Molecular Genetics of Bipolar Disorder

Thesis submitted for the degree of

Doctor of Philosophy

UCL

Radhika Kandaswamy

Molecular Psychiatry Laboratory

University College London

2007 – 2012

I, **Radhika Kandaswamy**, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Bipolar affective disorder has a strong genetic heritability. In the UCL laboratory, a locus on chromosome 1p36 was found to be linked to bipolar and related unipolar affective disorders with a log of the odds score above 3.00. This region was subjected to fine mapping using tests of allelic association in a case-control sample as part of this thesis in order to identify the genes involved in bipolar disorder. In addition, a GWAS was also employed to fine map other bipolar affective disorder susceptibility genes.

The tests of allelic association found evidence for the involvement of the PRKCZ gene. Markers D1S243 and rs3128396 at the PRKCZ gene were significantly associated with bipolar disorder with empirical P = 0.037 and P = 0.040, respectively. Other loci encoding brain expressed proteins found to be associated in the UCL GWAS sample were the genes - GRM3 and GRM7. Therefore, these genes were sequenced using PCR-based genomic sequencing. A 3'-UTR base pair change (rs56173829) in the GRM7 gene was found to be significantly associated with the disorder, although the minor allele was more frequent in controls.

A base pair mutation (rs148754219) was found in the GRM3 exon 1 two bases before the translation start codon (forming part of Kozak sequence) of a GRM3 receptor isoform. The mutation was associated with bipolar disorder (P = 0.0046, odds ratio 4.2 (95% CI 1.42-12.37). Transcription factor binding assays and cloning experiments comparing the gene expression activity of wild-type and mutant alleles showed that the mutant allele strongly affected the reporter gene activity in SH-SY5Y and HEK293 cells. If the GRM3 Kozak sequence mutation is confirmed as an

important mutation in the aetiology of bipolar disorder, then it could be used as a marker for personalised treatment for a genetic subtype of affective disorders.

Table of Contents

Acknowledgements

As it all draws to a close, I realise the enormity of direction and support I have been fortunate to receive throughout my research. In my debt of gratitude, the first and foremost of my thanks goes to Prof. Hugh Gurling for giving me this wonderful and exciting opportunity to carry out this PhD, which to be honest proved to be a bit tough at times. I would also like to thank Dr. Andrew McQuillin for all the support he has extended me throughout these years and for also answering all my questions, some of them are without doubt the silliest. I would also like to extend my sincerest thanks to Prof. Martin Orrell, without whose initiation I would have never landed up in the field of psychiatric research.

The next in line are my parents, whose unconditional love and affection are proof that there are indeed constants in this universe. I can never thank them enough for their support, without which my coming back from maternity period during the PhD could not have been the easy ride it turned out to be.

A very special thanks is in order to Sally for always trusting me and guiding me whenever I was lost in the mazes of science. She has been the one person cheering team who helped me cross the line. I would like to thank all my colleagues and ex-colleagues from my lab Yi, Alessia, John, Vinay, Ana, Sushmita, Mie and Khalid for their support.

Thanks are due to Ramya and Ashok, my sister and in-law for patiently listening to all my whining over the phone when my experiments failed without even the tiniest bit of exasperation and even in case they felt, for never having shown it. Nothing ever happens without family and friends. It is such a pleasure to have them in life and more importantly during a thesis preparation. Thank you all and you know who you are. A special mention to Krishna and Ganesh for helping with the enormous task of proofreading this thesis.

Last but not the least, I am extremely grateful to my husband Saravana Kumar for the never-ending emotional and financial support, for his cool and calm attitude, for

never doubting my abilities, and happily looking after our precious little son Karthik during the thesis writing period.

I would finally like to thank everyone once more for helping me make this extraordinary journey all the more memorable and highly rewarding.

List of Figures

List of Tables

Publications related to this thesis

Abstracts

Kandaswamy R, Bass N, Choudhury K, Puri V, McQuillin A, Lawrence J, Curtis D, Gurling H (2006) Fine mapping of a new bipolar and unipolar affective disorder locus on chromosome 1p36. Am J Med Genet B Neuropsychiatr Genet 141B(7): 747 P42, Proceedings of the 14th World Congress on Psychiatric Genetics.

Kandaswamy R, McQuillin A, Pereira ACP, Bass N, Lawrence J, Sklar P, Scolnick E, Purcell S, Curtis D, Gurling H (2008) Association and resequencing of the metabotropic glutamate receptor GRM7 in bipolar disorder. Proceedings of the 16th World Congress on Psychiatric Genetics.

Original articles

Kandaswamy R, McQuillin A, Curtis D, Gurling H (2012) Tests of linkage and allelic association between markers in the 1p36 PRKCZ (protein kinase C zeta) gene region and bipolar affective disorder. Am J Med Genet B Neuropsychiatr Genet 159B(2): 201-9.

Kandaswamy R, McQuillin A, Sharp S, Fiorentino A, Anjorin A, Blizard R, Curtis D, Gurling H (2012). Genetic association, mutation screening and functional analysis of a Kozak sequence variant in the GRM3 glutamate receptor gene in bipolar disorder. Archives of General Psychiatry. *Article accepted for publication*.

Abbreviations

(*RS*)-PPG	((RS)-4-phosphonophenylglycine)
ACPT-I	[(1S,3R,4S)-1-aminocyclo-pentane-1,3,4-tricarboxylic acid]
ACTH	Adrenocorticotropic hormone
AMN082	(N,N-Bis(diphenylmethyl)-1,2-ethanediamine)
AMPA	α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic
ANK3	Ankyrin 3
ASP	Affected sib-pairs
BDNF	Brain derived neurotrophic factor
BiGS	Bipolar Genome Study
BP I	Bipolar I disorder
BP II	Bipolar II disorder
BPAD	Bipolar affective disorder
BP-NOS	Bipolar disorder not otherwise specified
BSA	Bovine serum albumin
CACNA1C	Calcium channel, voltage-dependent, L type, alpha 1C subunit
CACNG5	Calcium channel, voltage-dependent, gamma subunit 5
CANMAT	Canadian Network for Mood and Anxiety Treatments
CDCV	Common Disease Common Variant
CDRV	Common Disease Rare Variant
CFTR	Cystic fibrosis transmembrane conductance regulator

CLOCK	Circadian Locomotor Output Cycles Kaput
CMLA	Comparative linkage meta-analysis
CNR	Copy Number Reference
CNS	Central Nervous System
CNV	Copy number variation
COMT	Catechol-O-methyltransferase
CPPG	[(RS)-alpha-cyclopropyl-4-phosphonophenyl glycine]
CRD	Cysteine-rich domain
CSF	Cerebrospinal fluid
DAOA	D-amino Acid oxidase Activator
DCTN5	Dynactin 5
DGKH	Diacyl glycerol kinase
DISC1	Disrupted in schizophrenia gene
DLPFC	Dorsolateral prefrontal cortex
DLR	Dual luciferase reporter
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	Dimethyl sulfoxide
dNTP	Deoxynucleoside triphosphate
DRD4	Dopamine Receptor D4
DSM	Diagnostic and Statistical Manual of Mental Disorders
DTT	Dithiothreitol

DZ	Dizygotic
ECA	Epidemiological Catchment Study
ECACC	European Collection of Animal Cell Cultures
ECT	Electroconvulsive therapy
EDTA	Ethylene diaminetetraacetic acid
EGFR	Epidermal growth factor receptor
EMSA	Electromobility shift assay
FBS	Foetal bovine serum
fMRI	Functional magnetic resonance imaging
FRET	Förster/fluorescence resonance energy transfer
GABA	γ-aminobutyric acid
GluR	Glutamate receptor
GPCR	G-protein-coupled receptor
GRIN2B	Glutamate Receptor Ionottropic N-Methyl D-aspartate 2B
GSK3β	Glycogen Synthase Kinase 3 Beta
GSMA	Genome scan meta-analysis
GWAS	Genome-wide association studies
HLA-DRA	Major histocompatibility complex, class II, DR alpha
HPA	Hypothalamic-pituitary-adrenal axis
HRM	High Resolution melting
HTR2A	Human Serotonin 2A Receptor gene

HTT	Huntington disease
HWE	Hardy-Weinberg equilibrium
IBD	Identity by descent
IBS	Identity by state
ICD	International Classification of Diseases
iGluR	Ionotropic glutamate receptor
IPTG	Isopropyl B-D-1-thiogalactopyranoside
IQR	Interquartile range
IRES	Internal ribosomal entry site
ISBD	International Society for Bipolar Disorders
KAR	Kainate receptor
LAR II	Luciferase Assay Reagent II
LD	Linkage Disequilibrium
LK	Lander and Kruglyak
LMAN2L	Lectin Mannose-binding 2-like
LNA	Locked Nucleic Acids
LOD	Logarithm of the odds
LTD	Long-term depression
LTP	Long-term potentiation
MAOI	Monoamine Oxidase Inhibitors
MBP	Myelin basic protein

MDD	Major depressive disorder
MFE	Minimum of free energy
mGluR	Metabotropic glutamate receptor
mRNA	messenger RNA
miRNA	microRNA
MMPIP	(6-(4-Methoxyphenyl)-5-methyl-3-(4-pyridinyl)-isoxazolo[4,5-c]pyridin-4(5*H*)-one hydrochloride)
MOAO	Monoamine Oxidase -A
MOD-score	LOD score maximized over genetic model parameters
MPEP	2-methyl-6-[phenylethynyl]-pyridine
MREC	Multicentre Research Ethics Committee
MSP	Multiple Scan Probability
MTEP	[(2-methyl-1,3-thiazol-4-yl)ethynyl]pyridine
MY05B	Myosin V B
MZ	Monozygotic
NAAG	N-acetylaspartateglutamate
NARI	Noradrenaline Reuptake Inhibitors
NCAM1	Neural Cell Adhesion Molecule 1
NCAN	Neurocan
NCS	National Comorbidity Study
NESARC	National Epidemiologic Survey on Alcohol and Related Conditions
NICE	National Institute for Health and Clinical Excellence

NIMH	National Institute of Mental Health
NMDA	N-methyl-D-aspartate
NMDAR	NMDA receptor
NRG1	Neuregulin 1
NTRK2	Neurotrophic tyrosine kinase, receptor, type 2
OFC	Olanzapine-fluoxetine combination
P2RX7	Purinergic receptor P2X, ligand-gated ion channel, 7
PAGE	Polyacrylamide Gel Electrophoresis
PBS	Phosphate buffered saline
PCP	Phencyclidine
PCR	Polymerase Chain Reaction
PD	Parkinson's Disease
PKC	Protein kinase C
PLB	Passive Lysis buffer
PTGFR	Prostaglandin F Receptor
PTSD	Post traumatic stress disorder
PVPP	Polyvinylpolypyrrolidine
QRT-PCR	Quantitative Real-Time PCR
RDC	Research Diagnostic Criteria
ROX	5-carboxy-X-rhodamine, succinimidyl ester
SADSL	Schizophrenia and Affective Disorders Schedule

SDS	Sodium dodecyl Sulfate
SERT	Serotonin Reuptake Transporter
SGLA	STRUCTURE-Guided Linkage Analysis
SNP	Single Nucleotide Polymorphism
SNRI	Serotonin-Noradrenaline Reuptake Inhibitors
SSRI	Selective Serotonin Reuptake Inhibitors
STR	Short tandem repeat
SYNE1	Spectrin repeat containing, nuclear envelope 1
TEMED	Tetramethylethylenediamine
TM	Transmembrane
TPH	Tryptophan hydroxylase
TRANK1 (LBA1)	Tetratricopeptide repeat and ankyrin repeat containing 1
TRPM2	Transient Receptor Potential Melastatin 2
TSPAN8	Tetraspanin 8
UPL	Universal Probe Library
UTR	Untranslated region
VFT	Venus Fly Trap
WFSBP	World Federation of Societies of Biological Psychiatry
WHO	World Health Organisation
WKY	Wistar Kyoto
WTCCC	Wellcome Trust Case Control Consortium

Chapter 1

General Introduction

1.1. Bipolar disorder

1.1.1. Overview

Bipolar disorder is a mood disorder characterized by impairing episodes of mania and depression. Although bipolar disorder has been described as a 'cyclic' illness, the manic episode is the defining feature of the disorder. A patient is diagnosed with bipolar disorder on the basis of having had a manic episode but they may not necessarily have had depressive episodes. Also, patients return to their normal state of well-being in the intervals between these episodes. Manic episodes are characterized by flight of ideas, pressure of speech, increased energy, decreased need for sleep, hyperactivity, grandiosity, increased sexual activity and impulsive behaviour. Predominantly, the mood of manic episodes is elevated or irritable. Psychotic symptoms such as delusions and hallucinations may or may not be present in a manic episode. Depressive episodes are characterised by depressed mood, loss of interest, decreased sexual interest, accompanied by reduced self-confidence, decreased energy and feelings of guilt and worthlessness. Severe forms of both mania and depression cause marked impairment in social and occupational functioning (WHO 1993; APA 2000). The morbidity and mortality associated with bipolar disorder is very high, with one study suggesting 20-fold greater risk for suicide compared to the general population (Perlis et al. 2010).

The personal and societal costs of bipolar disorder are enormous with relapse being common in adequately-treated individuals, and the disorder can be profoundly disabling resulting in serious economic burden (Kessler et al. 2005). With some variability by diagnoses, numerous epidemiological studies on mental disorders throughout the world have shown that in each year about one-third of the adult

population suffers from a mental disorder (Wittchen et al. 2011). The Global Burden of Disease studies have found that disorders of the brain are increasingly contributing to the disease burden. A recent analysis estimated that 13% of the global disease is due to the disorders of brain, surpassing both cardiovascular diseases and cancer (Collins et al. 2011; Wittchen et al. 2011). Every year over 38.2% of the European Union population (Wittchen et al. 2011) suffer from mental disorders costing the economy €798 billion (Olesen et al. 2012). Recent research has highlighted ongoing cognitive difficulties and multiple domains of dysfunction in patients suffering from bipolar disorder (Burdick et al. 2007; Gutierrez-Rojas et al. 2011). Furthermore, the rates of morbidity, mortality and suicide are higher in those with bipolar I disorder (BP I), a subtype of bipolar disorder (Yutzy et al. 2012). A recent World Mental Health Survey in seventeen countries conducted by World Health Organisation (WHO) (Kawakami et al. 2012) reported that early-onset mental disorders are associated with a significantly reduced household income in high and upper-middle income countries but not in low/lower-middle income countries, with associations consistently stronger among women than men.

1.1.2. History of Bipolar disorder

The ancient descriptions of the morbid states of depression and exaltation by the physicians and the philosophers in the pre-Hippocratic era make mania and melancholia two of the earliest described human diseases. Hippocrates (460 - 377 BC) was the first to systematically formulate the classification of mental disorders into melancholia, mania and paranoia (Angst and Marneros 2001). Later, in the first or the second Century AD, Areteus of Cappadocia conceived mania and melancholia as two different facets of one single disease, thereby, explicitly linking the two states

for the first time (Yutzy et al. 2012). During the seventeenth and eighteenth centuries, psychiatrists continued to acknowledge the association between mania and melancholia but it was the nineteenth century that witnessed the birth of the modern concept of bipolar disorder. In 1845, a German psychiatrist, Wilhelm Griesinger, described the relation of mania and melancholy using the image of a circle although it was Falret, in 1854, who finally drew the conclusion that bipolar disorder was an entity on its own based on his long-term observations of patients. He defined the sequential change from mania to melancholia and the interval in between as an independent disease on its own, namely the 'folie circulaire' (Angst and Marneros 2001). The foundation of psychiatric nosology and scientific psychiatry has its roots in the publications of Emil Kraeplin in 1899 that characterised psychoses into dementia praecox and manic-depressive insanity (Zivanovic and Nedic 2012). In 1957, Leonhard coined the term 'bipolar' for the patients with depression who also experienced mania. Hypomania is another term used quite often to discriminate between BP I and bipolar II (BP II) disorder based on DSM-IV classification (APA 1980). Accordingly, people with hypomania (BP II) have fewer and less severe symptoms that usually do not require hospitalization than those with mania (BP I). Erich Mendel described, conceptualized and coined the term 'hypomania' in 1881, although Dunner (1976) is credited for the creation of BP II category (Angst and Marneros 2001).

A gradual refinement of the terminology has occurred during the development of the concept of bipolar disorder moving from a traditional and broader concept of mania and melancholia to a more modern concept of circular disease. In recent years, the concept has broadened to include schizoaffective bipolar disorder type, mixed states,

bipolar spectrum etc. to encompass the complexity of degree of severity and a combination of symptoms.

1.1.3. Epidemiology of Bipolar disorder

Epidemiology is the study of the distribution of diseases in human populations and the variation of these distributions among different population subgroups. Prevalence is an epidemiological measure of the proportion of individuals that have a disorder at a specified time or during a specified period. The early family studies conducted from 1929 to 1954 did not distinguish between major depression and bipolar disorder due to the use of the term 'manic depression' encompassing both the disorders. The prevalence estimate of manic depression from these early studies was 14.6% among parents of probands and 10.9% among siblings of probands compared to the general population estimate of 0.7% (Smoller and Finn 2003).

Lifetime prevalence is the proportion of those in the population who had a disorder at some time in their life up to their age at the time of interview. A lifetime prevalence of 0.8% and 1.6% for manic episodes was reported in the Epidemiological Catchment Study (ECA) and the National Comorbidity Study (NCS) respectively (Regier et al. 1990; Kessler et al. 1994). A cross-national epidemiology study involving BP I patients from 10 countries, including United States, Canada, Puerto Rico, France, West Germany, Italy, Lebanon, Taiwan, Korea and New Zealand, reported lifetime prevalence rates ranging from 0.3% to 1.5% (Weissman et al.1996). European studies have reported lifetime prevalence rates for bipolar disorder ranging from 0.2% to 1.1% (Wittchen et al. 2011). A few studies report the wider range of bipolar spectrum disorders as having prevalence rates of 5.1% and 6% (Pini et al. 2005). Recent US and international studies have reported

the lifetime prevalence of bipolar I disorder ranging from 0.4% and 1.3% (Yutzy et al. 2012). Another review of the literature on the robust and replicable genetic findings reported in the last decade through genome-wide association studies (GWAS) and structural variation studies predicted the lifetime prevalence of bipolar disorder as 0.7% and the heritability as being 75% (Sullivan et al. 2012).

Even though BP II has neither been a subject of large-scale studies internationally nor is listed in the International Classification of Diseases (ICD)-9 or ICD-10 classification (WHO 1967; WHO 1993), a lifetime prevalence of 0.5% was reported for this disorder in the ECA study (Regier et al. 1990) and the Diagnostic and Statistical Manual of Mental Disorders (DSM) series (APA 1994; APA 2000). In a recent review, Merikangas and Lamers reported results from studies in adults across 11 countries as part of the World Mental Health Survey showing that lifetime prevalence rate of BP II was 0.4% and is also substantially greater in adolescents with rates approaching 3-4% (Merikangas and Lamers 2012).

Previous studies have also reported raised incidence of bipolar affective disorder (BPAD) and mania in African American, Asian Indian and many other immigrant groups (Leff et al. 1976; Lloyd et al. 2005).

Age at onset

Early-onset bipolar disorder has been associated with severe clinical symptoms, a stronger family history of affective disorder and poor clinical outcome. Previous studies on bipolar disorder have estimated the average age of onset between early-to-mid 20s to early 30s (Kupfer et al. 2002). Baldessarini et al. (2010) reported the estimated age of onset in a pooled sample of 1566 DSM type I or II bipolar disorder from six international sites. The mean age of onset (±IQR) across all the sites and

bipolar disorder subtypes was 25.2 ± 15.1 years. However, the median onset was nearly six years younger among BP I patients than BP II patients. Additionally, there was little (<3 years) difference in age of onset between men and women with men tending to be somewhat younger at onset among BP I cases with psychotic or mixed initial episodes (Baldessarini et al. 2010). Juvenile-onset (≤ 20 years) and childhood-onset (<13 years) bipolar disorder were more common in European samples than US samples (juvenile: 26.7% vs 15.9%; child: 3.33% vs 0%) (Baldessarini et al. 2010).

A significant association between the age of onset of bipolar disorder in probands and an increased risk of bipolar disorder in family members has been reported in several studies (Althoff et al. 2005). Comorbid depression (Post et al. 2010) and mania/hypomania (Bauer et al. 2010) have also been observed to occur in excess with early-onset bipolar disorder. Recently, an overall median age of onset of 24.0 years was reported in a study of 1665 BP I disorder patients from seven international centres along with an early age of onset being associated with the least favourable clinical outcomes and a greater familial risk (Baldessarini et al. 2012).

Gender

Existence of epidemiological and clinical differences between genders amongst patients with bipolar disorder has long been debated. Although compelling evidence from previous studies suggests that there is no major gender distinction in the lifetime prevalence of bipolar disorder, there are few studies that report higher prevalence rates in men (Diflorio and Jones 2010). According to the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC), men were twice as likely as women to have a history of unipolar mania amongst individuals

with BP I (Grant et al. 2005). However, a higher prevalence of BP II and hypomania in women has been reported in some but not all studies (Diflorio and Jones 2010).

Another controversial topic is the sex distribution of bipolar patients with rapid cycling. Two meta-analyses conducted between the years 1977 and 2002 reported that women with bipolar disorder showed greater rates of rapid cycling than men (Tondo and Baldessarini 1998; Kupka et al. 2003) whereas a recent prospective study reported an almost equal one year prevalence of rapid cycling in both the sexes (Schneck et al. 2008).

According to the earlier studies of bipolar disorder in the US (Viguera et al. 2001) and Europe (Raymont et al. 2003), women reportedly have later age at onset compared to men. In a study based in Camberwell, England (Kennedy et al. 2005), a strong association between gender and age at onset of first-episode mania and bipolar disorder was observed with men having a mean age at onset of 5.1 years and 4.4 years respectively earlier than women. In a childhood-onset bipolar sample, male gender was overrepresented in the group with an onset before 12 years (Biederman et al. 2004). In the NESARC study, men were reported to have a younger age of onset of manic and mixed but not depressive episodes in the BP I sample (Grant et al. 2005). Some studies have also indicated that women show a predominance of depressive polarity (Nivoli et al. 2010). Pregnancy and childbirth are considered to be predisposing factors to bipolar disorder. Women are suggested to have more than two-fold risk for bipolar disorder after childbirth (Altman et al. 2006). There is also consistent evidence of a gender effect for substance use disorders in bipolar patients with alcohol being more of a problem in men and eating disorders and Post

Traumatic Stress Disorder (PTSD) being more common in women (Diflorio and Jones 2010).

1.1.4. Diagnosis of Bipolar disorder

The two main modern systems of classification for the diagnosis of mental disorders are the DSM series published by the American Psychiatric Association (APA) and the ICD manual published by the WHO. The introduction of DSM-III in 1980 (APA 1980) initiated the conceptualization of mania and the associated disorders along with the development of a valid and reliable criteria. In 1972, Feighner and colleagues (Feighner et al. 1972) at the Washington University identified a set of criteria for mania that later formed the basis of the Research Diagnostic Criteria (RDC) (Spitzer et al. 1975; Spitzer et al. 1978) and DSM-III. No substantive changes have been added to the subsequent iterations of the DSM series (DSM-III-R, DSM-IV, DSM-IV-R) (Singerman 1981; APA 1987; APA 1994; APA 2000). DSM-IV classifies bipolar disorder into BP I and BP II whereas there is no such distinction in the ICD-10 classification. According to DSM-IV, at least one manic episode is required for the diagnosis of BP I disorder whereas at least one hypomanic episode and at least one major depressive episode are required for BP II diagnosis. The duration of a manic episode needs to be at least one week for a bipolar disorder diagnosis, while depressive episodes should last at least two weeks. Hypomania is a less severe form of mania where the patients are generally euphoric and very cheerful with an ‘infectious’ mood (APA 1980). For an ICD-10 diagnosis of bipolar disorder, a manic episode along with another mood disturbance should be experienced by the individual (WHO 1993). Patients with bipolar disorder may present with either a depressive or a manic episode but due to mania being the

important criteria for bipolar diagnosis, 40% of the patients are diagnosed initially with unipolar depression instead. Each episode may last from one week to several months varying across different patients. Also, a bipolar patient may suffer from many depressive episodes before having the first episode of mania (Akiskal 2005). While, the frequency of depressive episodes is usually higher than manic episodes, the course of illness is variable among individuals. A seasonal pattern has also been reported by some researchers with manic episodes more likely in the spring or early summer and depressive ones in the autumn or winter (APA 2000).

Other variants of bipolar disorder such as cyclothymia, mixed episodes, rapid cycling, schizoaffective disorder and seasonal affective disorder are specified as Bipolar disorder not otherwise specified (BP-NOS). Cyclothymia is a chronic condition in which the patient experiences short periods of mild depression and hypomania lasting anywhere between a few days to a few weeks again separated by short periods of normal mood. Mixed episodes are conditions when a patient experiences both manic and depressive episodes nearly every day over at least one week. Additionally, psychotic features may be present during mixed episodes (APA 1980). Schizoaffective disorder is a controversial subtype, which shows characteristics of both schizophrenia and bipolar disorder. In other words, it is a schizophrenia-bipolar boundary disorder. Rapid cycling is another condition in which the patient has at least four distinct episodes of mood disturbances in a single year separated by periods of remission. These individuals are usually resistant to many available treatments such as lithium therapy, and are thereby, clinically challenging to treat.

1.1.5. Treatment of Bipolar disorder

The treatment of bipolar disorder is considered on the basis of the course of the illness, i.e., specific for manic/hypomanic, mixed or depressive episodes (Fountoulakis and Vieta 2008). Mood stabilizers are the most important pharmacological treatments that ameliorate manic symptoms and also sometimes those of depression (lithium and lamotrigine) (Fountoulakis and Vieta 2008; Malhi et al. 2009a; Malhi et al. 2009b). Medications developed for other indications, such as anticonvulsants for epilepsy and antipsychotics for schizophrenia, have also been approved for use in bipolar disorder. Although lithium was first used as a successful treatment for bipolar disorder in 1947 (Cade 1949), it gained the US Food and Drug Administration approval only in 1970. Lithium has many effects in the brain. According to one interesting theory among several, lithium has been proposed to have a robust neuroprotective and neurotropic action leading to the upregulation of synaptic plasticity. A recent study found that a 4-week administration of lithium magnified the long-term potentiation of CA1 pyramidal cells thus upregulating the synaptic plasticity in the hippocampus (Shim et al. 2012). Reports also suggest that lithium decreases suicide risk (Muller-Oerlinghausen and Lewitzka 2010). Unfortunately, lithium has a narrow therapeutic index and lithium toxicity can be fatal in some occasions. Some other common mood stabilizers used in clinical practice are valproate, carbamazepine, gabapentine and lamotrigine.

It is well established that bipolar depression causes greater psychosocial impairment and disability in bipolar patients, who spend much more time being depressed than being manic (Fountoulakis et al. 2012). Almost every guideline on the treatment of bipolar depression reviewed by Fountoulakis et al. (2005) in 2005 suggested the use

of antidepressants only in combination with an antimanic agent to avoid manic switches. A number of antidepressants are in current use, including tricyclic antidepressants, Selective Serotonin Reuptake Inhibitors (SSRI), Monoamine Oxidase Inhibitors (MAOI), Noradrenaline Reuptake Inhibitors (NARI), Serotonin-Noradrenaline Reuptake Inhibitors (SNRI) and atypical antidepressants (Palucha and Pilc 2007). Among atypical antipsychotics, quetiapine, olanzapine and fluoxetine are routinely used in monotherapy or in combination for the management of bipolar depression. In a recent review, the CANMAT (Canadian Network for Mood and Anxiety Treatments), ISBD (International Society for Bipolar Disorders) and the WFSBP (World Federation of Societies of Biological Psychiatry) guidelines generally suggest lithium, lamotrigine or quetiapine monotherapy as the first-line choice. These guidelines also recommend the use of antidepressant agents in combination with antimanic agents such as lithium and valproate, as well as the use of Olanzapine-Fluoxetine Combination (OFC). The NICE (National Institute for Health and Clinical Excellence) guidelines also suggest a similar treatment strategy and in the presence of psychotic symptoms recommend adding an antipsychotic drug such as olanzapine, quetiapine or risperidone to the combination of antidepressant and antimanic agents (Nivoli et al. 2011).

Maintenance therapy in bipolar disorder is prophylactic, primarily to prevent relapses and improve residual symptoms. Atypical antipsychotics have been proven to be effective in maintenance treatment in addition to being effective against mania and bipolar depression. A placebo-controlled trial suggested that lamotrigine stabilized mood, which is useful in depressive-relapse prevention (Calabrese et al. 2003). Mood stabilisers namely, valproate, carbamazepine and lithium are widely

used due to their prophylactic effect. Electroconvulsive therapy (ECT) is an alternative treatment used for severely affected patients who fail to respond to pharmacotherapy. However, ECT is not routinely used lately.

In summary, treatment of bipolar disorder was very much improved when lithium was first introduced over 60 years ago. Lithium is only effective in a minority of cases. Mood stabilizers, antidepressants in combination with anti-manic agents, and antipsychotics are all currently used in the treatment of bipolar disorder. Monotherapy with antidepressants is contraindicated in mixed states, manic episodes and BP I. There is an urgent need for the development of novel drugs targeting bipolar disorder with greater efficacy, faster onset of action and with less burdensome adverse-effect profiles.

1.1.6. Pathophysiology of Bipolar disorder

Little is known about the specific pathophysiology of bipolar disorder. This limited understanding has prevented development of new treatments. Monoaminergic systems including serotonergic, noradrenergic and dopaminergic systems in the brain have been the most investigated in neurobiological studies of mood disorders making them valuable targets for the therapeutics.

Brain imaging studies (Savitz and Drevets 2012), transgenic mouse models (Lesch and Mossner 2006), genetic studies (Serretti and Mandelli 2008; Craddock and Sklar 2009) and gene expression studies (Elashoff et al. 2007) have all suggested that the monoaminergic and serotonergic systems are implicated in the pathophysiology of bipolar disorder. Studies investigating the effects of dopamine and serotonin antagonists on their respective neurotransmitter systems have also led to the

proposition that these systems are involved in the aetiology of bipolar disorder. Antipsychotics block dopamine D2 receptors and it is thought that some of the psychotic experiences in bipolar disorder are mediated through dopamine pathways (Brugue and Vieta 2007). Altered glutamate levels in the plasma, serum and cerebrospinal fluid (CSF) of individuals with bipolar disorder have also been well documented (Sanacora et al. 2008). Similarly, postmortem studies have also shown increased glutamate levels in diverse brain areas in individuals with mood disorders (Hashimoto et al. 2007). A recent study demonstrated altered expression of glutamate and dopamine transporters in postmortem frontal cortex from bipolar disorder patients (Rao et al. 2012). Elevated brain glutamate/glutamine ratio and reduced levels of N-methyl-D-aspartate (NMDA) receptor subunits have been reported in postmortem brains from bipolar disorder patients (Hashimoto et al. 2007; Rao et al. 2012). Inconsistent results have been documented regarding the levels of the Serotonin Reuptake Transporter (SERT) in the prefrontal cortex of bipolar disorder patients (Dean et al. 2001; Sun et al. 2001). Other neurotransmitter systems such as the γ-aminobutyric acid (GABA)-ergic (Fatemi et al. 2011; Luscher et al. 2011) and cholinergic systems (Machado-Vieira and Zarate 2011) may also be involved in the pathophysiology of bipolar disorder.

Atrophic changes in the brains of bipolar disorder patients have been reported since 1983 when Reider et al. (1983) first found a correlation between age and ventricle-to-brain ratios in a group of bipolar disorder patients. A recent meta-analysis found an inverse correlation between illness duration and total cerebral volume (Hallahan et al. 2011). The same study found increased hippocampal and amygdala volume in lithium treated bipolar disorder patients compared to lithium untreated patients and

normal subjects. Structural and functional imaging (fMRI) evidence and behavioural and cognitive deficits associated with bipolar disorder have implicated specific regions of the brain in the pathophysiology of the disorder. The ventral prefrontal cortex and the anterior cingulate cortex are the two regions with the most significant evidence for the involvement in bipolar disorder, although other regions have also been reported. Decreases in prefrontal and cingulate volume have been detected in studies involving youths with bipolar disorder (Schneider et al. 2012; Selvaraj et al. 2012). The rostral and lateral orbiotofrontal cortex subregions have reduced activation during manic episodes with a change in activation during depressed episodes (Jackowski et al. 2012). In studies involving twins discordant for bipolar disorder, volumetric abnormalities in the white matter of unaffected siblings was reported that included decreased left hemispheric white matter volume (Noga et al. 2001; van der Schot et al. 2009). These findings suggest that reduced white matter connectivity may represent an endophenotype for bipolar illness. Functional abnormalities displayed during performance of cognitive and emotional tasks across multiple cognitive domains, including emotional processing, working memory, inhibition control and reward processing have been reported in several studies (Schneider et al. 2012).

Dysregulation of the hypothalamic-pituitary-adrenal (HPA) axis, the primary hormonal stress response pathway, has been reported in individuals with bipolar disorder (Sinclair et al. 2012a, b). Additionally, disruptions in biological rhythms including sleep disturbances, social rhythms and life events have long been associated with bipolar disorder. Finally, neuroplastic changes at the functional, structural and cellular levels contribute to the dysregulation of key circuits in bipolar

disorder. These changes include impairments in neurotrophic and neuroprotective signaling cascades, altered glutamatergic, serotonergic, dopaminergic, cholinergic and glucocorticoid signaling along with changes in the rate of adult neurogenesis.

1.1.7. Heritability of Bipolar disorder

A strong genetic basis for psychosis has been apparent for more than five decades, with the mention of the hereditary factor in Kraepelin's observations involving patients and their families. Substantial evidence from a range of genetic methods has confirmed the high heritability of bipolar disorder with genetic factors explaining 60%-85% variance in risk (Smoller and Finn 2003). Although both genetic and environmental factors contribute to the aetiology of the disease, the family, twin and adoption studies have shown that there is increased risk of bipolar disorder among relatives of bipolar probands compared to normal controls. The observed familial risk and the inheritance pattern of psychiatric disorders is consistent with the multiple variants with modest effects or intermediate penetrance that cause the disease phenotype in an additive and/or multiplicative way (Craddock et al. 2007).

1.1.7.1. Family studies

In a family study, a systematic evaluation of families where illness is known to affect one member provides a relatively straightforward means of testing the hypothesis of genetic transmission. An impressive number of family studies on bipolar disorder indicate that the disease aggregates in families. Smoller and Finn (2003) reported that first-degree relatives (parents, siblings and offsprings) of bipolar disorder patients have approximately a 10-fold increased risk of the disorder compared to the unaffected controls. Another study reported that about 15%-20% of the first-degree

relatives of bipolar probands are at an increased risk of bipolar or unipolar disorder, with the preponderance of unipolar, compared to unaffected controls (Barnett and Smoller 2009). Additionally, the risk of a same-sex sibling developing a mood disorder is greater than for a sibling of the opposite sex (Tsuang MT 1990). Unipolar depression is the most common affective disorder in the families of bipolar probands whereas the first-degree relatives of unipolar patients have increased morbidity risk for bipolar disorder compared to the general population, suggesting that the overlap of risk across these phenotypic categories indicates the shared genetic susceptibility (Gershon et al. 1982; Barnett and Smoller 2009). While it has been estimated that second-degree relatives have a slightly greater risk for mood disorders than the general population (Targum 1988), some studies have suggested that the risk to second-degree relatives is similar to that of general population for unipolar illness (Gershon et al. 1982; Tsuang MT 1990). However, these studies used telephone interviews by students. An increased risk of bipolar related psychiatric disorders including unipolar depression and schizoaffective illness is found to be common in relatives of bipolar probands. Also, the risk of schizophrenia is sometimes higher in relatives of bipolar probands. The heritability of the occurrence of psychosis within families multiply affected by bipolar disorder has been reported in some family studies (Craddock and Sklar 2009). Greater family history with early onset of bipolar disorder is well known, suggesting that early-onset cases are genetically more homogeneous (Baldessarini et al. 2012). Thus, gene-environment interactions, as evident from the family and twin studies of bipolar and unipolar disorder, play an important role in the presentation of varied clinical features of the disorder.

1.1.7.2. Twin and adoption studies

The genetic and environmental contributions to a disorder can be estimated using a population of twins by comparing the concordance rate among monozygotic (MZ) or dizygotic (DZ) twins. Twins usually share a common environment while sharing either all of their genes (identical or MZ) or half of their genes on average (DZ). The concordance rate for bipolar disorder is significantly higher among MZ twin pairs than among DZ twin pairs (Shih et al. 2004). Assuming that the shared environmental influences on MZ twins are not different from the environmental influences of DZ twins, significantly higher concordance rates in MZ twins reflect the action of genes. Also, the high rate of concordance appears to hold true even if the twins are reared apart (Price 1968). In a comprehensive review of twin studies published before 1990, Tsuang and Faraone (Tsuang MT 1990) reported a proband-wise concordance rate of 78% for MZ twins and 29% for DZ twins. The heritability of bipolar disorder in a hospital register study of 67 twin pairs in the UK (McGuffin et al. 2003) and in a population register study of 19,124 same-sex twin pairs in Finland (Kieseppa et al. 2004) was reported to be 89% and 93% respectively. The proband-wise concordance rate estimated for bipolar disorder in the two studies was 67% and 43% for MZ twins and 19% and 6% for DZ twins respectively (McGuffin et al. 2003; Kieseppa et al. 2004). However, a Danish study reported equal rate of bipolar disorder in twins and singletons (Klaning et al. 2004). A Norwegian study investigated the heritability of three different bipolar spectrum disorders – BP I, BP II, cyclothymia and different combination of these disorders in 303 same-sex twin pairs (Edvardsen et al. 2008). On analysis, the concordance rates were higher among MZ than DZ twins for all single diagnoses and main combinations of the diagnoses.

Additionally, the largest difference in concordance rate between MZ and DZ pairs was observed when the frequency of BP I or BP II among the co-twins of BP I probands was compared, with 50% of the MZ co-twins of BP I probands having either BP I or BP II while none of the DZ co-twins received this diagnosis (Edvardsen et al. 2008). For more accurate results, the concordance can be corrected for the age of onset and age at interview of the twins, however, this is rarely done. Interestingly, the concordance rate of bipolar disorder in MZ twins is not 100%, thereby implying that genes alone cannot explain bipolar disorder, but environmental factors also contribute to the disorder. Adoption studies of bipolar disorder have been carried out and suggest genetic transmission (Smoller and Finn 2003).

1.2. Molecular genetic studies of Bipolar disorder

The identification of genes or the variation of a quantitative trait that contribute to the risk of a disease has long been one of the goals of human and medical genetics. The aetiology of bipolar disorder is complex, probably involving multiple genes and environmental factors. The molecular genetic studies of bipolar disorder will be discussed in four different stages: linkage studies, association studies, GWAS and copy number variation (CNV) studies. Traditionally, both family linkage and allelic association methods have been used to narrow the location of such genes to small regions of the genome with the hope that the gene and eventually the causal variant may be identified and characterized. Before the era of GWAS, molecular genetic studies can be theoretically divided into positional and candidate gene approaches. In the positional approach, the chromosomal positions of the susceptibility genes are

determined using linkage studies that require no prior knowledge of disease pathophysiology. In contrast, the candidate gene approach utilizes linkage and association studies to recognise the genes involved in the disease based on the assumption that there is sufficient understanding of the disease biology required for shortlisting the genes to be investigated (Craddock and Jones 1999). Recently, the study of common variations (GWAS) and structural variations (CNVs) across the whole of human genome has become possible due to advances in technology and reduction in genotyping costs. Before discussing the findings from the genetic studies in bipolar disorder, understanding the components contributing to the complexity of the disorder is very crucial. Even though several studies have been carried out on bipolar disorder with some success, the inconsistency of the findings among all these studies can be attributed to locus heterogeneity, inadequate sample sizes and phenotypic heterogeneity (Schulze and McMahon 2003).

1.2.1 Factors affecting genetic studies

Locus heterogeneity

Locus heterogeneity describes a state when the same phenotype is caused by mutations at more than one gene locus. Allelic heterogeneity is a phenomenon when more than one allele at a specific gene locus results in the same phenotype (Pericak-Vance 2006). While linkage analysis is immune to allelic heterogeneity, locus heterogeneity reduces the power of linkage and association analysis in detecting disease variants. Ascertaining study samples from genetically isolated populations could help deal with locus heterogeneity to some extent by reducing the number of different susceptibility genes.

Choice of statistical methods for linkage analysis

The lack of understanding of the mode of inheritance is a major hindrance in deciding the statistical approach to be used for the linkage data analysis, either parametric or non-parametric methods. It can be assumed on the basis of locus heterogeneity that the likelihood of the detection of a genetic locus will be maximised if multiply affected pedigrees that contain large numbers of cases rather than affected sib-pairs (ASP) is considered for linkage analysis. The lack of knowledge of allele frequency, the penetrance rate and the mode of inheritance of the bipolar disorder points in the direction of using non-parametric linkage method whereas much greater power is achieved using large multiply affected families and parametric analysis compared to sib-pairs (Alaerts and Del-Favero 2009).

Sample size

Another major problem in linkage and association studies of complex disorders such as bipolar disorder is the sample size. Several meta-analyses of worldwide genome - wide scans have been performed in the last decade resulting in the detection of weak but inconsistent linkage signals across studies (Craddock and Sklar 2009). Suarez et al. (1994) found that for an oligogenic trait simulated under reasonable parameters, when sampling families sequentially, the first true linkage to be detected will not likely be replicated until many further samples have been studied. In conclusion, the power to replicate the first linkage is present only when many samples several times larger than the initial sample have been collected and studied (Badner and Gershon 2002).

Phenotypic heterogeneity

The lack of specific biological markers in bipolar disorder can result in variation in the validity of the diagnosis. A clear phenotype best capturing the underlying mechanism of bipolar disorder is yet to be defined. Accurate and reliable identification of the affected phenotype for BP I is as high as any medical disease but BP II disorder is a less valid and reliable diagnosis (Craddock and Jones 1999). Different diagnostic categories such as 'narrow' and 'broad' phenotype, bipolar disorder spectrum, referring to different phenotypes have been previously used in linkage studies. The endophenotypes of psychiatric comorbidity and treatment response have also been used in linkage studies of bipolar disorder to reduce phenotypic heterogeneity (Turecki et al. 2001; Gottesman and Gould 2003). These approaches have not been successful and suggest that the bipolar spectrum phenotype for genetic analysis needs to be as broad as possible.

Anticipation

Several studies have shown that early-onset bipolar disorder is characterized by an increased number of psychotic symptoms and manic episodes, higher level of comorbidity with drug addiction and poor clinical outcomes (Craddock and Sklar 2009). One explanation might be a phenomenon called anticipation whereby a disease becomes more severe and/or presents itself with an earlier onset as it is transmitted through the generations of a family. Anticipation is known to involve either methylation effects on susceptibility genes or the expansion of segments of the genome consisting of trinucleotide repeats and one might assume that this explains the utilization of large trinucleotide repeats in early linkage studies in bipolar disorder probands. Clinical evidence for anticipation in BPAD has been reported

previously but more studies are required to confirm and elucidate its underlying mechanism (McInnis et al. 1993).

1.2.2. Genetic markers

The genome among humans is approximately 99.5% identical and the 0.1-0.5% difference between individuals is caused by several types of sequence variations and polymorphisms (Levy et al. 2007). Polymorphic genetic markers are a critical component of a disease gene mapping study. Short tandem repeats (STRs) or microsatellites, initially described by Weber and May (Weber and May 1989) and Litt and Luty (Litt and Luty 1989), have traditionally been the most utilized markers for linkage studies (Pericak-Vance 2006). Microsatellites are short di-, tri- or tetranucleotide tandem repeat sequences widely distributed in the genome and are highly polymorphic among individuals. Single nucleotide polymorphisms (SNPs) are bi-allelic markers that are densely spread throughout the human genome occurring approximately once every 1000 bp (Pericak-Vance 2006). SNPs have been widely used for fine mapping of positional and functional candidate genes. Recently, sets of hundreds of thousands of SNPs on a single array chip have been used for GWAS for screening of the whole human genome.

1.2.3. Methods of linkage analysis

Linkage analysis refers to a group of statistical methods that are used to map a gene to the region of the chromosome in which it is located. Genetic linkage is a violation of Mendel's law of independent assortment according to which the alleles at two chromosomal locations assort independently and are transmitted to offsprings in random combinations. Linkage analysis in humans is more difficult than in

experimental organisms such as *Drosophila* because of the limitations in family size, inability to do test crosses, long generation time and the lack of knowledge of phase in parents who are heterozygous at both loci being studied. Linkage studies have previously been successful in the identification of genes that cause Mendelian disorders such as cystic fibrosis (CFTR) and Huntington disease (HTT) (Bailey-Wilson and Wilson 2011). In addition, linkage studies of families with strong aggregation of specific complex diseases have also led to the identification of rare, high-penetrance risk alleles in certain genes that cause large increases in susceptibility to complex diseases, for example the BRCA1 and BRCA2 genes and breast cancer (Bailey-Wilson and Wilson 2011). Complex genetic disorders such as the psychiatric illnesses are characterised by disease heterogeneity with an unknown mode of transmission. Epistatic interactions between genes and among their gene products as well as their interactions with environmental risk factors are also considered to be a plausible explanation for the complexity (Bondy 2011).

Genetic recombination is an important process associated with all forms of linkage analysis. During meiosis, a free exchange of genetic material occurs between homologous chromosomes in an equal and reciprocal fashion resulting in the formation of recombinants that are chimeras of maternal and paternal genetic material. Linkage is the tendency of two or more genetic loci to be transmitted together during meiosis because they are physically close together on a chromosome. The strength of genetic linkage is measured by the recombination fraction denoted by θ, the proportion of recombinants among all potential haplotypes produced by a parent. When the genetic loci are completely linked, no recombination occurs and the recombination fraction $\theta = 0$. On the other hand, genes segregating independently

appear genetically unlinked $\theta = 1/2$ (Pericak-Vance 2006). Linkage analysis was developed to detect excess co-segregation of the putative alleles underlying a phenotype with the alleles at a marker locus in family data. Linkage studies are feasible with much less dense sets of genetic markers compared to the current GWAS and also these methods have the power to detect co-segregation over a much larger genetic distance (e.g. 40 million bases of DNA). Linkage methods are particularly powerful for the detection of variants with a large effect size, which are rare in the population. Power to detect such loci using linkage methods can be enhanced by ascertaining families with aggregation of the trait of interest (loaded families). Interest in these methods is undergoing a renaissance due to the availability of next generation sequencing and its promise to allow identification of the rare variants underlying linkage signals (Bailey-Wilson and Wilson 2011).

Morton (1955) coined the term LOD score and put forward the use of sequential test procedures in linkage analysis recognizing that the typical method of sampling small families was in fact sequential. The parametric LOD score method was originally designed for Mendelian disorders and is considered less suitable for complex disorders (Bailey-Wilson and Wilson 2011).

The family pedigree method or the LOD score linkage analysis involves the comparison of likelihoods for specific genetic linkage hypotheses. First, the likelihood of observing a specific pattern of markers in a pedigree is calculated assuming the null hypothesis of no genetic linkage to be true. Next, the likelihoods of observing the pattern of disease and marker alleles for each of several alternative hypotheses of linkage are calculated and compared with the likelihood of the null hypothesis by means of an 'odds ratio' (Bailey-Wilson and Wilson 2011). The odds

ratio is the likelihood of an alternative hypothesis divided by the likelihood of the null hypothesis. The alternative hypotheses of linkage are specified by different values of the recombination fraction: (1) linkage is so close that no recombinations have occurred, $\theta = 0$; (2) linkage is tight, but one-recombination event has occurred in 100 meioses ($\theta = 0.01$); (3) linkage is quite close, but relatively more recombinations have occurred ($\theta = 0.05$) and so on for increments of $\theta = 0.05$ until 0.45. The null hypothesis specifies that $\theta = 0.50$ (Schulze and McMahon 2003). A LOD score, the base 10 logarithm of the odds ratio, greater than 3.0 is taken as evidence of linkage (Lachman 1994).

Lander and Kruglyak (1994) reported the results of a simulation analysis of a fully informative map of markers suggesting LOD scores for significant and suggestive linkage. According to them, the LOD score of 3.3 was required to achieve a true genome-wide significance in terms of linkage and a LOD of 1.9 to be suggestive evidence for linkage. The significant and suggestive LOD scores for ASP model were recommended to be 3.6 and 2.2 respectively. However, some researchers have criticized these thresholds as overly conservative (Schulze and McMahon 2003).

Non-parametric linkage analysis is based on the comparison of observed vs. expected sharing of the same alleles, known as identity by descent (IBD), between pairs of affected relatives. The ASP method is a widely used non-parametric linkage analysis method. The sib-pair method would be the least powerful method for psychiatric disorders where there is locus heterogeneity. Initially in the sib-pair method, individual genetic markers were studied one at a time, thus failing to use the full inheritance information provided by multipoint linkage analysis. Later, Lander and Kruglyak (Kruglyak and Lander 1995) described methods for complete

multipoint analysis to be used for sib-pair analysis methods. However, a major weakness of sib-pair analyses is that they do not allow the estimation of the recombination fraction and cannot estimate what proportion of families are linked to a specific locus assuming heterogeneity (Schulze and McMahon 2003).

1.2.4. Linkage findings in Bipolar disorder

Since the 1990's, linkage analysis has been utilized in vast majority of endeavours to discover the genetic loci predisposing to mental disorders. The strategy involves employing genetic markers and testing for linkage among families with multiple affected members.

Baron et al. (1993) and Egeland et al. (1988) conducted two of the earliest linkage studies of bipolar disorder in Israeli and Old Order Amish pedigrees and implicated chromosomal regions on X and 11q respectively. Following its identification, the Xq28 locus became the focus of several linkage studies with several studies often supporting linkage in bipolar disorder at this region (Mendlewicz et al. 1987; Lucotte et al. 1992; Berrettini 1998; Massat et al. 2002). Early age at onset, a higher ratio of bipolar illness to unipolar illness and a more recurrent and severe form of unipolar illness were found in the earlier linkage study of X-linked families (Baron et al. 1990). Following on the X-linkage of bipolar disorder, a genome scan of a large Finnish family with bipolar disorder indicated evidence for linkage at Xq24-q27.1 with the highest LOD score for the marker DXS994 (Pekkarinen et al. 1995). Additionally, this locus was then followed up with increased marker density in the large family and 40 additional families reporting evidence of linkage in only a fraction of the families (Ekholm et al. 2002). Stine et al. (1997) carried out a genome scan using markers on selected chromosomes in 97 families from the NIMH

Genetics Initiative Bipolar Group and found evidence supporting linkage at Xq26-28.

It was evident right from the beginning that replication of findings in complex disorders was a major challenge. Since then numerous genome-wide scans of bipolar disorder have been carried out in different ethnic populations and to date, loci showing linkage have been found on every single chromosome (Serretti and Mandelli 2008). Previously, the numerous single locus linkage studies have been reviewed with some strikingly convincing replications on 21q and Xq (Gurling 1986; Gurling et al. 1988; Gurling 1998). Independent replication of linkage signals for bipolar disorder has been compelling on 12q, 21q 16p, Xq, 15q (Craddock and Jones 1999).

The problem of sample size in the linkage studies has been addressed with meta-analyses of bipolar disorder in which samples from several studies were combined to determine the significance of findings (Tang et al. 2011). Recently, two main analytical approaches, Multiple Scan Probability (MSP) and the rank-based genome scan meta-analysis (GSMA) have been utilized in the meta-analysis of bipolar disorder data. MSP is known to be efficient and more powerful in detecting large effects that may have high variance, while GSMA has higher power to detect effects with small variance across independent studies (Tang et al. 2011). For ease of understanding, these approaches will be discussed in detail in the context of the studies that used them.

To date, four meta-analyses of genome-wide linkage scans on bipolar disorder have been conducted (Badner and Gershon 2002; Segurado et al. 2003; McQueen et al. 2005; Tang et al. 2011).

Badner and Gershon (2002) performed the first meta-analysis of genome-wide linkage scans of bipolar disorder that included 11 studies using the MSP technique. In this approach, the probability values (*P*) were taken from multiple studies and combined after correcting each value for the size of the region containing the minimum P-value (Badner and Gershon 2002). Strong evidence for linkage with bipolar disorder was obtained at 13q ($P = 6 \times 10^{-6}$) and 22q ($P = 1 \times 10^{-5}$) chromosomal regions. Interestingly, the same regions were also significant in their meta-analysis of genome-scans of schizophrenia (Badner and Gershon 2002). The authors argued that this method was more conservative and powerful than applying a statistical criterion to be met by at least one of the several individual studies. Also, unlike sample pooling of the data, the MSP was more robust to the presence of study heterogeneity where linkage was present in some but not all studies.

The second meta-analysis of bipolar disorder genome scans was performed by Segurado et al. (2003). This study included all published and unpublished data available from bipolar disorder genome scans with >20 affected cases and also data from every marker in each scan to avoid publication bias. The GSMA method is a nonparametric approach used in the meta-analysis that tolerates a degree of variability in sampling and statistical methods, yielding largely empirical measures of significance rather than an effect size. In other words, GSMA is also capable of detecting significant cross-study results for regions that are weakly positive in many studies but not sufficiently positive to have been presented in any study (Segurado et al. 2003). In total, 18 bipolar disorder genome scan data sets were included in the analysis. No chromosomal regions achieved genome-wide statistical significance, albeit some support for linkage regions on chromosomes 14q, 9p-q, 10q, 18p-q, and

8q was obtained. One of the limitations of the study is that X- and Y-chromosome data were not accommodated by the GSMA procedure.

McQueen et al. (2005) performed the largest and the most comprehensive mega-analysis of bipolar disorder genome scans. The authors conducted a combined analysis using the original genotype data from 11 bipolar disorder genome-wide linkage scans comprising 5179 individuals from 1067 families and found significant linkage on chromosomes 6q and 8q. Genome-wide suggestive linkage on chromosomes 9p and 20p was also observed in the same study. Combining the original genotype data from all the studies rather than using only linkage statistics and/or P values provided more power and control over sources of heterogeneity unlike in previous studies (McQueen et al. 2005).

