	CTCCCTCTACTCCAGCCTCA	CCATTGCAGCAAGTCTTTCA	451bp	HS 62
MOG_18	Exon 8	GGTCACAGATAGGGCAAAGG	ACGCCACCAGAAGATACACC	515bp	HS 62
MOG_19	Exon 8	GTGCCTGGAATCCTCAAAGT	GGCAAGAAGAAAGAGCTTCG	545bp	HS 62

Table 5.3: PCR primers and conditions for the SOX 10 assays.

The table shows the fragment name, region of genomic sequence amplified, the Forward and Reverse primer sequences, the size of the fragment and the reaction conditions used to amplify all of the SOX 10 fragments utilised during the assaying of the gene.

Fragment Name	Spanning	Forward Primer (5' to 3')	Reverse Primer (5' to 3')	Size (bp)	Annealing Temperature (°C)
SOX 10 Prom 1	Putative Promoter	AGGAGAGGCCAAGGAAAGAG	CCCTCTGAGAAGTGCAGAGC	402bp	HS 60
SOX 10 Prom 2	Putative Promoter	GTCTGCAGAGAAGGGACCAG	CTTCATCCCCAACACTCAGG	499bp	HS 64
SOX 10 Ex 1	Exon 1	CCCAAGATAAGCCTCAGCAA	AACGGGGTTTAGAGGAGAGC	532bp	HS 60
SOX 10 Ex 2A	Exon 2	GATGGGTTTAGCTGGAGCAG	ACGGGGAACTTGTCATCGT	409bp	HS 60 DMSO
SOX 10 Ex 2B	Exon 2	AAGGTCAAGAAGGAGCAGCA	ACCTGGTCTTCCAGCCCTAT	408bp	HS 60
SOX 10 Ex 3	Exon 3	TCCAAGATGGACACTCAGAGG	CATCCAGCCATCTCCTGTCT	430bp	HS 60
SOX 10 Ex 4A	Exon 4	CACCTGCCTCTAACCTGCTT	TCTGTCTTCACCTGGGCTTT	438bp	HS 60
SOX 10 Ex 4B	Exon 4	GACACTCCGCCTGGATCTC	CAAGGAACAGGGCACACAG	481bp	HS 67
SOX 10 Ex 4C	Exon 4	CTACACGGCCATCTCTGACC	CTAGCCCTTAGGCACTGTGG	551bp	HS 60
SOX 10 Ex 4D	Exon 4	GGAGGCCTTACCACTCCTATG	GCTGTCACTCTCTGGGAAGC	515bp	HS 60
SOX 10 Ex 4E	Exon 4	CCAGCTCCCTCTTCCCTTAC	GAAGAGACAGGCCAGGTCAG	425bp	HS 60

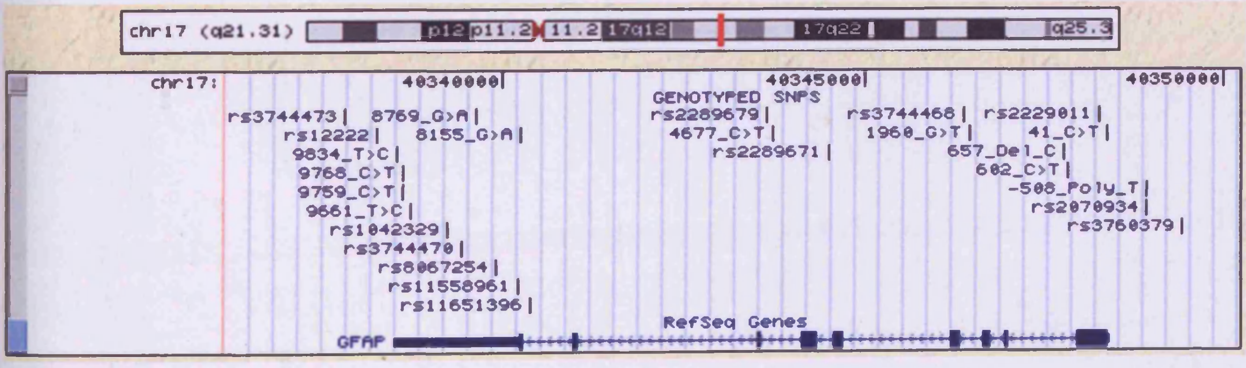
5.3 Results:

5.3.1 GFAP:

The GFAP intron/exon boundaries were defined by alignment of the cDNA sequence NM_002055 and the genomic sequence (http://genome.ucsc.edu). Mutation detection assays revealed twenty three SNPs, an insertion/deletion, and a poly T simple sequence repeat (See appendix 5.3 for sequencing results). There were two coding synonymous SNPs Y14Y (41 C>T) and P140P (rs2229011), ten 3'-UTR SNPs (8155 G>A, rs11558961, rs8067254, 8769 G>A, rs3744470, rs1042329, 9661 T>C, 9759 C>T, 9768 C>T and 9834 T>C), three 5'-flanking region polymorphisms (rs3760379, rs2070934 and -508 poly T) two 3'-flanking region SNPs (rs12222 and rs3744473) and eight intronic polymorphisms (602 C>T, 657 Del C, 1960 G>T, rs3744468, rs2289671, 4677 C>T, rs2289679 and rs11651396) (figure 5.1).

Figure 5.1. Schematic diagram of GFAP.

Diagram of the GFAP gene (adapted from the UCSC Genome Browser), indicating the positions of the identified polymorphisms. For numbering of polymorphisms, +1 indicates the 'A' of the ATG translation codon. The gene is displayed (3'-5').



Twenty one of the SNPs and the insertion/deletion were successfully genotyped in the pooled association sample. For two of the SNPs the minor allele frequencies were too rare to be assayed, whilst the poly-T sequence was incompatible with a primer extension genotyping assay. The extension primers used for pooled genotyping are listed in table 5.4. The calculated allele counts and frequencies for all of the genotyped polymorphisms are shown in table 5.5. Markers 140 G>A, 1960 G>T and 9834 T>C showed evidence for association in pools. However, all have a low minor allele frequency making the pooled genotyping assay more prone to inaccuracy and error. Given the possible inaccuracy in the pooled genotyping results I decided to type these three SNPs in 96 unrelated control subjects and 30 CEPH trios to a) confirm that the estimated pooled allele frequencies were approximately correct before continuing with individual genotyping of the whole association sample, b) determine the LD relationships for marker redundancy.

Table 5.4: Polymorphism and primer extension reaction details for GFAP.

Table of the polymorphisms identified in the GFAP gene, including the sequence flanking each of the polymorphisms, the extension primer used to assay the polymorphism by SNaPshot and the expected alleles in the primer extension reaction.

SNP number.	rs Number	SNP Flanking sequence	Extension primer	Expected Alleles
-1001 T>C	rs3760379	AGCCAGCTCATG[T/C]GTAACGGCTTTG	CCTTGACAGCTCCACAAAGCCGTTAC	A (Green) & G (Blue)
-518 T>A	rs2070934	TTCCCTTTTTTT[T/A]TTTTTTTTTTGT	GTACAGCCCTTCCTTCCCTTTTTTT	T (Red) & A (Green)
-508 Poly T	N/A	CCCTTTTTTTTTTTTTTTTTT[T/-]GTGAGACAAG	Unable to assay by primer extension.	
41 C>T	N/A	TCGCCGCTCCTA[C/T]GTCTCCTCAGGG	TCACCTCCGCTGCTCGCCGCTCCTA	C (Black) & T (Red)
140 G>A	rs2229011	CCCTCCACTCCC[G/A]ACCCGGGTGGAT	GCTCGAATGCCCCCTCCACTCCC	G (Blue) & A (Green)
602 C>T	N/A	CACTATCCTTGC[C/T]CCCCAACATGAT	CCCTACCTCTTGATGGGGCACTATCCTTGC	C (Black) & T (Red)
657 Del-C	N/A	GGGGATCTGATG[C/-]CCGCATGCCTTC	GACCCTGGACTCCTGGCAGAAGGCATGCGG	G (Blue) & C (Black)
1960 G>T	N/A	ATTCGAACTCCT[G/T]GCCGAGGGGCAG	GGAAAGAGACAAGAGAGCATTCGAACTCCT	G (Blue) & C (Black)
2044 C>T	rs3744468	TACCGCTTCTCT[C/T]ACCCTGGTCAGG	AGCCCTGGTACCGCTTCTCT	C (Black) & T (Red)
3895 A>G	rs2289671	GAAGGGCCAGGG[A/G]CGAGGACCTGGG	GCCCCGCCCTCGACCCAGGTCCTCG	T (Red) & C (Black)
4677 C>T	N/A	CTCCAGATAGCA[C/T]GTTGGCTGCTGG	CCTTGCAGACGTTTGGAACTCCAGATAGCA	C (Black) & T (Red)
4762 C>G	rs2289679	ATGATGGCCCCG[C/G]ACCAAGACCCAG	CGTTCAAGTTGTCATGATGGCCCCG	C (Black) & G (Blue)
8012 C>T	rs11651396	TCGGGGACTTCC[C/T]AAGGTTACACAG	CACCACCACGAGGCTGTGTAACCTT	G (Blue) & A (Green)
8155 G>A	N/A	AGGACCCACCTG[G/A]TGGCCTCTGCCC	AAGGATGTGATGTGAGGCAGGACCCACCTG	G (Blue) & A (Green)
8167 C>G	rs11558961	GTGGCCTCTGCC[C/G]CGTCTCATGAGG	CCTGCTTCTGCTCGGGCCCCTCATGAGACG	G (Blue) & C (Black)
8477 C>T	rs8067254	CATCGCCCTCCA[C/T]ATCCCTGATTCC	GCACCTACTACATCGCCCTCCA	C (Black) & T (Red)
8769 G>A	N/A	TGGAGCCACCAT[G/A]CCCAGCTCAGAG	GTAGCTGGGATTATAGATTGGAGCCACCAT	G (Blue) & A (Green)
8953 T>A	rs3744470	GCTGTGGGCAGG[T/A]GGGAGCTTGATT	GATCCCCCAAGTGCTGAGAATCAAGCTCCC	A (Green) & T (Red)
9140 G>A	rs1042329	AATTGGGATGGC[G/A]CTAGGCATACAA	CACCCACAACCCCTACTTGTATGCCTAG	C (Black) & T (Red)
9661 T>C	N/A	TGTGCCCAGAGA[T/C]GGGACTGGGAGG	GAGCAGACTGGTGGGATCTGTGCCCAGAGA	T (Red) & C (Black)
9759 C>T	N/A	TTTCCCTCCACT[C/T]CTTCCTGCCACC	CCAAGACTTGGTGTCCTTTCCCTCCACT	C (Black) & T (Red)
9768 C>T	N/A	ACTCCTTCCTGC[C/T]ACCTGCTGCTGC	CCTGAAGATTAGCAGCAGCAGCAGCAGCAGGT	G (Blue) & A (Green)
9834 T>C	N/A	GCTGAGGAAAAA[T/C]AAAGACAAATGC	CCTTTAGTCGCTGAGGAAAAA	T (Red) & C (Black)
10125 C>T	rs12222	CAGTGTTTACCA[C/T]TGGATAAATCTG	GAAAGTGTGAACTCAGATTTATCCA	G (Blue) & A (Green)
10566 T>C	rs3744473	AGGCTGGGTAG[T/C]GCCCATGAGATG	AGGATCTGAGAACGAGGCTGGGTAG	T (Red) & C (Black)

Table 5.5: Pooled genotyping data for GFAP.
Estimated minor allele frequencies and minor allele counts for the GFAP polymorphisms, with the corresponding chi-square and p-values. Significant p-values are displayed in bold.

SNP Data		Pooled Genotyping data				
SNP number.	rs Number	Minor Allele Cases (freq)	Minor Allele Controls (freq)	Difference	Chi-square	P-value
-1001 T>C	rs3760379	T= 465 (0.43)	T= 440 (0.4)	0.03	2.90	0.09
-518 T>A	rs2070934	A= 382 (0.36)	A= 391 (0.35)	0.01	0.02	0.89
-508 Poly T		Incompatible with a primer extension assay.				
41 C>T		T= 18 (0.02)	T= 17 (0.02)	0	0.07	0.79
140 G>A	rs2229011	A= 70 (0.07)	A= 49 (0.04)	0.03	4.65	**0.03**
602 C>T		T= 174 (0.16)	T= 176 (0.16)	0	0.04	0.84
657 Del-C		C-Ins= 178 (0.17)	C-Ins= 171 (0.16)	0.01	0.53	0.47
1960 G>T		T= 90 (0.08)	T= 53 (0.05)	0.03	11.53	**0.001**
2044 C>T	rs3744468	T= 119 (0.11)	T= 109 (0.1)	0.01	0.90	0.34
3895 A>G	rs2289671	G= 472 (0.44)	G= 493 (0.45)	0.01	0.07	0.79
4677 C>T		T= 145 (0.14)	T= 137 (0.12)	0.02	0.63	0.43
4762 C>G	rs2289679	G= 319 (0.3)	G= 350 (0.32)	0.02	0.91	0.34
8012 C>T	rs11651396	T= 301 (0.28)	T= 318 (0.29)	0.01	0.12	0.73
8155 G>A		A= 41 (0.04)	A= 52 (0.05)	0.01	1.02	0.31
8167 C>G	rs11558961	G= 306 (0.29)	G= 326 (0.3)	0.01	0.23	0.63
8477 C>T	rs8067254	T= 149 (0.14)	T= 147 (0.13)	0.01	0.17	0.68
8769 G>A		Minor allele frequency too rare to assay.				
8953 T>A	rs3744470	A= 130 (0.12)	A= 156 (0.14)	0.02	1.87	0.17
9140 G>A	rs1042329	A= 125 (0.12)	A= 115 (0.1)	0.02	0.89	0.35
9661 T>C		C= 307 (0.29)	C= 347 (0.31)	0.02	1.94	0.16
9759 C>T		T= 116 (0.11)	T= 103 (0.09)	0.02	1.37	0.24
9768 C>T		Minor allele frequency too rare to assay.				
9834 T>C		C= 30 (0.03)	C= 75 (0.07)	0.04	18.82	**0.001**
10125 C>T	rs12222	T= 324 (0.3)	T= 350 (0.32)	0.02	0.51	0.47
10566 T>C	rs3744473	C= 307 (0.29)	C= 336 (0.3)	0.01	0.79	0.37

Individual genotyping revealed the minor allele frequency for SNPs 140 G>A and 1960 G>T were consistent with the estimated allele frequency of the controls in the pooled genotyping assay (table 5.6). However, for 9834 T>C the minor allele frequency in the 96 controls and 30 trios was lower than in the pooling assay, suggesting that the polymorphism was not assayed accurately by pooled genotyping. LD analysis of the markers revealed the three polymorphisms were in D'=1 with each other but only had a value of r^2=<0.01.

Table 5.6 Control genotyping of 140 G>A, 1960 G>T and 9834 T>C.

Genotype and allele counts and frequencies for the GFAP SNPs 140 G>A, 1960 G>T and 9834 T>C after genotyping in 96 controls and 30 trios.

SNP	Genotypes			Alleles (freq)	
140 G>A	**G/G** 164	**G/A** 15	**A/A** 0	**G** 343 (0.96)	**A** 15 (0.04)
1960 G>T	**G/G** 153	**G/T** 9	**T/T** 0	**G** 315 (0.97)	**T** 9 (0.03)
9834 T>C	**T/T** 172	**T/C** 1	**C/C** 0	**T** 345 (0.995)	**C** 1 (0.005)

I therefore decided to genotype the two 140 G>A and 1960 G>T polymorphisms individually. This was achieved by performing two Amplifluor genotyping assays using the protocol as described in section 2.10.3. (See appendix 5.4 for amplifluor assay design).

The genotyping results are shown (tables 5.7 and 5.8). No significant differences were found in the cases and controls. The failure to replicate the pooling results is not totally unexpected as with small minor allele frequencies a very small change in the allele frequency can make a large difference to the statistical significance of the result. In this case the minor allele frequency of 140 G>A was reduced by only 0.02 in the cases, whilst there was a reduction of 0.04 and 0.02 in the cases and controls respectively for 1960 G>T, but these small differences were enough to make the two results non-significant.

Table 5.7: Individual genotyping results for SNP 140 G>A.

Individual genotype and allele counts and frequencies for GFAP 140 G>A after genotyping in the schizophrenia association sample. Both genotypic and allelic p-values are shown as calculated using CLUMP.

SNP	**Genotypes**			**Alleles**	
140 G>A	**G/G**	**G/A**	**A/A**	**G**	**A**
Cases	557	63	0	1177 (0.95)	63 (0.05)
Controls	635	59	1	1329 (0.96)	61 (0.04)
	χ^2 = 1.96 2df	P = 0.34		χ^2 = 0.69	P = 0.40

Table 5.8: Individual genotyping results for SNP 1960 G>T.

Individual genotype and allele counts and frequencies for GFAP 1960 G>T after genotyping in the schizophrenia association sample. Both genotypic and allelic p-values are shown as calculated using CLUMP.

SNP	**Genotypes**			**Alleles**	
1960 G>T	**G/G**	**G/T**	**T/T**	**G**	**T**
Cases	548	48	2	1218 (0.96)	52 (0.04)
Controls	656	47	0	1359 (0.97)	47 (0.03)
	χ^2 = 2.62 2df	P = 0.29		χ^2 = 1.06	P = 0.30

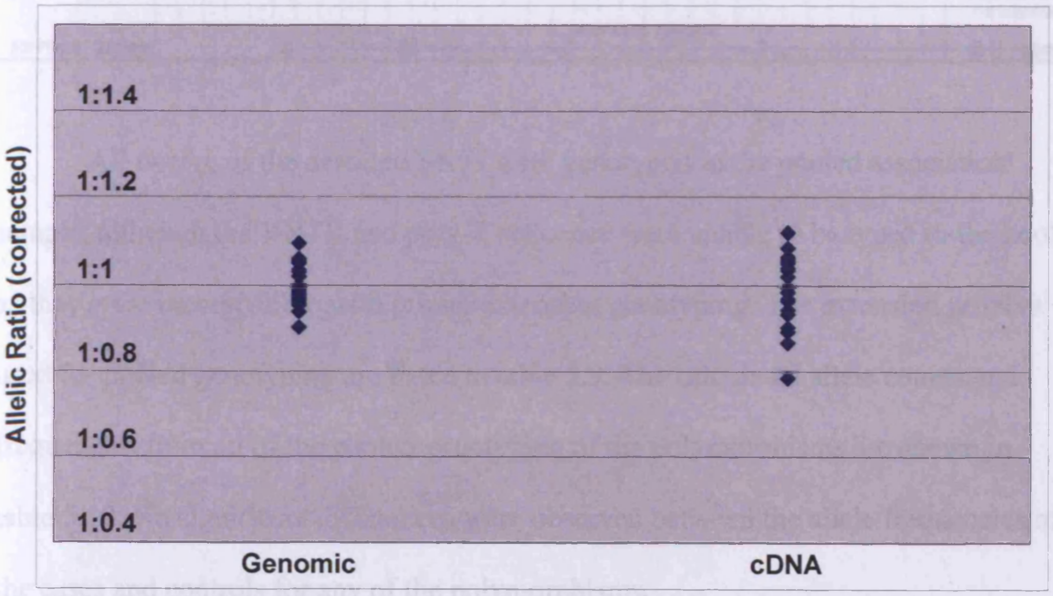
Haplotype analysis was performed on the two individually genotyped markers. The analysis revealed no evidence of association (global p value =0.38) with no individual haplotypes displaying a significant difference between the cases and controls.

5.3.1.1 Allelic expression assay of GFAP.

The G>C polymorphism rs11558961 was selected as a tag for analysis because of its suitable minor allele frequency of 0.3. Of the 60 subjects for whom brain was available, 27 were heterozygous. This subset of 27 individuals provided a power of approximately 93% to detect the effect of a regulatory variant present at a frequency of 0.05, and a power of 99% to detect a variant at a frequency of 0.1. Of the 27 GFAP heterozygous individuals, one showed a relative expression difference of ≥ 20% compared to the corresponding average genomic ratios (see Graph 5.1), the cDNA presenting with a corrected ratio of 1:0.78. The data points from each individual cDNA sample showed excellent reproducibility, with an average SD of the mean of 0.03. The assay therefore provides evidence for the presence of a rare *cis*-acting variant affecting the allelic expression of the GFAP gene. Whether this *cis*-acting variant is responsible for the results of previous expression studies (Tkachev, Mimmack et al. 2003; Fatemi, Laurence et al. 2004; Webster, O'Grady et al. 2005) is unclear though given its frequency, it is unlikely unless it is dramatically more common in cases. It seems more likely that the observation of altered GFAP expression results from *trans*-acting factors.

Graph 5.1: GFAP Allelic Expression: rs11558961.

Corrected genomic and cDNA ratios for rs11558961. Comparison between the observed corrected genomic ratios and the corrected cDNA ratios assayed for GFAP (n=27). Data are expressed as the mean of the ratio of G:C for two measurements of each genomic DNA sample and two measurements of each cDNA sample.



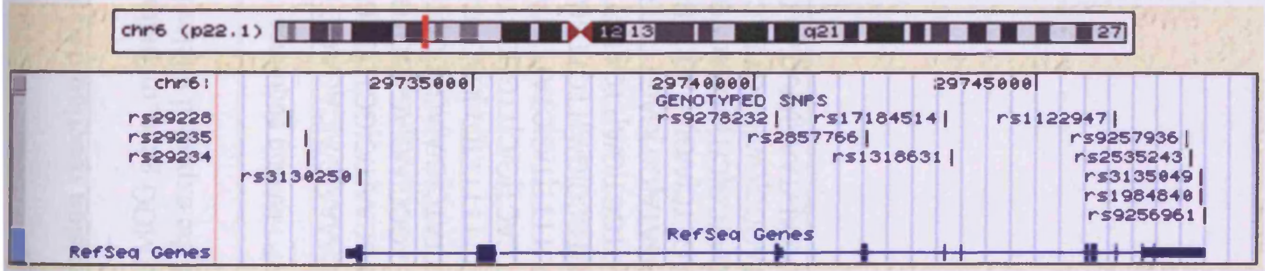
5.3.2 MOG:

Determination of the MOG intron/exon boundaries was based on alignment of the cDNA sequence NM_002433 and the genomic sequence available from the UCSC Genome Bioinformatics website (http://genome.ucsc.edu).

Mutation detection identified twelve SNPs, along with a VNTR and a poly T simple sequence repeat (See appendix 5.3 for sequencing results). There was one non-synonymous SNP V171L (rs2857766), one synonymous SNP S5S (rs3130250), five 3'-UTR polymorphisms (rs9257936, rs2535243, rs3135049, rs1984840 and rs9256961), three 5' flanking SNPs (rs29228, rs29235 and rs29234) and four intronic polymorphisms (rs9278232, rs17184514, rs1318631 and rs1122947) (figure 5.2).

Figure 5.2. Schematic diagram of MOG.

Diagram of the MOG gene (adapted from the UCSC Genome Browser), indicating the positions of the identified polymorphisms. For numbering of polymorphisms, +1 indicates the 'A' of the ATG translation codon. The gene is displayed (5'-3').



All twelve of the detected SNPs were genotyped in the pooled association sample, although the VNTR and poly-T sequence were unable to be typed in the pools as they were incompatible with primer extension genotyping. The extension primers used for pooled genotyping are listed in table 5.9. The calculated allele counts and frequencies from all of the pooled genotyping of the polymorphisms are shown in table 5.10. No significant differences were observed between the allele frequencies of the cases and controls for any of the polymorphisms.

Table 5.9: Polymorphism and primer extension reaction details for MOG.

Table of the polymorphisms identified in the MOG gene, including the sequence flanking each of the polymorphisms, the extension primer used to assay the polymorphism by SNaPshot and the expected alleles in the primer extension reaction.

SNP number.	rs Number	SNP Flanking sequence	Extension primer	Expected Alleles
-1909 G>A	rs29228	TCTCGGACAAAA[G/A]CAGAAATGTATG	CTCCTTACGCCCATCTCGGACAAAA	G (Blue) & A (Green)
-909 T>C	rs29235	ACCTGCTACCAA[T/C]GCTGGTCTCACA	TGAGAATTGCAAGTGTGAGACCAGC	A (Green) & G (Blue)
-875 A>C	rs29234	CTCAGCTGAGCC[A/C]AGAGGTGAGAGA	AAAATGACCTTTCTCTCACCTCT	T (Red) & G (Blue)
14 G>A	rs3130250	GGCAAGCTTATC[G/A]AGACCCTCTCTG	GAACAGTAGAGATGGCAAGCTTATC	G (Blue) & A (Green)
7412 Poly T	rs9278232	GAATTTTTTTTTTTTTTT[T/-]GAGACAGTGTCT	Unable to assay by primer extension.	
9016 G>C	rs2857766	CTGCAGATCACT[G/C]TTGGCCTCGTCT	GCAGAGGAAGACGAGGCCAA	C (Black) & G (Blue)
10423 T>G	rs17184514	ATGTTGCCTTTT[T/G]CTATTTTAGGAA	TCTGCTCGAAGTTTTCCTAAAATAG	A (Green) & C (Black)
10521 G>A	rs1318631	GGAAATGGTCCC[G/A]TTCTTGGATCGT	TTTTTGGCATAACGGAAATGGTCCC	G (Blue) & A (Green)
13447 G>A	rs1122947	GGGAAAGGTGCT[G/A]TTCATCTTCCAC	ATAGGGAGGATGTGGGAAAGGTGCT	G (Blue) & A (Green)
14789 C>T	rs9257936	TTAATAAAGATA[C/T]GAGTTTTGGCCG	AAACAAGCTAGGTGCTAATTAATAAAGATA	C (Black) & T (Red)
14875 A>G	rs2535243	GGTCAGGAGTTC[A/G]AGACCAGCCTGG	GGGCGAATCACGAGGTCAGGAGTTC	A (Green) & G (Blue)
15034 C>T	rs3135049	GAGCTGAGATCG[C/T]GCCACTGCACTC	CAGAGGTTGCAGTGAGCTGAGATCG	C (Black) & T (Red)
15052 G>A	rs1984840	TGCACTCTAGCC[G/A]GAGTGACAGAGT	CCACTGCACTCTAGCC	G (Blue) & A (Green)
15126 VNTR	rs9256961	TAAATAAATAAA[-/TAAA]AAATAATAATTAC	Unable to assay by primer extension.	

Table 5.10: Pooled genotyping data for MOG.

Estimated minor allele frequencies and minor allele counts for all of the pooled genotyping of the MOG polymorphisms, along with the corresponding chi-square and p-values.

SNP Data		Pooled Genotyping data				
SNP number.	rs Numbers	Minor Allele Cases (freq)	Minor Allele Controls (freq)	Difference	Chi-square	P-value
-1909 G>A	rs29228	A= 329 (0.24)	A= 326 (0.23)	0.01	0.45	0.50
-909 T>C	rs29235	C= 65 (0.05)	C= 64 (0.04)	0.01	0.09	0.76
-875 A>C	rs29234	C= 76 (0.05)	C= 72 (0.05)	0	0.32	0.57
14 G>A	rs3130250	A= 228 (0.17)	A= 240 (0.17)	0	0.03	0.87
7412 Poly T	rs9278232	Incompatible with a primer extension assay.				
9016 G>C	rs2857766	C= 381 (0.28)	C= 372 (0.26)	0.02	0.95	0.33
10423 T>G	rs17184514	G= 93 (0.07)	G= 105 (0.07)	0	0.38	0.54
10521 G>A	rs1318631	A= 349 (0.25)	A= 334 (0.23)	0.02	1.47	0.23
13447 G>A	rs1122947	G= 236 (0.17)	G= 254 (0.18)	0.01	0.19	0.66
14789 C>T	rs9257936	T= 156 (0.11)	T= 158 (0.11)	0	0.05	0.82
14875 A>G	rs2535243	G= 247 (0.18)	G= 270 (0.19)	0.01	0.43	0.51
15034 C>T	rs3135049	T= 169 (0.16)	T= 160 (0.15)	0.01	0.71	0.39
15052 G>A	rs1984840	A= 218 (0.2)	A= 223 (0.2)	0	0.01	0.92
15126 VNTR	rs9256961	Incompatible with a primer extension assay.				

As the 15126 VNTR (rs9256961) could not be assayed using the pooled genotyping primer extension method I decided to genotype the polymorphism individually using a Genescan fluorescent PCR assay (as described in section 2.10.5). The VNTR was made up of variable number of 4bp TAAA repeats. The calculated allele counts and frequencies for the VNTR genotyping are shown in table 5.11 below. No significant difference was observed between the allele counts in the cases and the controls.

Table 5.11: Allelic genotyping results for the MOG 15126 VNTR polymorphism.

Allele counts and frequencies observed when the MOG 15126 VNTR was genotyped using a fluorescent PCR. There were a total of ten alleles observed each 4bp apart. Chi-square and p-values calculated using CLUMP.

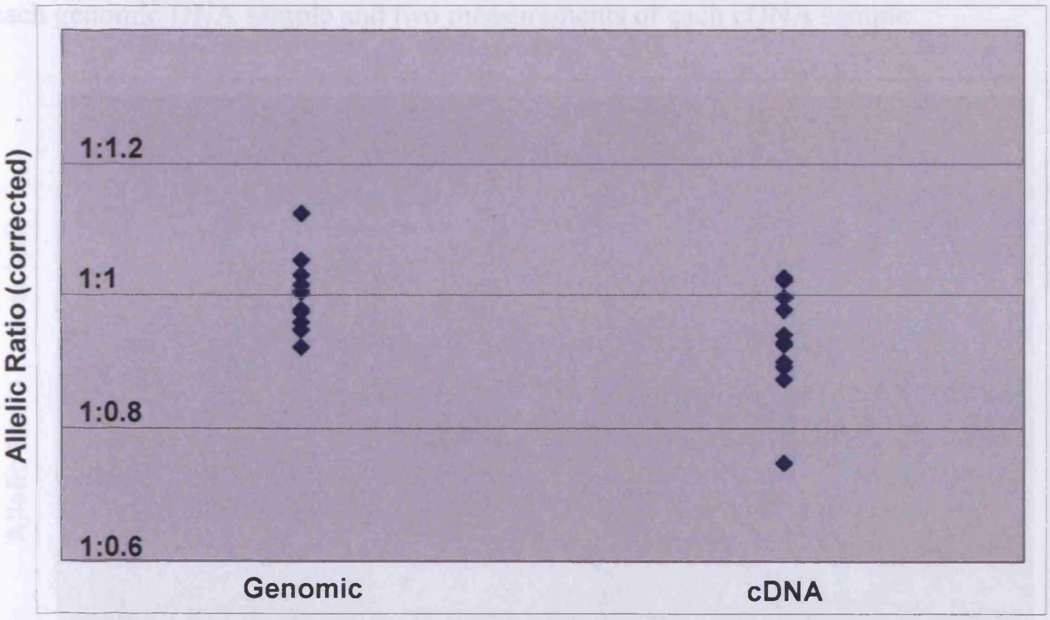
15126 VNTR	**Cases**		**Controls**	
Allele No (Size bp):	**Allele Count**	**Allele Freq**	**Allele Count**	**Allele Freq**
1 (232bp)	0	0	2	0.001
2 (236bp)	2	0.002	1	0.0007
3 (240bp)	7	0.005	11	0.008
4 (244bp)	83	0.064	90	0.063
5 (248bp)	68	0.052	77	0.054
6 (252bp)	288	0.22	272	0.19
7 (256bp)	368	0.28	408	0.29
8 (260bp)	251	0.19	278	0.20
9 (264bp)	230	0.18	278	0.20
10 (268bp)	7	0.005	7	0.005
			χ^2 = 7.232	**P= 0.642**

5.3.2.1 Allelic expression assay of MOG.

The MOG 14 G>A polymorphism (rs3130250) was selected as a tag marker for the assay due to its suitable minor allele frequency (0.17). Of the 60 subjects, 12 were heterozygous for rs3130250. This subset of 12 individuals provided a power of approximately 70% to detect the effect of a regulatory variant present at a frequency of 0.05, and a power of 90% to detect a variant at a frequency of 0.1. Of the 12 heterozygous individuals, only one showed a relative expression difference of ≥ 20% compared to the corresponding average genomic ratios (see Graph 5.2), with the cDNA presenting with a corrected ratio of 1:0.74. This corresponds to a 35% increase in expression of the A allele relative to the G allele. The data points from each individual cDNA sample showed excellent reproducibility, with an average SD of the mean of 0.03. Although only one individual showed an expression difference greater than the arbitrary 20% threshold described by (Bray, Buckland et al. 2003a), there was an interesting trend in the rest of the data. When cDNA data from all of the individuals was combined, the A allele was significantly over-expressed (p = 0.03, 99%CI), with an average 8% increase of the A allele relative to the G allele.

Graph 5.2: MOG Allelic Expression A: rs3130250.

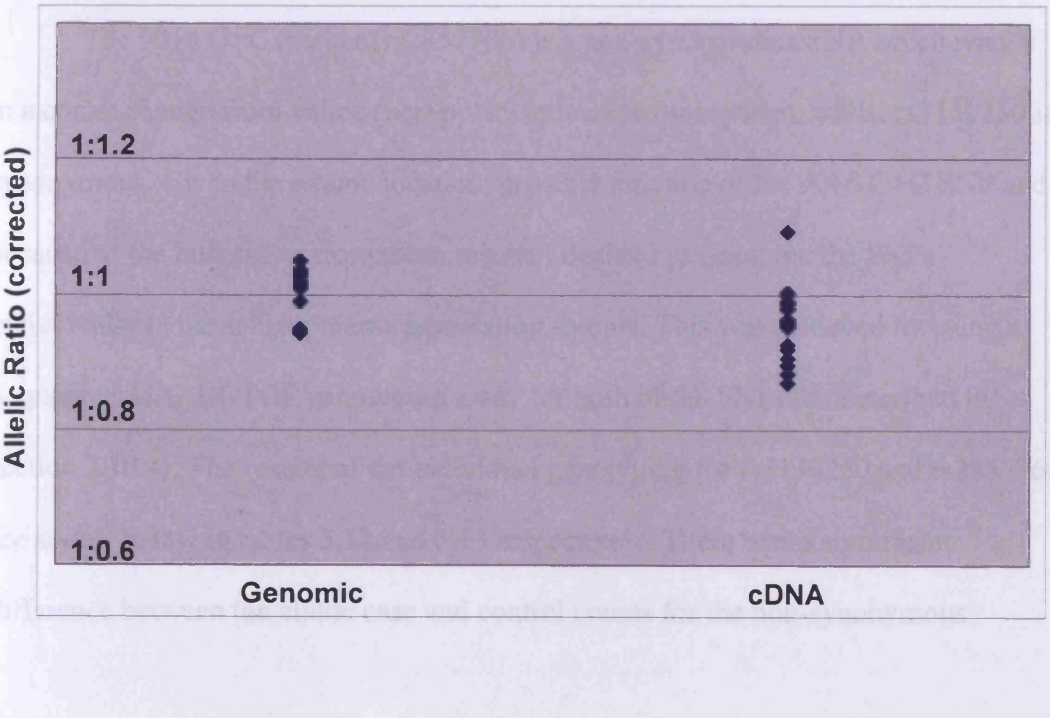
Corrected genomic and cDNA ratios for rs3130250. Comparison between the observed corrected genomic ratios and the corrected cDNA ratios assayed for MOG (n=12). Data are expressed as the mean of the ratio of A:G for two measurements of each genomic DNA sample and two measurements of each cDNA sample.



Given the significant overall expression difference observed I performed a second assay using a different tag SNP to try and replicate the expression findings from the first experiment. The MOG 9016 G>C polymorphism (rs2857766) was selected as a marker for the assay. Of the 60 subjects for whom brain was available, 18 were heterozygous for rs2857766 and informative for analysis. Of the 18 heterozygous individuals, none showed a relative expression difference of ≥ 20% compared to the corresponding average genomic ratios (see Graph 5.3). The data points from each individual cDNA sample showed good reproducibility, with an average SD of the mean of 0.03. When the cDNA data from all of the individuals were combined, the C allele was significantly over-expressed ($p = 0.006$, 99%CI), displaying an average 5% increase of the C allele relative to the G allele.

Graph 5.3: MOG Allelic Expression B: rs2857766.

Corrected genomic and cDNA ratios for rs2857766. Comparison between the observed corrected genomic ratios and the corrected cDNA ratios assayed for MOG (n=18). Data are expressed as the mean of the ratio of C:G for two measurements of each genomic DNA sample and two measurements of each cDNA sample.



Alternative transcripts of MOG exist. However, the rs3130250 tag SNP is a synonymous mutation in all of the known transcripts, whilst the rs2857766 tag SNP is a non-synonymous mutation in the majority of transcripts and is located in the 3'UTR of the remaining transcripts. The two marker SNPs are both therefore conserved in the exonic sequence of all known transcripts of MOG and hence the allelic expression assays should detect the presence of altered total expression within the different gene transcripts. LD analysis of the two allelic expression markers revealed the markers to be related (D' = 1.0, $r^2 = 0.04$). Further analysis revealed the presence of three haplotypes; the common G-G and the two rarer A-G and G-C haplotypes. The two rare alleles (A-C) are not observed together as they appear to be on different haplotype backgrounds. Therefore the expression of both the rare alleles haplotypes is independently increased compared to the common haplotype. Or conversely the common haplotype is always under expressed compared to the two rare haplotype backgrounds. The results of the assays would suggest the existence of one or more *cis*-acting variants which have a modest effect on the expression of the MOG gene.

The 9016 G>C marker (rs2857766) is a non-synonymous SNP which results in a codon change from valine (non-polar) to leucine (non-polar), whilst rs3130250 is synonymous. Given the exonic location, possible function of the 9016 G>C SNP and because of the interesting expression results I decided to genotype the SNPs individually in the schizophrenia association sample. This was achieved by using a Sequenom MALDI-TOF genotyping assay for both of the SNPs (as described in section 2.10.4). The results of the individual genotyping for rs3130250 and rs2857766 are shown below in tables 5.12 and 5.13 respectively. There was a significant difference between the allelic case and control counts for the non-synonymous rs2857766 (table 5.13), caused by an increased frequency of the minor C allele in the cases (0.27) compared to the controls (0.24). No significant difference between the genotype or allele counts in the cases and controls was detected for rs3130250. In the expression assays the minor alleles of both markers displayed a relative increase in expression. Therefore the results of the individual genotyping, with an increased frequency of the minor alleles in schizophrenic cases are consistent with increased MOG expression in schizophrenia. However, the results of previous micro-array studies (Tkachev, Mimmack et al. 2003; Katsel, Davis et al. 2005b) have actually reported a reduced expression of MOG in schizophrenia, a finding which is contradictory to these results. This reported down-regulation must therefore result from *trans*-acting factors and is not likely to be aetiological. Though, the non-synonymous SNP might be of aetiological importance if it affects a *trans*-acting factor to influence expression and may still account for the reduced expression.

Table 5.12: Individual genotyping results for rs3130250.

Genotype and allele counts and frequencies for MOG SNP rs3130250 in the schizophrenia association sample. Both genotypic and allelic p-values are shown.

SNP	**Genotypes**			**Alleles**	
14 G>A	**G/G**	**G/A**	**A/A**	**G**	**A**
Cases	430	153	20	1013 (0.84)	193 (0.16)
Controls	492	171	15	1155 (0.85)	201 (0.15)
	$\chi^2 = 1.498$	P = 0.473		$\chi^2 = 0.683$	P = 0.408

Table 5.13: Individual genotyping results for rs2857766.

Genotype and allele counts and frequencies for MOG SNP rs2857766 in the schizophrenia association sample. Both genotypic and allelic p-values are shown.

SNP	**Genotypes**			**Alleles**	
9016 G>C	**G/G**	**G/C**	**C/C**	**G**	**C**
Cases	302	225	43	829 (0.73)	311 (0.27)
Controls	368	230	34	966 (0.76)	298 (0.24)
	$\chi^2 = 4.422$	P = 0.110		$\chi^2 = 4.349$	P = **0.037**

Haplotype analysis was performed on all of the individually genotyped markers including the VNTR polymorphism. None of the global haplotype p-values for the two or three marker haplotypes revealed any evidence of association and there were no individual haplotypes that displayed a significant difference between the cases and controls (data not shown).

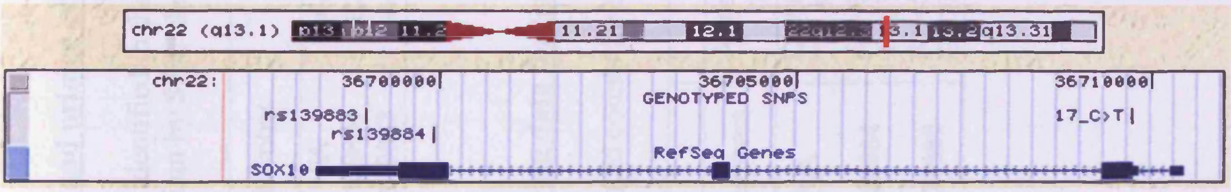
5.3.3 SOX 10:

Determination of the SOX 10 intron/exon boundaries was based on the alignment of the cDNA sequence NM_006941 with genomic sequence available from the UCSC Genome Bioinformatics website (http://genome.ucsc.edu).

Mutation detection assays revealed three SNPs; two coding synonymous SNPs D6D (17 C>T) and H309H (rs139884), and one 3'-UTR SNP (rs139883) (figure 5.3) (See appendix 5.3 for sequencing results).

Figure 5.3. Schematic diagram of SOX 10.

Diagram of the SOX 10 gene (adapted from the UCSC Genome Browser), indicating the positions of the identified polymorphisms. For numbering of polymorphisms, +1 indicates the 'A' of the ATG translation codon. The gene is displayed (3'-5').



All three of the SNPs were genotyped in the pooled association sample (table 5.15). The extension primers used for pooled genotyping are listed in table 5.14. No significant differences were observed between the allele frequencies of the cases and controls.

Table 5.14: Polymorphism and primer extension reaction details for SOX 10.

Table of the polymorphisms identified in the SOX 10 gene, including the sequence flanking each of the polymorphisms, the extension primer used to assay the polymorphism by SNaPshot and the expected alleles in the primer extension reaction.

SNP number.	rs Number	SNP Flanking sequence	Extension primer	Expected Alleles
17 C>T	N/A	GGAGGAGCAGGA[C/T]CTATCGGAGGTG	CTCAGCTCCACCTCCGATAG	G (Blue) & A (Green)
9815 C>T	rs139884	GCACCCAGGCCA[C/T]GTGAGCAGCTAC	GCCCAATGGGCACCCAGGCCA	C (Black) & T (Red)
10764 G>A	rs139883	GAGACTGTGCCC[G/A]GAGCTGGCCGTG	GCCACCCAGAGACTGTGCCC	G (Blue) & A (Green)

Table 5.15: Pooled genotyping data for Sox 10.

Estimated allele frequencies and counts for the SOX 10 polymorphisms, along with the corresponding p-values.

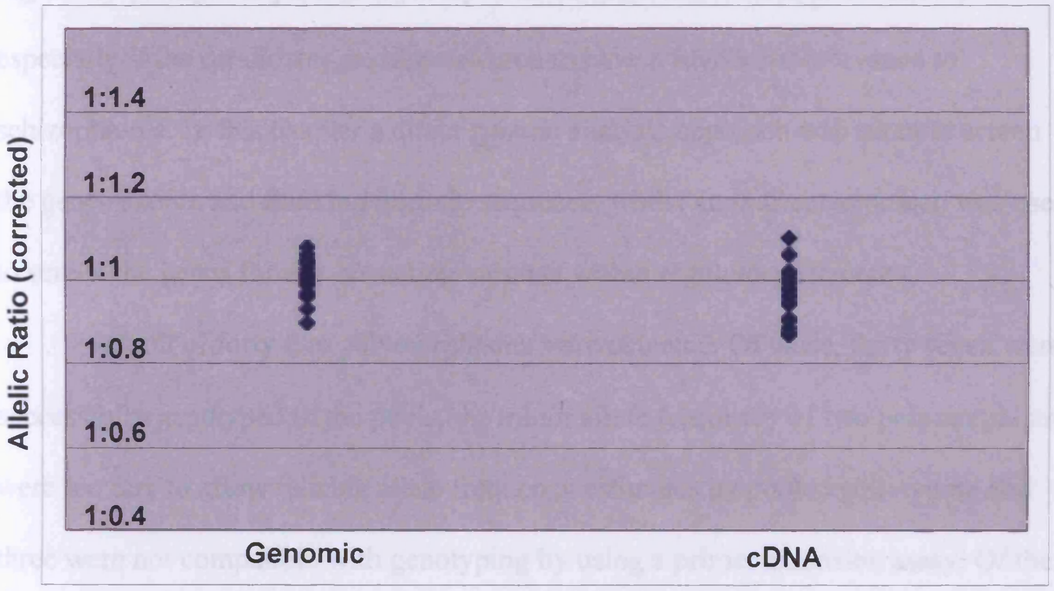
SNP Data		Pooled Genotyping data				
SNP number.	rs Numbers	Alleles Cases (freq)	Alleles Controls (freq)	Difference	Chi-square	P-value
17 C>T	N/A	C= 680 (0.92) T= 56 (0.08)	C= 686 (0.93) T= 50 (0.07)	0.01	0.37	0.55
9815 C>T	rs139884	C= 462 (0.63) T= 274 (0.37)	C= 462 (0.63) T= 274 (0.37)	0	0	1
10764 G>A	rs139883	A= 458 (0.62) G= 278 (0.38)	A= 467 (0.63) G= 269 (0.37)	0.1	0.24	0.63

5.3.3.1 Allelic expression assay of SOX 10.

The G>A marker rs139884 was selected for analysis given its informative minor allele frequency of 0.37. Of the 60 subjects for whom brain was available, 24 individuals were heterozygous for SNP rs139884. This subset of 24 individuals provided a power of approximately 91% to detect the effect of a regulatory variant present at a frequency of 0.05, and a power of 99% to detect a variant at a frequency of 0.1. Of the 24 SOX 10 heterozygous individuals, none showed a relative expression difference of ≥ 20% compared to the corresponding average genomic ratios (see Graph 5.4), the cDNA ratios varied from 1:1.1 to 1:0.87. The data points from each individual cDNA sample showed good reproducibility, with an average SD of the mean of 0.025. The assay therefore provides no evidence for the presence of a *cis*-acting variant affecting the allelic expression of the Sox 10 gene, and suggests that the down-regulation observed in micro-array studies are caused by the effects of environmental or other *trans*-acting factors acting upon the gene. This is inconsistent with the reports of increased methylation reducing the genes expression (Iwamoto et al, 2005b). However, it is also possible that the results of the micro-array studies were influenced by confounding factors such as methodological artefacts, experimental variation and sample differences (Katsel, Davis et al. 2005c). Although it is of note that the down regulation of SOX 10 has been reported in independent studies (Tkachev, Mimmack et al. 2003; Iwamoto, Bundo et al. 2005b; Katsel, Davis et al. 2005b) providing support for the observed expression differences as a genuine finding.

Graph 5.4: SOX 10 Allelic Expression: rs139884.

Corrected genomic and cDNA ratios for rs139884. Comparison between the observed corrected genomic ratios and the corrected cDNA ratios assayed for SOX 10 (n=24). Data are expressed as the mean of the ratio of G:A for two measurements of each genomic DNA sample and two measurements of each cDNA sample.



5.4 Discussion:

In this chapter I report analysis of three genes (GFAP, MOG and SOX 10) relevant to myelin function for which there is altered expression data, and are all located within reported linkage regions. By testing candidate genes located within regions of linkage the prior probability of detecting a risk variant is enhanced, especially if the candidates are also selected to have a functional relevance to schizophrenia. In this chapter a direct genetic analysis approach was taken to screen the genes exonic and flanking intronic sequence, whilst an indirect approach was used to screen the genes for any *cis*-acting variants within regulatory elements.

A total of forty two polymorphisms were detected. Of these, thirty seven were successfully genotyped in the pools, the minor allele frequency of two polymorphisms were too rare to allow reliable allele frequency estimates by pooled genotyping and three were not compatible with genotyping by using a primer extension assay. Of the thirty seven polymorphisms successfully typed in the pools only two GFAP markers showed significant differences between the case and control pools, p=<0.05 in each case: GFAP 140 G>A and GFAP 1960 G>T. However, when these two SNPs were individually genotyped in the association sample, no significant difference was observed in the allele frequencies of either SNP between cases and controls (see tables 5.7 and 5.8). The detection of a false positive result is not be unexpected as with a level of significance of 0.05 you would expect a positive finding to occur by statistical chance once in every twenty polymorphisms genotyped. Though if the findings were attributable to chance then the results would also be expected to replicate when the identical samples were genotyped individually. However, in this case the pooled and individual genotyping were performed in slightly different sized samples, as at the time of study only 535 cases and 552 controls of the association samples were available in a pooled format, thus the observed differences in allele frequency might be attributable to the slightly different sizes of the genotyped samples. It is also possible that the positive findings were caused by technical error in the SNaPshot pooled genotyping. This technical error is likely to have been inflated in this case as both polymorphisms had low minor allele frequencies making accurate allele frequency estimation by pooled genotyping more difficult.

As well as screening the three genes for genetic association to schizophrenia I have also assayed the genes to look for the presence of *cis*-acting sequence variants which affect the expression of the genes. In the case of the SOX 10 gene there was no evidence for a genetic variant affecting expression significantly, whilst there was evidence for modest altered allelic expression of the GFAP gene caused by a rare *cis*-acting variant. The results of my SOX 10 allelic expression assay does not support the findings of previous studies (Tkachev, Mimmack et al. 2003; Iwamoto, Bundo et al. 2005b; Katsel, Davis et al. 2005b) which reported a marked two-fold down regulation of the gene in schizophrenia. The failure to detect any common *cis*-acting variants affecting SOX 10 expression can not be explained by a relative lack of power, as the subset of 24 individuals used to perform the assay provided a power of approximately 91% to detect the effect of a regulatory variant present at a frequency of 0.05, and a power of 99% to detect a variant at a frequency of 0.1. The observed expression differences in the micro-array studies may reflect the effects of *trans*-acting factors. However the findings of Iwamoto and colleagues (Iwamoto, Bundo et al. 2005b) who reported a higher frequency of methylated SOX 10 alleles correlated with the genes reduced expression might provide support for the existence of *cis*-acting effects. It is also of note that none of the brain samples used in my allelic expression assay were from schizophrenic patients. Thus if the altered expression of SOX 10 is caused by an epigenetic process highly specific to schizophrenic brain, the failure to detect any significant allelic expression differences in my assay would not be unexpected.

The results of the MOG allelic expression assay were potentially more interesting. Whilst there was no large expression differences observed between haplotypes there was a relative increase in expression of the minor alleles observed for two independent assays, although the two markers were related (D'=1.0, r^2=0.04). This on its own is not particularly interesting, but when the two marker SNPs were subsequently genotyped individually in the case-control sample the minor allele frequency of the non-synonymous marker rs2857766 was increased in schizophrenics compared to controls, with the C allele being significantly associated (uncorrected) with schizophrenia (p=0.03) (table 5.13). Haplotype analysis of individually typed markers revealed no stronger association with schizophrenia at global haplotype or specific haplotype level. That the expression assays displayed a relative increase of the minor alleles expression, when taken with the results of the individual genotyping which displayed an increased frequency of the minor alleles in schizophrenia, would support a theory of MOG expression being increased in schizophrenia.

However, it should be noted that my results indicating a possible increase in expression of the MOG gene in schizophrenia contradict with the results of the previous micro-array studies (Tkachev, Mimmack et al. 2003; Katsel, Davis et al. 2005b), which report down-regulation of the gene in schizophrenic brain. My data suggests that common *cis*-acting factors are not individually responsible for the large 2-3 fold expression differences observed in these studies and that the micro-array findings are likely to be attributable to *trans*-acting factors or environmental effects such as medication usage.

The results of my association and allelic expression analysis of the MOG gene along with the results of the previous studies by (Liu, Qin et al. 2005; Zai, King et al. 2005) are not compelling and provide little evidence for the involvement of MOG in the aetiology of schizophrenia. However, further examination of the non-synonymous SNP is probably warranted to independently replicate the results.

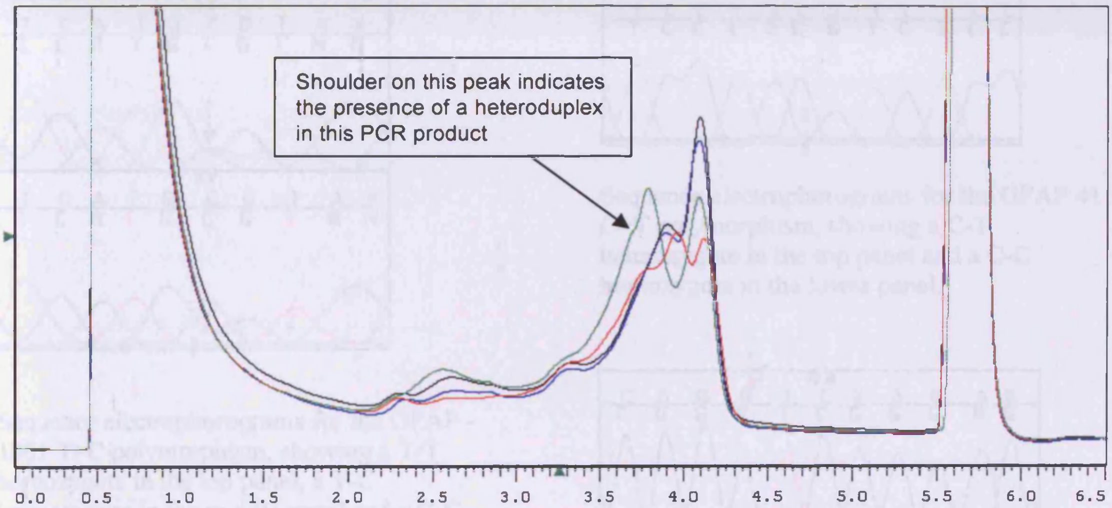
With regard to the successfully typed polymorphisms in GFAP and Sox 10, I conclude that there is no evidence of association with schizophrenia. However, as I discussed earlier (chapter 4, section 4.4) I cannot completely exclude these genes from being involved in susceptibility to schizophrenia.

5.5 Appendix:

Examples of experimental data for chapter 5.

Appendix 5.1.

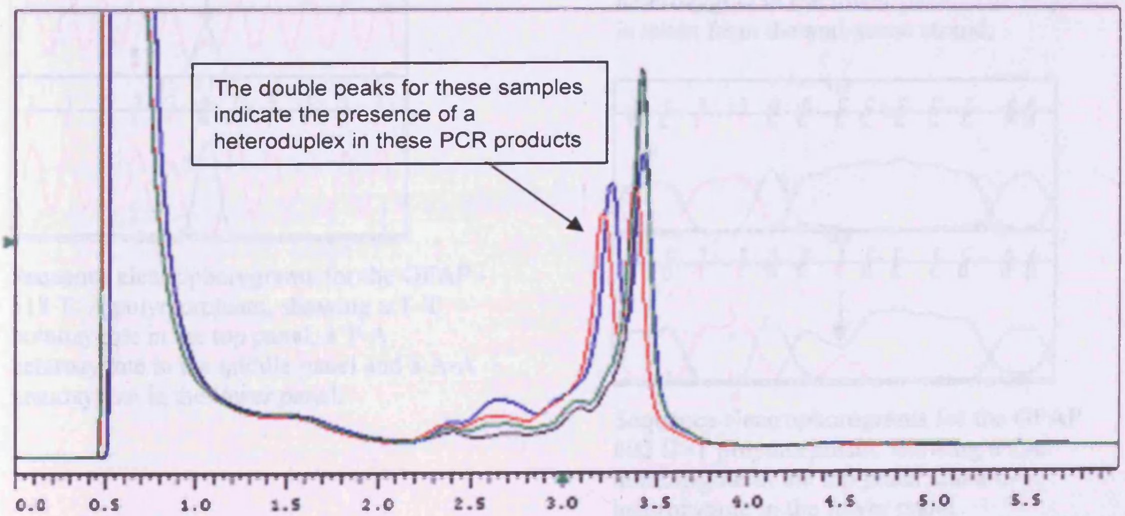
Example of DHPLC chromatographs produced when GFAP amplimer 4 was analysed at 65˚C, for screening set samples 1, 2, 3 and 5. The chromatograph shows the peaks representing these samples PCR products, with changes in the peak shape indicating the presence of a heteroduplex.



Shoulder on this peak indicates the presence of a heteroduplex in this PCR product

Appendix 5.2.

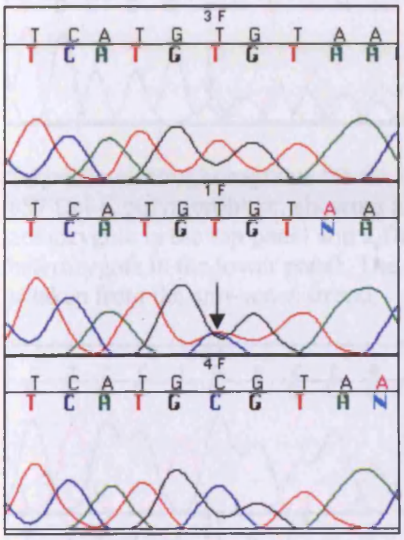
Example of the DHPLC chromatographs produced when MOG amplimer 6 was analysed at 62˚C, for screening set samples 11, 12, 13 and 14. The chromatograph shows the peaks representing these samples PCR products, with changes in the peak shape indicating the presence of a heteroduplex.



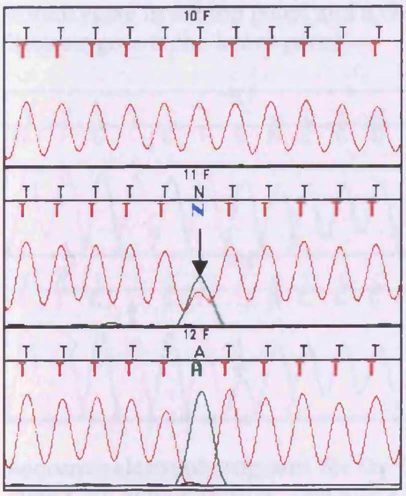
The double peaks for these samples indicate the presence of a heteroduplex in these PCR products

Appendix 5.3:

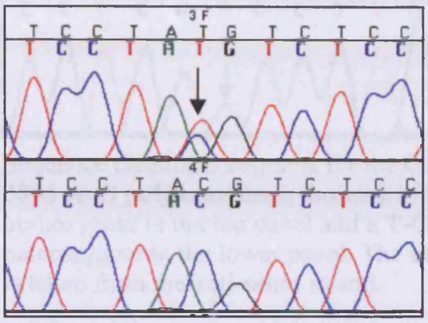
Examples of the different sequencing electrophorograms produced for each of the polymorphisms that were detected as a result of screening the three myelination genes in this project. (Polymorphisms are indicated with an arrow).



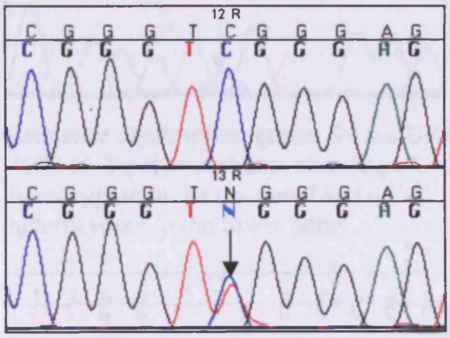
Sequence electrophorograms for the GFAP -1001 T>C polymorphism, showing a T-T homozygote in the top panel, a T-C heterozygote in the middle panel and a C-C homozygote in the lower panel.



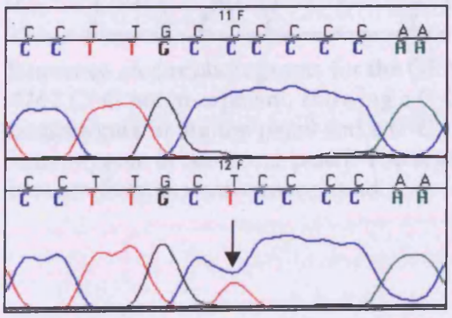
Sequence electrophorograms for the GFAP -518 T>A polymorphism, showing a T-T homozygote in the top panel, a T-A heterozygote in the middle panel and a A-A homozygote in the lower panel.



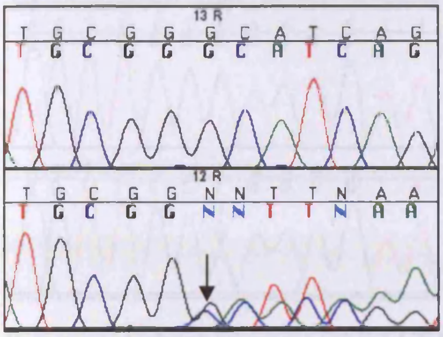
Sequence electrophorograms for the GFAP 41 C>T polymorphism, showing a C-T heterozygote in the top panel and a C-C homozygote in the lower panel.



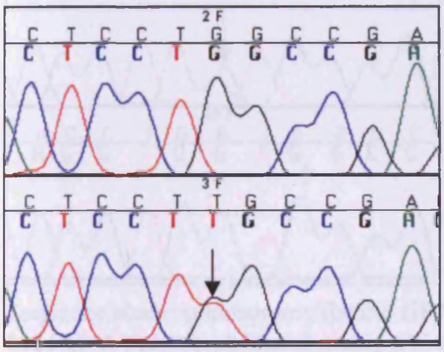
Sequence electrophorograms for the GFAP 140 G>A polymorphism, showing a C-C homozygote in the top panel and a C-T heterozygote in the lower panel. The sequence is taken from the anti-sense strand.



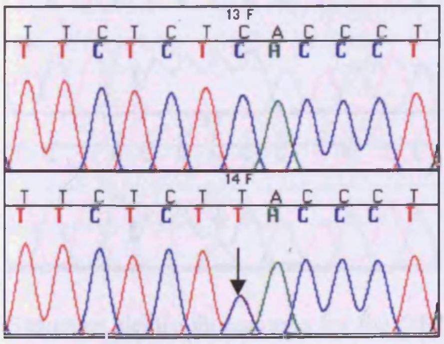
Sequence electrophorograms for the GFAP 602 C>T polymorphism, showing a C-C homozygote in the top panel and a C-T heterozygote in the lower panel.



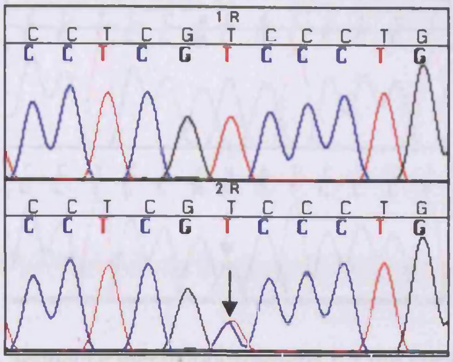
Sequence electrophorograms for the GFAP 657 Del-C polymorphism, showing a G-G homozygote in the top panel and a Del-G heterozygote in the lower panel. The sequence is taken from the anti-sense strand.



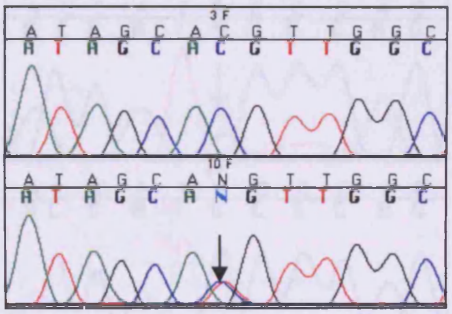
Sequence electrophorograms for the GFAP 1960 G>T polymorphism, showing a G-G homozygote in the top panel and a G-T heterozygote in the lower panel.



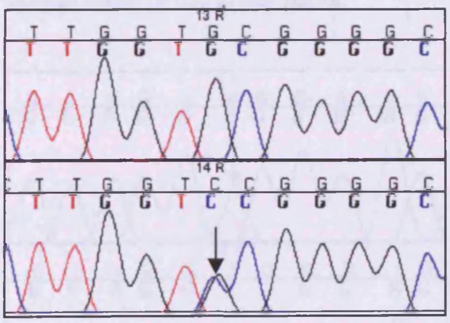
Sequence electrophorograms for the GFAP 2044 C>T polymorphism, showing a C-C homozygote in the top panel and a C-T heterozygote in the lower panel.



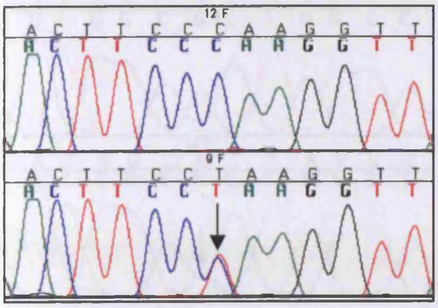
Sequence electrophorograms for the GFAP 3895 A>G polymorphism, showing a T-T homozygote in the top panel and a T-C heterozygote in the lower panel. The sequence is taken from the anti-sense strand.



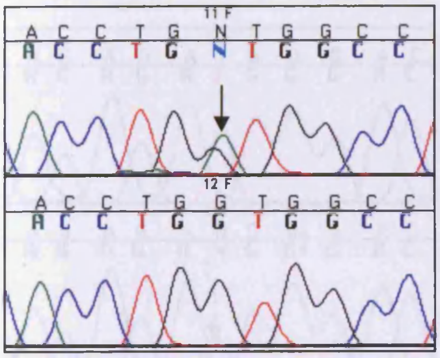
Sequence electrophorograms for the GFAP 4677 C>T polymorphism, showing a C-C homozygote in the top panel and a C-T heterozygote in the lower panel.



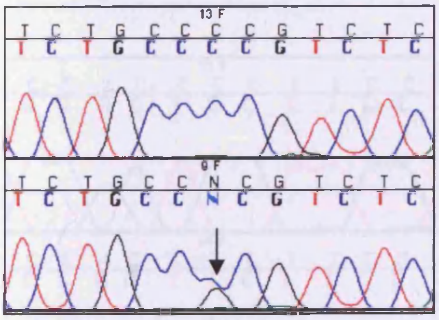
Sequence electrophorograms for the GFAP 4762 C>G polymorphism, showing a G-G homozygote in the top panel and a G-C heterozygote in the lower panel. The sequence is taken from the anti-sense strand.