Recently, Tang et al. (2011) used comparative linkage meta-analysis (CMLA) to compare the meta-analytic findings obtained under MSP and GSMA methods using results from published English language genome-wide linkage scans of bipolar disorder published between 2000 and 2010. The authors also examined the potential implications of convergent and discrepant results for the underlying genetic architectures of bipolar disorder given other genetic evidence from GWAS available. A total of 13 English-language genome-wide linkage scans were included in the MSP and GSMA analyses. The most significant MSP result was found on 5q14.3-q23.3, with 14q11.2-q13.1 and 5p13.3-q13.3 attaining Lander and Kruglyak (LK) significance criteria, and 14q12-q22.3 and 6p23-p21.1 retaining LK-suggestive criteria. As a result of the GSMA analysis, significant results were obtained for the regions 3p25.3-3p22.1 and 10p14-q11.21, and nominally significant results for 10q11.21-q22.1. Surprisingly, only 10p11.21-q22.1 and 10q22.1-q24.1 produced

partially-overlapping MSP-GSMA significant results. Interestingly, 10p11.21-q22.1 region contains the Ankyrin 3 (ANK3) gene which is a well replicated gene association finding in bipolar disorder. Evidence for association within *ANK3* gene comes from a bipolar GWAS (Lee et al. 2012), a bipolar GWAS meta-analysis (Ferreira et al. 2008), a combined bipolar and major depressive disorder (MDD) GWAS (Liu et al. 2011) and two recent combined GWAS meta-analysis (Wang et al. 2010; Sklar et al. 2011) suggesting that this gene is very likely among the contributors to linkage signals.

Several European linkage studies have been reported since the last meta-analysis of bipolar disorder linkage studies (McQueen et al. 2005).

A genome-wide linkage analysis of 52 families of Spanish, Romany and Bulgarian descent confirmed linkage on the previously implicated bipolar disorder susceptibility loci, 4q31 and 6q24 and also found borderline significance for 1p35-36 region based on the MOD-score (LOD score maximized over genetic model parameters) analysis method (Schumacher et al. 2005). The gene-gene interaction at the association level is a potential strategy in the identification of risk genes for complex disorders. With this approach in mind, the first genome-wide interaction and locus-heterogeneity linkage scan in BPAD using the same set of linkage data was performed and the strongest interaction was reported between genes on chromosomes 2q22-q24 and 6q23-q24, albeit evidence of locus heterogeneity was obtained at chromosome 2q (Abou Jamra et al. 2007).

Etain et al. (2006) performed a genome-wide linkage scan in 87 sib-pairs ascertained through an early-onset BP I disorder type proband and non parametric multi-point analysis found the most significant linkage at 3p14 with a LOD score of 3.51, which

increased to 3.83 on additional genotyping with increased marker density. Another whole genome linkage scan of bipolar disorder in an Irish population was reported by Cassidy et al.(2007) using STRUCTURE-Guided Linkage Analysis (SGLA), which is designed to reduce genetic heterogeneity and increase the power to detect linkage. The most significant linkage was obtained at 14q24 (NPL LOD = 3.27) for BP I only model of affection and suggestive evidence at 4q21, 9p21, 12q24 and 16p13 was also reported.

Hamshere et al. (2009) reported results of a genome-wide linkage scan in European families multiply affected by bipolar spectrum mood disorder, including individuals diagnosed with DSM-IV BP I or schizoaffective disorder, bipolar type. Covariate linkage analysis showed genome-wide suggestive evidence for linkage on chromosomes 1q32.3, 7p13 and 20q13.31. Interestingly, previous studies on bipolar disorder (Curtis et al. 2003) and schizophrenia (Ekelund et al. 2001; Ekelund et al. 2004) have reported linkage signals in the 1q region, close to Disrupted in schizophrenia gene (DISC I), that lies at 1q42. Also, evidence for significant familiality of lifetime occurrence of mood-incongruent psychosis, i.e., delusions or hallucinations that are not understandable within the context of predominant mood state, was observed in individuals with bipolar disorder in the study.

X-chromosome linkage again came to attention with a genome scan in sibling pairs with juvenile-onset mood disorder from 146 Hungarian nuclear families by Wigg et al. (2009) which documented evidence for linkage at Xq28 with a non-parametric multipoint LOD score of 2.10.

Mathieu et al. (2010) investigated the relationship between phenotypic heterogeneity in bipolar disorder and genetic heterogeneity among the eight susceptible loci

reported in a previous study (Etain et al. 2006) using additional markers and an extended sample of families. Analysis of the extended sample of families supported linkage in four regions, including 2q14, 3p14, 16p23 and 20p12. Genetic heterogeneity was observed between early- and late-onset BP I in the 2q14 region with only the early form of bipolar disorder being linked to this region. These results showed that stratification according to age of onset may be more relevant than DSM-IV classifications of mood disorders for the 2q14 region and for the identification of genetic susceptibility variations.

A recent method for analysing genome-wide SNP genotype data and family linkage analysis uses SNPs to identify extended segments showing increased IBD sharing within case-case pairs, relative to case-control or control-control pairs. A significant linkage signal on 19q was observed in cases reporting a positive family history of bipolar disorder or schizophrenia (Francks et al. 2010).

In the current study, I have investigated the 1p36 region with positive linkage in the family study (Curtis et al. 2003) using tests of linkage and allelic association. Therefore, in the next section I will summarize the linkage studies reporting findings on the 1p locus in bipolar disorder.

1.2.4.1. Summary of linkage studies at 1p locus

Evidence for linkage at the 1p35-p36 locus with a four-point LOD score of 2 was reported in a study involving UCL bipolar family pedigree sample and furthermore, a single pedigree reporting a LOD of 3.1 for the same region (Curtis et al. 2003). The highest nonparametric linkage score of 3.97 was obtained for chromosomal region 1p35-p36 in the Spanish sample on performing a genome-wide linkage study in 52

families of Spanish, Romany and Bulgarian descent and further fine mapping of the found susceptibility loci in additional 56 German families. MOD-score analysis provided borderline significance for the 1p region (Schumacher et al. 2005). Support for linkage at 1p36 has also been reported in isolated populations. Gypsies are a population comprising multiple genetically differentiated sub-isolates with strong founder effects and limited genetic diversity. A follow-up study (Kaneva et al. 2009) in Bulgarian Gypsy families with BPAD involving linkage analyses of the 12 genomic regions previously showing linkage (Schumacher et al. 2005) showed nominally significant support for the 1p36 region in individual families as well as in the super-pedigree with 181 members.

Cheng et al. (2006) conducted a genome-wide linkage scan in 1,060 individuals from 154 multiplex families from NIMH initiative and performed linkage analyses using standard diagnostic models and also using three phenotypic subtypes, psychosis, suicidal behaviour and panic disorder. Suggestive linkage was found at 1p13 and 1p21 for the psychosis sub-phenotype, although no signal was observed for these loci using standard diagnostic models. Studies suggesting linkage at 1p with bipolar disorder have also been reported in Italian (Vazza et al. 2007), Portuguese (Fanous et al. 2012), Colombian (Kremeyer et al. 2010) and Australian population (Fullerton et al. 2008; Fullerton et al. 2009). In a genome-wide linkage scan (Fullerton et al. 2008) on 15 Australian extended families with bipolar disorder support for 1p13-31 was observed. Further study on pooled genotype data from 65 Australian extended pedigrees detecting evidence of genetic interaction reported robust interchromosomal interaction exceeding Bonferroni correction between regions 4q32-35 and 1p36 (Fullerton et al. 2009). Linkage analysis of 47 Portuguese families with sib-pairs

concordant for schizophrenia, bipolar disorder or psychosis revealed significant linkage for chromosome 1p36 with a LOD of 3.51 with bipolar disorder (Fanous et al. 2012). Also a meta-analysis of schizophrenia and bipolar datasets as part of a population based linkage analysis provided evidence for linkage on 1p34 (Francks et al. 2010).

The initial excitement generated by the identification of linkage markers for bipolar illness dissipated quickly due to the lack of replication of many of the early linkage results. Susceptibility loci for bipolar loci recognised by analysing the meta-analyses and the European linkage studies include 8q, 14q, 10q, 1q, 3p and 2q. A recent study confirmed replication of susceptibility loci 6q25 and 17p12 implicated previously (Kelsoe et al. 1996; Lachman et al. 1997; Kelsoe et al. 2001), in an extended and independent sample of 34 pedigrees segregating bipolar disorder (Greenwood et al. 2012).

1.2.5. Candidate gene association studies of Bipolar disorder

A meta-analysis of genetic association studies on bipolar disorder using random-effects meta-analysis approach on all polymorphisms reported in three or more case-control studies was recently published by Seifuddin et al. (2012). The results from the meta-analysis were then compared with the findings from a mega-analysis of eleven GWAS in bipolar disorder (Sklar et al. 2011). Nominally significant associations ($P<0.05$) were found for polymorphisms in Brain derived neurotrophic factor (BDNF), Dopamine Receptor D4 (DRD4), D-amino acid oxidase activator (DAOA) and Neuronal tryptophan hydroxylase (TPH1), although the findings did not survive multiple testing correction. Only the most replicated and consistent association findings will be discussed further as a comprehensive review of all the association studies is beyond the scope of this thesis.

One of the most widely investigated genes in bipolar disorder is the serotonin transporter, SLC6A4, located on 17q reported to be positive in two linkage studies. Several association studies and four meta-analyses have also been reported with positive findings (Serretti and Mandelli 2008). TPH2 gene encoding rate-limiting enzyme in the biosynthesis of serotonin in the central nervous system has been positively associated with bipolar disorder in several independent studies, albeit negative associations have also been reported in different populations (Cichon et al. 2008; Serretti and Mandelli 2008; Barnett and Smoller 2009; Choi et al. 2010). Serotonin receptor genes HTR1A, HTR2A, HTR2C and HTR7 have also been investigated in a few studies with some evidence of association (Serretti and Mandelli 2008).

Dopaminergic genes DRD1, DRD3, DRD4 and SLC6A3 have been subjected to investigations in several studies in different populations. Positive reports for a promoter polymorphism and haplotype association along with negative SNP associations have been reported for DRD1 (Nothen et al. 1992; Cichon et al. 1996; Severino et al. 2005; Dmitrzak-Weglarz et al. 2006). DRD3 has been reported to be positive in only one study. There is considerable support for the involvement of DRD4 and SLC6A3 genes in bipolar disorder based on several studies (Serretti and Mandelli 2008).

Catechol-O-methyltransferase (COMT) is another promising candidate gene that encodes for an enzyme catalysing the degradation of catecholamines including dopamine, epinephrine and norepinephrine. The functional Val158Met polymorphism in this gene has been widely investigated with mostly negative results (Shifman et al. 2004; Serretti and Mandelli 2008). A recent study has reported interaction between DRD3 Ser9Gly polymorphism and COMT Val158Met polymorphism in BP I patients (Lee et al. 2011b), but this is unreplicated.

BDNF is a gene selected for having evidence for both positional and functional involvement in bipolar disorder. This gene is located at 11p13, a region consistently linked to bipolar disorder in previous linkage studies which also contains the DRD4 locus (Pato et al. 2004; Craddock and Sklar 2009). BDNF codes for a prosurvival factor induced by cortical neurons that is necessary for survival of striatal neurons in the brain. Reports suggesting effects of the Val158Met polymorphism on BDNF secretion and memory have been published (Egan et al. 2001). Variants in BDNF have also been reported to be associated with working memory and schizophrenia (Diaz-Asper et al. 2008). Two meta-analyses have confirmed association between

BDNF and bipolar disorder, particularly in Caucasians (Serretti and Mandelli 2008; Craddock and Sklar 2009).

DAOA is a gene pertaining to the glutamatergic pathway that has been associated with bipolar disorder in several studies. Located in a positive linkage region, 13q33.2, DAOA plays an important role in the activation of NMDA receptors, which are implicated in schizophrenia and bipolar disorder (Chumakov et al. 2002; Williams et al. 2006). Several independent Caucasian studies have reported positive associations with different SNPs in bipolar disorder (Hattori et al. 2003; Chen et al. 2004; Schumacher et al. 2004; Prata et al. 2008; Bass et al. 2009; O'Donovan et al. 2009; Cherlyn et al. 2010; Nothen et al. 2010).

DISC1 was implicated cytogenetically in psychosis and has been found to be associated with bipolar disorder in several studies with an inconsistent pattern of association (Serretti and Mandelli 2008). Initially, DISC1 was found to be disrupted by a translocation that co-segregated with major psychiatric illness including schizophrenia, bipolar disorder, recurrent depression and childhood-onset behavioural disorder (St Clair et al. 1990). Since then multiple studies have confirmed associations between DISC1 and bipolar disorder, however, a pathogenic variant is still to be detected (Craddock and Sklar 2009; Hennah et al. 2009).

Several other genes with replicated positive findings have been identified including TRPM2, GSK3β, DTN bipolar disorder1, NRG1, NCAM1, GRIN2B, MOAO, P2RX7 and GABA receptor genes (McQuillin et al. 2006; Lachman et al. 2007; Serretti and Mandelli 2008; Craddock and Sklar 2009). Genes involved in circadian rhythms have been investigated, including CLOCK and BMAL1 (Barnett and Smoller 2009). The evidence for allelic association with bipolar disorder at various

genes is convincing. The problem of locus heterogeneity gives rise to the "Winner's curse" problem in which an initial report of a genetic association takes a long time to replicate (Zollner and Pritchard 2007). In view of the problem of locus and allelic heterogeneity, extremely large samples are required for robust identification of genes influencing risk and expression of bipolar disorder.

1.2.6. Genome-wide association studies in Bipolar disorder

A GWAS involves a comprehensive interrogation of the common variations across the entire human genome to identify genetic associations with observable traits without the need for prior hypothesis about the location of the variants. Due to the extensive correlation or LD between SNPs, the majority of all common genetic variation is captured by genotyping only a small proportion of the total number of known SNPs using this technique. The genetic architecture of psychiatric disorders has proven to be complex with two contrasting hypotheses being routinely used to explain it - Common Disease Common Variant (CDCV) and Common Disease Rare Variant (CDRV) hypotheses. According to CDCV, the genetic risk in an individual and in the population is attributable to many high-frequency variants each conferring modest level of risk (Risch and Merikangas 1996). Alternately, the CDRV hypothesis states that a large number of rare variants with relatively large effects contribute to the disease risk such that hundreds of genes, each with many rare variants, contribute to the phenotype (Bodmer and Bonilla 2008).

Based on the conception that psychiatric disorders have a negative effect on the fecundity, natural selection would prevent the causal alleles from reaching common frequencies in a population, which supports the CDRV hypothesis (Alaerts and Del-Favero 2009). Successful identification of common risk variants in complex diseases by GWAS has transformed the field of medicine. The list of genetic variants for common somatic diseases has expanded considerably within the last few years leading to novel genetic findings resulting in a better understanding of diseases, such as cancer (Elliott et al. 2010) and obesity (Scherag et al. 2010). As a result of

GWAS, several new insights into the field of cancer biology are already being translated into clinical benefits for the patients and the potential findings of GWAS might soon be part of risk screening programmes in other diseases in the future (Bondy 2011).

GWAS combine the merits of association studies in the power to detect small effects with that of linkage studies in that they do not require any knowledge of the pathogenesis of the disease. The GWAS approach requires correction for extensive multiple testing and for population stratification. The large samples required for performing GWAS provide power to detect variants with small or large effects. Meta-analyses of bipolar disorder GWAS (Chen et al. 2011) can focus on high or lower ranking signals suggesting susceptibility loci rather than on the 'top hits' of individual studies. Locus and/or allelic heterogeneity will play an important role in diluting true signals even with the usage of large samples (Pritchard and Cox 2002). A disadvantage of using large samples is that the LD variation between populations could be confounded with disease associations. Population structure/stratification is being routinely assessed and corrected for in the GWAS using principal component analysis method (Price et al. 2006), logistic regression, identical by state (IBS) clustering (Purcell et al. 2007) or linear mixed models (Alaerts and Del-Favero 2009).

Imputation methods are used in GWAS to improve the coverage of the human genome across different genotyping platforms. Imputation methods (Marchini et al. 2007; Purcell et al. 2007) impute and test ungenotyped SNPs for association in a sample thus adding to the information content. Analysis of GWAS data using more sophisticated statistical methods may provide additional power leading to new

results and a better understanding of the best ways to interpret the GWAS data. Some of the alternative methods previously employed in the analysis of GWAS data include multipoint methods combining information across markers with more powerful tests for indirect association with untyped variants, simultaneous analysis of all SNPs to identify the subset that best predicts disease outcome, and weighted analyses incorporating prior information and pathway-based approaches, which jointly consider multiple variants in interacting or related genes (Alaerts and Del-Favero 2009).

1.2.6.1. Findings of GWAS

Several GWAS on bipolar disorder have been published in the last five years revealing several interesting loci, which are not found in all samples consistently. A catalogue of published GWAS is available at http://www.genome.gov/gwastudies/. The first GWAS on bipolar disorder was reported by Baum et al. (2008) using 461 BP I patients and 562 healthy controls of European Caucasian background that were obtained from NIMH. A German replication sample consisting of 772 BP I patients and 876 matched controls was also used in the study. All patients met DSM-IV criteria for BP I. DNA pooling was used to reduce the cost of genotyping in this study and over 550,000 SNPs were genotyped in the two samples. Pooling of DNA samples reduces the power to detect genetic association but is a cost effective method to genotype large number of samples. The strongest association signal was detected within the first intron of DGKH (diacyl glycerol kinase) ($P = 1.5 \times 10^{-8}$, OR=1.59).

The Wellcome Trust Case Control Consortium (WTCCC) (2007) conducted a joint GWAS to identify genes linked to seven different diseases including bipolar disorder in a British population. This bipolar GWAS was a result of the collaboration of three different research institutions. The Affymetrix GeneChip 500K Mapping Array set was used in this study to genotype ~2000 cases of bipolar disorder and a shared set of 3000 controls. Although several regions associated with bipolar disorder were found, none reached genome-wide significance level. The strongest region showing association with bipolar disorder was at chromosome 16p12 that included genes PALB2 (partner and localizer of BRCA2), NDUFAB1 (NADH dehydrogenase (ubiquinone) 1, alpha/beta subcomplex, 1) and DCTN5 (dynactin 5). The study also found association with KCNC2, which encodes the Shaw-related voltage-gated potassium channel, suggesting a role for ion channelopathies in bipolar disorder.

The second GWAS by a consortium was reported by Sklar et al. (2008) that included non-overlapping Caucasian case-control samples from several research groups. The study sample consisted of 1461 BP I cases and 2008 controls from systematic treatment enhancement program (STEP-BD) study and UCL. A total of 372,193 SNPs were genotyped in the study using Affymetrix GeneChip Human Mapping 500K Array Set. Again, none of the SNPs met genome-wide significance on association analysis but the top SNPs associated with bipolar disorder were in genes MY05B ($P = 1.66 \times 10^{-7}$), TSPAN8 ($P = 6.11 \times 10^{-7}$) and EGFR ($P = 3.26 \times 10^{-5}$). Replication was attempted by genotyping 304 SNPs in family-based NIMH samples (n=409 trios) and University of Edinburgh case-control samples (N = 365 cases, 351 controls) and the results showed that 9 out of the 13 SNPs that showed association were replicated. The other interesting genes suggesting association with bipolar

disorder reported in this study were CACNA1C, MBP, DFNB31, GRM3 and HLA-DRA. This study did not identify any association between DGKH and bipolar disorder as shown previously in a bipolar disorder GWAS (Baum et al. 2008).

In another collaborative approach, large sample sets from STEP-BD, University of Edinburgh and Trinity College, Dublin (ED-DUB-STEP2), were put together to perform a GWAS on non-overlapping samples (Ferreira et al. 2008). Over 300,000 SNPs were genotyped in 1098 bipolar cases and 1267 controls and on analysis no chromosomal regions were found to be associated at genome-wide significance level. Interestingly, CACNA1C was one of the fourteen regions that showed association at $P < 5x10^{-5}$. Furthermore, the genotyping data from the WTCCC, STEP-BD-UCL, and ED-DUB-STEP2 studies (4,387 cases and 6,209 controls) was combined with additional imputation data from 1,444,258 HapMap SNPs using PLINK (Purcell et al. 2007). The strongest associations were observed for the imputed SNP in ANK3 (rs10994336, $P=9.1 \times 10^{-9}$), a SNP in the third intron of CACNA1C (rs1006737, $P=7.0x10^{-8}$) and for the 15q14 region containing C15orf53 and RSGRP1 (rs12899449, $P = 3.5 \times 10^{-7}$). These results again pointed in the direction of ion channelopathies being important in the pathogenicity of bipolar disorder.

Scott et al. (2009) performed a GWAS by genotyping 2,076 bipolar cases and 1,676 controls of European ancestry obtained from the research groups - NIMH/Pritzker, and samples collected in London, Toronto and Dundee sponsored by Glaxo Smith Kline research and development (GSK) using the Illumina HumanHap550 Beadchip. SNP imputations were performed and association analysis was carried out in each sample individually followed by a 2-study meta-analysis. None of the SNPs

achieved genome-wide significance. A 3-study meta-analysis performed by adding genotype data from WTCCC resulting in an overall non-overlapping sample of 3,683 cases and 14,507 controls resulted in 3 regions with association $P \approx 10^{-7}$, 1p31.1, 3p21 and 5q15. Additionally, previously reported association between bipolar disorder and the ANK3 SNP rs10994336 (Ferreira et al. 2008) was replicated in the non-overlapping GSK sample (P = 0.042; OR = 1.37).

Around the same time, Smith et al. (2009) conducted two separate GWAS in individuals of European ancestry (EA; n= 1001 cases, 1033 controls) and of African ancestry (AA; n = 345 cases, 670 controls). A genetic component of the Bipolar Consortium (sponsored by the US National Institute of Mental Health), termed the 'Bipolar Genome Study (BiGS)' and funded by the Foundation for the National Institutes of Health Genetic Information Association Network (GAIN) initiative, contributed to the EA samples in this GWAS and some samples have been used in other GWAS (Baum et al. 2008; Sklar et al. 2008; Scott et al. 2009). Although not an independent sample set, the study found the strongest association signals at Xq27.1 ($P=1.6x10^{-6}$) and 2q21.1 (NAP5; $P=9.8x10^{-6}$) for the EA sample, and 19q13.11 (DPY19L3; $P=1.5x10^{-6}$) and 9q21.33 (NTRK2; $P=4.5x10^{-5}$) for the AA sample. Again, ANK3 region showed consistent but modest replication in the EA sample (rs1938526, P=0.036, OR = 1.31) and borderline support for replication was observed for the 15q14 region containing C15orf53. However, the study failed to show any support for CACNA1C association.

Recently, the Psychiatric GWAS Consortium (PGC) (Sklar et al. 2011) reported the results of a combined GWAS of 7481 bipolar disorder cases and 9250 controls. This dataset was obtained by dividing the primary genotype and phenotype data for

overlapping case-control samples from seven previously published studies (Baum et al. 2008; Ferreira et al. 2008; Sklar et al. 2008; Scott et al. 2009; Smith et al. 2009; Djurovic et al. 2010; Cichon et al. 2011) into 11 different case and control groupings and assigning each individual to only one group. On the basis of diagnosis, the cases were made up of 84% BP I, 11% BP II, 4% schizoaffective disorder and 1% with other bipolar diagnoses. The two regions reaching genome-wide significance in this sample were ANK3 on chromosome 10q21 (imputed SNP rs10994397, $P = 7.1 \times 10^{-9}$, OR = 1.35) and SYNE1 on chromosome 6q25 (rs9371601, $P = 4.3 \times 10^{-8}$, OR = 1.15). Furthermore, two other association signals nearing genome-wide significance were observed on chromosome 12 and in the region of ODZ4. For replication, unpublished data from research groups on an independent sample consisting of 4,496 bipolar disorder cases and 42,422 controls was used. Although ANK3 and SYNE1 associations failed to replicate, SNPs mapping to CACNA1C and C11orf80 were found to be significantly associated with bipolar disorder following a correction for multiple testing. A fixed-effects meta-analysis performed on the combined sample of 11,974 cases and 51,792 controls identified strong genome-wide association signals in the regions containing CACNA1C ($P=1.52 \times 10^{-8}$, OR = 1.14) and ODZ4 ($P=4.40 \times x10^{-8}$, OR=0.89). A pathway analysis on the combined sample showed significant enrichment of genes encoding calcium channel subunits, CACNA1C and CACNA1D. Sklar et al. (2011) also reported a strong association for SNPs at the CACNA1C locus and in the region of NEK4-ITIH1-ITIH3-ITIH4 based on the combined GWAS analysis of schizophrenia and bipolar disorder.

Smith et al. (2011) performed a GWAS on a total of 2191 cases and 1434 controls of European ancestry collected as part of the BiGS-TGEN and the GAIN initiative.

This study sample comprised of 1,190 new bipolar disorder cases from the BiGS study and 401 controls (TGEN sample), along with 1,001 cases and 1033 controls previously genotyped through the GAIN initiative (Smith et al. 2009). No genome-wide significant associations for any region were reported. Also, the authors did not find any genome-wide significant associations when they performed a meta-analysis of SNP genotype data from the current study and WTCCC. Based on the results, they hypothesised that the presence of likely common genetic variations associated with bipolar disorder near exons (±10 kb) could be identified in larger studies and provide a framework for assessing the potential for replication when combining results from multiple studies (Smith et al. 2011). A further need is for a Bayesian analysis of the GWAS data which can take the place of corrections for multiple testing.

A new bipolar disorder susceptibility candidate, Neurocan (NCAN) has recently been reported reaching genome-wide significance ($P=3.02 \times 10^{-8}$, OR=1.31) in a GWAS conducted by Cichon et al. (2011) on European samples. The association was replicated in two follow-up steps in the study in independent samples from Europe, USA and Australia surpassing genome-wide significance threshold and the combined analysis of all the samples with 8441 cases and 35,362 controls also confirmed the association ($P=2.14 \times 10^{-9}$, OR=1.17).

GWAS on bipolar disorder in Han Chinese (Lee et al. 2011a), Norwegian (Djurovic et al. 2010), Japanese (Hattori et al. 2009) and Bulgarian (Yosifova et al. 2011) samples have also been reported. Although, none of these studies found any SNPs showing genome-wide level significance, replication in the form of a modest

association was reported for ANK3 in the Chinese study (Lee et al. 2011a) and DGKH in the Bulgarian GWAS (Yosifova et al. 2011).

A meta-analysis of 5 major mood disorder case-control samples including bipolar disorder and MDD of European ancestry found genome-wide significant association at 3p21.1 (rs2251219, $P = 3.63 \times 10^{-8}$, OR = 0.87) (McMahon et al. 2010). Another meta-analysis of genome-wide association data of bipolar disorder and MDD was conducted by Liu et al. (2011). The bipolar cases were from the US, UK and Ireland (WTCCC 2007; Sklar et al. 2008) and the MDD subjects (Penninx et al. 2008) were from the Netherlands. This study provided evidence for the role of CACNA1C variants in bipolar and unipolar major mood disorders but failed to confirm the previous association of bipolar risk locus ANK3 in the meta-analysis. Even larger collection of bipolar disorder GWAS signals was analysed for age of onset and psychotic symptoms and two sub-phenotypes of bipolar disorder. These included GWAS data from GAIN-BP, BiGS, and a German study but it failed to report any significant genome-wide associations (Belmonte Mahon et al. 2011).

Wang et al. (2010) performed a meta-analysis for GWAS data in 653 bipolar cases and 1034 controls, and 1172 schizophrenia cases and 1379 controls and found two intergenic SNPs in the 9q33 region showing genome-wide significance. No replications of previously associated SNPs were reported (Wang et al. 2010). Another study comparing the marker allele frequencies between schizophrenia and bipolar disorder cases rather than with controls reported a significant association for CACNG5 that supports the role of voltage-dependent calcium channel genes such as CACNA1C in the susceptibility to bipolar disorder (Curtis et al. 2011).

Crossing the intercontinental boundaries, Chen et al. (2011) recently performed a meta-analysis of 7773 bipolar disorder cases and 9883 controls of European and Asian ancestry and found novel significant associations near the genes TRANK1 (LBA1), LMAN2L and PTGFR. *In vitro* studies supporting a functional role for TRANK1 showed that valproic acid markedly increased TRANK1 mRNA expression in a dose- and time-dependent manner in SH-SY5Y, HeLa, and HEK293 cells. Additionally, this study also reported associations between ANK3 and bipolar disorder as well as markers on a chromosome 3p21 locus.

The current generation of GWAS technologies is best able to detect association of genetic variants present in 1% or more of the population. SNP variants in ANK3, ODZ4, SYNE1, CACNA1C, DGKH, and C15orf53 have consistently been associated with bipolar disorder in different GWAS. Another approach termed as gene-wide analysis of GWAS data sets, which avoids correction for multiple SNPs within a gene, has also provided additional evidence for association to these genes reported in the previous GWAS (Moskvina et al. 2009). With the growing body of evidence for the 'common disease multiple rare variants' (Gershon et al. 2011) hypothesis of bipolar disorder, there is an urgent need to pursue approaches designed to detect rare variants contributing to bipolar disorder. Broadening the spectrum of detectable risk alleles can be achieved by next generation GWAS technologies aimed at enhancing detection of rare variants and polymorphisms present at <1%, thereby enhancing the disease association detection. Further functional studies to elucidate the mechanism of risk contributed by common and rare variants are warranted as well.

1.2.7. CNVs in Bipolar disorder

CNVs are structural variations of the DNA that include insertions, deletions, inversions and duplications that may vary in length from a few kilobases to several megabases (Lee et al. 2012). Two human chromosomes in a population differ at a rate of 0.1% according to the early surveys of genetic variation (International Hapmap Consortium 2005; Marian 2012). Although SNPs account for the most numerous variants in the genome, CNVs have been reported to involve upto 12% of the human genome contributing to a sizeable amount of phenotypic variation within normal individuals and complex diseases including psychiatric disorders like schizophrenia and bipolar disorder. CNVs are classified as de novo and inherited. De novo CNVs can occur at the rates of up to four folds greater than single nucleotide substitution rates (Lupski 2007). In comparison to SNPs, the overall genomic change as measured in nucleotides is high in the case of CNVs, in the order of 1000 bp per generation, leading to a larger functional impact per site (Malhotra and Sebat 2012). CNV deletions and duplications possibly interfere with the regulatory regions or coding sequences of various genes thereby altering their genetic makeup and biological functions. Additionally, the presence of CNVs can result in the up- or downregulation of dosage sensitive genes, thereby, contributing to disease susceptibility variations (Lee et al. 2012). Based on the sheer size and potential to impact genes or multiple genes, structural variations are more pathogenic on average and de novo CNVs might be more enriched in variants that have a large effect on disease risk (Malhotra and Sebat 2012). CNV studies are a formal test of the CDRV hypothesis. A widespread distribution of submicroscopic variations (<500 kb in size) in

DNA copy numbers in normal human genomes was demonstrated in two landmark studies conducted in 2004 (Iafrate et al. 2004; Sebat et al. 2004). Also, CNV deletions accounting for the variations in intermediate phenotypes within complex neuropsychiatric illnesses such as cognitive impairment or physiological measures have previously been reported (Friedman et al. 2008).

The literature available on the studies examining structural variations in bipolar disorder is very limited. In a postmortem brain DNA screening study by Wilson et al. (2006), several loci including GRM7, AKAP5 and CACNG2 were reported to contain CNVs. These proteins are highly expressed in the brain and are involved in glutamatergic signalling which is integral to the normal neuronal function (Wilson et al. 2006). Another study by Zhang et al. (2009) analyzed CNVs (≥100 kb) in 1001 bipolar patients and 1034 controls as part of a genome-wide CNV survey. Although the singleton deletions were present at a higher frequency in patients (16.2%) compared to controls (12.3%), the difference was more marked in patients with early age at onset of mania (Zhang et al. 2009). A possible limitation of the study was the recruitment strategy employed for enrolling cases and controls. The bipolar cases were recruited at 11 data collection sites using DSM-IV criteria, whereas the controls were recruited separately through a volunteer panel and a web-based psychiatric interview.

A genome-wide survey of large, rare CNVs in a case-control sample from WTCCC (Grozeva et al. 2010) found that the rate of CNVs was higher in controls than in bipolar cases. The absence of rare and large structural deletions associated with bipolar disorder and their presence in 5%-10% of schizophrenia cases shows that the two disorders are

genetically distinct for this type of genetic variation. Similar results were obtained in another study utilizing the UCL bipolar disorder research sample (McQuillin et al. 2011), in which the rare CNV burden was also slightly lower in the cases compared to controls. Both the studies (Grozeva et al. 2010; McQuillin et al. 2011) also failed to report any association between age-at-onset and singleton deletions as reported previously (Zhang et al. 2009). The fact that both studies found higher rates of CNVs in controls is difficult to explain but it could be the presence of schizophrenia CNV carriers or undiagnosed cases in the controls that are responsible.

A combined analysis of CNVs in three case-control affective disorder samples from Denmark, Norway and Iceland reported no association between the CNVs and affective disorders (Olsen et al. 2011). CNVs at 10 genomic loci were investigated in 1223 unipolar cases, 463 bipolar cases and 11,231 controls. A meta-analysis with the inclusion of CNV data from the WTCCC study (Grozeva et al. 2010) for the selected loci showed nominal association for only a 9q31.1 locus ($P = 0.041$, $OR = 3.8$).

In the first genome-wide survey involving family trios for the identification of CNVs in bipolar disorder (Malhotra et al. 2011), the frequencies of de novo CNVs were significantly higher in bipolar disorder patients compared to controls ($n= 185$, $P = 0.009$, $OR = 4.8$). Furthermore, the enrichment of these de novo CNVs was more evident among the cases with early age-at-onset ($\leq$18 years). This study also included schizophrenia cases and concluded that these cases were also enriched with de novo CNVs. In contrast, the results of this study were inconsistent with other family-based studies of schizophrenia (Xu et al. 2008) and autism (Levy et al. 2011) that have found a

strong genetic effect of de novo mutations and a weak genetic effect for inherited variants. Recently, a GWAS in a Swedish population reported support for a greater CNV involvement in schizophrenia compared to bipolar disorder (Bergen et al. 2012). Cumulative evidence on complex psychiatric diseases also suggests that de novo CNVs are being found at a higher frequency among sporadic cases, whereas inherited CNVs are enriched among familial cases (Levy et al. 2012).

Priebe et al. (2012) performed a search for the presence of CNVs using genome-wide SNP data obtained from a bipolar disorder GWAS (Cichon et al. 2011) conducted in 882 patients and 872 controls. Genome-wide burden analyses revealed that the frequency and size of CNVs did not differ substantially between cases and controls. However, a separate analysis of patients with early-onset bipolar disorder (age-at-onset $\leq$21 years) and late-onset bipolar disorder (age-at-onset $\geq$21 years) revealed that the frequency of microduplications was significantly higher ($P = 0.0004$) and the average size of singleton microdeletions was significantly larger ($P=0.0056$) in patients with early-onset bipolar disorder compared to controls. Two microduplications - 160 kb on 10q11 and 248 kb on 6q27 - were overrepresented in the early-onset subgroup of bipolar disorder patients compared to controls (Priebe et al. 2012). These results provided further support for the influence of CNVs on early-onset bipolar disorder, and also suggested the presence of an aetiological difference between early-onset and late-onset bipolar disorder.

In conclusion, the potential role of CNVs in bipolar disorder is highlighted by some publications reporting overrepresentation of CNVs (Zhang et al. 2009; Malhotra et al.

2011) in patients with bipolar disorder and others being negative for any associations (Grozeva et al. 2010; McQuillin et al. 2011). The effect of de novo CNVs on age-at-onset in bipolar disorder has consistently been nominally significant in several studies (Zhang et al. 2009; Malhotra et al. 2011; Priebe et al. 2012) suggesting that individuals with early-onset bipolar disorder may represent a subclass of the disorder in which there is a greater contribution from rare CNV alleles of large effect. Support for the notion that segregation of early-onset bipolar disorder in families is consistent with major gene effects. However, familial segregation of bipolar disorder and unipolar disorders shows a wide range of ages of onset and it is not possible to identify clear cut differences in age of onset between families. In conclusion, several lines of evidence suggest that the overall rare CNV burden is more modest for bipolar disorder, with schizophrenia showing a large increase in the number of CNVs found to be associated compared to controls (Girirajan and Eichler 2011).

1.3. Glutamate receptors

Glutamate is the primary excitatory neurotransmitter in the human brain. Under normal conditions, glutamate plays an important role in synaptic plasticity, learning and memory, but in pathological conditions it is known to be a potent neuronal excitotoxin, triggering either rapid or delayed neurotoxicity (Sanacora et al. 2008). Glutamate acts on presynaptic and postsynaptic neurons and glia, the three different cell compartments characterizing the 'tripartite glutamatergic synapse'. Disturbances in glutamate-mediated transmission have been increasingly documented in a range of neuropsychiatric disorders including mood disorders, Alzheimer's disease, schizophrenia, substance abuse and autism spectrum disorders (Moghaddam and Javitt 2011). In the Central Nervous System (CNS), glutamate binds to both ion channel-associated (ionotropic glutamate receptors, iGluRs) and G-protein-coupled receptor subtypes (metabotropic glutamate receptors, mGluRs) mediating direct as well as fast excitatory and second-messenger evoked transmission respectively (Cartmell and Schoepp 2000). Thus, neuronal signalling can be modified by presynaptic glutamate receptors that directly regulate glutamate release, or by glutamate receptors localized postsynaptically on glutamatergic and nonglutamatergic neurons. Presynaptically, the iGluRs and mGluRs act by controlling the strength of synaptic transmission by altering the likelihood that the synaptic vesicles would be released in response to an incoming action potential (Pinheiro and Mulle 2008).

1.3.1. Ionotropic glutamate receptors

Ionotropic glutamate receptors include NMDA, α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic (AMPA) and kainate receptors. The NMDA receptors (NMDAR) have a slower and more prolonged postsynaptic current than the AMPA or kainate receptors (KAR). All iGluRs exhibit fast receptor deactivation and dissociation of glutamate. NMDA receptors are ligand-gated ion channels activated in a voltage-dependent manner and made up of an obligate NMDAR subunit 1 (NR1) together with 1 of 4 subunits of NR2A-D. The activation of NMDARs requires glutamate and glycine along with membrane depolarization in order to lift the magnesium block. An intracellular site for non-competitive antagonists such as phencyclidine (PCP), ketamine and MK-801 is also present in NMDARs (Cherlyn et al. 2010; Machado-Vieira et al. 2012). Reduced levels of NMDARs have been reported in postmortem brain from bipolar disorder patients (Rao et al. 2012). The AMPA receptors participate in fast excitatory transmission and their activation opens the pore permitting the inward flow of sodium resulting in the depolarization of the neuronal membrane. Co-expression of AMPA receptors with NMDA receptors has been observed at mature synapses, suggesting that they both contribute to synaptic plasticity and neuroprotection. However, AMPA receptors have a lower affinity for glutamate compared to NMDA receptors, which allows for a more rapid dissociation of glutamate and a fast deactivation of the AMPA receptors (Zarate et al. 2010). Decreased levels of AMPA receptor subunits (GluR1, GluR2 and GluR3) have been reported in the prefrontal cortex and striatum of subjects with mood disorders and transgenic animals with lower GluR1 expression exhibit increased depressive-like behaviour (Zarate et al. 2010). Kainate receptors activate postsynaptic receptors and

inhibit neurotransmission by regulating GABA release. The distribution of kainate receptors in the brain is limited and their effect on synaptic signalling and plasticity is less than AMPA receptors (Machado-Vieira et al. 2012). In preclinical studies, a GluR6 knockout mouse with no kainate receptor subunit genes showed increased risk-taking and aggressive behaviours as well as hyperactivity in response to amphetamine. These manic-like behaviours decreased after chronic lithium treatment (Shaltiel et al. 2008).

1.3.2. Metabotropic glutamate receptors

Metabotropic glutamate receptors are members of the most abundant receptor gene family in the human genome, the G-protein-coupled receptor (GPCR) superfamily. The mGluRs are classified into three groups based on their pharmacology, sequence homology, G-protein coupling and association with specific second-messenger systems. Group I includes mGluR1 and mGluR5, group II includes mGluR2 and mGluR3 and group III includes mGluR4, mGluR6, mGluR7 and mGluR8 (Conn and Pin 1997) (Table 1.1). In general, mGluRs contain a large extracellular N-terminal domain that contains the glutamate binding site, seven putative transmembrane-spanning domains separated by short intra- and extracellular loops, and an intracellular C-terminal domain that is variable in length and less conserved among the members of the same group. Sequence analysis revealed that the N-terminal extracellular domain of mGluRs is made up of a Venus Fly Trap (VFT) domain similar to the periplasmic bacterial leucine/isoleucine/valine binding protein. This VFT domain is linked to the first of the seven transmembrane (TM) domains via a cysteine-rich domain (CRD) containing 9 highly conserved cysteine residues (O'Hara et al. 1993) (Figure 1.1). The structures of mGluR3 and mGluR7 VFT

domains confirm these findings (Muto et al. 2007a; Muto et al. 2007b). The VFT domain is composed of two lobes each made of α-helices around a large β-sheet with the glutamate binding site located in the cleft between the two lobes (Conn and Pin 1997). The heptahelical domains of mGluRs share very low homology compared to other GPCRs and it has been proposed that the second and the third intracellular loops, as well as helix 8 are important in G-protein coupling and are also involved in coupling sensitivity (Pin et al. 2003). The majority of characterized positive- and negative-allosteric modulators of mGluRs affecting glutamate activity bind within the heptahelical domain. The C-termini of mGluRs modulate G-protein coupling and are subject to alternative splicing, regulation by phosphorylation and modulatory protein-protein interactions (Niswender and Conn 2010).

Earlier studies have predicted that the mGluRs only form homomers. However, energy transfer technologies along with systems to control the receptor subunit composition have recently shown that mGluRs form heterodimers as well (Nicoletti et al. 2011). A recent report by Doumazane et al. (2010) demonstrated that subtypes of the same group of mGluRs can form intragroup heterodimers (e.g., mGluR1 with mGluR5 and mGluR2 with mGluR3 receptors). Interestingly, group II and group III mGluRs can form intergroup heterodimers (mGluR2 with mGluR4). No heterodimers can be formed between group I and Group II/III mGluRs suggesting that only mGluR subtypes coupled to the same G-protein can form heterodimers. The existence of mGluR heterodimers in the brain is evident by the co-localization of different mGluR subunits in specific subdomains in neurons (Doumazane et al. 2010).


Third party copyright material removed

1.3.2.1. Group I mGluRs

Group I mGluRs are positively linked to phospholipase C and direct activation of mGlu1/5 receptors results in increased phosphoinositide turnover. The majority of localization studies of Group I mGluRs have revealed the postsynaptic terminal localization of mGLuR1 and mGluR5 (Table 1.1). The activation of these receptors often leads to cell depolarization and increase in neuronal excitability (Niswender and Conn 2010). The group I mGluRs can also act presynaptically to either increase or decrease neurotransmitter release mediated by postsynaptic group I mGluRs and release of retrograde messengers, such as endocannabinoids. Group I mGluRs play important roles in induction of long-lasting forms of synaptic plasticity, including long-term depression (LTD) and long-term potentiation (LTP) of transmission at multiple glutamatergic synapses and induce long-lasting changes in neuronal excitability (Niswender and Conn 2010).

The activation of group I mGluRs leads to excitatory effects in the brain, such as activation of NMDA receptors or an increase in glutamate release. mGluR1 is abundantly present in the cerebellum, the CA3 region of hippocampus, the thalamus, the dentate gyrus and the medial central gray whereas mGluR5 is highly expressed in the CA1 and CA3 regions of the hippocampus, in the septum, the basal ganglia, the amygdala and the nucleus accumbens. As noted above, the abundance of group I mGluRs in the limbic structures that are involved in the processes of emotions and motivation suggests a role for these receptors in affective disorders (Palucha et al. 2007). Genetic association studies on group I mGluRs have implicated mGluR5 (GRM5) in schizophrenia (Devon et al. 2001) but failed to show any association for

mGluR1 (GRM1) in bipolar disorder (Kaneva et al. 2009). Recently, GRM5 has been associated with ADHD, with an intronic SNP showing second best association in a German sample GWAS (Hinney et al. 2011) and the presence of deletions in a genome-wide CNV study (Elia et al. 2011). Reduction in the level of mGluR5 protein in the prefrontal cortex of patients with MDD was demonstrated in a study using postmortem brain samples. Several findings have suggested a pathophysiological role for mGluR5 in MDD as well as a role in the action of antidepressants (Hashimoto 2011).

The first mGluR to be knocked out in mice was mGluR1 and these animals show normal gross brain morphology but exhibit several phenotypes. In a recent study recording functional changes taking place at the hippocampal CA3-CA1 synapse during the acquisition of an associative task in mGluR1 knockout mice, the animals were unable to learn the task and showed significant impairments in the ability to induce LTP (Gil-Sanz et al. 2008). Deficits in prepulse inhibition, a measure of sensorimotor gating that is impaired in schizophrenic patients and reversed by antipsychotics, have been reported in both mGluR1 and mGluR5 knockout animals (Brody et al. 2003; Brody et al. 2004). Also, in models of drug abuse and addiction, mGluR5 knockout mice failed to self-administer cocaine or increased their locomotor activity after cocaine administration (Chiamulera et al. 2001). Group I mGluR antagonists have shown potential therapeutic effects in mood disorders. The group I mGluR5 antagonists MPEP (2-methyl-6-[phenylethynyl]-pyridine) and MTEP [(2-methyl-1, 3-thiazol-4-yl)ethynyl]pyridine) showed antidepressant-like activity in animal models. Another mGluR5 antagonist fenobam was associated with

significant psychostimulant effects and also showed robust anxiolytic efficacy in a double-blind, placebo-controlled trial (Machado-Vieira et al. 2012).

1.3.2.2. Group II mGluRs

Group II mGluRs are negatively coupled to adenylyl cyclase and upon activation inhibit forskolin-stimulated cyclin AMP formation (Cartmell and Schoepp 2000). These receptors are largely present on presynaptic membranes and on glial cells where they are believed to modulate glutamatergic neurotransmission by sensing glutamate spillover and regulating transmitter release. These receptors are highly expressed in brain structures apparently related to emotional states, including the forebrain and limbic areas, such as the amygdala, hippocampus and prefrontal cortex

(Palucha and Pilc 2007). Thus, the group II receptors mGluR2/3 are extremely interesting targets for antidepressant drug development based on their role in modulating glutamate neurotransmission. Presynaptic group II mGluRs also regulate the release of other neurotransmitters such as GABA, dopamine, serotonin, purines and others (Figure 1.2) (Cartmell and Schoepp 2000). Postmortem brain studies have demonstrated an elevated level of mGluR2/3 protein in the prefrontal cortex of patients with MDD (Feyissa et al. 2010).

The majority of genetic association studies on mGluRs in psychiatric disorders revolve around the group II receptors, which include mGluR2 (GRM2) and mGluR3 (GRM3). Multiple GWAS have investigated the association of GRM2 with bipolar disorder but no meta-analysis has been carried out. Equal number of studies recording positive and negative associations of GRM3 with schizophrenia has been reported, but there has been no evidence of the involvement of GRM3 in bipolar disorder (Joo et al. 2001; Marti et al. 2002; Cherlyn et al. 2010).

Evidence from several studies suggests that mGluR2/3 -targeted drugs possess both antidepressant and anxiolytic properties (Sanacora et al. 2008). With the exception of N-acetylaspartateglutamate (NAAG) that selectively activates mGluR3, no orthosteric agonists or antagonists can differentiate between mGluR2 and mGluR3. Both mGluR2 and mGluR3 have been separately deleted from the mouse genome and these animals have contributed greatly to the understanding of the role of these receptors. Severe impairment in terms of LTD induced by low-frequency stimulation at synapse was observed in mGluR2 knockout mice, along with normal regulation of basal synaptic transmission and LTP at mossy fiber-CA3 synapses in the hippocampus (Niswender and Conn 2010). mGluR2 knockout mice also showed

increased responsiveness to cocaine suggesting an important role in drug addiction and abuse (Morishima et al. 2005) and this receptor has also been shown to selectively mediate the beneficial effects of group II agonists in rodent models of psychosis (Fell et al. 2008). In a study involving mixed cultures of cortical neurons and astrocytes from wild-type mGluR2 and mGluR3 knockout mice, the neuroprotective effect of LY379268, a mGluR2/3 agonist, was lost when mGluR3 was missing from the astrocytes in the culture, whereas activation of mGluR2 were found to be excitotoxic (Corti et al. 2007). Phase II clinical trials in the treatment of generalized anxiety disorder for another group II selective agonist, LY354740 were underway before discontinuation on the basis of findings of convulsions in preclinical studies (Dunayevich et al. 2008). MGS0039 is a selective group II antagonist that has been shown to exert antidepressant effects in animal models. Other effects consistent with antidepressant activity including increasing dopamine release in the nucleus accumbens and neurogenesis in the hippocampal dentate gyrus have been reported on treatment with MGS0039 (Nicoletti et al. 2011). Unfortunately, research into drugs with fast onset of antidepressant activity has not yielded any results. Interestingly, low doses of the mGLuR2/3 agonist, LY379268, shorten the temporal latency of classical antidepressants in reducing the expression of β1-adrenergic receptors in the hippocampus (Matrisciano et al. 2005; Matrisciano et al. 2007). Both positive and negative symptoms in schizophrenic patients were shown to be reduced in a phase II clinical study involving 21-day treatment with LY2140023 (oral prodrug of the mGluR2/3 agonist LY404039) and this reduction was comparable with that caused by the antipsychotic, olanzapine (Patil et al. 2007).

1.3.2.3. Group III mGluRs

Group III mGluRs are the largest and the least investigated family of the mGluRs, which are classified into four subtypes: mGluR4, mGluR6, mGluR7 and mGluR8. The group III mGluRs are similar to group II mGluRs in that they are negatively coupled to adenylyl cyclase and their activation inhibits stimulated cAMP formation. Unlike all these group III receptors which are expressed in several regions of the CNS, mGluR6 subtype is limited to the retina (Nakajima et al. 1993). Both neuronal and glial cells express group III mGluRs and their localization on asymmetrical and symmetrical synapses have been well documented. Due to their predominant localization on axon terminals these receptors control the synaptic availability of glutamate as well as GABA, dopamine and serotonin (Figure 1.2) (Hashimoto 2011). mGluR4 receptors are localized mainly in the cerebellum, and the receptor knockout mice exhibit disturbed motor coordination in the rotarod test suggesting that mGluR4 may be responsible for motor function (Palucha and Pilc 2007). The mGluR7 is highly expressed in the neocortical regions, cingulate and piriphorm cortices, CA1, CA3 and dentate gyrus regions of hippocampus, amygdala, hypothalamic and thalamic nuclei (Ohishi et al. 1995). The expression of mGluR7 is restricted to the presynaptic zones and the sites of synaptic vesicle fusion. The receptor is the main one responsible for the regulation of glutamate release under normal physiological conditions (Palucha and Pilc 2007). Knockout of mGluR7 results in animals having an epileptic phenotype and mGluR7-/- mice show abnormalities in learning tasks suggesting an important role mediating amygdala-dependent learning. In addition, these mice have demonstrated roles for mGluR7 in anxiety and depression (Niswender and Conn 2010). mGluR8 expression seems to be dominant in

presynaptic terminals in the olfactory bulbs, piriform cortex, hippocampus and cerebellum (Shigemoto et al. 1996; Shigemoto et al. 1997). The mGluR8 knockout mice also show enhanced anxiety and weight gain compared to control animals suggesting a role for mGluR8 activators in the treatment of anxiety (Linden et al. 2002; Duvoisin et al. 2005).

Only one study has reported positive association of mGluR4 (GRM4) with bipolar disorder and schizophrenia (Fallin et al. 2005). Genetic studies on GRM8 are limited with inconsistent positive association reports with schizophrenia (Bolonna et al. 2001; Takaki et al. 2004). Elia et al. (2011) investigated CNVs in ADHD and reported deletions in GRM8 that were found only in patients.

Several studies have investigated the relationship between GRM7 and neuropsychiatric diosrders. Linkage studies (Pulver et al. 1995) and meta-analysis (Lewis et al. 2003) have implicated the GRM7 region, 3p26.1-p25.1 in susceptibility to schizophrenia. Furthermore, independent studies have confirmed the association between GRM7 and schizophrenia (Bolonna et al. 2001; Ohtsuki et al. 2008; Ganda et al. 2009; Shibata et al. 2009). In a recent GWAS of personality traits in bipolar patients GRM7 was significantly associated with neuroticism and anxiety (Alliey-Rodriguez et al. 2011). Although not genome-wide significant, GRM7 was found to be nominally and consistently associated with bipolar disorder in several GWAS (WTCCC 2007; Baum et al. 2008; Ferreira et al. 2008; Sklar et al. 2008). A recent GWAS of response to methylphenidate in children with ADHD found suggestive evidence of association with a GRM7 SNP (Mick et al. 2008). CNV studies in schizophrenia, ADHD and bipolar disorder have also repeatedly implicated GRM7

(WTCCC 2007; Shibata et al. 2009; Zhang et al. 2009; Saus et al. 2010; Elia et al. 2011; McQuillin et al. 2011).

Genetic studies have demonstrated that mice carrying the $GRM7^{BALB/cJ}$ gene variant had higher abundance of GRM7 mRNA in the brain, and lower level of alcohol preference drinking when compared with mice carrying the $GRM7^{B6By}$ variant on the same genetic background (Vadasz et al. 2007). A study by Simonyi et al. (2000) revealed that chronic exposure of rats to alcohol caused a substantial decrease in the mGluR7 levels in the CA3 region of the hippocampus whereas another study by Gyetvai et al.(2011) showed that GRM7 knockout mice expressed increased alcohol consumption. These findings suggest that constitutively low GRM7 expression may contribute to the predisposition to excessive alcohol self-administration in these animals. mGluR7 has also been implicated in animal models of addiction (Li et al. 2010a).