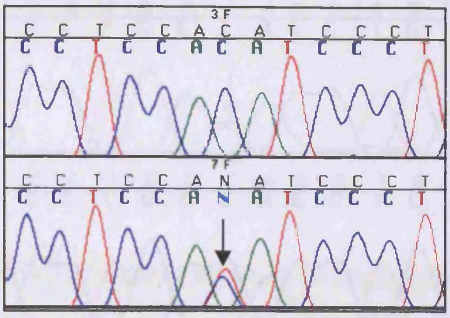
Sequence electrophorograms for the GFAP 8012 C>T polymorphism, showing a C-C homozygote in the top panel and a C-T heterozygote in the lower panel.



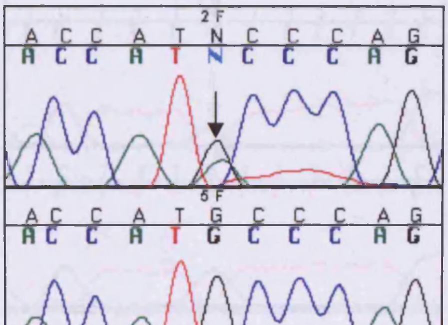
Sequence electrophorograms for the GFAP 8155 G>A polymorphism, showing a G-A heterozygote in the top panel and a G-G homozygote in the lower panel.



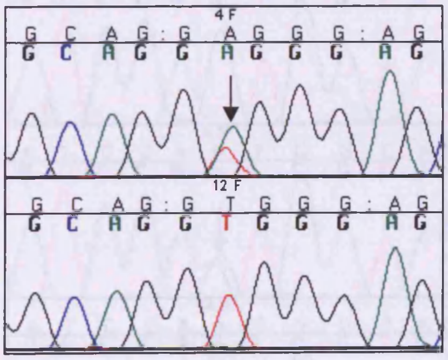
Sequence electrophorograms for the GFAP 8167 C>G polymorphism, showing a C-C homozygote in the top panel and a C-G heterozygote in the lower panel.



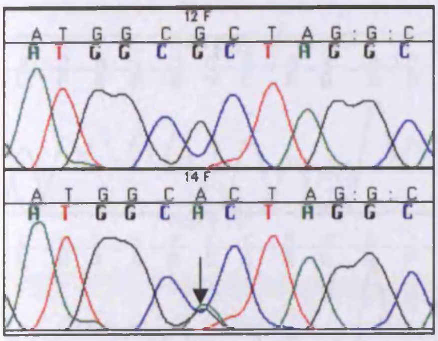
Sequence electrophorograms for the GFAP 8477 C>T polymorphism, showing a C-C homozygote in the top panel and a C-T heterozygote in the lower panel.



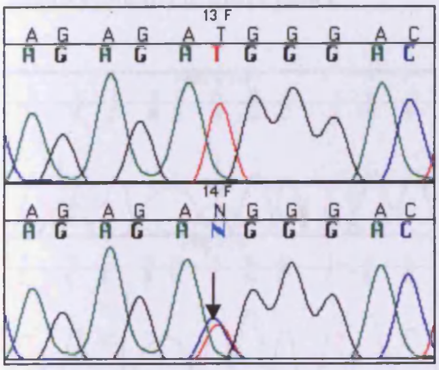
Sequence electrophorograms for the GFAP 8769 G>A polymorphism, showing a G-A heterozygote in the top panel and a G-G homozygote in the lower panel.



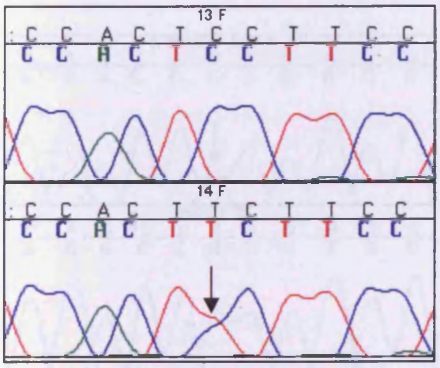
Sequence electrophorograms for the GFAP 8953 T>A polymorphism, showing a T-A heterozygote in the top panel and a T-T homozygote in the lower panel.



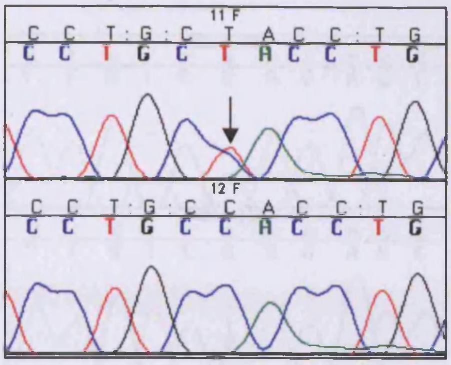
Sequence electrophorograms for the GFAP 9140 G>A polymorphism, showing a G-G homozygote in the top panel and a G-A heterozygote in the lower panel.



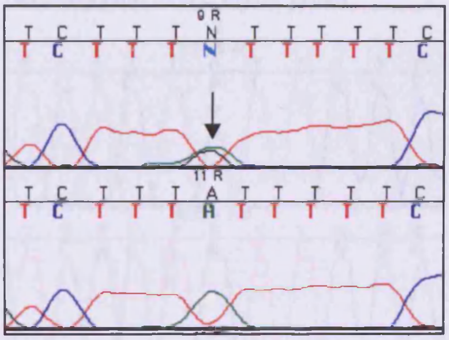
Sequence electrophorograms for the GFAP 9661 T>C polymorphism, showing a T-T homozygote in the top panel and a T-C heterozygote in the lower panel.



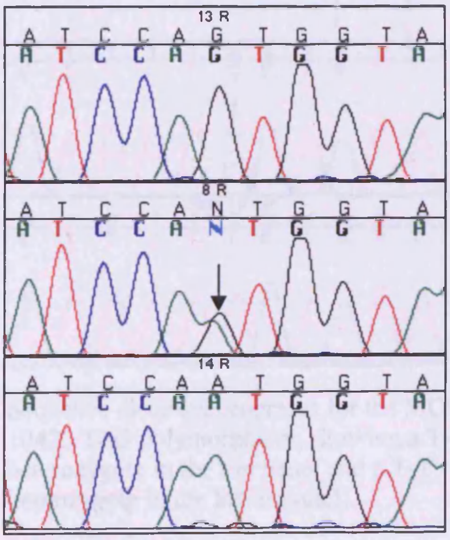
Sequence electrophorograms for the GFAP 9759 C>T polymorphism, showing a C-C homozygote in the top panel and a C-T heterozygote in the lower panel.



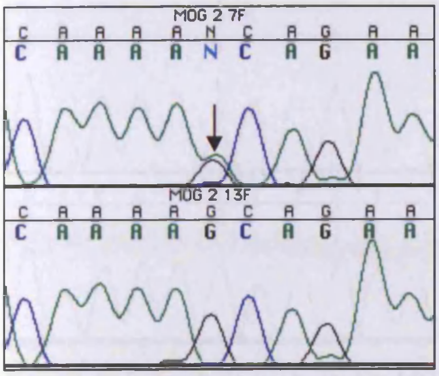
Sequence electrophorograms for the GFAP 9768 C>T polymorphism, showing a C-T heterozygote in the top panel and a C-C homozygote in the lower panel.



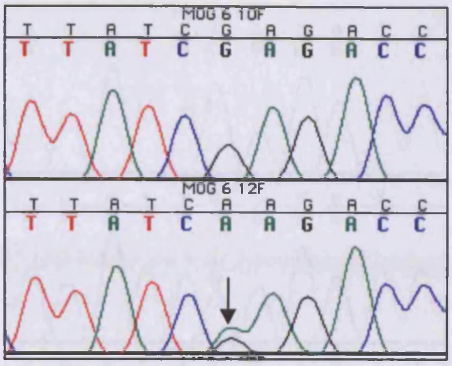
Sequence electrophorograms for the GFAP 9834 T>C polymorphism, showing a A-G heterozygote in the top panel and a A-A homozygote in the lower panel. The sequence is taken from the anti-sense strand.



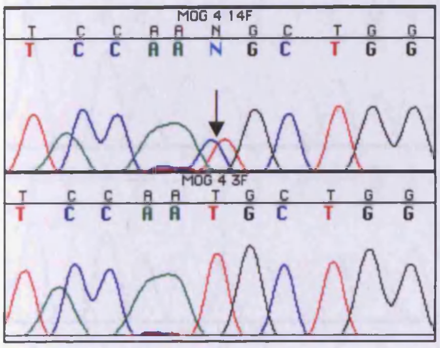
Sequence electrophorograms for the GFAP 10125 C>T polymorphism, showing a G-G homozygote in the top panel, a G-A heterozygote in the middle panel and a A-A homozygote in the lower panel. The sequence is taken from the anti-sense strand.



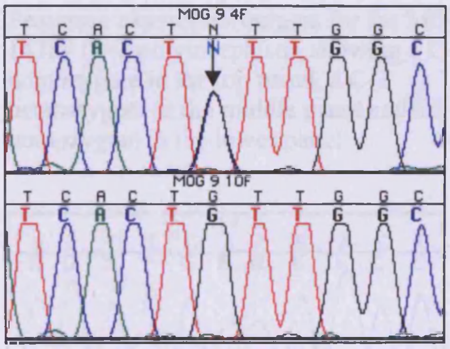
Sequence electrophorograms for the MOG -1909 G>A polymorphism, showing a G-A heterozygote in the top panel and a G-G homozygote in the lower panel.



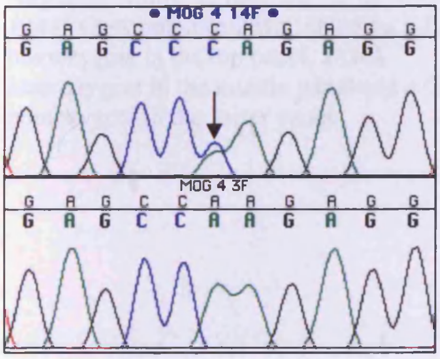
Sequence electrophorograms for the MOG 14 G>A polymorphism, showing a G-G homozygote in the top panel and a G-A heterozygote in the lower panel.



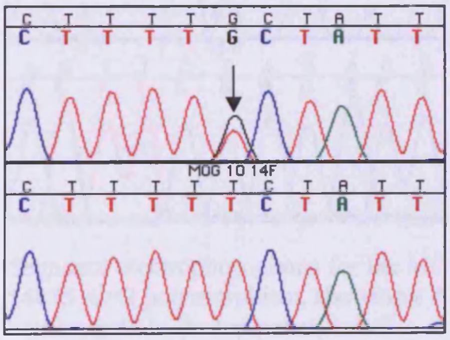
Sequence electrophorograms for the MOG -909 T>C polymorphism, showing a T-C heterozygote in the top panel and a T-T homozygote in the lower panel.



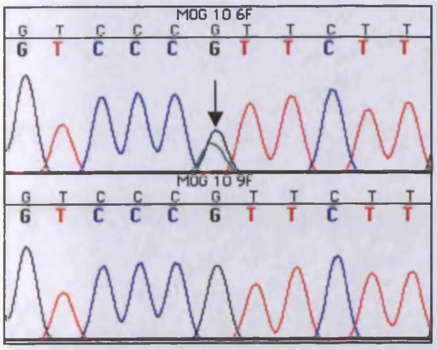
Sequence electrophorograms for the MOG 9016 G>C polymorphism, showing a G-C heterozygote in the top panel and a G-G homozygote in the lower panel.



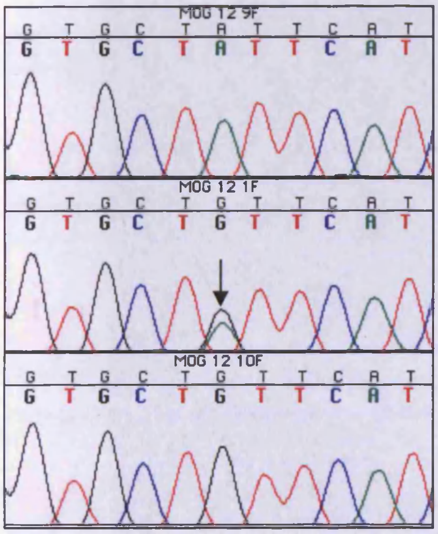
Sequence electrophorograms for the MOG -875 A>C polymorphism, showing a A-C heterozygote in the top panel and a A-A homozygote in the lower panel.



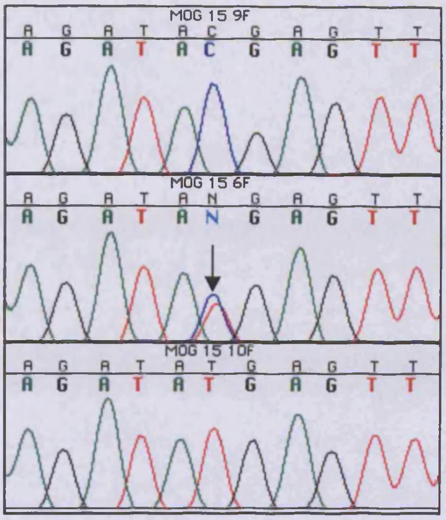
Sequence electrophorograms for the MOG 10423 T>G polymorphism, showing a T-G heterozygote in the top panel and a T-T homozygote in the lower panel.



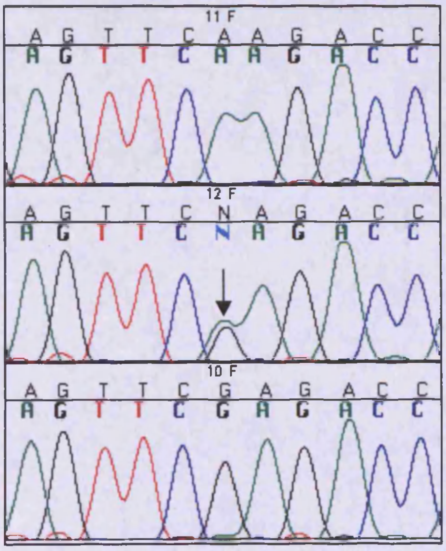
Sequence electrophorograms for the MOG 10521 G>A polymorphism, showing a G-A heterozygote in the top panel and a G-G homozygote in the lower panel.



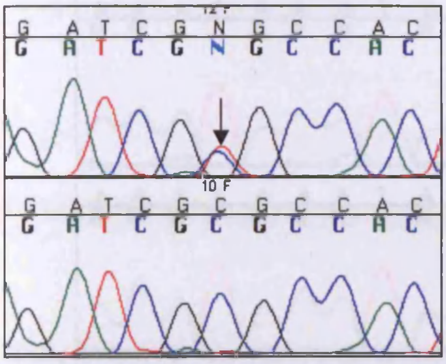
Sequence electrophorograms for the MOG 13447 G>A polymorphism, showing a A-A homozygote in the top panel, a G-A heterozygote in the middle panel and a G-G homozygote in the lower panel.



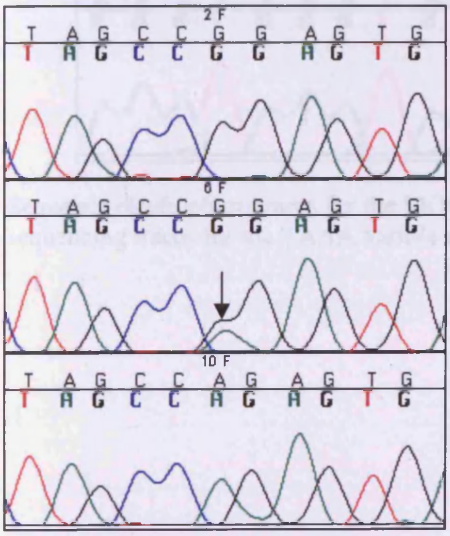
Sequence electrophorograms for the MOG 14789 C>T polymorphism, showing a C-C homozygote in the top panel, a C-T heterozygote in the middle panel and a T-T homozygote in the lower panel.



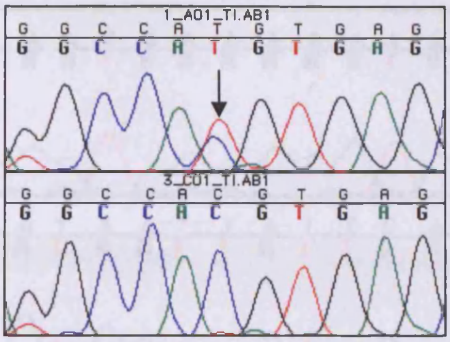
Sequence electrophorograms for the MOG 14875 A>G polymorphism, showing a A-A homozygote in the top panel, a A-G heterozygote in the middle panel and a G-G homozygote in the lower panel.



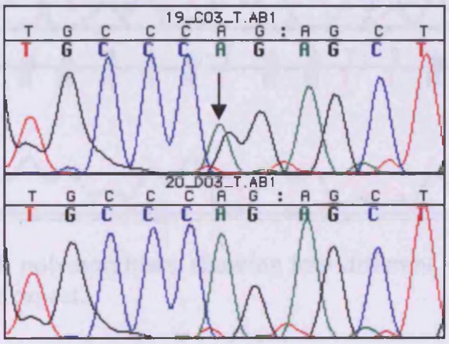
Sequence electrophorograms for the MOG 15034 C>T polymorphism, showing a C-T heterozygote in the top panel and a C-C homozygote in the lower panel.



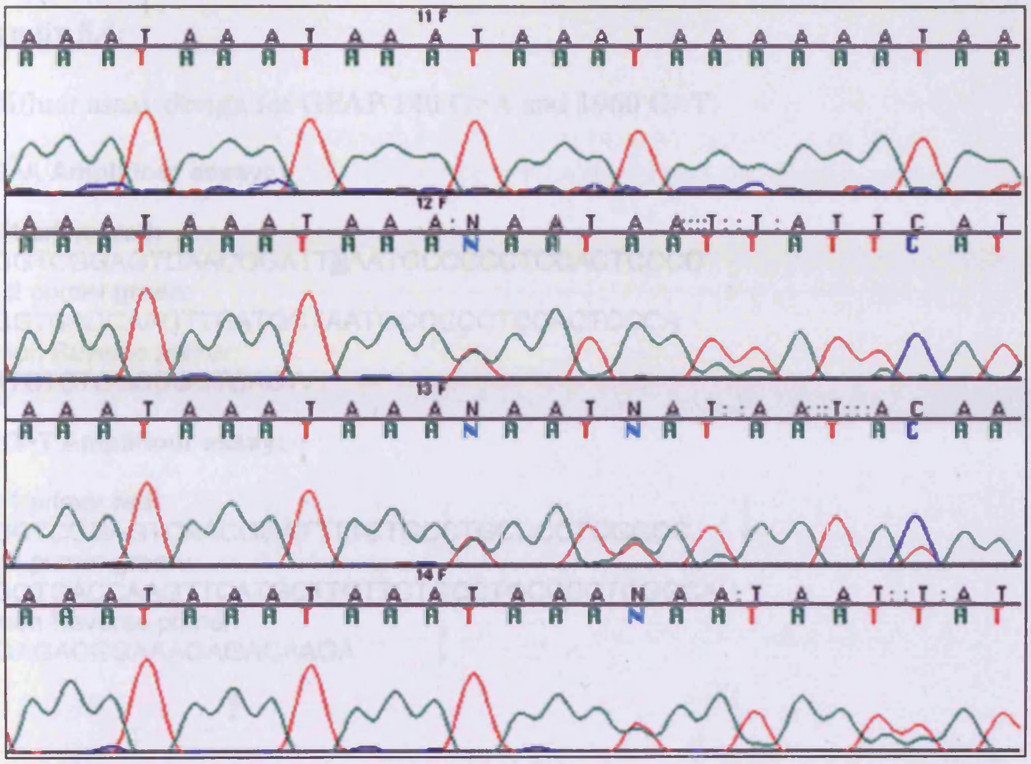
Sequence electrophorograms for the MOG 15052 G>A polymorphism, showing a G-G homozygote in the top panel, a G-A heterozygote in the middle panel and a A-A homozygote in the lower panel.



Sequence electrophorograms for the SOX 10 9815 C>T polymorphism, showing a C-T heterozygote in the top panel and a C-C homozygote in the lower panel.



Sequence electrophorograms for the SOX 10 10764 G>A polymorphism, showing a G-A heterozygote in the top panel and a A-A homozygote in the lower panel.



Sequence electrophorograms for the MOG 15126 VNTR polymorphism, showing four different length sequencing traces for the TAAA varible number tandem repeat.

Appendix 5.4:

Amplifluor assay design for GFAP 140 G>A and 1960 G>T.

140 G>A Amplifluor assay:

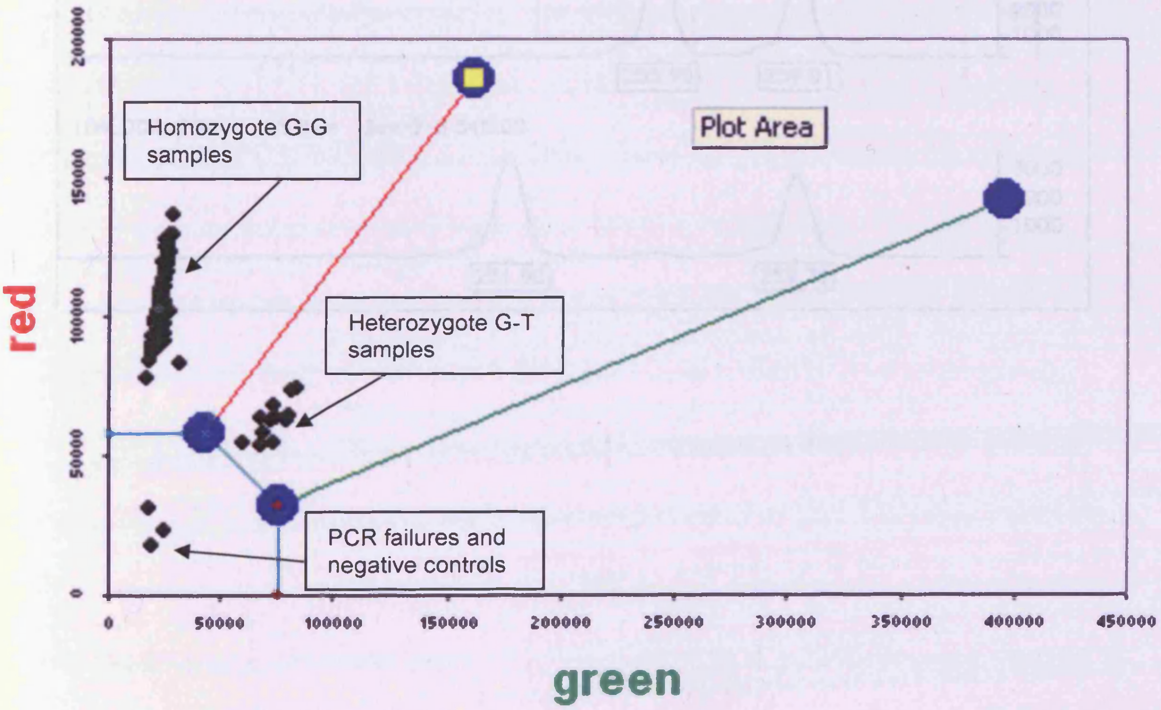
Allele 1 primer **red:**
GAAGGTCGGAGTCAACGGATTTAATGCCCCTCCACTCCCG
Allele 2 primer **green:**
GAAGGTGACCAAGTTCATGCTAATGCCCCCTCCACTCCCA
Common Reverse primer:
ATCATCTCTGCCCGCTCACT

1960 G>T Amplifluor assay:

Allele 1 primer **red:**
GAAGGTCGGAGTCAACGGATTTTCTGCCTGCCCCTCGGCC
Allele 2 primer **green:**
GAAGGTGACCAAGTTCATGCTTCTTCTGCCTGCCCCTCGGCA
Common Reverse primer:
AGGGAGACGGAAAGAGACAAGA

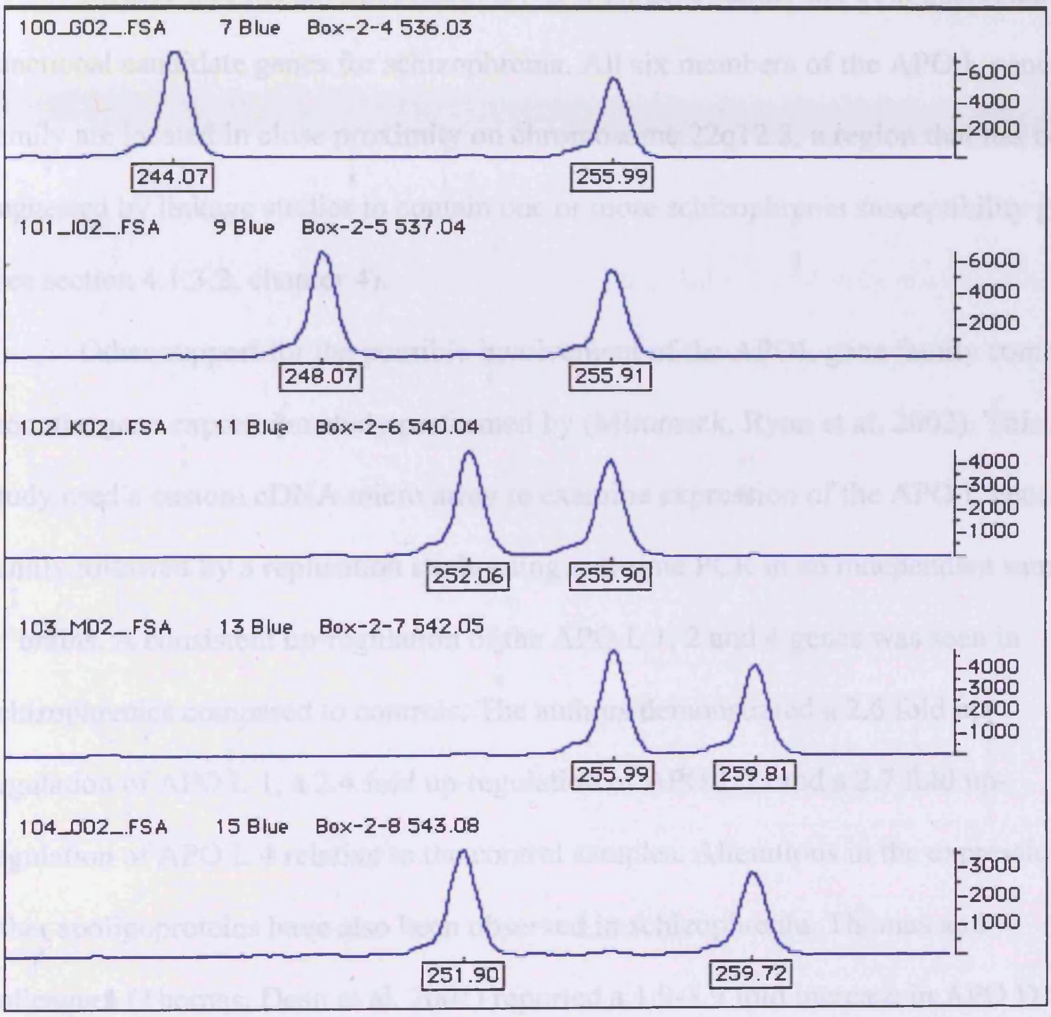
Appendix 5.5:

Example of an Amplifluor assay graph produced whilst genotyping the GFAP 1960 G>T SNP in the schizophrenia association sample.



Appendix 5.6:

Example of the Genescan electrophorograms produced when genotyping the MOG VNTR polymorphism. The numbers below the electrophorograms correspond to the fragment sizes in base pairs. When a electrophorogram has two peaks it represents a heterozygote of the two peaks corresponding repeat lengths.



Chapter 6. The APO L gene family and schizophrenia.

6.1 The APO L gene family:

Members of the apolipoprotein L (APO L) gene family are both positional and functional candidate genes for schizophrenia. All six members of the APO L gene family are located in close proximity on chromosome 22q12.3, a region that has been suggested by linkage studies to contain one or more schizophrenia susceptibility genes (see section 4.1.3.2, chapter 4).

Other support for the possible involvement of the APOL gene family comes from the gene expression study performed by (Mimmack, Ryan et al. 2002). This study used a custom cDNA micro array to examine expression of the APO L gene family followed by a replication study using real-time PCR in an independent sample of brains. A consistent up-regulation of the APO L 1, 2 and 4 genes was seen in schizophrenics compared to controls. The authors demonstrated a 2.6 fold up-regulation of APO L 1, a 2.4 fold up-regulation of APO L 2, and a 2.7 fold up-regulation of APO L 4 relative to the control samples. Alterations in the expression of other apolipoproteins have also been observed in schizophrenia. Thomas and colleagues (Thomas, Dean et al. 2001) reported a 1.9-3.9 fold increase in APO D expression levels in the dorso lateral prefrontal cortex (DLPFC) of schizophrenic patients compared to controls, and also in the prefrontal cortex of bipolar patients. The function of APO D is not clear but it has been proposed to play a neuroprotective role, to be involved with lipid transport in the CNS during maintenance and tissue repair, and also to play a role in the transport of cholesterol in the CNS (Thomas, Dean et al. 2001). A further study (Dean, Laws et al. 2003) reported a significant increase in the levels of APO E protein in the prefrontal cortex of schizophrenic patients compared to controls. When taken together these three studies on APO D, E and L which all report increased levels of their respective apolipoproteins in the prefrontal cortex of schizophrenic patients, would support the hypothesis that apolipoproteins play a role in the aetiology of schizophrenia.

APO L proteins belong to the high density lipoproteins. Monajemi and colleagues (Monajemi, Fontijn et al. 2002) have shown that the APO L gene family has arisen relatively recently in evolution, with APO Ls 1-4 probably arising as a result of tandem duplication events. The APO L gene family are thought to be involved in the transport of high-density lipoproteins (HDLs) in cell membranes and to have an important role in the development and maintenance of cell membrane structure and function (Sutcliffe and Thomas 2002; McGhee, Morris et al. 2005). Lipoproteins provide cells with lipids vital to normal cell functioning. Lipids are used to generate energy, as building blocks for biomembranes and play an important role in cell signalling (Danik, Champagne et al. 1999). Most lipoproteins are widely expressed in tissues, including the brain (Danik, Champagne et al. 1999).

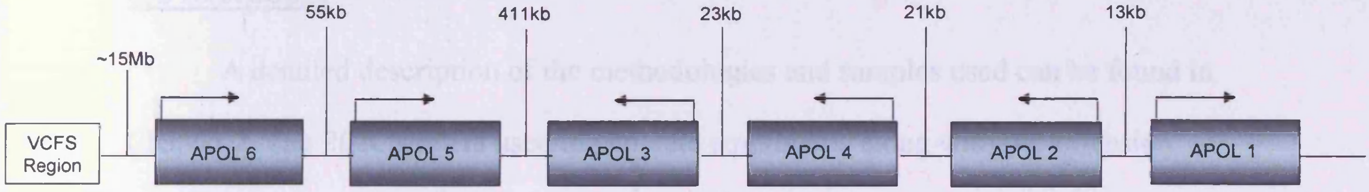
APO L proteins are functionally associated with apolipoprotein A1 (APO A1), the main protein component of high density lipoprotein (Sutcliffe and Thomas 2002). APO A1 plays an important role in cholesterol transport (Dietschy and Turley 2001; Sutcliffe and Thomas 2002). Given their association with APO A1, APO L proteins are also thought to be of importance in cholesterol transport. The cholesterol content of membranes is important in cellular processes such as modifying gene transcription and signal transduction both in the adult brain and during neurodevelopment (Mimmack, Ryan et al. 2002). The nervous system is thought to contain almost a quarter of the total body content of unesterified cholesterol. These cholesterols are largely located in two pools, the membranes of glial cells and neurons and secondly in myelin sheaths (Dietschy and Turley 2001). Brain cholesterol is thought to be almost exclusively synthesized in situ in the brain (Dietschy and Turley 2001). It has been proposed that any changes in the cholesterol content of the CNS may affect the integrity of neurons and myelin (Dietschy and Turley 2001).

It is thought that myelin is assembled by vesicular trafficking of membranes that contain myelin specific proteins, glycolipids and cholesterol. It has been suggested that myelin membranes result from the accumulation of 'myelin-specific rafts' in which cholesterol is closely related with myelin specific membrane proteins (Saher, Brugger et al. 2005). A myelin sheath is essentially a spiral extension of the oligodendrocytes plasma membrane and appears in cross-section as a compact, multilayered stack of membranes. In addition to the myelin specific membrane proteins, myelin contains a very high content of lipids (about 70% of its dry weight). Of this lipid content, around 25% is cholesterol, a figure which is higher than the 20% content level that is seen in normal plasma membranes (Saher, Brugger et al. 2005). In a recent study, Saher and colleagues (Saher, Brugger et al. 2005) reported that cholesterol plays an essential function in central nervous system (CNS) myelination. They reported that mutant mice lacking the ability to synthesize cholesterol in oligodendrocytes had severely defective myelin synthesis. However the mutant mice survived as mutant oligodendrocytes were able to use cholesterol from neighbouring wild-type cells of other types. The authors showed that myelin cannot be synthesized without cholesterol, not even at the slightly reduced levels of cholesterol content found in normal plasma membranes. They therefore concluded that cholesterol availability is a critical prerequisite and an essential rate limiting factor of myelin membrane growth during brain maturation.

Given the location of the APOL gene family on chromosome 22q12.3, the evidence pointing to the altered expression of the genes in schizophrenia and the possible involvement of the gene family in cholesterol transport, a vital component of successful myelination and brain development, I decided to study the gene family.

Figure 6.1: Genomic organisation of the APO L gene family.

Schematic diagram showing the genomic organisation of the APOL gene family on chromosome 22q12.3 the genes position relative to the VCFS region and the direction of transcription (indicated by the arrows).



6.1.2 Strategy for this study;

A different approach has been applied in the study of the APOL genes than has been employed in the rest of this project. Like the other genes studied the APO Ls have been proposed as candidates essentially on the basis of their altered expression in tissues originating from schizophrenic patients compared to controls and their location in a region of the genome previously implicated in schizophrenia. Thus the first step was to look for evidence of polymorphic *cis*-acting influences on the genes expression. This was to determine whether the expression differences previously seen in micro-array studies were potentially attributable to common genetic variation or were caused by other environmental or *trans*-acting factors. Where these initial studies demonstrated that expression of the genes was influenced by *cis*-acting sequence variation, genetic studies were performed to test the genes for association with schizophrenia. Unlike the other genes studied as part of this thesis the APO Ls were tested for association by genotyping the case control pools through a 5kb grid of marker SNP's spanning the gene. This indirect marker grid approach was used rather than the direct method employed for other genes in this thesis as the genetic variation causing expression differences is likely to be located outside exonic sequence. By using a grid approach the aim was to detect association to the true susceptibility variants by virtue of linkage disequilibrium (LD) between these and the marker SNPs.

6.2 Methods:

A detailed description of the methodologies and samples used can be found in Chapter 2. The PCR primers used to generate amplimers, along with the extension primers used to perform primer extension reactions for the allelic expression and association analysis of each gene, are listed below in tables 6.1, 6.2, 6.3, 6.4 and 6.5 respectively.

Table 6.1: PCR and primer details for the APOL 1 assays:

This table shows the rs numbers of the SNPs used in the APOL 1 assays, the forward and reverse PCR primers used to amplify the sequence around these polymorphisms, the size of the PCR fragment and the PCR conditions used. It also shows the sequence flanking each of the SNPs, the extension primers used in the SNaPshot reaction and the expected alleles in the reaction.

Fragment Name	rs number	Forward PCR Primer (5' to 3')	Reverse PCR Primer (5' to 3')	Size bp	Annealing Temp	SNP Flanking sequence	Extension primer (5' to 3')	Expected Alleles
Apol 1_A	rs2227089	tgtgttccaaaattgtacgaga	tccgatggcacaatctgtag	230bp	HS 60	ctttac[A/T]tctcaa	tttgttatctgcaaaaagtgtttactttac	A (Green) & T (Red)
Apol 1_B	rs136150	ccccatgttgtcttgaggat	gaggcagtcatccctggtaa	215bp	HS 60	ggctggt[A/T]cccatc	acaggggaccagtagaggggatggg	T (Red) & A (Green)
Apol 1_C	rs136154	aagggagaatgcagagacca	ccggccacctcattcttaat	194bp	HS 60	ttcttc[A/T]tctgtg	ctctcgctgtctggattttactttcttc	A (Green) & T (Red)
Apol 1_D	rs713753	aaggccttagaggcaactcc	tcccactcacatgtcctctg	243bp	HS 60	aaccaa[C/T]accaag	cctttgggcaggatcttggt	G (Blue) & A (Green)
Apol 1_Exp	rs2239785	aggcagatgagctccgtaaa	tgagagagccagacaccaca	231bp	HS 60	gaaaagt[A/G]agcttg	agccttcttatgttatcctcaagct	T (Red) & C (Black)
Apol 1_E	rs2012928	cagctctccccatcttgtgt	aggtctcagggaaatgacca	220bp	HS 60	caatca[A/G]cactga	gtcctggagacttgggtccaatca	A (Green) & G (Blue)

Table 6.2: PCR and primer details for the APOL 2 assays:

This table shows the rs numbers of the SNPs used in the APOL 2 assays, the forward and reverse PCR primers used to amplify the sequence around these polymorphisms, the size of the PCR fragment and the PCR conditions used. It also shows the sequence flanking each of the SNPs, the extension primers used in the SNaPshot reaction and the expected alleles in the reaction.

Fragment Name	rs number	Forward PCR Primer (5' to 3')	Reverse PCR Primer (5' to 3')	Size bp	Annealing Temp	SNP Flanking sequence	Extension primer (5' to 3')	Expected Alleles
Apol 2_A	rs1974458	ctgccactgctttcctgtct	cggaggacgtgtctggttat	245bp	HS 60	attgca[A/G]gaataa	gtgatctacccccatgtccttattc	T (Red) & C (Black)
Apol 2_B	rs1557533	ggggaaacagaccaagatga	tggaggatttccgtgacttt	249bp	HS 60	ttttct[A/G]gaaact	gtttatttgatcatcaccttttagtttc	T (Red) & C (Black)
Apol 2_C	rs2157249	agaactcacgcactgcctct	tgaactcctcccctctctga	225bp	HS 60	tcagaga[C/T]gccaag	aatgagggcaacaggagatcagaga	C (Black) & T (Red)
Apol 2_D	rs2003813	ctgcaggtcaacctcctctc	ctgcatcctgacctctcctc	187bp	HS 60	ctgggc[C/T]gtgtca	tcctgacctctcctcggggtgacac	G (Blue) & A (Green)
Apol 2_E	rs5995263	gctccatgcgatctctatcc	gtgttgacctgcccttgagt	234bp	HS 60	atggaa[A/T]taatta	caatcaaatagaaacaaattttacaaatggaa	A (Green) & T (Red)
Apol 2_Exp	rs2017689	actggagagtgtggggaatg	tacacccccagaaggacaag	295bp	HS 60	ggggtc[A/G]gtagtg	gagaagaaacagcagctccactac	T (Red) & C (Black)
Apol 2_F	rs2016988	agttcctgggtggatcactg	tcagccctcatggtcttttc	238bp	HS 60	agtttg[C/G]aaaaac	aaagcagacatctacctctagtttg	C (Black) & G (Blue)

Table 6.3: PCR and primer details for the APOL 3 assays:

This table shows the rs numbers of the SNPs used in the APOL 3 assays, the forward and reverse PCR primers used to amplify the sequence around these polymorphisms, the size of the PCR fragment and the PCR conditions used. It also shows the sequence flanking each of the SNPs, the extension primers used in the SNaPshot reaction and the expected alleles in the reaction.

Fragment Name	rs number	Forward Primer (5' to 3')	Reverse Primer (5' to 3')	Size bp	Annealing Temp	SNP Flanking sequence	Extension primer (5' to 3')	Expected Alleles
Apol 3_A	rs132673	tggatgaagctgatccctta	tgctcctgctcatactggaa	141bp	HS 60	tttact[A/G]ggaggc	tagttggtgtgagtgggaagcctcc	T (Red) & C (Black)
Apol 3_B	rs132661	gcagaaatcattccgctctc	ctgcactgggatggattctt	183bp	HS 60	gagagt[A/G]tattac	gcaaagtgtattgaagcacagtaata	T (Red) & C (Black)
Apol 3_C	rs132656	ctggaaaagtcacgctcctc	ccctgaccccttcttcctaa	230bp	HS 60	gggggg[C/T]ggatgt	gtctttgtcttttactcccacatcc	G (Blue) & A (Green)
Apol 3_D	rs2097465	gtcctccttaggcgttgttg	ggagctgctttgctgagagt	255bp	HS 60	gtcctc[C/T]ctgggc	ttgaccaaactgcaggcttgcccag	G (Blue) & A (Green)
Apol 3_E	rs2017329	caggaggtggagtttgtggt	tcccttcccaattaaaagca	138bp	HS 50 DMSO	agaaaa[A/T]aaaaaa	gctaattttgatgttcagattttccttttttt	A (Green) & T (Red)
Apol 3_F	rs132644	ggctcccattttgaaacaga	cacttgggcatagagtgcaa	216bp	HS 52 DMSO	tgctta[C/T]tgtaaa	aaaatattgactcatgcatacatgctta	C (Black) & T (Red)
Apol 3_G	rs132635	cagcctgtggcgttttacta	ccactgatgggtcgaaactc	238bp	HS 60	accatg[C/T]tttctc	tggaatatttcccgtacatggggagaaa	G (Blue) & A (Green)
Apol 3_H	rs80578	tctctttggttggggacttg	gtgagaagagggaggcactg	321bp	HS 60	tcagaa[C/T]gcccca	gtcccaaatggcagttcagaa	C (Black) & T (Red)
Apol 3_Exp	rs132618	ttacgaagccacacaaacca	cacatccagtgcaaggaaga	211bp	HS 60	ttgacc[A/T]ccactt	cgaatctcagctggaagtgg	T (Red) & A (Green)
Apol 3_J	rs132617	gccctcctcctgtgacataa	ccaaacagccacctgaattt	219bp	HS 60	agcaag[C/T]gattga	agaagaagagtgtgtgtgcagtagcaag	C (Black) & T (Red)

Table 6.4: PCR and primer details for the APOL 4 assays:

This table shows the rs numbers of the SNPs used in the APOL 4 assays, the forward and reverse PCR primers used to amplify the sequence around these polymorphisms, the size of the PCR fragment and the PCR conditions used. It also shows the sequence flanking each of the SNPs, the extension primers used in the SNaPshot reaction and the expected alleles in the reaction.

Fragment Name	rs number	Forward PCR Primer (5' to 3')	Reverse PCR Primer (5' to 3')	Size bp	Annealing Temp	SNP Flanking sequence	Extension primer (5' to 3')	Expected Alleles
Apol 4_A	rs132743	ggttcaaatggggtgaaatg	ggacttgggcaacacattct	237bp	HS 62	gaaaat[C/G]taattt	agggagaaagtgaggaagaaaatta	G (Blue) & C (Black)
Apol 4_B	rs2005590	caggcctgctcttgttgaag	tctagtccctccaccctgtg	224bp	HS 62	gctcat[C/T]gttcag	caccacctggaaaaatcatgctcat	C (Black) & T (Red)
Apol 4_C	rs1812023	ggagggagcaatcagactca	accagctggatcttgctcag	226bp	HS 62	gcctgg[A/G]tgcaaa	agtggtagagccgggatttgca	T (Red) & C (Black
Apol 4_D	rs132734	ccatcctccttggtcattgt	ggagccatgtgggtttttct	182bp	HS 62	gacaca[A/G]tgcgtc	gactgttagcctcaactaggacaca	A (Green) & G (Blue)
Apol 4_E	rs6000179	acaaggctaaggtccggtct	ccaaactggaggaggttttg	215bp	HS 62	tttatc[A/C]ctggaa	aattaggacattatttgattagttccag	T (Red) & G (Blue)
Apol 4_F	rs2007468	gaggttacctgggcaattca	ggaggattaacaccgaacca	226bp	HS 62	acgcac[A/G]aatctc	gcgatgaagcctggaagagatt	T (Red) & C (Black
Apol 4_Exp	rs132700	gaaaggcttcgtgtcattgc	caattggtcagtgctggttg	279bp	HS 60	ccaaca[C/T]aacgcc	gcatcctgtctgtcattggcgtt	G (Blue) & A (Green)
Apol 4_G	rs2227167	ggttccgggctacttttctc	actgaccaattggaggcatt	243bp	HS 62	tgcaaa[A/G]gaaagc	tggcttcgtcaaaatcaagtgcaaa	A (Green) & G (Blue)
Apol 4_H	rs916336	ttggaagcttttgcaggagt	caggattgacatgggtcaca	185bp	HS 62	ctgtct[C/G]tctctg	tccatgagacatagagggacagaga	G (Blue) & C (Black)
Apol 4_J	rs1966267	acagaccatgagggctgaat	aggtgtgctccactgctctt	165bp	HS 62	ctcaac[A/G]tctctc	ggaaagcacagcagaatttgagaga	T (Red) & C (Black
Apol 4_K	rs132691	ggagcacctgctgtcatctt	ggcgaggacgaaatttatca	203bp	HS 62	cgctgg[C/G]acctga	gtggaaaggggcaaggaacgctgg	C (Black) & G (Blue)
Apol 4_L	rs132683	ttatcaggcccatacccaaa	ttggccttttatggggatta	188bp	HS 62	tatgct[A/G]gtttat	tgattagtacaggaagggtataaac	T (Red) & C (Black

Table 6.5: PCR primer details for APOL 5 assays:

This table shows the rs numbers of the SNPs used in the APOL 5 assays, the forward and reverse PCR primers used to amplify the sequence around these polymorphisms, the size of the PCR fragment and the PCR conditions used. It also shows the sequence flanking each of the SNPs, the extension primers used in the SNaPshot reaction and the expected alleles in the reaction.

Fragment Name	rs number	Forward PCR Primer (5' to 3')	Reverse PCR Primer (5' to 3')	Size bp	Annealing Temp	SNP Flanking sequence	Extension primer (5' to 3')	Expected Alleles
Apol 5_A	rs1540297	atgggtccagttttgtggtc	caagccagaagggtaaaacg	197bp	HS 62	tgtggt[C/T]gagaat	cactgttatctggttgtgtggt	C (Black) & T (Red)
Apol 5_B	rs5999971	tgcaggagaactaggggtga	agtccctgctgaatgacacc	238bp	HS 62	aagaag[C/T]tgcaat	ttcagacactctctctcacattgca	G (Blue) & A (Green)
Apol 5_C	rs2009168	ttgctggttttctcccaatc	ggcagcagagtttgttgtga	196bp	HS 62	gctgca[A/G]caatag	ctctcccacagagctcagctgca	A (Green) & G (Blue)
Apol 5_D	rs2076671	atccaactctggcttcatgg	tgctctcagttcctctgctg	246bp	HS 62	tctgga[C/T]ggctag	ctctctgcacccctctagcc	G (Blue) & A (Green)
Apol 5_Exp	rs2076672	tggggctggcttcttactta	cacacgggatcctcgaac	226bp	HS 60	caagga[C/T]ggagac	cctggaggatggggcaagga	C (Black) & T (Red)
Apol 5_E	rs2016586	catgccagacctagcagtga	gttcagtggtgacagggtga	197bp	HS 62	gaaaag[G/T]actagg	tcattcttcactttgcacacctagt	C (Black) & A (Green)
Apol 5_F	rs5995165	gttggaaacaaccctctgga	ctttccctcacaccacaggt	150bp	HS 62	gagggg[A/G]tggtga	ccacactggttcaggtcacca	T (Red) & C (Black)
Apol 5_G	rs879680	gaaaggcctgtgcaacctaa	aaggaggcaagtgtgcaaag	226bp	HS 62	aattgc[C/T]gtctgt	aaattggtttcccggctcacagac	G (Blue) & A (Green)

6.3 Results:

6.3.1 Allelic Expression assays:

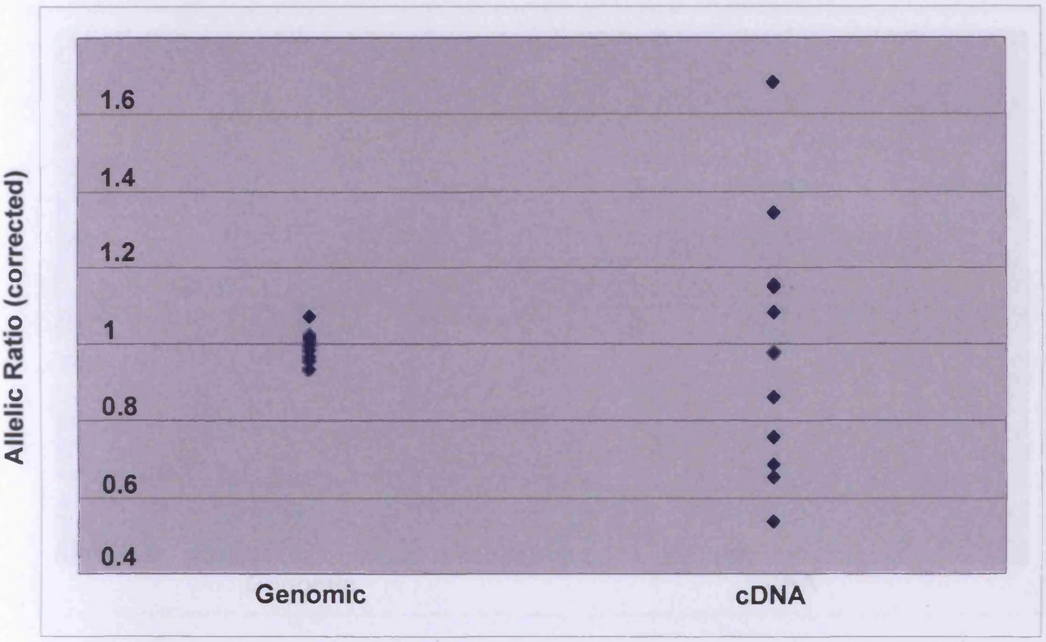
Of the six genes in the APOL gene family, only APOL 1-5 contained a suitable database polymorphism at the time of this study for use as a tag SNP in the allelic expression assay. The known polymorphisms in APOL 6 were either not in the exonic sequence or did not have an informative minor allele frequency. Therefore the gene was not screened for altered allelic expression.

APOL 1:

The G>A polymorphism rs2239785 was selected for analysis. Of the 60 subjects for whom brain was available, 11 were heterozygous. The 11 heterozygotes yield a power of approximately 67% to detect the effect of a regulatory variant present at a frequency of 0.05, and a power of 89% to detect a variant at a frequency of 0.1. Of the 11 heterozygous individuals, six showed a relative expression difference of ≥ 20% (see Graph 6.1). The data points from each individual cDNA sample showed moderate reproducibility, with an average SD of the mean of 0.13. The cDNA allelic expression ratio has departed from the genomic 1:1 ratio in both directions with the expression difference ranging from a 68% increase of the A allele relative to the G allele to an 88% decrease of the A allele relative to the G allele. That the expression differences go in both directions on the graph might reflect the effect of a *cis*-acting variant that is only in low LD with the tag SNP that was used or that there is more than one variant affecting the genes expression.

Graph 6.1: Apol 1 Allelic Expression, rs2239785.

Corrected genomic and cDNA ratios for rs2239785. Comparison between the observed corrected genomic ratios and the corrected cDNA ratios assayed for APOL 1 (n=11). Data are expressed as the mean of the ratio of G:A for two measurements for each genomic DNA sample and four measurements of each cDNA sample.

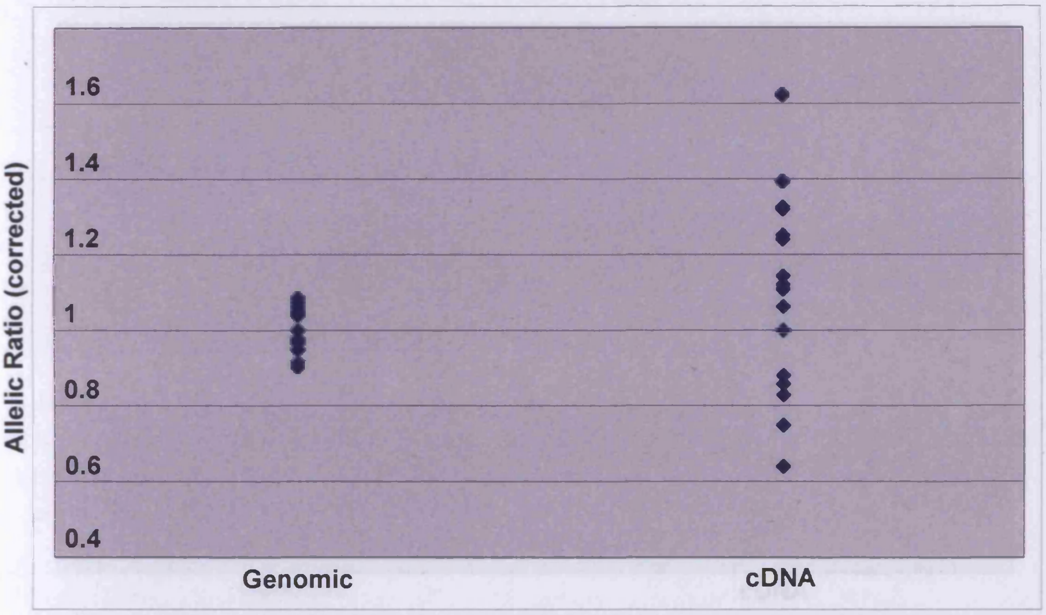


APOL 2:

The A>G polymorphism rs2017689 was selected for analysis. The 16 heterozygotes for this marker yields a power of approximately 80% to detect the effect of a regulatory variant present at a frequency of 0.05, and a power of 96% to detect a variant at a frequency of 0.1. Of the 16 heterozygous individuals, nine showed a relative expression difference of ≥ 20% compared to the corresponding average genomic ratios (see Graph 6.2). The data points from each individual cDNA sample showed good reproducibility, with an average SD of the mean of 0.08. The allelic expression ratio has departed from the genomic 1:1 ratio in both directions with the expression difference ranging from a 62% increase of the A allele relative to the G allele to a 56% decrease of the A allele relative to the G allele.

Graph 6.2: Apol 2 Allelic Expression, rs2017689.

Corrected genomic and cDNA ratios for rs2017689. Comparison between the observed corrected genomic ratios and the corrected cDNA ratios assayed for APOL 2 (n=16). Data are expressed as the mean of the ratio of G:A for two measurements for each genomic DNA sample and four measurements of each cDNA sample.

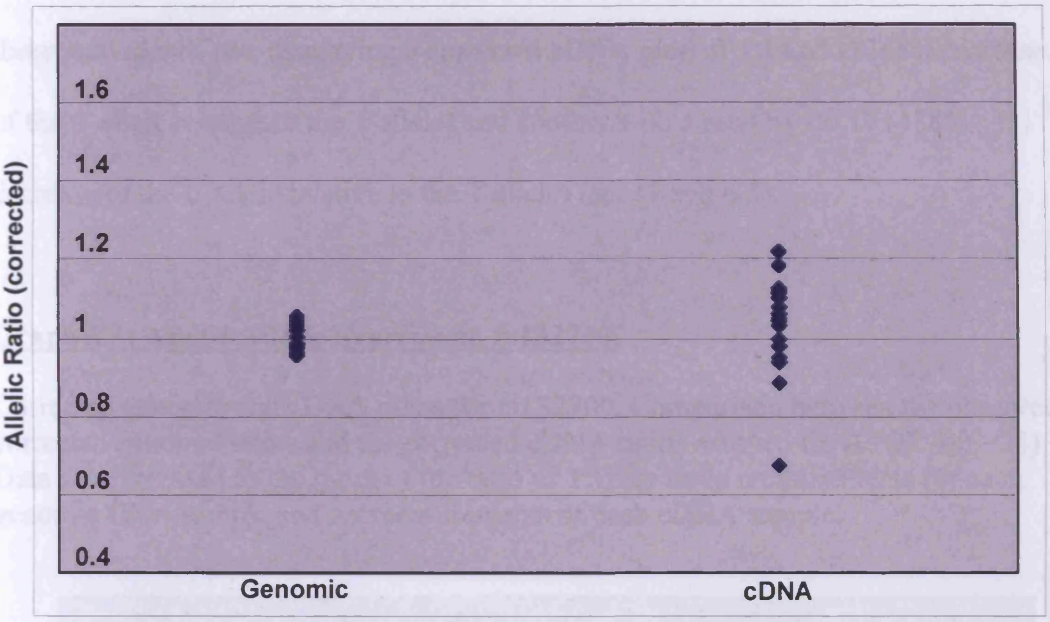


APOL 3:

The T>A polymorphism rs132618 was selected for analysis. Of the 60 subjects for whom brain was available, 20 were heterozygous. The 20 heterozygotes yields a power of approximately 86% to detect the effect of a regulatory variant present at a frequency of 0.05, and a power of 98% to detect a variant at a frequency of 0.1. Of the 20 heterozygous individuals, two showed a relative expression difference of ≥ 20% compared to the corresponding average genomic ratios (see Graph 6.3). The data points from each individual cDNA sample showed reasonable reproducibility, with an average SD of the mean of 0.11. The cDNA ratio departed from the genomic 1:1 ratio in both directions with the expression difference ranging from a 21% increase of the T allele relative to the A allele to a 49% decrease of the T allele relative to the A allele.

Graph 6.3: Apol 3 Allelic Expression, rs132618.

Corrected genomic and cDNA ratios for rs132618. Comparison between the observed corrected genomic ratios and the corrected cDNA ratios assayed for APOL 3 (n=20). Data are expressed as the mean of the ratio of A:T for two measurements for each genomic DNA sample and four measurements of each cDNA sample.

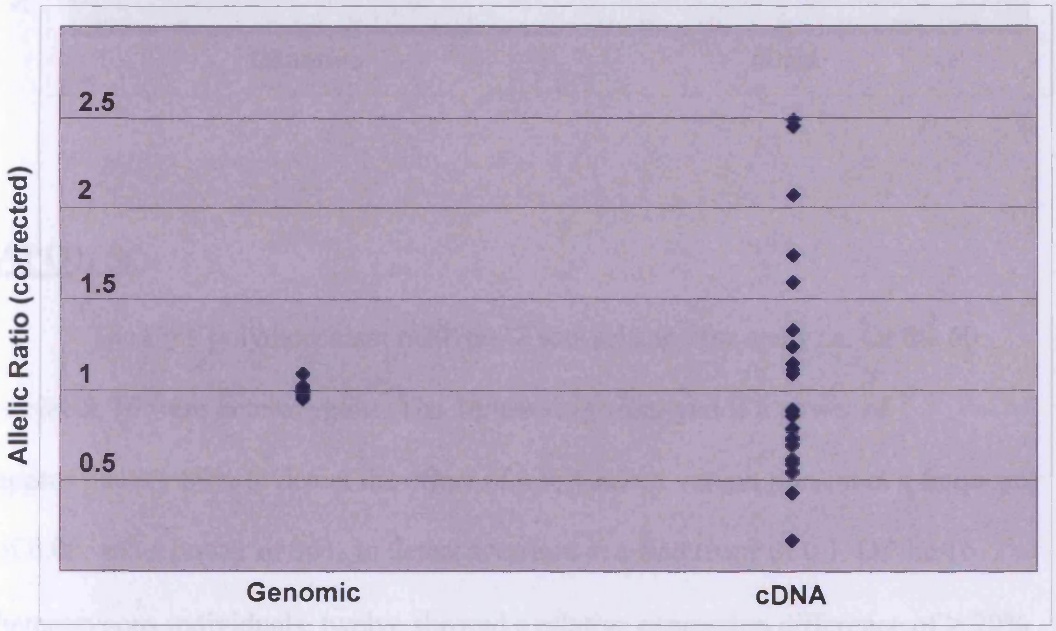


APOL 4:

The C>T polymorphism rs132700 was selected for analysis. The 24 heterozygotes for this marker yields a power of approximately 91% to detect the effect of a regulatory variant present at a frequency of 0.05, and a power of 99% to detect a variant at a frequency of 0.1. Of the 24 heterozygous individuals, seventeen showed a relative expression difference of ≥ 20% compared to the corresponding average genomic ratio (see Graph 6.4 and 6.5). The data points from each individual cDNA sample showed poor reproducibility, with an average SD of the mean of 0.48. Given the observation of almost mono-allelic expression for some individuals it was necessary to remove these samples when plotting the standard allelic expression assay graph (graph 6.4) to enable the observation of the smaller relative expression differences present. The samples displaying large expression differences were then visualised by plotting all of the assayed individuals on an allelic expression log scale graph (graph 6.5). The cDNA ratio has departed from the genomic 1:1 ratio in both directions with evidence of nearly complete mono-allelic expression observed for three individuals, two displaying a corrected cDNA ratio of 1:14.65 (1368% increase of the C allele relative to the T allele) and another with a ratio of 1:0.17 (488% decrease of the C allele relative to the T allele) (see Graph 6.5).

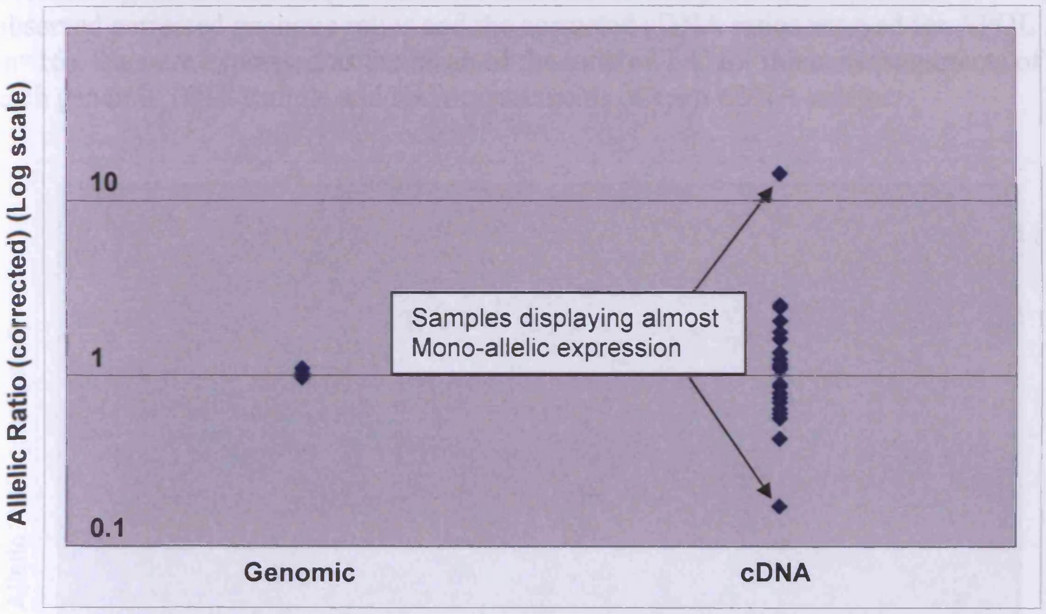
Graph 6.4: Apol 4 Allelic Expression, rs132700.

Corrected genomic and cDNA ratios for rs132700. Comparison between the observed corrected genomic ratios and the corrected cDNA ratios assayed for APOL 4 (n=24). Data are expressed as the mean of the ratio of T:C for three measurements for each genomic DNA sample and six measurements of each cDNA sample.



Graph 6.5: Apol 4 Allelic Expression, Log scale of rs132700 assay.

The graph shows the corrected genomic and cDNA ratios for rs132700 plotted on a log scale graph. Comparison between the observed corrected genomic ratios and the corrected cDNA ratios assayed for APOL 4 (n=24). Data are expressed as the mean of the ratio of T:C for three measurements for each genomic DNA sample and six measurements of each cDNA sample. The three individuals showing evidence of mono-allelic expression are indicated at the top and bottom of the cDNA ratio scale.



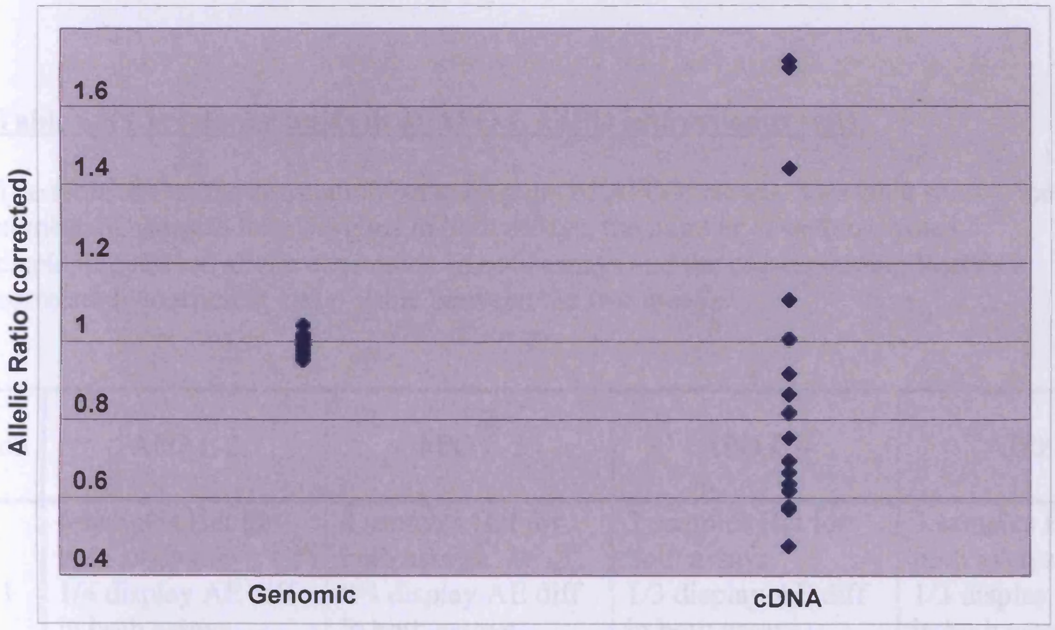
APOL 5:

The C>T polymorphism rs2076672 was selected for analysis. Of the 60 subjects, 16 were heterozygous. The 16 heterozygotes yields a power of approximately 80% to detect the effect of a regulatory variant present at a frequency of 0.05, and a power of 96% to detect a variant at a frequency of 0.1. Of the 16 heterozygous individuals, twelve showed a relative expression difference of ≥ 20% compared to the corresponding average genomic ratios (see Graph 6.6). The data points from each individual cDNA sample showed poor reproducibility, with an average SD of the mean of 0.57. The cDNA ratio departed from the genomic 1:1 ratio in both directions with the relative expression differences ranging from a 72%

increase of the C allele relative to the T allele, to a 108% decrease of the C allele relative to the T allele.

Graph 6.6: Apol 5 Allelic Expression, rs2076672.