It is widely hypothesised that the development of depression and anxiety disorders can be traced back to chronic exposure to stress and stressful life events (O'Connor et al. 2010). Additionally, chronic exposure to stress may induce long-term vulnerability to cognitive deficits in later life (Lupien et al. 2009) whereas acute activation of the stress systems during learning can facilitate memory consolidation (Sandi 2004). It is possible that a number of converging cellular targets, including mGluR7, influence the manifestation of plasticity in specific brain regions coincident with stress- and glucocorticoid-related cognitive and psychiatric alterations (Lupien et al. 2009). Experiments using mGluR7 agonist AMN082 (N, N_-Bis (diphenylmethyl)-1,2-ethanediamine) in rats showed that mGluR7 activation attenuated LTP in the amygdala. Another interesting observation is that extinction of

fear is due to the formation of new memories, and mGluR7 knockout animal studies suggest that the receptor plays an essential role in fear extinction (Callaerts-Vegh et al. 2006). RNA interference mediated mGluR7 knockdown led to impairment in the extinction of conditioned taste aversion learning (Fendt et al. 2008). Interestingly, AMN082 administration causes both impairment in the acquisition of fear and fear extinction, processes of active learning (Fendt et al. 2008). These contradictory results might be explained by differential expression of mGluR7 in brain regions, the amygdala being responsible for acquisition of fear while the prefrontal cortex for extinction learning (O'Connor et al. 2010).

Pharmacological studies of group III mGluRs are limited because of the lack of subtype specific and bioavailable pharmacological tools. Thus, the role of these receptors in anxiety and depression along with the potential anxiolytic and antidepressant-like effects of their ligands require more development. Antidepressant-like effects were observed for the group III mGluR agonist ACPT-I [(1S,3R,4S)-1-aminocyclo-pentane-1,3,4-tricarboxylic acid] in the forced-swim test and these effects were antagonized by group III mGluR antagonist CPPG [(RS)-alpha-cyclopropyl-4-phosphonophenyl glycine] (Hashimoto 2011). Administration of a systemically active, potent and specific mGluR7 agonist, AMN082, elevated plasma levels of corticosterone and adrenocorticotropic hormone (ACTH) in control animals and not in mGluR7 deficient mice (Mitsukawa et al. 2005). However, in the same study increased levels of glucocorticoid receptors and 5-HT_{1A} receptors in the hippocampus of mGluR7 deficient mice were reported suggesting increased negative feedback action of the HPA axis leading to suppression of HPA hyperactivity and to an antidepressant and anxiolytic phenotype (Mitsukawa et al. 2005). In another

similar study, AMN082 showed antidepressant-like effects without altering the behaviour of the mice suggesting that mGluR7 is involved in the antidepressant-like activity of AMN082 (Palucha et al. 2007). The first selective mGluR7 antagonist MMPIP (6-(4-Methoxyphenyl)-5-methyl-3-(4-pyridinyl)-isoxazolo[4,5-c]pyridin-4(5*H*)-one hydrochloride) was reported by Suzuki and colleagues but no *in vivo* studies of this drug have been carried out (Suzuki et al. 2007). Furthermore, evidence from clinical and preclinical studies have emphasized the role of group III presynaptic mGluRs in reversing motor deficits in Parkinson's Disease (PD) and dopamine-deficient mice (Feeley Kearney and Albin 2003; Lopez et al. 2007). In a study by Greco et al. (2009), AMN082 administration at efficacious doses reduced the duration of haloperidol-induced catalepsy in a mGLuR7 dependent manner in wild-type but not mGluR7 knockout mice. However, higher doses of AMN082 had no effect on the same models of PD (Greco et al. 2009).

mGluR8 agonist (*RS*)-PPG ((RS)-4-phosphonophenylglycine) also showed dose-dependent antidepressant-like effects in the forced-swim test in rats (Palucha et al. 2004).

In summary, our understanding of the mGluRs has grown exponentially within the past decade. An analysis of mGluR genes using published GWAS data is yet to be carried out. The magnitude and extent of the abnormalities associated with bipolar disorder in the glutamatergic system require further clarification. Preclinical studies on mGluR5 negative allosteric modulators (MPEP, MTEP), group II mGluR agonists (LY354740, LY544344) and group III mGluR ligands (AMN082) have been very encouraging regarding their antidepressant-like and anxiolytic effects. However, problems with efficacy and adverse effects need to be further investigated.

An improved understanding of the function, anatomy and localization of different glutamate receptors in the brain would help in developing subunit-selective agents necessary for improved drug treatments.

Chapter 2

Aims of the Thesis

The main aim of the thesis was to elucidate the genetic basis of bipolar disorder by employing two strategies. The positional approach was exploited for localizing the susceptibility gene/loci in the linkage region identified in the UCL bipolar family linkage study. In the following sections of the thesis the association of two mGluRs in the UCL1 bipolar GWAS was investigated further in a replication sample (UCL2) and the genes were resequenced for the identification of novel rare variants associated with bipolar disorder. The specific aims of the thesis were:

1. To fine map the 1p36 linkage region with increased marker density in order to identify the gene conferring susceptibility to bipolar affective disorder at this locus in the UCL1 bipolar research sample.

2. To replicate the association of the three most significantly associated GRM7 SNPs in the UCL1 bipolar GWAS, in the UCL2 replication sample.

3. To resequence the GRM7 gene in a sub-sample of bipolar cases selected for carrying a disease haplotype in the UCL1 GWAS data and matching controls for the identification of rare allelic variants in GRM7 associated with bipolar disorder.

4. To validate the CNVs found in the UCL1 bipolar GWAS using QRT-PCR.

5. To replicate the UCL1 Bipolar disorder GWAS association of GRM3 gene by genotyping the three most significantly associated SNPs in the UCL2 replication sample.

6. To identify rare variants associated with bipolar disorder in the GRM3 gene by resequencing the gene in a sub-sample of bipolar cases carrying a disease haplotype in the UCL1 GWAS data and matching controls.

7. To examine the possible role of any aetiological base pair changes found to be associated with bipolar disorder in the three genes investigated in this study using functional assays.

Chapter 3

Materials and Methods

3.1. General methods

3.1.1. Research subjects

3.1.1.1. UCL bipolar research sample

The UCL bipolar research sample consisted of 1099 affected bipolar research subjects and 1152 normal comparison subjects. The research subjects were sampled in two cohorts and all of them gave informed signed consent. The first cohort called UCL1 comprised of 506 BP I cases and 510 screened normal comparison subjects (Sklar et al. 2008; Sklar et al. 2011; Dedman et al. 2012). The second cohort (UCL2) comprised of 409 BP I (69%) cases, 184 BP II cases and 642 screened and unscreened (75%) normal research subjects (Dedman et al. 2012). The sample of 1152 normal comparison subjects consisted of 672 subjects screened for the absence of psychiatric disorders and with no first-degree relatives with a psychiatric disorder, and 480 randomly selected unrelated British normal subjects whose DNA was obtained from the European Collection of Animal Cell Cultures (ECACC, UK). The cases were recruited from National Health Service (NHS) psychiatric services and from volunteers who were members of a support organisation for sufferers of bipolar disorder living in South East England. The screened control subjects were recruited from London branches of the National Blood Service, from local NHS family doctor clinics and from university student volunteers. In order to reduce the likelihood of population stratification between the case and control samples, the bipolar disorder case and control recruitment was strictly based on an ancestry questionnaire. For selection into the study, the ancestry of both parents and all four grandparents must have been Irish, Welsh, Scottish or English. Jewish subjects were excluded. Subjects

were also included if one of the four grandparents was of European Union (EU) ancestry before the 2004 EU enlargement. All case and control subjects were interviewed using the Lifetime Version of the Schizophrenia and Affective Disorders Schedule (SADSL) (Spitzer 1977). This information was supplemented by material from case-notes. Diagnoses were assigned using the Research Diagnostic Criteria (RDC) (Spitzer et al. 1978). A second psychiatrist confirmed the diagnoses. All the bipolar research subjects were also rated with the 90-item OPCRIT checklist (McGuffin et al. 1991). The research was approved by the UK NHS Multicentre Research Ethics Committee (MREC) and National Research Ethics Service. The sample has been used for several previous allelic and haplotypic association studies as well as GWAS of bipolar disorder (Ferreira et al. 2008; Bass et al. 2009; McQuillin et al. 2009; Vine et al. 2009; Nyegaard et al. 2010; Curtis et al. 2011; Dedman et al. 2012; Kandaswamy et al. 2012). The use of microarrays with up to 500,000 SNP markers in the GWAS allowed for an analysis of the genetic stratification between UCL1 cases and controls, which showed that no correction for population stratification was necessary (Ferreira et al. 2008; Sklar et al. 2008). DNA was obtained from blood samples for the cases and controls in UCL1 and the remaining screened controls in UCL2, and from saliva samples for the cases in UCL2.

3.1.1.2. UCL schizophrenia research sample

The UCL schizophrenia research sample consisted of 617 affected research subjects recruited from the UK NHS psychiatric services. The subjects were selected based on a strict ancestry questionnaire used for selecting UCL bipolar cases described in section 3.1.1.1. Ethical approval was obtained from the UK NHS MREC and local

research ethics centre. All subjects signed an approved consent after reading an information sheet. All the schizophrenia cases were selected for having an ICD-10 (WHO 1993) diagnosis of schizophrenia recorded in the NHS medical case-notes. Research subjects with short-term drug-induced psychoses, learning disabilities, head injuries and other symptomatic psychoses were excluded. All the affected research subjects were interviewed by a psychiatrist using the Lifetime Version of the SADS-L (Spitzer 1977). The cases were also rated using the 90-item OPCRIT checklist (McGuffin et al. 1991) and the diagnoses were assigned using RDC (Spitzer et al. 1978). Blood was collected from the subjects and the DNA was extracted.

3.1.1.3. UCL alcohol dependence research sample

The alcohol dependence sample consisted of 976 research subjects sampled in two cohorts and all of them gave informed signed consent. The first collection of subjects was recruited from the Centre for Hepatology at the Royal Free Hampstead NHS Trust between 1997 and 2008. The second wave of samples were recruited from several community substance misuse services including the Bexley Substance Misuse Service, South London and Maudsley Alcoholism Service, East Herts Community Drug Action Team and the Max Glatt Unit. All the patients fulfilled the criteria for a diagnosis of alcohol dependence using DSM-IV (APA 1994) or ICD-10 (WHO 1993). None of the patients was currently misusing illegal drugs. The selection of the alcohol dependence subjects was also based on a strict ancestry questionnaire similar to the UCL bipolar research sample described in section 3.1.1.1.

3.1.2. DNA extraction

3.1.2.1. DNA extraction from blood samples

Genomic DNA was extracted from whole blood samples using a standard phenol/chloroform extraction method that ran over two days. Isolation of leukocytes followed by lysis to release the genomic material was performed on the first day of extraction. Following an overnight digestion of proteins with proteinase K, the genomic DNA was extracted on the subsequent day.

Day I

The blood samples stored at -80^{o}C were allowed to gently thaw in a water bath at 30-35^{o}C. This was done to prevent the cells from lysing due to shock and from the release of damaging enzymes such as DNase. The thawed blood was then stored on ice followed by a transfer to 50 ml falcon tubes and topped up to 50 ml with 1X lysis buffer. The tube was inverted several times to ensure mixing. The lysis buffer breaks up the whole red blood cells (erythrocytes) whilst leaving the DNA containing white blood cells (leukocytes) intact. The lysate was then spun in a centrifuge at 2000 *g* at 4^{o}C for 15 minutes to pellet out the white blood cells. The pellet was resuspended in 50 ml of 1X lysis buffer to remove any remaining red blood cells and again centrifuged at 2000 *g* for 15 minutes. The cleaned white blood cell pellet was resuspended in 10 ml of proteinase K buffer and 500 µl of 10% sodium dodecyl sulfate (SDS). The proteinase K buffer provides stable optimum conditions for the proteinase K enzyme while SDS breaks open the cells by disrupting the lipid membrane, thus releasing the cell's contents including the genomic DNA into the solution. 50 µl of proteinase K was added to the lysate, which aids in the efficient

removal of complex proteins by digesting them and also deactivates any damaging enzymes such as DNases and RNases. The lysate was then incubated in a water bath at 55^{o}C (just below the optimum temperature of proteinase K at 65^{o}C) on a shaking platform overnight. Meanwhile, 1 g of PVPP (Polyvinylpolypyrrolidine) was weighed out into a fresh falcon tube and 5ml TE was added to it. These PVPP-TE containing tubes were also incubated overnight in a water bath at 55^{o}C.

Day II

The cell lysate was transferred to the falcon tube containing PVPP dissolved in TE. To this mixture, 5 ml of buffered phenol along with 5ml of 1:24 Isoamyl alcohol: chloroform mix was added. The phenol-chloroform mixture dissolves the lipids and agglomerates the proteins. The mixture was then centrifuged at room temperature for 15 min at 2000 *g* that allows the mixture to separate out into an organic solvent (phenol-chloroform) layer at the bottom (containing lipids), a protein interphase, and a top aqueous layer containing the nucleic acid. On centrifugation, PVPP holds down the protein interphase along with the organic solvents allowing the maximum removal of the aqueous layer without contamination. This aqueous layer was carefully transferred into a fresh and labelled 50 ml falcon tube for DNA precipitation. The organic solvents were disposed off safely in accordance with UCL's health and safety disposal of hazardous solvents protocol. The DNA was precipitated from this aqueous layer by adding 1500 µl of 3M sodium acetate and 30 ml of absolute ethanol. The tube was inverted gently until the DNA precipitated out of the solution into a condensed white clump. The DNA clump was removed with a plastic rod and then washed in 70% ethanol to remove as much sodium acetate as possible to prevent interference with the Polymerase Chain Reaction (PCR). The

DNA clump was then transferred into a labelled 1.5 ml microcentrifuge tube containing 500 µl of low EDTA TE. The DNA samples were stored at room temperature for about a week to allow the DNA to dissolve into the solution before quantification.

3.1.2.2. DNA extraction from saliva samples

The Oragene DNA Self-Collection Kit (DNA Genotek) is a non-invasive method of collecting DNA samples from patients. This is an all-in-one system for the collection, preservation, transportation and purification of DNA from saliva. Donor patients give a saliva sample into the base of an Oragene sample collection kit. The lid of the container has a preservative which stabilizes the DNA, and allows the samples to be kept up to 5 years at room temperature. The preservative in the lid is immediately released when the lid is tightened after the saliva collection.

DNA from the saliva samples was extracted according to the manufacturer's instructions. The saliva sample container was incubated in a water bath at $50^{o}C$ overnight or for a minimum of 1 hour. This heat-treatment step maximizes DNA yield and ensures that nucleases are permanently inactivated. 160 µl of Oragene DNA Purifier was added to 4 ml of sample and the mixture was incubated on ice for 10 minutes. The samples were centrifuged at 13000 rpm for 10 minutes at room temperature. The clear supernatant was transferred to a fresh centrifuge tube and the pellet was discarded. Equal volume of room temperature absolute ethanol was added to the supernatant and the contents were mixed gently by inversion. The samples were mixed well until a DNA precipitate could be seen and the sample was left standing at room temperature for 10 minutes to allow the DNA to fully precipitate.

The samples were then spun in a centrifuge at 3000 *g* for 10 minutes at room temperature. The supernatant was discarded and the DNA pellet was washed using 1 ml of 70% ethanol and then left to dry. The DNA was rehydrated in 500µl of TE containing microcentrifuge tube and stored at room temperature for a week.

3.1.3. DNA quantification

All DNA samples were quantified using Picogreen (Molecular Probes) on an LC480 (Roche Diagnostics Ltd, UK). Picogreen is an ultra-sensitive fluorescent nucleic acid stain that specifically binds to double stranded DNA (dsDNA) and aids in the quantification of dsDNA. Dilutions of the DNA samples were prepared by adding 2 µl of the stock DNA to 78 µl of TE and used as a working stock. In a 384-well Roche PCR plate, 2 µl of the working stock of DNA sample was dispensed into wells containing 8 µl of TE. Alongside, known concentrations of DNA derived from calf thymus were dispensed in a series of wells in order to produce a standard curve. The known concentrations used to prepare the standard curve were 0, 1, 2, 5, 10, 15, 20 ng/µl. Fresh PicoGreen working solution was prepared everytime by diluting the PicoGreen reagent 1:200 in TE. Equal volume (10 µl) of PicoGreen working stock was added to each well containing known and unknown DNA samples and mixed well. The 384-well plate was then sealed and spun in a centrifuge at 1000 rpm for 1 minute and read on an LC480 to quantify the DNA. The fluorescence is directly proportional to the quantity of DNA present. The standard curve was then plotted and used to derive the concentrations of the unknown DNA samples. Based on the results, dilutions of 25 ng/µl were prepared for all DNA samples and stored under 5^{o}C, thereby forming the working stock. The remaining DNA was stored as stock at -20^{o}C.

3.1.4. Polymerase chain reaction

3.1.4.1. Primer design guidelines

1. The optimal length of the primers is 18-22 bp.

2. The forward primer sequence is complimentary to the reverse strand of DNA and the reverse primer to the leading strand.

3. The GC content of the primer sequence is kept within 40-60%.

4. The absence of stop codons (either TTA at the 3' end or TAG/TAA/TGA at the 5' end) is maintained.

5. The primer does not end with a T because the *Taq* DNA polymerase does not bind well and therefore is prone to mispriming.

6. The melting temperature (T_m) of the primer is the temperature at which one half of the DNA duplex dissociates to become single stranded, which indicates duplex stability. The optimal T_m is 55-65^oC. Ideally, the T_ms of the two primers is within 5^oC of each other. The T_m is calculated using the formula:

$$T_m = [(2^oC \times AT) + (4^oC \times CG)] - 2$$

7. GC clamp is the presence of G or C bases within the last five bases from the 3′ end of the primers and it helps to promote specific binding at the 3′ end due to stronger bonding of G and C bases. More than 3 G's or C's should be avoided in the last 5 bases at the 3′ end of the primer.

8. Absence of hairpins or dimerization in primers is required.

9. Repeats along with long runs of a single base are avoided.

Primers were designed using Primer 3 software 3 (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi). NetPrimer (Premier Biosoft) (http://www.premierbiosoft.com/servlet/com.pbi.crm.clientside.FreeToolLoginServlet#), a free primer analysis software that analyzes primers for amplification related properties such as Tm, GC %, secondary structures, 3′ end stability etc., was used to determine the quality of the primers. For difficult regions such as GC rich regions and regions with repeats, the primers were designed by using the general guidelines described above.

3.1.4.2. General PCR principle

A conventional PCR consists of three steps - DNA denaturation, primer annealing and polymerase extension. The denaturation step, usually performed at 95^{o}C for about 30 seconds to 5 minutes, renders all the dsDNA in the reaction mix to become single stranded. The optimal annealing temperature is about 5^{o}C lower than the primer-template T_m and the extension temperature is calculated based on one minute per kb of DNA to be amplified. Optimal primer sequences and primer concentrations are essential for the maximal specificity and efficiency of PCR. Limited primer concentration adversely affects the efficiency of PCR whereas too high primer concentration increases the chance of mispriming and the subsequent extension of misprimed molecules results in nonspecific PCR products. The optimal primer concentration for most PCR applications is between 0.1 and 0.5 µM. PCR buffers contain cations such as K^+ and NH_4^+ that neutralize the negatively charged phosphate

groups on the DNA backbone thereby, weakening the electrorepulsive forces between the DNA strands. This reduction of the repulsive forces facilitates the annealing process between the primer and the template and maintains a high ratio of specific to nonspecific primer-template binding over a wide temperature range. Addition of Magnesium ions (Mg^{2+}) to the reaction mix stabilizes primer annealing, influences enzyme activity and increases the T_m of dsDNA. Mg^{2+} also forms soluble complexes with (deoxynucleoside triphosphate) dNTPs in the reaction mix to produce the actual substrate that the polymerase recognizes. The optimal concentration for each dNTP to minimize polymerase error is 200 µM. Adjuvants such as dimethyl sulfoxide (DMSO), betaine, and glycerol are sometimes used to improve the amplification efficiency by melting the secondary structures and decreasing non-specific products.

3.1.4.3. General PCR optimisation

The optimisation of primers was performed using four common master mix conditions to optimise $MgCl_2$ concentration and the presence or absence of betaine (Table 3.1) on MWG-HT Primus 96 thermocycler. Three standard PCR programs, Standard 55^{o}C, Standard 60^{o}C and Touchdown, were used to amplify the PCR products. The annealing temperature and the number of cycles were adjusted in these protocols in order to obtain the best PCR condition to amplify the target region. The three cycling conditions used for PCR optimisation are described in Table 3.2.

Table 3.1 PCR primer optimisation layout

Component	Volume (µl)			
	2 mM $MgCl_2$ + Betaine	**2 mM $MgCl_2$ + No Betaine**	**2.5 mM $MgCl_2$ + Betaine**	**2.5 mM $MgCl_2$ + No Betaine**
10 X Buffer	2.5	2.5	2.5	2.5
Betaine 5M	5	0	5	0
50 mM $MgCl_2$	1	1	1.25	1.25
25 mM dNTP	0.2	0.2	0.2	0.2
F primer (10 pmol/µl)	1	1	1	1
R primer (10 pmol/µl)	1	1	1	1
Taq polymerase (1 U/µl)	0.5	0.5	0.5	0.5
DNA 25 ng/µl	2	2	2	2
Water	11.8	16.8	11.55	16.55
Total volume	25	25	25	25

Table 3.2 The PCR cycling conditions used for optimisation

Standard 55°C (STD55)	**Standard 60°C (STD60)**	**Touch Down (MHTD)**
Lid heated to 105 °C	Lid heated to 105 °C	Lid heated to 105 °C
Products denatured at 94 °C for 5 minutes	Products denatured at 94 °C for 5 minutes	Products denatured at 94 °C for 5 minutes
35 cycles of	35 cycles of	3 cycles of
94°C - 30 seconds	94°C - 30 seconds	94°C - 30 seconds
55°C - 30 seconds	60°C - 30 seconds	63°C - 30 seconds
72°C - 30 seconds	72°C - 30 seconds	72°C - 30 seconds
		3 cycles of
Hold at 72°C - 10 minutes	Hold at 72°C - 10 minutes	94°C - 30 seconds
Store at 4°C	Store at 4°C	60°C - 30 seconds
		72°C - 30 seconds
		3 cycles of
		94°C - 30 seconds
		57°C - 30 seconds
		72°C - 30 seconds
		3 cycles of
		94°C - 30 seconds
		54°C - 30 seconds
		72°C - 30 seconds
		3 cycles of
		94°C - 30 seconds
		51°C - 30 seconds
		72°C - 30 seconds
		3 cycles of
		94°C - 30 seconds
		48°C - 30 seconds
		72°C - 30 seconds
		Hold at 72°C - 10 minutes
		Store at 4°C

3.1.4.4. Detection of PCR products using agarose gel electrophoresis

During electrophoresis, the mobility of a molecule in an electric field is inversely proportional to molecular friction, which is the result of its molecular size and shape, and directly proportional to the voltage and charge of the molecule. For the gel, appropriate amount of agarose (1% w/v) was weighted into a conical flask and the required volume of 1X TBE was added. The solution was boiled in a microwave oven for one minute to dissolve the agarose completely and then cooled down to around 50^{o}C - 60^{o}C. A 10 mg/ml stock solution of the gel staining dye ethidium bromide was prepared and added to the gel at a final concentration of 0.5 µg/ml. The gel was then poured into a casting tray containing selected combs and allowed to set for 20-30 minutes. Upon solidification, the combs were removed and the gel was transferred to the electrophoresis gel tank and 1X TBE was added to fill the buffer tanks just enough to cover the gel. Biotaq red DNA polymerase contains a red dye that eliminates the need for loading buffer and provides sufficient density to the samples so that they can be loaded directly onto the agarose gel. Thus, 5 µl of the PCR products and Hyperladder IV or Hyperladder I (Bioline, UK), the molecular size marker, were pipetted into separate wells. The electophoresis apparatus was then connected toa power supply and run at 120 V for 30 minutes. After the electrophoresis, the gel was visualized under a UV transilluminator (UVP Gel-doc-it Imaging systems, UK) and the image was captured using a camera attached to the gel doc system.

3.1.4.5. Purification of general PCR procucts

The PCR products were purified using MicroCLEAN (Microzone Limited, UK). MicroCLEAN is a DNA clean-up reagent that purifies or concentrates dsDNA from reaction buffers, enzymes, primers, primer dimers or dNTPs that can interfere during the sequencing reaction.

MicroCLEAN protocol for tubes

An equal volume of microCLEAN was added to the DNA sample. The solution was mixed using a vortex and left at room temperature for 5 minutes. The DNA-microCLEAN solution in the tubes was spun in a microcentrifuge at 13000 rpm for 7 minutes. The supernatant was removed and the tubes were spun again briefly to remove all the liquid. The pellet was resuspended in 5 µl of PCR water (Sigma-Aldrich, UK) and allowed to rehydrate for 5 minutes at room temperature for further usage.

MicroCLEAN protocol for 96 well plates

An equal volume of microCLEAN was added to the DNA sample and the solution was mixed using a Mixmate (Eppendorf, Germany). The sample mix was spun in a centrifuge at 4000 *g* for 40 minutes. The supernatant was removed by placing the plates upside down on a tissue in the centrifuge holder and spinning at 1000 *g* for 30 seconds. The pellet was resuspended in 5 µl PCR water and left at room temperature to rehydrate.

3.1.5. DNA Sequencing

3.1.5.1. Big Dye sequencing reaction

Big Dye terminator sequencing, similar to the Sanger-Coulson sequencing, is a fluorescence-based cycle sequencing method. This method uses fluorescent dyes to label the extension products, where the components are combined in a reaction that is subjected to cycles of annealing, extension, and denaturation in a thermal cycler. Thermal cycling of the sequencing reactions creates and amplifies extension products that are terminated by one of the four dideoxynucleotides. The ratio of deoxynucleotides to dideoxynucleotides is optimized to produce a balanced population of long and short extension products. Because each dye emits a unique wavelength when excited by light, the fluorescent dye on the extension product identifies the 3′ terminal dideoxynucleotide as A, C, G or T (Figure 3.1).

The region of interest was amplified under optimum conditions using forward and reverse M13 tailed primers and the products were purified using microCLEAN (described in section 3.1.4.5) and resuspended in 5 µlof PCR water (Sigma-Aldrich, UK). Sequencing PCR was carried out using the resuspended amplified products and

the Big Dye Terminator v3.1 sequencing kit (Applied Biosystems, UK) according to manufacturer's instructions. Two reactions for each sample, one using M13 forward primer and one with M13 reverse primer were performed for each sample using the reaction mix (Table 3.3) and the cycling conditions described below.

Table 3.3 Reaction mix used for Big Dye sequencing reaction

Reagent	**Final concentration**	**Volume (µl)**
Terminator ready reaction mix	1X	1
BigDye Sequencing buffer 5X	1X	1.5
10µM M13 primer F/R	0.32 pmol	0.32
Template (0.5-1 kb PCR product)	(5-20ng)	1
Water		6.18
Total volume		10

The cycling conditions used for Big Dye sequencing were:

- Heat the lid to 105^{o}C
- Initial denaturation

  94^{o}C for 5 minutes
- 25 Cycles of

  96^{o}C - 10 seconds

  50^{o}C - 5 seconds

  60^{o}C - 4 minutes
- Store the reactions at 4^{o}C

3.1.5.2. Purification of extension products from sequencing PCR

The salts and unincorporated dye terminators were removed from the extension products prior to electrophoresis to obtain clean sequencing data. Purification of the extended products in 96-well plates was carried out using the ethanol/EDTA precipitation method. To each 10 µl reaction, 2.5µl of 125 mM EDTA and 30 µl of absolute ethanol were added. The reaction mix was mixed using a MixMate and left at room temperature for 10 minutes to allow the precipitation of the extension products. The reactions were spun in a centrifuge at 3000 *g* for 60 minutes. The supernatant was discarded by inverting the reaction plate on a paper towel and spinning at 1000 *g* for 1 minute. A 70% wash step was performed on the reaction plate by adding 30 µl of 70% ethanol to each pellet and spinnig at 3000 *g* for 10 minutes. The supernatant was again discarded as described above using a paper towel. The reaction plate was left at room temperature for 15 minutes to evaporate off any residual ethanol. The pellets were resuspended in 15 µl of Hi-Di™ formamide per well and sent off for sequencing at the Center of Comparative Genomics, Department of Biology, UCL and the NHS North East Thames Regional Genetics Laboratory, London. The samples were run on the 3730*xl* DNA Analyzer (Applied Biosystems, UK) (Figure 3.2).

3.1.5.3. Analysis of sequencing data

The sequencing data was analysed using the Staden Package (Staden 1996). The assembly program consists of Gap4 and Pregap4. Gap4 performs assembly, contig joining, assembly checking, repeat searching, experiment suggestion, read pair analysis and contig editing. Pregap4 provides a graphical user interface to set up the processing required to prepare trace data for assembly or analysis.

3.1.6 Genotyping of genetic markers

3.1.6.1. Genotyping of microsatellite markers

Primer design for amplification of microsatellites

Tailed primer approach was used for the amplification of the STRs or the microsatellites. One of the primers was synthesized with an M13 forward or reverse primer sequence on the 5' end. The PCR products were labeled during amplification by incorporating an IRDye-labeled M13 (700 or 800 nm) primer that was included in the PCR reaction. Li-Cor M13 forward and reverse primer sequences are as follows: Forward IRDye 800-labeled primer, 19-mer: 5'-CACGACGTTGTAAAACGAC-3′ Reverse IRDye 800-labeled primer, 20-mer: 5'-GGATAACAATTTCACACAGG-3′

Amplification of microsatellites by PCR

Genomic DNA was amplified by PCR under specific temperature and cycling conditions for each pair of oligonucleotide primers. The PCR reaction for each DNA sample was prepared by mixing all the components described in Table 3.4.

Table 3.4 Reaction mix for amplification of microsatellites

Component	Final concentration	Volume (µl)
10X buffer	1X	1
Betaine		2
50 mM $MgCl_2$	2.5 mM	0.5
25 mM dNTPs	0.25 mM	1
10 µM F primer	1 µM	1
10 µM R primer	10 µM	1
5 µM M13 primer (700/800)	0.5 µM	1
Taq 1 U/µl	0.2 U	0.2
DNA (25 ng/µl)	50 ng	2
dH_2O		0.3
Total Volume		10

The cycling conditions STD55 or MHTD were used for the amplification of the microsatellites as described previously in section 3.1.4.3. Where possible, markers were designed to be pooled together so that markers with differing allele size ranges could be run simultaneously on each wavelength. This strategy was not only cost-effective but also less time-consuming. PCR products were then prepared for electrophoresis.

Detection of amplified microsatellites using polyacrylamide gel electrophoresis (PAGE)

SequaGel XR system, used for preparing the polyacrylamide gel, consists of two components:

(i) SequaGel XR monomer solution containing urea as well as acrylamide and acrylamide derivatives in a propreitary ratio dissolved in deionized distilled water, and

(ii) SequaGel complete buffer solution containing 5X TBE and TEMED in deionized distilled water.

The two components were mixed together to form the SequaGel mix, which was used within one month.

Microsatellite allele sizes were resolved by PAGE on dual laser LiCor 4200L (LiCor, UK) sequencers that detect IRDye 700 and IRDye 800 at the same time without spectral overlap between detection channels. The IRDye infrared dye-labeled products separate according to their sizes on an acrylamide gel. A solid-state diode laser excites the infrared dye on the DNA as it migrates past the detector window. A focusing fluorescence microscope containing a solid-state silicon

avalanche photodiode scans back and forth across the width of the gel collecting data in real time.

Assembling the gel apparatus

The 25 cm glass plates and 2 mm spacers were thoroughly cleaned and rinsed with deionized water and ethanol to remove any grease spots. The plates were air dried and assembled with spacers in a gel caster. The gel solution was prepared by adding 200 µl of 10% APS to 25 ml SequaGel XR mix. This solution was stirred well and poured into the plates using a 50 ml syringe and a rectangular comb was gently inserted that later forms a trough. The gel was allowed to set for 30 minutes. Once the gel was set, the comb was removed, trough was cleaned and the sharktooth comb was inserted for loading the samples. The plates were again cleaned using deionized water and ethanol and placed in the LiCor 4200 sequencer. The upper and lower buffer tanks were filled with 1X TBE and the lids were placed. The connectors on the buffer tanks were fully inserted into their respective places and the gel was pre-run for 15 minutes. During the pre-electrophoresis run, Quick SequenceIR software automates the processes of focusing, autogain and setting the electrophoresis parameters.

Preparation and loading of the samples

The gel was pre-heated to 45^{o}C and run at a constant voltage of 1200V for 15 minutes to prevent the formation of secondary structure and allow the linear DNA to run in proportion to its size. An equal volume of 0.001 % fuchsin (w/v fuchsin dissolved in formaldehyde) (Sigma-Aldrich, UK) loading buffer was added to the samples and the samples were denatured at 95^{o}C for 3 minutes immediately before loading. Using an 8 channel Hamilton syringe, the denatured samples and an IRD-

labeled DNA size standard MicroSTEP20a (Microzone, UK) were loaded into the sample wells. The DNA size standard was loaded every 15 samples to assist the sizing of allelic fragments. The fragment sizes of the molecular weight standard were - 70, 90, 100, 120, 140, 160, 180, 200, 220, 240, 260, 280, 300, 310, 330, 340, 350, 360, 380 and 400 bp.

Running the gel

The gel was run at 1200V for approximately 2 hrs.

Genotyping data analysis

The fluorescent images were analysed using SAGA-GT software (LiCor, Cambridge, UK) (Figure 3.3). SAGA^{GT} is an advanced genotyping software that automates electrophoresis, lane finding, location of standards, calibration of band sizes, and allele scoring. Genotypes were read by two independent assessors blind to each other's genotypes and blind to diagnosis. The genotypes that could not be agreed upon were repeated. The genotyping data for each marker from SAGA^{GT} was finally input into Microsoft Access database and analysed for conflicts arising from individuals who have been genotyped more than once. After resolving the conflicts, the final data was used for association and haplotype analysis.

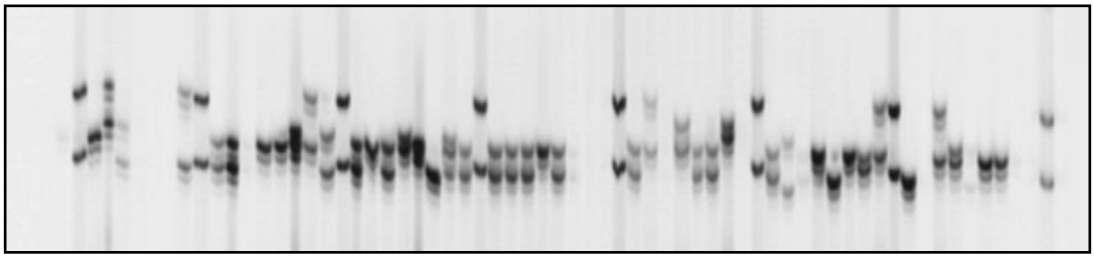


Figure 3.3 An example of a genotyping gel for a microsatellite marker

3.1.6.2. SNP genotyping

SNP genotyping for common SNPs was performed by allele-specific PCR at KBiosciences using KASPar reagents (Kbiosciences, Hoddesdon, UK). SNP genotyping for rare variants was performed in-house using KASPar and High Resolution Melting Curve analysis (HRM) techniques.

3.1.6.2.1. KASPar genotyping assay

Primer design for KASPar assay

The primers for SNPs were designed using the primer picker software provided by KBiosciences. A sequence containing the SNP along with 25 bases on each side was pasted into the sequence input window and analysed using the software. The results provided by the primer picker software consisted of two forward primers for the alternate alleles of the SNP and two reverse primers. The reverse primer was optimised to find the best reverse primer for the KASPar assay.

KASPar reaction

KASP (KBioscience Competitive Allele-Specific PCR) genotyping system is a homogeneous, fluorescent FRET based, endpoint-genotyping technology. The KASP genotyping system consists of two components:

1. *Assay mix:* The assay mix is comprised of three unlabelled oligonucleotides, two allele-specific forward primers (one for each allele of the SNP), each labelled with a unique unlabelled tail sequence at the 5' end and one common reverse primer. The composition of the allele mix is described in Table 3.5.

Table 3.5 Allele mix for KASPar assay

Component	Concentration in Assay Mix (µM)	Volume in Assay Mix (µl)
Allele specific primer F1 (100µM)	12	12
Allele specific primer F2 (100µM)	12	12
Common reverse primer R (100µM)	30	30
H_2O		46
Total		100

2. *Reaction mix (2X)*: The reaction mix contains *Taq* polymerase enzyme and the passive reference dye 5-carboxy-X-rhodamine, succinimidyl ester (ROX), two 5' fluor-labelled oligos, one labelled with FAM and the other with CAL Fluor Orange 560, two oligos with quenchers bound at the 3′ ends.

Principle

During the initial stage of PCR, the allele-specific forward primer binds to the complementary region directly upstream of the SNP, the 5' fluor-labelled oligos inturn interact with the sequences of the tails of the allele-specific primers whereas the oligos with quenchers bound at the 3′ ends are complementary to those of the fluor-labelled oligos, thereby, quenching the fluorescent signal until required. The common reverse primer also binds to the DNA. As PCR proceeds further, one of the fluor-labelled oligos corresponding to the amplified allele also gets incorporated into the template, and is hence no longer bound to its quencher-bound complement. As the fluor is no longer quenched, the appropriate fluorescent signal is generated and detected (Figure 3.4).

Method

The individual primer pairs were optimised as described in Table 3.6.

The optimum conditions thus derived were used to carry out the PCR reaction for cases and controls on 384-well plates containing dried DNA. Previously, case and control DNA (10 ng) was dispensed individually into 384-well plates using Epmotion 5075 (Eppendorf, UK) and dried in a hot air oven at 50^{o}C. For each KASPar assay, the mastermix was prepared in bulk and dispensed into the 384-well plates. The reactions were mixed using a Mixmate and spun in a centrifuge briefly before loading into LC480. The cycling conditions used for the KASPar assay are described in Table 3.7.

Table 3.6 KASPar optimisation layout

Component	Volume (µl)					
	1.8 mM $MgCl_2$	**2.2 mM $MgCl_2$**	**2.5 mM $MgCl_2$**	**2.8 mM $MgCl_2$**	**1.8 mM $MgCl_2$ + 5% DMSO**	**1.8 mM $MgCl_2$ + 10% DMSO**
DNA (10 ng/µl)	1	1	1	1	1	1
KASP Reaction mix (2x)	2	2	2	2	2	2
Assay mix	0.055	0.055	0.055	0.055	0.055	0.055
$MgCl_2$	0	0.032	0.56	0.8	0	0
H_2O	0.945	0.913	0.385	0.145	0.745	0.545
DMSO	0	0	0	0	0.2	0.4
Total volume	4	4	4	4	4	4

Table 3.7 KASPar cycling conditions - Dual colour hydrolysis probe/UPL probe program on LC480

Program name	Cycles	Analysis mode	Temp (°C)	Acquisition mode	Hold	Ramp rate (°C/s)	Acquisitions (per °C)	Sec Target (°C)	Step Size (°C)	Step delay (cycles)
Hot start activation	1	None	94	None	15 min					
1st amplification	20	Quantification	94	None	10 s	2.5		0	0	0
			57	Single	5 s	2.5		0	0	0
			72	None	10 s	4.8		0	0	0
2nd amplification	18	Quantification	94	None	10 s	2.5		0	0	0
			57	Single	20 s	2.5		0	0	0
			72	None	40 s	4.8		0	0	0
Reading	1	Melting	40	None	1 s	2.5				
		Curves	41	Continuous		0.06	5			
3rd amplification	3	Quantification	94	None	10 s	2.5				
			57	Single	20 s	2.5				
			72	None	40 s	4.8				
Reading 1	1	Melting	40	None	1 s	2.5				
		Curves	41	Continuous		0.06	5			
4th amplification	3	Quantification	94	None	10 s	2.5				
			57	Single	20 s	2.5				
			72	None	40 s	4.8				
Reading 2	1	Melting	40	None	1 s	2.5				
		Curves	41	Continuous		0.06	5			
Cooling	1	None	40	None	10 s	2.5		0	0	0

Endpoint genotyping analysis

The KASP assays were run on an LC480 (Roche Diagnostics Ltd, UK) and analyzed using the Endpoint genotyping LC480 software (Figure 3.5).

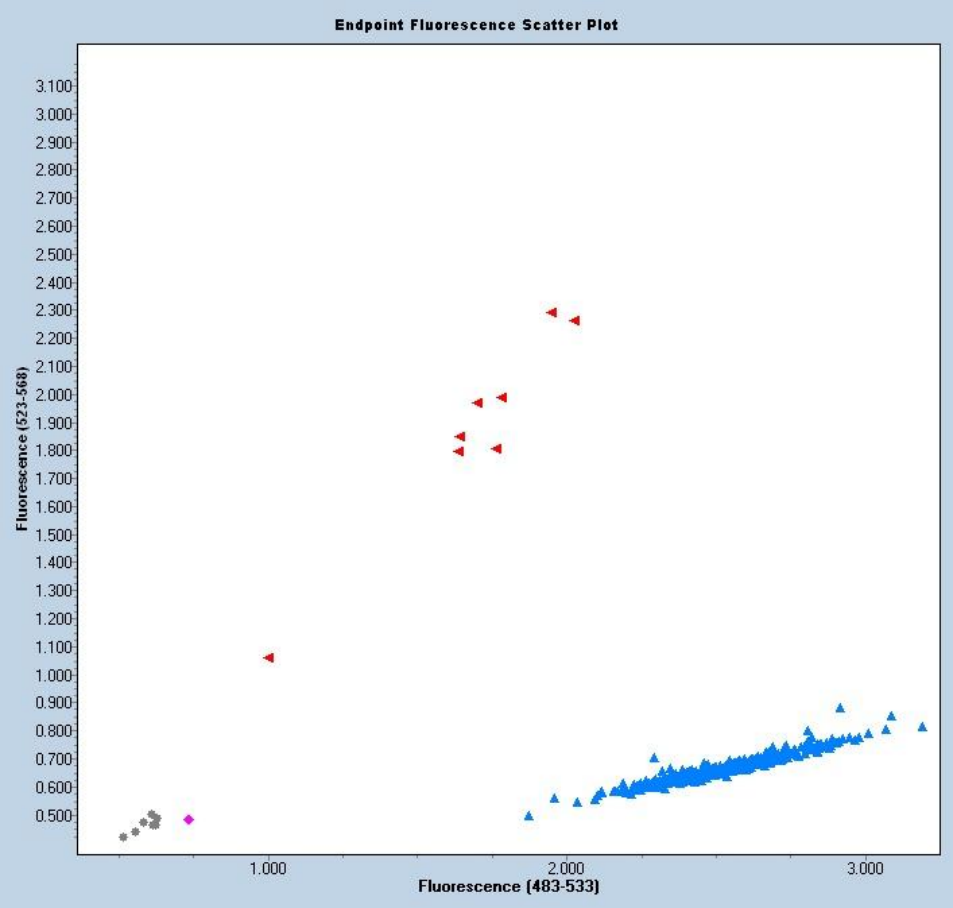


Figure 3.5 An example of the endpoint analysis performed after KASPar PCR reaction.

3.1.6.2.2. HRM genotyping assay

Primer design

Primers with melting temperatures of around 60^{o}C were designed according to the standard protocol described in section 3.1.4.1. The primers were designed to amplify around 100 -250 bp of the DNA sequence surrounding the SNP and then optimised using available HRM mastermixes in the laboratory according to the manufacturer's instructions.

HRM reaction

HRM analysis is a post-PCR analysis method used for fast genotyping and high-throughput mutation scanning of disease-related genes (Herrmann et al. 2006; Vossen et al. 2009; Wittwer 2009). This method involves PCR in the presence of saturating dsDNA binding dyes such as LCGreen, EvaGreen, SyBr Green etc., followed by a short melting step and subsequent analysis of the melting profile of the amplicons to detect sequence variations. Using optimum conditions and HRM mastermix, the reactions were performed on a 384-well plate containing 10 ng dried DNA. 5 µl of HRM mastermix was dispensed and the reactions were mixed using a Mixmate. The 384-well plate containing the reaction mixes was briefly spun in a centrifuge and loaded into the LC480 and the HRM program was run. The cycling conditions used for HRM are described in Table 3.8. During the high-resolution melting step following amplification, the LC480 records a large number of fluorescence data points per change in temperature with high precision. This change of fluorescence during an experiment can be used to measure temperature- induced DNA dissociation during high resolution melting. The melting profile of the amplicon depends on the GC content, sequence, length and heterozygosity. Therefore, any variations in the sequence will result in heteroduplexes with a different melt profile compared to the wild-type sequence.

HRM gene scanning analysis

The data was analysed using the LC480 Gene Scanning software (Roche Diagnostics Ltd, UK). The software analysed the different melting profiles of the amplicons by first normalising the data and then temperature shifting the curves such that the differences in the melting temperatures between both homozygous and heterozygous

could be noticed. In the end, the software groups the samples according to the three possible genotypes (Figure 3.6).

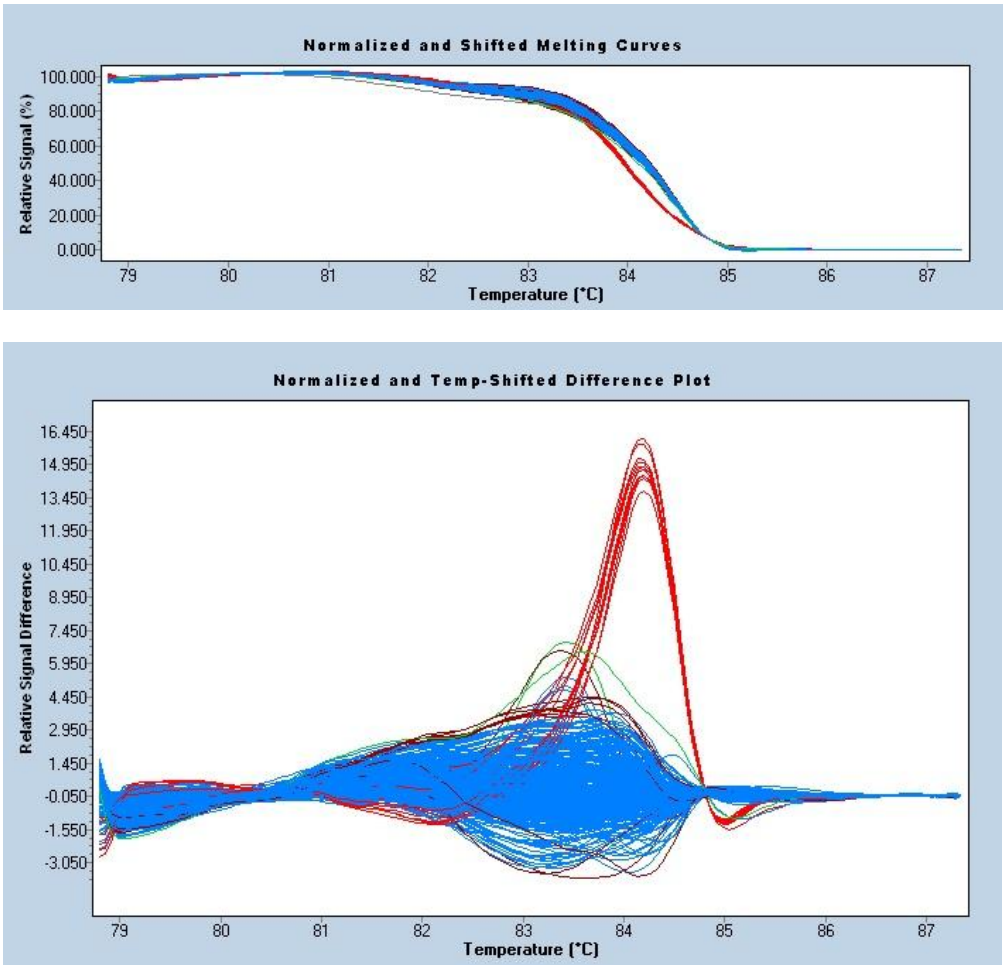


Figure 3.6 An example of HRM genotyping analysis

Table 3.8 HRM cycling conditions

Program name	Cycles	Analysis mode	Temp (°C)	Acquisition mode	Hold	Ramp rate (°C/s)	Acquisitions (per °C)	Sec Target (°C)	Step Size (°C)	Step delay (cycles)
Pre-Incubation	1	None	95	None	10 min		None	0	0	0
Amplification	45	Quantification	95	None	10 s	4.8		0	0	0
			65	None	15 s	2.5		53	0.5	1
			72	Single	10 s	4.8		0	0	0
High Resolution Melting	1	Melting Curves	95	None	1 min	4.8				
			40	None	1 min	2.5				
			65	None	1 s	1				
			95	Continuous		0.02	25			
Cooling	1	None	40	None	10 s	2.5		0	0	0

3.1.6.2.3. Bioinformatic analysis of novel SNPs

Bioinformatic analysis was carried out using the UCSC genome browser (http://genome.ucsc.edu/), PolyPhen-2 (genetics.bwh.harvard.edu/pph2/) (Adzhubei et al.) RESCUE-ESE (http://genes.mit.edu/burgelab/rescue-ese/) (Fairbrother et al. 2002), TESS (http://www.cbil.upenn.edu/cgi-bin/tess/tess) and UTRsite (http://utrsite.ba.itb.cnr.it/) (Grillo et al. 2010) to identify potentially novel functional SNPs found by resequencing genes. Bioinformatic analyses using polyphen 2 (Adzhubei et al. 2010) were used to predict the potential damage from coding mutations. mFold (http://mfold.rna.albany.edu/?q=mfold) and RNAfold (rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi) were used for predicting the RNA structure and RegRNA (regrna.mbc.nctu.edu.tw/) was used for predicting the miRNA binding sites.

3.1.7. GWAS data

The UCL1 research sample has been genotyped using Affymetrix GeneChip Human Mapping 500K Array in a collaborative GWAS (Sklar et al. 2008). A total of 506 BP I cases and 510 controls recruited as part of the UCL bipolar research sample were genotyped and the genotypes were called using the Bayesian Robust Linear Model with Mahalanobis distance classifier. Data analysis and quality control were performed using PLINK software package (http://pngu.mgh.harvard.edu/purcell/plink/).

3.1.8. Imputation using 1000 genomes data

To refine the association of the genes with bipolar disorder in the current study, I utilized the data from the European ancestry samples in the 1000 genomes

sequencing panel (www.1000genomes.org) and performed proxy association using PLINK. In this approach, flanking markers and haplotypes that are in strong linkage disequilibrium with a reference SNP are detected and tested for association with the disease within a haplotype-based framework. The data for SNPs in the exons and exon/intron boundaries of the genes was extracted from the European ancestry samples in the 1000 genomes database located at FTP directory /vol1/ftp/release/20110521 at ftp.1000genomes.ebi.ac.uk using a windows version of tabix. Tabix indexes a TAB-delimited genome position file '*in.tab.bgz'* from a position sorted and compressed input data file. Tabix was run with the following command that was modified according to the gene:

tabix -h

ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/release/20110521/ALL.chr.phase1_release_v3.20101123.snps_indels_svs.genotypes.vcf.gz chr:position from-to file name.tabix

The extracted data was pruned and the genotypes were correctly decoded and the codes were replaced with the alleles of the SNP. The ped and map files were created and merged with the SNP data from all other resources for the genes and the *proxyassocall* command on PLINK was performed.

3.1.9. Association and haplotypic analysis programs

3.1.9.1. Haplotypic and statistical analysis for microsatellite markers

Genotyping quality control was assured by genotyping 17% duplicates for cases and controls. A minimum call rate of 90% for the microsatellite markers and 95% for the SNPs was set as the standard. The final data was analysed to confirm Hardy-Weinberg equilibrium (HWE). The markers with lack of HWE were rejected and repeated. Genotypic and allelic associations for SNPs were tested using χ^2 tests. The microsatellite data was analyzed for allelic association with bipolar disorder using CLUMP which employs an empirical Monte Carlo test of significance and does not require correction for multiple alleles (Sham and Curtis 1995). Subtests of the CLUMP programme are:

- T1 - Pearson's 2 statistic of the raw contingency table
- T2 - the 2 statistic of a table with rare alleles grouped together to prevent small expected cell counts
- T3 - the largest of the 2 statistics of 2 × 2 tables each of which compares one allele against the rest grouped together
- T4 - the largest of the 2 statistics of all possible 2 × 2 tables comparing any combination of alleles against the rest.

Tests of haplotypic association were carried out using SCANASSOC and GENECOUNTING. A preliminary analysis of the haplotypes was done using

SCANASSOC. SCANASSOC is part of the GENECOUNTING program that computes the maximum likelihood estimates of haplotypic frequencies from phase-unknown data similar to GENECOUNTING. However, it does not perform permutation testing (Zhao et al. 2002; Curtis et al. 2006). This means that SCANASSOC gives a raw result but does not calculate the empirical significance of those results. GENECOUNTING, in addition, performs permutation testing, thereby, producing an empirical significance P-value (Zhao et al. 2002; Curtis et al. 2006). The analysis of haplotype block structure was performed using HAPLOVIEW (Barrett et al. 2005). GENECOUNTING was also used to calculate pairwise linkage disequilibrium between all markers and their relationship was viewed on LocusView 2.0 (Petryshen and Kirby, unpublished software; http://www.broad.mit.edu/mpg/locusview). COMBASSOC was used to combine all the markers to implicate association at the level of the whole gene rather than just for individual markers or haplotypes. This program combines the P values from all single loci and/or multilocus analyses of different markers according to the formula of Fisher, $X = \Sigma(-2\ln(pi))$, and then assesses the empirical significance of this statistic using permutation testing (Curtis et al. 2008).