Corrected genomic and cDNA ratios for rs2076672. Comparison between the observed corrected genomic ratios and the corrected cDNA ratios assayed for APOL 5 (n=16). Data are expressed as the mean of the ratio of T:C for three measurements of each genomic DNA sample and six measurements of each cDNA sample.



6.3.1.1 Correlation of allelic expression assay results:

Given the evidence for *cis*-acting sequence variants affecting the expression of all five APO L genes, I analysed the data to see if the observed relative allelic expression differences were correlated between assays, indicating a polymorphism in a master control region. To achieve this I analysed samples which were heterozygous for pairs of assays and tested these for a correlation in the relative expression differences observed. As the phase of the markers was unable to be established the allelic expression differences would not be expected in one direction. Therefore the analyses were performed after converting all of the relative expression differences to be positive. This revealed no evidence of a correlation between the relative expression differences observed for any combination of two assays (table 6.6). Interestingly there was no correlation detected between the relative expression for any of the genes transcribed in the same direction (for example, see graph 6.7). There was also no evidence revealed of any samples displaying deviation/non-deviation in allelic expression across all five assays.

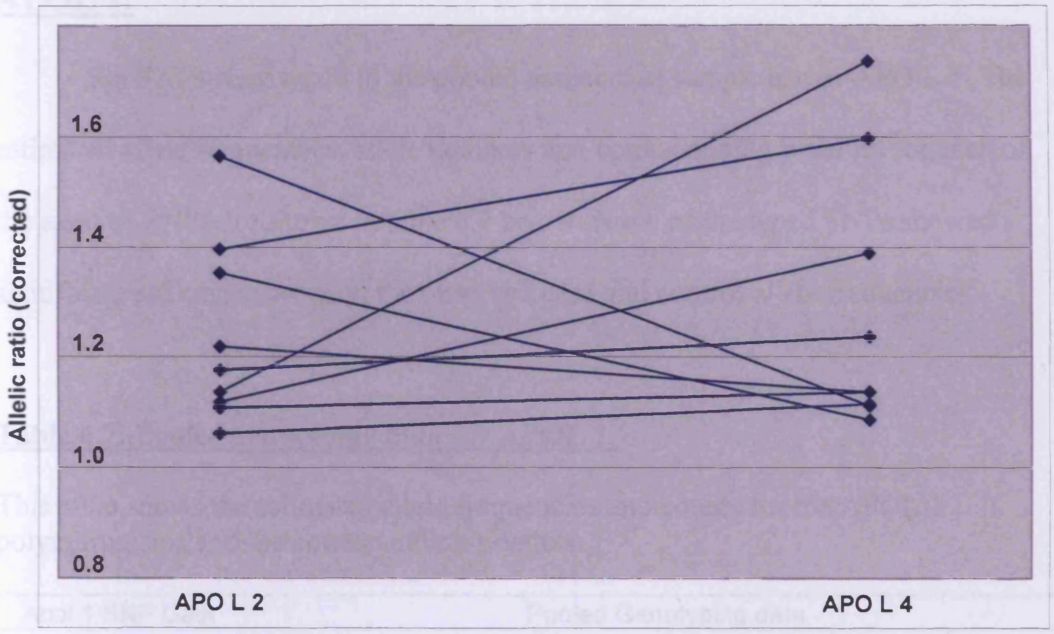
Table 6.6 Correlation analysis of APO L allelic expression results.

The table shows the correlation between pairs of APO L assays. The table shows, the number of samples heterozygous in both assays, the number of heterozygotes displaying altered allelic expression in both assays and the corresponding Pearsons correlation coefficient and p-value between the two assays.

Allelic Expression Assay	APO L 2	APO L 3	APO L 4	APO L 5
APO L 1	4 samples Het for both assays. 1/4 display AE diff in both assays. corr= 0.19, p= 0.81	4 samples Het for both assays. 0/4 display AE diff in both assays. corr= 0.50, p=0.49	3 samples Het for both assays. 1/3 display AE diff in both assays. corr= 0.51, p= 0.66	3 samples Het for both assays. 1/3 display AE diff in both assays. corr= 0.91, p= 0.27
APO L 2		7 samples Het for both assays. 0/7 display AE diff in both assays. corr= 0.33, p= 0.47	9 samples Het for both assays. 1/9 display AE diff in both assays. corr= 0.11, p= 0.78	4 samples Het for both assays. 0/4 display AE diff in both assays. corr= 0.88, p= 0.12
APO L 3			8 samples Het for both assays. 0/8 display AE diff in both assays. corr= 0.19, p= 0.66	5 samples Het for both assays. 0/5 display AE diff in both assays. corr= 0.25, p= 0.69
APO L 4				9 samples Het for both assays. 7/9 display AE diff in both assays. corr= 0.21, p= 0.59

Graph 6.7. APO L 2 and 4 Correlation analysis.

Graph displaying the lack of correlation between the relative allelic expression differences observed for APO L genes transcribed in the same direction. Corrected cDNA ratios are plotted for samples heterozygous in both the APO L 2 and APO L 4 allelic expression assays, displaying no evidence of a correlation between assays.



6.3.2 Indirect genetic analysis of the APOL gene family:

Given the results of the allelic expression assays indicating the presence of *cis*-acting sequence variants causing altered allelic expression in the APO L genes, I decided to study the gene family further by screening the genes to look for evidence of association with schizophrenia. This was achieved by designing a grid of marker SNPs approximately 5kb apart spanning each of the genes (see figure 6.2) and then typing these SNPs in the pooled association sample to test for association with schizophrenia. All of the grids were designed to include the marker SNP used in the allelic expression assays. As well as typing all of the SNPs in the pooled association sample the SNPs were additionally typed in 96 unrelated control samples in order to calculate D' and r^2 values and the LD structure of all of the markers typed across the APO L genomic region (see figure 6.2).

APOL 1:

Six SNPs were typed in the pooled association sample across APO L 1. The estimated allele frequencies, allele numbers and corresponding p-values for each of the assayed SNPs are shown in table 6.7 below. None of the typed SNPs showed a significant difference between the observed case and control allele frequencies.

Table 6.7: Pooled genotyping data for APOL 1.

This table shows the estimated allele frequencies and counts for the APOL 1 polymorphisms and the corresponding p-values.

Apol 1 SNP Data	Pooled Genotyping data				
SNP number.	Alleles Cases (freq)	Alleles Controls (freq)	Difference	Chi-square	P-value
rs2227089	A= 678 (0.61) T= 426 (0.38)	A= 677 (0.61) T= 427 (0.38)	0	0.002	0.97
rs136150	T= 259 (0.23) A= 845 (0.77)	T= 237 (0.22) A= 867 (0.78)	0.01	1.26	0.67
rs136154	A= 312 (0.23) T= 1070 (0.77)	A= 305 (0.21) T= 1129 (0.79)	0.02	0.71	0.41
rs713753	C= 451 (0.41) T= 653 (0.59)	C= 422 (0.38) T= 682 (0.62)	0.03	1.59	0.21
rs2239785	G= 212 (0.2) A= 892 (0.8)	G= 191 (0.17) A= 913 (0.83)	0.03	1.34	0.25
rs2012928	G= 877 (0.79) A= 227 (0.21)	G= 881(0.8) A= 223 (0.2)	0.01	0.05	0.83

APOL 2:

Seven SNPs were typed in the pooled association sample across APOL 2. The estimated allele frequencies and allele numbers for each of the assayed SNPs is shown in table 6.8 below. None of the SNPs showed a significant difference between the observed case and control allele frequencies.

Table 6.8: Pooled genotyping data for APOL 2.

This table shows the estimated allele frequencies and counts for the APOL 2 polymorphisms and the corresponding p-values.

Apol 2 SNP Data	Pooled Genotyping data				
SNP number.	Alleles Cases (freq)	Alleles Controls (freq)	Difference	Chi-square	P-value
rs1974458	A= 819 (0.74) G= 285 (0.26)	A= 800 (0.72) G= 304 (0.28)	0.02	0.84	0.36
rs1557533	A= 613 (0.56) G= 491 (0.44)	A= 603 (0.55) G= 501 (0.45)	0.01	0.18	0.67
rs2157249	C= 404 (0.37) T= 700 (0.63)	C= 413 (0.37) T= 691 (0.63)	0	0.16	0.69
rs2003813	C= 711 (0.64) T= 393 (0.36)	C= 696 (0.63) T= 408 (0.37)	0.01	0.456	0.52
rs5995263	A= 508 (0.46) T= 596 (0.54)	A= 510 (0.46) T= 594 (0.54)	0	0.007	0.93
rs2017689	G= 821 (0.74) A= 283 (0.26)	G= 806 (0.73) A= 298 (0.27)	0.01	0.53	0.47
rs2016988	C= 1016 (0.92) G= 88 (0.08)	C= 1020 (0.92) G= 84 (0.08)	0	0.10	0.75

APOL 3:

Ten SNPs were typed in the pooled association sample across APOL 3. The estimated allele frequencies and allele numbers for each of the assayed SNPs is shown in table 6.9 below. None of the typed SNPs showed a significant difference between the observed case and control allele frequencies.

Table 6.9: Pooled genotyping data for APOL 3.

This table shows the estimated allele frequencies and counts for the APOL 3 polymorphisms and the corresponding p-values.

Apol 3 SNP Data	Pooled Genotyping data				
SNP number.	Alleles Cases (freq)	Alleles Controls (freq)	Difference	Chi-square	P-value
rs132673	G= 647 (0.59) A= 457 (0.41)	G= 635 (0.58) A= 469 (0.42)	0.01	0.27	0.61
rs132661	A= 665 (0.6) G= 439 (0.4)	A= 688 (0.62) G= 416 (0.38)	0.02	1.01	0.32
rs132656	T= 579 (0.52) C= 525 (0.48)	T= 573 (0.52) C= 531 (0.48)	0	0.07	0.79
rs2097465	C= 943 (0.68) T= 439 (0.32	C= 946 (0.66) T= 488 (0.34)	0.02	1.64	0.20
rs2017329	A= 579 (0.52) T= 525 (0.48)	A= 606 (0.55) T= 498 (0.45)	0.03	1.33	0.25
rs132644	C= 883 (0.8) T= 221 (0.2)	C= 888 (0.8) T= 216 (0.2)	0	0.07	0.79
rs132635	C= 748 (0.68) T= 356 (0.32)	C= 766 (0.69) T= 338 (0.31)	0.01	0.68	0.41
rs80578	C= 665 (0.6) T= 439 (0.4)	C= 668 (0.6) T= 436 (0.4)	0	0.02	0.89
rs132618	T= 608 (0.55) A= 496 (0.45)	T= 621 (0.56) A= 483 (0.44)	0.01	0.31	0.58
rs132617	C= 707 (0.64) T= 397 (0.36)	C= 728 (0.66) T= 376 (0.34)	0.02	0.88	0.35

APOL 4:

Twelve SNPs were typed in the pooled association sample across APOL 4. The estimated allele frequencies and allele numbers for each of the assayed SNPs is shown in table 6.10 below. None of the typed SNPs showed a significant difference between the observed case and control allele frequencies.

Table 6.10: Pooled genotyping data for APOL 4.

This table shows the estimated allele frequencies and counts for the APOL 4 polymorphisms and the corresponding p-values.

Apol 4 SNP Data	Pooled Genotyping data				
SNP number.	Alleles Cases (freq)	Alleles Controls (freq)	Difference	Chi-square	P-value
rs132743	G= 973 (0.88) C= 131 (0.12)	G= 972 (0.88) C= 132 (0.12)	0	0.004	0.95
rs2005590	C= 778 (0.7) T= 326 (0.3)	C= 771 (0.7) T= 333 (0.3)	0	0.11	0.75
rs1812023	G= 886 (0.64) A= 496 (0.36)	G= 926 (0.65) A= 508 (0.35)	0.01	0.07	0.74
rs132734	G= 860 (0.62) A= 522 (0.38)	G= 899 (0.63) A= 535 (0.37)	0.01	0.06	0.8
rs6000179	A= 977 (0.89) C= 127 (0.11)	A= 966 (0.87) C= 138 (0.13)	0.02	0.52	0.47
rs2007468	A= 558 (0.51) G= 546 (0.49)	A= 558 (0.51) G= 546 (0.49)	0	0	1
rs132700	T= 793 (0.72) C= 311 (0.28)	T= 790 (0.28) C= 314 (0.28)	0	0.02	0.89
rs2227167	A= 577 (0.52) G= 527 (0.48)	A= 574 (0.52) G= 530 (0.48)	0	0.02	0.89
rs916336	C= 561 (0.51) G= 543 (0.49)	C= 560 (0.51) G= 544 (0.49)	0	0.002	0.97
rs1966267	A= 564 (0.51) G= 540 (0.49)	A= 569 (0.51) G= 535 (0.49)	0	0.05	0.83
rs132691	C= 797 (0.72) G= 307 (0.28)	C= 792 (0.72) G= 312 (0.28)	0	0.06	0.81
rs132683	A= 806 (0.73) G= 298 (0.27)	A= 809 (0.73) G= 295 (0.27)	0	0.02	0.89

APOL 5:

Eight SNPs were typed in the pooled association sample across APOL 5. The estimated allele frequencies and allele numbers for each of the assayed SNPs is shown in table 6.11 below. None of the typed SNPs showed a significant difference between the observed case and control allele frequencies.

Table 6.11: Pooled genotyping data for APOL 5.

This table shows the estimated allele frequencies and counts for the APOL 1 polymorphisms and the corresponding p-values.

Apol 5 SNP Data	Pooled Genotyping data				
SNP number.	Alleles Cases (freq)	Alleles Controls (freq)	Difference	Chi-square	P-value
rs1540297	T= 679 (0.61) C= 425 (0.39)	T= 667 (0.6) C= 437 (0.4)	0.01	0.27	0.60
rs5999971	T= 748 (0.68) C= 356 (0.32)	T= 736 (0.67) C= 368 (0.33)	0.01	0.29	0.59
rs2009168	A= 752 (0.68) G= 352 (0.32)	A= 729 (0.66) G= 375 (0.34)	0.02	1.09	0.29
rs2076671	C= 888 (0.64) T= 494 (0.36)	C= 952 (0.66) T= 482 (0.34)	0.02	1.41	0.23
rs2076672	C= 801 (0.73) T= 303 (0.27)	C= 788 (0.71) T= 316 (0.29)	0.02	0.38	0.54
rs2016586	G= 787 (0.71) T= 317 (0.29)	G= 760 (0.69) T= 344 (0.31)	0.02	1.57	0.21
rs5995165	A= 793 (0.72) G= 311 (0.28)	A= 792 (0.72) G= 312 (0.28)	0	0.002	0.96
rs879680	C= 794 (0.72) T= 310 (0.28)	C= 768 (0.7) T= 336 (0.3)	0.02	1.48	0.22

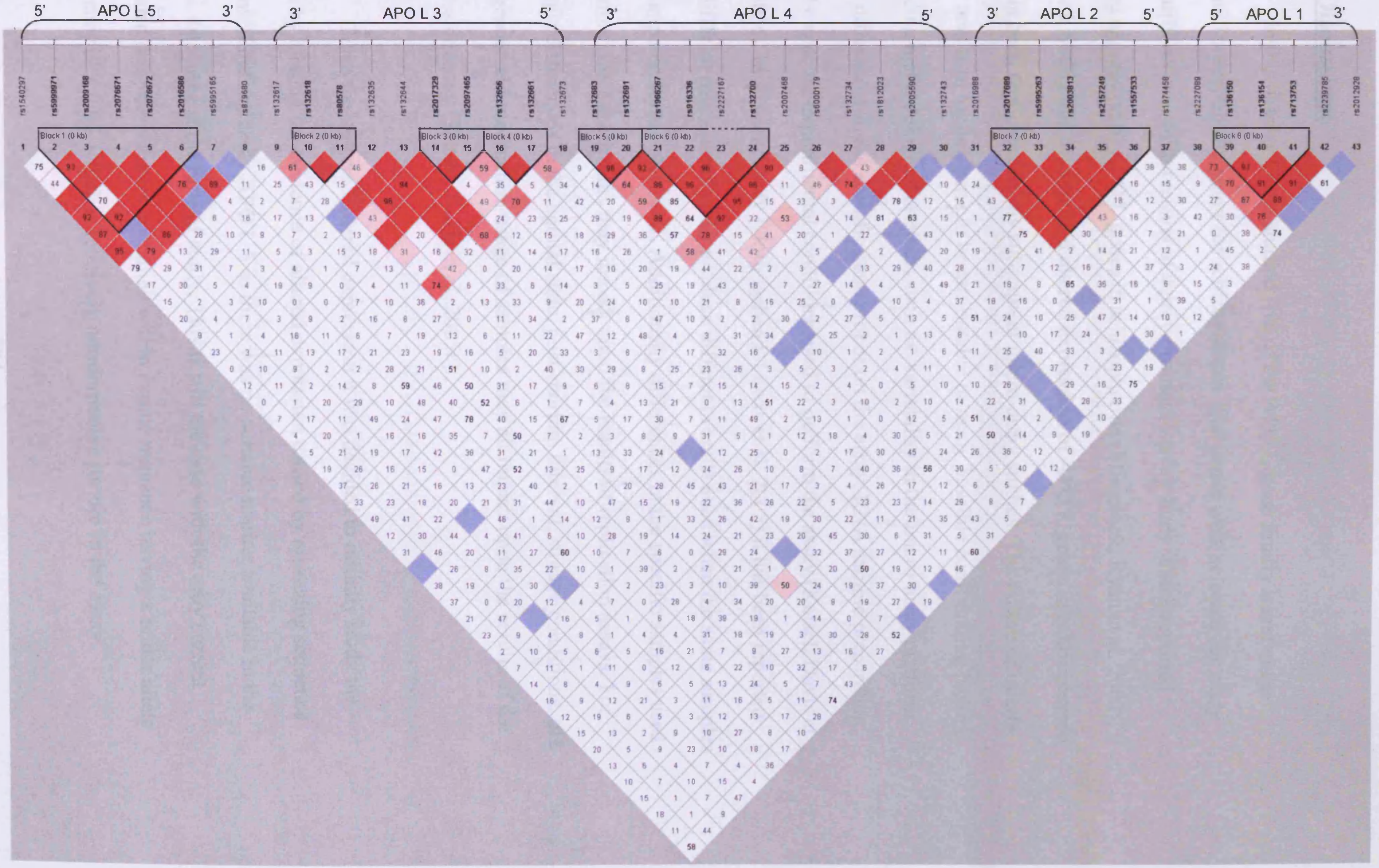
6.3.2.1 LD analysis of the APOL genomic region.

Following genotyping of all APO L markers in the 96 unrelated control samples the LD structure across the APO L genomic region was calculated using the program Haploview (http://www.broad.mit.edu/mpg/haploview/). The resultant Haploview LD plot is presented below (figure 6.2). The Haploview program defines blocks of LD using the method developed by Gabriel and colleagues (Gabriel, Schaffner et al. 2002).

The analysis revealed strong LD across APO L 2 and APO L 5, with a block of LD present in APO L 2 between markers rs2017689 and rs1557533 and a block of LD in APO L 5 between markers rs5999971 and rs2016586. There was also revealed to be strong LD present across most of APO L 1 and modest levels of LD present across APO L 3 and APO L 4 (figure 6.2). The strong LD might explain why although 43 SNPs were tested for association, none were significant, when 2 SNPs would have been expected to show significance by chance at the p=0.05 level. The LD may have reduced the number of effective SNPs being tested for association, as markers in LD are not independent from each other.

Figure 6.2 LD structure of the APOL genomic region.

Haploview LD plot of the APO L genomic region displaying the D' values between the APO L markers, the LD blocks present as defined by Gabriel and colleagues (2002) and the position of the five APO L genes (not to scale).



6.4 Discussion:

In this chapter I have studied five of the APO L gene family members as potential susceptibility genes for schizophrenia. The genes were selected for study primarily because of the results of a previous micro-array study which reported altered expression of APO L genes in schizophrenia (Mimmack, Ryan et al. 2002). It is worth noting that the reported altered expression of APO L genes in schizophrenia has not been replicated in any subsequent micro-array studies. The failure of micro-array results to replicate might be attributable to many possible confounding factors including methodological artefacts such as: variability in the micro-array platform used, differences in the specific gene probes on the array, RNA sample preparation, and the specific experimental conditions used to perform the assay (Katsel, Davis et al. 2005c). In a recent study (Hollingshead, Lewis et al. 2005) it has been reported that different micro-arrays may produce different results even when the same samples are compared across platforms. Additionally differences in the specific samples used to perform the experiments may impact upon the results of replication assays. For example differences in the agonal state, post-mortem interval, storage conditions, pH, age, gender and medication usage of samples may all impact upon the results of the expression assay (Katsel, Davis et al. 2005c).

Given the reported altered expression it was decided to initially study the genes by looking for evidence of altered expression caused by *cis*-acting sequence variants. I did not assay APO L 6 as there was no suitable marker available in the public databases at the time of study. This is still the case with the only current validated database markers contained within exonic sequence having a minor allele frequency of ≤ 0.03, which is relatively uninformative for use in the assay.

Of the five genes that were assayed all displayed evidence of altered allelic expression caused by *cis*-acting sequence variation. There are two general explanations for allelic expression differences: a) DNA sequence variant(s) in a region of the genomic sequence that affects the transcription, splicing or stability of the genes mRNA transcript, or b) Epigenetic effects (e.g Methylation etc) which result in the preferential transcription of one parental chromosome. The results of the expression assays were slightly surprising in that all five of the assayed genes displayed evidence of altered allelic expression. It has been previously reported that on average about 50% of genes display altered allelic expression attributable to common cis-acting variation (Bray, Buckland et al. 2003a; Pastinen, Sladek et al. 2004). It is therefore unusual that all five of the studied APO L genes show evidence of differential allelic expression. Moreover the significant differences in allelic expression were seen in a large proportion of the samples assayed for each of the genes, along with variations in the magnitude of observed expression differences.

These observations might be explained by the fact that all five of the genes are located together on chromosome 22, where it is possible that they are all controlled by a common master regulatory element, which contains a *cis*-acting sequence variant responsible for the observed allelic expression differences. However, the genes are transcribed in different directions (figure 6.1) and it is unlikely that they are all under the control of a single promoter. Furthermore the extent of the altered expression differs both within and between the assays, ranging from the fairly minor differences observed for APO L 3 to what appears to be almost mono-allelic expression in some of the samples in the APO L 4 assay. Finally analysis of the data revealed no evidence of correlation between the levels of altered expression observed for the same individuals across the different assays (table 6.6). For instance the three individuals that exhibit almost mono-allelic expression in the APO L 4 assay do not display mono-allelic or highly altered expression in any of the other assays. Additionally, analysis of the expression data for the genes which are transcribed in the same direction (APO L 1 and 5) and (APO L 2, 3, and 4) revealed no evidence of a significant correlation between either the samples displaying altered expression or the extent of the expression differences observed in the assays (table 6.6). When taken together this would tend to suggest that the expression differences observed are not due to a common variant affecting the expression of all five genes.

The APO L 3 gene was the only assay not to show large expression differences. This is not explained by a relative lack of power compared to the analysis of the other genes, as 20 heterozygous samples were assayed compared with 11 heterozygotes in the case of the APO L 1 assay. The subset of 20 heterozygotes assayed provides a power of approximately 86% to detect the affect of a regulatory variant present in the general population at a frequency of 0.05. The lack of large expression differences can also not be explained by the specific samples used to assay APO L 3 as several of the samples displayed large expression differences in other assays including one individual which displays almost mono-allelic expression for APOL 4.

Possibly the most interesting result from all of the assays apart from the existence of altered allelic expression in all of the genes is the apparent mono-allelic expression of three individuals in APO L 4. The individuals displayed allelic ratios of >1:10, which would represent a >1000% difference in the relative allelic expression. These results possibly suggest a *cis*-acting variant with a substantial effect upon the gene expression, or epigenetic effects such as methylation, imprinting, or histone acetylation etc. It has been proposed that polymorphic imprinting of certain genes may play a role in common mental disorders such as schizophrenia, attention deficit hyper-activity disorder (ADHD) and bipolar disorder given the evidence suggesting the disorders are inherited preferentially from one parent and/or the related linkage/association findings that are parent of origin dependant (Isles and Wilkinson 2000; Davies, Isles et al. 2005; Morison, Ramsay et al. 2005). It is however worth noting that only about 1% of genes have been estimated to be subject to epigenetic modification (Sleutels, Barlow et al. 2000; Davies, Isles et al. 2005). A high proportion of the imprinted genes that have been discovered are highly expressed in the CNS, and it is thought that they might play a role in growth and development processes. It has also been proposed that the neurodevelopmental process may provide a mechanism for imprinted genes in the CNS to influence brain function (Davies, Isles et al. 2005).

The allelic expression assay can only measure the effects of *cis*-acting phenomena but is unable to identify either the specific functional variant or its genomic location. Unfortunately, in the absence of a specific causal or functional variant it is not possible to determine definitively whether the observed expression differences are the results of cis-acting sequence variants or are caused by epigenetic effects.

Given the evidence compatible with *cis*-acting effects from the allelic expression assays I decided to study the APO L genes genetically using an indirect association study design. As I had not been able to assay the allelic expression of APO L 6 I decided to also not screen the gene for association with schizophrenia. A total of forty three polymorphisms were genotyped through the pooled case-control association sample, spanning a total of seventy four kilobases of genomic sequence.

None of the polymorphisms genotyped in the pools displayed any evidence of association with schizophrenia. These results are consistent with the findings of a recent study (McGhee, Morris et al. 2005) who also reported no evidence for association between schizophrenia and the APO L gene family. In their study they genotyped 51 polymorphic markers across all six of the APOL genes in a large Irish case-control pooled sample. Although no evidence for association was detected in my study or the previous study (McGhee, Morris et al. 2005) it is possible that genetic variants associated with schizophrenia and/or that affect the genes expression are located either proximal or distal to the gene, and were therefore undetected in this study. It is also possible that the variants might have effect sizes prohibitively small to allow them to have been detected with the power of this study.

One of the major drawbacks of using a pooled genotyping strategy is the inability to be able to perform haplotype analysis. These analyses could potentially increase the power of the study to detect any associated variants or haplotypes that are present. However, the high density of markers and the relatively high levels of LD in this study would suggest that any associated variants or haplotypes have a high chance of detection.

An alternative method to the one employed in this project would have been to use Tag markers selected from the HapMap. These markers are selected using bioinformatics programs in order to 'tag' the genetic information contained within a gene. These tags can be selected using different criteria, for example extracting the information from untyped markers on the basis of pairwise r^2 or 2-3 marker haplotypes. If this technique had been applied to study APO L 1-5 based on the HapMap (version 2.0) using a pairwise r^2 tagging method based on the parameters (r^2=0.8, MAF=0.1) a total of 54 tags markers would need to be genotyped in order to capture all of the information in the APO L genes currently contained within the HapMap. Only seventeen of the 54 recommended tags or proxies for these markers have actually been typed in this study. Only two of the eight markers required to screen APO L 1 were typed, three of thirteen for APO L 2, four out of sixteen for APO L 3, and four out of thirteen for APO L 4. In the case of APO L 5 all of the tag markers have been genotyped.

In conclusion this study has shown strong evidence of altered allelic expression attributable to *cis*-acting sequence variation or epigenetics in five of the six members of the APO L gene family. However, it was not possible in this study to relate this altered expression directly to schizophrenia. When tested none of the APOL genes displayed evidence for genetic association to schizophrenia. Thus my study does not explain the results of the previous study which reported altered expression of the APO L gene family in schizophrenia (Mimmack, Ryan et al. 2002). This may be because for the reason of power, marker density or measurement error in pools, I failed to detect true effects. Alternatively, it is possible that the altered expression observed in that study was not directly related to the aetiology of schizophrenia but was attributable to artefact or secondary effects caused by environmental or *trans*-acting factors affecting the genes expression, such as diet, the effects of medication, or the alteration of transcription factor efficiency.

Chapter 7. Final Discussion.

Schizophrenia is a severe psychiatric disorder with onset in adolescence or early adulthood, characterised by a diverse range of symptoms that have an impact on many areas of mental function. Although its aetiology is not well understood, it is widely accepted that the disorder has a significant genetic component. The mode of transmission is most likely complex involving a number of genetic loci with a variable but mostly small effect on susceptibility for the disorder. However, studies have shown that the risk of schizophrenia can not be explained exclusively by genetic factors, but also has a significant environmental component. It seems most likely that the disorder is ultimately caused by the co-action of multiple genetic factors together with interaction with various environmental factors. Thus the challenge facing the field of schizophrenia research is not only to detect susceptibility genes but also to determine how these genes interact with each other and the environment to result in disease susceptibility.

Numerous hypotheses have so far been formulated to explain the biological origins of the disorder, but schizophrenia has increasingly been viewed as a neurodevelopmental disorder, caused by abnormalities occurring in the brain during development and before the onset of schizophrenic symptoms.

Recently, and compatible with a developmental hypothesis, it has been proposed that oligodendrocyte and/or myelination abnormalities play a role in the aetiology of schizophrenia. Support for this has been provided by the results of several areas of research including neuropathological studies, brain imaging analysis, phenotypic overlap with demyelinating diseases, age related changes in myelin and from the results of the micro-array expression studies performed on post-mortem schizophrenic brain which have displayed altered expression of OMR genes in schizophrenia.

7.1 My Research.

The series of investigations described in this thesis were principally candidate gene studies of schizophrenia. A specific hypothesis was investigated, which proposed that variation in genes related to oligodendrocyte and/or myelination function play a role in the aetiology of schizophrenia. Genes were selected as candidates for study on the basis of their documented altered expression in schizophrenic post-mortem brain supported in many cases by their location in regions of the genome reported to be linked to schizophrenia.

In total I studied thirteen genes as candidates for involvement in schizophrenia: CNP, NOGO, NgR, OMG, NGFR, GFAP, MOG, SOX 10 and APOL 1-5. A prior requirement for candidate gene studies is the availability of polymorphisms to test for association to disease. For all of the research presented in this thesis, except for the study of the APOL gene family, polymorphisms were identified through mutation detection using denaturing high performance liquid chromatography (DHPLC). During the investigations 147 polymorphisms were identified: 137 SNPs, 5 insertion/deletions, 2 VNTRs, 2 variable length poly T sequences and 1 variable length poly A sequence. Of the polymorphisms, 55 were exonic SNPs of which 10 were synonymous and 7 were non-synonymous SNPs.

It is notable that of the identified polymorphisms, 121 are now independently described in the public databases, reflecting the huge increase in the genetic information which has become available over the past five years. Given these resources now available to researchers it is becoming increasingly more appropriate to select SNPs from a database rather than to use extensive mutation screening to detect them experimentally, except perhaps in the case of the examination of rare susceptibility variants. Indeed it might be argued that the examination of grids of markers based upon the known LD structure is more likely to detect association from non-coding variants than the approach utilised in the investigations presented here of screening all of the variation in the exonic and limited 5'-flanking sequence.

I attempted to genotype all the polymorphisms identified in these studies that were compatible with a primer extension assay in DNA pools. Pooled genotyping was attempted for 142 polymorphisms, of which, 7 had a minor allele frequency too rare to be genotyped in pools, resulting in pooled genotyping data for 135 polymorphisms. As no attempt was made to genotype by an alternative method polymorphisms that were either too rare or incompatible with primer extension, direct associations due to these variants may have gone undetected, unless being detected through LD with another typed marker. This potential problem might have been rectified by selecting additional markers from the public databases rather than relying on variants identified by DHPLC, although it should be noted that if the risk variant is rare, and only common markers are tested for association, then the difference in allele frequencies at the risk and marker loci will significantly reduce the power to detect association, even if the two markers are in complete LD (as measured by D').

During the series of studies presented here I detected modest evidence for association with three genes (CNP, NOGO and MOG) and also showed that the expression of these genes is under the influence of *cis*-acting polymorphisms. The genes are good candidates for study as they are differentially expressed in post-mortem schizophrenic brain, located in reported schizophrenia linkage regions and functionally involved with the development of the CNS.

CNP had not been previously reported to be associated with schizophrenia and whilst the result is supported by both the allelic expression assay and the observations in a multiply affected family showing linkage to the region, the finding certainly requires independent replication in large suitably powered association samples. Several studies have previously reported NOGO to be associated with schizophrenia, although there have also been several negative studies published. It should also be noted that there are inconsistencies with the alleles and allele frequencies of the reported associations. These differences and the negative studies are possibly caused by different LD structures or alleles being responsible for association in different populations. The results of my association and allelic expression analysis of MOG, along with the findings of previous association studies are not compelling and provide weak evidence for the involvement of MOG in schizophrenia aetiology. However, further examination of the associated non-synonymous SNP is probably warranted to independently replicate the results. Whilst, the results for the three genes are modest, would not survive correction for the total number of tests performed and are awaiting further replication, they do provide some support for the general hypothesis that oligodendrocyte and/or myelination abnormalities play a role in schizophrenia aetiology.

There was no evidence for association detected with the other ten genes studied as part of this thesis. However for the reasons I discussed earlier in chapter 4, I cannot exclude any of these genes from playing a role in the aetiology of schizophrenia. To be able to do this, at a minimum requires genotyping of all the non-redundant variation contained within the gene in a sufficiently powerful association sample representative of the general population. The problems associated with excluding a gene from disease aetiology are not easily remedied since there is always an effect size that the largest available sample is unable to detect.