3.1.9.2. Haplotypic and statistical analysis for SNP markers

PLINK and Haploview were used for the allelic association and haplotypic association analyses of SNP markers. PLINK is a free, open-source whole genome association analysis toolset developed by Shaun Purcell to perform basic, large-scale analyses in a computationally efficient manner

(http://pngu.mgh.harvard.edu/~purcell/plink/index.shtml). Association analysis, merging datasets, and meta-analyses were performed using PLINK (Purcell et al. 2007). Haploview (Barrett et al. 2005) was used for the haplotype analysis of SNPs. Haploview is a software designed to simplify and expedite the process of haplotype analysis by providing a common interface to several tasks relating to such analyses.

3.1.9.3. Genetic power calculation for replication analysis

In this thesis, UCL2 sample was used as the replication sample for testing association of the previously associated GWAS SNPs in UCL1 sample. Power calculation for the UCL2 sample performed using Genetic Power Calculator (http://pngu.mgh.harvard.edu/~purcell/gpc/cc2.html) (Purcell et al. 2003) for different combinations of odds ratios and minor allele frequencies are reported in appendix VI.

3.2. Methods for Chapter 4. Fine mapping of a new bipolar affective disorder locus on chromosome 1p36

3.2.1. Research subjects

Six hundred research subjects with bipolar disorder and 605 supernormal control subjects from the UCL1 cohort along with additional samples from the UCL2 cohort that were available during the study were used for the fine mapping of 1p36 region.

3.2.2. Selection of genetic markers for fine mapping of 1p36

Fine mapping of the 1p36 region was carried out in two stages in this study. In stage 1, a follow-up of the previously identified linkage region, 1p36 was carried out using increased marker density. Four microsatellite markers on chromosome 1p36, close to and including the marker D1S243, which had the maximum LOD score in the UCL bipolar disorder family linkage study were selected. Markers PRKCZ $(GT)_{17}$, PRKCZ$(CA)_{13}$ and PRKCZ$(AC)_{24}$ were identified from the UCSC Genome Browser database May 2004 assembly. In stage 2, single nucleotide polymorphisms (SNPs) in LD with PRKCZ were selected for genotyping from the HAPMAP database (Barrett et al. 2005; Barrett 2009) for further fine mapping of the region.

3.2.3. Genotyping of genetic markers in 1p36 region

The four microsatellite markers were genotyped using the standard protocol described previously in section 3.1.6.1. The primers for the markers PRKCZ $(GT)_{17}$,

PRKCZ$(CA)_{13}$ and PRKCZ$(AC)_{24}$ were designed using Primer 3. The sequences are listed in appendix III. The primer sequence for D1S243 was obtained from Ensembl. All four microsatellite markers were genotyped in 440 controls and 600 BPAD patients.

The SNPs were genotyped in two batches as this study was carried out during a period of sample collection. The first batch of six SNPs (rs908742, rs3107156, rs3753242, rs384726, rs12134873 and rs262688) was genotyped in 600 bipolar disorder patients and 440 controls and the second batch comprising SNPs rs3128296, rs2503706 and rs12132341 was genotyped in 600 bipolar disorder patients and 554 controls. The SNPs were genotyped using allele-specific PCR method at KBiosciences (Kbiosciences, Hoddesdon, UK) as described in section 3.1.6.2.1.

3.2.4. GWAS data

Additional data from eight SNPs genotyped in the UCL case-control sample as part of a GWAS (Sklar et al. 2008) in the region surrounding *PRKCZ* was included in the association analysis with bipolar disorder.

3.2.5. Gender-based association analysis

Gender based association analysis of all the microsatellite markers and SNPs was carried out using CLUMP and PLINK. For female-only and male-only analyses female-female case-control pairs and male-male case-control pairs were analysed respectively.

3.2.6. Association and haplotypic analysis

Tests of allelic association were carried out using CLUMP (Sham and Curtis 1995) and Haploview (Barrett et al. 2005) described in section 3.1.9. Haplotype association analysis was carried out using the program Haploview for SNPs and GENECOUNTING (Curtis et al. 2006) for combinations of SNPs and microsatellites.

3.3. Methods for Chapter 5. Mutation screening of the GRM7 glutamate receptor gene in Bipolar disorder

3.3.1. Research subjects

The UCL1 and UCL2 cohorts were used for the association study of GRM7 in bipolar disorder. These samples have been previously described in section 3.1.1.

3.3.2. Selection of subjects for resequencing GRM7

The most significantly associated GRM7 marker in the UCL1 GWAS was the SNP rs1508724 [P = 0.0028; OR (95% CI) = 1.33 (1.10 - 1.60)]. In the absence of any strongly associated haplotype, DNA samples from 32 bipolar research subjects homozygous or heterozygous for the rare allele of rs1508724 and also with an early age of onset of bipolar disorder were selected for sequencing. On the basis of previous studies by different research groups, I hypothesised that cases selected with early age-at-onset would provide a more homogeneous sample for the detection of rare variants. Also, 32 random control subjects were selected for the sequencing of GRM7.

3.3.3. Resequencing GRM7

Sequencing was carried out on 15 exons included in all the GRM7 isoforms (UCSC, March 2006 assembly), exon/intron boundaries, ~3 kb of the promoter region of the isoforms ENST00000486284(main isoform) and ENST00000463676, and 1 kb promoter region of the isoform ENST00000458641, 5'-UTR and 3′-UTR of the GRM7 gene. The sequence of primers used for sequencing is listed in appendix II.

Sequencing was performed and the data was analysed using the standard protocol described in sections 3.1.4 and 3.1.5.

3.3.4. SNP genotyping

The three most significantly associated GRM7 markers in the UCL1 GWAS dataset were genotyped in the UCL2 bipolar sample at Kbiosciences (Kbiosciences, Hoddesdon, UK). Novel rare varaints and potentially aetiological SNPs based on bioinformatic analyses described in section 3.1.6.2.3 were genotyped in the complete UCL sample. HRM and KASPar assays were used for genotyping SNPs as described in section 3.1.6.2. The primers used for the genotyping assays are listed in appendix III.

3.3.5. Association and haplotype analysis

Genotype data for the GRM7 SNPs was extracted from the UCL1 GWAS dataset and used in the association analysis. Genotypic and allelic associations along with haplotypic tests of association for SNPs were tested as described in section 3.1.9. Imputation of GRM7 SNPs from European sub-sample of the 1000 genomes dataset was performed using protocol described in section 3.1.8.

3.3.6. CNV validation using QRT-PCR

CNVs detected in the UCL1 samples were validated using QRT-PCR. TaqMan® RNase P Copy Number Reference (CNR) Assay, Human (Applied Biosystems, UK) was used for the CNV validation. In the duplex real-time PCR experiment, a UPL primer-probe (Universal Probe Library, Roche diagnostics Ltd, UK) combination

was used for the detection of the target gene (GRM7) and the TaqMan® RNase P CNR assay was used for the detection of a reference gene (RNase P). The reference gene, Ribonuclease P RNA component H1 (H1RNA) gene (RPPH1) located on 14q11.2, is known to be present in two copies in a diploid genome. The CNR assay was comprised of two components:

1. two unlabeled primers for amplifying the reference sequence and
2. one TaqMan TAMRA™ probe including the VIC® reporter dye attached to the 5' end and a TAMRA™ quencher attached to the 3′ end for detecting the reference sequence.

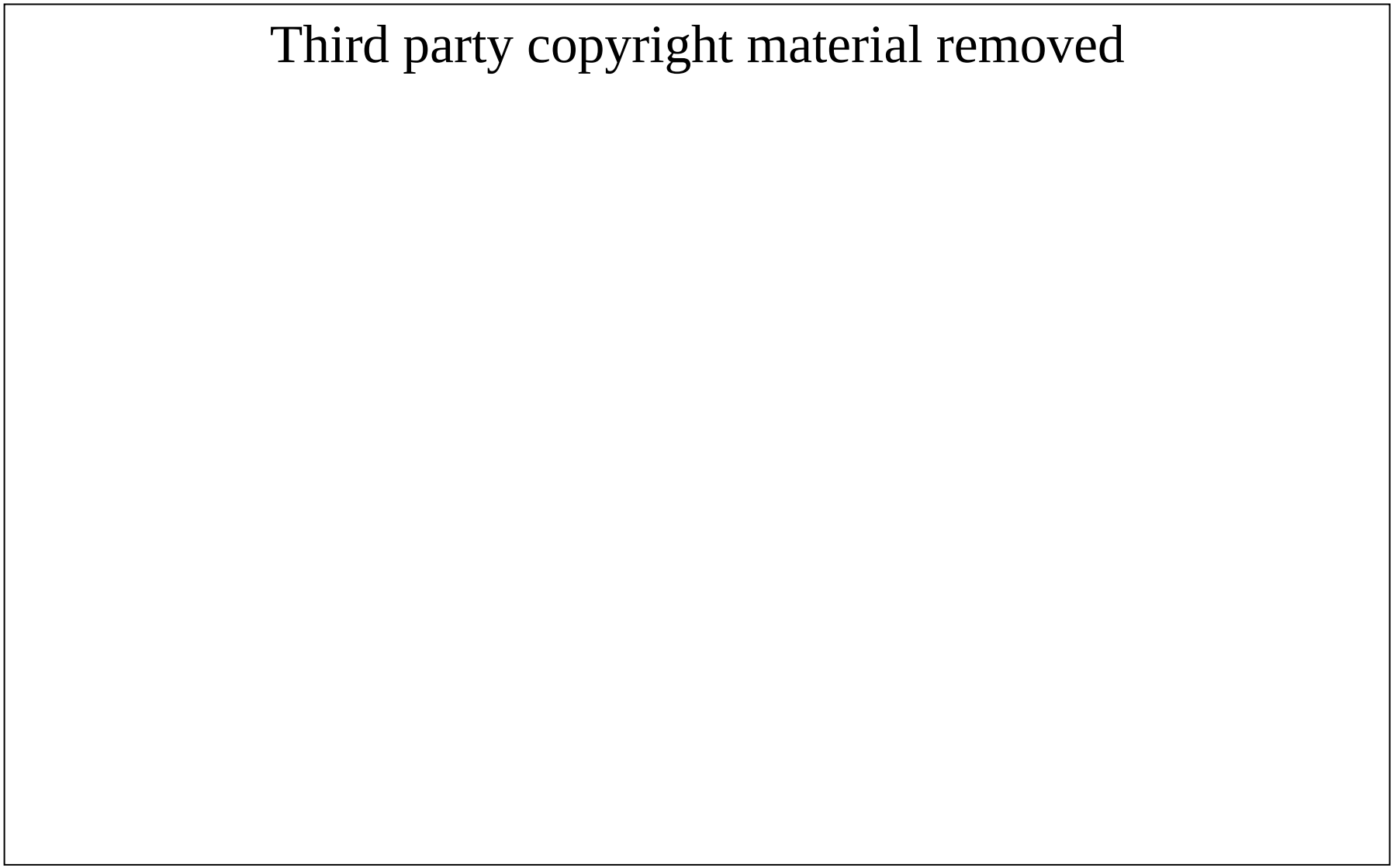
The primer-probe pair for the amplification and detection of GRM7 was designed using the UPL assay design service (https://www.roche-applied-science.com) (Appendix IV). The UPL probe is a hydrolysis probe formed of 8-9 nucleotides and labelled with fluorescein (FAM) at the 5' end and a quencher at the 3′ end. Incorporation of Locked Nucleic Acids (LNA), DNA nucleotide analogues with increased binding strengths compared to standard DNA nucleotides, into the UPL probes helped maintain the specificity and T_m required by the qPCR probes. Each transcript in the human transcriptome is detected by approximately 16 different probes. The UPL assay design program designs an optimal set of PCR primers and a probe for the region of interest. For the QRT-PCR reaction, LightCycler® probes master reaction mix containing FastStart *Taq* DNA polymerase for hot start PCR and optimal $MgCl_2$ concentration was used. During PCR, the genomic DNA template was denatured and the assay primers annealed to the target sequence. The probes annealed specifically to the complementary sequence between forward and reverse primer binding sites and when intact, the reporter dye signal was quenched due to its

proximity to the quencher dye. With each round of PCR, the target and the reference sequences were amplified by the FastStart *Taq* DNA polymerase resulting in the cleavage of probes hybridized to each amplicon. As a result, the quencher was separated from the reporter dye, thereby, increasing the fluorescence of the reporter. This increase in the fluorescence of each reporter dye at each PCR cycle was monitored and this was proportional to the accumulated PCR products (Figure 3.7). For each reaction in a 384-well plate, 10 ng of template DNA, 0.2 µl each of 10 µM forward and reverse primers and 10 µM probe, 0.5 µl of RNase P CNR assay mix and PCR water were added to a total volume of 10 µl. The reaction contents were mixed using a MixMate and spun briefly in the centrifuge at 1000 *g*. The plate was loaded into the LC480 and the UPL Dual hydrolysis probe reaction was run. The cycling conditions used are described in Table 3.9.

Table 3.9 The cycling conditions for the CNV validation assay

Stage	**Temperature**	**Time**
Hold	95^0C	10 min
Cycle (40 Cycles)	95^0C	15 sec
	60^0C	60 sec

Data analysis was performed using the LightCycler relative quantification method. The relative quantification method measures the relative change in mRNA levels of the target and reference genes. The threshold cycle (C_t) or the crossing point is the cycle at which there is a significant detectable increase in fluorescence and is a function of the amount of starting template.


Third party copyright material removed

In the $\Delta\Delta C_t$ method of QRT-PCR data analysis, the difference between the C_t values (ΔC_t) of the target gene and the housekeeping gene is calculated for each experimental sample and then the difference in the ΔC_t values between the experimental and control samples ($\Delta\Delta C_t$) is calculated. The fold-change in expression of the target gene between the two samples is then equal to $2^{-\Delta\Delta Ct}$.

$$\Delta\Delta C_t = (C_t^{target} - C_t^{reference})_{time\ x} - (C_t^{target} - C_t^{reference})_{time0}$$

where, $time_x$ is any time point and $time_0$ represents the 1X expression of the target gene normalized to the reference gene (Livak and Schmittgen 2001)

3.4. Methods for Chapter 6. Mutation screening of the GRM3 glutamate receptor gene in bipolar disorder

3.4.1. Genetic association of GRM3 with bipolar disorder

3.4.1.1. Research subjects

The combined sample of UCL1 and UCL2 cohorts was used for the association study of GRM3 in bipolar disorder. Schizophrenia and alcohol dependence syndrome samples were also used for association analysis of the 5′-UTR variant rs148754219. All these samples have been described in section 3.1.1.

3.4.1.2. Selection of subjects for resequencing GRM3

The G allele of rs2237563 was the most strongly associated SNP in the UCL1 sample ($P = 3.85 \times 10^{-5}$) (Ferreira et al. 2008; Sklar et al. 2008). In the absence of any strongly associated haplotype, DNA from 32 bipolar research volunteers were selected for being homozygous or heterozygous for the G allele of rs2237563 and 32 random normal comparison subjects were also selected for sequencing.

3.4.1.3. Resequencing GRM3

Sequencing was carried out on the promoter region 1000 bp upstream of the transcriptional start site, 5' untranslated region (UTR), the exons and intron/exon

junctions and the entire 3′-UTR of the GRM3 isoform 1 (NM_000840) which contains all 6 exons. The primer sequences used for sequencing GRM3 are listed in appendix II. The PCR products were amplified using the standard protocol and then purified using MicroCLEAN described in section 3.1.4. Sequencing was done using the Big Dye terminator v3.1 Cycle Sequencing kit (Applied Biosystems, UK) on an ABI 3730*xl* DNA Analyzer using the standard protocol described in section 3.1.5.

3.4.1.4. SNP genotyping

SNP genotyping for the three most significantly associated GRM3 SNPs, rs2237563, rs274621, and rs2158786, in the UCL1 bipolar disorder GWAS was performed in the UCL2 samples by allele-specific PCR at KBiosciences (Kbiosciences, UK). Novel rare varaints and potentially aetiological SNPs based on bioinformatic analyses described in section 3.1.6.2.3 were genotyped in the complete UCL sample using KASPar assay.

3.4.1.5. Association and haplotype analysis

Genotype data for the GRM3 SNPs was extracted from the UCL1 GWAS dataset and used in the association analysis. Genotypic and allelic associations along with haplotypic tests of association for SNPs were tested as described in section 3.1.9. Imputation of GRM7 SNPs from European sub-sample of the 1000 genomes dataset was performed using protocol described in section 3.1.8.

Genotypic and allelic associations of SNPs were tested using χ^2 tests and PLINK. Haplotype tests of association were performed using Haploview (Barrett et al. 2005). The protocol used for these analyses is described in section 3.1.9.

3.4.2. Functional characterization of GRM3 Kozak sequence variant

3.4.2.1. Animal cell-culture

Cell-lines

Two human cell-lines were used in this study, HEK293 and SH-SY5Y. HEK293 cells were a kind gift from Dr Josef Kittler and SH-SY5Y cells were obtained from European Collection of Animal Cell Cultures (ECACC, UK). HEK293, first described in 1977, is a cell-line generated by the transformation of human embryonic kidney cells with sheared adenovirus 5 DNA (Graham et al. 1977). Interestingly, speculations by scientists including Van der Eb predict that the cells may be neuronal in origin because of the presence of the mRNA and other proteins normally expressed in neuronal lineage cells (Graham et al. 1977). HEK293 cells do not adhere to the substrate at room temperature but re-attach to the flask over a period of several days in culture at 37^{o}C. SH-SY5Y cell-line is a thrice-cloned neuroblastoma sub-line of bone marrow biopsy-derived SK-N-SH and the cells grow as a mixture of floating and adherent cells. These cells grow as clusters of neuroblastic cells with multiple, short, fine cell processes (neurites) and have a 48 hr doubling time.

Subculturing the cells

HEK293 and SH-SY5Y cells were subcultured in 75cm^2 T-flasks (Nunc, UK) after reaching 70 - 80% confluency. The growth medium was removed and the cells were washed with 2ml of 1X PBS. To detach the adherent cells from the flask surface, 1 to 2 ml of trypsin-EDTA solution was added and the cells were incubated at 37^{o}C with 5% CO_2 for 1-2 minutes. The sides of the flask were gently tapped to dislodge the cells and the detached cells were observed under a microscope to confirm complete dislodging. Immediately, 5 ml of fresh growth medium was added to the flask to stop the trypsinisation reaction and the cell suspension was aspirated into a fresh 15 ml falcon tube. The cell suspension was spun in a centrifuge at 1500 rpm for 5 minutes and the pellet was resuspended in 5 ml of fresh growth medium. Appropriate aliquots of this cell suspension were dispensed into fresh culture flasks containing 10 ml final volume of complete growth medium and incubated at 37^{o}C with 5% CO_2. The cells were weekly subcultured in 1:5 to 1:10 ratio and the medium was changed every 3-5 days.

Counting of cells using haemocytometer

A haemocytometer is a specialised microscope slide used for estimating the number of cells in a measured volume under a microscope (usually blood cells). An improved Neubauer haemocytometer is a thick glass slide with two counting chambers and the main divisions separate each chamber into nine large squares (Figure 3.8).

Each square has a surface area of 1 mm^2 and the depth of the chamber is 0.1 mm. With the coverslip in place, each square of the haemocytometer represents a total volume of 0.1 mm^3 or 10^{-4} cm^3 (1 cm^3 = 1 ml). The cell suspension used for counting should be dilute enough to avoid clumps of cells and mixed thoroughly to ensure uniform distribution. The cell concentration per ml was determined using the following calculations:

Cells per ml = the average count of cells per square x the dilution factor x 10^4

Cell counting

The surface of the haemocytometer and the coverslip was washed with 70% alcohol. The sides of the coverslip were moistened by breathing on the surface and aligning it

over the counting chamber using gentle downward pressure. The cell suspension was prepared using the standard protocol used for subculturing and 100 µl of the uniform suspension was dispensed into a sterile microcentrifuge tube. Equal volume of trypan blue solution (dilution factor of 2) was added to the suspension in the microcentrifuge tube and mixed gently by pipetting. Trypan blue is a vital dye that does not interact with the cells unless the membrane is damaged. Viable cells exclude the dye and remain colorless and cytoplasm of non-viable or dead cells take up the dye and turn blue. A thoroughly mixed aliquot (10-20 µl) of cell suspension was dispensed at the edge of the chamber allowing to fill with capillary action. The same process was followed for the second chamber and the haemocytometer was observed under the microscope using 20X magnification. Systematically, the cells in the four corner squares were counted excluding the cells touching the outer perimeter.

Viable Cell count = Average live cell count x dilution factor x 10^4

Non-viable cell count = Average dead cell count x dilution factor x 10^4

Percentage Viability = Viable cell count / Total cell count x 100

Cryogenic storage and recovery of cells

Cryopreservation of cells ensures continuous supply and back up of cells in case of contamination and to have a lower passage number stock of cells. It is good practice to cryopreserve cells when they are at their maximum growth rate to aid in good recovery during the thawing of cells. The freezing medium was prepared using DMSO (Sigma-Aldrich, UK), a cryoprotective agent that reduces the freezing point of the medium and also allows a slower cooling rate thereby, reducing the risk of ice crystal formation which can damage cells. The cells were harvested using the

standard protocol used for subculturing and suspended in 2.5 ml of Dulbecco's modified eagle medium (DMEM) supplemented with 20% foetal bovine serum (FBS) and 1% Penstrep. In a fresh falcon tube, 20% DMSO was added to 2.5ml of DMEM supplemented with 20% FBS and 1X Penstrep, and this medium was then added drop by drop to the tube containing cells. After mixing the culture to form a uniform suspension 1ml aliquots of this cell suspension were dispensed into 1.5 ml cryogenic vials (Nunc, UK) and stored in a freezer container, Nalgene®, Mr.Frosty (Sigma-Aldrich, UK) at -80^{o}C. The polycarbonate freezer container contains a high-density polyethylene tube holder and foam insert. 100% isopropyl alcohol was added to the container and the tubes containing cells were placed into the tube holder. The container provides the critical and repeatable, 1^{o}C/min cooling rate that is required for successful cryopreservation of cells. Mr. Frosty container was kept at -80^{o}C undisturbed for a minimum of 4 hrs and then transferred to liquid nitrogen containers for long-term storage. The cells can be stored at -80^{o}C for 6 months or in liquid nitrogen for 2 years.

Thawing process

The thawing of the cells from -80^{o}C was done quickly to maximise the recovery of cells and minimise the loss of cells due to the presence of DMSO in the freezing medium. The cryovials were removed from -80^{o}C freezer and immediately placed in a 37^{o}C water bath for less than a minute or until around 80% of the suspension had thawed. The contents of the cryovial were pipetted into a fresh 15 ml falcon tube containing 5 ml pre-warmed growth medium and mixed gently by pipetting. The cells were centrifuged at 1500 rpm for 3 minutes and the supernatant discarded. The pellet was resuspended in 1 ml of growth medium and the cells were seeded in T-75

flask containing 9 ml of complete growth medium and incubated at 37^{o}C with 5% CO_2.

3.4.2.2. Electromobility Shift Assay (EMSA)

Preparation of nuclear extract from SH-SY5Y cells

SH-SY5Y cells were grown in 75 cm^2 T-flasks to 70-80 % confluency in the complete growth medium. The growth medium was removed and the cells were washed with cold 1X phosphate buffered saline (PBS). Two millilitres of cytoplasmic lysis buffer was added to each flask and the cells were harvested into pre-chilled 1.5 ml eppendorf tubes using cell scrapers (Nunc, UK). The cells were lysed by adding 90 µl of 10 % NP-40 solution to each eppendorf tube containing 1 ml of the cells and vortexing vigorously for 10 seconds. The homogenate was then centrifuged at 2000 *g* for 10 minutes at 4^{o}C and the supernatant containing the cytoplasmic proteins was discarded. The nuclear pellets were resuspended in 32 µl nuclear lysis buffer and the nuclei were disrupted by three freeze-thaw cycles in dry ice with ethanol and a water bath at 37^{o}C. The cellular debris was removed by centrifugation at 13000 *g* for 10 minutes and the supernatant (nuclear extract) was stored in aliquots of 20 µl at -80^{o}C.

Protein quantification using Bradford assay

The protein concentration of the nuclear lysate was quantified using Bradford assay reagent (Sigma-Aldrich, UK) according to the manufacturer's instructions. The reagent contains Brilliant Blue G dye that forms a complex with the proteins in the solution causing a shift in the absorption maximum of the dye from 465 to 595 nm. The amount of absorption is thus, proportional to the protein present in the solution.

The Bradford reagent was brought to room temperature before performing the assay and bovine serum albumin (BSA) was used as the protein standard. Protein standards of 1 µg/ml, 2 µg/ml, 5 µg/ml, 10 µg/ml, and 20 µg/ml were prepared by diluting 10 mg/ml BSA in nuclear lysis buffer. One part of the protein sample was mixed with 30 parts of the Bradford reagent in a disposable cuvette and incubated at room temperature for 5 minutes. The absorbance of the colored solution was measured at 595 nm using a UV spectrophotometer (Pye Unicam PU8600 UV/VIS spectrophotometer, Philips) within 60 minutes of adding the Bradford reagent as the protein-dye complex is stable upto 60 minutes only. A standard curve was created to determine the protein concentrations of the nuclear lysate by plotting the net absorbance at 595 nm versus the protein standard concentrations (Figure 3.9).

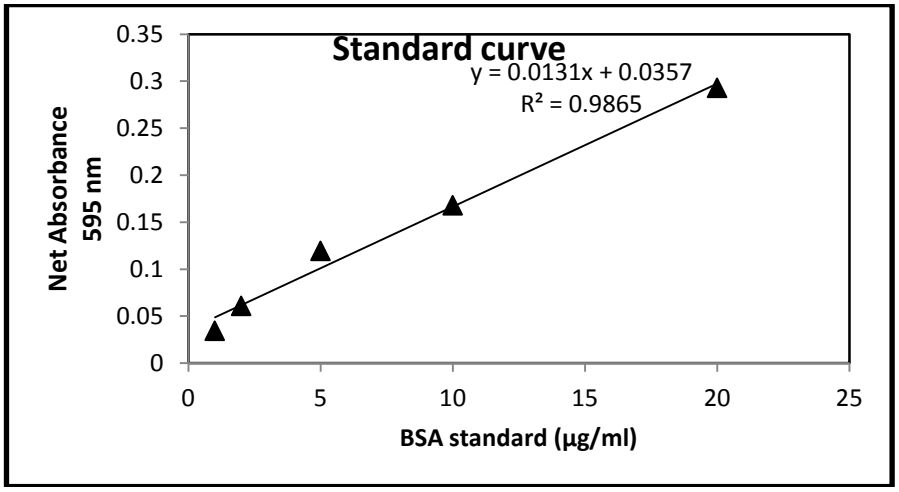


Figure 3.9 Standard curve for the quantification of protein concentration in the nuclear lysate using Bradford assay reagent

Preparation of 4% native polyacrylamide gel

In a 100 ml glass beaker, 5 ml of 40% polyacrylamide stock solution, 2.5 ml of 10X TBE, 80 µl of TEMED and 300 µl of 10% APS were added and MilliQ water was added to a final volume of 50 ml. The solution was mixed by swirling and poured

between 40 cm glass plates with 0.75 mm spacers and comb and allowed to set for 1 hr.

Preparation of oligos

Oligonucleotides with the wild-type (G) and the mutant (A) allele of rs148754219 were 5′ end-labelled with infrared dye DY-682 (Eurofins MWG, Ebersberg, Germany). Aliquots of 5 µl 20 pmol/µl of each pair of labelled oligonucleotides were denatured at 100^{o}C for 3 min and slowly cooled to room temperature to generate dsDNA.

Binding reaction

The binding reaction was performed in dark at room temperature by incubating 50 fmol of labelled double-stranded oligonucleotides with 10 µg of nuclear extract in 1X binding buffer to a final volume of 20 µl for 30 min. Poly (dI.dC) in the binding buffer prevents non-specific binding of proteins in the nuclear extract to the oligonucleotide probes thereby reducing background.

Electrophoresis

One microlitre of Orange dye (10X) was added to the binding reactions and mixed by pipetting. The samples were then resolved by electrophoresis on a 4% native polyacrylamide gel at 10 V/cm at 4^{o} C in 0.5 X TBE buffer. The DNA-protein complex was detected on a LiCor 4200L sequencer (LiCor, Lincoln, Nebraska, USA). A further follow-up experiment was carried out to compete out the band formed by the mutant allele and determine the binding specificity by incubating varying amounts of unlabelled competitor probe (10X, 20X, 50X, 100X, 200X) for the wild-type and mutant alleles in separate tubes. The principle behind the competition reaction is that when excess of unlabelled DNA is added to the binding

reaction, it competes with the labelled DNA for binding sites and eliminates labeled oligo binding causing the removal of the band shift, thereby, indicating that the binding reaction is specific. The competition reaction experiment was repeated three times.

3.4.2.3. Construction of reporter plasmids

3.4.2.3.1. Plasmid vectors

Vectors are carrier DNA molecules that can replicate autonomously or independent of host genome replication in host cells and facilitate the manipulation of a newly created recombinant DNA molecule. Other properties of vectors include containing a multiple cloning site formed of a number of unique restriction endonuclease cleavage sites and carrying a selectable marker to distinguish the host cells carrying the vectors from the host cells that do not contain a vector.

pJET1.2 cloning vector: The pJET1.2/blunt cloning vector (Figure 3.10) is part of the CloneJET^{TM} PCR Cloning kit that enables cloning of blunt and sticky ended DNA or PCR products generated with *Pfu* DNA polymerase, *Taq* DNA polymerase or any other thermostable DNA polymerases. This vector is a linearized vector that accepts inserts from 6 bp to 10 kb and the 5' ends of its cloning site contain phosphoryl groups making phosphorylation of the PCR primers unnecessary. The vector contains a lethal gene, which is disrupted only by successful ligation of a DNA insert into the cloning site due to which only recombinant plasmids are able to propagate. This cloning vector eliminates the need for blue/white screening as recircularized vector cannot propogate due to the expression of the lethal restriction enzyme.

pGL3 Luciferase reporter vectors: The pGL3 luciferase reporter vectors were used in this study to quantitatively analyse the functional effects of a polymorphism in GRM3 in regulating mammalian gene expression. The backbone of the pGL3 luciferase reporter vectors is designed for increased expression compared to pGL2 vectors and contains a modified coding region for firefly (*Photinus pyralis*) luciferase that has been optimized for monitoring transcriptional activity in transfected eukaryotic cells. The pGL3 luciferase reporter vectors used in this study are - pGL3-basic and pGL3-promoter.

pGL3-basic vector: pGl3-basic vector lacks the eukaryotic promoter and enhancer sequences thereby, allowing maximum flexibility in cloning putative regulatory sequences (Figure 3.11). The functional effect of a promoter polymorphism can be determined by observing the expression of luciferase activity in cells transfected with the wild-type and mutant plasmid constructs.

pGL3-promoter vector: The pGL3-promoter vector is similar to the pGL3-basic vector in all respects but additionally contains an SV-40 promoter upstream of the luciferase gene (Figure 3.12). DNA fragments containing putative enhancer elements can be inserted upstream or downstream of the promoter-*luc*+ transcriptional unit and the effect of polymorphisms in these regions can be explored.

3.4.2.3.2 Amplification of PCR products for cloning into reporter plasmids

The rs148754219 variant is located in the first exon of the gene, which is transcribed but not translated in the main isoform of GRM3. In the study by Corti et al. (2001) there is a suggestion of a silencer element being present in exon 1 of the GRM3 gene. Furthermore, rs148754219 is at -2 position from the ATG start codon of the GRM3 isoform transcript (RefSeq accession no. NM_000840) forming part of the Kozak sequence for that particular transcript. Therefore, five different plasmid constructs were constructed to explore the possibility of the effect of this SNP on the transcriptional and post-transcriptional/translational regulation of the gene.

(a) Amplification of 5'-UTR of GRM3

A 972 bp PCR product comprising the whole of Exon 1 of GRM3 was amplified using the primers F: 5′-CTCGCAGTGTGCAGTTGAGT-3′ and R: 5′-CTTACCTCGCACCGACAACT - 3′ (Sigma-Aldrich, UK). In a 96-well plate, master mix comprising of 2.5 µl of 10X Pfx amplification buffer, 5µl of 5 M betaine, 2.5 µl of 10X enhancer buffer, 1 µl of 50 mM $MgSO_4$, 0.2 µl of 25 mM dNTPs, 1 µl each of 10pmol/µl F and R primers, 0.4 µl Platinum® Pfx DNA polymerase (Invitrogen, UK) and PCR water was added to a final volume of 23 µl and dispensed into a well containing 2µl of of genomic DNA (50 ng). Platinum® Pfx DNA polymerase, a highly processive polymerase that contains a 3′-5' proofreading exonuclease and lacks a 5'-3′ exonuclease activity, was used in the PCR reaction to improve the fidelity and specificity of the PCR product. All the reagents, except 25 mM dNTPs (Bioline, UK) and betaine for the PCR, were obtained from Invitrogen, UK. The PCR program MHTD.pyex_1k was used to amplify the product and the cycling conditions are described in Table 3.10.

Table 3.10 Cycling conditions used for PCR amplification of GRM3 inserts for cloning

MHTD.pyex_1k	MHTD.pyex_2k
Lid heated to 105 ^{o}C	Lid heated to 105 ^{o}C
Hot start at 94 ^{o}C for 3 minutes	Hot start at 94 ^{o}C for 3 minutes
3 cycles of	3 cycles of
94^{o}C - 30 seconds	94^{o}C - 30 seconds
63^{o}C - 30 seconds	63^{o}C - 30 seconds
72^{o}C - 1 minute	72^{o}C - 2 minutes 30 seconds
3 cycles of	3 cycles of
94^{o}C - 30 seconds	94^{o}C - 30 seconds
60^{o}C - 30 seconds	60^{o}C - 30 seconds
72^{o}C - 1 minute	72^{o}C - 2 minutes 30 seconds
3 cycles of	3 cycles of
94^{o}C - 30 seconds	94^{o}C - 30 seconds
57^{o}C - 30 seconds	57^{o}C - 30 seconds
72^{o}C - 1 minute	72^{o}C - 2 minutes 30 seconds
3 cycles of	3 cycles of
94^{o}C - 30 seconds	94^{o}C - 30 seconds
54^{o}C - 30 seconds	54^{o}C - 30 seconds
72^{o}C - 1 minute	72^{o}C - 2 minutes 30 seconds
3 cycles of	3 cycles of
94^{o}C - 30 seconds	94^{o}C - 30 seconds
51^{o}C - 30 seconds	51^{o}C - 30 seconds
72^{o}C - 1 minute	72^{o}C - 2 minutes 30 seconds
33 cycles of	33 cycles of
94^{o}C - 30 seconds	94^{o}C - 30 seconds
48^{o}C - 30 seconds	48^{o}C - 30 seconds
72^{o}C - 1 minute	72^{o}C - 2 minutes 30 seconds
Hold at 72^{o}C - 20 minutes	Hold at 72^{o}C - 20 minutes
Store at 8^{o}C	Store at 8^{o}C

(b) Amplification of a 972 bp 5'-UTR with HindIII and NcoI restriction sites

A 972 bp GRM3 5'-UTR PCR product with *Hind*III and *Nco*I restriction sites was amplified using the primers (Sigma-Aldrich, UK)

F- 5′- CCTCTCAAGCTTCTCGCAGTGTGCAGTTGAGT - 3′

R - 5′- CTCCTCTCCATGGCTTACCTCGCACCGACAACT - 3′

using the standard protocol described in section 3.4.2.3.2 (a).

(c) Amplification of 1495 bp PCR fragment comprising promoter and 5'-UTR of GRM3

A 1495 bp PCR product comprising the whole of Exon 1 of GRM3 and a 523 bp minimal promoter containing the CCAAT and the TATA box required for GRM3 gene transcription as deduced from the Corti et al. (2001) was amplified using the primers F: 5'-TCTGTAGCTGGCAGCCTATT -3′ and R: 5'-CTTACCTCGCACCGACAACT - 3′. Master mix comprising of 2.5 µl of 10X Pfx amplification buffer, 5 µl of 5 M betaine, 2.5 µl of 10X enhancer buffer, 1 µl of 50 mM $MgSO_4$, 0.2 µl of 25 mM dNTPs, 1 µl each of 10 pmol/µl F and R primers, 0.4 µl Platinum® Pfx DNA polymerase (Invitrogen) and PCR water was added to a final volume of 23 µl and dispensed in a 96-well plate containing 2µl of of genomic DNA (50 ng). The PCR program MHTD.pyex_2k was used to amplify the product and the cycling conditions are described in Table 3.10.

3.4.2.3.3 Gel extraction of PCR products for cloning into reporter plasmids

The size of the PCR products was confirmed by running the products on 1% agarose gel with ethidium bromide and hyperladder I as the size marker. Upon verification, the PCR product band was extracted from the gel using Qiagen MinElute Gel

Extraction Kit (Qiagen, UK). The rehydrated purified PCR product was quantified using a Nano1000 (Thermo scientific, UK) and delivered to be sequenced at the Scientific Support Services at the Wolfson Institute for Biomedical Reasearch (WIBR), UCL to verify the sequence of the PCR product.

3.4.2.3.4. Cloning of PCR fragments into reporter plasmids

(A) Cloning of the PCR fragment into pJet1.2/blunt vector

Ligation: Initially the blunt end 972 bp 5'-UTR PCR product was cloned into the pjet1.2 blunt cloning vector (Fermentas) according to the manufacturer's instructions. A 5 minute ligation reaction was carried out at room temperature with 3:1 insert/vector ratio (0.15/0.05 pmol ends), 2X reaction buffer, T4 DNA ligase and PCR water to a final volume of 20 µl.

Transformation: An aliquot of 100 µl of Max efficiency DH5α competent cells (Invitrogen) was transformed with 2.5 µl of the ligation mixture according to the manufacturer's instructions. 100 µl of competent cells were thawed on ice and 2.5 µl of the ligation mix was added to the cells gently. The cells were then incubated on ice for 30 minutes and then heat shocked at 42^{o}C water bath for 45 seconds. The cells were immediately placed on ice for 2 minutes and then 0.9 ml of room temperature S.O.C medium was added to the cells. The tubes containing the cells were then placed in a shaker incubator at 37^{o}C for one hour at 225 rpm. After the incubation period, aliquots of 50 µl, 100 µl and 250 µl of cells were spread on LB agar plates containing 100 µg/ml ampicillin and the plates were incubated overnight at 37^{o}C. Also, another aliquot of 100 µl of competent cells were transformed in parallel with 50 pg of control DNA (pUC19) to determine the transformation efficiency. As the vector contains a lethal gene that is interrupted by the successful

cloning of an insert, only transformed colonies survive on LB-amp agar plates. On the next day, the transformed colonies were screened for the presence of the cloned insert using colony PCR.

Colony PCR: Colony PCR is a protocol designed for quick screening of the plasmid inserts directly from bacterial colonies. A single colony from the transformation plate was picked using a sterile 10 µl pipette tip and dipped in a 96-well plate containing 5 µl of PCR water and the tip was then inoculated into a fresh 15 ml falcon tube containing 3 ml of LB broth with ampicillin. The master mix containing all the PCR reagents described in section 3.4.2.3.2 for the amplification of 972 bp and 1495 bp PCR product was added to the 96-well plate containing cells and the appropriate cycling conditions were used based on the insert of interest. At the end of the PCR, the products were run on 1% agarose gel to determine their specificity (Figure 3.13).

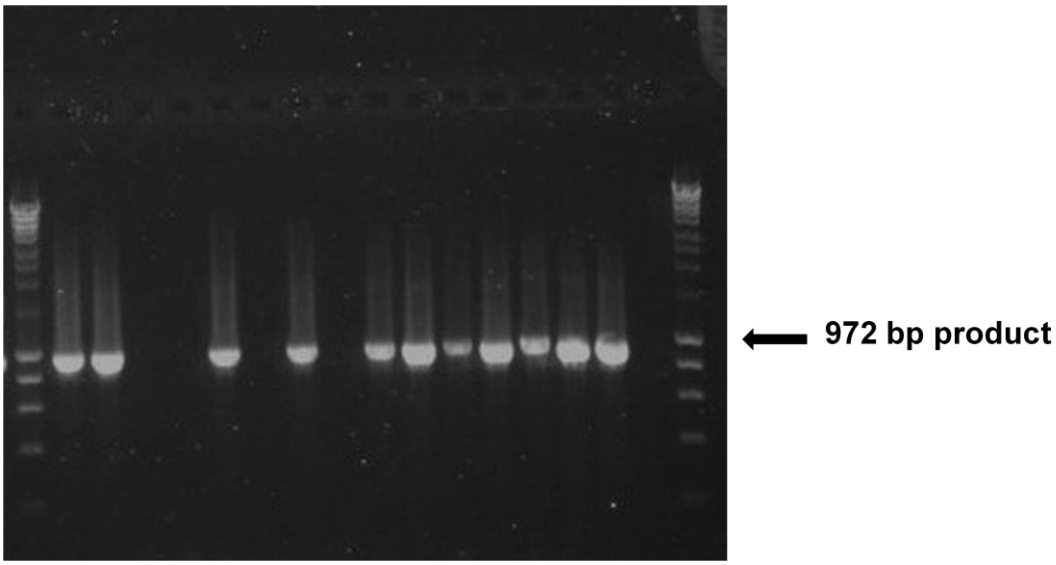


Figure 3.13 Colony PCR for identifying pJet1.2 clones containing 972 bp GRM3 5'-UTR.

Lanes 1 and 18, Hyperladder I; Lanes 2-17, randomly selected pJet1.2 clones for screening

3.4.2.3.5. Isolation, quantification and sequence verification of plasmid DNA

Meanwhile, the falcon tubes inoculated with single transformed colonies containing the vector along with the insert were selected and grown overnight in a shaker incubator at 37^{o}C and 225 rpm. On the next day, the cultures were subjected to plasmid isolation using QIAprep Spin Miniprep Kit (Qiagen, UK) based on the manufacturer's protocol. The plasmid DNA isolated was quantified using a NanoDrop® ND-1000 or NanoDrop 2000 Spectrophotometer (Thermo-Scientific, UK) and delivered to be sequenced at WIBR, UCL using pJET1.2 F and R primers (Appendix IV).

3.4.2.3.6. Sub-cloning of the GRM3 inserts into pGL3 vectors

(i) Sub-cloning of the GRM3 inserts from pjet1.2 vector into pGL3 vectors (Clones A, B and C)

Upon sequence verification of the pJet1.2 clones containing 972 bp (5'-UTR) and 1495 bp (promoter and 5'-UTR) of GRM3 gene, these inserts were subcloned into pGL3-basic and pGL3-promoter vectors resulting in three different constructs. I chose to utilize these two vector types as they differ in the basal transcriptional level of the luciferase, which is very low in the case of the pGL3-basic due to the lack of a constitutive promoter sequence and higher in the pGL3-promoter because of the presence of an SV-40 derived promoter. Two double digestion reactions using restriction enzymes (*Xho*I and *Bgl*II) common to the pJet1.2 and pGL3 vector of interest were performed - one was to excise the insert from the pJet1.2 vector and the other to cleave the pGL3 vector. The restriction enzyme digest was set up by adding 2 µl of 10 X restriction enzyme buffer, 0.2 µl of acetylated BSA (10 µg/µl), 1 µl of plasmid DNA (1 µg/µl) and sterile deionized water to a final volume of 15.3 µl and

finally adding 0.5 µl of each restriction enzyme (10 U/ µl).

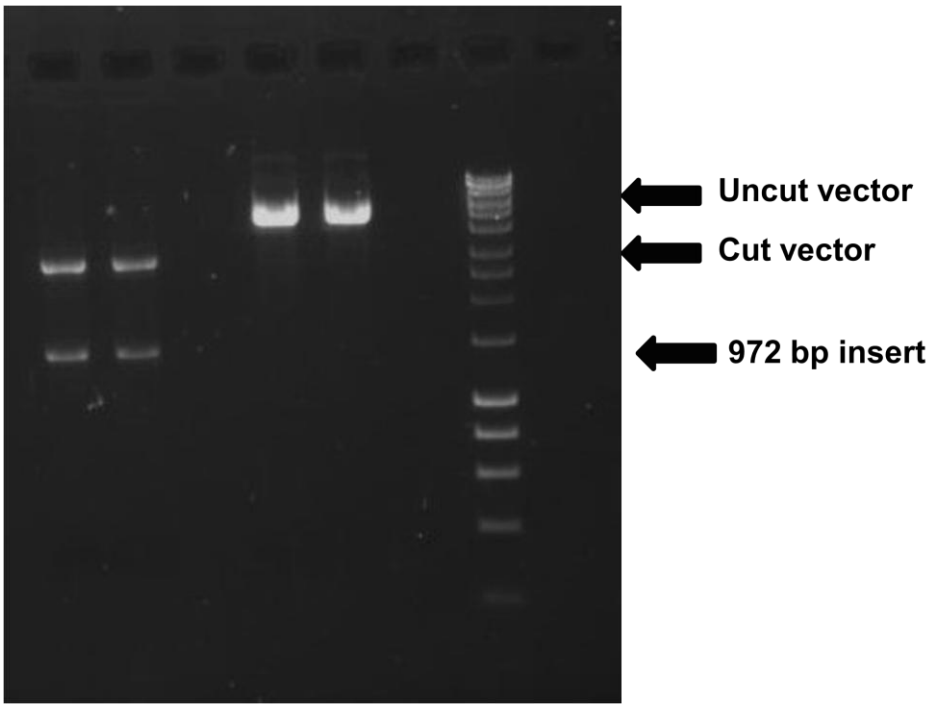


Figure 3.14 Restriction digestion of the pJet1.2 construct for subcloning

Lanes 1 and 2, pJet1.2/blunt construct containing 972 bp GRM3 5'-UTR insert treated with *Xho*I and *Bgl*II, Lanes 4 and 5, undigested pJet1.2/blunt clone containing the 972 bp GRM3 5'-UTR insert, Lane 7, Hyperladder I.

The contents in the tube were mixed gently by pipetting and spun briefly and incubated at 37^{o}C for 4 hours. BSA in the digestion reaction stabilizes the enzymes by protecting them from proteases, non-specific adsorption and harmful environmental factors such as heat, surface tension and interfering substances. In double restriction digestion reactions the most optimal buffer or the compatible buffer for the activity of both the enzymes was used. The digested reaction mix was then run on 1% agarose gel to confirm the completion of the cleavage of the plasmid DNA and separate the vector and the insert. On successful separation, two bands of appropriate sizes for the vector and the insert were visible on the gel for pJet1.2 vector digest (Figure 3.14) and one band for the cleaved pGL3 vector. The cleaved

and linear pGL3 vector ran faster than the circular uncleaved vector on the gel. The 972 bp or 1495 bp insert and the digested pGL3 vector were then extracted from the gel using QIAquick Gel Extraction Kit (Qiagen, UK). The DNA was quantified using Nano1000 and ligation reactions were carried out using T4 DNA ligase (Promega, UK) according to manufacturer's instructions. A 3:1 molar ratio of insert:vector DNA was used for ligation reaction and the amount of vector and insert DNA was calculated using this formula:

ng of insert = (ng of vector x kb size of insert/kb size of the vector) x (molar ratio of insert/vector)

The ligation reactions were performed at 4^{o}C overnight and the following day, 2 µl of the ligation mix was used to transform 50 µl of maximum efficiency DH5α/XL-10 Gold competent cells according to the standard protocol described in section 3.4.2.3.4. The transformed cells were spread on LB agar plates containing 100 µg/ml of ampicillin. The following day the transformed colonies were screened for the presence of the insert of interest by colony PCR. Plasmid DNA was extracted for colonies containing the correct insert the following day and delivered for sequencing at WIBR, UCL. The three resulting constructs and there compositions were:

1. **Clone A**: The 972 bp GRM3 5'-UTR was subcloned into pGL3-basic vector using *Xho*I and *Bgl*II restriction sites upstream of the luciferase gene.

2. **Clone B**: The 1495 bp PCR product comprising the whole of Exon 1 of GRM3 and a 523 bp minimal promoter was subcloned into pGL3-basic vector using the *Xho*I and *Bgl*II restriction sites.

3. **Clone C**: The 972 bp GRM3 5'-UTR was cloned upstream of the SV-40 promoter in the pGL3-promoter vector using *Xho*I and *Bgl*II restriction sites.

(ii) Sub-cloning of the 972 bp GRM3 5'-UTR PCR product into pGL3-promoter vector (Clone D)

The 972 bp GRM3 5'-UTR PCR fragment amplified using primers with *Hind*III and *Nco*I restriction sites was used for the construction of Clone D. Therefore, in Clone D, the GRM3 5'-UTR is cloned just outside and downstream of the SV-40 promoter and upstream and in-frame to the luciferase gene. A restriction digestion reaction was set up for the PCR product and pGL3-promoter vector in separate tubes using *Hind*III and *Nco*I restriction enzymes (Promega, UK) according to the standard protocol described before in section 3.4.2.3.6. The pGL3-promoter vector was dephosphorylated using calf intestinal alkaline phosphatase at 37^{o}C for 30 minutes. The digested insert and the vector were run on 1% agarose gel and extracted using the QIAquick Gel Extraction Kit. An overnight ligation reaction at 4^{o}C was set up using 3:1 molar ratio of insert: vector DNA and 2 µl of the ligation mix was used for transformation the following day. The transformed colonies were screened using colony PCR and restriction digestion with *Hind*III and *Nco*I to confirm the size and orientation of the insert. The colonies containing the insert of the correct size and orientation were processed for plasmid isolation and sequence verification.

(iii) Mutating the A base of the ATG of luciferase gene in Clone D to produce Clone E

In Clone E, the A base of the start codon of the luciferase gene of Clone D was replaced with the G base using the QuikChange II XL system (Strategene, LA Jolla, CA) according to the manufacturer's instructions for site-directed mutagenesis

described in section 3.4.2.3.7. In this clone, the translation of the gene is under the control of in-frame ATG start codon of GRM3 972 bp 5′ UTR. The sequence and orientation of the experimental construct was verified by DNA sequencing.

3.4.2.3.7. Site-directed mutagenesis

Single base mutations were introduced in all the constructs by site-directed mutagenesis utilizing the QuikChange II XL system (Strategene, UK) according to the manufacturer's instructions. The primers for site-directed mutagenesis were designed using web-based QuikChange Primer Design Program available online at www.agilent.com/genomics/qcpd. The HPLC purified oligos (Sigma-Aldrich, UK) were reconstituted with PCR water on arrival. The mutant strand synthesis reaction was set up by adding 5 µl of 10X reaction buffer, 10 ng of wild-type plasmid DNA, 125 ng of F and R primers, 1 µl of dNTP mix, 3 µl of Quicksolution and PCR water to a final volume of 50 µl in a 96-well PCR plate. Finally 1 µl of *PfuUltra* HF DNA polymerase (2.5 U/ul) was added to the reaction mix and mixed well by pipetting. The samples were then subjected to PCR and the cycling conditions used are described in Table 3.11.

Table 3.11 Cycling conditions for site directed mutagenesis

Cycles	Temperature	Time
1	95^{o}C	1 minute
18	95^{o}C	50 seconds
	60^{o}C	50 seconds
	68^{o}C	1 minute/kb of plasmid length
1	68^{o}C	7 minutes

After the PCR, the reactions were cooled by keeping on ice for 2 minutes. To the reaction mix, 1 µl of *Dpn*I restriction enzyme (10 U/µl) was added and gently mixed

by pipetting. The reaction mix was spun down in a centrifuge briefly and incubated at 37^{o}C for 1 hour to digest the parental supercoiled dsDNA. Following the incubation, XL10Gold ultracompetent cells were transformed using 2 µl of the *Dpn*I digested DNA in a 14 ml BD round bottom falcon tube using the standard protocol described in the kit manual. S.O.C medium was added to the falcon tubes after heat shock and the cells were incubated at 37^{o}C for 1 hr with shaking at 225 rpm. Aliquots of cells were plated on LB-ampicillin plates and the plates were incubated at 37^{o}C overnight. A mutagenesis control was performed in parallel using pWhitescript 4.5-kb control plasmid and the competent cells transformed with this plasmid were plated on LB-ampicillin plates containing 80 µg/ml X-gal and 20 mM IPTG. A successful pWhitescript control mutagenesis reaction resulted in blue colonies. The transformed colonies were screened using colony PCR and plasmid was isolated the following day and sequence verified for the presence of the mutation.

3.4.2.4. Large-scale plasmid preparation

For transfection experiments, plasmid DNA was isolated on a large-scale using 25 ml of cultures of the clones inoculated in LB-ampicillin medium and grown overnight. The plasmid was extracted using QIAfilter Plasmid Midi Kit (Qiagen, UK) according to the manufacturer's instructions.

3.4.2.5. Preparation of glycerol stocks for long-term storage of clones

The culture of the clones was streaked on LB-ampicillin plates for isolating single colonies. The following day, a single colony was inoculated in 3 ml LB-ampicillin medium and incubated overnight at 37^{o}C. After the incubation, 0.5 ml of the culture was added to 0.5 ml of 80% sterile glycerol in a sterile screw cap tube. This glycerol stock was mixed by vortexing and then stored at -80^{o}C. The plasmid was isolated from 1 ml of the same culture using the Qiagen miniprep kit and sequence verified for confirmation.

3.4.2.6. Transient transfection

Day before transfection

Twenty four hours before transfection, HEK293 and SH-SY5Y cells were seeded in 24 well plates at approximately 0.5 x 10^5 and 1 x 10^5 cells/well for HEK293 and SH-SY5Y cells respectively in 500 µl of antibiotic-free medium (DMEM + 10% FBS). The plates were incubated at 37^{o}C with 5% CO_2 and were 90-95% confluent at the time of transfection.