Since this series of investigations was undertaken, the field of molecular genetics has moved on at an impressive speed, meaning that many of the techniques utilised in these studies are no longer relevant or appropriate for use. For instance the use of DHPLC mutation detection has now become a very expensive and slow process and given the wealth of available genetic information almost obsolete. Even where mutation detection is still required, such as for the identification of rare sequence variants, new technology has been developed making the screening of hundreds even thousands of samples feasible (Li, Atmaca-Sonmez et al. 2006; Maller, George et al. 2006). The recent advances in the bioinformatic resources available to molecular biologists such as the HapMap combined with tag SNP selection programs and higher throughput genotyping platforms have also changed the way that candidate gene studies are designed. Increasingly, genes and large regions of the genome are being examined for association in at least moderate depth using indirect methods relying upon linkage disequilibrium.

Given the progress that has being made in molecular genetics, it is encouraging that recently molecular genetic studies of schizophrenia have started to produce evidence implicating several genes in the aetiology of the disorder (described in chapter 1). However, the majority of the genetic risk for schizophrenia still remains to be accounted for. In order to increase the chance of detecting the remaining susceptibility genes, strategies that increase the statistical power to detect small effects need to be employed, along with the use of other complementary solutions such as refined phenotype analysis to reduce heterogeneity. The power to detect novel loci by investigating gene-gene and gene-environment interactions while promoted by some, with apparent success (Caspi, Sugden et al. 2003; Thapar, Langley et al. 2005), remains controversial (Wang, Barratt et al. 2005; Todd 2006). However, at a minimum such studies based on known genetic risk factors may make significant contributions to our knowledge of pathogenesis. Such studies, based in part on the work described in this thesis, have recently provided further evidence for the involvement of CNP and other OMR genes in schizophrenia susceptibility (Georgieva, Moskvina et al. 2006) and further support the hypothesis that oligodendrocyte and/or myelination abnormalities play a role in the aetiology of schizophrenia.

Chapter 8. Bibliography.

Abecasis G R, Noguchi E, Heinzmann A, Traherne J A, Bhattacharyya S, Leaves N I, Anderson G G, Zhang Y, Lench N J, Carey A, Cardon L R, Moffatt M F and Cookson W O (2001) Extent and distribution of linkage disequilibrium in three genomic regions. *Am J Hum Genet* 68 (1) 191-197.

Aberg K, Saetre P, Jareborg N and Jazin E (2006a) Human QKI, a potential regulator of mRNA expression of human oligodendrocyte-related genes involved in schizophrenia. *Proc Natl Acad Sci U S A* 103 (19) 7482-7487.

Aberg K, Saetre P, Lindholm E, Ekholm B, Pettersson U, Adolfsson R and Jazin E (2006b) Human QKI, a new candidate gene for schizophrenia involved in myelination. *Am J Med Genet B Neuropsychiatr Genet* 141 (1) 84-90.

Ackerman H, Usen S, Mott R, Richardson A, Sisay-Joof F, Katundu P, Taylor T, Ward R, Molyneux M, Pinder M and Kwiatkowski D P (2003) Haplotypic analysis of the TNF locus by association efficiency and entropy. *Genome Biol* 4 (4) R24.

Addington A M, Gornick M, Sporn A L, Gogtay N, Greenstein D, Lenane M, Gochman P, Baker N, Balkissoon R, Vakkalanka R K, Weinberger D R, Straub R E and Rapoport J L (2004) Polymorphisms in the 13q33.2 gene G72/G30 are associated with childhood-onset schizophrenia and psychosis not otherwise specified. *Biol Psychiatry* 55 (10) 976-980.

Agartz I, Andersson J L and Skare S (2001) Abnormal brain white matter in schizophrenia: a diffusion tensor imaging study. *Neuroreport* 12 (10) 2251-2254.

Akaho R, Matsushita I, Narita K, Okazaki Y, Okabe Y, Matsushita M, Hohjoh H, Tokunaga K and Sasaki T (2000) Support for an association between HLA-DR1 and schizophrenia in the Japanese population. *Am J Med Genet* 96 (6) 725-727.

Akbarian S, Bunney W E, Jr., Potkin S G, Wigal S B, Hagman J O, Sandman C A and Jones E G (1993a) Altered distribution of nicotinamide-adenine dinucleotide phosphate-diaphorase cells in frontal lobe of schizophrenics implies disturbances of cortical development. *Arch Gen Psychiatry* 50 (3) 169-177.

Akbarian S, Vinuela A, Kim J J, Potkin S G, Bunney W E, Jr. and Jones E G (1993b) Distorted distribution of nicotinamide-adenine dinucleotide phosphate-diaphorase neurons in temporal lobe of schizophrenics implies anomalous cortical development. *Arch Gen Psychiatry* 50 (3) 178-187.

Akil M and Lewis D A (1997) Cytoarchitecture of the entorhinal cortex in schizophrenia. *Am J Psychiatry* 154 (7) 1010-1012.

Al Halabiah H, Delezoide A L, Cardona A, Moalic J M and Simonneau M (2005) Expression pattern of NOGO and NgR genes during human development. *Gene Expr Patterns* 5 (4) 561-568.

Andreasen N C (1995) Symptoms, signs, and diagnosis of schizophrenia. *Lancet* 346 (8973) 477-481.

Aoki Y, Haginoya K, Munakata M, Yokoyama H, Nishio T, Togashi N, Ito T, Suzuki Y, Kure S, Iinuma K, Brenner M and Matsubara Y (2001) A novel mutation in glial fibrillary acidic protein gene in a patient with Alexander disease. *Neurosci Lett* 312 (2) 71-74.

Ardlie K G, Kruglyak L and Seielstad M (2002) Patterns of linkage disequilibrium in the human genome. *Nat Rev Genet* 3 (4) 299-309.

Arnold S E, Hyman B T, Van Hoesen G W and Damasio A R (1991) Some cytoarchitectural abnormalities of the entorhinal cortex in schizophrenia. *Arch Gen Psychiatry* 48 (7) 625-632.

Arnold S E, Ruscheinsky D D and Han L Y (1997) Further evidence of abnormal cytoarchitecture of the entorhinal cortex in schizophrenia using spatial point pattern analyses. *Biol Psychiatry* 42 (8) 639-647.

American Psychiatric Association (1994) *Diagnostic and statistical manual of mental disorders (DSM-IV).* Washington, DC: American Psychiatric Association.

Aston C, Jiang L and Sokolov B P (2004) Microarray analysis of postmortem temporal cortex from patients with schizophrenia. *J Neurosci Res* 77 (6) 858-866.

Aston C, Jiang L and Sokolov B P (2005) Transcriptional profiling reveals evidence for signaling and oligodendroglial abnormalities in the temporal cortex from patients with major depressive disorder. *Mol Psychiatry* 10 (3) 309-322.

Baare W F, van Oel C J, Hulshoff Pol H E, Schnack H G, Durston S, Sitskoorn M M and Kahn R S (2001) Volumes of brain structures in twins discordant for schizophrenia. *Arch Gen Psychiatry* 58 (1) 33-40.

Badner J A and Gershon E S (2002) Meta-analysis of whole-genome linkage scans of bipolar disorder and schizophrenia. *Mol Psychiatry* 7 (4) 405-411.

Bagary M S, Symms M R, Barker G J, Mutsatsa S H, Joyce E M and Ron M A (2003) Gray and white matter brain abnormalities in first-episode schizophrenia inferred from magnetization transfer imaging. *Arch Gen Psychiatry* 60 (8) 779-788.

Barrett J C and Cardon L R (2006) Evaluating coverage of genome-wide association studies. *Nat Genet* 38 (6) 659-662.

Barrett T B, Hauger R L, Kennedy J L, Sadovnick A D, Remick R A, Keck P E, McElroy S L, Alexander M, Shaw S H and Kelsoe J R (2003) Evidence that a single nucleotide polymorphism in the promoter of the G protein receptor kinase 3 gene is associated with bipolar disorder. *Mol Psychiatry* 8 (5) 546-557.

Bassett A S, Hodgkinson K, Chow E W, Correia S, Scutt L E and Weksberg R (1998) 22q11 deletion syndrome in adults with schizophrenia. *Am J Med Genet* 81 (4) 328-337.

Baumann N and Pham-Dinh D (2001) Biology of oligodendrocyte and myelin in the mammalian central nervous system. *Physiol Rev* 81 (2) 871-927.

Benes F M (2000) Emerging principles of altered neural circuitry in schizophrenia. *Brain Res Brain Res Rev* 31 (2-3) 251-269.

Benes F M, Turtle M, Khan Y and Farol P (1994) Myelination of a key relay zone in the hippocampal formation occurs in the human brain during childhood, adolescence, and adulthood. *Arch Gen Psychiatry* 51 (6) 477-484.

Bennett P, Segurado R, Jones I, Bort S, McCandless F, Lambert D, Heron J, Comerford C, Middle F, Corvin A, Pelios G, Kirov G, Larsen B, Mulcahy T, Williams N, O'Connell R, O'Mahony E, Payne A, Owen M, Holmans P, Craddock N and Gill M (2002) The Wellcome trust UK-Irish bipolar affective disorder sibling-pair genome screen: first stage report. *Mol Psychiatry* 7 (2) 189-200.

Bifulco M, Laezza C, Stingo S and Wolff J (2002) 2',3'-Cyclic nucleotide 3'-phosphodiesterase: a membrane-bound, microtubule-associated protein and membrane anchor for tubulin. *Proc Natl Acad Sci U S A* 99 (4) 1807-1812.

Blackwood D H, Fordyce A, Walker M T, St Clair D M, Porteous D J and 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) 428-433.

Bleuler E (1911) *Dementia Praecox or the Group of Schizophrenias*. Leipzig: Deutike.

Blouin J L, Dombroski B A, Nath S K, Lasseter V K, Wolyniec P S, Nestadt G, Thornquist M, Ullrich G, McGrath J, Kasch L, Lamacz M, Thomas M G, Gehrig C, Radhakrishna U, Snyder S E, Balk K G, Neufeld K, Swartz K L, DeMarchi N, Papadimitriou G N, Dikeos D G, Stefanis C N, Chakravarti A, Childs B, Housman D E, Kazazian H H, Antonarakis S and Pulver A E (1998) Schizophrenia susceptibility loci on chromosomes 13q32 and 8p21. *Nat Genet* 20 (1) 70-73.

Bogerts B, Falkai P, Haupts M, Greve B, Ernst S, Tapernon-Franz U and Heinzmann U (1990) Post-mortem volume measurements of limbic system and basal ganglia structures in chronic schizophrenics. Initial results from a new brain collection. *Schizophr Res* 3 (5-6) 295-301.

Bogerts B, Meertz E and Schonfeldt-Bausch R (1985) Basal ganglia and limbic system pathology in schizophrenia. A morphometric study of brain volume and shrinkage. *Arch Gen Psychiatry* 42 (8) 784-791.

Braun P E, Lee, J, Gravel, M. (2004) 2', 3'-Cyclic Nucleotide 3'-Phosphodiesterase: Structure, Biology and Function.*Myelin Biology and Disorders.*, Vol. 1st. Elsevier.

Bray N J, Buckland P R, Owen M J and O'Donovan M C (2003a) Cis-acting variation in the expression of a high proportion of genes in human brain. *Hum Genet* 113 (2) 149-153.

Bray N J, Buckland P R, Williams N M, Williams H J, Norton N, Owen M J and O'Donovan M C (2003b) A haplotype implicated in schizophrenia susceptibility is associated with reduced COMT expression in human brain. *Am J Hum Genet* 73 (1) 152-161.

Bray N J, Preece A, Williams N M, Moskvina V, Buckland P R, Owen M J and O'Donovan M C (2005) Haplotypes at the dystrobrevin binding protein 1 (DTNBP1) gene locus mediate risk for schizophrenia through reduced DTNBP1 expression. *Hum Mol Genet* 14 (14) 1947-1954.

Britsch S, Goerich D E, Riethmacher D, Peirano R I, Rossner M, Nave K A, Birchmeier C and Wegner M (2001) The transcription factor Sox10 is a key regulator of peripheral glial development. *Genes Dev* 15 (1) 66-78.

Brown R, Colter N, Corsellis J A, Crow T J, Frith C D, Jagoe R, Johnstone E C and Marsh L (1986) Postmortem evidence of structural brain changes in schizophrenia. Differences in brain weight, temporal horn area, and parahippocampal gyrus compared with affective disorder. *Arch Gen Psychiatry* 43 (1) 36-42.

Brzustowicz L M, Hodgkinson K A, Chow E W, Honer W G and Bassett A S (2000) Location of a major susceptibility locus for familial schizophrenia on chromosome 1q21-q22. *Science* 288 (5466) 678-682.

Brzustowicz L M, Honer W G, Chow E W, Little D, Hogan J, Hodgkinson K and Bassett A S (1999) Linkage of familial schizophrenia to chromosome 13q32. *Am J Hum Genet* 65 (4) 1096-1103.

Buchanan R W, Vladar K, Barta P E and Pearlson G D (1998) Structural evaluation of the prefrontal cortex in schizophrenia. *Am J Psychiatry* 155 (8) 1049-1055.

Byne W, Kidkardnee S, Tatusov A, Yiannoulos G, Buchsbaum M S and Haroutunian V (2006) Schizophrenia-associated reduction of neuronal and oligodendrocyte numbers in the anterior principal thalamic nucleus. *Schizophr Res* 85 (1-3) 245-253.

Cahn W, Hulshoff Pol H E, Lems E B, van Haren N E, Schnack H G, van der Linden J A, Schothorst P F, van Engeland H and Kahn R S (2002) Brain volume changes in first-episode schizophrenia: a 1-year follow-up study. *Arch Gen Psychiatry* 59 (11) 1002-1010.

Calaora V, Rogister B, Bismuth K, Murray K, Brandt H, Leprince P, Marchionni M and Dubois-Dalcq M (2001) Neuregulin signaling regulates neural precursor growth and the generation of oligodendrocytes in vitro. *J Neurosci* 21 (13) 4740-4751.

Callicott J H, Straub R E, Pezawas L, Egan M F, Mattay V S, Hariri A R, Verchinski B A, Meyer-Lindenberg A, Balkissoon R, Kolachana B, Goldberg T E and Weinberger D R (2005) Variation in DISC1 affects hippocampal structure and function and increases risk for schizophrenia. *Proc Natl Acad Sci U S A* 102 (24) 8627-8632.

Cannon M, Jones P B and Murray R M (2002) Obstetric complications and schizophrenia: historical and meta-analytic review. *Am J Psychiatry* 159 (7) 1080-1092.

Canoll P D, Musacchio J M, Hardy R, Reynolds R, Marchionni M A and Salzer J L (1996) GGF/neuregulin is a neuronal signal that promotes the proliferation and survival and inhibits the differentiation of oligodendrocyte progenitors. *Neuron* 17 (2) 229-243.

Cantor-Graae E and Selten J P (2005) Schizophrenia and migration: a meta-analysis and review. *Am J Psychiatry* 162 (1) 12-24.

Capon F, Allen M H, Ameen M, Burden A D, Tillman D, Barker J N and Trembath R C (2004) A synonymous SNP of the corneodesmosin gene leads to increased mRNA stability and demonstrates association with psoriasis across diverse ethnic groups. *Hum Mol Genet* 13 (20) 2361-2368.

Cardno A G and 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) 12-17.

Cardon L R and Abecasis G R (2003) Using haplotype blocks to map human complex trait loci. *Trends Genet* 19 (3) 135-140.

Carlson C S, Eberle M A, Kruglyak L and Nickerson D A (2004) Mapping complex disease loci in whole-genome association studies. *Nature* 429 (6990) 446-452.

Carlsson A (1978) Does dopamine have a role in schizophrenia? *Biol Psychiatry* 13 (1) 3-21.

Caspi A, Sugden K, Moffitt T E, Taylor A, Craig I W, Harrington H, McClay J, Mill J, Martin J, Braithwaite A and Poulton R (2003) Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science* 301 (5631) 386-389.

Chan J R, Cosgaya J M, Wu Y J and Shooter E M (2001) Neurotrophins are key mediators of the myelination program in the peripheral nervous system. *Proc Natl Acad Sci U S A* 98 (25) 14661-14668.

Chen M S, Huber A B, van der Haar M E, Frank M, Schnell L, Spillmann A A, Christ F and Schwab M E (2000) Nogo-A is a myelin-associated neurite outgrowth inhibitor and an antigen for monoclonal antibody IN-1. *Nature* 403 (6768) 434-439.

Chen Q Y, Chen Q, Feng G Y, Wan C L, Lindpaintner K, Wang L J, Chen Z X, Gao Z S, Tang J S, Li X W and He L (2006) Association between the brain-derived neurotrophic factor (BDNF) gene and schizophrenia in the Chinese population. *Neurosci Lett* 397 (3) 285-290.

Chen W, Gu N, Duan S, Sun Y, Zheng Y, Li C, Pan Y, Xu Y, Feng G and He L (2004a) No association between the genetic polymorphisms within RTN4 and schizophrenia in the Chinese population. *Neurosci Lett* 365 (1) 23-27.

Chen Y S, Akula N, Detera-Wadleigh S D, Schulze T G, Thomas J, Potash J B, DePaulo J R, McInnis M G, Cox N J and McMahon F J (2004b) Findings in an independent sample support an association between bipolar affective disorder and the G72/G30 locus on chromosome 13q33. *Mol Psychiatry* 9 (1) 87-92; image 85.

Chowdari K V, Mirnics K, Semwal P, Wood J, Lawrence E, Bhatia T, Deshpande S N, B K T, Ferrell R E, Middleton F A, Devlin B, Levitt P, Lewis D A and Nimgaonkar V L (2002) Association and linkage analyses of RGS4 polymorphisms in schizophrenia. *Hum Mol Genet* 11 (12) 1373-1380.

Chumakov I, Blumenfeld M, Guerassimenko O, Cavarec L, Palicio M, Abderrahim H, Bougueleret L, Barry C, Tanaka H, La Rosa P, Puech A, Tahri N, Cohen-Akenine A, Delabrosse S, Lissarrague S, Picard F P, Maurice K, Essioux L, Millasseau P, Grel

P, Debailleul V, Simon A M, Caterina D, Dufaure I, Malekzadeh K, Belova M, Luan J J, Bouillot M, Sambucy J L, Primas G, Saumier M, Boubkiri N, Martin-Saumier S, Nasroune M, Peixoto H, Delaye A, Pinchot V, Bastucci M, Guillou S, Chevillon M, Sainz-Fuertes R, Meguenni S, Aurich-Costa J, Cherif D, Gimalac A, Van Duijn C, Gauvreau D, Ouellette G, Fortier I, Raelson J, Sherbatich T, Riazanskaia N, Rogaev E, Raeymaekers P, Aerssens J, Konings F, Luyten W, Macciardi F, Sham P C, Straub R E, Weinberger D R, Cohen N and Cohen D (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) 13675-13680.

Clements C S, Reid H H, Beddoe T, Tynan F E, Perugini M A, Johns T G, Bernard C C and Rossjohn J (2003) The crystal structure of myelin oligodendrocyte glycoprotein, a key autoantigen in multiple sclerosis. *Proc Natl Acad Sci U S A* 100 (19) 11059-11064.

Conrad A J, Abebe T, Austin R, Forsythe S and Scheibel A B (1991) Hippocampal pyramidal cell disarray in schizophrenia as a bilateral phenomenon. *Arch Gen Psychiatry* 48 (5) 413-417.

Consortium. I H (2005) A haplotype map of the human genome. *Nature* 437 (7063) 1299-1320.

Coon H, Myles-Worsley M, Tiobech J, Hoff M, Rosenthal J, Bennett P, Reimherr F, Wender P, Dale P, Polloi A and Byerley W (1998) Evidence for a chromosome 2p13-14 schizophrenia susceptibility locus in families from Palau, Micronesia. *Mol Psychiatry* 3 (6) 521-527.

Cope N, Harold D, Hill G, Moskvina V, Stevenson J, Holmans P, Owen M J, O'Donovan M C and Williams J (2005) Strong evidence that KIAA0319 on chromosome 6p is a susceptibility gene for developmental dyslexia. *Am J Hum Genet* 76 (4) 581-591.

Corder E H, Saunders A M, Strittmatter W J, Schmechel D E, Gaskell P C, Small G W, Roses A D, Haines J L and Pericak-Vance M A (1993) Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science* 261 (5123) 921-923.

Corfas G, Roy K and Buxbaum J D (2004) Neuregulin 1-erbB signaling and the molecular/cellular basis of schizophrenia. *Nat Neurosci* 7 (6) 575-580.

Corvin A P, Morris D W, McGhee K, Schwaiger S, Scully P, Quinn J, Meagher D, Clair D S, Waddington J L and Gill M (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) 208-213.

Cosgaya J M, Chan J R and Shooter E M (2002) The neurotrophin receptor p75NTR as a positive modulator of myelination. *Science* 298 (5596) 1245-1248.

Cotter D R, Pariante C M and Everall I P (2001) Glial cell abnormalities in major psychiatric disorders: the evidence and implications. *Brain Res Bull* 55 (5) 585-595.

Covault J, Lee J, Jensen K and Kranzler H (2004) Nogo 3'-untranslated region CAA insertion: failure to replicate association with schizophrenia and demonstration of marked population difference in frequency of the insertion. *Brain Res Mol Brain Res* 120 (2) 197-200.

Craddock N, O'Donovan M C and Owen M J (2006a) Genes for schizophrenia and bipolar disorder? Implications for psychiatric nosology. *Schizophr Bull* 32 (1) 9-16.

Craddock N, Owen M J and O'Donovan M C (2006b) The catechol-O-methyl transferase (COMT) gene as a candidate for psychiatric phenotypes: evidence and lessons. *Mol Psychiatry* 11 (5) 446-458.

Daly M J, Rioux J D, Schaffner S F, Hudson T J and Lander E S (2001) High-resolution haplotype structure in the human genome. *Nat Genet* 29 (2) 229-232.

Danik M, Champagne D, Petit-Turcotte C, Beffert U and Poirier J (1999) Brain lipoprotein metabolism and its relation to neurodegenerative disease. *Crit Rev Neurobiol* 13 (4) 357-407.

David A S, Malmberg A, Brandt L, Allebeck P and Lewis G (1997) IQ and risk for schizophrenia: a population-based cohort study. *Psychol Med* 27 (6) 1311-1323.

Davidson M, Reichenberg A, Rabinowitz J, Weiser M, Kaplan Z and Mark M (1999) Behavioral and intellectual markers for schizophrenia in apparently healthy male adolescents. *Am J Psychiatry* 156 (9) 1328-1335.

Davies W, Isles A R and Wilkinson L S (2005) Imprinted gene expression in the brain. *Neurosci Biobehav Rev* 29 (3) 421-430.

Davis K L, Stewart D G, Friedman J I, Buchsbaum M, Harvey P D, Hof P R, Buxbaum J and Haroutunian V (2003) White matter changes in schizophrenia: evidence for myelin-related dysfunction. *Arch Gen Psychiatry* 60 (5) 443-456.

Dawson E, Abecasis G R, Bumpstead S, Chen Y, Hunt S, Beare D M, Pabial J, Dibling T, Tinsley E, Kirby S, Carter D, Papaspyridonos M, Livingstone S, Ganske R, Lohmussaar E, Zernant J, Tonisson N, Remm M, Magi R, Puurand T, Vilo J, Kurg A, Rice K, Deloukas P, Mott R, Metspalu A, Bentley D R, Cardon L R and Dunham I (2002) A first-generation linkage disequilibrium map of human chromosome 22. *Nature* 418 (6897) 544-548.

de Bakker P I, Yelensky R, Pe'er I, Gabriel S B, Daly M J and Altshuler D (2005) Efficiency and power in genetic association studies. *Nat Genet* 37 (11) 1217-1223.

Dean B, Laws S M, Hone E, Taddei K, Scarr E, Thomas E A, Harper C, McClean C, Masters C, Lautenschlager N, Gandy S E and Martins R N (2003) Increased levels of apolipoprotein E in the frontal cortex of subjects with schizophrenia. *Biol Psychiatry* 54 (6) 616-622.

DeLisi L E, Shaw S H, Crow T J, Shields G, Smith A B, Larach V W, Wellman N, Loftus J, Nanthakumar B, Razi K, Stewart J, Comazzi M, Vita A, Heffner T and Sherrington R (2002) A genome-wide scan for linkage to chromosomal regions in 382 sibling pairs with schizophrenia or schizoaffective disorder. *Am J Psychiatry* 159 (5) 803-812.

den Dunnen J T and Antonarakis S E (2001) Nomenclature for the description of human sequence variations. *Hum Genet* 109 (1) 121-124.

Dick D M, Foroud T, Flury L, Bowman E S, Miller M J, Rau N L, Moe P R, Samavedy N, El-Mallakh R, Manji H, Glitz D A, Meyer E T, Smiley C, Hahn R, Widmark C, McKinney R, Sutton L, Ballas C, Grice D, Berrettini W, Byerley W, Coryell W, DePaulo R, MacKinnon D F, Gershon E S, Kelsoe J R, McMahon F J, McInnis M, Murphy D L, Reich T, Scheftner W and Nurnberger J I, Jr. (2003) Genomewide linkage analyses of bipolar disorder: a new sample of 250 pedigrees from the National Institute of Mental Health Genetics Initiative. *Am J Hum Genet* 73 (1) 107-114.

Dietschy J M and Turley S D (2001) Cholesterol metabolism in the brain. *Curr Opin Lipidol* 12 (2) 105-112.

Done D J, Crow T J, Johnstone E C and Sacker A (1994) Childhood antecedents of schizophrenia and affective illness: social adjustment at ages 7 and 11. *Bmj* 309 (6956) 699-703.

Dracheva S, Davis K L, Chin B, Woo D A, Schmeidler J and Haroutunian V (2006) Myelin-associated mRNA and protein expression deficits in the anterior cingulate cortex and hippocampus in elderly schizophrenia patients. *Neurobiol Dis* 21 (3) 531-540.

Duan J, Wainwright M S, Comeron J M, Saitou N, Sanders A R, Gelernter J and Gejman P V (2003) Synonymous mutations in the human dopamine receptor D2 (DRD2) affect mRNA stability and synthesis of the receptor. *Hum Mol Genet* 12 (3) 205-216.

Ewald H, Wikman F P, Teruel B M, Buttenschon H N, Torralba M, Als T D, El Daoud A, Flint T J, Jorgensen T H, Blanco L, Kruse T A, Orntoft T F and Mors O (2005) A genome-wide search for risk genes using homozygosity mapping and microarrays with 1,494 single-nucleotide polymorphisms in 22 eastern Cuban families with bipolar disorder. *Am J Med Genet B Neuropsychiatr Genet* 133 (1) 25-30.

Fatemi S H, Laurence J A, Araghi-Niknam M, Stary J M, Schulz S C, Lee S and Gottesman, II (2004) Glial fibrillary acidic protein is reduced in cerebellum of subjects with major depression, but not schizophrenia. *Schizophr Res* 69 (2-3) 317-323.

Fernandez P A, Tang D G, Cheng L, Prochiantz A, Mudge A W and Raff M C (2000) Evidence that axon-derived neuregulin promotes oligodendrocyte survival in the developing rat optic nerve. *Neuron* 28 (1) 81-90.

Filbin M T (2003) Myelin-associated inhibitors of axonal regeneration in the adult mammalian CNS. *Nat Rev Neurosci* 4 (9) 703-713.

Flynn S W, Lang D J, Mackay A L, Goghari V, Vavasour I M, Whittall K P, Smith G N, Arango V, Mann J J, Dwork A J, Falkai P and Honer W G (2003) Abnormalities of myelination in schizophrenia detected in vivo with MRI, and post-mortem with analysis of oligodendrocyte proteins. *Mol Psychiatry* 8 (9) 811-820.

Foong J, Symms M R, Barker G J, Maier M, Woermann F G, Miller D H and Ron M A (2001) Neuropathological abnormalities in schizophrenia: evidence from magnetization transfer imaging. *Brain* 124 (Pt 5) 882-892.

Fournier A E, GrandPre T and Strittmatter S M (2001) Identification of a receptor mediating Nogo-66 inhibition of axonal regeneration. *Nature* 409 (6818) 341-346.

Francks C, Paracchini S, Smith S D, Richardson A J, Scerri T S, Cardon L R, Marlow A J, MacPhie I L, Walter J, Pennington B F, Fisher S E, Olson R K, DeFries J C, Stein J F and Monaco A P (2004) A 77-kilobase region of chromosome 6p22.2 is associated with dyslexia in families from the United Kingdom and from the United States. *Am J Hum Genet* 75 (6) 1046-1058.

Frank M (2000) MAL, a proteolipid in glycosphingolipid enriched domains: functional implications in myelin and beyond. *Prog Neurobiol* 60 (6) 531-544.

Friston K J (1998) The disconnection hypothesis. *Schizophr Res* 30 (2) 115-125.

Funke B, Finn C T, Plocik A M, Lake S, DeRosse P, Kane J M, Kucherlapati R and Malhotra A K (2004) Association of the DTNBP1 locus with schizophrenia in a U.S. population. *Am J Hum Genet* 75 (5) 891-898.

Gabriel S B, Schaffner S F, Nguyen H, Moore J M, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero S N, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander E S, Daly M J and Altshuler D (2002) The structure of haplotype blocks in the human genome. *Science* 296 (5576) 2225-2229.

Geddes J R, Verdoux H, Takei N, Lawrie S M, Bovet P, Eagles J M, Heun R, McCreadie R G, McNeil T F, O'Callaghan E, Stober G, Willinger U and Murray R M (1999) Schizophrenia and complications of pregnancy and labor: an individual patient data meta-analysis. *Schizophr Bull* 25 (3) 413-423.

Georgieva L, Moskvina V, Peirce T, Norton N, Bray N J, Jones L, Holmans P, Macgregor S, Zammit S, Wilkinson J, Williams H, Nikolov I, Williams N, Ivanov D, Davis K L, Haroutunian V, Buxbaum J D, Craddock N, Kirov G, Owen M J and O'Donovan M C (2006) Convergent evidence that oligodendrocyte lineage transcription factor 2 (OLIG2) and interacting genes influence susceptibility to schizophrenia. *Proc Natl Acad Sci U S A* 103 (33) 12469-12474.

Giedd J N, Blumenthal J, Jeffries N O, Castellanos F X, Liu H, Zijdenbos A, Paus T, Evans A C and Rapoport J L (1999) Brain development during childhood and adolescence: a longitudinal MRI study. *Nat Neurosci* 2 (10) 861-863.

Gill M, Vallada H, Collier D, Sham P, Holmans P, Murray R, McGuffin P, Nanko S, Owen M, Antonarakis S, Housman D, Kazazian H, Nestadt G, Pulver A E, Straub R E, MacLean C J, Walsh D, Kendler K S, DeLisi L, Polymeropoulos M, Coon H, Byerley W, Lofthouse R, Gershon E, Read C M and et al. (1996) A combined analysis of D22S278 marker alleles in affected sib-pairs: support for a susceptibility locus for schizophrenia at chromosome 22q12. Schizophrenia Collaborative Linkage Group (Chromosome 22). *Am J Med Genet* 67 (1) 40-45.

Goldman-Rakic P S and Selemon L D (1997) Functional and anatomical aspects of prefrontal pathology in schizophrenia. *Schizophr Bull* 23 (3) 437-458.

GrandPre T, Nakamura F, Vartanian T and Strittmatter S M (2000) Identification of the Nogo inhibitor of axon regeneration as a Reticulon protein. *Nature* 403 (6768) 439-444.

Grant S F, Thorleifsson G, Reynisdottir I, Benediktsson R, Manolescu A, Sainz J, Helgason A, Stefansson H, Emilsson V, Helgadottir A, Styrkarsdottir U, Magnusson K P, Walters G B, Palsdottir E, Jonsdottir T, Gudmundsdottir T, Gylfason A, Saemundsdottir J, Wilensky R L, Reilly M P, Rader D J, Bagger Y, Christiansen C, Gudnason V, Sigurdsson G, Thorsteinsdottir U, Gulcher J R, Kong A and Stefansson K (2006) Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat Genet* 38 (3) 320-323.

Gravel M, Di Polo A, Valera P B and Braun P E (1998) Four-kilobase sequence of the mouse CNP gene directs spatial and temporal expression of lacZ in transgenic mice. *J Neurosci Res* 53 (4) 393-404.

Gravel M, Peterson J, Yong V W, Kottis V, Trapp B and Braun P E (1996) Overexpression of 2',3'-cyclic nucleotide 3'-phosphodiesterase in transgenic mice alters oligodendrocyte development and produces aberrant myelination. *Mol Cell Neurosci* 7 (6) 453-466.

Green E K, Raybould R, Macgregor S, Hyde S, Young A H, O'Donovan M C, Owen M J, Kirov G, Jones L, Jones I and Craddock N (2006) Genetic variation of brain-derived neurotrophic factor (BDNF) in bipolar disorder: case-control study of over 3000 individuals from the UK. *Br J Psychiatry* 188 21-25.

Gregorio S P, Mury F B, Ojopi E B, Sallet P C, Moreno D H, Yacubian J, Tavares H, Santos F R, Gattaz W F and Dias-Neto E (2005) Nogo CAA 3'UTR Insertion polymorphism is not associated with Schizophrenia nor with bipolar disorder. *Schizophr Res* 75 (1) 5-9.