Day of transfection

Transfection was performed using Lipofectamine^{TM} 2000 (Invitrogen, UK) according to the manufacturer's protocol. For each well to be transfected, two microcentrifuge tubes were prepared. In one tube, 0.8 µg of the reporter plasmid DNA along with pRL-SV40 renilla luciferase reporter vector at a ratio of 50:1 (16 ng/well) was diluted in 50 µl of Optim-MEM®I reduced serum medium. pRL-SV40

plasmid DNA was used as a control in the experiment to determine the transfection efficiency. In the second tube, 2 µl and 4 µl of Lipofectamine™ 2000 for HEK293 and SH-SY5Y cells respectively, was diluted in 50 µl of Optim-MEM®I medium and incubated for 5 minutes at room temperature. At the end of the incubation period, the contents of tube one and two were mixed together gently to form transfection complexes and incubated for 20 minutes at room temperature. The 100 µl of DNA-lipofectamine complexes were then added to each well to be transfected and mixed gently by rocking the plate back and forth. The cells were incubated at 37°C in the CO_2 incubator. The medium was changed after four hours of incubation with complete growth medium and the cells were incubated for further 48 hours at 37°C with 5% CO_2.

3.4.2.7. Luciferase reporter assay

The Dual-Luciferase® Reporter (DLR™) Assay system (Promega Corp, USA) was used to quantify the reporter activity after transfection. In the DLR™ Assay, the firefly and *renilla* luciferase activities can be measured sequentially from a single sample. The components of this system are:

- Passive Lysis buffer (PLB, 5X): PLB promotes rapid lysis of cultured mammalian cells, prevents foaming during lysis and provides optimum performance and stability of firefly and Renilla luciferase reporter enzymes. Working stock of PLB was prepared by adding one volume of 5X PLB to 4 volumes of distilled water and mixing well. The diluted PLB (1X) can be stored at 4°C for up to one month.

- Luciferase Assay Reagent II: The firefly luciferase reporter activity was measured using LAR II. Working stock of LAR II was prepared by resuspending the lyophilized luciferase assay substrate in 10 ml of luciferase assay buffer II. Aliquots of 1 ml LAR II were stored at -80^{o}C.

- Stop & Glo® Reagent: This reagent quenches the firefly luciferase luminescence and concomitantly activates *Renilla* luciferase. Fresh Stop & Glo® reagent was prepared everytime by adding one volume of 50X Stop & Glo® substrate to 50 volumes of Stop & Glo® buffer.

Preparation of cell lysate

The cells were harvested 48 hours after transfection and the activity of firefly and renilla luciferase was measured using the Dual Luciferase Reporter® Assay System (Promega) on a Turbo 20/20 luminometer. The growth medium was removed from the wells and the cells were washed with 1X PBS. To each well, 100 µl of 1X PLB was added and the plate was placed on an orbital shaker at room temperature for 15 minutes. The lysate was then transferred to a microcentrifuge tube and the lysate was cleared by centrifugation at 13000 rpm for 30 seconds at 4^{o}C. The clear lysates were transferred to fresh microcentrifuge tubes.

Measurement of reporter activity using a luminometer

For each sample to be assayed, 100 µl of LAR II was predispensed into a microcentrifuge tube. The DLR II Promega program was selected on the Turbo 20/20 luminometer (Turner Biosystems, UK) for reporter activity measurement. The luminometer was programmed to perform a 2 second premeasurement delay followed by a 10 second measurement period for each reporter assay. In the tube containing LAR II, 20 µl of the cell lysate was transfered carefully and mixed by

pipetting and the tube was placed in the luminometer for reading. This was the luciferase activity measurement for the sample. After recording the initial reading, the tube was removed and 100 µl of Stop & Glo® reagent was added and mixed by pipetting. The tube was replaced in the luminometer and the reporter activity recorded. This reading was the *Renilla* luciferase activity measurement for the sample.

Statistical analysis

Transfections were performed in triplicate and repeated three times. Firefly luciferase luminescence was expressed as a ratio to that of the *Renilla* luciferase activity and the normalized data was compared with a Student's t-test or Mann-Whitney U-test (SPSS, version 18.0). A cut-off significance of $P < 0.05$ was used.

3.4.2.8. mRNA expression analysis

Transfection

One day before transfection, HEK293 and SH-SY5Y cells were seeded in 6-well plates at approximately 5×10^5 and 1×10^6 cells/well in an antibiotic-free medium. Transfections were performed for HEK293 and SH-SY5Y cells using 10 µl/ well and 15 µl/well Lipofectamine 2000 respectively following the standard protocol described in section 3.4.2.4. 2.4 µg of Clone D reporter construct and 48 ng/well of pRL-SV40 at a ratio of 50:1 per well were transfected into each well containing the cells. The medium was changed to a complete growth medium after four hours and the plates were incubated for 24 hours at 37^oC with 5% CO_2.

RNA extraction

Total RNA was extracted after 24 h of incubation using Illustra RNAspin Mini RNA Isolation Kit (GE Healthcare Life Sciences, UK) and subjected to DNAse treatment using RQ1 RNase-free DNase (Promega, UK) for 1 hour at 37^{o}C in order to eliminate plasmid DNA according to the manufacturer's instructions. The RNA was quantified using Nano2000 (Thermo scientific, UK) by using 1 µl of the sample.

cDNA preparation

DNAse treated RNA was used to make cDNA in the presence of Superscript^{TM} III Reverse Transcriptase (Life Technologies, UK) according to the manufacturer's instructions. In a nuclease-free microcentrifuge tube, 2 µg of total RNA, 500 ng of oligo(dT)$_{15}$ and 1 µl of 10 mM dNTP mix were added and PCR water was added to a final volume of 13 µl. The mixture was heated in a thermocycler at 65^{o}C for 5 minutes and immediately incubated on ice for 1 minute. The tube was then centrifuged briefly and to each tube 4 µl of 5X first-strand buffer, 1 µl of 0.1 M DTT, 1 µl of Recombinant Rnasin® Ribonuclease inhibitor (Promega, UK) and 1 µl of SuperScript^{TM} III RT (Invitrogen, UK) was added. The contents in the microcentrifuge tube were mixed gently by pipetting and incubated at 55^{o}C for 60 minutes. The reaction was inactivated by heating the tube at 70^{o}C for 15 minutes. A negative control or a non-RT reaction was maintained by excluding the RT enzyme in the reaction contents. The resulting cDNA was then used for Quantitative real time PCR (QRT-PCR).

Quantitative real time PCR

Firefly luciferase mRNA was quantified using primers F – CCTTCGATAGGGACAAGACAA, R- AGCGACACCTTTAGGCAGAC and

Universal Probe Library probe #82 (UPL; Roche Diagnostics Ltd, UK) using an LC480. *Renilla* luciferase mRNA was also quantified using primers F-TGTGTCAGTTAGGGTGTGGAA, R- CACCTGGTTGCTGACTAATTGA and UPL probe #29 (Roche Diagnostics Ltd, UK) to take into account differences in transfection efficiencies between constructs. A housekeeping gene, *RPLPO*, was quantified as an internal control using primers F - CGACAATGGCAGCATCTAC and R - GGTAGCCAATCTGCAGAC, and UPL probe #6 (Roche Diagnostics Ltd, UK). In a 384 well plate, 9 µl master mix containing 5 µl of probes master mix (Roche Diagnostics Ltd, UK), 100 nM UPL probe primer, 200 nM forward and reverse primers was added to 1 µl of cDNA sample and each sample was amplified in quadruplicate. The dual hydrolysis probe/UPL probe program was run on the LC480. Firefly luciferase was normalized to that of renilla luciferase to control the transfection efficiency. The $\Delta\Delta C_t$ method was employed by the LC480 software for the analysis of the results and is described in section 3.3.6.

Chapter 4

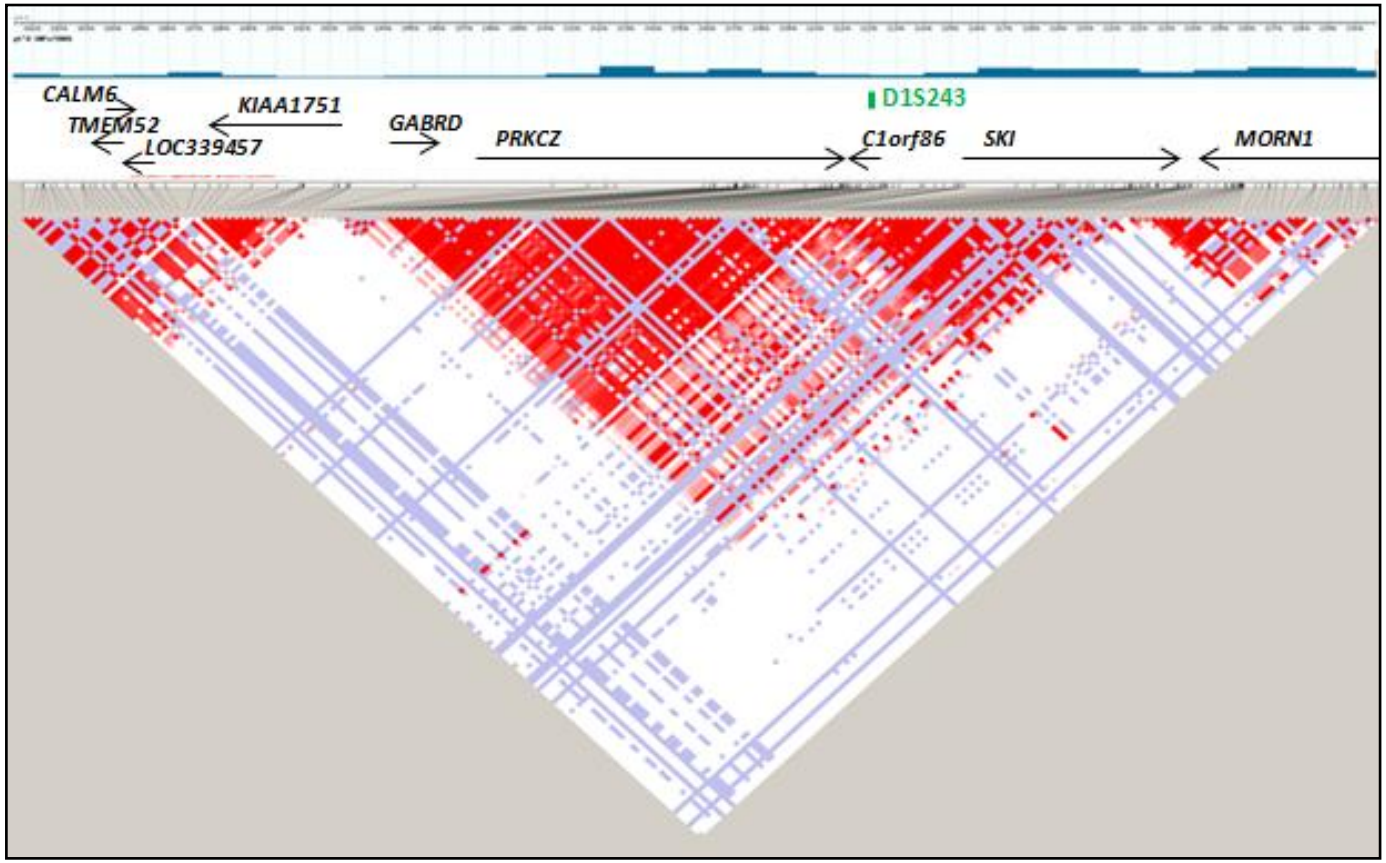
Fine mapping of a new Bipolar affective disorder locus on chromosome 1p36

4.1. Introduction

Accumulating evidence from family linkage studies suggests the presence of a BPAD susceptibility locus on 1p36. The linkage findings implicating chromosome 1p have been previously discussed in section 1.7.4.1. The 1p36 linkage region that included the positive marker of the linkage study (Curtis et al. 2003), D1S243, and 10 Mbp on either side of the marker spans about 187 RefSeq genes. Some of the candidate genes in this region are GABRD, GPR153, FBXO, PER3, MTHFR and PRKCZ. On close inspection of the linkage disequilibrium (LD) pattern, the marker D1S243 lies in an LD block comprising genes PRKCZ, C1orf86 and SKI (Figure 4.1). Of these, protein kinase C, zeta type (PRKCZ) was one of the promising candidate genes as it is very highly expressed in all parts of the brain including the prefrontal cortex, amygdala and hypothlalamus. PRKCZ is a member of the PKC family of serine/threonine kinases, which are involved in cell proliferation, differentiation and secretion. PRKCZ protein has an N-terminal regulatory domain followed by a hinge region and a C-terminal catalytic domain. The second messengers stimulate PKCs by binding to the regulatory domain, translocating the enzyme from cytosol to membrane and producing a conformational change that removes autoinhibition of the PKC catalytic activity. PRKCZ has been implicated in cardiac muscle function (Wu and Solaro 2007), B-cell function (Martin et al. 2002), and insulin-stimulated glucose transport (Bandyopadhyay et al. 2002). Allelic and haplotypic association studies were carried out in the UCL bipolar disorder case-control sample for this thesis in order to confirm the role of the PRKCZ in genetic susceptibility to affective disorders. The method relies on the detection of evolutionarily determined relationship of linkage disequilibrium (LD) between

polymorphic markers within or very close to the PRKCZ gene and unknown aetiological base pair changes within PRKCZ. Positive evidence for LD is found by discovering allelic or haplotypic association between markers and disease status compared to ancestrally matched unaffected comparison subjects.

Figure 4.1 RefSeq genes in the linkage region including the positive marker D1S243 and 10 Mbp on either side of D1S243



Linkage region implicated in the UCL family linkage study (Curtis et al. 2003)

4.2. Hypothesis

On the basis of modest evidence from family linkage studies I hypothesised that the chromosome 1p36 locus contains a BPAD susceptibility gene.

4.3. Aims

The aim of this study was to fine map the 1p36 region using microsatellite markers. Furthermore, tagging SNPs were sought to improve the association signal between *PRKCZ* and bipolar disorder. The previously reported gender-based association of 1p36 region with bipolar disorder was also investigated in the UCL research sample.

4.4. Results

Genotyping of four microsatellite markers was initially carried out in a total of 440 controls and 600 BPAD patients. Marker D1S243 showed positive allelic association with bipolar disorder (P = 0.037; CLUMP T1) as shown in Table 4.1. The three other microsatellite markers – PRKCZ $(GT)_{17}$, PRKCZ$(CA)_{13}$ and PRKCZ$(AC)_{24}$ (Table 4.1) failed to show any association with bipolar disorder (P = 0.555, 0.472, 0.68 respectively). The Hapmap LD structure of all the SNPs between the markers PRKCZ $(AC)_{24}$ and PRKCZ$(CA)_{13}$ is shown in Figure 4.1. It is evident from Figure 4.1 that D1S243 falls into an LD block which spans the genes PRKCZ, SKI and C1orf86 and there is a distinct break in the LD outside this region. Therefore, this region was fine mapped by genotyping nine haplotype tagging SNPs spanning the PRKCZ gene (Figure 4.2). All the SNP markers were in HWE. SNP rs3128296 was significantly associated with bipolar disorder (P = 0.040) and another SNP rs2503706 (P = 0.057) showed a trend towards association (Table 4.2).

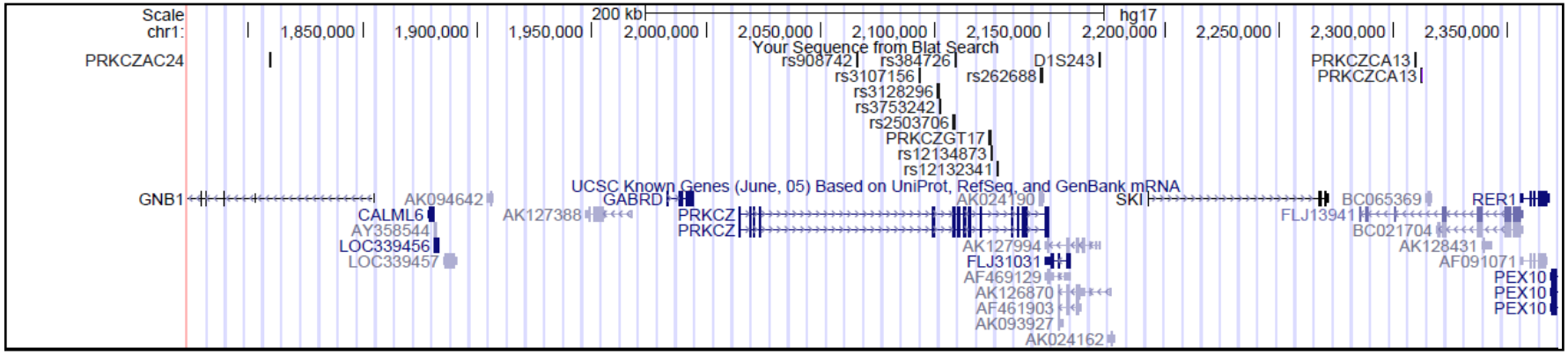


Figure 4.2 Graphical representation of the markers genotyped for the fine mapping of 1p36 locus

Table 4.1 Tests of allelic association of microsatellite markers with Bipolar affective disorder at the 1p36 region

Marker	Marker Location	Bp from prior marker	Number of samples (N)	Allelic bases or fragment sizes with observed allele counts (allele frequency %) below	χ^2	P[a]-value
PRKCZ$(AC)_{24}$	1809303			232 234 236 248 250 252 254 256 258 260 262 264 266 268 270 272 274		
Case			563	1 331 11 0 1 6 23 224 50 180 215 42 20 12 9 0 1	13.53	0.634
				(0.1) (29.4) (1) (0) (0.1) (0.5) (2.0) (19.9)(4.4) (16) (19.1)(3.7) (1.8) (1.1) (0.8) (0) (0.1)		
Control			397	0 224 8 1 0 3 12 142 41 132 182 23 15 7 3 1 0		
				(0) (28.2) (1) (0.1) (0) (0.4) (1.5) (17.9) (5.2)(16.6)(23.0)(2.9)(1.9) (0.9) (0.5) (0.1) (0)		
PRKCZ$(GT)_{17}$	2123652	314349		295 297 299 301 303 305 307 309 311 313 315		
Case			558	1 31 281 24 574 93 13 11 85 3 0	8.956	0.536
				(0.1)(2.8) (25.2)(2.2) (51.4) (8.3) (1.2) (1.1)(7.6) (0.3) (0)		
Control			380	0 20 193 15 381 59 11 12 60 8 1		
				(0) (2.6)(25.4) (2.1) (50.1)(7.8) (1.5) (1.6) (7.9) (1.1) (0.1)		
D1S243	2171436	47784		159 161 163 165 167 169 171 173 175 177 179 181 183 185 187 189		
Case			595	355 1 1 1 7 98 220 66 23 55 75 189 78 21 0 0	26.13	**0.037**
				(29.8)(0.1) (0.1) (0.1)(0.6) (8.2) (18.5) (5.6) (1.9)(4.6) (6.3) (15.9)(6.6) (1.8) (0) (0)		
Control			436	240 4 0 0 1 91 138 51 10 56 56 128 68 27 1 1		
				(27.5)(0.5) (0) (0) (0.1)(10.4)(15.8)(5.9) (1.2) (6.4) (6.4) (14.7)(7.8) (3.1) (0.1) (0.1)		
PRKCZ$(CA)_{13}$Case	2309397	137961		231 233 235 237 239 241 243 245 247 249 251 253 255		
			581	2 2 287 342 65 254 6 11 119 51 16 1 6	11.66	0.473
Control				(0.2) (0.2)(24.7)(29.4)(5.6)(21.9)(0.5) (0.9)(10.2)(4.4) (1.4) (0.1) (0.5)		
			394	1 0 195 233 42 150 11 6 91 38 18 1 2		
				(0.1) (0) (24.7)(29.6)(5.3)(19.0)(1.4) (0.8)(11.5)(4.8) (2.3) (0.1) (0.3)		

[a] Two-tailed significance from 2 x 2 χ^2, with 1 df; df, degrees of freedom, from CLUMP Monte Carlo subtest T1. Bold values signify P<0.05, Marker positions on chr 1 based on NCBI35 build hg17.

No significant differences were observed in the allele distributions of SNPs rs908472, rs3107156, rs384726, rs12134873, rs12132341 and rs262688 between patients and controls. After the initial fine mapping study, data from eight more SNPs in the region spanning PRKCZ gene was retrieved from the UCL1 bipolar GWAS (Ferreira et al. 2008; Sklar et al. 2008) and included in the final allelic and haplotypic association analysis. These SNPs are marked with an asterisk in Table 4.2. The SNP rs3753242 was genotyped in both the studies i.e. fine mapping at UCL and UCL1 GWAS, and the merged data from these two studies is reported in Table 4.2. One SNP rs3128309 approached significance in the association analysis with bipolar disorder (P = 0.060). The results of pairwise LD statistics between all 21 markers (microsatellites and SNPs) genotyped are shown in Figure 4.3. Strong LD was observed between the positive marker D1S243 and the positive SNPs rs3128296, rs3128309 and rs2503706 near the PRKCZ gene. The SNP rs3128296 was not in LD with rs2503706 (D' = 0.553, r^2 = 0.02) but was in moderate LD with rs3128309 (D'=0.943, r^2=0.747). SNPs rs3128309 and rs2503706 are in LD (D' = 1, r^2 = 0.055). The UCL controls recruited in this study are supernormal because they have been psychiatrically screened. To determine if the LD pattern in this group of controls represents the normal population, the LD structure of the SNP data of the controls genotyped in the UCL1 bipolar GWAS was compared to that of the Hapmap controls. The results suggested that the LD pattern of the two groups was similar (Figure 4.4).

Table 4.2 Genotype and allele frequencies of single nucleotide polymorphisms in the PRKCZ gene in the UCL research sample.

SNP	*Position*	*N*	*Genotype counts*			*Allele counts (frequency)*		χ^2	P^a*-value*
rs7513222*	2060063		AA	AG	GG	A	G	0.553	0.457
Case		543	62	245	236	369 (0.34)	717 (0.66)		
Control		510	50	231	229	331 (0.33)	689 (0.67)		
rs908742	2065418		AA	GA	GG	A	G	0.006	0.940
Case		576	58	263	255	379 (0.33)	773 (0.67)		
Control		428	46	191	191	283 (0.33)	573 (0.67)		
rs3107146*	2079746		AA	AG	GG	A	G	1.446	0.230
Case		543	5	114	424	124 (0.11)	962 (0.89)		
Control		510	9	116	385	134 (0.13)	886 (0.87)		
rs3107156	2092792		CC	CT	TT	C	T	0.415	0.520
Case		577	217	271	89	705 (0.61)	449 (0.39)		
Control		434	147	224	63	518 (0.60)	350 (0.40)		
rs3107157*	2094131		AA	AG	GG	A	G	0.352	0.553
Case		537	94	258	185	446 (0.41)	628 (0.59)		
Control		501	88	253	160	429 (0.43)	573 (0.57)		
rs3128296	2101068		GG	TG	TT	G	T	4.23	**0.040**[2]
Case		584	8	113	463	129 (0.11)	1039 (0.89)		
Control		543	10	131	402	151(0.14)	935 (0.86)		
rs3753242*[1]	2101843		CC	CT	TT	C	T	1.702	0.192
Case		591	508	80	3	1096 (0.93)	86 (0.07)		
Control		605	535	68	2	1138 (0.94)	72 (0.06)		
rs2503706	2107413		CC	CT	TT	C	T	3.62	0.057
Case		580	286	239	55	811 (0.70)	349 (0.30)		
Control		550	236	256	58	728 (0.66)	372 (0.34)		
rs384726	2108217		AA	GA	GG	A	G	0.138	0.710
Case		579	3	80	496	86 (0.07)	1072 (0.93)		
Control		429	2	56	371	60 (0.07)	798 (0.93)		
rs385039*	2109571		CC	CT	TT	C	T	0.251	0.616
Case		543	44	224	275	312 (0.29)	774 (0.71)		
Control		510	35	213	262	283 (0.28)	737 (0.72)		
rs12134873	2123984		AA	GA	GG	A	G	0.595	0.441
Case		570	0	6	564	6 (0.01)	1134 (0.99)		
Control		435	0	7	428	7 (0.01)	863 (0.99)		
rs12132341	2126694		AA	AG	GG	A	G	0.355	0.551
Case		580	0	23	557	23 (0.02)	1137 (0.98)		
Control		554	1	24	529	26 (0.02)	1082 (0.98)		
rs262641*	2137143		CC	CT	TT	C	T	0.240	0.624
Case		535	44	221	270	309 (0.29)	761 (0.71)		
Control		507	37	209	261	283 (0.28)	731 (0.72)		
rs3128309*	2138040		AA	AG	GG	A	G	3.532	0.060
Case		542	1	98	443	100 (0.09)	984 (0.91)		
Control		507	6	107	394	119 (0.12)	895 (0.88)		
rs626479*	2142422		AA	AG	GG	A	G	0.011	0.916
Case		530	38	208	284	284 (0.27)	776 (0.73)		
Control		504	33	202	269	268 (0.27)	740 (0.73)		
rs262688	2145727		AA	CA	CC	A	C	0.064	0.801
Case		571	293	235	43	821 (0.72)	321 (0.28)		
Control		433	222	183	28	627 (0.72)	239 (0.28)		
rs262683*	2188031		CC	CT	TT	C	T	2.286	0.131
Case		536	386	131	19	903 (0.84)	169 (0.16)		
Control		503	382	107	14	871 (0.87)	135 (0.13)		

Position on Chr 1 (bp)(UCSC May 2004 release); N, Number of samples; *SNPs genotyped in the GWAS (Ferreira et al. 2008; Sklar et al. 2008); *[1] SNPs genotyped in the GWAS and fine mapping study; [2]The P-value does not survive multiple testing

[a]Two-tailed significance P-value from 2 x 2 χ^2 1 df; df, degrees of freedom; Bold values signify P<0.05

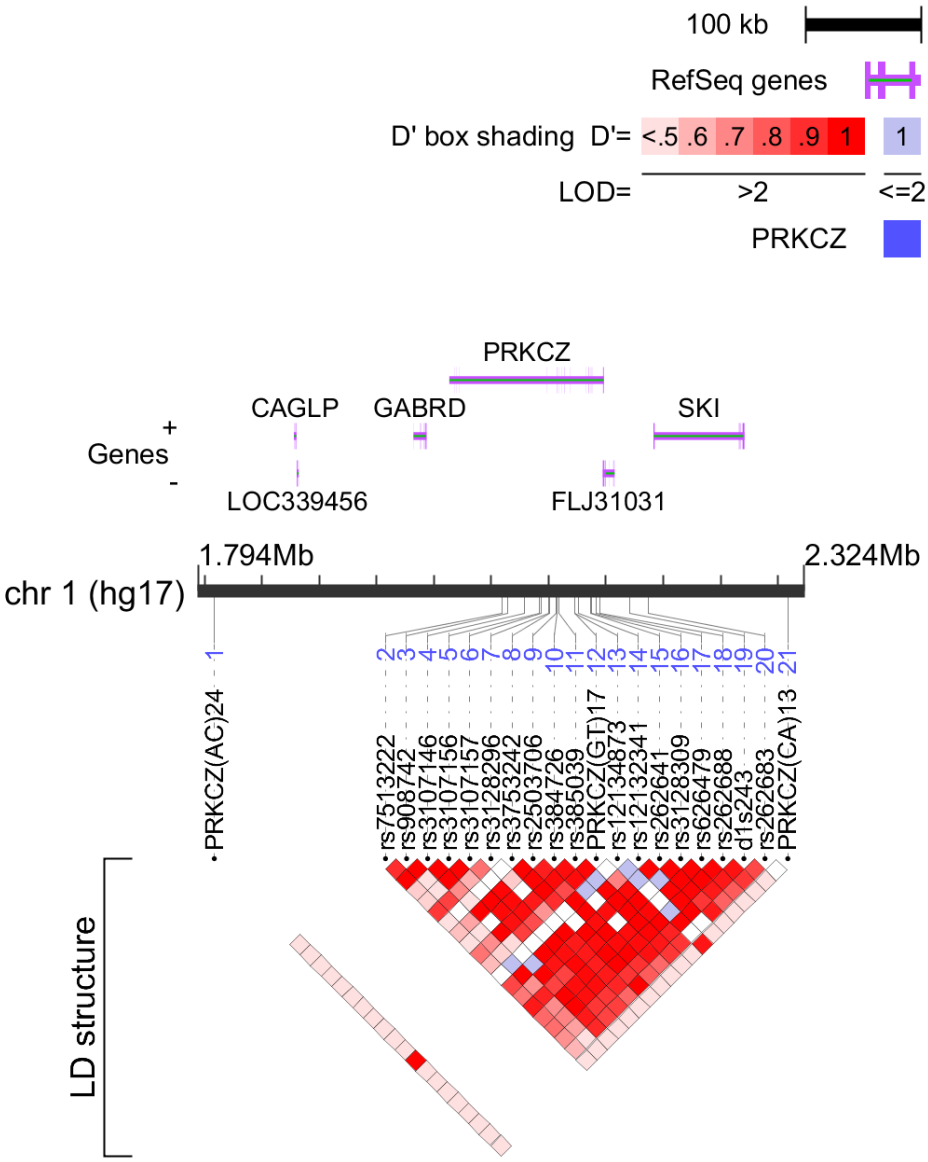


Figure 4.3 Linkage disequilibrium between all markers genotyped within and near the PRKCZ gene locus in the UCL sample

The figure was generated using Locusview 2.0. Genomic positions are in accordance with the NCBI information build 35 hg 17 of the human genome assembly.

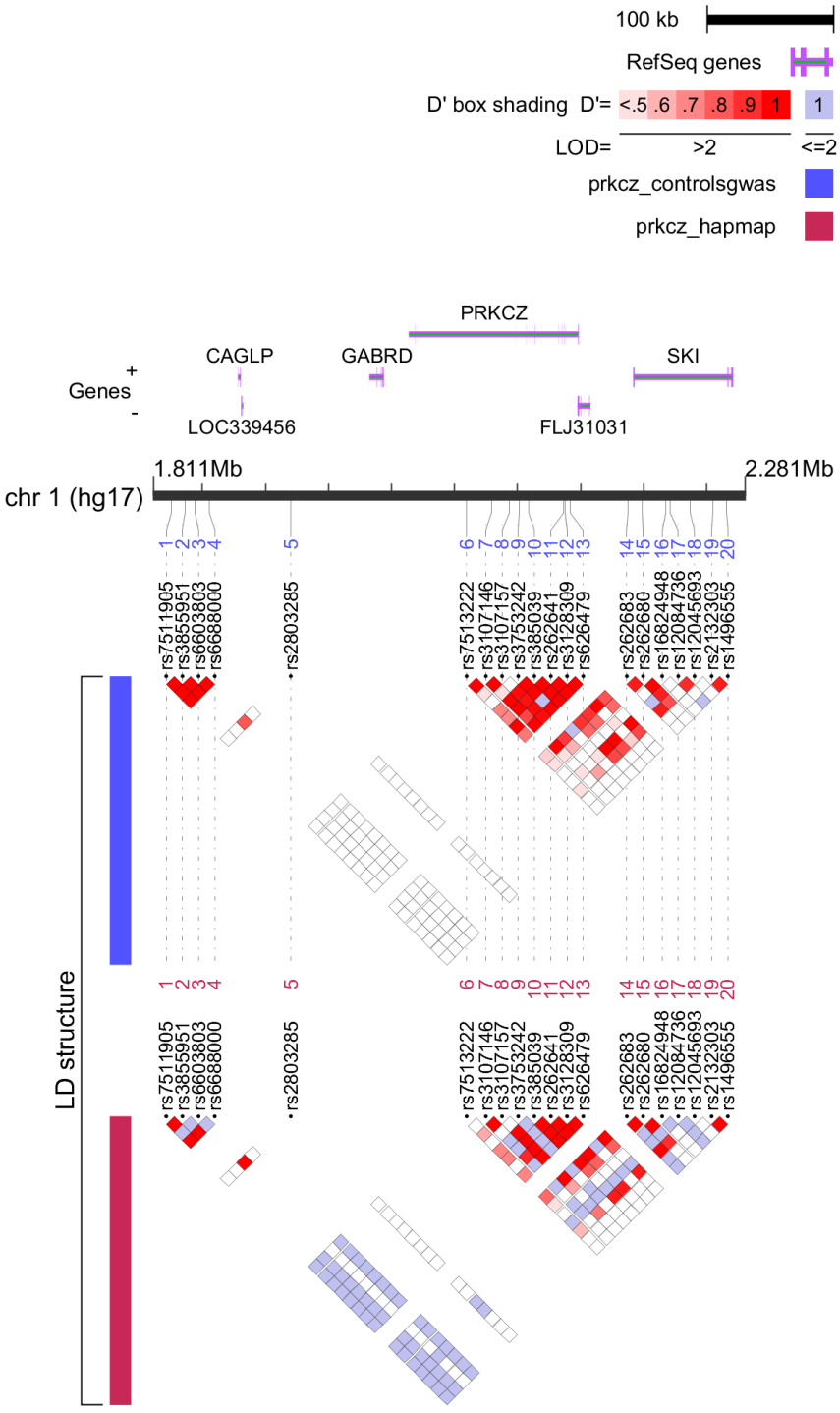


Figure 4.4 Comparison of LD between all SNP markers genotyped within and near the PRKCZ gene locus in the UCL control sample and Hapmap LD between the same SNP markers

The figure was generated using Locusview 2.0. Genomic positions are in accordance with the NCBI information build 35 hg 17 of the human genome assembly.

In the combined haplotype analysis of all the 17 SNPs, using the solid spine of D' = 1 rule in Haploview, there were three main haplotype blocks but none remained associated with bipolar disorder after permutation testing. Therefore, further exploratory haplotypic association analysis was undertaken. I found evidence for haplotypic association at PRKCZ when the three SNPs rs3128296, rs2503706 and rs3128309 were analysed as a single haplotype in Haploview (empirical P = 0.004) (Table 4.3). Tests of haplotypic association using GENECOUNTING were carried out. Two locus and three locus combinations of the associated SNPs and D1S243 were analysed. A trend towards haplotypic association was observed when D1S243 marker data was included with the SNP data from all the three associated SNPs (empirical P = 0.06).

Table 4.3 Haplotypic association tests with BPAD at the 1p36 region with markers within and near PRKCZ gene

No. of markers and haplotype reference	Haplotype	Alleles	Frequency		χ^2	Global empirical P-value*
			Cases (%)	Controls (%)		
2 marker haplotypes Hap 1	rs3128296 – rs2503706	T-C	60.3	54.4	8.20	0.009
Hap2	rs2503706 – rs3128309	C-G	60.4	54.9	7.57	0.008
Hap3	rs3128296 – rs3128309	T-G	88.5	85.7	6.53	0.016
3 marker haplotypes Hap4	rs3128296-rs2503706-rs3128309	T-C-G	59.9	54.0	8.48	0.004

*P**, Haplotype-permutation test empirical *P*, based on 9,999 permutations.

Association analysis of female cases with female controls using PLINK showed that SNPs rs3128296 and rs3128309, which were previously associated in the whole sample, became more significantly associated (P = 0.0035; OR = 0.6197; 95% CI

0.45 – 0.86 and P =0.0159; OR = 0.646; 95% CI 0.45 – 0.92 respectively) (Table 4.4). In the females only analysis an additional SNP rs3107146 became significant (P = 0.044; OR = 0.703; 95% CI 0.50 – 0.99). None of these SNPs survived multiple correction. None of the microsatellite markers showed gender-specific association with bipolar disorder in the female-only analysis (Table 4.4). Additionally, only one of the markers rs2503706 showed association with bipolar disorder when males only case-control pairs were analysed (Table 4.5).

Table 4.4 Association analysis of PRKCZ SNPs with bipolar disorder after excluding males in the UCL sample

Marker	MAF in cases	MAF in controls	χ^2	OR	P-value
PRKCZ$(AC)_{24}$			8.051		0.947[a]
rs7513222	0.357	0.322	1.668	1.166	0.197
rs908742	0.344	0.324	0.506	1.094	0.477
rs3107146	0.101	0.137	4.085	0.703	**0.043**
rs3107156	0.385	0.400	0.250	0.941	0.617
rs3107157	0.419	0.433	0.243	0.945	0.622
rs3128296	0.105	0.158	8.507	0.62	**0.004**
rs3753242	0.066	0.056	0.598	1.189	0.440
rs2503706	0.310	0.320	0.138	0.957	0.710
rs384726	0.077	0.066	0.479	1.175	0.489
rs385039	0.299	0.278	0.666	1.107	0.414
PRKCZ$(GT)_{17}$			11.977		0.287[b]
rs12134873	0.007	0.013	0.871	0.571	0.351
rs12132341	0.023	0.027	0.285	0.829	0.594
rs262641	0.301	0.278	0.802	1.119	0.370
rs3128309	0.089	0.131	5.809	0.646	**0.016**
rs626479	0.280	0.266	0.316	1.074	0.574
rs262688	0.289	0.266	0.751	1.122	0.386
D1S243			20.022		0.172[c]
rs262683	0.168	0.135	2.722	1.301	0.099
PRKCZ$(CA)_{13}$			7.257		0.840[d]

MAF, minor allele frequency; Bold values signify P<0.05 , and these do not withstand multiple correction; [a] Empirical Clump T1 significance (16 df); [b] Empirical Clump T1 significance (10 df); [c] Empirical Clump T1 significance (15 df) ; [d] Empirical Clump T1 significance (12 df); df, degrees of freedom.

Table 4.5 Association analysis of PRKCZ SNPs with Bipolar disorder after excluding females in the UCL sample.

Marker	MAF in cases	MAF in controls	χ^2	OR	P-vlaue
PRKCZ$(AC)_{24}$			16.417		0.219
rs7513222	0.313	0.328	0.212	0.934	0.646
rs908742	0.305	0.339	1.088	0.856	0.297
rs3107146	0.136	0.123	0.300	1.119	0.584
rs3107156	0.396	0.408	0.129	0.950	0.720
rs3107157	0.410	0.422	0.124	0.952	0.725
rs3128296	0.120	0.115	0.066	1.054	0.798
rs3753242	0.083	0.064	1.371	1.332	0.242
rs2503706	0.286	0.363	6.162	0.703	**0.013**
rs384726	0.070	0.075	0.056	0.938	0.812
rs385039	0.269	0.277	0.057	0.964	0.812
PRKCZ$(GT)_{17}$			9.485		0.487
rs12134873	0.002	0.003	0.008	0.882	0.929
rs12132341	0.016	0.019	0.140	0.827	0.708
rs262641	0.270	0.280	0.121	0.947	0.728
rs3128309	0.097	0.098	0.003	0.988	0.959
rs626479	0.249	0.266	0.302	0.916	0.582
rs262688	0.269	0.289	0.402	0.906	0.526
D1S243			18.579		0.234
rs262683	0.140	0.134	0.076	1.057	0.782
PRKCZ$(CA)_{13}$			18.194		0.11

MAF, minor allele frequency; Bold values signify P<0.05 , and these do not withstand multiple correction;
[a] Empirical Clump T1 significance (16 df); [b] Empirical Clump T1 significance (10 df); [c] Empirical Clump T1 significance (15 df) ; [d] Empirical Clump T1 significance (12 df); df, degrees of freedom.

4.5. Discussion

In the present study, I show nominally significant associations between genetic markers in the PRKCZ gene and BPAD. In the fine mapping study, two markers D1S243 and rs3128296 showed association with bipolar disorder with P-values of 0.037 (permuted P-value) and 0.040 respectively. A trend towards association with bipolar disorder was observed for SNP rs2503706 (P = 0.057). None of these SNPs remained significant after testing for multiple correction. Additionally, data from eight more SNPs from the UCL1 bipolar GWAS was added to the association analysis, resulting in the SNP rs3128309 (P = 0.060) showing a trend towards association. The markers outside of the PRKCZ gene failed to show any association with bipolar disorder.

Exploratory haplotypic association analysis of all three associated SNPs rs3128296, rs2503706 and rs3128309 resulted in a significant association with bipolar disorder (global empirical P = 0.004). All associated SNPs are intronic in the PRKCZ gene. They are, however, in the regulatory or promoter region of the two shorter isoforms of the gene with SNP rs2503706 being approximately 640bp from the 5'-UTR of one of the short isoforms (UCSC Human genome GRCh37/hg19 Feb 2009 assembly). Transcription factor binding analysis using TESS software (http://www.cbil.upenn.edu/) predicted that the mutant allele of the SNP rs2503706 bound to new and more transcription factors like Egr-1, CACCC-binding factor, gammaCAC1, gammaCAC2 and Sp1 compared to the wild-type allele. The SNPs in gene regulatory regions such as promoter and enhancer regions can alter gene expression by altering binding affinities to transcription factors (Chorley et al. 2008).

Thus, it is possible that these SNPs regulate expression of the alternative isoforms of PRKCZ.

In an earlier whole genome linkage scan of recurrent depressive disorder, there is suggestive evidence for linkage on chromosome 1p36 where the LOD score for female-only cases exceeded 3.00 (McGuffin et al. 2005). This region is slightly different from the chromosome 1 region, reported by Curtis et al. (Curtis et al. 2003) to be linked to bipolar and unipolar affective disorders, being approximately 5Mb from the linkage region reported in the McGuffin et al. (2005) study. Another genome-wide linkage analysis of 972 bipolar pedigrees of European ancestry failed to report any evidence of linkage at 1p (Badner et al. 2011). Interestingly, female-only analysis of all the microsatellite and SNP markers in the current study resulted in two previously associated SNPs rs3128296 and rs3128309 becoming more significantly associated, whereas only one of the other markers, rs2503706 showed association in male-only analysis, suggesting a 'sex' effect. The frequency of men and women suffering from neuropsychiatric disorders is all but proportionally distributed with women being more susceptible than men in the precipitation of depressive symptomatology (Pitychoutis et al. 2010). Some studies associate this sex-specific vulnerability to the pronounced genetic predisposition that women may present towards the development of depressive disorders. Additionally, clinical evidence has shown that women may have a better outcome when treated with SSRIs compared to tricyclic antidepressants suggesting that antidepressant response is also characterised by sex-specific manifestations (Kornstein et al. 2000; Khan et al. 2005). Therefore, the genetic makeup of an individual, in terms of 'sex', needs to be adequately addressed for the development of pharmacogenetic drugs in psychiatry.

Association studies of the PRKCZ gene with bipolar disorder have been done as part of GWAS but no meta-analyses have been carried out. MTHFR is another interesting candidate gene located on chromosome 1p36 which has been studied in bipolar disorder, but it lies almost 12Mb from the positive marker D1S243 in the UCL linkage study. GABRD is localised close to D1S243 and a recent study has reported association of GABRD gene with childhood-onset mood disorders (Feng et al. 2010). GABRD is at the distal end of the telomere which is prone to higher rates of recombination. The signal from the markers in this study does not seem to be from GABRD because there is almost negligible LD between the two genes PRKCZ and GABRD. Previous GWAS have failed to report any association between PRKCZ and bipolar disorder (WTCCC 2007; Baum et al. 2008).

PRKCZ (OMIM176982) encodes an atypical protein kinase C (PKCz) that is necessary for mediating axonal differentiation induced through Wnt signaling (Zhang et al. 2007). PKC zeta exhibits a kinase activity which is independent of calcium and diacylglycerol but not of phosphatidylserine. It has only a single zinc finger module and alternative splicing results in multiple transcript variants encoding different isoforms. A brain-specific, constitutively active isoform of the kinase, PKMzeta is necessary and sufficient for long-term potentiation at synapses that are responsible for memory (Hernandez et al. 2003; Sajikumar et al. 2005). Nerve injury triggers long-term plastic changes along sensory pathways, which in turn contribute to pain, perception, fear and memory. Li et al. (2010b) found that PKMzeta contributes to neuropathic pain by modulating excitatory synaptic transmission within the anterior cingulated cortex thereby, providing a potential new therapeutic target for chronic pain. A study by Navakkode et al. (2010) identified PKMzeta as a

dopamine-induced plasticity-related-protein, which exerted its action at activated synaptic inputs through processes of synaptic tagging. Further analysis and fine mapping of the PRKCZ gene in other samples is needed. In the future, DNA sequencing of PRKCZ may identify aetiological base pair changes affecting the expression or function of PRKCZ.

Chapter 5

Mutation screening of the GRM7 glutamate receptor gene in Bipolar disorder

5.1. Introduction

Metabotropic glutamate receptors are emerging novel targets for the disorders of brain associated with the dysfunction of the prefrontal cortex, a region critical for cognitive and emotional processes. The mGluRs are involved in LTP and LTD, two prominent forms of long-term synaptic plasticity in the hippocampus which play an important role in learning and memory as well as in disease development (Sanderson and Dell'Acqua 2011). mGluR7, encoded by GRM7, is a highly conserved group III mGluR and is highly expressed in the cingulate cortex, frontal cortex, amygdala, hippocampus and locus coeruleus, the brain regions associated with emotional learning (Shigemoto et al. 1997; O'Connor et al. 2010). Presynaptic mGluR7 modulates the release of L-glutamate and GABA and may also be involved in excitability levels in specific neuronal circuits, thereby, influencing different emotional states such as anxiety and depression as well as cognitive dysfunction associated with such disorders (Swanson et al. 2005).

GRM7 is located on 3p26.1 and previous linkage studies have implicated chromosomal region 3p in bipolar disorder (Kelsoe et al. 2001; Fallin et al. 2005; Etain et al. 2006). A recent comparative linkage meta-analysis of bipolar disorder genome-wide linkage datasets using GSMA and MSP methods also found evidence of association at 3p25.3-3p22.1 under GSMA narrow (P = 0.0099) and broad (P = 0.0060) models (Tang et al. 2011). Independent studies by Breen et al. (2011b) and Pergadia et al. (2011) reported evidence for genetic linkage between MDD and the choromosome 3p26-p25 locus, the region containing GRM7, with LOD scores of 4.0

and 4.14 respectively. However, both the studies failed to replicate their initial findings in their respective replication samples.

Genome-wide significant associations have not been reported for GRM7 in any bipolar GWAS. However, some of these studies have reported modest support for association with GRM7 (WTCCC 2007; Sklar et al. 2008; Sklar et al. 2011). In the Psychiatric GWAS Consortium (PGC) bipolar study (Figure 5.1), the most associated GRM7 marker was rs11915371 ($P = 1.13 \times 10^{-4}$) (Sklar et al. 2011) (Figure 5.1). A recent GWAS of personality traits in bipolar patients found association for a GRM7 SNP with the neuroticism-anxiety scale of the Zuckerman-Kuhlman Personality Questionnaire (rs13080594, $P = 7.68 \times 10^{-7}$). Scott et al. (2009) performed a meta-analysis of non-overlapping samples from three bipolar disorder GWAS (NIMH/Pritzker, GSK and WTCCC) and found significant association at 3p21, which was one of the three regions showing association at the level of $P \approx 10^{-7}$. GWAS of major depression along with meta-analyses have also demonstrated association with GRM7 (Sullivan et al. 2009; Muglia et al. 2010; Shi et al. 2011; Shyn et al. 2011). A meta-analysis of major mood disorders, including bipolar disorder and MDD samples from GAIN-MDD, NIMH- BP, WTCCC, Step-BD and an unpublished German bipolar disorder study found genome-wide significant evidence for SNPs in 3p21.1 region (rs2251219, $P = 3.63 \times 10^{-8}$) (McMahon et al. 2010). Breen et al. (2011a) reanalysed the data from the McMahon et al. (2010) study and performed further replication in additional MDD samples from three studies: the RADIANT (Recurrent depressive disorder and treatment study), deCODE Genetics, and the MDD2000+ community sample (Breen et al. 2011a).

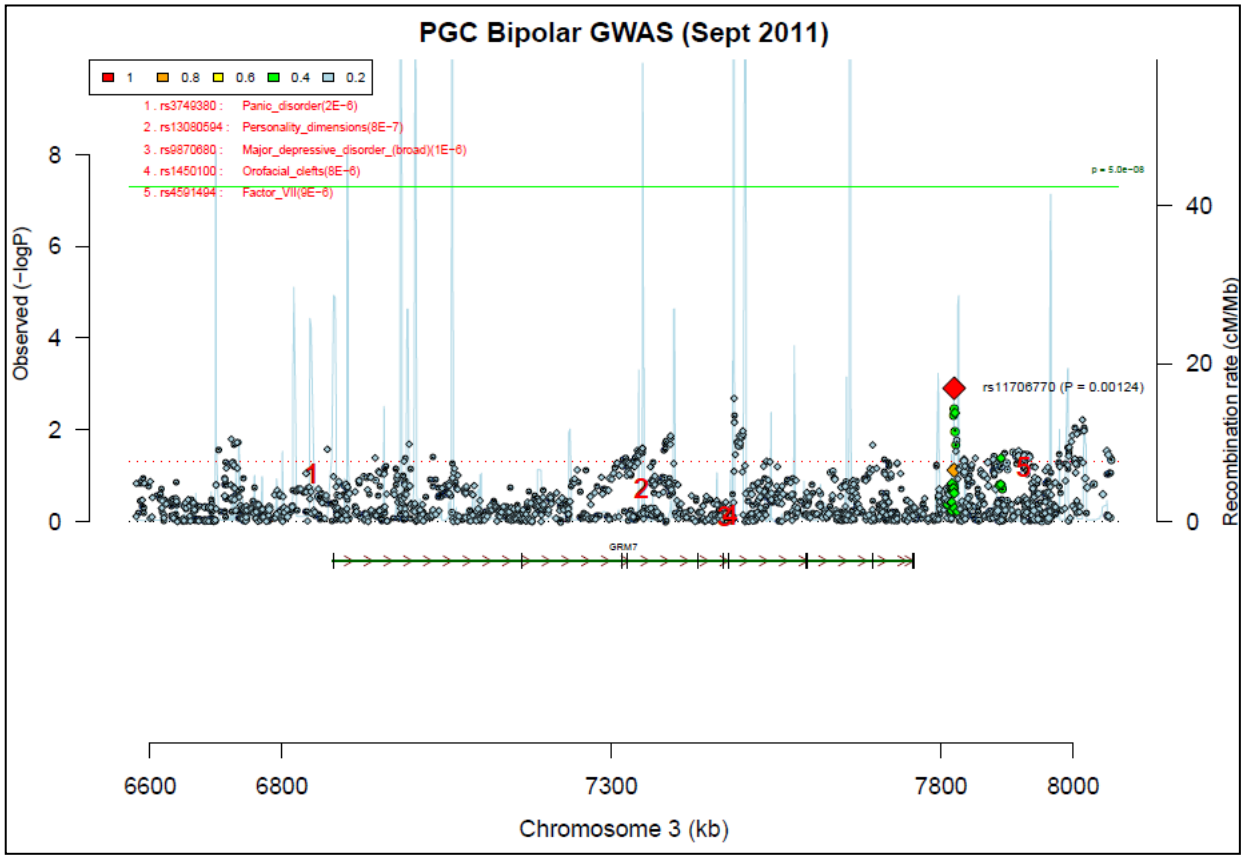


Figure 5.1 Association of GRM7 SNPs in the PGC-Bipolar GWAS generated using Ricopili.

Random-effect meta-analysis of the McMahon et al. (2010) study showed that the SNP rs2251219 was a genome-wide significant locus for bipolar disorder and not MDD whereas the replication study failed to find any association between rs2251219 and MDD (P = 0.14) (Breen et al. 2011a).

The contribution of CNVs in the psychiatric disorders has been appreciated only recently with the implementation of array technologies that enable genome-wide determination of CNVs. In a systematic study of CNVs in bipolar disorder, although an increase in overall CNV load was not reported, a nominally significant increase in singleton CNVs in bipolar disorder cases compared to controls was found (Zhang et al. 2009). GRM7 was one of the genes affected by singleton deletions in this study and the results have been discussed in detail in section 1.2.7. Deletions and duplications in GRM7 were found in UCL1 bipolar sample as well (McQuillin et al. 2011). Rare CNVs occurring within GRM7 have been documented in patients with other psychiatric disorder patients such as mood disorder (Saus et al. 2010) and schizophrenia (Walsh et al. 2008).

Structurally, mGluR7 is composed of a large N-terminal extracellular domain of approximately 590 amino acids, a 7 transmembrane (TM) domain of around 260 amino acids and an intracellular C-terminus consisting of 30-80 amino acids similar to the other group III mGluRs (Figure 5.2). The VFT domain is responsible for the binding of L-glutamate and also contains a highly evolutionarily conserved orthosteric binding site. On binding to L-glutamate or an orthosteric agonist, the VFT domain undergoes a conformational change which is then propagated via an extracellular cysteine rich region to the 7 TM and cytoplasmic domains resulting in

G-protein activation, second messenger formation and ion channel modulations (Niswender and Conn 2010). L-AP4 was the first prototypic group III agonist used for defining the physiological functions of these receptors *in vivo*, although it was not subtype specific. AMN082 is the first selective allosteric agonist for mGluR7 that does not require the presence of any orthosteric ligand and has been extensively used in the pharmacological studies. Interestingly, AMN082 activates mGluR7 in a pathway-selective manner such that only a subset of the receptor's signal transduction pathways is activated (Flor and Acher 2012).

A growing body of evidence suggests a strong relationship between mGluR7 function and the action of antidepressants and mood-stabilizers. mGluR7 mRNA was upregulated in the dentate gyrus of Wistar Kyoto (WKY) rats, animal models of depression (O'Mahony et al. 2010). In animal models, AMN082 decreased

immobility in standard experimental paradigms like forced swim test and tail suspension tests whereas these antidepressant-like effects were blocked when GRM7-deficient mice were used (Palucha et al. 2007). Furthermore, chronic treatment with citalopram decreased mGluR7 immunoreactivity in rat hippocampus and frontal cortex (Wieronska et al. 2007). Fabbri et al. (2012) investigated the modulation of early antidepressant efficacy by glutamatergic gene variants in the STAR*D sample and found that a GRM7 SNP rs1083801 was associated with an early response to treatment under a recessive model (GG genotype observed in 14.34% of early responders vs 5.25% of late responders). Also the effect of the rs1083801 variant was more pronounced in their female subsample, confirming previous results which suggested that a gender-specific effect may be not infrequent in antidepressant pharmacogenetics (Pitychoutis et al. 2010)

In summary, good evidence from genetic, pharmacological and animal studies suggests a connection between GRM7 and bipolar disorder and other neuropsychiatric disorders.

5.2. Hypothesis

GRM7 was significantly associated in the UCL1 bipolar GWAS. Analysis of deletions and duplications in the GWAS data revealed the presence of CNVs in GRM7, thereby strengthening the connection between GRM7 and bipolar disorder in the UCL1 sample. I hypothesised on the basis of these results and results from previous published linkage, GWA, CNV and expression studies that GRM7 is a strong susceptibility gene for bipolar disorder.

5.3. Aims

The aim of this study was to identify novel variants in GRM7 associated with bipolar disorder by resequencing GRM7 in bipolar cases selected for inheriting a disease haplotype in the UCL1 GWAS. On the detection of novel associated functional variants, further characterisation of the variant would be undertaken to understand the mechanism of action of the mutation in the disorder. Also, the CNVs identified in the UCL1 sample would be validated using QRT-PCR.

5.4. Results

5.4.1. Association analysis

In the UCL1 GWAS (Sklar et al. 2008), eight SNPs in GRM7 showed allelic association with bipolar disorder and the most significant SNP associated at this locus was rs1508724 (P = 0.0014) (Appendix V). Also, GRM7 has been found to be associated in other bipolar disorder samples. In the WTCCC sample, 35 SNP markers within GRM7 were found to be associated with bipolar disorder and in the STEP-BD sample a further 10 SNPs were associated. Furthermore, only one SNP, rs13070476, was associated with bipolar disorder in the UCL and WTCCC sample suggesting the presence of heterogeneity even within populations of similar ancestry.

I failed to replicate the association of the three most significantly associated UCL1 GWAS SNPs, rs1508724, rs11710946, and rs6769814 in the UCL2 sample (Table 5.1). On combining the UCL1 and UCL2 genotype data for the three SNPs using PLINK, two SNPs, rs1508724 and rs6769814, remained significant (Table 5.2). I also performed a meta-analysis to assess the association of these SNPs in the

combined sample keeping account of the heterogeneity. Only one SNP rs6769814 remained significant as a result of meta-analysis (Cochrane Q-statistic *P*-value of 0.2166, Fixed effects model; *P*-value= 0.03506 and OR= 1.1544) (Table 5.2). The heterogeneity index for the other two SNPs rs1508724 and rs11710946 was significantly high at 76.83 and 81.67 respectively. This suggested that the UCL1 and UCL2 samples are not completely homogeneous although both were selected based on strict ancestry-based questionnaires.

Table 5.1 SNP association results with GRM7 in the UCL bipolar samples

SNP	Position (NCBI35/hg19)	UCL1 Allele counts (MAF)		P-value	OR	UCL2 Alleles counts (MAF)		P-value	OR
		BP	CON			BP	CON		
rs1508724	7241745	A 344 (0.34) G 668 (0.66)	A 280 (0.28) G 740 (0.72)	**0.001**	1.36	A 297 (0.29) G 723 (0.71)	A 270 (0.29) G 652 (0.71)	0.979	1.00
rs11710946	7246241	A 359 (0.36) G 653 (0.64)	A 430 (0.42) G 590 (0.58)	**0.002**	0.75	A 399 (0.40) G 615 (0.60)	A 351 (0.38) G 567 (0.62)	0.634	1.05
rs6769814	7251433	G 349 (0.35) A 657 (0.65)	G 297 (0.29) A 721 (0.71)	**0.008**	1.29	G 324(0.31) A 716 (0.69)	G 275 (0.30) A 639 (0.70)	0.571	1.06

MAF, minor allele frequency; BP, bipolar; CON, control; OR, odds-ratio

Table 5.2 Combined analysis of GRM7 SNPs in UCL1 and UCL2 samples

SNP	Position (NCBI35/hg19)	Combined analysis Allele counts (MAF)				Meta-analysis Fixed effects model		Random effects model		Q	I^2
		BP	CON	*P*-value	OR	*P* Value	OR	*P*-value	OR		
rs1508724	7241745	A 623 (0.31) G 1373 (0.69)	A 550 (0.28) G 1392 (0.72)	**0.043**	1.15	**0.032**	1.16	0.32	1.15	0.038	76.83
rs11710946	7246241	A 751 (0.38) G 1241 (0.62)	A 781 (0.40) G 1157 (0.60)	0.092	0.90	0.080	0.89	0.477	0.90	0.020	81.67
rs6769814	7251433	G 652 (0.33) A 1354 (0.67)	G 572 (0.30) A 1360 (0.70)	**0.045**	1.15	**0.035**	1.15	0.091	1.15	0.217	34.5

MAF, minor allele frequency; BP, bipolar; CON, control; OR, odds ratio; Q, p-value for Cochrane's Q statistic; I^2, I^2 heterogeneity index (0-100)

5.4.2. Sequencing

Detection of novel variants in GRM7 associated with bipolar disorder was pursued by resequencing all the exons in all the GRM7 isoforms (Figure 5.3), exon/intron boundaries, ~3 kb of the promoter region of the isoforms ENST00000486284 and ENST00000463676, and 1 kb of the promoter region of the isoform ENST00000458641, 5'-UTR and 3′-UTR of the GRM7 gene in 32 haplotype selected cases and 32 random controls. In total, I found 71 previously published and 10 novel mutations in GRM7 (Table 5.3). Most of the published SNPs were of similar frequency in the sequenced cases and controls (Table 5.3), and therefore, these were not genotyped. Also, SNPs that were in complete LD with the UCL1 bipolar GWAS SNPs were not genotyped in the complete sample. Of the novel mutations, three were single base substitutions, two STRs, and five insertion deletions (Table 5.3). Bioinformatic analysis was carried out for the novel SNPs to determine their effect on the structure and/or function of GRM7. Finally, based on an increased frequency in cases compared to controls in the sequencing panel, 18 SNPs were further genotyped in the complete UCL sample (UCL1 and UCL2). Among these, only rs56173829 was significantly associated with bipolar disorder in the UCL sample (P = 0.035; OR = 0.4829) (Table 5.4), although the minor allele was less common in cases than controls. The rare allele of the SNP rs56173829 was present in 25/1821 alleles (1.4%) in control subjects versus 12/1810 (0.66%) in bipolar cases (Table 5.4). The SNP rs56173829 is present in the 3′-UTR of the long isoform of GRM7 and may play an important role in microRNA binding.

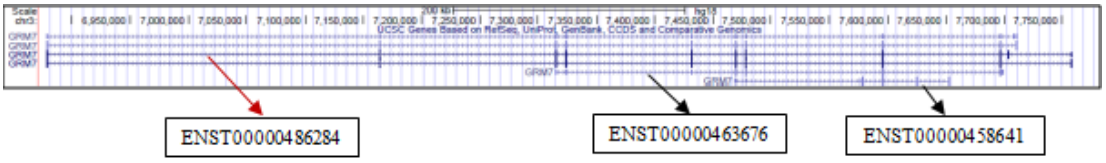


Figure 5.3 Isoforms of GRM7

(Generated using UCSC genome browser Feb. 2006 assembly). The main isoform of GRM7 is shown by the red arrow and the other two shorter isoforms are shown by black arrows.

Additionally, some SNPs were present only in cases, rs138571076 (P = 0.0826) in three bipolar disorder patients, GRM7_3f_7313045 in two bipolar disorder patients and GRM7_nPb_7467774 in only one bipolar disorder patient. One of the novel SNPs, GRM7_9c_7698252 was found in only one control. All these SNPs are intronic in the GRM7 gene. The SNP rs712774 was predicted to introduce a splice acceptor site using the Splice Site Prediction by Neural Network (http://www.fruitfly.org/seq_tools/splice.html) but was not associated with bipolar disorder in the UCL sample (P=0.853). Imputation using European sample data for the GRM7 exonic SNPs and SNPs in the intron/exon boundary in the 1000 genomes sequencing panel did not reveal any synonymous or functional SNPs that were associated with bipolar disorder (Table 5.5).

Table 5.3. SNPs found by resequencing GRM7

Region	Coding/non-coding	SNP ID	Position (NCBI35/hg19)	Alleles	MAF
Promoter	promoter	3_6900524_1000G	6900524	CT/C	0.18*
Promoter	promoter	rs183111337	6900527	T/C	0.01*
Promoter	promoter	rs62237226	6900528	C/T	0.19*
Promoter	promoter	rs115717493	6900573	G/A	0.01*
Promoter	promoter	rs340653	6900663	C/T	0.3*
Promoter	promoter	rs163422	6900744	C/T	0.08*
Promoter	promoter	rs371097	6900910	C/A	0.1*
Promoter	promoter	rs371841	6900912	C/T	0.29*
Promoter	promoter	rs62237227	6901071	C/T	0.19*
Promoter	promoter	rs339807	6901090	G/A	0.26
Promoter	promoter	rs63470962	6901176	C/CAAG	0.09*
Promoter	promoter	rs114774914	6901783	A/T	0.03*
Promoter	promoter	3_6901914_1000G	6901976	AAGA/-	NA
Promoter	promoter	rs62237228	6902167	A/G	0.19*
Promoter	promoter	3_6902624_1000G	6902624	T/TC	0.25*
Exon 1	coding-synonymous	rs3749380	6903297	C/T	0.45
Intron	intronic	rs342034	6903601	A/G	0.04
Intron	intronic	rs3828472	7188006	G/T	0.33
Intron	intronic	rs9868134	7188033	T/C	0.22
Intron	intronic	rs3749450	7188063	G/A	0.33
Intron	intronic	rs3749449	7188072	A/G	0.18
Intron	intronic	rs3749448	7188116	G/A	0.18
Exon 2	coding-synonymous	rs35106713	7188180	T/C	0.02
Intron	intronic	rs140139253	7188396	T/C	0.001*
Promoter of ENST00000463676 isoform	intronic	rs6768211	7337515	G/T	NA
Promoter of ENST00000463676 isoform	intronic	rs185220306	7337786	T/G	0.01*
Promoter of ENST00000463676 isoform	intronic	rs9850091	7337788	G/T	0.44*
Promoter of ENST00000463676 isoform	intronic	rs116361217	7337904	C/T	0.07*
Promoter of ENST00000463676 isoform	intronic	GRM7_3f_7313045	7338045	G/A	NA
Promoter of ENST00000463676 isoform	intronic	rs7638548	7338112	G/A	0.37*
Promoter of ENST00000463676 isoform	intronic	rs78210196	7338179	C/A	0.01*
Promoter of ENST00000463676 isoform	intronic	rs7640834	7338353	G/A	0.37*
Promoter of ENST00000463676 isoform	intronic	rs61060115	7338521	-/A	NA *
Promoter of ENST00000463676 isoform	intronic	rs75618003	7338776	C/T	0.18*
Promoter of ENST00000463676 isoform	intronic	rs138571076	7338985	A/G	0.01*
Promoter of ENST00000463676 isoform	intronic	rs138761134	7339328	ATGTT/A	0.01*
Promoter of ENST00000463676 isoform	intronic	rs78677232	7339409	C/T	0.18*
Promoter of ENST00000463676 isoform	intronic	rs192193072	7339471	A/G	0.002*
Promoter of ENST00000463676 isoform	intronic	rs2291867	7340164	C/T	0.39

Region	Coding/non-coding	SNP ID	Position (NCBI35/hg19)	Alleles	MAF
Intron	affects splicing	rs712774	7456675	C/T	0.26
Promoter of ENST00000458641 isoform	intronic	GRM7_nPb_7467774	7492774	C/T	NA
Promoter of ENST00000458641 isoform	intronic	rs712792	7492779	C/T	0.26
Promoter of ENST00000458641 isoform	intronic	nPa_7493030	7493030	-/C indel	NA
Promoter of ENST00000458641 isoform	intronic	rs73015547	7493347	C/T	0.13*
Promoter of ENST00000458641 isoform	intronic	rs712793	7493351	A/G	0.34
Exon 6	coding-non-synonymous	rs2229902	7494417	A/T	0.44
Exon 8	coding-synonymous	rs1485174	7595828	A/G	0.29
Intron	intronic	nex4_7600830	7600830	T indel	NA
Intron	intronic	rs1143739	7601039	G/T	0.26
Exon of ENST00000458641 isoform	3' UTR	rs1965222	7603194	C/T	0.13
Exon of ENST00000458641 isoform	3' UTR	rs3828429	7603367	C/G	0.01*
Exon 8	coding synonymous	rs34373930	7620168	A/G	0.3*
Exon 8	coding synonymous	rs7614915	7620382	C/T	0.26
Exon 8	coding synonymous	rs1485175	7620789	C/T	0.43
Intron	intronic	rs1485173	7621093	A/T	0.3*
Intron	intronic	rs1485172	7621158	C/T	0.43
Intron	intronic	rs2139187	7649591	G/T	0.34*
Exon of ENST00000458641 isoform	frameshift	nex7_7649745	7649745	Insertion of T	NA
Intron	intronic	rs114582026	7649815	G/T	0.16*
Intron	intronic	rs115155482	7649819	C/T	0.3*
Exon of ENST00000458641 isoform	3' UTR	rs7611935	7678010	A/G	0.34*
Exon of ENST00000458641 isoform	3' UTR	rs7612048	7678090	A/G	0.3
Intron	intronic	rs17047754	7678134	C/T	0.3
Intron	intronic	nex8_7678357	7678357	G/T	NA
Intron	intronic	rs162802	7721613	T/C	0.5
Exon 9	3' UTR	rs2280739	7721997	G/A	0.06
Exon 9 3′-UTR	3' UTR	rs10514663	7723179	G/A	0.08
Exon 9 3′-UTR	3' UTR	GRM7_9c_7698252	7723252	G/A	NA
Exon 9 3′-UTR	3' UTR	rs162801	7723744	G/A	0.5*
Exon 9 3′-UTR	3' UTR	rs140995942	7723896	G/A	0.01*
Intron	intronic	STR	7724024	Simple T repeat	
Exons 11 and 12	3' UTR	rs5846531	7732725	-/T	0.02*
Intron	intronic	rs2279840	7735770	A/G	0.2
Intron	intronic	Ex14_Indel	7736114	18 bp insertion	NA
Intron	intronic	rs162777	7736246	C/T	0.26
Exon 15, 3′-UTR	3' UTR	rs75721571	7782367	C/G	0.07*
Exon 15, 3′-UTR	3' UTR	rs9826579	7782371	C/T	0.24*
Exon 15, 3′-UTR	3' UTR	rs56173829	7782494	A/T	0.01*

Region	Coding/non-coding	SNP ID	Position (NCBI35/hg19)	Alleles	MAF
Exon 15, 3′-UTR	3' UTR	rs17726576	7782551	C/T	0.02
Exon 15, 3′-UTR	3' UTR	rs60445645	7782703	-/T	NA
Outside the gene, 3′ end	outside the gene	rs150288969	7783347	G/A	0.01*

MAF, minor allele frequency; NA, not available; *MAF in European samples derived from the 1000 genomes project

Table 5.4 Tests of allelic association of GRM7 SNPs, found by resequencing, in the UCL sample

SNP	Position (NCBI35/ hg19)	Min/Maj Allele	MAF	Genotype counts	Allele counts	P-value	OR (95% CI)
rs114774914	6901783						
Cases		A/T	0.014	0/26/897	26/1820	0.268	0.75 (0.45-1.25)
Controls			0.019	0/35/901	35/1837		
3_6901914	6901914						
Cases		C/CTCTT	0.022	1/39/885	41/1809	0.766	1.07 (0.69-1.67)
Controls			0.021	0/38/878	38/1794		
rs62237228	6902167						
Cases		C/A	0.190	30/290/601	350/1492	0.228	1.11 (0.94-1.31)
Controls			0.175	27/267/625	321/1517		
rs342034	6903601						
Cases		A/G	0.040	4/64/837	72/1738	0.351	1.18 (0.83-1.67)
Controls			0.034	1/60/852	62/1764		
rs35106713	7188180						
Cases		C/T	0.008	0/14/915	14/1844	0.188	0.64 (0.33-1.25)
Controls			0.012	0/22/914	22/1850		
rs140139253	7188396						
Cases		T/C	0.004	0/7/927	7/1861	0.787	1.16 (0.39-3.47)
Controls			0.003	0/6/924	6/1854		
GRM7_3f_7313045	7338045						
Cases		G/A	0.001	0/2/927	2/1856	0.157	NA
Controls			0.000	0/0/932	0/1864		
rs138571076	7338986						
Cases		G/A	0.002	0/3/926	3/1855	0.083	NA
Controls			0.000	0/0/932	0/1864		
rs192193072	7339471						
Cases		A/G	0.001	0/2/926	2/1854	0.156	0.33 (0.07-1.65)
Controls			0.003	0/6/919	6/1844		
rs712774	7456675						
Cases		C/T	0.263	81/392/580	554/1552	0.853	0.99 (0.86-1.14)
Controls			0.266	65/353/491	483/1335		
GRM7_nPb_7467774	7492774						
Cases		T/C	0.001	0/1/923	1/1847	0.321	NA
Controls			0.000	0/0/909	0/1818		
rs2229902	7494417						
Cases		T/A	0.447	185/445/282	815/1009	0.227	0.92 (0.81-1.05)
Controls			0.467	213/442/275	868/992		
rs1965222	7603194						
Cases		T/C	0.128	18/235/805	271/1845	0.713	1.04 (0.86-1.25)
Controls			0.124	11/210/713	232/1636		
GRM7_9c_7698252	7723252						
Cases		A/G	0.000	0/0/924	0/1848	0.319	NA
Controls			0.001	0/1/928	1/1857		
rs140995942	7723896						
Cases		A/G	0.013	0/24/909	24/1842	0.371	0.78 (0.46-1.34)
Controls			0.016	0/30/885	30/1800		
rs56173829	7782494						
Cases		A/T	0.007	0/12/899	12/1810	**0.035**	0.48 (0.24-0.96)
Controls			0.014	0/25/898	25/1821		

SNP	Position (NCBI35/ hg19)	Min/Maj Allele	MAF	Genotype counts	Allele counts	P-value	OR (95% CI)
rs17726576	7782551						
Cases		T/C	0.027	1/47/874	49/1795	0.763	1.07 (0.71-1.6)
Controls			0.025	0/46/874	46/1794		
rs150288969	7783347						
Cases		A/G	0.008	0/14/910	14/1834	0.682	1.18 (0.54-2.55)
Controls			0.006	0/12/918	12/1848		

Min, minor; Maj, major; MAF, minor allele frequency; OR, odds-ratio; CI, confidence interva

Table 5.5 Significantly associated GRM7 SNPs (P ≤ 0.05) using imputation from 1000G GRM7 SNPs

SNP	Position	A1	A2	GENO	NPRX	INFO	F_A	F_U	OR	P
rs10510354	7252190	C	G	0.408	2	0.479	0.375	0.408	0.87	0.003
rs1352411	7799831	A	G	0.389	2	0.423	0.079	0.092	0.839	0.017
rs340659	6906622	C	T	0.408	2	0.54	0.065	0.052	1.27	0.021
rs11708019	7229619	G	A	0.408	2	0.441	0.343	0.365	0.908	0.027
rs9823996	7244509	G	C	0.408	2	0.475	0.316	0.294	1.11	0.028
rs9820417	7104786	G	A	0.387	2	0.438	0.238	0.258	0.899	0.033
rs17288442	6866379	A	C	0.407	2	0.479	0.302	0.324	0.903	0.034
rs13070476	7259313	C	T	0.403	2	0.45	0.226	0.245	0.901	0.034
rs1508724	7241745	A	G	0.763	2	0.82	0.311	0.284	1.14	0.035
rs6769814	7251433	G	A	0.763	2	0.802	0.324	0.297	1.13	0.035
rs17047149	7380288	G	A	0.407	2	0.353	0.007	0.011	0.623	0.052
rs11717750	7464013	T	C	0.408	2	0.846	0.507	0.479	1.12	0.053

A1, Allele 1; A2, Allele 2; GENO, Genotyping for the reference SNP; NPRX, Number of proxy SNPs used to tag reference SNP; INFO, Information metric for each reference SNP; F_A; Reference SNP allele frequency in cases; F_U; Reference SNP allele frequency in controls; OR, odds-ratio; P, Asymptotic p-value for test of association. All the positions are correct according to NCBI35/hg19 assembly

5.4.3. Haplotype analysis

I performed haplotype analysis on the combined data including all the SNPs genotyped in GRM7 (UCL1 GWAS SNPs, the three GWAS SNPs in UCL2, and SNPs found by mutation screening in GRM7) in the present study using Haploview (Figure 5.4) (Barrett et al. 2005). The most significant haplotype (HAP_GRM7) was associated with normal controls comprising SNPs rs1400166, rs2875257, rs10510353, rs11708019, rs1963265, rs1508724, rs9823996, rs11710946 and rs6769814 (P = 0.007) (2.2% and 4.3% in cases and controls respectively). Particularly, the SNPs driving the association of this haplotype were the three most significantly associated SNPs in UCL1 GWAS, rs1508724, rs11710946 and rs6769814. All the other associated haplotypes, including HAP_GRM7, did not survive permutation testing.

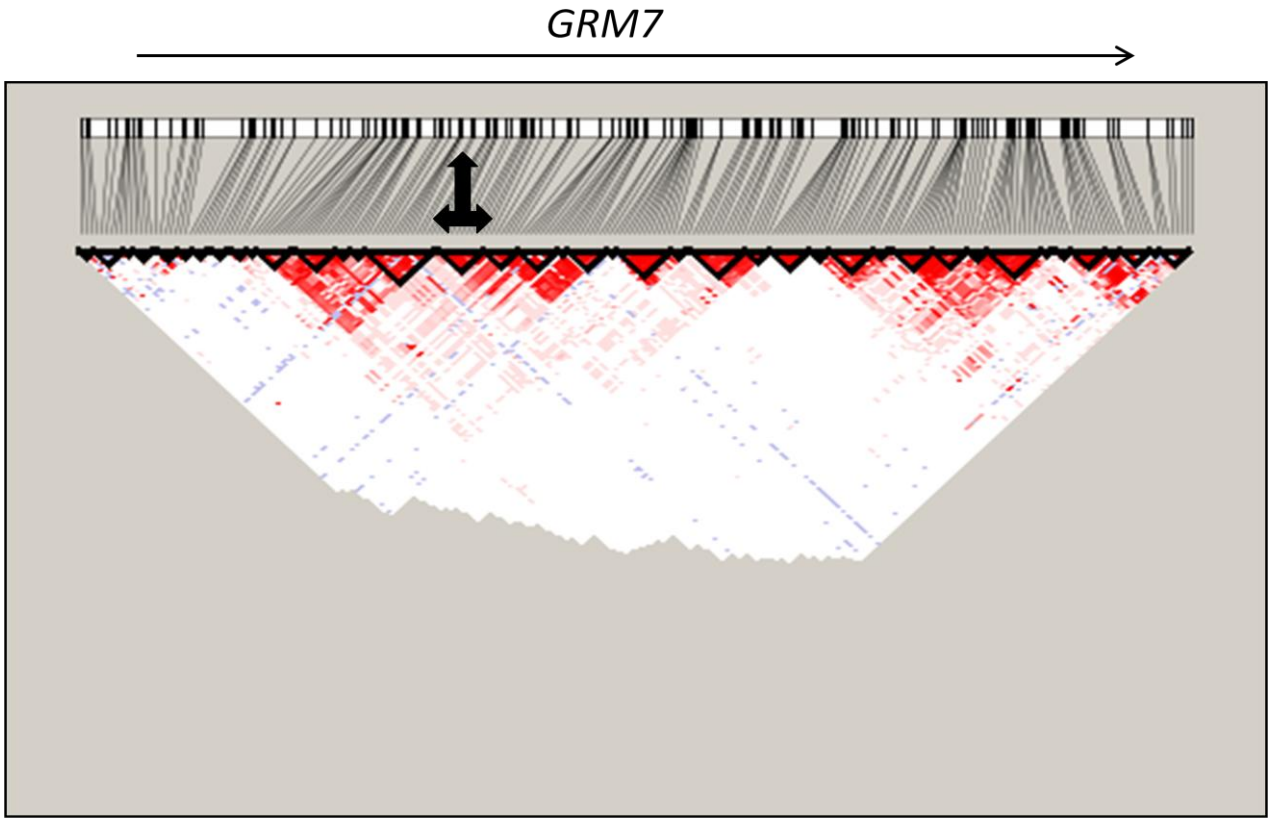


Figure 5.4 LD structure between GRM7 SNPs genotyped in the UCL research sample

The most significantly associated haplotype that also contains the significant GWAS SNPS is marked with a a three sided arrow.

5.4.4. CNV validation

McQuillin et al. (2011) performed an analysis of the deletions and the duplications in the UCL1 bipolar research sample and found that the overall rate of CNVs was significantly lower in the cases compared to controls. Both deletions and duplications of size >100 kb were detected in GRM7 in the UCL1 sample (Table 5.6). Singleton deletions were present in two cases and one control whereas duplication was present in only one case but not in any controls (Figure 5.5).

Table 5.6 CNVs in GRM7 detected in UCL1 samples

Location	CNV	Sample
Chr3:7347535 - 7733108	Duplication	Case
Chr3:7277350 - 7558603	Deletion	Case
Chr3:7001273 - 7107443	Deletion	Case
Chr3:7040496 - 7208936	Deletion	Control

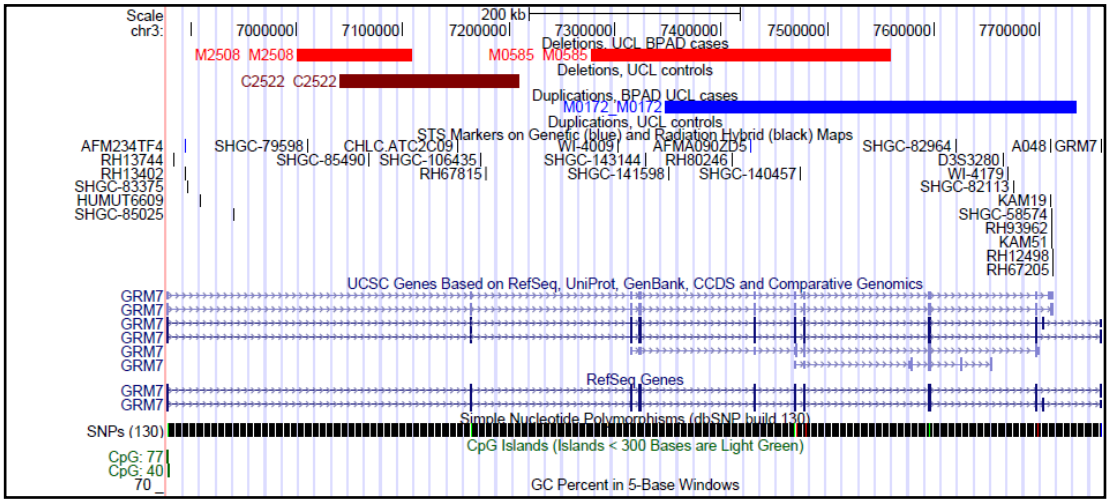


Figure 5.5 CNVs occurring in GRM7 in UCL1 GWAS sample

Figure generated using data from (McQuillin et al. 2011). UCSC genome browser March 2009 assembly

I selected a deletion and a duplication involving a common genomic region for the validation using QRT-PCR assay. TaqMan® RNase P Copy Number Reference (CNR) Assay, Human (Applied Biosystems, UK) was used for the CNV validation. In the RNase P CNR assay, four individuals, one each with a deletion and a duplication, and two CNV negative control individuals were tested using primer pairs designed to detect duplication and deletion in the same assay. Replicable results were obtained with all the four primer pairs tested and the CNVs were confirmed in the respective UCL samples (Table 5.7) (Figure 5.6 and Figure 5.7).

Table 5.7 CNV validation using RNaseP copy number reference assay

Sample	Target	Reference	Mean Cp	Cp Error	Mean Cp	Cp Error	Target/Ref	Target/Ref Error
Duplication	GRM7	RNaseP	26.65	0.121	28.10	0.205	2.729	0.4022
Control	GRM7	RNaseP	27.41	0.046	28.47	0.118	2.083	0.1613
Control	GRM7	RNaseP	27.24	0.177	28.37	0.071	2.192	0.2593
Deletion	GRM7	RNaseP	27.39	0.165	27.89	0.079	1.415	0.1674

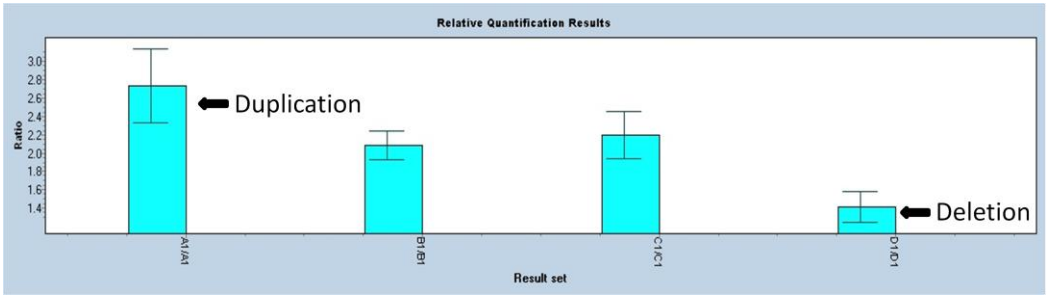


Figure 5.6 CNV validation with one primer pair using TaqMan® RNase P copy number reference assay

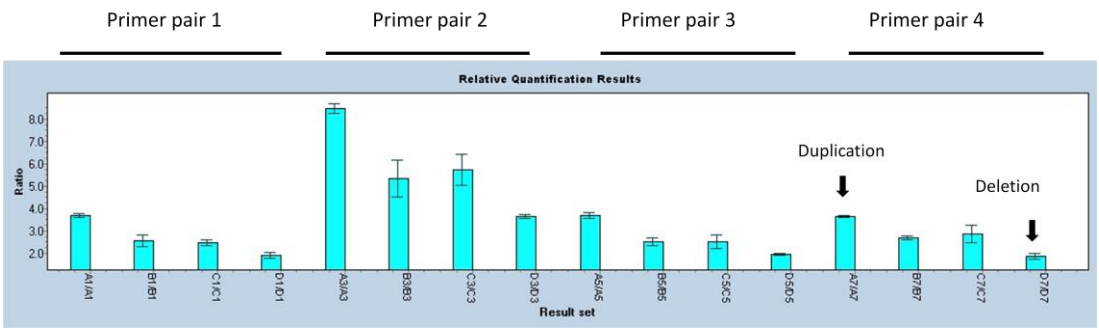


Figure 5.7 CNV validation with four different primer pairs using TaqMan® RNase P copy number reference assay

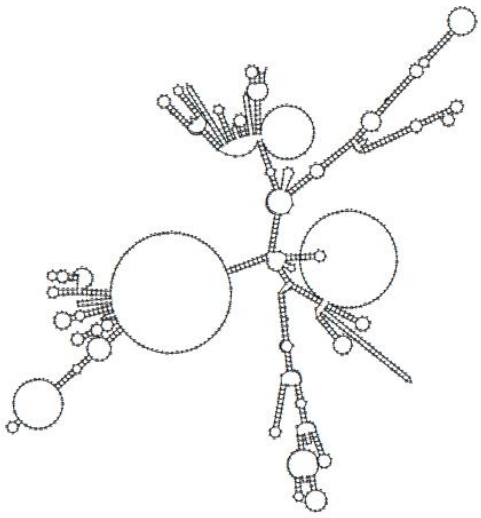
5.5. Discussion

The association of GRM7 in the UCL1 sample with bipolar disorder was not replicated in the UCL2 sample. This might be explained by the presence of heterogeneity even within a single ancestrally selected group of bipolar cases. The population based case-control samples are useful in detecting the sporadic cases containing common risk variants conferring susceptibility to a complex disorder like bipolar disorder, thereby sometimes failing to identify the rare variants associated with the disorder. Another limitation may be that the replication sample was not large enough to detect the expected level of association with GRM7, having less than 80% power to detect the association for all the three SNPs (refer to the graph in Appendix VI). Furthermore, the presence and the proportion of several other disease haplotypes in GRM7 causing bipolar disorder in different populations makes it more difficult to replicate findings. Another possibility is that the previous association results are false positives. Nonetheless, two of the previously associated SNPs, rs1508724 and rs6769814, in UCL1 remained significant when I combined genotype data of the UCL1 and UCL2 samples using PLINK. Resequencing of GRM7 in the haplotype selected cases detected a 3′-UTR variant, rs56173829, with the minor

allele significantly more common in controls compared to cases. Variants in the 3′-UTR play an important role in the regulation of translation by means of microRNA (miRNA) binding sites. Additionally, I found two novel and one previously published case-only base-pair changes in the UCL sample, however, their frequencies were extremely low.

Alternative splicing in GRM7 results in five isoforms characterised by different C-terminals. The rs56173829 variant is located in the 3′-UTR of the long isoform of GRM7 (NM_181874; ENST00000486284). RNA webserver was used to predict the secondary structure of the wild-type and mutant 3′-UTR containing SNP rs56173829. The RNA webserver (http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi) produces two structures - MFE and centroid secondary structure. The MFE structure of an RNA sequence is the secondary structure that contributes a minimum of free energy (MFE) whereas the centroid structure of the same is the secondary structure with minimal base pair distance to all other secondary structures in the Boltzmann ensemble. The predicted centroid secondary structures for the wild-type 3′-UTR was altered by the introduction of the SNP rs56173829, however, the MFE structures for both were not dissimilar (Figure 5.8 and Figure 5.9). The cis-regulatory elements in the 3′-UTR of mRNAs have been reported to influence translation and cause genetic diseases. Translation de-regulation and disorders resulting from mutations affecting the termination codon, polyadenylation signal and secondary structure of 3′-UTR of mRNA have also been documented (Chatterjee and Pal 2009). A correlation between the functionality of 3′-UTR variants and alterations in the predicted mRNA secondary structure was reported in a study by Chen et al. (2006) on 83 disease associated 3′-UTR variants of various human mRNAs.

(a)



(b)

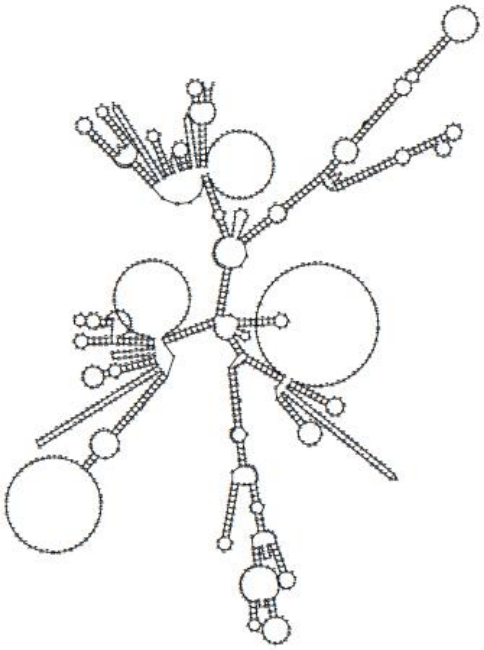
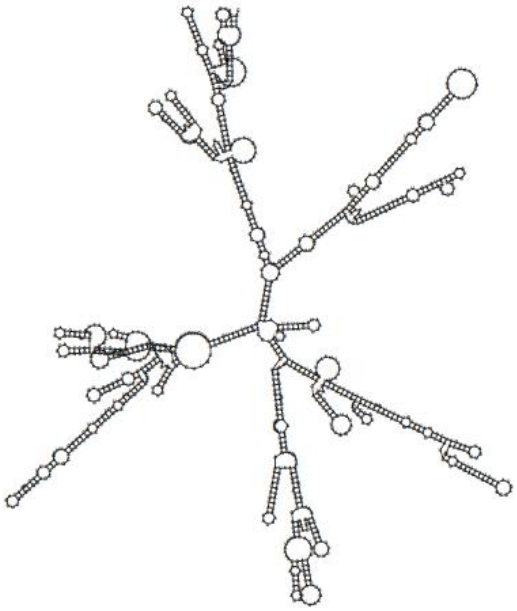


Figure 5.8 Centroid secondary structure for the wild-type (a) and mutant (b) 3'-UTR containing SNP rs56173829 predicted using RNAfold webserver

(a)



(b)

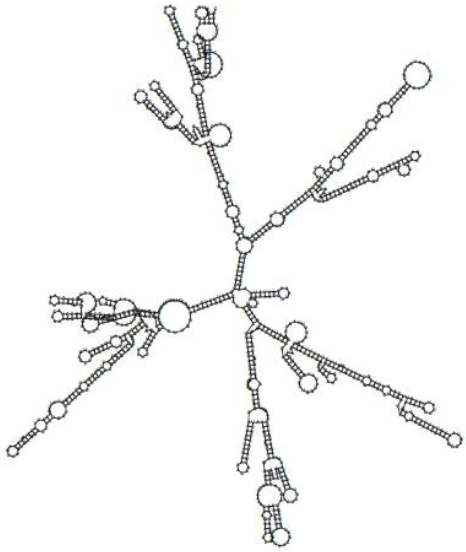


Figure 5.9 MFE secondary structure for the wild-type (a) and mutant (b) 3'-UTR containing SNP rs56173829 predicted using RNAfold webserver

miRNAs are 21-22 nucleotides long, single-stranded and bind to 3′-UTRs of particular mRNAs through partially complementary sequences and prevent the mRNAs from being translated into proteins. Multiple types of miRNAs can cooperate to suppress translation of a single mRNA, and a single miRNA can interact with multiple types of miRNAs, thereby regulating the protein expression of different genes (Filipowicz et al. 2008). RegRNA (http://regrna.mbc.nctu.edu.tw/html/prediction.html) is a web server that designs a variety of interface to facilitate the analysis for the homologs of the regulatory RNA motifs. This software predicted that the wild-type and mutant 3′-UTR were bound to different miRNAs (Appendix VII). However, target scan and UCSC genome browser did not predict any miRNA binding site involving the nucleotide and the sequence around it. In a study by Zhou et al. (2009) chronic treatment of primary cultures with the mood stabilizers, VPA or lithium, elevated levels of GRM7 and lowered levels of miR-34a. Incubation of primary hippocampal culture with miR-34a precursor significantly reduced GRM7 protein levels whereas, with anti-miR miRNA-34a inhibitor significantly increased GRM7 protein levels. These results suggested that the mood stabilizers produce a part of their behavioural effects through mechanisms including modulation of protein levels through miRNAs such as miR-34a (Zhou et al. 2009).

CNVs have been implicated in the pathogenesis of psychiatric disorders such as autism (Sebat et al. 2007; Christian et al. 2008; Glessner et al. 2009), schizophrenia (Bergen et al. 2012), bipolar disorder (Zhang et al. 2009)and ADHD (Elia et al. 2011). I confirmed the presence of previously reported singleton deletion and

duplication in the UCL1 sample. Functional studies investigating the possible role of these CNVs are required to shed more light on their involvement in the disorder.

In conclusion, I found a 3′-UTR variant in GRM7 that was significantly associated with bipolar disorder in the UCL sample but with increased representation in controls compared to cases. Bioinformatic analyses predicted that the rs56173829 variant influenced the centroid secondary structure of the RNA and also altered the binding site of microRNAs. However, results from these *in silico* analyses should be interpreted with caution as these are based on mathematical algorithms instead of expression studies. Thus, functional assays need to be carried out to assess the role of the 3′-UTR variant on GRM7 gene expression. The case-only novel mutations reported in the current study also warrant further investigation in other populations to understand their distribution and involvement in the aetiology of bipolar disorder.

Chapter 6

Mutation screening of the GRM3 glutamate receptor gene in Bipolar disorder

6.1. Introduction

The GRM3 gene encodes the mGluR3 protein which is a group II metabotropic glutamate receptor similar to mGluR2. mGluR3 is a G protein-coupled receptor which acts by inhibition of adenylate cyclase and reducing cyclic AMP production (Cartmell and Schoepp 2000). The group II mGluRs are involved in learning, memory, anxiety and perception of pain. mGluR3 is found on pre- and postsynaptic neurons in synapses of the hippocampus, cerebellum, cerebral cortex as well as other parts of the brain and in peripheral tissues. The mGluR3 receptors modulate glutamate concentrations by sensing excess synaptic glutamate release and providing a negative feedback mechanism which downregulates glutamate transmitter release. The activation of postsynaptic mGluR2/3 receptors can reduce neuronal excitability and plasticity through intracellular mechanisms such as modulation of ion channels and induction of long-term synaptic depression (Anwyl 1999). Studies have also suggested that mGluR2/3 receptors may function as heteroreceptors controlling the release of GABA and other neurotransmitters (Feyissa et al. 2010).

GRM3 is located on 7q21. Several bipolar linkage studies provide support for linkage on 7q (Baron 2002; Liu et al. 2003; Cassidy et al. 2007). Two meta-analyses of genome scans were also supportive under some models (Segurado et al. 2003; Tang et al. 2011). A recent study following up previously reported linkage signals with enlarged samples and enhanced marker density demonstrated strong linkage for bipolar disorder on 7q34 with a LOD of 3.53 (Xu et al. 2011). The linkage studies and GWAS are discussed in detail in sections 1.2.4 and 1.2.6.

The GRM3 gene has been investigated in BPAD as part of several GWAS, but failed to reach genome-wide significance in any of these studies. In a collaborative study (Ferreira et al. 2008; Sklar et al. 2008) fifteen single nucleotide polymorphism (SNP) markers within 50 kb of GRM3 gene showed allelic association with bipolar disorder. The SNP rs2237563 was the most significantly associated marker ($P = 3.85 \times 10^{-5}$) in the UCL1 sub-sample. Fourteen other markers within 125 kb of the gene also showed nominally significant associations with bipolar disorder in the UCL1 sub-sample. In the WTCCC sample (WTCCC 2007), thirteen SNP markers within 125 kb of GRM3 were found to be associated with bipolar disorder and in the STEP-BD sample, a further four SNPs were associated. Of these, three markers, rs10258008, rs11974622 and rs1557665 were consistently found to be associated with bipolar disorder in the UCL and WTCCC sample, whereas only one marker, rs12704286, was associated with bipolar disorder in both the STEP-BD and UCL samples. In the PGC bipolar study, the most associated GRM3 marker was rs17161018 ($P = 9.32 \times 10^{-4}$) (Figure 6.1) (Sklar et al. 2011). A total of 32 different SNP markers within 125 kb of GRM3 were nominally associated with bipolar disorder. Several other association studies in different populations also provide evidence of association between GRM3 markers and bipolar disorder. In a study of Ashkenazi bipolar trios, GRM3 was ranked as showing "highly suggestive" transmission disequilibrium between a four SNP marker haplotype and bipolar disorder (Fallin et al. 2005). A study of bipolar disorder in South Africa found that heterozygotes for the GRM3 SNP rs6465084 had a four-fold increased risk of lifetime history of psychotic symptoms (Dalvie et al. 2010). However, a German study found no evidence of association between GRM3

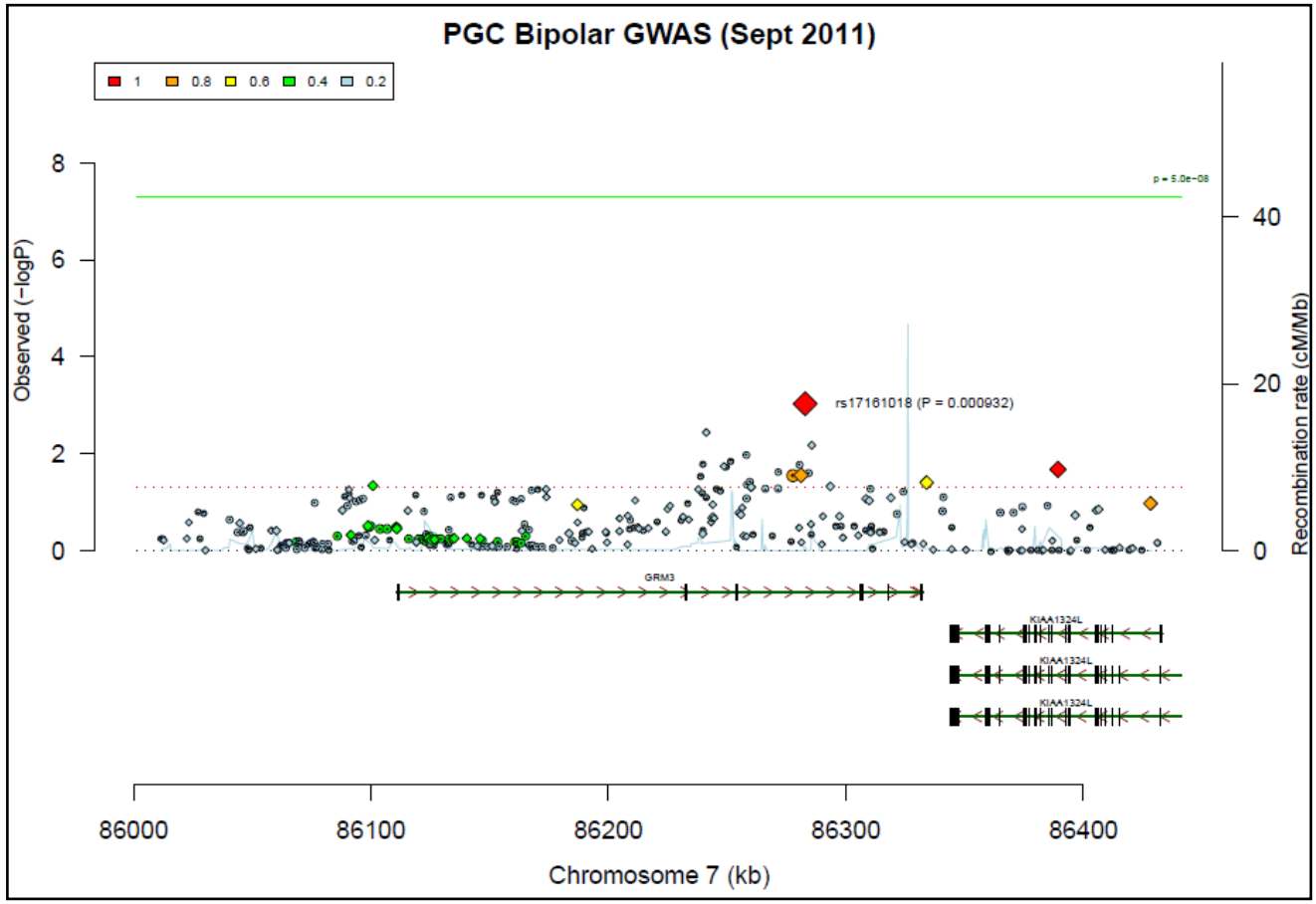


Figure 6.1 Association of GRM3 SNPs in the PGC-BP GWAS generated using Ricopili

and bipolar disorder (Marti et al. 2002). A study of MDD revealed an allelic association between the SNP rs6465084 in GRM3 and depression (P = 0.037) (Tsunoka et al. 2009).

The human GRM3 gene consists of six exons that give rise to four transcript variants, the main transcript being NM_000840. Four alternatively spliced transcripts of GRM3 have been previously reported in human brain, including a full length transcript, a transcript with exon 2 deleted, a transcript with exons 2 and 3 deleted and a transcript with exon 4 deleted (Sartorius et al. 2006) (Figure 6.2). Molecular characterization of the promoter of GRM3 (Corti et al. 2001) provided evidence of a silencing element being present in exon 1. Clinical studies of the effect of receptor agonists and antagonists of mGluR2/mGluR3 have been carried out for the treatment of anxiety, schizophrenia, depression and withdrawal from morphine and nicotine (Schoepp et al. 2003; Patil et al. 2007; Dunayevich et al. 2008).

6.2. Hypothesis

A positive association was observed for GRM3 in the UCL1 bipolar GWAS and several other bipolar GWAS have also reported nominal associations for the gene. On the basis of the results of GWAS and linkage studies, I hypothesised that GRM3 is a strong susceptibility candidate gene for bipolar disorder.

6.3. Aims

The aim of this study was to identify novel variants in GRM3 associated with bipolar disorder by resequencing GRM3 in bipolar cases selected for inheriting a disease haplotype in the GWAS. Further functional characterisation of the novel variants would be undertaken to understand the mechanism of action of the mutation in the disorder.

6.4. Results

6.4.1. GRM3 association analysis

GRM3 was associated with bipolar disorder in the UCL1 GWAS (Appendix V). I tested association of the three most significantly associated GRM3 SNPs, rs2237563, rs274621 and rs2158786 from the UCL1 bipolar sample in the UCL2 case-control sample. None of these SNPs showed association with bipolar disorder in the UCL2 case-control sample (Table 6.1). When the genotyping data of the three SNPs in UCL1 and UCL2 was combined using PLINK (Purcell et al. 2007), two of the three SNPs remained significantly associated (rs2237563: P=0.0099, OR=1.222 and rs2158786: P= 0.025, OR=0.852) with bipolar disorder (Table 6.2).

Table 6.1 SNP association results with GRM3 in the UCL bipolar sample

SNP	Position (NCBI35/hg19)	UCL1[1]				UCL2[2]			
		Allele counts (freq)		*P* Value	OR	Allele counts (freq)		*P* Value	OR
		BP	CON			BP	CON		
rs2237563	86443168	G 264 (0.27)	G 190 (0.19)	**3.85 x 10^{-5}**	1.56	G 217 (0.21)	G 195 (0.21)	0.95	1.01
		A 734 (0.73)	A 822 (0.81)			A 813 (0.79)	A 729 (0.79)		
rs274621	86272799	C 388 (0.38)	C 332 (0.33)	**0.007**	1.29	C 365 (0.35)	C 329 (0.36)	0.82	0.98
		T 622 (0.62)	T 688 (0.67)			T 677 (0.65)	T 599 (0.64)		
rs2158786	86406018	G 764 (0.76)	G 723 (0.71)	**0.01**	0.78	G 757 (0.72)	G 661 (0.72)	0.67	0.96
		A 242 (0.24)	A 295 (0.29)			A 285 (0.28)	A 263 (0.28)		

SNP, single nucleotide polymorphism; OR, odds ratio;[1] UCL1 data from our GWA study (Sklar et al. 2008); [2] UCL2 is a replication sample; Additional genotyping performed in UCL1 after the GWAS has been added to UCL1&2; BP, bipolar cases, and CON, controls; Significant P values are in bold P < 0.05

Table 6.2 Combined analysis of GRM3 SNPs in UCL1 and UCL2 samples

SNP	Position (NCBI35/hg19)	Combined analysis Allele counts (MAF)				Meta-analysis Fixed effects model		Meta-analysis Random effects model		Q	I^2
		BP	CON	P-value	OR	*P* Value	OR	P-value	OR		
rs2237563	7251433	G 465 (0.23)	G 385 (0.20)	**0.01**	1.22	0.00323	1.2536	0.2989	1.2439	0.0062	86.64
		A 1547 (0.77)	A 1551 (0.80)								
rs274621	7241745	C 728 (0.36)	C 662 (0.34)	0.18	1.09	0.07854	1.1225	0.3981	1.119	0.0428	75.62
		T 1296 (0.64)	T 1288 (0.66)								
rs2158786	7246241	G 1453 (0.64)	G 1440 (0.71)	**0.03**	0.85	**0.02134**	0.8498	0.1715	0.8504	0.0938	64.38
		A 517 (0.26)	A 558 (0.29)								

MAF, minor allele frequency; BP, bipolar; CON, control; OR, odds ratio; Q, p-value for Cochrane's Q statistic; I^2, I^2 heterogeneity index (0-100)

To assess the effect of heterogeneity on the association of these SNPs in the combined sample, I performed a meta-analysis of these SNPs using PLINK. Interestingly, only one of the SNPs, rs2158786 remained significant in the meta-analysis (Cochrane Q-statistic *P*-value of 0.1405, Fixed effects model; *P*-value= 0.021 and OR= 0.850) (Table 6.2). The heterogeneity index for the two other SNPs, rs2237563 and rs274621 was very high at 86.64 and 75.62 respectively, suggesting that UCL1 and UCL2 are heterogeneous even though they were selected using the same ancestry-based methods (Table 6.2). Of note is the impact of heterogeneity on replication of allelic associations even within well-defined ancestral populations, but the meta-analysis in this case is underpowered with only two studies being analysed. Therefore, there is support for the association between the GWAS associated SNP rs2237563 in GRM3 and bipolar disorder in the UCL2 samples despite evidence of heterogeneity.

6.4.2. Detection and evaluation of new variants

A total of 11 SNPs including a putative Kozak sequence variant (rs148754219) were detected by sequence analysis across the promoter region, 5'-UTR, exons, intron/exon junctions and the 3′-UTR of GRM3 (Table 6.3). These included the synonymous and non-synonymous coding base pair changes, rs2228595 and rs17161026; a SNP in intron 2, rs139639092; three SNPs in the exon 1 5′-UTR, rs184681725, rs2073549 and rs148754219; and five promoter SNPs, rs274617, rs274618, rs166677, rs274619, and rs274622. Bioinformatic analysis of the promoter region SNPs for altered transcription factor binding indicated that the mutant alleles of rs148754219 and rs166677 were likely to bind to an increased range of

Table 6.3 SNPs found by sequencing GRM3

GRM3 Gene region	SNP	Predicted effect on GRM3 structure/function for novel SNPs	Chr 7 position (NCBI37/hg19)	BP MAF	CON MAF	HapMap MAF	1000 genome MAF (Mar 2012)
Promoter	rs274617		86271830	0.42	0.57	0.32	0.26
Promoter	rs274618		86272016	0.42	0.60	0.32	0.24
Promoter	rs166677	Alters binding of two transcription factors (TESS)	86272294	0.29	0.47	0.40	0.35
Promoter	rs274619		86272487	0.46	0.34	0.25	0.26
Promoter	rs274622		86272940	0.42	0.60	0.32	0.27
Exon 1 5′-UTR	rs184681725		86273267	0.02	0.03	na	0.003
Exon 1 5′-UTR	rs2073549		86273584	0.07	0.05	0.01	0.02
Exon 1 5′-UTR	rs148754219		86274087	0.03	0	na	0.01
Intron 2	rs139639092		86395015	0.03	0	na	0.01
Exon 3	rs2228595	Coding synonymous (A293A)	86415987	0.05	0.02	0.06	0.06
Exon 4	rs17161026	Coding non-synonymous (G475D)	86468254	0.02	0	0.02	0.01

Chr, Chromosome; MAF, minor allele frequency; BP, bipolar cases; CON, Controls; UTR, untranslated region; na, not available.

transcription factors compared to their respective common alleles. Prediction of the 5′-UTR secondary structure of GRM3 mRNA by the Mfold program showed that the rs148754219 base pair change was not likely to alter the mRNA structure (Figure 6.3). The variant rs148754219 is located two bases upstream of the ATG start codon of the GRM3 isoform GRM3Δ2Δ3 (ENST00000546348) encoding ENSP00000441407, and is therefore in the Kozak sequence for this isoform (Figure 6.2).