Grupe A, Li Y, Rowland C, Nowotny P, Hinrichs A L, Smemo S, Kauwe J S, Maxwell T J, Cherny S, Doil L, Tacey K, van Luchene R, Myers A, Wavrant-De Vrieze F, Kaleem M, Hollingworth P, Jehu L, Foy C, Archer N, Hamilton G, Holmans P, Morris C M, Catanese J, Sninsky J, White T J, Powell J, Hardy J, O'Donovan M, Lovestone S, Jones L, Morris J C, Thal L, Owen M, Williams J and Goate A (2006) A scan of chromosome 10 identifies a novel locus showing strong association with late-onset Alzheimer disease. *Am J Hum Genet* 78 (1) 78-88.

Gurling H M, Kalsi G, Brynjolfson J, Sigmundsson T, Sherrington R, Mankoo B S, Read T, Murphy P, Blaveri E, McQuillin A, Petursson H and Curtis D (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) 661-673.

Hafner H, an der Heiden W (2003) Course and outcome in schizophrenia. In: Hirsh SR, W D (ed.) *Schizophrenia.* Oxford: Blackwell Scientific.

Hakak Y, Walker J R, Li C, Wong W H, Davis K L, Buxbaum J D, Haroutunian V and Fienberg A A (2001) Genome-wide expression analysis reveals dysregulation of myelination-related genes in chronic schizophrenia. *Proc Natl Acad Sci U S A* 98 (8) 4746-4751.

Hall D, Gogos J A and Karayiorgou M (2004) The contribution of three strong candidate schizophrenia susceptibility genes in demographically distinct populations. *Genes Brain Behav* 3 (4) 240-248.

Harrison P J (1999) The neuropathology of schizophrenia. A critical review of the data and their interpretation. *Brain* 122 (Pt 4) 593-624.

Harrison P J and Law A J (2006) Neuregulin 1 and schizophrenia: genetics, gene expression, and neurobiology. *Biol Psychiatry* 60 (2) 132-140.

Hasegawa Y, Fujitani M, Hata K, Tohyama M, Yamagishi S and Yamashita T (2004) Promotion of axon regeneration by myelin-associated glycoprotein and Nogo through divergent signals downstream of Gi/G. *J Neurosci* 24 (30) 6826-6832.

Hattori E, Liu C, Badner J A, Bonner T I, Christian S L, Maheshwari M, Detera-Wadleigh S D, Gibbs R A and Gershon E S (2003) Polymorphisms at the G72/G30 gene locus, on 13q33, are associated with bipolar disorder in two independent pedigree series. *Am J Hum Genet* 72 (5) 1131-1140.

He Z and Koprivica V (2004) The Nogo signaling pathway for regeneration block. *Annu Rev Neurosci* 27 341-368.

Hennah W, Varilo T, Kestila M, Paunio T, Arajarvi R, Haukka J, Parker A, Martin R, Levitzky S, Partonen T, Meyer J, Lonnqvist J, Peltonen L and Ekelund J (2003) Haplotype transmission analysis provides evidence of association for DISC1 to schizophrenia and suggests sex-dependent effects. *Hum Mol Genet* 12 (23) 3151-3159.

Hinds D A, Stuve L L, Nilsen G B, Halperin E, Eskin E, Ballinger D G, Frazer K A and Cox D R (2005) Whole-genome patterns of common DNA variation in three human populations. *Science* 307 (5712) 1072-1079.

Hirschhorn J N and Daly M J (2005) Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet* 6 (2) 95-108.

Ho B C, Andreasen N C, Nopoulos P, Arndt S, Magnotta V and Flaum M (2003) Progressive structural brain abnormalities and their relationship to clinical outcome: a longitudinal magnetic resonance imaging study early in schizophrenia. *Arch Gen Psychiatry* 60 (6) 585-594.

Hodgkinson C A, Goldman D, Jaeger J, Persaud S, Kane J M, Lipsky R H and Malhotra A K (2004) Disrupted in schizophrenia 1 (DISC1): association with schizophrenia, schizoaffective disorder, and bipolar disorder. *Am J Hum Genet* 75 (5) 862-872.

Hof P R, Haroutunian V, Copland C, Davis K L and Buxbaum J D (2002) Molecular and cellular evidence for an oligodendrocyte abnormality in schizophrenia. *Neurochem Res* 27 (10) 1193-1200.

Hof P R, Haroutunian V, Friedrich V L, Jr., Byne W, Buitron C, Perl D P and Davis K L (2003) Loss and altered spatial distribution of oligodendrocytes in the superior frontal gyrus in schizophrenia. *Biol Psychiatry* 53 (12) 1075-1085.

Hollingshead D, Lewis D A and Mirnics K (2005) Platform influence on DNA microarray data in postmortem brain research. *Neurobiol Dis* 18 (3) 649-655.

Honer W G, Hurwitz T, Li D K, Palmer M and Paty D W (1987) Temporal lobe involvement in multiple sclerosis patients with psychiatric disorders. *Arch Neurol* 44 (2) 187-190.

Hoogendoorn B, Norton N, Kirov G, Williams N, Hamshere M L, Spurlock G, Austin J, Stephens M K, Buckland P R, Owen M J and O'Donovan M C (2000) Cheap, accurate and rapid allele frequency estimation of single nucleotide polymorphisms by primer extension and DHPLC in DNA pools. *Hum Genet* 107 (5) 488-493.

Hsu T M, Chen X, Duan S, Miller R D and Kwok P Y (2001) Universal SNP genotyping assay with fluorescence polarization detection. *Biotechniques* 31 (3) 560, 562, 564-568, passim.

Huber A B, Kolodkin A L, Ginty D D and Cloutier J F (2003) Signaling at the growth cone: ligand-receptor complexes and the control of axon growth and guidance. *Annu Rev Neurosci* 26 509-563.

Hyde T M, Ziegler J C and Weinberger D R (1992) Psychiatric disturbances in metachromatic leukodystrophy. Insights into the neurobiology of psychosis. *Arch Neurol* 49 (4) 401-406.

Inoue K, Khajavi M, Ohyama T, Hirabayashi S, Wilson J, Reggin J D, Mancias P, Butler I J, Wilkinson M F, Wegner M and Lupski J R (2004) Molecular mechanism for distinct neurological phenotypes conveyed by allelic truncating mutations. *Nat Genet* 36 (4) 361-369.

Ioannidis J P, Ntzani E E, Trikalinos T A and Contopoulos-Ioannidis D G (2001) Replication validity of genetic association studies. *Nat Genet* 29 (3) 306-309.

Isles A R and Wilkinson L S (2000) Imprinted genes, cognition and behaviour. *Trends Cogn Sci* 4 (8) 309-318.

Iwamoto K, Bundo M and Kato T (2005a) Altered expression of mitochondria-related genes in postmortem brains of patients with bipolar disorder or schizophrenia, as revealed by large-scale DNA microarray analysis. *Hum Mol Genet* 14 (2) 241-253.

Iwamoto K, Bundo M, Yamada K, Takao H, Iwayama-Shigeno Y, Yoshikawa T and Kato T (2005b) DNA methylation status of SOX10 correlates with its downregulation and oligodendrocyte dysfunction in schizophrenia. *J Neurosci* 25 (22) 5376-5381.

Jablensky A (1997) The 100-year epidemiology of schizophrenia. *Schizophr Res* 28 (2-3) 111-125.

Jablensky A, Sartorius N, Ernberg G, Anker M, Korten A, Cooper J E, Day R and Bertelsen A (1992) Schizophrenia: manifestations, incidence and course in different cultures. A World Health Organization ten-country study. *Psychol Med Monogr Suppl* 20 1-97.

Jakob H and Beckmann H (1986) Prenatal developmental disturbances in the limbic allocortex in schizophrenics. *J Neural Transm* 65 (3-4) 303-326.

Jaskiw G E, Juliano D M, Goldberg T E, Hertzman M, Urow-Hamell E and Weinberger D R (1994) Cerebral ventricular enlargement in schizophreniform disorder does not progress. A seven year follow-up study. *Schizophr Res* 14 (1) 23-28.

Jeffreys A J, Kauppi L and Neumann R (2001) Intensely punctate meiotic recombination in the class II region of the major histocompatibility complex. *Nat Genet* 29 (2) 217-222.

Johns T G and Bernard C C (1999) The structure and function of myelin oligodendrocyte glycoprotein. *J Neurochem* 72 (1) 1-9.

Johnson G C, Esposito L, Barratt B J, Smith A N, Heward J, Di Genova G, Ueda H, Cordell H J, Eaves I A, Dudbridge F, Twells R C, Payne F, Hughes W, Nutland S, Stevens H, Carr P, Tuomilehto-Wolf E, Tuomilehto J, Gough S C, Clayton D G and Todd J A (2001) Haplotype tagging for the identification of common disease genes. *Nat Genet* 29 (2) 233-237.

Johnston-Wilson N L, Sims C D, Hofmann J P, Anderson L, Shore A D, Torrey E F and Yolken R H (2000) Disease-specific alterations in frontal cortex brain proteins in schizophrenia, bipolar disorder, and major depressive disorder. The Stanley Neuropathology Consortium. *Mol Psychiatry* 5 (2) 142-149.

Jones A C, Austin J, Hansen N, Hoogendoorn B, Oefner P J, Cheadle J P and O'Donovan M C (1999) Optimal temperature selection for mutation detection by denaturing HPLC and comparison to single-stranded conformation polymorphism and heteroduplex analysis. *Clin Chem* 45 (8 Pt 1) 1133-1140.

Jones P, Rodgers B, Murray R and Marmot M (1994) Child development risk factors for adult schizophrenia in the British 1946 birth cohort. *Lancet* 344 (8934) 1398-1402.

Karnezis T, Mandemakers W, McQualter J L, Zheng B, Ho P P, Jordan K A, Murray B M, Barres B, Tessier-Lavigne M and Bernard C C (2004) The neurite outgrowth inhibitor Nogo A is involved in autoimmune-mediated demyelination. *Nat Neurosci* 7 (7) 736-744.

Katsel P, Davis K L, Gorman J M and Haroutunian V (2005a) Variations in differential gene expression patterns across multiple brain regions in schizophrenia. *Schizophr Res* 77 (2-3) 241-252.

Katsel P, Davis K L and Haroutunian V (2005b) Variations in myelin and oligodendrocyte-related gene expression across multiple brain regions in schizophrenia: A gene ontology study. *Schizophr Res.*

Katsel P L, Davis K L and Haroutunian V (2005c) Large-scale microarray studies of gene expression in multiple regions of the brain in schizophrenia and Alzheimer's disease. *Int Rev Neurobiol* 63 41-82.

Kim J E, Li S, GrandPre T, Qiu D and Strittmatter S M (2003) Axon regeneration in young adult mice lacking Nogo-A/B. *Neuron* 38 (2) 187-199.

Kim J S, Kornhuber H H, Schmid-Burgk W and Holzmuller B (1980) Low cerebrospinal fluid glutamate in schizophrenic patients and a new hypothesis on schizophrenia. *Neurosci Lett* 20 (3) 379-382.

Kirov G, Ivanov D, Williams N M, Preece A, Nikolov I, Milev R, Koleva S, Dimitrova A, Toncheva D, O'Donovan M C and Owen M J (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) 971-975.

Klei L, Bacanu S A, Myles-Worsley M, Galke B, Xie W, Tiobech J, Otto C, Roeder K, Devlin B and Byerley W (2005) Linkage analysis of a completely ascertained sample of familial schizophrenics and bipolars from Palau, Micronesia. *Hum Genet* 117 (4) 349-356.

Korostishevsky M, Kaganovich M, Cholostoy A, Ashkenazi M, Ratner Y, Dahary D, Bernstein J, Bening-Abu-Shach U, Ben-Asher E, Lancet D, Ritsner M and Navon R (2004) Is the G72/G30 locus associated with schizophrenia? single nucleotide polymorphisms, haplotypes, and gene expression analysis. *Biol Psychiatry* 56 (3) 169-176.

Korostishevsky M, Kremer I, Kaganovich M, Cholostoy A, Murad I, Muhaheed M, Bannoura I, Rietschel M, Dobrusin M, Bening-Abu-Shach U, Belmaker R H, Maier W, Ebstein R P and Navon R (2006) Transmission disequilibrium and haplotype analyses of the G72/G30 locus: suggestive linkage to schizophrenia in Palestinian Arabs living in the North of Israel. *Am J Med Genet B Neuropsychiatr Genet* 141 (1) 91-95.

Kovelman J A and Scheibel A B (1984) A neurohistological correlate of schizophrenia. *Biol Psychiatry* 19 (12) 1601-1621.

Kraepelin E (1896) *Compendium der Psychiatrie*. 5th ed. Leipzig: Barth.

Krimer L S, Herman M M, Saunders R C, Boyd J C, Hyde T M, Carter J M, Kleinman J E and Weinberger D R (1997) A qualitative and quantitative analysis of the entorhinal cortex in schizophrenia. *Cereb Cortex* 7 (8) 732-739.

Kruglyak L (1999) Prospects for whole-genome linkage disequilibrium mapping of common disease genes. *Nat Genet* 22 (2) 139-144.

Kruglyak L (2005) Power tools for human genetics. *Nat Genet*.

Kruys V, Marinx O, Shaw G, Deschamps J and Huez G (1989) Translational blockade imposed by cytokine-derived UA-rich sequences. *Science* 245 (4920) 852-855.

Kubicki M, McCarley R, Westin C F, Park H J, Maier S, Kikinis R, Jolesz F A and Shenton M E (2005a) A review of diffusion tensor imaging studies in schizophrenia. *J Psychiatr Res*.

Kubicki M, Park H, Westin C F, Nestor P G, Mulkern R V, Maier S E, Niznikiewicz M, Connor E E, Levitt J J, Frumin M, Kikinis R, Jolesz F A, McCarley R W and Shenton M E (2005b) DTI and MTR abnormalities in schizophrenia: analysis of white matter integrity. *Neuroimage* 26 (4) 1109-1118.

Kuersten S and Goodwin E B (2003) The power of the 3' UTR: translational control and development. *Nat Rev Genet* 4 (8) 626-637.

Kumra S, Ashtari M, McMeniman M, Vogel J, Augustin R, Becker D E, Nakayama E, Gyato K, Kane J M, Lim K and Szeszko P (2004) Reduced frontal white matter integrity in early-onset schizophrenia: a preliminary study. *Biol Psychiatry* 55 (12) 1138-1145.

Kwok P Y (2002) SNP genotyping with fluorescence polarization detection. *Hum Mutat* 19 (4) 315-323.

Lane A, Kinsella A, Murphy P, Byrne M, Keenan J, Colgan K, Cassidy B, Sheppard N, Horgan R, Waddington J L, Larkin C and O'Callaghan E (1997) The anthropometric assessment of dysmorphic features in schizophrenia as an index of its developmental origins. *Psychol Med* 27 (5) 1155-1164.

Lappe-Siefke C, Goebbels S, Gravel M, Nicksch E, Lee J, Braun P E, Griffiths I R and Nave K A (2003) Disruption of Cnp1 uncouples oligodendroglial functions in axonal support and myelination. *Nat Genet* 33 (3) 366-374.

Larocque D, Galarneau A, Liu H N, Scott M, Almazan G and Richard S (2005) Protection of p27(Kip1) mRNA by quaking RNA binding proteins promotes oligodendrocyte differentiation. *Nat Neurosci* 8 (1) 27-33.

Law A J, Lipska B K, Weickert C S, Hyde T M, Straub R E, Hashimoto R, Harrison P J, Kleinman J E and Weinberger D R (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) 6747-6752.

Lawrie S M and Abukmeil S S (1998) Brain abnormality in schizophrenia. A systematic and quantitative review of volumetric magnetic resonance imaging studies. *Br J Psychiatry* 172 110-120.

Lawrie S M, Whalley H, Kestelman J N, Abukmeil S S, Byrne M, Hodges A, Rimmington J E, Best J J, Owens D G and Johnstone E C (1999) Magnetic resonance imaging of brain in people at high risk of developing schizophrenia. *Lancet* 353 (9146) 30-33.

Lee J, Gravel M, Zhang R, Thibault P and Braun P E (2005) Process outgrowth in oligodendrocytes is mediated by CNP, a novel microtubule assembly myelin protein. *J Cell Biol* 170 (4) 661-673.

Lee J, O'Neill R C, Park M W, Gravel M and Braun P E (2006) Mitochondrial localization of CNP2 is regulated by phosphorylation of the N-terminal targeting signal by PKC: implications of a mitochondrial function for CNP2 in glial and non-glial cells. *Mol Cell Neurosci* 31 (3) 446-462.

Levine M and Tjian R (2003) Transcription regulation and animal diversity. *Nature* 424 (6945) 147-151.

Lewis C M, Levinson D F, Wise L H, DeLisi L E, Straub R E, Hovatta I, Williams N M, Schwab S G, Pulver A E, Faraone S V, Brzustowicz L M, Kaufmann C A, Garver D L, Gurling H M, Lindholm E, Coon H, Moises H W, Byerley W, Shaw S H, Mesen A, Sherrington R, O'Neill F A, Walsh D, Kendler K S, Ekelund J, Paunio T, Lonnqvist J, Peltonen L, O'Donovan M C, Owen M J, Wildenauer D B, Maier W, Nestadt G, Blouin J L, Antonarakis S E, Mowry B J, Silverman J M, Crowe R R, Cloninger C R, Tsuang M T, Malaspina D, Harkavy-Friedman J M, Svrakic D M, Bassett A S, Holcomb J, Kalsi G, McQuillin A, Brynjolfson J, Sigmundsson T, Petursson H, Jazin E, Zoega T and Helgason T (2003) Genome scan meta-analysis of schizophrenia and bipolar disorder, part II: Schizophrenia. *Am J Hum Genet* 73 (1) 34-48.

Li D, Collier D A and He L (2006a) Meta-analysis shows strong positive association of the neuregulin 1 (NRG1) gene with schizophrenia. *Hum Mol Genet* 15 (12) 1995-2002.

Li M, Atmaca-Sonmez P, Othman M, Branham K E, Khanna R, Wade M S, Li Y, Liang L, Zareparsi S, Swaroop A and Abecasis G R (2006b) CFH haplotypes without the Y402H coding variant show strong association with susceptibility to age-related macular degeneration. *Nat Genet* 38 (9) 1049-1054.

Li R, Johnson A B, Salomons G, Goldman J E, Naidu S, Quinlan R, Cree B, Ruyle S Z, Banwell B, D'Hooghe M, Siebert J R, Rolf C M, Cox H, Reddy A, Gutierrez-Solana L G, Collins A, Weller R O, Messing A, van der Knaap M S and Brenner M (2005) Glial fibrillary acidic protein mutations in infantile, juvenile, and adult forms of Alexander disease. *Ann Neurol* 57 (3) 310-326.

Liddle P a P, C. (2003) Brain imaging in schizophrenia. In: Hirsch, S R a W D (ed.) *Schizophrenia.* (2nd edn.) Oxford: Blackwell Science Ltd.

Liedtke W, Edelmann W, Bieri P L, Chiu F C, Cowan N J, Kucherlapati R and Raine C S (1996) GFAP is necessary for the integrity of CNS white matter architecture and long-term maintenance of myelination. *Neuron* 17 (4) 607-615.

Lim K O, Hedehus M, Moseley M, de Crespigny A, Sullivan E V and Pfefferbaum A (1999) Compromised white matter tract integrity in schizophrenia inferred from diffusion tensor imaging. *Arch Gen Psychiatry* 56 (4) 367-374.

Lindholm E, Ekholm B, Shaw S, Jalonen P, Johansson G, Pettersson U, Sherrington R, Adolfsson R and Jazin E (2001) A schizophrenia-susceptibility locus at 6q25, in one of the world's largest reported pedigrees. *Am J Hum Genet* 69 (1) 96-105.

Liu B P, Fournier A, GrandPre T and Strittmatter S M (2002) Myelin-associated glycoprotein as a functional ligand for the Nogo-66 receptor. *Science* 297 (5584) 1190-1193.

Liu X, Qin W, He G, Yang Y, Chen Q, Zhou J, Li D, Gu N, Xu Y, Feng G, Sang H, Hao X, Zhang K, Wang S and He L (2005) A family-based association study of the MOG gene with schizophrenia in the Chinese population. *Schizophr Res* 73 (2-3) 275-280.

Lohoff F W, Sander T, Ferraro T N, Dahl J P, Gallinat J and Berrettini W H (2005) Confirmation of association between the Val66Met polymorphism in the brain-derived neurotrophic factor (BDNF) gene and bipolar I disorder. *Am J Med Genet B Neuropsychiatr Genet* 139 (1) 51-53.

Low S C and Berry M J (1996) Knowing when not to stop: selenocysteine incorporation in eukaryotes. *Trends Biochem Sci* 21 (6) 203-208.

Maller J, George S, Purcell S, Fagerness J, Altshuler D, Daly M J and Seddon J M (2006) Common variation in three genes, including a noncoding variant in CFH, strongly influences risk of age-related macular degeneration. *Nat Genet* 38 (9) 1055-1059.

Marenco S and Weinberger D R (2000) The neurodevelopmental hypothesis of schizophrenia: following a trail of evidence from cradle to grave. *Dev Psychopathol* 12 (3) 501-527.

Marta C B, Montano M B, Taylor C M, Taylor A L, Bansal R and Pfeiffer S E (2005) Signaling cascades activated upon antibody cross-linking of myelin oligodendrocyte glycoprotein: potential implications for multiple sclerosis. *J Biol Chem* 280 (10) 8985-8993.

Martinez M, Goldin L R, Cao Q, Zhang J, Sanders A R, Nancarrow D J, Taylor J M, Levinson D F, Kirby A, Crowe R R, Andreasen N C, Black D W, Silverman J M, Lennon D P, Nertney D A, Brown D M, Mowry B J, Gershon E S and Gejman P V (1999) Follow-up study on a susceptibility locus for schizophrenia on chromosome 6q. *Am J Med Genet* 88 (4) 337-343.

McGhee K A, Morris D W, Schwaiger S, Nangle J M, Donohoe G, Clarke S, Meagher D, Quinn J, Scully P, Waddington J L, Gill M and Corvin A (2005) Investigation of the apolipoprotein-L (APOL) gene family and schizophrenia using a novel DNA pooling strategy for public database SNPs. *Schizophr Res* 76 (2-3) 231-238.

McGrath J, Saha S, Welham J, El Saadi O, MacCauley C and Chant D (2004) A systematic review of the incidence of schizophrenia: the distribution of rates and the influence of sex, urbanicity, migrant status and methodology. *BMC Med* 2 (1) 13.

McGuffin P, Owen M J and Farmer A E (1995) Genetic basis of schizophrenia. *Lancet* 346 (8976) 678-682.

McKerracher L, David S, Jackson D L, Kottis V, Dunn R J and Braun P E (1994) Identification of myelin-associated glycoprotein as a major myelin-derived inhibitor of neurite growth. *Neuron* 13 (4) 805-811.

Menn B, Garcia-Verdugo J M, Yaschine C, Gonzalez-Perez O, Rowitch D and Alvarez-Buylla A (2006) Origin of oligodendrocytes in the subventricular zone of the adult brain. *J Neurosci* 26 (30) 7907-7918.

Middleton F A, Mirnics K, Pierri J N, Lewis D A and Levitt P (2002) Gene expression profiling reveals alterations of specific metabolic pathways in schizophrenia. *J Neurosci* 22 (7) 2718-2729.

Millar J K, Wilson-Annan J C, Anderson S, Christie S, Taylor M S, Semple C A, Devon R S, Clair D M, Muir W J, Blackwood D H and Porteous D J (2000) Disruption of two novel genes by a translocation co-segregating with schizophrenia. *Hum Mol Genet* 9 (9) 1415-1423.

Mimmack M L, Ryan M, Baba H, Navarro-Ruiz J, Iritani S, Faull R L, McKenna P J, Jones P B, Arai H, Starkey M, Emson P C and Bahn S (2002) Gene expression analysis in schizophrenia: reproducible up-regulation of several members of the apolipoprotein L family located in a high-susceptibility locus for schizophrenia on chromosome 22. *Proc Natl Acad Sci U S A* 99 (7) 4680-4685.

Mingorance A, Fontana X, Sole M, Burgaya F, Urena J M, Teng F Y, Tang B L, Hunt D, Anderson P N, Bethea J R, Schwab M E, Soriano E and del Rio J A (2004) Regulation of Nogo and Nogo receptor during the development of the entorhino-hippocampal pathway and after adult hippocampal lesions. *Mol Cell Neurosci* 26 (1) 34-49.

Mirnics K, Middleton F A, Marquez A, Lewis D A and Levitt P (2000) Molecular characterization of schizophrenia viewed by microarray analysis of gene expression in prefrontal cortex. *Neuron* 28 (1) 53-67.

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

Moghaddam B a K, JH. (2003) The neurochemistry of schizophrenia. In: Hirsch, S R a W D (ed.) *Schizophrenia.* (2nd edn.) Oxford: Blackwell Science Ltd.

Mollaaghababa R and Pavan W J (2003) The importance of having your SOX on: role of SOX10 in the development of neural crest-derived melanocytes and glia. *Oncogene* 22 (20) 3024-3034.

Monajemi H, Fontijn R D, Pannekoek H and Horrevoets A J (2002) The apolipoprotein L gene cluster has emerged recently in evolution and is expressed in human vascular tissue. *Genomics* 79 (4) 539-546.

Morison I M, Ramsay J P and Spencer H G (2005) A census of mammalian imprinting. *Trends Genet* 21 (8) 457-465.

Mortensen P B, Pedersen C B, Westergaard T, Wohlfahrt J, Ewald H, Mors O, Andersen P K and Melbye M (1999) Effects of family history and place and season of birth on the risk of schizophrenia. *N Engl J Med* 340 (8) 603-608.

Mueser K T and McGurk S R (2004) Schizophrenia. *Lancet* 363 (9426) 2063-2072.

Munafo M R, Thiselton D L, Clark T G and Flint J (2006) Association of the NRG1 gene and schizophrenia: a meta-analysis. *Mol Psychiatry* 11 (6) 539-546.

Munroe D and Jacobson A (1990) mRNA poly(A) tail, a 3' enhancer of translational initiation. *Mol Cell Biol* 10 (7) 3441-3455.

Murphy K C (2002) Schizophrenia and velo-cardio-facial syndrome. *Lancet* 359 (9304) 426-430.

Murphy K C, Jones L A and Owen M J (1999) High rates of schizophrenia in adults with velo-cardio-facial syndrome. *Arch Gen Psychiatry* 56 (10) 940-945.

Murray R M, Jones P. B, Susser E. S, van Os J, Cannon M. (2002) *The Epidemiology of Schizophrenia.* Cambridge: Cambridge University Press.

Murray R M and Van Os J (1998) Predictors of outcome in schizophrenia. *J Clin Psychopharmacol* 18 (2 Suppl 1) 2S-4S.

Myakishev M V, Khripin Y, Hu S and Hamer D H (2001) High-throughput SNP genotyping by allele-specific PCR with universal energy-transfer-labeled primers. *Genome Res* 11 (1) 163-169.

Neves-Pereira M, Cheung J K, Pasdar A, Zhang F, Breen G, Yates P, Sinclair M, Crombie C, Walker N and St Clair D M (2005) BDNF gene is a risk factor for schizophrenia in a Scottish population. *Mol Psychiatry* 10 (2) 208-212.

Niculescu A B, 3rd, Segal D S, Kuczenski R, Barrett T, Hauger R L and Kelsoe J R (2000) Identifying a series of candidate genes for mania and psychosis: a convergent functional genomics approach. *Physiol Genomics* 4 (1) 83-91.

Norton N, Williams N M, Williams H J, Spurlock G, Kirov G, Morris D W, Hoogendoorn B, Owen M J and O'Donovan M C (2002) Universal, robust, highly quantitative SNP allele frequency measurement in DNA pools. *Hum Genet* 110 (5) 471-478.

Novak G, Kim D, Seeman P and Tallerico T (2002) Schizophrenia and Nogo: elevated mRNA in cortex, and high prevalence of a homozygous CAA insert. *Brain Res Mol Brain Res* 107 (2) 183-189.

Numakawa T, Yagasaki Y, Ishimoto T, Okada T, Suzuki T, Iwata N, Ozaki N, Taguchi T, Tatsumi M, Kamijima K, Straub R E, Weinberger D R, Kunugi H and Hashimoto R (2004) Evidence of novel neuronal functions of dysbindin, a susceptibility gene for schizophrenia. *Hum Mol Genet* 13 (21) 2699-2708.

Nunes S O, Borelli S D, Matsuo T, Watanabe M A and Itano E N (2005) The association of the HLA in patients with schizophrenia, schizoaffective disorder, and in their biological relatives. *Schizophr Res* 76 (2-3) 195-198.

O'Donovan M C, Oefner P J, Roberts S C, Austin J, Hoogendoorn B, Guy C, Speight G, Upadhyaya M, Sommer S S and McGuffin P (1998) Blind analysis of denaturing high-performance liquid chromatography as a tool for mutation detection. *Genomics* 52 (1) 44-49.

O'Neill R C, Minuk J, Cox M E, Braun P E and Gravel M (1997) CNP2 mRNA directs synthesis of both CNP1 and CNP2 polypeptides. *J Neurosci Res* 50 (2) 248-257.

Ordovas J M, Corella D, Demissie S, Cupples L A, Couture P, Coltell O, Wilson P W, Schaefer E J and Tucker K L (2002) Dietary fat intake determines the effect of a common polymorphism in the hepatic lipase gene promoter on high-density lipoprotein metabolism: evidence of a strong dose effect in this gene-nutrient interaction in the Framingham Study. *Circulation* 106 (18) 2315-2321.

Owen M J, Craddock N and O'Donovan M C (2005) Schizophrenia: genes at last? *Trends Genet* 21 (9) 518-525.

Owen M J, Williams N M and O'Donovan M C (2004) The molecular genetics of schizophrenia: new findings promise new insights. *Mol Psychiatry* 9 (1) 14-27.

Pagani F, Raponi M and Baralle F E (2005) Synonymous mutations in CFTR exon 12 affect splicing and are not neutral in evolution. *Proc Natl Acad Sci U S A* 102 (18) 6368-6372.

Parwani A, Duncan E J, Bartlett E, Madonick S H, Efferen T R, Rajan R, Sanfilipo M, Chappell P B, Chakravorty S, Gonzenbach S, Ko G N and Rotrosen J P (2000) Impaired prepulse inhibition of acoustic startle in schizophrenia. *Biol Psychiatry* 47 (7) 662-669.

Pastinen T, Sladek R, Gurd S, Sammak A, Ge B, Lepage P, Lavergne K, Villeneuve A, Gaudin T, Brandstrom H, Beck A, Verner A, Kingsley J, Harmsen E, Labuda D, Morgan K, Vohl M C, Naumova A K, Sinnett D and Hudson T J (2004) A survey of genetic and epigenetic variation affecting human gene expression. *Physiol Genomics* 16 (2) 184-193.

Pearlson G D and Marsh L (1999) Structural brain imaging in schizophrenia: a selective review. *Biol Psychiatry* 46 (5) 627-649.

Pedersen C B and Mortensen P B (2001) Evidence of a dose-response relationship between urbanicity during upbringing and schizophrenia risk. *Arch Gen Psychiatry* 58 (11) 1039-1046.

Peirano R I, Goerich D E, Riethmacher D and Wegner M (2000) Protein zero gene expression is regulated by the glial transcription factor Sox10. *Mol Cell Biol* 20 (9) 3198-3209.

Peirce T R, Bray N J, Williams N M, Norton N, Moskvina V, Preece A, Haroutunian V, Buxbaum J D, Owen M J and O'Donovan M C (2006) Convergent evidence for

2',3'-cyclic nucleotide 3'-phosphodiesterase as a possible susceptibility gene for schizophrenia. *Arch Gen Psychiatry* 63 (1) 18-24.

Petryshen T L, Middleton F A, Kirby A, Aldinger K A, Purcell S, Tahl A R, Morley C P, McGann L, Gentile K L, Rockwell G N, Medeiros H M, Carvalho C, Macedo A, Dourado A, Valente J, Ferreira C P, Patterson N J, Azevedo M H, Daly M J, Pato C N, Pato M T and Sklar P (2005) Support for involvement of neuregulin 1 in schizophrenia pathophysiology. *Mol Psychiatry* 10 (4) 366-374, 328.

Pfaff C L, Parra E J, Bonilla C, Hiester K, McKeigue P M, Kamboh M I, Hutchinson R G, Ferrell R E, Boerwinkle E and Shriver M D (2001) Population structure in admixed populations: effect of admixture dynamics on the pattern of linkage disequilibrium. *Am J Hum Genet* 68 (1) 198-207.

Phillips M S, Lawrence R, Sachidanandam R, Morris A P, Balding D J, Donaldson M A, Studebaker J F, Ankener W M, Alfisi S V, Kuo F S, Camisa A L, Pazorov V, Scott K E, Carey B J, Faith J, Katari G, Bhatti H A, Cyr J M, Derohannessian V, Elosua C, Forman A M, Grecco N M, Hock C R, Kuebler J M, Lathrop J A, Mockler M A, Nachtman E P, Restine S L, Varde S A, Hozza M J, Gelfand C A, Broxholme J, Abecasis G R, Boyce-Jacino M T and Cardon L R (2003) Chromosome-wide distribution of haplotype blocks and the role of recombination hot spots. *Nat Genet* 33 (3) 382-387.

Pingault V, Bondurand N, Le Caignec C, Tardieu S, Lemort N, Dubourg O, Le Guern E, Goossens M and Boespflug-Tanguy O (2001) The SOX10 transcription factor: evaluation as a candidate gene for central and peripheral hereditary myelin disorders. *J Neurol* 248 (6) 496-499.