Ten of these SNPs were selected for genotyping in the complete sample (UCL1 and UCL2) on the basis of increased frequency in the sequenced cases compared to sequenced controls or on predicted functional effects (Table 6.3). Of these, the Kozak sequence variant rs148754219 was found to be significantly associated with bipolar disorder (P = 0.005; OR, 4.20; 95% CI, 1.43-12.37) (Table 6.4) after correction for multiple markers (P = 0.047). In the UCL research sample, 19 bipolar cases and only four controls were found to be heterozygous for the variant. The association of the rs148754219 variant in the UCL research sample remained significant when only Bipolar I cases were included in the analysis (P = 0.006). None of the other new SNPs were associated with bipolar disorder in the complete UCL sample (Table 6.4). Imputation using genotype data for GRM3 exonic SNPs and SNPs in the intron/exon boundary, of European samples in the 1000 genomes project did not find any functional SNPs associated with bipolar disorder (Table 6.5). All the variants showing proxy association as a result of imputation, excluding the rs148754219 variant, were intronic in GRM3 or outside the gene.

Table 6.4 Tests of association with potentially aetiological GRM3 SNPs in the UCL sample

SNP ID	Position (NCBI37/hg19)	N	Allele counts (MAF)	Genotype counts	P-value	OR (95% CI)
rs148754219						
Case	86274087	1062	A 19 (0.01); G 2105	AA 0; AG 19; GG 1043	0.005	4.20 (1.43-12.37)
Control		932	A 4 (0.002) ; G 1860	AA 0; AG 4; GG 928		
rs166677						
Case	86272294	1048	G 960 (0.46); A 1136	GG 221; AG 518; AA 310	0.19	1.09 (0.96-1.23)
Control		908	G 794 (0.44); A 1022	GG 176; AG 442; AA 290		
rs17161026						
Case	86415987	1056	T 30 (0.01); C 2082	TT 0; CT 30; CC 1026	0.51	1.21 (0.69-2.10)
Control		931	T 22 (0.01); C 1840	TT 0; CT 22; CC 909		
rs139639092						
Case	86395015	1057	A 21 (0.01); G 2093	AA 0; AG 21; GG 1036	0.53	1.24 (0.64-2.41)
Control		934	A 15 (0.01); G 1853	AA 0; AG 15; GG 919		
rs2228595						
Case	86468254	1053	T 118 (0.06); C1988	TT 1; CT 116; CC 936	0.68	0.94 (0.72-1.23)
Control		931	T 110 (0.06); C 1752	TT 6; CT 98; CC 826		
rs2073549						
Case	86273584	1061	T 76 (0.04); A 2046	TT 1; AT 74; AA 986	0.85	0.97 (0.69-1.35)
Control		932	T 69 (0.04); A 1795	TT 1; AT 67; AA 865		
rs184681725						
Case	86273267	1056	A 25 (0.01); C 2087	AA 0; AC 25; CC 1031	0.87	0.96 (0.54-1.69)
Control		927	A 23 (0.01); C 1831	AA 0; AC 23; CC 904		

SNP, single nucleotide polymorphism; N, number of samples; MAF, minor allele frequency; OR, odds ratio; CI, confidence interval; NA, not applicable

Table 6.5 Tests of association with GRM3 SNPs in the UCL sample using imputation from 1000 genomes European samples

SNP	BP	A1	A2	GENO	NPRX	INFO	F_A	F_U	OR	P
rs2237563	86443168	G	A	0.752	2	0.816	0.231	0.2	1.2	0.00627
rs148754219	86274087	A	G	0.908	2	0.988	0.0089	0.00216	4.14	0.00672
rs2299231	86489625	G	A	0.403	2	0.403	0.117	0.103	1.16	0.0164
rs2158786	86406018	A	G	0.757	2	0.784	0.258	0.285	0.869	0.0194
rs12704286	86374941	A	G	0.398	2	0.494	0.253	0.275	0.893	0.0217
rs2237554	86378896	G	T	0.402	2	0.499	0.253	0.275	0.894	0.0227
rs10236047	86512637	G	A	0.402	2	0.403	0.115	0.102	1.15	0.0251
rs10258008	86535446	A	C	0.403	2	0.417	0.115	0.101	1.16	0.0251
rs723631	86401592	C	G	0.403	2	0.469	0.251	0.271	0.901	0.0273
rs2519713	86544667	C	A	0.402	2	0.445	0.428	0.45	0.913	0.0292
rs11974622	86536377	A	G	0.402	2	0.416	0.115	0.102	1.15	0.0295
rs7804907	86428208	T	C	0.403	2	0.48	0.25	0.268	0.911	0.0507

A1, Allele 1; A2, Allele 2; GENO, Genotyping for the reference SNP; NPRX, Number of proxy SNPs used to tag reference SNP; INFO, Information metric for each reference SNP; F_A; Reference SNP allele frequency in cases; F_U; Reference SNP allele frequency in controls; OR, odds-ratio; P, Asymptotic p-value for test of association. All the positions are correct according to NCBI35/hg19 assembly

(a) **(b)**

Figure 6.3 Secondary structures of mRNA predicted using mFold webserver for wild-type and mutant 5′-UTR of GRM3 containing rs148754219 variant

I combined data from all the SNPs genotyped in the GRM3 gene in the present study (UCL1 GWAS SNPs, the three GWAS SNPs in UCL2, and ten SNPs found by mutation screening) and performed haplotype analysis using Haploview (Barrett et al. 2005). The haplotype block comprising seven SNPs rs2299221, rs12704289, rs7804907, rs2237563, rs7806785, rs2299224 and rs17697609 showed the strongest association with bipolar disorder, which survived permutation testing ($P = 7.11 \times 10^{-5}$, empirical $P = 0.0009$) (Figure 6.4) (Table 6.6). The haplotype driving this association was the same haplotype identified by SNP rs2237563 in the UCL1 GWAS study (Sklar et al. 2008). No other haplotypes were significantly associated with bipolar disorder. The Kozak sequence variant rs148754219 was located outside this haplotype block and showed negligible LD with SNP rs2237563 ($D' = 0.57$, $r^2 = 0.001$). The SNPs that were of similar frequency in the case and control sequencing panel samples and those that were in complete LD with the UCL1 bipolar GWAS SNPs were not genotyped in the whole sample

Table 6.6 GRM3 haplotypes associated with bipolar disorder

Haplotype	Alleles	Frequency of the haplotype	Frequency		Chi-square	P-value
			Case	Control		
rs1405881-rs1405880-rs274637-rs274621-rs802437-rs802439-rs802441-rs802459	TGCCAGCT	0.349	0.375	0.321	6.86	0.0088
rs802467-rs2189812-rs10238436-rs16888204	ACGA	0.297	0.323	0.270	7.095	0.0077
rs2237553-rs12704286-rs2237554-SNP_A-2126810-rs2888551-rs723631-rs17160978-rs17697445-rs2158786-rs16888210-rs2237559	CGTAAGCGGCG	0.702	0.727	0.677	6.299	0.0121
	TAGAACCGACG	0.133	0.117	0.150	4.993	0.0255
rs2299221-rs12704289-rs7804907-rs2237563-rs7806785-rs2299224-rs17697609	ATCGCAA	0.222	0.257	0.185	15.782	7.11E-05
	ATTAGAA	0.143	0.126	0.161	5.133	0.0235
rs17161024-rs166677-rs184681725-rs2073549-rs10952891-rs148754219	GACACG	0.467	0.436	0.501	8.075	0.0045
rs7788115-rs6955452-rs6955917-rs2299231-rs10236047-rs10258008-rs7781178-rs11974622-rs2519713	TCGAACAGA	0.287	0.311	0.265	5.339	0.0209

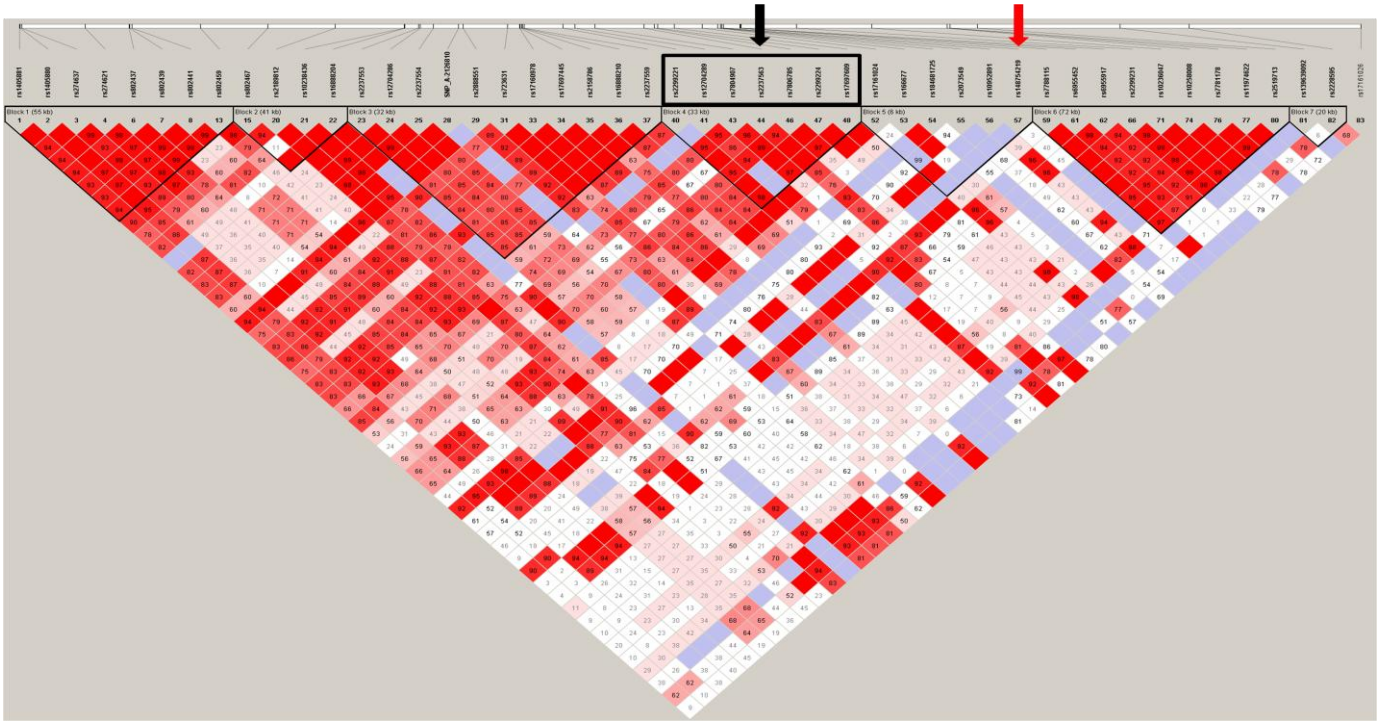


Figure 6.4 Linkage disequilibrium between all the GRM3 **SNPs genotyped in UCL research sample**

The figure was generated using Haploview. The most significantly associated GWAS SNP, rs2237563 is marked with a black arrow; the 5′-UTR variant, rs148754219 is marked with a red arrow and the associated haplotype is marked using a box.

6.4.3. Genotyping of the rs148754219 variant in a bipolar family, and unrelated alcohol dependence and schizophrenia cohorts

I sequenced the DNA of the family members available for one of the bipolar patients who was heterozygous for the rs148754219 variant to examine if the mutation segregated with psychiatric illness in the family. DNA samples available from a family containing members with unipolar disorder included that of a same-sex sibling (F076004) and, one alternate sex (F076002) and two same-sex third-degree relatives (F076001 and F076003), all of whom have a unipolar disorder diagnosis except F076002 (Figure 6.5). Sequencing analysis revealed that none of the family members of the proband (heterozygous for rs148754219 variant) carried the mutation.

I further genotyped the rs148754219 variant in the alcohol dependence and schizophrenia samples to detect the frequency of this variant in other psychiatric disorder samples available in the Molecular Psychiatry Laboratory, UCL. The variant showed significant association with schizophrenia and a trend towards association with alcoholism (P = 0.024 and 0.0869 respectively) (Table 6.7). When the genotype data from all the bipolar disorder, alcohol dependence and schizophrenia samples was combined and compared to that of the UCL controls, the SNP remained significant at P = 0.012 [OR (95%CI) = 3.47 (1.24-9.73)].

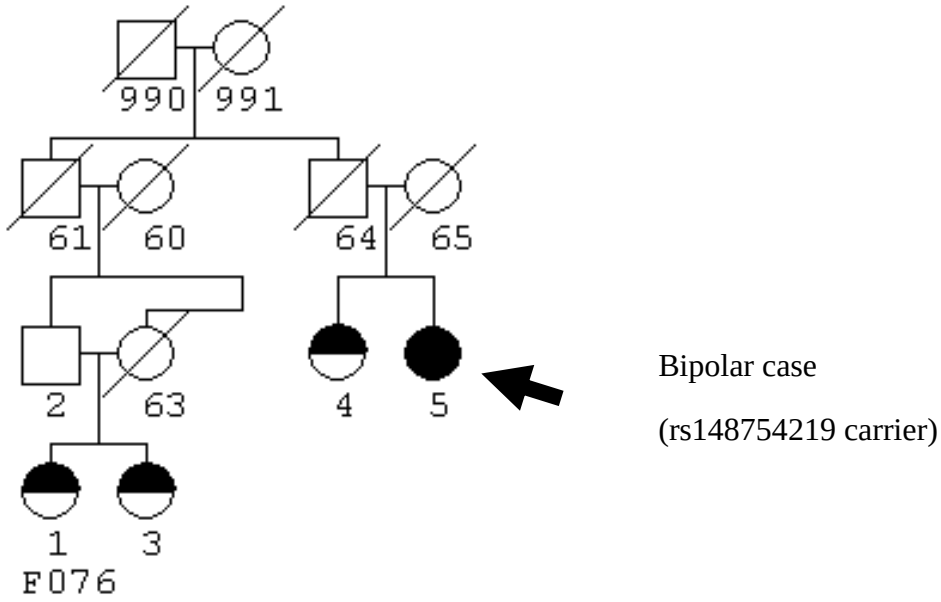


Figure 6.5 Pedigree of the bipolar case heterozygous for the rs148754219 variant

Solid shapes represent bipolar disorder diagnosis, half-filled shapes represent unipolar disorder diagnosis and clear shapes represent normal individual.

Table 6.7 Association results of rs148754219 variant in psychiatric disorder samples

Sample	Allele counts		Genotype counts			P-value	OR (95% CI)
	A	G	AA	AG	GG		
CON	4	1860	0	4	928	NA	NA
ALC	11	1941	1	9	966	0.085	2.64 (0.84-8.30)
BP	19	2105	0	19	1043	**0.005**	4.20 (1.43-12.37)
SCZ	9	1175	0	9	583	**0.024**	3.56 (1.09-11.59)
ALC+BP+SCZ/CON	39	5221	1	37	2592	**0.012**	3.47 (1.24-9.73)

CON, controls; ALC, alcohol dependence; BP bipolar; SCZ, schizophrenia; OR, odds ratio; CI, confidence interval.

6.4.4. DNA-Protein complex formation by rs148754219 variant

TESS analysis predicted that the rare allele of rs148754219 variant created transcription factor binding sites recognised by transcription factors including AP-1, NF-X3, RFX2, E12 and myogenin and the wild-type allele created binding sites for RAF-1 and GCR1 (Figure 6.6). Electrophoretic mobility shift assays (EMSAs) of the allelic variants of rs148754219 on transcription factor binding showed that under the experimental conditions used the mutant A allele caused gel shifts compared to the wild-type G base sequence. This effect could be completely abolished by the addition of a 200-fold excess of unlabelled A probe but not by unlabelled competitor wild-type G probe, indicating that the DNA-protein binding was exclusive to the mutant allele (Figure 6.7). Similar results were obtained in three independent experiments. Supporting the TESS prediction results, the rare A allele of rs148754219 appears to bind more strongly to transcription factors within the nuclear lysate than the wild-type G allele, suggesting that this variant creates a transcription factor binding site likely to change GRM3 mRNA expression compared to the wild-type sequence.

GRM3_rs148754219_wt

```
00001  GCCAGTAAGC TACCTCTTTT GTGTCGGATG AGGAGGACCA ACCATGAGCC  00050
       ====(8.00) NF-1 R01681
                  ====(8.00) Dof2,Dof3,MNB1a,PBF
R08443,R08440,R08441,R08442
                  =====(8.00) LEF-1,TCF-1(P),TCF-1,TCF-1A,TCF-1B,TCF-
1C,TCF-1E,TCF-1F,TCF-1G,TCF-2alpha
R02248,R02248,R02248,R02248,R02248,R02248,R02248,R02248,R02248,R02248
                    = =====(12.00) GR R03537
                        ====(8.00) RAF R00256
                          =====(10.00) GCR1 R03806
                              = ====(10.00) GAL4 R00492
                                  ======(8.9089) H-2RIIBP I00178
                                    =====(10.00) H4TF-2 R00681
                                    =====(7.9248) H4TF2 I00180
```

GRM3_rs148754219_mut

```
00001  GCCAGTAAGC TACCTCTTTT GTGTCAGATG AGGAGGACCA ACCATGAGCC  00050

       ====(8.00) NF-1 R01681
                  ====(8.00) Dof2,Dof3,MNB1a,PBF
R08443,R08440,R08441,R08442
                  =====(8.00) LEF-1,TCF-1(P),TCF-1,TCF-1A,TCF-1B,TCF-
1C,TCF-1E,TCF-1F,TCF-1G,TCF-2alpha
R02248,R02248,R02248,R02248,R02248,R02248,R02248,R02248,R02248,R02248
                    = ======(14.00) Zta R00299
                    = ======(9.3987) v-Jun I00417
                    = ======(14.00) AP-1,c-Jun,v-Jun,YAP1
R01412,R01412,R01412,R01412
                    = =====(12.00) GR R03537
                        ======(9.6721) NF-X3 I00397
                        ======(9.6721) RFX2 I00406
                        ======(12.00) E12,myogenin R08512,R08512
                            = ====(10.00) GAL4 R00492
                                ======(8.9089) H-2RIIBP I00178
                                  =====(10.00) H4TF-2 R00681
                                  =====(7.9248) H4TF2 I00180
```

Figure 6.6 TESS results for rs148754219 wild-type and mutant alleles

The alleles are shown in red and the transcription factors predicted to bind the alleles are highlighted.

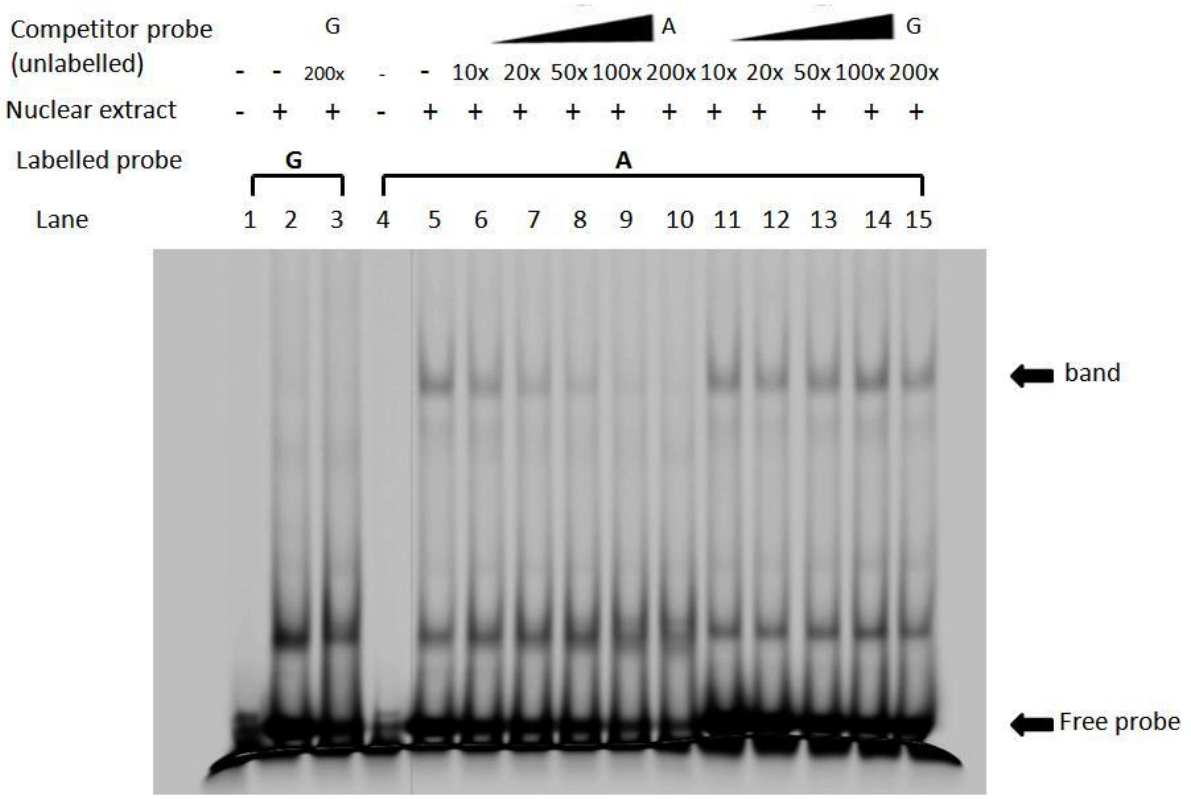


Figure 6.7 Electrophoretic mobility shift assay for the rs148754219 variant

EMSAs performed with G (1-3) or A allele probes (4-15) and SH-SY5Y cell nuclear extract. DNA-protein complex (band) formed with A-allele was competed in a concentration-dependent manner by unlabelled A-allele (5-10), but not by unlabelled G-allele (11-15).

6.4.5. Effect of the rs148754219 variant on GRM3 promoter activity and its function as a transcriptional silencer

The ability of the rs148754219 variant to modulate transcription was tested using luciferase reporter gene constructs (Figure 6.8). Clone A tested the effect of rs148754219 on the ability of GRM3 exon 1 to act as a promoter for luciferase reporter gene activity. In HEK293 cells there was no difference in the relative luciferase expression levels of wild-type and mutant Clone A and B constructs (Figure 6.9 and 6.10). I observed significant reductions of 61% (Student's *t*-test, $P<0.001$) (Figure 6.10) in the level of luciferase expression with the mutant Clone A construct compared to the wild-type Clone A construct, whereas, there was no difference in the expression levels of wild-type and mutant Clone B constructs in SH-SY5Y cells (Figure 6.10). Clone B tested whether similar effects on promoter activity with the rs148754219 alleles could be observed when the GRM3 putative minimal promoter was combined with exon 1. Since the difference in expression levels between wild-type and mutant Clone A constructs is not apparent in the presence of the minimal promoter in the Clone B constructs, minimal GRM3 promoter activity appears to mask any effect the mutation would have on the expression of GRM3. Therefore, given the considerable loss in Clone A construct promoter activity with the mutant compared to the wild-type allele in SH-SY5Y cells, I further tested the ability of the rs148754219 mutation to negatively regulate expression of GRM3.

Clone C with GRM3 exon 1 cloned upstream of the SV-40 promoter followed by the luciferase gene was used to test whether rs148754219 alleles affected the silencing element present in exon 1 (Corti et al. 2001). After 48 hours of transfection, the mutant Clone C construct reduced the expression of luciferase by 37% (Mann-Whitney U test, P < .0001) compared to the wild-type construct in HEK293 cells (Figure 6.9). In contrast, this effect was reversed in SH-SY5Y cells with an increase in luciferase expression in the mutant construct of 11% (Student's *t*-test, *P*-value = 0.04) compared to the wild-type construct (Figure 6.10). Together, these findings indicate that the rs148754219 variant exhibits characteristics of a tissue-specific transcription enhancer element binding site in SH-SY5Y cells and a silencer element in HEK293 cells.

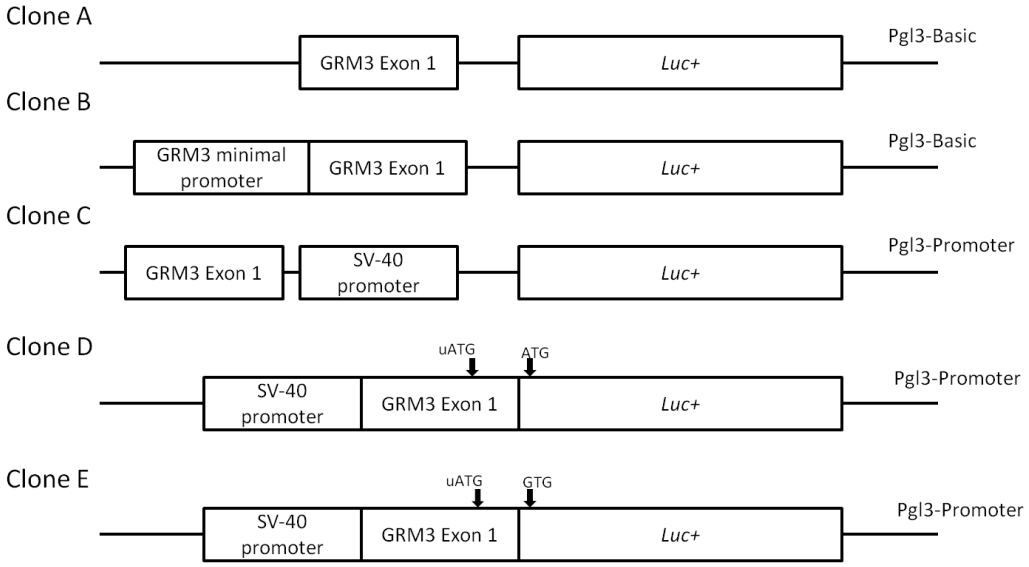


Figure 6.8 Luciferase (Luc+) reporter gene constructs designed to assess the impact of rs148754219 on transcription and/or translation

In clone E the native luciferase start codon is mutated so that translation can only begin at the start codon in GRM3 exon 1.

6.4.6. Elimination of luciferase expression driven by post-transcriptional/translational activity of the rs148754219 variant

In order to investigate whether the rs148754219 variant regulated GRM3 gene expression at the translational level, I cloned the GRM3 Exon 1 upstream of the firefly luciferase gene and just downstream of the SV-40 promoter, making the insert the 5′-UTR of the firefly luciferase gene. The wild-type and mutant constructs (Clone D) were transfected in HEK293 and SH-SY5Y cells. In both cell-lines, the wild-type rs148754219 variant in the pGL3-promoter backbone of Clone D resulted in a decrease in luciferase reporter activity while the mutant led to total elimination of luciferase activity (Figure 6.9 and 6.10). To elucidate that the nullifying effect of the mutant allele of rs148754219 on protein production was related to the in-frame ATG 2 bp downstream of the mutation in the GRM3 5′-UTR insert and not the ATG of the firefly luciferase gene, I created Clone E. In Clone E, the A base of the start codon of luciferase gene was mutated to G to give GTG, leaving the ATG immediately downstream of the GRM3 5′-UTR insert as the only in-frame start codon for initiating the translation of firefly luciferase. Again, following the transfection of wild-type and mutant Clone E constructs, the mutant Clone E construct did not produce any reporter activity in either HEK293 or SH-SY5Y cells as observed with Clone D (Figures 6.8 and 6.9). These results suggest that the mutant A allele of rs148754219 polymorphism abolishes upstream post-transcriptional/translational activity.

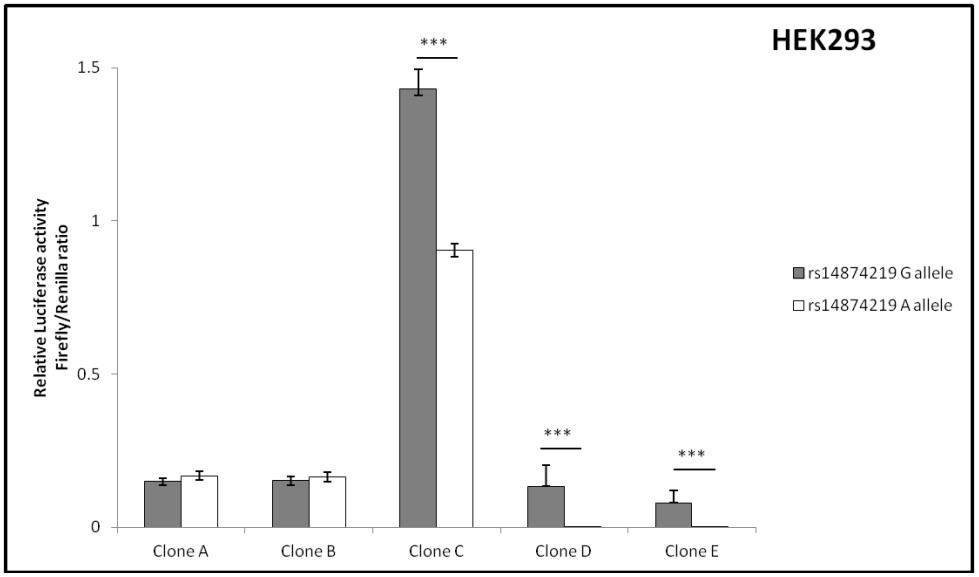


Figure 6.9 Relative luciferase activity of reporter gene assays for the experimental constructs in HEK293 cells

Data expressed as mean ± SEM of three independent experiments performed in triplicate (n = 9). *P < 0.05; ***P < 0.001.

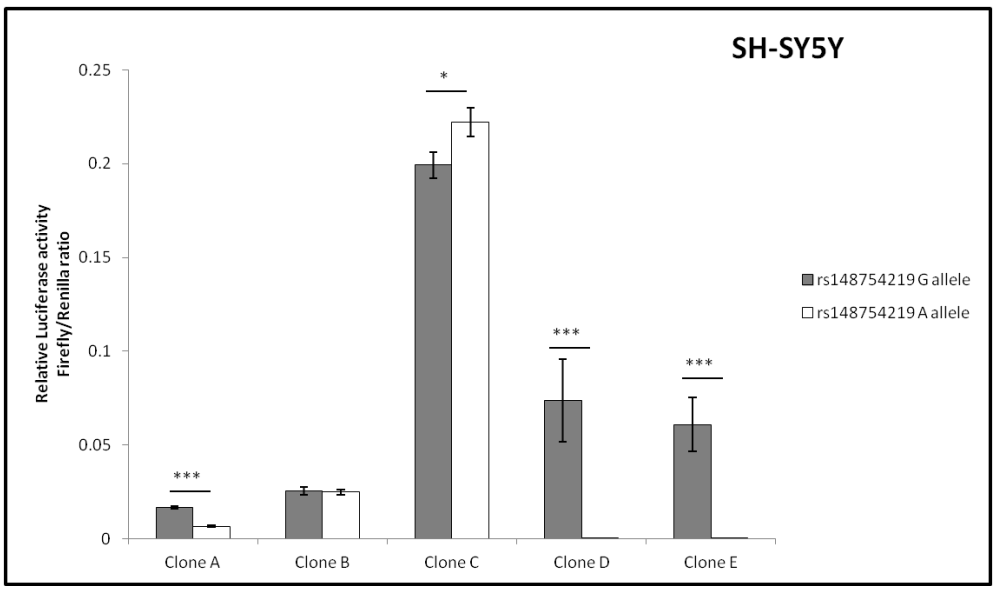


Figure 6.10 Relative luciferase activity of reporter gene assays for the experimental constructs in SH-SY5Y cells

Data expressed as mean ± SEM of three independent experiments performed in triplicate (n = 9). *P < 0.05; ***P < 0.001.

To gain further insight into the molecular impact of the rs148754219 variant on gene regulation, I quantified the luciferase mRNA levels of the wild-type and mutant Clone D constructs and compared the results to that of the gene reporter luminescence assays described above. There was no significant difference in the relative mRNA levels of the wild-type and mutant clones in the HEK293 cell-line (wild-type; mean [SEM], 0.8517 [0.0274]; mutant, 0.8543 [0.0513]) (*n*=3) and the SH-SY5Y cell-line (wild-type, 0.8303 [0.0087]; mutant, 0.7889 [0.0291]) (*n*=3) (Figure 6.11). Therefore, the abundance of luciferase mRNA produced by the Clone D construct with the mutant allele suggests that the luciferase gene is being transcribed. However, this did not correlate with the relative reporter activity of this clone whereby luciferase luminescence was absent, indicating that the luciferase protein was not translated. Thus, these results suggest that the bipolar associated GRM3 5′-UTR variant rs148754219 can be functionally distinguished from the wild-type in different gene reporter assays and has an effect at the post-transcriptional and/or translational level.

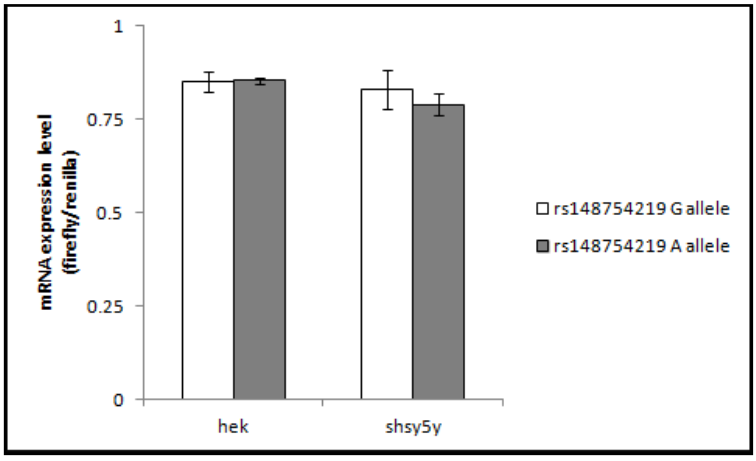


Figure 6.11 Relative luciferase mRNA quantification of wild-type and mutant Clone D constructs using QRT-PCR

6.5. Discussion

The failure to find genetic association in the UCL2 cohort for the markers most strongly associated in the UCL1 sample (Sklar et al. 2008) is typical in the field of complex disease genetics. The difficulty in obtaining clear cut replications is due to the presence of low frequency disease alleles and high degree of aetiological genetic heterogeneity. The UCL2 replication sample was under powered for detecting the association between the SNPs rs274621 and rs2158786 and bipolar disorder but had 95% power for detecting that of SNP rs2237563 (Appendix VI). Therefore, it is possible that the initial association signals were chance findings. Significant association was observed for the rs148754219 variant with BP I disorder in the combined UCL1 and UCL2 sample. The small number of BP II cases did not permit a test of association with this variant in the UCL sample. The rs148754219 variant has not been previously reported in any other sample other than the 1000 genomes sample (www.1000genomes.org). The rs148754219 variant did not segregate in unipolar affective disorder family members of one of the heterozygous bipolar patients for the variant. One explanation may be that this variant is a pure bipolar disorder susceptibility variant and thus could not be detected in the unipolar family members. Also this variant was significantly associated with schizophrenia and was present at a higher frequency in the alcohol dependence syndrome samples compared to the control samples. It should be noted that the same controls samples were used for testing allelic association in the bipolar disorder, alcohol dependence and schizophrenia samples. Thus, caution needs to be excercised while interpreting the results of the association analyses of the Kozak sequence variant with other psychiatric disorders.

The rs148754219 variant is located in the first exon of GRM3, which is transcribed and not translated in the main isoform of the gene (NM_000840). Bioinformatic analysis predicted that the mutant allele of rs148754219 bound to several more transcription factors than the wild-type allele. The results from the EMSA assays provided experimental evidence for this as shown by the protein/DNA band shifts with the mutant allele. In this study, I did not use super-gel shift assays to verify the composition of the DNA-protein complex because antibodies specific for all the transcription factors altered by the mutant allele were not available commercially. Based on the bioinformatic analysis and the EMSA results I provide evidence that the rs148754219 variant affects gene expression via the basal transcription apparatus that contains partial sequence downstream of the transcription start site as reported for several other genes (Minet et al. 1999; Singh et al. 2002; Berardi et al. 2003; Coppotelli et al. 2006). I also showed that the mutant A allele could cause upregulation of reporter activity in neuroblastoma cells and downregulation in kidney derived cells. These differences are likely due to the variation in the presence of different transcription factors expressed in the two cell types.

To further elucidate the mechanism of action of this variant on the transcriptional regulation of the gene, I generated three constructs using two different reporter vectors – pGL3-Basic and pGL3-promoter vector. The resultant clones were transfected into two different cell-lines, one that does not express GRM3 endogenously and another expressing GRM3 endogenously, HEK293 and SH-SY5Y respectively. With the introduction of new transcription factor binding sites specific to the mutant allele as predicted by TESS analysis, I expected to see a change in expression in the promoter assays with the mutant allele. As the differences observed

with the two clones (Clones A and B) are inconsistent and vary by cell-type, it adds to the complexity of the role this variant has in controlling mRNA transcription. One explanation for the inconsistency between the expression level results of Clone A and Clone B constructs is that the minimal GRM3 promoter activity in Clone B constructs masks any effect the mutation has on the expression of GRM3, implying an interaction between the promoter and the 5′-UTR.

Experimental evidence suggests that exon 1 of GRM3 contains a silencer binding sequence resulting in a negative regulation of transcriptional activity (Corti et al. 2001). Interestingly, contrasting results in the two cell-lines were obtained with these constructs designed to assess the effect of rs148754219 variant on the silencing effect of exon 1. In accordance with these results, it is apparent that this mutation might affect the transcriptional regulation of GRM3 in a cell-specific manner by creating a transcription enhancer element binding site in the SH-SY5Y neuroblastoma cells and a transcription silencer element binding site in HEK293 cells. Together, the results from all the three constructs (Clone A, Clone B, and Clone C) designed for investigating the *in vitro* effect of rs148754219 polymorphism on the transcriptional regulation of GRM3 gene suggest that the variant may regulate gene expression by altering the binding affinity of the 5′-UTR towards transcription factors, in turn affecting the promoter activity.

Four alternatively spliced transcripts of GRM3 have been previously reported in human brain corresponding to full-length GRM3, transcripts with a deletion of exon 2, exons 2 and 3 deleted or with exon 4 deleted (Figure 6.1) (Sartorius et al. 2006). In GRM3Δ2Δ3 isoform, the SNP rs148754219 is at -2 position from the translation

initiation codon ATG forming part of the Kozak consensus motif, which is a consensus sequence found in eukaryotic mRNA regulating protein translation (Kozak 1999). Based on its location in the GRM3 5′-UTR, I hypothesised that rs148754219 variant might influence the translation efficiency of an alternatively spliced isoform, GRM3Δ2Δ3 and levels of the truncated form of protein for GRM3. Luciferase assays with the mutant 5′-UTR cloned downstream of a strong SV-40 promoter and upstream of the firefly luciferase cDNA (Clone D) showed that the mutant allele caused a total suppression of luciferase expression in both the cell-lines suggesting that the mutation might be harmful. Confirmation that the null reporter activity caused by the mutant allele resulted from the effect on translation initiation from the ATG codon immediately following the rs148754219 variant was found by mutating the luciferase gene start codon in Clone E. Thus, the AUG codon 2 bp downstream of the rs148754219 allele was the only in-frame AUG present in the Clone E construct that could result in the translation of a full-length, functional luciferase protein. Similar to the Clone D constructs, the mutant Clone E construct did not produce any luciferase expression compared to the wild-type clone. Prediction of the 5′-UTR secondary structure of GRM3 mRNA by Mfold program showed that the rs148754219 polymorphism does not alter the mRNA structure. The mRNA levels of the mutant and wild-type alleles of Clone D were consistent suggesting that the variant was acting at the post-transcriptional / translational level and not at the level of transcription.

GRM3 has previously been reported to be genetically associated with schizophrenia (Harrison and Weinberger 2005; Straub and Weinberger 2006; Mossner et al. 2008). Altered levels of dimeric forms of GRM3 with no change in total GRM3 was

reported in the prefrontal cortex in schizophrenia patients (Corti et al. 2007). The GRM3 splice isoform GRM3Δ4 encodes a truncated variant of the receptor lacking the trans-membrane domain and has been detected in the human brain (Sartorius et al. 2006). Sartorius et al. (2008) hypothesised that the truncated form of the receptor encoded by GRM3Δ4 might interact with the full-length receptor by forming hetero-dimers or act as a 'decoy receptor' soaking up glutamate in active competition with the full-length receptor. Expression data results of GRM3 and GRM3Δ4 transcripts in the dorsolateral prefrontal cortex (DLPFC) and hippocampus of two case-control cohorts of schizophrenia patients and controls, CBDB/NIMH collection and Stanley Array Collection, were inconclusive as they found that the GRM3 expression was increased in schizophrenia patients only in the CBDB/NIMH collection (Sartorius et al. 2008).

Multiple proteins can be synthesized from one mRNA by alternative splicing or by the choice of a translation initiation codon. The choice as well as efficiency of translation is controlled by mRNA-specific elements such as the GC content of the 5′-UTR, the size of the 5′-UTR and the location of potentially active upstream initiation codons (Meijer and Thomas 2002). Mutations in the 5′-UTR of genes affecting translation efficiency have been extensively documented (Signori et al. 2001; Bisio et al. 2010; Lukowski et al. 2011). According to Kozak, the optimal context around an initiation codon required for efficient translation in higher eukaryotes is CCRCCAUGG, with the most conserved nucleotides at -3 position (R) being A or G, and the G at +4 considering the A of the AUG as +1 (Kozak 1999). The nucleotides at positions -6, -5, -4, -2, -1 could also influence translation efficiency when the transcripts contain a weak Kozak motif (Kozak 1986). Kozak

SNPs at positions -1 and -5 of genes, *CD40* and *GP1BA* (glycoprotein Ib, α polypeptide) and +4 position of *CHAT* (choline acetyl transferase) and *SCARB1* (scavenger receptor class B, member 1) have been reported to be associated with Grave's disease (Jacobson et al. 2005), ischemic stroke (Baker et al. 2001; Maguire et al. 2008), Alzheimer's disease (Ozturk et al. 2006) and Type 2 diabetes (Osgood et al. 2003) respectively. The 148754219 polymorphism in exon 1 is located upstream of the initiation codons of two alternatively spliced transcripts of GRM3 gene, GRM3Δ2 and GRM3Δ2Δ3. I propose that this mutation might affect GRM3 expression by altering the expression level(s) of either or both these transcripts. Firstly, the GRM3Δ2Δ3 transcript does not have a strong Kozak sequence as the triplet sequence preceding the initiation codon is CGG that lacks a purine at the -3 position. It is possible that the substitution of G allele with the A allele in the rs148754219 polymorphism at -2 position alters the strength of its Kozak sequence leading to altered expression levels of this transcript in bipolar disorder patients. Luciferase activity by the wild-type Clones D and E demonstrates that the in-frame ATG of the GRM3Δ2Δ3 isoform is efficiently recognised by scanning ribosomes, which might then initiate translation of this transcript to produce a truncated protein. Furthermore, evidence from the bioinformatic analysis and EMSA results suggests that it is likely that the mutant allele in these constructs binds to a translational repressor, thereby, suppressing reporter activity. These results provide preliminary experimental evidence that GRM3Δ2Δ3 isoform is translatable *in vivo*. Existence of a truncated form of GRM3 protein resulting from the translation of GRM3Δ2Δ3 isoform can only be confirmed by carrying out further *in vivo* localization and expression experiments.

Secondly, in the luciferase assays with Clone D constructs, the mutant allele had a debilitating impact on the reporter activity, whereby, the mRNA was being transcribed at a similar level for the wild-type and mutant constructs. Again, this might be explained by ribosomal scanning process, which is involved in the IRES (internal ribosomal entry site) -mediated cap-independent translational control of mRNA (Kozak 1978). During the scanning process, it is plausible that the rs148754219 mutation in the 5′-UTR of GRM3 gene causes the ribosomes to ignore the ATG of the GRM3Δ2Δ3 transcript, a process called 'leaky-scanning', and recognise the downstream ATG of the GRM3Δ2 transcript and initiate translation of this transcript. In the Clone E constructs, the downstream ATG of the GRM3Δ2 transcript was out-of-frame and the ATG of GRM3Δ2Δ3 transcript was the only in-frame codon to initiate translation of the full-length luciferase protein. Perhaps, this mutation causes a switch in the usage of the alternative initiation codon on exon 1 resulting in the translation to start at the downstream and out-of-frame ATG of GRM3Δ2 transcript in the mutant clone, thereby, failing to produce a functional protein. This switching of the initiation codons might be applicable to the other two downstream AUGs located in exon 2 of GRM3 corresponding to that of full-length GRM3 and GRM3Δ4 as well. Because in the present study I did not incorporate the downstream sequence from exon 2 in the plasmid constructs (to prevent complications in the methodology and interpretation of the results), it is rather difficult to comment on the impact of this mutation on the translation efficiency of full-length GRM3 and GRM3Δ4. A significant regional variation in the mRNA expression and splicing of the MAPT (microtubule-associated protein tau) gene within the human brain has been reported in a study including brain samples from

439 normal individuals (Trabzuni et al. 2012). The Trabzuni et al. (2012) study found that the H1/H2 haplotypes were associated with the expression of exon 3 containing isoforms in human brain. Thus, evidence is emerging suggesting that genetic factors contributing to neurodegenerative diseases may operate by changing mRNA splicing in different brain regions as opposed to the overall expression of the MAPT gene. Therefore it is possible that altered transcript regulation caused by rs148754219 might be one of the potential molecular mechanisms behind the genetic association of this variant in GRM3 with bipolar disorder.

The modulation of excitatory neurotransmission via mGluRs in treating anxiety and stress-related disorders has been increasingly recognised in the last decade (Spooren et al. 2001; Schoepp et al. 2003; Linden et al. 2005). Altered expression levels of mGluR3 are well described in the pathology of several psychiatric disorders such as depression, anxiety and schizophrenia. Preliminary expression data in animals suggest that mGluR2/3 is reduced in the hippocampus in animal models of depression such as the olfactory bulbectomy in mice (Wieronska et al. 2008), and in spontaneously depressed flinders sensitive line rats when compared to controls (Matrisciano et al. 2008). In a recent study by Feyissa et al. (2011), significantly increased levels of mGluR2/3 protein was observed in the prefrontal cortex of MDD subjects compared to controls. Additionally, reports on elevated (Gupta et al. 2005) and reduced expression (Gonzalez-Maeso et al. 2008; Ghose et al. 2009) levels of mGluR2/mGluR3 in the frontal cortex of schizophrenia patients have previously been reported. Using a systematic approach combining exon capture and massive parallel sequencing of GPCR family genes in melanoma, Prickett et al. (2011)

identified mutations in GRM3 activating MEK, leading to increased anchorage-independent growth and migration of melanoma cells.

Several mGluR2/3 agonists and antagonists have been reported. Clinical studies on mGLuR2/3 agonist LY354740 in the treatment of anxiety (Schoepp et al. 2003; Dunayevich et al. 2008) and its LY404309 pro-drug LY2140023 in the treatment of schizophrenia (Patil et al. 2007) have been conducted with positive outcomes. MGS0039 is a selective group II mGluR antagonist that has antidepressant-like activity in rat models (Chaki et al. 2004). In the UCL sample, the Kozak sequence variant was convincingly over-represented in bipolar disorder cases compared to controls with an odds ratio above 4, suggestive of moderate to high risk to bipolar disorder. Further screening of normal comparison subjects in the general population is needed to establish the distribution of the mutation. This is the first study to date which reports a possible, though unconfirmed, rare and functional base pair change that is significantly associated with bipolar disorder. The finding needs confirmation in additional samples of both unipolar and BPAD as well as in alcoholism where unipolar affective disorder is known to be a primary aetiological factor. The knowledge of expression of the alternatively spliced transcripts of GRM3 and their respective protein products in bipolar disorder, schizophrenia, alcohol dependent and normal individuals may provide more insight into the regulation of this gene under these pathological conditions. Thus, further expression and immunochemical studies on bipolar disorder, schizophrenia and alcohol dependence should be designed to quantify and localize the alternative transcripts of GRM3 gene, which might be altered in these patients.

Chapter 7

General Discussion

Bipolar disorder is a highly heritable disorder of the brain with increasing evidence pointing towards the involvement of multiple risk variants in the aetiology of the disease. Even with ever advancing technology, the underlying genetic and pathophysiological components of the disorder remain elusive. Intensive research in the field of bipolar disorder, including numerous family linkage studies, association studies and GWAS, has implicated genes on almost all the chromosomes in the pathogenicity of bipolar disorder. In this thesis, I aimed to investigate the genetic basis of bipolar disorder by examining a susceptibility locus based on linkage results and two metabotropic glutamate receptor genes based on GWAS results. In reference to the aims of this thesis, the following findings are presented:

1. The fine mapping study showed positive association with the microsatellite marker D1S243 (P=0.037) leading us to further investigate the most proximal gene to this marker, PRKCZ, with increased marker density. Marker rs3128396 in PRKCZ was significantly associated with bipolar disorder in the UCL1 sample and the association improved further in bipolar disorder females for the two SNPs, rs3128296 and rs3128309 on gender-specific association analysis. However, a different marker rs2503706, which was not associated in the female-only analysis showed association in the male-only analysis. Haplotype association analysis also revealed significantly associated disease haplotypes in the gene.
2. The replication of the three most significantly associated GRM7 SNPs in the UCL1 GWAS, rs1508724, rs11710946, and rs6769814, was not successful in the UCL2 sample. Analysis of the combined genotype data along with a meta-analysis of the two study samples found that two of the SNPs,

rs1508724 and rs6769814, remained significant in the combination analysis whereas only rs6769814 reached significance under the fixed-effects model.

3. Resequencing of the GRM7 gene in the haplotype selected cases and random controls identified a 3'-UTR variant rs56173829 which was significantly associated with bipolar disorder in the complete sample (UCL1 and UCL2), however, it was more common in controls compared to cases. Bioinformatic analyses suggested that the variant altered the centroid secondary structure of the RNA and also changed the miRNA binding sites. Three bipolar-only mutations (2 novel and one previously published) and one novel control-only mutation in GRM7 with extremely low frequency were also identified in this study.
4. The CNVs identified in the UCL1 bipolar GWAS were successfully validated using QRT-PCR.
5. The replication of the three most significantly associated GRM3 SNPs in the UCL1 bipolar GWAS, rs2237563, rs274621 and rs2158786, was not successful in the UCL2 sample. Genotype data from both the studies was subjected to a combined analysis and a meta-analysis resulting in SNPs rs2237563 and rs2158786, whereas only SNP rs2158786 becoming significantly associated with bipolar disorder respectively.
6. Resequencing GRM3 gene in haplotype selected cases and random controls identified a 5′-UTR variant rs148754219, which on genotyping in the complete sample was found to be significantly associated with bipolar. Bioinformatic analysis of the variant suggested that the rare allele of the rs148754219 variant altered the binding of transcription factors.

7. Electromobility shift assays and luciferase reporter assays were employed to investigate the mechanism underlying the effect of rs148754219 variant on GRM3 gene expression. A specific DNA-protein complex was observed for the mutant allele of the variant and not for the common allele in the EMSA experiments confirming the predictions of the bioinformatic analysis that the rare allele of the variant introduced new transcription factor binding sites. Further characterisation of the rs148754219 variant using luciferase reporter assays suggested that, though the mutation affected gene expression via both transcriptional and post-translational mechanisms, it affected more evidently via a post-translational mechanism.

The most promising finding of this thesis is the identification of a possible functional variant in GRM3, rs148754219 that is significantly associated with bipolar disorder. Apart from the current study, the rs148754219 variant has only been reported in the 1000 genomes data suggestive of its low frequency. This reinforces the theory of multiple rare variants in the causation of complex disorders like bipolar disorder. It is plausible that owing to its low frequency and its absence in the microarray chips used in GWAS the discovery of this mutation was not possible in other studies. The location of this variant in the GRM3 gene is of particular interest, being present in the 5′-UTR (Exon 1) of the gene that has been previously reported to contain a silencing element and also forming part of the promoter region. Additionally, this mutation is present in the Kozak sequence of a GRM3 isoform suggesting that it might play an important role in the translation of this particular transcript.

On the basis of EMSA and reporter assays, I predict that the rare allele of the rs148754219 variant binds to a repressor protein/transcription factor causing

downregulation or complete abolishment of gene expression. However, it is also possible that this variant alters the translation efficiency of the isoforms of GRM3 gene causing changes in the relative expression of these isoforms *in vivo*. The results from these *in vitro* studies need to be supported by primary culture and animal studies before arriving at a firm conclusion.

Alternative splicing resulting in four isoforms of GRM3 gene has been documented in human brain. Previously, altered expression of a GRM3 isoform GRM3Δ4 has been observed in schizophrenic patients. However, the relative expression of all the GRM3 isoforms in a normal individual and in an individual with bipolar disorder is yet to be investigated. Interestingly, the rs148754219 variant showed association with schizophrenia and trend towards association with alcoholism (P = 0.024 and 0.087 respectively) in this study. It is known that bipolar disorder is highly comorbid with alcohol dependence and schizophrenia, making a possible association with these disorders not unexpected. Confirmation of the association of the rs148754219 variant in other bipolar disorder samples as well as in schizophrenia and alcohol dependence syndrome is required. Further expression and immunochemical studies on bipolar disorder and schizophrenia ought to be designed to quantify and localize the alternative transcripts of GRM3 gene, which might be altered in these patients. In future, lymphoblastoid, fibroblast, genome-edited and iPS cell-lines from affected patients can be cultured and used for the study of the neurobiology of mGluR3 and targeted treatments for mutation carriers. Additionally, the findings in PRKCZ and GRM7 also warrant further investigation in other samples. Functional assays aimed at understanding the role of the 3'-UTR variant in GRM7, rs56173829, in the pathology of bipolar disorder also need to be undertaken.