Prabakaran S, Swatton J E, Ryan M M, Huffaker S J, Huang J T, Griffin J L, Wayland M, Freeman T, Dudbridge F, Lilley K S, Karp N A, Hester S, Tkachev D, Mimmack M L, Yolken R H, Webster M J, Torrey E F and Bahn S (2004) Mitochondrial dysfunction in schizophrenia: evidence for compromised brain metabolism and oxidative stress. *Mol Psychiatry* 9 (7) 684-697, 643.

Prinjha R, Moore S E, Vinson M, Blake S, Morrow R, Christie G, Michalovich D, Simmons D L and Walsh F S (2000) Inhibitor of neurite outgrowth in humans. *Nature* 403 (6768) 383-384.

Pritchard J K and Przeworski M (2001) Linkage disequilibrium in humans: models and data. *Am J Hum Genet* 69 (1) 1-14.

Qin W, Gao J, Xing Q, Yang J, Qian X, Li X, Guo Z, Chen H, Wang L, Huang X, Gu N, Feng G and He L (2005) A family-based association study of PLP1 and schizophrenia. *Neurosci Lett* 375 (3) 207-210.

Raisman G (2004) Myelin inhibitors: does NO mean GO? *Nat Rev Neurosci* 5 (2) 157-161.

Rajakumar N, Leung L S, Ma J, Rajakumar B and Rushlow W (2004) Altered neurotrophin receptor function in the developing prefrontal cortex leads to adult-onset dopaminergic hyperresponsivity and impaired prepulse inhibition of acoustic startle. *Biol Psychiatry* 55 (8) 797-803.

Rajkowska G, Miguel-Hidalgo J J, Makkos Z, Meltzer H, Overholser J and Stockmeier C (2002) Layer-specific reductions in GFAP-reactive astroglia in the dorsolateral prefrontal cortex in schizophrenia. *Schizophr Res* 57 (2-3) 127-138.

Rapoport J L, Addington A M, Frangou S and Psych M R (2005) The neurodevelopmental model of schizophrenia: update 2005. *Mol Psychiatry* 10 (5) 434-449.

Rasband M N, Tayler J, Kaga Y, Yang Y, Lappe-Siefke C, Nave K A and Bansal R (2005) CNP is required for maintenance of axon-glia interactions at nodes of Ranvier in the CNS. *Glia* 50 (1) 86-90.

Raz S and Raz N (1990) Structural brain abnormalities in the major psychoses: a quantitative review of the evidence from computerized imaging. *Psychol Bull* 108 (1) 93-108.

Rees M I, Fenton I, Williams N M, Holmans P, Norton N, Cardno A, Asherson P, Spurlock G, Roberts E, Parfitt E, Mant R, Vallada H, Dawson E, Li M W, Collier D A, Powell J F, Nanko S, Gill M, McGuffin P and Owen M J (1999) Autosome search for schizophrenia susceptibility genes in multiply affected families. *Mol Psychiatry* 4 (4) 353-359.

Reich D E, Cargill M, Bolk S, Ireland J, Sabeti P C, Richter D J, Lavery T, Kouyoumjian R, Farhadian S F, Ward R and Lander E S (2001) Linkage disequilibrium in the human genome. *Nature* 411 (6834) 199-204.

Reischies F M, Baum K, Brau H, Hedde J P and Schwindt G (1988) Cerebral magnetic resonance imaging findings in multiple sclerosis. Relation to disturbance of affect, drive, and cognition. *Arch Neurol* 45 (10) 1114-1116.

Reynolds G (2002) Neurochemistry of Schizophrenia. *Psychiatry* 1 (9) 22-25.

Risch N (1990a) Genetic linkage and complex diseases, with special reference to psychiatric disorders. *Genet Epidemiol* 7 (1) 3-16; discussion 17-45.

Risch N (1990b) Linkage strategies for genetically complex traits. II. The power of affected relative pairs. *Am J Hum Genet* 46 (2) 229-241.

Risch N and Merikangas K (1996) The future of genetic studies of complex human diseases. *Science* 273 (5281) 1516-1517.

Risch N J (2000) Searching for genetic determinants in the new millennium. *Nature* 405 (6788) 847-856.

Saher G, Brugger B, Lappe-Siefke C, Mobius W, Tozawa R, Wehr M C, Wieland F, Ishibashi S and Nave K A (2005) High cholesterol level is essential for myelin membrane growth. *Nat Neurosci* 8 (4) 468-475.

Saleh M C, Espinosa de los Monteros A, de Arriba Zerpa G A, Fontaine I, Piaud O, Djordjijevic D, Baroukh N, Garcia Otin A L, Ortiz E, Lewis S, Fiette L, Santambrogio P, Belzung C, Connor J R, de Vellis J, Pasquini J M, Zakin M M, Baron B and Guillou F (2003) Myelination and motor coordination are increased in transferrin transgenic mice. *J Neurosci Res* 72 (5) 587-594.

Sanfilipo M, Lafargue T, Rusinek H, Arena L, Loneragan C, Lautin A, Feiner D, Rotrosen J and Wolkin A (2000) Volumetric measure of the frontal and temporal lobe regions in schizophrenia: relationship to negative symptoms. *Arch Gen Psychiatry* 57 (5) 471-480.

Sasaki T, Matsushita M, Nanko S, Fukuda R, Kennedy J L and Tokunaga K (1999) Schizophrenia and the HLA-DRB1 gene in the Japanese population. *Am J Psychiatry* 156 (5) 771-773.

Satoh J, Onoue H, Arima K and Yamamura T (2005) Nogo-A and nogo receptor expression in demyelinating lesions of multiple sclerosis. *J Neuropathol Exp Neurol* 64 (2) 129-138.

Schaal T D and Maniatis T (1999) Multiple distinct splicing enhancers in the protein-coding sequences of a constitutively spliced pre-mRNA. *Mol Cell Biol* 19 (1) 261-273.

Schumacher J, Jamra R A, Becker T, Ohlraun S, Klopp N, Binder E B, Schulze T G, Deschner M, Schmal C, Hofels S, Zobel A, Illig T, Propping P, Holsboer F, Rietschel M, Nothen M M and Cichon S (2005) Evidence for a relationship between genetic variants at the brain-derived neurotrophic factor (BDNF) locus and major depression. *Biol Psychiatry* 58 (4) 307-314.

Schumacher J, Jamra R A, Freudenberg J, Becker T, Ohlraun S, Otte A C, Tullius M, Kovalenko S, Bogaert A V, Maier W, Rietschel M, Propping P, Nothen M M and Cichon S (2004) Examination of G72 and D-amino-acid oxidase as genetic risk factors for schizophrenia and bipolar affective disorder. *Mol Psychiatry* 9 (2) 203-207.

Schwab S G, Albus M, Hallmayer J, Honig S, Borrmann M, Lichtermann D, Ebstein R P, Ackenheil M, Lerer B, Risch N and et al. (1995) Evaluation of a susceptibility gene for schizophrenia on chromosome 6p by multipoint affected sib-pair linkage analysis. *Nat Genet* 11 (3) 325-327.

Schwab S G, Eckstein G N, Hallmayer J, Lerer B, Albus M, Borrmann M, Lichtermann D, Ertl M A, Maier W and Wildenauer D B (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) 156-160.

Schwab S G, Hallmayer J, Albus M, Lerer B, Eckstein G N, Borrmann M, Segman R H, Hanses C, Freymann J, Yakir A, Trixler M, Falkai P, Rietschel M, Maier W and Wildenauer D B (2000) A genome-wide autosomal screen for schizophrenia susceptibility loci in 71 families with affected siblings: support for loci on chromosome 10p and 6. *Mol Psychiatry* 5 (6) 638-649.

Schwab S G, Hallmayer J, Albus M, Lerer B, Hanses C, Kanyas K, Segman R, Borrman M, Dreikorn B, Lichtermann D, Rietschel M, Trixler M, Maier W and Wildenauer D B (1998) Further evidence for a susceptibility locus on chromosome 10p14-p11 in 72 families with schizophrenia by nonparametric linkage analysis. *Am J Med Genet* 81 (4) 302-307.

Schwab S G, Knapp M, Mondabon S, Hallmayer J, Borrmann-Hassenbach M, Albus M, Lerer B, Rietschel M, Trixler M, Maier W and Wildenauer D B (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) 185-190.

Sham P a M, P. (2003) Linkage and Association. In: McGuffin, P, Owen, MJ, and Gottesman, II. (ed.) *Psychiatric Genetics and Genomics.* (1st edn.) Oxford: Oxford University Press.

Sharpley M, Hutchinson G, McKenzie K and Murray R M (2001) Understanding the excess of psychosis among the African-Caribbean population in England. Review of current hypotheses. *Br J Psychiatry Suppl* 40 s60-68.

Shenton M E, Dickey C C, Frumin M and McCarley R W (2001) A review of MRI findings in schizophrenia. *Schizophr Res* 49 (1-2) 1-52.

Sherman D L and Brophy P J (2005) Mechanisms of axon ensheathment and myelin growth. *Nat Rev Neurosci* 6 (9) 683-690.

Shih R A, Belmonte P L and Zandi P P (2004) A review of the evidence from family, twin and adoption studies for a genetic contribution to adult psychiatric disorders. *Int Rev Psychiatry* 16 (4) 260-283.

Sigmundsson T, Suckling J, Maier M, Williams S, Bullmore E, Greenwood K, Fukuda R, Ron M and Toone B (2001) Structural abnormalities in frontal, temporal, and limbic regions and interconnecting white matter tracts in schizophrenic patients with prominent negative symptoms. *Am J Psychiatry* 158 (2) 234-243.

Simonen M, Pedersen V, Weinmann O, Schnell L, Buss A, Ledermann B, Christ F, Sansig G, van der Putten H and Schwab M E (2003) Systemic deletion of the myelin-associated outgrowth inhibitor Nogo-A improves regenerative and plastic responses after spinal cord injury. *Neuron* 38 (2) 201-211.

Sinibaldi L, De Luca A, Bellacchio E, Conti E, Pasini A, Paloscia C, Spalletta G, Caltagirone C, Pizzuti A and Dallapiccola B (2004) Mutations of the Nogo-66 receptor (RTN4R) gene in schizophrenia. *Hum Mutat* 24 (6) 534-535.

Sklar P, Gabriel S B, McInnis M G, Bennett P, Lim Y M, Tsan G, Schaffner S, Kirov G, Jones I, Owen M, Craddock N, DePaulo J R and Lander E S (2002) Family-based association study of 76 candidate genes in bipolar disorder: BDNF is a potential risk locus. Brain-derived neutrophic factor. *Mol Psychiatry* 7 (6) 579-593.

Sleutels F, Barlow D P and Lyle R (2000) The uniqueness of the imprinting mechanism. *Curr Opin Genet Dev* 10 (2) 229-233.

Sow A, Lamant M, Bonny J M, Larvaron P, Piaud O, Lecureuil C, Fontaine I, Saleh M C, Garcia Otin A L, Renou J P, Baron B, Zakin M and Guillou F (2006) Oligodendrocyte differentiation is increased in transferrin transgenic mice. *J Neurosci Res* 83 (3) 403-414.

Spielman R S, McGinnis R E and Ewens W J (1993) Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet* 52 (3) 506-516.

St Clair D, Blackwood D, Muir W, Carothers A, Walker M, Spowart G, Gosden C and Evans H J (1990) Association within a family of a balanced autosomal translocation with major mental illness. *Lancet* 336 (8706) 13-16.

St Clair D, Xu M, Wang P, Yu Y, Fang Y, Zhang F, Zheng X, Gu N, Feng G, Sham P and He L (2005) Rates of adult schizophrenia following prenatal exposure to the Chinese famine of 1959-1961. *Jama* 294 (5) 557-562.

Stark A K, Uylings H B, Sanz-Arigita E and Pakkenberg B (2004) Glial cell loss in the anterior cingulate cortex, a subregion of the prefrontal cortex, in subjects with schizophrenia. *Am J Psychiatry* 161 (5) 882-888.

Stefansson H, Sarginson J, Kong A, Yates P, Steinthorsdottir V, Gudfinnsson E, Gunnarsdottir S, Walker N, Petursson H, Crombie C, Ingason A, Gulcher J R, Stefansson K and St Clair D (2003) Association of neuregulin 1 with schizophrenia confirmed in a Scottish population. *Am J Hum Genet* 72 (1) 83-87.

Stefansson H, Sigurdsson E, Steinthorsdottir V, Bjornsdottir S, Sigmundsson T, Ghosh S, Brynjolfsson J, Gunnarsdottir S, Ivarsson O, Chou T T, Hjaltason O, Birgisdottir B, Jonsson H, Gudnadottir V G, Gudmundsdottir E, Bjornsson A, Ingvarsson B, Ingason A, Sigfusson S, Hardardottir H, Harvey R P, Lai D, Zhou M, Brunner D, Mutel V, Gonzalo A, Lemke G, Sainz J, Johannesson G, Andresson T, Gudbjartsson D, Manolescu A, Frigge M L, Gurney M E, Kong A, Gulcher J R, Petursson H and Stefansson K (2002) Neuregulin 1 and susceptibility to schizophrenia. *Am J Hum Genet* 71 (4) 877-892.

Stephens J C, Schneider J A, Tanguay D A, Choi J, Acharya T, Stanley S E, Jiang R, Messer C J, Chew A, Han J H, Duan J, Carr J L, Lee M S, Koshy B, Kumar A M, Zhang G, Newell W R, Windemuth A, Xu C, Kalbfleisch T S, Shaner S L, Arnold K, Schulz V, Drysdale C M, Nandabalan K, Judson R S, Ruano G and Vovis G F (2001) Haplotype variation and linkage disequilibrium in 313 human genes. *Science* 293 (5529) 489-493.

Stewart D G and Davis K L (2004) Possible contributions of myelin and oligodendrocyte dysfunction to schizophrenia. *Int Rev Neurobiol* 59 381-424.

Stolt C C, Rehberg S, Ader M, Lommes P, Riethmacher D, Schachner M, Bartsch U and Wegner M (2002) Terminal differentiation of myelin-forming oligodendrocytes depends on the transcription factor Sox10. *Genes Dev* 16 (2) 165-170.

Stolt C C, Schmitt S, Lommes P, Sock E and Wegner M (2005) Impact of transcription factor Sox8 on oligodendrocyte specification in the mouse embryonic spinal cord. *Dev Biol* 281 (2) 309-317.

Stram D O (2004) Tag SNP selection for association studies. *Genet Epidemiol* 27 (4) 365-374.

Straub R E, Jiang Y, MacLean C J, Ma Y, Webb B T, Myakishev M V, Harris-Kerr C, Wormley B, Sadek H, Kadambi B, Cesare A J, Gibberman A, Wang X, O'Neill F A, Walsh D and Kendler K S (2002a) 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) 337-348.

Straub R E, MacLean C J, Ma Y, Webb B T, Myakishev M V, Harris-Kerr C, Wormley B, Sadek H, Kadambi B, O'Neill F A, Walsh D and Kendler K S (2002b) Genome-wide scans of three independent sets of 90 Irish multiplex schizophrenia families and follow-up of selected regions in all families provides evidence for multiple susceptibility genes. *Mol Psychiatry* 7 (6) 542-559.

Straub R E, MacLean C J, Martin R B, Ma Y, Myakishev M V, Harris-Kerr C, Webb B T, O'Neill F A, Walsh D and Kendler K S (1998) A schizophrenia locus may be located in region 10p15-p11. *Am J Med Genet* 81 (4) 296-301.

Straub R E, MacLean C J, O'Neill F A, Burke J, Murphy B, Duke F, Shinkwin R, Webb B T, Zhang J, Walsh D and et al. (1995) A potential vulnerability locus for schizophrenia on chromosome 6p24-22: evidence for genetic heterogeneity. *Nat Genet* 11 (3) 287-293.

Straub R E, MacLean C J, O'Neill F A, Walsh D and Kendler K S (1997) Support for a possible schizophrenia vulnerability locus in region 5q22-31 in Irish families. *Mol Psychiatry* 2 (2) 148-155.

Strittmatter W J, Saunders A M, Schmechel D, Pericak-Vance M, Enghild J, Salvesen G S and Roses A D (1993) Apolipoprotein E: high-avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease. *Proc Natl Acad Sci U S A* 90 (5) 1977-1981.

Suddath R L, Christison G W, Torrey E F, Casanova M F and Weinberger D R (1990) Anatomical abnormalities in the brains of monozygotic twins discordant for schizophrenia. *N Engl J Med* 322 (12) 789-794.

Sugai T, Kawamura M, Iritani S, Araki K, Makifuchi T, Imai C, Nakamura R, Kakita A, Takahashi H and Nawa H (2004) Prefrontal abnormality of schizophrenia revealed by DNA microarray: impact on glial and neurotrophic gene expression. *Ann N Y Acad Sci* 1025 84-91.

Sun Z Y, Wei J, Xie L, Shen Y, Liu S Z, Ju G Z, Shi J P, Yu Y Q, Zhang X, Xu Q and Hemmings G P (2004) The CLDN5 locus may be involved in the vulnerability to schizophrenia. *Eur Psychiatry* 19 (6) 354-357.

Susser E, Neugebauer R, Hoek H W, Brown A S, Lin S, Labovitz D and Gorman J M (1996) Schizophrenia after prenatal famine. Further evidence. *Arch Gen Psychiatry* 53 (1) 25-31.

Sutcliffe J G and Thomas E A (2002) The neurobiology of apolipoproteins in psychiatric disorders. *Mol Neurobiol* 26 (2-3) 369-388.

Szeszko P R, Ardekani B A, Ashtari M, Kumra S, Robinson D G, Sevy S, Gunduz-Bruce H, Malhotra A K, Kane J M, Bilder R M and Lim K O (2005) White matter abnormalities in first-episode schizophrenia or schizoaffective disorder: a diffusion tensor imaging study. *Am J Psychiatry* 162 (3) 602-605.

Tamminga C A and Holcomb H H (2005) Phenotype of schizophrenia: a review and formulation. *Mol Psychiatry* 10 (1) 27-39.

Tan E C, Chong S A, Wang H, Chew-Ping Lim E and Teo Y Y (2005) Gender-specific association of insertion/deletion polymorphisms in the nogo gene and chronic schizophrenia. *Brain Res Mol Brain Res.*

Tang J X, Chen W Y, He G, Zhou J, Gu N F, Feng G Y and He L (2004) Polymorphisms within 5' end of the Neuregulin 1 gene are genetically associated with schizophrenia in the Chinese population. *Mol Psychiatry* 9 (1) 11-12.

Tang J X, Zhou J, Fan J B, Li X W, Shi Y Y, Gu N F, Feng G Y, Xing Y L, Shi J G and He L (2003) Family-based association study of DTNBP1 in 6p22.3 and schizophrenia. *Mol Psychiatry* 8 (8) 717-718.

Terwilliger J D and Hiekkalinna T (2006) An utter refutation of the "Fundamental Theorem of the HapMap". *Eur J Hum Genet* 14 (4) 426-437.

Terwilliger J D and Ott J (1992) A haplotype-based 'haplotype relative risk' approach to detecting allelic associations. *Hum Hered* 42 (6) 337-346.

Thapar A, Langley K, Fowler T, Rice F, Turic D, Whittinger N, Aggleton J, Van den Bree M, Owen M and O'Donovan M (2005) Catechol O-methyltransferase gene variant and birth weight predict early-onset antisocial behavior in children with attention-deficit/hyperactivity disorder. *Arch Gen Psychiatry* 62 (11) 1275-1278.

Thomas E A, Dean B, Pavey G and Sutcliffe J G (2001) Increased CNS levels of apolipoprotein D in schizophrenic and bipolar subjects: implications for the pathophysiology of psychiatric disorders. *Proc Natl Acad Sci U S A* 98 (7) 4066-4071.

Thomson P A, Wray N R, Millar J K, Evans K L, Hellard S L, Condie A, Muir W J, Blackwood D H and Porteous D J (2005) Association between the TRAX/DISC locus and both bipolar disorder and schizophrenia in the Scottish population. *Mol Psychiatry* 10 (7) 657-668, 616.

Tkachev D, Mimmack M L, Ryan M M, Wayland M, Freeman T, Jones P B, Starkey M, Webster M J, Yolken R H and Bahn S (2003) Oligodendrocyte dysfunction in schizophrenia and bipolar disorder. *Lancet* 362 (9386) 798-805.

Todd J A (2006) Statistical false positive or true disease pathway? *Nat Genet* 38 (7) 731-733.

Tomas C, Canellas F, Rodriguez V, Picornell A, Lafau O, Nadal M, Roca M, Serrano M J, Castro J A and Ramon M M (2006) Genetic linkage study for bipolar disorders on chromosomes 17 and 18 in families with a high expression of mental illness from the Balearic Islands. *Psychiatr Genet* 16 (4) 145-151.

Torrey E F, Miller J, Rawlings R and Yolken R H (1997) Seasonality of births in schizophrenia and bipolar disorder: a review of the literature. *Schizophr Res* 28 (1) 1-38.

Uranova N, Orlovskaya D, Vikhreva O, Zimina I, Kolomeets N, Vostrikov V and Rachmanova V (2001) Electron microscopy of oligodendroglia in severe mental illness. *Brain Res Bull* 55 (5) 597-610.

Uranova N A, Vostrikov V M, Orlovskaya D D and Rachmanova V I (2004) Oligodendroglial density in the prefrontal cortex in schizophrenia and mood disorders: a study from the Stanley Neuropathology Consortium. *Schizophr Res* 67 (2-3) 269-275.

van Rossum J M (1966) The significance of dopamine-receptor blockade for the mechanism of action of neuroleptic drugs. *Arch Int Pharmacodyn Ther* 160 (2) 492-494.

Vartanian T, Fischbach G and Miller R (1999) Failure of spinal cord oligodendrocyte development in mice lacking neuregulin. *Proc Natl Acad Sci U S A* 96 (2) 731-735.

Vawter M P, Crook J M, Hyde T M, Kleinman J E, Weinberger D R, Becker K G and Freed W J (2002) Microarray analysis of gene expression in the prefrontal cortex in schizophrenia: a preliminary study. *Schizophr Res* 58 (1) 11-20.

Vella A, Cooper J D, Lowe C E, Walker N, Nutland S, Widmer B, Jones R, Ring S M, McArdle W, Pembrey M E, Strachan D P, Dunger D B, Twells R C, Clayton D G and Todd J A (2005) Localization of a type 1 diabetes locus in the IL2RA/CD25 region by use of tag single-nucleotide polymorphisms. *Am J Hum Genet* 76 (5) 773-779.

Vourc'h P, Dessay S, Mbarek O, Marouillat Vedrine S, Muh J P and Andres C (2003a) The oligodendrocyte-myelin glycoprotein gene is highly expressed during the late stages of myelination in the rat central nervous system. *Brain Res Dev Brain Res* 144 (2) 159-168.

Vourc'h P, Martin I, Marouillat S, Adrien J L, Barthelemy C, Moraine C, Muh J P and Andres C (2003b) Molecular analysis of the oligodendrocyte myelin glycoprotein gene in autistic disorder. *Neurosci Lett* 338 (2) 115-118.

Wahlbeck K, Forsen T, Osmond C, Barker D J and Eriksson J G (2001) Association of schizophrenia with low maternal body mass index, small size at birth, and thinness during childhood. *Arch Gen Psychiatry* 58 (1) 48-52.

Wall J D and Pritchard J K (2003a) Assessing the performance of the haplotype block model of linkage disequilibrium. *Am J Hum Genet* 73 (3) 502-515.

Wall J D and Pritchard J K (2003b) Haplotype blocks and linkage disequilibrium in the human genome. *Nat Rev Genet* 4 (8) 587-597.

Wan C, Yang Y, Feng G, Gu N, Liu H, Zhu S, He L and Wang L (2005) Polymorphisms of myelin-associated glycoprotein gene are associated with schizophrenia in the Chinese Han population. *Neurosci Lett* 388 (3) 126-131.

Wang F, Sun Z, Cui L, Du X, Wang X, Zhang H, Cong Z, Hong N and Zhang D (2004a) Anterior cingulum abnormalities in male patients with schizophrenia determined through diffusion tensor imaging. *Am J Psychiatry* 161 (3) 573-575.

Wang J F and Young L T (1996) Differential display PCR reveals increased expression of 2',3'-cyclic nucleotide 3'-phosphodiesterase by lithium. *FEBS Lett* 386 (2-3) 225-229.

Wang K C, Kim J A, Sivasankaran R, Segal R and He Z (2002a) P75 interacts with the Nogo receptor as a co-receptor for Nogo, MAG and OMgp. *Nature* 420 (6911) 74-78.

Wang K C, Kopriva V, Kim J A, Sivasankaran R, Guo Y, Neve R L and He Z (2002b) Oligodendrocyte-myelin glycoprotein is a Nogo receptor ligand that inhibits neurite outgrowth. *Nature* 417 (6892) 941-944.

Wang W Y, Barratt B J, Clayton D G and Todd J A (2005) Genome-wide association studies: theoretical and practical concerns. *Nat Rev Genet* 6 (2) 109-118.

Wang X, He G, Gu N, Yang J, Tang J, Chen Q, Liu X, Shen Y, Qian X, Lin W, Duan Y, Feng G and He L (2004b) Association of G72/G30 with schizophrenia in the Chinese population. *Biochem Biophys Res Commun* 319 (4) 1281-1286.

Webster M J, O'Grady J, Kleinman J E and Weickert C S (2005) Glial fibrillary acidic protein mRNA levels in the cingulate cortex of individuals with depression, bipolar disorder and schizophrenia. *Neuroscience* 133 (2) 453-461.

Wei Q, Miskimins W K and Miskimins R (2004) Sox10 acts as a tissue-specific transcription factor enhancing activation of the myelin basic protein gene promoter by p27Kip1 and Sp1. *J Neurosci Res* 78 (6) 796-802.

Weinberger D R (1987) Implications of normal brain development for the pathogenesis of schizophrenia. *Arch Gen Psychiatry* 44 (7) 660-669.

Weiss K M and Clark A G (2002) Linkage disequilibrium and the mapping of complex human traits. *Trends Genet* 18 (1) 19-24.

World Health Organisation (1992) *The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines.* Geneva: World Health Organisation.

Williams H J, Glaser B, Williams N M, Norton N, Zammit S, MacGregor S, Kirov G K, Owen M J and O'Donovan M C (2005) No association between schizophrenia and polymorphisms in COMT in two large samples. *Am J Psychiatry* 162 (9) 1736-1738.

Williams N M, Green E K, Macgregor S, Dwyer S, Norton N, Williams H, Raybould R, Grozeva D, Hamshere M, Zammit S, Jones L, Cardno A, Kirov G, Jones I, O'Donovan M C, Owen M J and Craddock N (2006) Variation at the DAOA/G30 locus influences susceptibility to major mood episodes but not psychosis in schizophrenia and bipolar disorder. *Arch Gen Psychiatry* 63 (4) 366-373.

Williams N M, Norton N, Williams H, Ekholm B, Hamshere M L, Lindblom Y, Chowdari K V, Cardno A G, Zammit S, Jones L A, Murphy K C, Sanders R D, McCarthy G, Gray M Y, Jones G, Holmans P, Nimgaonkar V, Adolfson R, Osby U, Terenius L, Sedvall G, O'Donovan M C and Owen M J (2003a) A systematic genomewide linkage study in 353 sib pairs with schizophrenia. *Am J Hum Genet* 73 (6) 1355-1367.

Williams N M, O'Donovan M C and Owen M J (2002) Genome scans and microarrays: converging on genes for schizophrenia? *Genome Biol* 3 (4) REVIEWS1011.

Williams N M, Preece A, Morris D W, Spurlock G, Bray N J, Stephens M, Norton N, Williams H, Clement M, Dwyer S, Curran C, Wilkinson J, Moskvina V, Waddington J L, Gill M, Corvin A P, Zammit S, Kirov G, Owen M J and O'Donovan M C (2004) Identification in 2 independent samples of a novel schizophrenia risk haplotype of the dystrobrevin binding protein gene (DTNBP1). *Arch Gen Psychiatry* 61 (4) 336-344.

Williams N M, Preece A, Spurlock G, Norton N, Williams H J, Zammit S, O'Donovan M C and Owen M J (2003b) Support for genetic variation in neuregulin 1 and susceptibility to schizophrenia. *Mol Psychiatry* 8 (5) 485-487.

Wing J K, Babor T, Brugha T, Burke J, Cooper J E, Giel R, Jablenski A, Regier D and Sartorius N (1990) SCAN. Schedules for Clinical Assessment in Neuropsychiatry. *Arch Gen Psychiatry* 47 (6) 589-593.

Wong A H and Van Tol H H (2003) Schizophrenia: from phenomenology to neurobiology. *Neurosci Biobehav Rev* 27 (3) 269-306.

Wong S T, Henley J R, Kanning K C, Huang K H, Bothwell M and Poo M M (2002) A p75(NTR) and Nogo receptor complex mediates repulsive signaling by myelin-associated glycoprotein. *Nat Neurosci* 5 (12) 1302-1308.

Woolf C J (2003) No Nogo: now where to go? *Neuron* 38 (2) 153-156.

Woolf C J and Bloechlinger S (2002) Neuroscience. It takes more than two to Nogo. *Science* 297 (5584) 1132-1134.

Wright I C, Rabe-Hesketh S, Woodruff P W, David A S, Murray R M and Bullmore E T (2000) Meta-analysis of regional brain volumes in schizophrenia. *Am J Psychiatry* 157 (1) 16-25.

Wright P, Nimgaonkar V L, Donaldson P T and Murray R M (2001) Schizophrenia and HLA: a review. *Schizophr Res* 47 (1) 1-12.

Xi Z R, Qin W, Yang Y F, He G, Gao S H, Ren M S, Peng Y W, Zhang Z and He L (2004) Transmission disequilibrium analysis of the GSN gene in a cohort of family trios with schizophrenia. *Neurosci Lett* 372 (3) 200-203.

Xie X a O, J. (1993) Testing linkage disequilibrium between a disease gene and marker loci. *American Journal of Human Genetics.* 53 1107.

Xiong L, Rouleau G A, Delisi L E, St-Onge J, Najafee R, Riviere J B, Benkelfat C, Tabbane K, Fathalli F, Danics Z, Labelle A, Lal S and Joober R (2005) CAA insertion polymorphism in the 3'UTR of Nogo gene on 2p14 is not associated with schizophrenia. *Brain Res Mol Brain Res* 133 (1) 153-156.

Yan H, Yuan W, Velculescu V E, Vogelstein B and Kinzler K W (2002) Allelic variation in human gene expression. *Science* 297 (5584) 1143.

Yang J Z, Si T M, Ruan Y, Ling Y S, Han Y H, Wang X L, Zhou M, Zhang H Y, Kong Q M, Liu C, Zhang D R, Yu Y Q, Liu S Z, Ju G Z, Shu L, Ma D L and Zhang D (2003) Association study of neuregulin 1 gene with schizophrenia. *Mol Psychiatry* 8 (7) 706-709.

Yang Y F, Qin W, Shugart Y Y, He G, Liu X M, Zhou J, Zhao X Z, Chen Q, La Y J, Xu Y F, Li X W, Gu N F, Feng G Y, Song H, Wang P and He L (2005) Possible association of the MAG locus with schizophrenia in a Chinese Han cohort of family trios. *Schizophr Res* 75 (1) 11-19.

Yin X, Peterson J, Gravel M, Braun P E and Trapp B D (1997) CNP overexpression induces aberrant oligodendrocyte membranes and inhibits MBP accumulation and myelin compaction. *J Neurosci Res* 50 (2) 238-247.

Zai G, Bezchlibnyk Y B, Richter M A, Arnold P, Burroughs E, Barr C L and Kennedy J L (2004) Myelin oligodendrocyte glycoprotein (MOG) gene is associated with obsessive-compulsive disorder. *Am J Med Genet B Neuropsychiatr Genet* 129 (1) 64-68.

Zai G, King N, Wigg K, Couto J, Wong G W, Honer W G, Barr C L and Kennedy J L (2005) Genetic study of the myelin oligodendrocyte glycoprotein (MOG) gene in schizophrenia. *Genes Brain Behav* 4 (1) 2-9.

Zhang Y and Feng Y (2001) Distinct molecular mechanisms lead to diminished myelin basic protein and 2',3'-cyclic nucleotide 3'-phosphodiesterase in qk(v) dysmyelination. *J Neurochem* 77 (1) 165-172.

Zhao J H, Curtis D and Sham P C (2000) Model-free analysis and permutation tests for allelic associations. *Hum Hered* 50 (2) 133-139.

Zhao X, Shi Y, Tang J, Tang R, Yu L, Gu N, Feng G, Zhu S, Liu H, Xing Y, Zhao S, Sang H, Guan Y, St Clair D and He L (2004) A case control and family based association study of the neuregulin1 gene and schizophrenia. *J Med Genet* 41 (1) 31-34.

Zheng B, Ho C, Li S, Keirstead H, Steward O and Tessier-Lavigne M (2003) Lack of enhanced spinal regeneration in Nogo-deficient mice. *Neuron* 38 (2) 213-224.

Zou F, Li C, Duan S, Zheng Y, Gu N, Feng G, Xing Y, Shi J and He L (2005) A family-based study of the association between the G72/G30 genes and schizophrenia in the Chinese population. *Schizophr Res* 73 (2-3) 257-261.