Recent GWAS in bipolar disorder have identified common non-coding polymorphisms in some genes, such as ANK3 and CACNA1C that have been replicated in several samples. However, bipolar disorder being a complex disorder, allelic heterogeneity must play a strong role in the genetic aetiology of bipolar disorder and larger samples are needed. Lack of complete frequent replications in association is to be expected if disease allele frequencies of 1%-2% are the norm and in many samples they may not be detectable. Epigenetic mechanisms causing modifications in the genome also increase the complexity of the genomic variation and provide a mechanism for modification of the bipolar disorder phenotype. Combined efforts from different research groups involving collaborative GWAS and meta-analysis have aided the search for genetic variants contributing to bipolar disorder. New analytical approaches such as pathway analysis of the genes in related systems involved in the disorder are currently being employed to identify novel susceptibility genes. Furthermore, next-generation sequencing seems to be the way forward for tackling allelic heterogeneity underlying these disorders. With the discovery of susceptibility genes and their susceptibility base pair changes, the development of treatments better targeted towards individual patients is already a viable proposition.

References

Abou Jamra, R., R. Fuerst, et al. (2007). "The first genomewide interaction and locus-heterogeneity linkage scan in bipolar affective disorder: strong evidence of epistatic effects between loci on chromosomes 2q and 6q." Am J Hum Genet **81**(5): 974-86.

Adzhubei, I. A., S. Schmidt, et al. (2010). "A method and server for predicting damaging missense mutations." Nat Methods **7**(4): 248-9.

Akiskal, H. S. (2005). Mood disorders: Clinical features. New York, Lippincott Williams & Wilkins.

Alaerts, M. and J. Del-Favero (2009). "Searching genetic risk factors for schizophrenia and bipolar disorder: learn from the past and back to the future." Hum Mutat **30**(8): 1139-52.

Alliey-Rodriguez, N., D. Zhang, et al. (2011). "Genome-wide association study of personality traits in bipolar patients." Psychiatr Genet **21**(4): 190-4.

Althoff, R. R., S. V. Faraone, et al. (2005). "Family, twin, adoption, and molecular genetic studies of juvenile bipolar disorder." Bipolar Disord **7**(6): 598-609.

Altman, S., S. Haeri, et al. (2006). "Predictors of relapse in bipolar disorder: A review." J Psychiatr Pract **12**(5): 269-82.

Angst, J. and A. Marneros (2001). "Bipolarity from ancient to modern times: conception, birth and rebirth." J Affect Disord **67**(1-3): 3-19.

Anwyl, R. (1999). "Metabotropic glutamate receptors: electrophysiological properties and role in plasticity." Brain Res Brain Res Rev **29**(1): 83-120.

APA (1980). Diagnostic and statistical manual of mental disorders. Washington DC., American Psychiatric Association.

APA (1987). Diagnostic and Statistical Manual of Mental Disorders (DSM-IIIR).

APA (1994). Diagnostic and Statistical Manual of Mental Disorders (DSM-IV).

APA (2000). Diagnostic and Statistical Manual of Mental disorders, Text revision. Washington DC.,, American Psychiatric Association.

Badner, J. A. and E. S. Gershon (2002). "Meta-analysis of whole-genome linkage scans of bipolar disorder and schizophrenia." Mol Psychiatry **7**(4): 405-11.

Badner, J. A., D. Koller, et al. (2011). "Genome-wide linkage analysis of 972 bipolar pedigrees using single-nucleotide polymorphisms." Mol Psychiatry **17**(8): 818-26.

Bailey-Wilson, J. E. and A. F. Wilson (2011). "Linkage analysis in the next-generation sequencing era." Hum Hered **72**(4): 228-36.

Baker, R. I., J. Eikelboom, et al. (2001). "Platelet glycoprotein Ibalpha Kozak polymorphism is associated with an increased risk of ischemic stroke." Blood **98**(1): 36-40.

Baldessarini, R. J., L. Bolzani, et al. (2010). "Onset-age of bipolar disorders at six international sites." J Affect Disord **121**(1-2): 143-6.

Baldessarini, R. J., L. Tondo, et al. (2012). "Age at onset versus family history and clinical outcomes in 1,665 international bipolar-I disorder patients." World Psychiatry **11**(1): 40-6.

Bandyopadhyay, G., M. P. Sajan, et al. (2002). "PKC-zeta mediates insulin effects on glucose transport in cultured preadipocyte-derived human adipocytes." J Clin Endocrinol Metab **87**(2): 716-23.

Barnett, J. H. and J. W. Smoller (2009). "The genetics of bipolar disorder." Neuroscience **164**(1): 331-43.

Baron, M. (2002). "Manic-depression genes and the new millennium: poised for discovery." Mol Psychiatry **7**(4): 342-58.
Baron, M., N. F. Freimer, et al. (1993). "Diminished support for linkage between manic depressive illness and X-chromosome markers in three Israeli pedigrees." Nat Genet **3**(1): 49-55.
Baron, M., R. Hamburger, et al. (1990). "The impact of phenotypic variation on genetic analysis: application to X-linkage in manic-depressive illness." Acta Psychiatr Scand **82**(3): 196-203.
Barrett, J. C. (2009). "Haploview: Visualization and analysis of SNP genotype data." Cold Spring Harb Protoc **2009**(10): pdb ip71.
Barrett, J. C., B. Fry, et al. (2005). "Haploview: analysis and visualization of LD and haplotype maps." Bioinformatics **21**(2): 263-5.
Bass, N. J., S. R. Datta, et al. (2009). "Evidence for the association of the DAOA (G72) gene with schizophrenia and bipolar disorder but not for the association of the DAO gene with schizophrenia." Behav Brain Funct **5**: 28.
Bauer, M., T. Glenn, et al. (2010). "Association between age of onset and mood in bipolar disorder: comparison of subgroups identified by cluster analysis and clinical observation." J Psychiatr Res **44**(16): 1170-5.
Baum, A. E., N. Akula, et al. (2008). "A genome-wide association study implicates diacylglycerol kinase eta (DGKH) and several other genes in the etiology of bipolar disorder." Mol Psychiatry **13**(2): 197-207.
Belmonte Mahon, P., M. Pirooznia, et al. (2011). "Genome-wide association analysis of age at onset and psychotic symptoms in bipolar disorder." Am J Med Genet B Neuropsychiatr Genet **156B**(3): 370-8.
Bengra, C., T. E. Mifflin, et al. (2002). "Genotyping of essential hypertension single-nucleotide polymorphisms by a homogeneous PCR method with universal energy transfer primers." Clin Chem **48**(12): 2131-40.
Berardi, P., M. Meyyappan, et al. (2003). "A novel transcriptional inhibitory element differentially regulates the cyclin D1 gene in senescent cells." J Biol Chem **278**(9): 7510-9.
Bergen, S. E., C. T. O'Dushlaine, et al. (2012). "Genome-wide association study in a Swedish population yields support for greater CNV and MHC involvement in schizophrenia compared with bipolar disorder." Mol Psychiatry (Epub).
Berrettini, W. (1998). "Progress and pitfalls: bipolar molecular linkage studies." J Affect Disord **50**(2-3): 287-97.
Biederman, J., A. Kwon, et al. (2004). "Absence of gender differences in pediatric bipolar disorder: findings from a large sample of referred youth." J Affect Disord **83**(2-3): 207-14.
Bisio, A., S. Nasti, et al. (2010). "Functional analysis of CDKN2A/p16INK4a 5'-UTR variants predisposing to melanoma." Hum Mol Genet **19**(8): 1479-91.
Bodmer, W. and C. Bonilla (2008). "Common and rare variants in multifactorial susceptibility to common diseases." Nat Genet **40**(6): 695-701.
Bolonna, A. A., R. W. Kerwin, et al. (2001). "Polymorphisms in the genes for mGluR types 7 and 8: association studies with schizophrenia." Schizophr Res **47**(1): 99-103.
Bondy, B. (2011). "Genetics in psychiatry: are the promises met?" World J Biol Psychiatry **12**(2): 81-8.
Breen, G., C. M. Lewis, et al. (2011a). "Replication of association of 3p21.1 with susceptibility to bipolar disorder but not major depression." Nat Genet **43**(1): 3-5

Breen, G., B. T. Webb, et al. (2011b). "A genome-wide significant linkage for severe depression on chromosome 3: the depression network study." Am J Psychiatry **168**(8): 840-7.
Brody, S. A., F. Conquet, et al. (2003). "Disruption of prepulse inhibition in mice lacking mGluR1." Eur J Neurosci **18**(12): 3361-6.
Brody, S. A., S. C. Dulawa, et al. (2004). "Assessment of a prepulse inhibition deficit in a mutant mouse lacking mGlu5 receptors." Mol Psychiatry **9**(1): 35-41.
Brugue, E. and E. Vieta (2007). "Atypical antipsychotics in bipolar depression: neurobiological basis and clinical implications." Prog Neuropsychopharmacol Biol Psychiatry **31**(1): 275-82.
Burdick, K. E., R. J. Braga, et al. (2007). "Cognitive dysfunction in bipolar disorder: future place of pharmacotherapy." CNS Drugs **21**(12): 971-81.
Cade, J. F. (1949). "Lithium salts in the treatment of psychotic excitement." Med J Aust **2**(10): 349-52.
Calabrese, J. R., C. L. Bowden, et al. (2003). "A placebo-controlled 18-month trial of lamotrigine and lithium maintenance treatment in recently depressed patients with bipolar I disorder." J Clin Psychiatry **64**(9): 1013-24.
Callaerts-Vegh, Z., T. Beckers, et al. (2006). "Concomitant deficits in working memory and fear extinction are functionally dissociated from reduced anxiety in metabotropic glutamate receptor 7-deficient mice." J Neurosci **26**(24): 6573-82.
Cartmell, J. and D. D. Schoepp (2000). "Regulation of neurotransmitter release by metabotropic glutamate receptors." J Neurochem **75**(3): 889-907.
Cassidy, F., C. Zhao, et al. (2007). "Genome-wide scan of bipolar disorder and investigation of population stratification effects on linkage: support for susceptibility loci at 4q21, 7q36, 9p21, 12q24, 14q24, and 16p13." Am J Med Genet B Neuropsychiatr Genet **144B**(6): 791-801.
Chaki, S., R. Yoshikawa, et al. (2004). "MGS0039: a potent and selective group II metabotropic glutamate receptor antagonist with antidepressant-like activity." Neuropharmacology **46**(4): 457-67.
Chatterjee, S. and J. K. Pal (2009). "Role of 5'- and 3'-untranslated regions of mRNAs in human diseases." Biol Cell **101**(5): 251-62.
Chen, D. T., X. Jiang, et al. (2011). "Genome-wide association study meta-analysis of European and Asian-ancestry samples identifies three novel loci associated with bipolar disorder." Mol Psychiatry.
Chen, J. M., C. Ferec, et al. (2006). "A systematic analysis of disease-associated variants in the 3' regulatory regions of human protein-coding genes II: the importance of mRNA secondary structure in assessing the functionality of 3' UTR variants." Hum Genet **120**(3): 301-33.
Chen, Y. S., N. Akula, et al. (2004). "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
Cheng, R., S. H. Juo, et al. (2006). "Genome-wide linkage scan in a large bipolar disorder sample from the National Institute of Mental Health genetics initiative suggests putative loci for bipolar disorder, psychosis, suicide, and panic disorder." Mol Psychiatry **11**(3): 252-60.
Cherlyn, S. Y., P. S. Woon, et al. (2010). "Genetic association studies of glutamate, GABA and related genes in schizophrenia and bipolar disorder: a decade of advance." Neurosci Biobehav Rev **34**(6): 958-77.

Chiamulera, C., M. P. Epping-Jordan, et al. (2001). "Reinforcing and locomotor stimulant effects of cocaine are absent in mGluR5 null mutant mice." Nat Neurosci **4**(9): 873-4.
Choi, K. Y., H. K. Yoon, et al. (2010). "Association between Serotonin-Related Polymorphisms in 5HT2A, TPH1, TPH2 Genes and Bipolar Disorder in Korean Population." Psychiatry Investig **7**(1): 60-7.
Chorley, B. N., X. Wang, et al. (2008). "Discovery and verification of functional single nucleotide polymorphisms in regulatory genomic regions: current and developing technologies." Mutat Res **659**(1-2): 147-57.
Christian, S. L., C. W. Brune, et al. (2008). "Novel submicroscopic chromosomal abnormalities detected in autism spectrum disorder." Biol Psychiatry **63**(12): 1111-7.
Chumakov, I., M. Blumenfeld, et al. (2002). "Genetic and physiological data implicating the new human gene G72 and the gene for D-amino acid oxidase in schizophrenia." Proc Natl Acad Sci U S A **99**(21): 13675-80.
Cichon, S., T. W. Muhleisen, et al. (2011). "Genome-wide association study identifies genetic variation in neurocan as a susceptibility factor for bipolar disorder." Am J Hum Genet **88**(3): 372-81.
Cichon, S., M. M. Nothen, et al. (1996). "Systematic screening for mutations in the 5'-regulatory region of the human dopamine D1 receptor (DRD1) gene in patients with schizophrenia and bipolar affective disorder." Am J Med Genet **67**(4): 424-8.
Cichon, S., I. Winge, et al. (2008). "Brain-specific tryptophan hydroxylase 2 (TPH2): a functional Pro206Ser substitution and variation in the 5'-region are associated with bipolar affective disorder." Hum Mol Genet **17**(1): 87-97.
Collins, P. Y., V. Patel, et al. (2011). "Grand challenges in global mental health." Nature **475**(7354): 27-30.
Conn, P. J. and J. P. Pin (1997). "Pharmacology and functions of metabotropic glutamate receptors." Annu Rev Pharmacol Toxicol **37**: 205-37.
Coppotelli, G., A. Summers, et al. (2006). "Functional characterization of the 258 A/G (D2-ORFa-Gly3Asp) human type-2 deiodinase polymorphism: a naturally occurring variant increases the enzymatic activity by removing a putative repressor site in the 5' UTR of the gene." Thyroid **16**(7): 625-32.
Corti, C., G. Battaglia, et al. (2007). "The use of knock-out mice unravels distinct roles for mGlu2 and mGlu3 metabotropic glutamate receptors in mechanisms of neurodegeneration/neuroprotection." J Neurosci **27**(31): 8297-308.
Corti, C., J. H. Xuereb, et al. (2001). "Identification and characterization of the promoter region of the GRM3 gene." Biochem Biophys Res Commun **286**(2): 381-7.
Craddock, N. and I. Jones (1999). "Genetics of bipolar disorder." J Med Genet **36**(8): 585-94.
Craddock, N., M. C. O'Donovan, et al. (2007). "Phenotypic and genetic complexity of psychosis. Invited commentary on ... Schizophrenia: a common disease caused by multiple rare alleles." Br J Psychiatry **190**: 200-3.
Craddock, N. and P. Sklar (2009). "Genetics of bipolar disorder: successful start to a long journey." Trends Genet **25**(2): 99-105.
Curtis, D., G. Kalsi, et al. (2003). "Genome scan of pedigrees multiply affected with bipolar disorder provides further support for the presence of a susceptibility locus on chromosome 12q23-q24, and suggests the presence of additional loci on 1p and 1q." Psychiatr Genet **13**(2): 77-84.
Curtis, D., J. Knight, et al. (2006). "Program report: GENECOUNTING support programs." Ann Hum Genet **70**(Pt 2): 277-9.

Curtis, D., A. E. Vine, et al. (2008). "A simple method for assessing the strength of evidence for association at the level of the whole gene." Adv Appl Bioinform Chem **1**: 115-20.
Curtis, D., A. E. Vine, et al. (2011). "Case-case genome-wide association analysis shows markers differentially associated with schizophrenia and bipolar disorder and implicates calcium channel genes." Psychiatr Genet **21**(1): 1-4.
Dalvie, S., N. Horn, et al. (2010). "Psychosis and relapse in bipolar disorder are related to GRM3, DAOA, and GRIN2B genotype." Afr J Psychiatry (Johannesbg) **13**(4): 297-301.
Dean, B., G. Pavey, et al. (2001). "A change in the density of [(3)H]flumazenil, but not [(3)H]muscimol binding, in Brodmann's Area 9 from subjects with bipolar disorder." J Affect Disord **66**(2-3): 147-58.
Dedman, A., A. McQuillin, et al. (2012). "Sequencing of the ANKYRIN 3 gene (ANK3) encoding ankyrin G in bipolar disorder reveals a non-conservative amino acid change in a short isoform of ankyrin G." Am J Med Genet B Neuropsychiatr Genet **159B**(3): 328-35.
Devon, R. S., S. Anderson, et al. (2001). "The genomic organisation of the metabotropic glutamate receptor subtype 5 gene, and its association with schizophrenia." Mol Psychiatry **6**(3): 311-4.
Diaz-Asper, C. M., T. E. Goldberg, et al. (2008). "Genetic variation in catechol-O-methyltransferase: effects on working memory in schizophrenic patients, their siblings, and healthy controls." Biol Psychiatry **63**(1): 72-9.
Diflorio, A. and I. Jones (2010). "Is sex important? Gender differences in bipolar disorder." Int Rev Psychiatry **22**(5): 437-52.
Djurovic, S., O. Gustafsson, et al. (2010). "A genome-wide association study of bipolar disorder in Norwegian individuals, followed by replication in Icelandic sample." J Affect Disord **126**(1-2): 312-6.
Dmitrzak-Weglarz, M., J. K. Rybakowski, et al. (2006). "Dopamine receptor D1 gene -48A/G polymorphism is associated with bipolar illness but not with schizophrenia in a Polish population." Neuropsychobiology **53**(1): 46-50.
Doumazane, E., P. Scholler, et al. (2010). "A new approach to analyze cell surface protein complexes reveals specific heterodimeric metabotropic glutamate receptors." Faseb J **25**(1): 66-77.
Dunayevich, E., J. Erickson, et al. (2008). "Efficacy and tolerability of an mGlu2/3 agonist in the treatment of generalized anxiety disorder." Neuropsychopharmacology **33**(7): 1603-10.
Dunner, D. L., J. L. Fleiss, et al. (1976). "The course of development of mania in patients with recurrent depression." Am J Psychiatry **133**(8): 905-8.
Duvoisin, R. M., C. Zhang, et al. (2005). "Increased measures of anxiety and weight gain in mice lacking the group III metabotropic glutamate receptor mGluR8." Eur J Neurosci **22**(2): 425-36.
Edvardsen, J., S. Torgersen, et al. (2008). "Heritability of bipolar spectrum disorders. Unity or heterogeneity?" J Affect Disord **106**(3): 229-40.
Egan, M. F., T. E. Goldberg, et al. (2001). "Effect of COMT Val108/158 Met genotype on frontal lobe function and risk for schizophrenia." Proc Natl Acad Sci U S A **98**(12): 6917-22.
Egeland, J. A. (1988). "A genetic study of manic-depressive disorder among the old order Amish of Pennsylvania." Pharmacopsychiatry **21**(2): 74-5.
Ekelund, J., W. Hennah, et al. (2004). "Replication of 1q42 linkage in Finnish schizophrenia pedigrees." Mol Psychiatry **9**(11): 1037-41.

Ekelund, J., I. Hovatta, et al. (2001). "Chromosome 1 loci in Finnish schizophrenia families." Hum Mol Genet **10**(15): 1611-7.
Ekholm, J. M., P. Pekkarinen, et al. (2002). "Bipolar disorder susceptibility region on Xq24-q27.1 in Finnish families." Mol Psychiatry **7**(5): 453-9.
Elashoff, M., B. W. Higgs, et al. (2007). "Meta-analysis of 12 genomic studies in bipolar disorder." J Mol Neurosci **31**(3): 221-43.
Elia, J., J. T. Glessner, et al. (2011). "Genome-wide copy number variation study associates metabotropic glutamate receptor gene networks with attention deficit hyperactivity disorder." Nat Genet **44**(1): 78-84.
Elliott, K. S., E. Zeggini, et al. (2010). "Evaluation of association of HNF1B variants with diverse cancers: collaborative analysis of data from 19 genome-wide association studies." PLoS One **5**(5): e10858.
Etain, B., F. Mathieu, et al. (2006). "Genome-wide scan for genes involved in bipolar affective disorder in 70 European families ascertained through a bipolar type I early-onset proband: supportive evidence for linkage at 3p14." Mol Psychiatry **11**(7): 685-94.
Fabbri, C., A. Drago, et al. (2012). "Early antidepressant efficacy modulation by glutamatergic gene variants in the STAR()D." Eur Neuropsychopharmacol (Epub).
Fairbrother, W. G., R. F. Yeh, et al. (2002). "Predictive identification of exonic splicing enhancers in human genes." Science **297**(5583): 1007-13.
Fallin, M. D., V. K. Lasseter, et al. (2005). "Bipolar I disorder and schizophrenia: a 440-single-nucleotide polymorphism screen of 64 candidate genes among Ashkenazi Jewish case-parent trios." Am J Hum Genet **77**(6): 918-36.
Fanous, A. H., F. A. Middleton, et al. (2012). "Genetic overlap of schizophrenia and bipolar disorder in a high-density linkage survey in the Portuguese Island population." Am J Med Genet B Neuropsychiatr Genet **159B**(4): 383-91.
Fatemi, S. H., T. D. Folsom, et al. (2011). "Deficits in GABA(B) receptor system in schizophrenia and mood disorders: a postmortem study." Schizophr Res **128**(1-3): 37-43.
Feeley Kearney, J. A. and R. L. Albin (2003). "mGluRs: a target for pharmacotherapy in Parkinson disease." Exp Neurol **184 Suppl 1**: S30-6.
Feighner, J. P., E. Robins, et al. (1972). "Diagnostic criteria for use in psychiatric research." Arch Gen Psychiatry **26**(1): 57-63.
Fell, M. J., K. A. Svensson, et al. (2008). "Evidence for the role of metabotropic glutamate (mGlu)2 not mGlu3 receptors in the preclinical antipsychotic pharmacology of the mGlu2/3 receptor agonist (-)-(1R,4S,5S,6S)-4-amino-2-sulfonylbicyclo[3.1.0]hexane-4,6-dicarboxylic acid (LY404039)." J Pharmacol Exp Ther **326**(1): 209-17.
Fendt, M., S. Schmid, et al. (2008). "mGluR7 facilitates extinction of aversive memories and controls amygdala plasticity." Mol Psychiatry **13**(10): 970-9.
Feng, Y., K. Kapornai, et al. (2010). "Association of the GABRD gene and childhood-onset mood disorders." Genes Brain Behav **9**(6): 668-72.
Ferreira, M. A., M. C. O'Donovan, et al. (2008). "Collaborative genome-wide association analysis supports a role for ANK3 and CACNA1C in bipolar disorder." Nat Genet **40**(9): 1056-8.
Feyissa, A. M., W. L. Woolverton, et al. (2010). "Elevated level of metabotropic glutamate receptor 2/3 in the prefrontal cortex in major depression." Prog Neuropsychopharmacol Biol Psychiatry **34**(2): 279-83.

Feyissa, A. M., W. L. Woolverton, et al. (2011). "Elevated level of metabotropic glutamate receptor 2/3 in the prefrontal cortex in major depression." Prog Neuropsychopharmacol Biol Psychiatry **34**(2): 279-83.
Filipowicz, W., S. N. Bhattacharyya, et al. (2008). "Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight?" Nat Rev Genet **9**(2): 102-14.
Flor, P. J. and F. C. Acher (2012). "Orthosteric versus allosteric GPCR activation: The great challenge of group-III mGluRs." Biochem Pharmacol **84**(4): 414-24.
Fountoulakis, K. N., J. R. Kelsoe, et al. (2012). "Receptor targets for antidepressant therapy in bipolar disorder: an overview." J Affect Disord **138**(3): 222-38.
Fountoulakis, K. N. and E. Vieta (2008). "Treatment of bipolar disorder: a systematic review of available data and clinical perspectives." Int J Neuropsychopharmacol **11**(7): 999-1029.
Fountoulakis, K. N., E. Vieta, et al. (2005). "Treatment guidelines for bipolar disorder: a critical review." J Affect Disord **86**(1): 1-10.
Francks, C., F. Tozzi, et al. (2010). "Population-based linkage analysis of schizophrenia and bipolar case-control cohorts identifies a potential susceptibility locus on 19q13." Mol Psychiatry **15**(3): 319-25.
Friedman, J. I., T. Vrijenhoek, et al. (2008). "CNTNAP2 gene dosage variation is associated with schizophrenia and epilepsy." Mol Psychiatry **13**(3): 261-6.
Fullerton, J. M., J. A. Donald, et al. (2009). "Two-dimensional genome scan identifies multiple genetic interactions in bipolar affective disorder." Biol Psychiatry **67**(5): 478-86.
Fullerton, J. M., Z. Liu, et al. (2008). "Genome screen of 15 Australian bipolar affective disorder pedigrees supports previously identified loci for bipolar susceptibility genes." Psychiatr Genet **18**(4): 156-61.
Ganda, C., S. G. Schwab, et al. (2009). "A family-based association study of DNA sequence variants in GRM7 with schizophrenia in an Indonesian population." Int J Neuropsychopharmacol **12**(9): 1283-9.
Gershon, E. S., N. Alliey-Rodriguez, et al. (2011). "After GWAS: searching for genetic risk for schizophrenia and bipolar disorder." Am J Psychiatry **168**(3): 253-6.
Gershon, E. S., J. Hamovit, et al. (1982). "A family study of schizoaffective, bipolar I, bipolar II, unipolar, and normal control probands." Arch Gen Psychiatry **39**(10): 1157-67.
Ghose, S., K. A. Gleason, et al. (2009). "Differential expression of metabotropic glutamate receptors 2 and 3 in schizophrenia: a mechanism for antipsychotic drug action?" Am J Psychiatry **166**(7): 812-20.
Gil-Sanz, C., J. M. Delgado-Garcia, et al. (2008). "Involvement of the mGluR1 receptor in hippocampal synaptic plasticity and associative learning in behaving mice." Cereb Cortex **18**(7): 1653-63.
Girirajan, S. and E. E. Eichler (2011). "De novo CNVs in bipolar disorder: recurrent themes or new directions?" Neuron **72**(6): 885-7.
Glessner, J. T., K. Wang, et al. (2009). "Autism genome-wide copy number variation reveals ubiquitin and neuronal genes." Nature **459**(7246): 569-73.
Gonzalez-Maeso, J., R. L. Ang, et al. (2008). "Identification of a serotonin/glutamate receptor complex implicated in psychosis." Nature **452**(7183): 93-7.
Gottesman, II and T. D. Gould (2003). "The endophenotype concept in psychiatry: etymology and strategic intentions." Am J Psychiatry **160**(4): 636-45.
Graham, F. L., J. Smiley, et al. (1977). "Characteristics of a human cell line transformed by DNA from human adenovirus type 5." J Gen Virol **36**(1): 59-74.

Grant, B. F., F. S. Stinson, et al. (2005). "Prevalence, correlates, and comorbidity of bipolar I disorder and axis I and II disorders: results from the National Epidemiologic Survey on Alcohol and Related Conditions." J Clin Psychiatry **66**(10): 1205-15.

Greco, B., S. Lopez, et al. (2009). "Metabotropic glutamate 7 receptor subtype modulates motor symptoms in rodent models of Parkinson's disease." J Pharmacol Exp Ther **332**(3): 1064-71.

Greenwood, T. A., C. M. Nievergelt, et al. (2012). "Further evidence for linkage of bipolar disorder to chromosomes 6 and 17 in a new independent pedigree series." Bipolar Disord **14**(1): 71-9.

Grillo, G., A. Turi, et al. (2010). "UTRdb and UTRsite (RELEASE 2010): a collection of sequences and regulatory motifs of the untranslated regions of eukaryotic mRNAs." Nucleic Acids Res **38**(Database issue): D75-80.

Grozeva, D., G. Kirov, et al. (2010). "Rare copy number variants: a point of rarity in genetic risk for bipolar disorder and schizophrenia." Arch Gen Psychiatry **67**(4): 318-27.

Gupta, D. S., R. E. McCullumsmith, et al. (2005). "Metabotropic glutamate receptor protein expression in the prefrontal cortex and striatum in schizophrenia." Synapse **57**(3): 123-31.

Gurling, H. (1986). "Candidate genes and favoured loci: strategies for molecular genetic research into schizophrenia, manic depression, autism, alcoholism and Alzheimer's disease." Psychiatr Dev **4**(4): 289-309.

Gurling, H. (1998). "Chromosome 21 workshop." Psychiatr Genet **8**(2): 109-13.

Gurling, H. M., R. P. Sherrington, et al. (1988). "Molecular genetics and heterogeneity in manic depression." Mol Neurobiol **2**(2): 125-32.

Gutierrez-Rojas, L., D. Jurado, et al. (2011). "Factors associated with work, social life and family life disability in bipolar disorder patients." Psychiatry Res **186**(2-3): 254-60.

Gyetvai, B., A. Simonyi, et al. (2011). "mGluR7 genetics and alcohol: intersection yields clues for addiction." Neurochem Res **36**(6): 1087-100.

Hallahan, B., J. Newell, et al. (2011). "Structural magnetic resonance imaging in bipolar disorder: an international collaborative mega-analysis of individual adult patient data." Biol Psychiatry **69**(4): 326-35.

Hamshere, M. L., T. G. Schulze, et al. (2009). "Mood-incongruent psychosis in bipolar disorder: conditional linkage analysis shows genome-wide suggestive linkage at 1q32.3, 7p13 and 20q13.31." Bipolar Disord **11**(6): 610-20.

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

Hashimoto, K. (2011). "The role of glutamate on the action of antidepressants." Prog Neuropsychopharmacol Biol Psychiatry **35**(7): 1558-68.

Hashimoto, K., A. Sawa, et al. (2007). "Increased levels of glutamate in brains from patients with mood disorders." Biol Psychiatry **62**(11): 1310-6.

Hattori, E., C. Liu, et al. (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): 1131-40.

Hattori, E., T. Toyota, et al. (2009). "Preliminary genome-wide association study of bipolar disorder in the Japanese population." Am J Med Genet B Neuropsychiatr Genet **150B**(8): 1110-7.

Hennah, W., P. Thomson, et al. (2009). "DISC1 association, heterogeneity and interplay in schizophrenia and bipolar disorder." Mol Psychiatry **14**(9): 865-73.

Hernandez, A. I., N. Blace, et al. (2003). "Protein kinase M zeta synthesis from a brain mRNA encoding an independent protein kinase C zeta catalytic domain. Implications for the molecular mechanism of memory." J Biol Chem **278**(41): 40305-16.

Herrmann, M. G., J. D. Durtschi, et al. (2006). "Amplicon DNA melting analysis for mutation scanning and genotyping: cross-platform comparison of instruments and dyes." Clin Chem **52**(3): 494-503.

Hinney, A., A. Scherag, et al. (2011). "Genome-wide association study in German patients with attention deficit/hyperactivity disorder." Am J Med Genet B Neuropsychiatr Genet **156B**(8): 888-97.

Iafrate, A. J., L. Feuk, et al. (2004). "Detection of large-scale variation in the human genome." Nat Genet **36**(9): 949-51.

International Hapmap Consortium (2005). "A haplotype map of the human genome." Nature **437**(7063): 1299-320.

Jackowski, A. P., G. M. Filho, et al. (2012). "The involvement of the orbitofrontal cortex in psychiatric disorders: an update of neuroimaging findings." Rev Bras Psiquiatr **34**(2): 207-12.

Jacobson, E. M., E. Concepcion, et al. (2005). "A Graves' disease-associated Kozak sequence single-nucleotide polymorphism enhances the efficiency of CD40 gene translation: a case for translational pathophysiology." Endocrinology **146**(6): 2684-91.

Joo, A., H. Shibata, et al. (2001). "Structure and polymorphisms of the human metabotropic glutamate receptor type 2 gene (GRM2): analysis of association with schizophrenia." Mol Psychiatry **6**(2): 186-92.

Kandaswamy, R., A. McQuillin, et al. (2012). "Tests of linkage and allelic association between markers in the 1p36 PRKCZ (protein kinase C zeta) gene region and bipolar affective disorder." Am J Med Genet B Neuropsychiatr Genet **159B**(2): 201-9.

Kaneva, R., V. Milanova, et al. (2009). "Bipolar disorder in the Bulgarian Gypsies: genetic heterogeneity in a young founder population." Am J Med Genet B Neuropsychiatr Genet **150B**(2): 191-201.

Kawakami, N., E. A. Abdulghani, et al. (2012). "Early-life mental disorders and adult household income in the world mental health surveys." Biol Psychiatry **72**(3): 228-37.

Kelsoe, J. R., A. D. Sadovnick, et al. (1996). "Possible locus for bipolar disorder near the dopamine transporter on chromosome 5." Am J Med Genet **67**(6): 533-40.

Kelsoe, J. R., M. A. Spence, et al. (2001). "A genome survey indicates a possible susceptibility locus for bipolar disorder on chromosome 22." Proc Natl Acad Sci U S A **98**(2): 585-90.

Kennedy, N., J. Boydell, et al. (2005). "Gender differences in incidence and age at onset of mania and bipolar disorder over a 35-year period in Camberwell, England." Am J Psychiatry **162**(2): 257-62.

Kessler, R. C., W. T. Chiu, et al. (2005). "Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication." Arch Gen Psychiatry **62**(6): 617-27.

Kessler, R. C., K. A. McGonagle, et al. (1994). "Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey." Arch Gen Psychiatry **51**(1): 8-19.

Khan, A., A. E. Brodhead, et al. (2005). "Sex differences in antidepressant response in recent antidepressant clinical trials." J Clin Psychopharmacol **25**(4): 318-24.

Kieseppa, T., T. Partonen, et al. (2004). "High concordance of bipolar I disorder in a nationwide sample of twins." Am J Psychiatry **161**(10): 1814-21.

Klaning, U., T. M. Laursen, et al. (2004). "Is the risk of bipolar disorder in twins equal to the risk in singletons? A nationwide register-based study." J Affect Disord **81**(2): 141-5.
Kornstein, S. G., A. F. Schatzberg, et al. (2000). "Gender differences in treatment response to sertraline versus imipramine in chronic depression." Am J Psychiatry **157**(9): 1445-52.
Kozak, M. (1978). "How do eucaryotic ribosomes select initiation regions in messenger RNA?" Cell **15**(4): 1109-23.
Kozak, M. (1986). "Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes." Cell **44**(2): 283-92.
Kozak, M. (1999). "Initiation of translation in prokaryotes and eukaryotes." Gene **234**(2): 187-208.
Kremeyer, B., J. Garcia, et al. (2010). "Genome-wide linkage scan of bipolar disorder in a Colombian population isolate replicates Loci on chromosomes 7p21-22, 1p31, 16p12 and 21q21-22 and identifies a novel locus on chromosome 12q." Hum Hered **70**(4): 255-68.
Kruglyak, L. and E. S. Lander (1995). "Complete multipoint sib-pair analysis of qualitative and quantitative traits." Am J Hum Genet **57**(2): 439-54.
Kupfer, D. J., E. Frank, et al. (2002). "Demographic and clinical characteristics of individuals in a bipolar disorder case registry." J Clin Psychiatry **63**(2): 120-5.
Kupka, R. W., D. A. Luckenbaugh, et al. (2003). "Rapid and non-rapid cycling bipolar disorder: a meta-analysis of clinical studies." J Clin Psychiatry **64**(12): 1483-94.
Lachman, D. F. P. a. H. M. (1994). Genetic studies in affective disorders: Overview of basic methods, current directions and critical research issues, Wiley-Interscience Publication, John Wiley & Sons, Inc.
Lachman, H. M., J. R. Kelsoe, et al. (1997). "Linkage studies suggest a possible locus for bipolar disorder near the velo-cardio-facial syndrome region on chromosome 22." Am J Med Genet **74**(2): 121-8.
Lachman, H. M., E. Pedrosa, et al. (2007). "Increase in GSK3beta gene copy number variation in bipolar disorder." Am J Med Genet B Neuropsychiatr Genet **144B**(3): 259-65.
Lander, E. S. and N. J. Schork (1994). "Genetic dissection of complex traits." Science **265**(5181): 2037-48.
Lee, K. W., P. S. Woon, et al. (2012). "Genome wide association studies (GWAS) and copy number variation (CNV) studies of the major psychoses: what have we learnt?" Neurosci Biobehav Rev **36**(1): 556-71.
Lee, M. T., C. H. Chen, et al. (2011a). "Genome-wide association study of bipolar I disorder in the Han Chinese population." Mol Psychiatry **16**(5): 548-56.
Lee, S. Y., S. L. Chen, et al. (2011b). "The COMT and DRD3 genes interacted in bipolar I but not bipolar II disorder." World J Biol Psychiatry **12**(5): 385-91.
Leff, J. P., M. Fischer, et al. (1976). "A cross-national epidemiological study of mania." Br J Psychiatry **129**: 428-42.
Lesch, K. P. and R. Mossner (2006). "Inactivation of 5HT transport in mice: modeling altered 5HT homeostasis implicated in emotional dysfunction, affective disorders, and somatic syndromes." Handb Exp Pharmacol(175): 417-56.
Levy, D., M. Ronemus, et al. (2011). "Rare de novo and transmitted copy-number variation in autistic spectrum disorders." Neuron **70**(5): 886-97.
Levy, R. J., B. Xu, et al. (2012). "Copy number variation and psychiatric disease risk." Methods Mol Biol **838**: 97-113.

Levy, S., G. Sutton, et al. (2007). "The diploid genome sequence of an individual human." PLoS Biol **5**(10): e254.
Lewis, C. M., D. F. Levinson, et al. (2003). "Genome scan meta-analysis of schizophrenia and bipolar disorder, part II: Schizophrenia." Am J Hum Genet **73**(1): 34-48.
Li, X., J. Li, et al. (2010a). "Activation of mGluR7s inhibits cocaine-induced reinstatement of drug-seeking behavior by a nucleus accumbens glutamate-mGluR2/3 mechanism in rats." J Neurochem **114**(5): 1368-80.
Li, X. Y., H. G. Ko, et al. (2010b). "Alleviating neuropathic pain hypersensitivity by inhibiting PKMzeta in the anterior cingulate cortex." Science **330**(6009): 1400-4.
Linden, A. M., B. G. Johnson, et al. (2002). "Increased anxiety-related behavior in mice deficient for metabotropic glutamate 8 (mGlu8) receptor." Neuropharmacology **43**(2): 251-9.
Linden, A. M., H. Shannon, et al. (2005). "Anxiolytic-like activity of the mGLU2/3 receptor agonist LY354740 in the elevated plus maze test is disrupted in metabotropic glutamate receptor 2 and 3 knock-out mice." Psychopharmacology (Berl) **179**(1): 284-91.
Litt, M. and J. A. Luty (1989). "A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene." Am J Hum Genet **44**(3): 397-401.
Liu, J., S. H. Juo, et al. (2003). "Evidence for a putative bipolar disorder locus on 2p13-16 and other potential loci on 4q31, 7q34, 8q13, 9q31, 10q21-24, 13q32, 14q21 and 17q11-12." Mol Psychiatry **8**(3): 333-42.
Liu, Y., D. H. Blackwood, et al. (2011). "Meta-analysis of genome-wide association data of bipolar disorder and major depressive disorder." Mol Psychiatry **16**(1): 2-4.
Livak, K. J. and T. D. Schmittgen (2001). "Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method." Methods **25**(4): 402-8.
Lloyd, T., N. Kennedy, et al. (2005). "Incidence of bipolar affective disorder in three UK cities: results from the AESOP study." Br J Psychiatry **186**: 126-31.
Lopez, S., N. Turle-Lorenzo, et al. (2007). "Targeting group III metabotropic glutamate receptors produces complex behavioral effects in rodent models of Parkinson's disease." J Neurosci **27**(25): 6701-11.
Lucotte, G., A. Landoulsi, et al. (1992). "Manic depressive illness is linked to factor IX in a French pedigree." Ann Genet **35**(2): 93-5.
Lukowski, S. W., C. Bombieri, et al. (2011). "Disrupted post-transcriptional regulation of the cystic fibrosis transmembrane conductance regulator (CFTR) by a 5'UTR mutation is associated with a CFTR-related disease." Hum Mutat **32**(10): E2266-82.
Lupien, S. J., B. S. McEwen, et al. (2009). "Effects of stress throughout the lifespan on the brain, behaviour and cognition." Nat Rev Neurosci **10**(6): 434-45.
Lupski, J. R. (2007). "An evolution revolution provides further revelation." Bioessays **29**(12): 1182-4.
Luscher, B., Q. Shen, et al. (2011). "The GABAergic deficit hypothesis of major depressive disorder." Mol Psychiatry **16**(4): 383-406.
Machado-Vieira, R., L. Ibrahim, et al. (2012). "Novel glutamatergic agents for major depressive disorder and bipolar disorder." Pharmacol Biochem Behav **100**(4): 678-87.
Machado-Vieira, R. and C. A. Zarate, Jr. (2011). "Proof of concept trials in bipolar disorder and major depressive disorder: a translational perspective in the search for improved treatments." Depress Anxiety **28**(4): 267-81.

Maguire, J. M., A. Thakkinstian, et al. (2008). "Polymorphisms in platelet glycoprotein 1balpha and factor VII and risk of ischemic stroke: a meta-analysis." Stroke **39**(6): 1710-6.
Malhi, G. S., D. Adams, et al. (2009a). "Medicating mood with maintenance in mind: bipolar depression pharmacotherapy." Bipolar Disord **11 Suppl 2**: 55-76.
Malhi, G. S., D. Adams, et al. (2009b). "The management of individuals with bipolar disorder: a review of the evidence and its integration into clinical practice." Drugs **69**(15): 2063-101.
Malhotra, D., S. McCarthy, et al. (2011). "High frequencies of de novo CNVs in bipolar disorder and schizophrenia." Neuron **72**(6): 951-63.
Malhotra, D. and J. Sebat (2012). "CNVs: harbingers of a rare variant revolution in psychiatric genetics." Cell **148**(6): 1223-41.
Marchini, J., B. Howie, et al. (2007). "A new multipoint method for genome-wide association studies by imputation of genotypes." Nat Genet **39**(7): 906-13.
Marian, A. J. (2012). "Molecular genetic studies of complex phenotypes." Transl Res **159**(2): 64-79.
Marti, S. B., S. Cichon, et al. (2002). "Metabotropic glutamate receptor 3 (GRM3) gene variation is not associated with schizophrenia or bipolar affective disorder in the German population." Am J Med Genet **114**(1): 46-50.
Martin, P., A. Duran, et al. (2002). "Role of zeta PKC in B-cell signaling and function." Embo J **21**(15): 4049-57.
Massat, I., D. Souery, et al. (2002). "Excess of allele1 for alpha3 subunit GABA receptor gene (GABRA3) in bipolar patients: a multicentric association study." Mol Psychiatry **7**(2): 201-7.
Mathieu, F., M. H. Dizier, et al. (2010). "European collaborative study of early-onset bipolar disorder: Evidence for genetic heterogeneity on 2q14 according to age at onset." Am J Med Genet B Neuropsychiatr Genet **153B**(8): 1425-33.
Matrisciano, F., A. Caruso, et al. (2008). "Defective group-II metaboropic glutamate receptors in the hippocampus of spontaneously depressed rats." Neuropharmacology **55**(4): 525-31.
Matrisciano, F., I. Panaccione, et al. (2007). "Group-II metabotropic glutamate receptor ligands as adjunctive drugs in the treatment of depression: a new strategy to shorten the latency of antidepressant medication?" Mol Psychiatry **12**(8): 704-6.
Matrisciano, F., S. Scaccianoce, et al. (2005). "Metabotropic glutamate receptors and neuroadaptation to antidepressants: imipramine-induced down-regulation of beta-adrenergic receptors in mice treated with metabotropic glutamate 2/3 receptor ligands." J Neurochem **93**(5): 1345-52.
McGuffin, P., A. Farmer, et al. (1991). "A polydiagnostic application of operational criteria in studies of psychotic illness. Development and reliability of the OPCRIT system." Arch Gen Psychiatry **48**(8): 764-70.
McGuffin, P., J. Knight, et al. (2005). "Whole genome linkage scan of recurrent depressive disorder from the depression network study." Hum Mol Genet **14**(22): 3337-45.
McGuffin, P., F. Rijsdijk, et al. (2003). "The heritability of bipolar affective disorder and the genetic relationship to unipolar depression." Arch Gen Psychiatry **60**(5): 497-502.
McInnis, M. G., F. J. McMahon, et al. (1993). "Anticipation in bipolar affective disorder." Am J Hum Genet **53**(2): 385-90.
McMahon, F. J., N. Akula, et al. (2010). "Meta-analysis of genome-wide association data identifies a risk locus for major mood disorders on 3p21.1." Nat Genet **42**(2): 128-31.

McQueen, M. B., B. Devlin, et al. (2005). "Combined analysis from eleven linkage studies of bipolar disorder provides strong evidence of susceptibility loci on chromosomes 6q and 8q." Am J Hum Genet **77**(4): 582-95.
McQuillin, A., N. Bass, et al. (2011). "Analysis of genetic deletions and duplications in the University College London bipolar disorder case control sample." Eur J Hum Genet **19**(5): 588-92.
McQuillin, A., N. J. Bass, et al. (2009). "Case-control studies show that a non-conservative amino-acid change from a glutamine to arginine in the P2RX7 purinergic receptor protein is associated with both bipolar- and unipolar-affective disorders." Mol Psychiatry **14**(6): 614-20.
McQuillin, A., N. J. Bass, et al. (2006). "Fine mapping of a susceptibility locus for bipolar and genetically related unipolar affective disorders, to a region containing the C21ORF29 and TRPM2 genes on chromosome 21q22.3." Mol Psychiatry **11**(2): 134-42.
Meijer, H. A. and A. A. Thomas (2002). "Control of eukaryotic protein synthesis by upstream open reading frames in the 5'-untranslated region of an mRNA." Biochem J **367**(Pt 1): 1-11.
Mendlewicz, J., P. Simon, et al. (1987). "Polymorphic DNA marker on X chromosome and manic depression." Lancet **1**(8544): 1230-2.
Merikangas, K. R. and F. Lamers (2012). "The 'true' prevalence of bipolar II disorder." Curr Opin Psychiatry **25**(1): 19-23.
Mick, E., B. Neale, et al. (2008). "Genome-wide association study of response to methylphenidate in 187 children with attention-deficit/hyperactivity disorder." Am J Med Genet B Neuropsychiatr Genet **147B**(8): 1412-8.
Minet, E., I. Ernest, et al. (1999). "HIF1A gene transcription is dependent on a core promoter sequence encompassing activating and inhibiting sequences located upstream from the transcription initiation site and cis elements located within the 5'UTR." Biochem Biophys Res Commun **261**(2): 534-40.
Mitsukawa, K., R. Yamamoto, et al. (2005). "A selective metabotropic glutamate receptor 7 agonist: activation of receptor signaling via an allosteric site modulates stress parameters in vivo." Proc Natl Acad Sci U S A **102**(51): 18712-7.
Moghaddam, B. and D. Javitt (2011). "From revolution to evolution: the glutamate hypothesis of schizophrenia and its implication for treatment." Neuropsychopharmacology **37**(1): 4-15.
Morishima, Y., T. Miyakawa, et al. (2005). "Enhanced cocaine responsiveness and impaired motor coordination in metabotropic glutamate receptor subtype 2 knockout mice." Proc Natl Acad Sci U S A **102**(11): 4170-5.
Morton, N. E. (1955). "Sequential tests for the detection of linkage." Am J Hum Genet **7**(3): 277-318.
Moskvina, V., N. Craddock, et al. (2009). "Gene-wide analyses of genome-wide association data sets: evidence for multiple common risk alleles for schizophrenia and bipolar disorder and for overlap in genetic risk." Mol Psychiatry **14**(3): 252-60.
Mossner, R., A. Schuhmacher, et al. (2008). "Further evidence for a functional role of the glutamate receptor gene GRM3 in schizophrenia." Eur Neuropsychopharmacol **18**(10): 768-72.
Muglia, P., F. Tozzi, et al. (2010). "Genome-wide association study of recurrent major depressive disorder in two European case-control cohorts." Mol Psychiatry **15**(6): 589-601.

Muller-Oerlinghausen, B. and U. Lewitzka (2010). "Lithium reduces pathological aggression and suicidality: a mini-review." Neuropsychobiology **62**(1): 43-9.
Muto, T., D. Tsuchiya, et al. (2007a). "Expression, purification, crystallization and preliminary X-ray analysis of the ligand-binding domain of metabotropic glutamate receptor 7." Acta Crystallogr Sect F Struct Biol Cryst Commun **63**(Pt 7): 627-30.
Muto, T., D. Tsuchiya, et al. (2007b). "Structures of the extracellular regions of the group II/III metabotropic glutamate receptors." Proc Natl Acad Sci U S A **104**(10): 3759-64.
Nakajima, Y., H. Iwakabe, et al. (1993). "Molecular characterization of a novel retinal metabotropic glutamate receptor mGluR6 with a high agonist selectivity for L-2-amino-4-phosphonobutyrate." J Biol Chem **268**(16): 11868-73.
Navakkode, S., S. Sajikumar, et al. (2010). "Protein kinase Mzeta is essential for the induction and maintenance of dopamine-induced long-term potentiation in apical CA1 dendrites." Learn Mem **17**(12): 605-11.
Nicoletti, F., J. Bockaert, et al. (2011). "Metabotropic glutamate receptors: from the workbench to the bedside." Neuropharmacology **60**(7-8): 1017-41.
Niswender, C. M. and P. J. Conn (2010). "Metabotropic glutamate receptors: physiology, pharmacology, and disease." Annu Rev Pharmacol Toxicol **50**: 295-322.
Nivoli, A. M., F. Colom, et al. (2011). "New treatment guidelines for acute bipolar depression: a systematic review." J Affect Disord **129**(1-3): 14-26.
Nivoli, A. M., I. Pacchiarotti, et al. (2010). "Gender differences in a cohort study of 604 bipolar patients: the role of predominant polarity." J Affect Disord **133**(3): 443-9.
Noga, J. T., K. Vladar, et al. (2001). "A volumetric magnetic resonance imaging study of monozygotic twins discordant for bipolar disorder." Psychiatry Res **106**(1): 25-34.
Nothen, M. M., J. Erdmann, et al. (1992). "Lack of association between dopamine D1 and D2 receptor genes and bipolar affective disorder." Am J Psychiatry **149**(2): 199-201.
Nothen, M. M., V. Nieratschker, et al. (2010). "New findings in the genetics of major psychoses." Dialogues Clin Neurosci **12**(1): 85-93.
Nyegaard, M., J. E. Severinsen, et al. (2010). "Support of association between BRD1 and both schizophrenia and bipolar affective disorder." Am J Med Genet B Neuropsychiatr Genet **153B**(2): 582-91.
O'Connor, R. M., B. C. Finger, et al. (2010). "Metabotropic glutamate receptor 7: at the interface of cognition and emotion." Eur J Pharmacol **639**(1-3): 123-31.
O'Donovan, M. C., N. J. Craddock, et al. (2009). "Genetics of psychosis; insights from views across the genome." Hum Genet **126**(1): 3-12.
O'Hara, P. J., P. O. Sheppard, et al. (1993). "The ligand-binding domain in metabotropic glutamate receptors is related to bacterial periplasmic binding proteins." Neuron **11**(1): 41-52.
O'Mahony, C. M., J. A. Bravo, et al. (2010). "Comparison of hippocampal metabotropic glutamate receptor 7 (mGlu7) mRNA levels in two animal models of depression." Neurosci Lett **482**(2): 137-41.
Ohishi, H., C. Akazawa, et al. (1995). "Distributions of the mRNAs for L-2-amino-4-phosphonobutyrate-sensitive metabotropic glutamate receptors, mGluR4 and mGluR7, in the rat brain." J Comp Neurol **360**(4): 555-70.
Ohtsuki, T., M. Koga, et al. (2008). "A polymorphism of the metabotropic glutamate receptor mGluR7 (GRM7) gene is associated with schizophrenia." Schizophr Res **101**(1-3): 9-16.
Olesen, J., A. Gustavsson, et al. (2012). "The economic cost of brain disorders in Europe." Eur J Neurol **19**(1): 155-62.

Olsen, L., T. Hansen, et al. (2011). "Copy number variations in affective disorders and meta-analysis." Psychiatr Genet **21**(6): 319-22.
Osgood, D., D. Corella, et al. (2003). "Genetic variation at the scavenger receptor class B type I gene locus determines plasma lipoprotein concentrations and particle size and interacts with type 2 diabetes: the framingham study." J Clin Endocrinol Metab **88**(6): 2869-79.
Ozturk, A., S. T. DeKosky, et al. (2006). "Genetic variation in the choline acetyltransferase (CHAT) gene may be associated with the risk of Alzheimer's disease." Neurobiol Aging **27**(10): 1440-4.
Palucha, A., K. Klak, et al. (2007). "Activation of the mGlu7 receptor elicits antidepressant-like effects in mice." Psychopharmacology (Berl) **194**(4): 555-62.
Palucha, A. and A. Pilc (2007). "Metabotropic glutamate receptor ligands as possible anxiolytic and antidepressant drugs." Pharmacol Ther **115**(1): 116-47.
Palucha, A., E. Tatarczynska, et al. (2004). "Group III mGlu receptor agonists produce anxiolytic- and antidepressant-like effects after central administration in rats." Neuropharmacology **46**(2): 151-9.
Patil, S. T., L. Zhang, et al. (2007). "Activation of mGlu2/3 receptors as a new approach to treat schizophrenia: a randomized Phase 2 clinical trial." Nat Med **13**(9): 1102-7.
Pato, C. N., M. T. Pato, et al. (2004). "Genome-wide scan in Portuguese Island families implicates multiple loci in bipolar disorder: fine mapping adds support on chromosomes 6 and 11." Am J Med Genet B Neuropsychiatr Genet **127B**(1): 30-4.
Pekkarinen, P., J. Terwilliger, et al. (1995). "Evidence of a predisposing locus to bipolar disorder on Xq24-q27.1 in an extended Finnish pedigree." Genome Res **5**(2): 105-15.
Penninx, B. W., A. T. Beekman, et al. (2008). "The Netherlands Study of Depression and Anxiety (NESDA): rationale, objectives and methods." Int J Methods Psychiatr Res **17**(3): 121-40.
Pergadia, M. L., A. L. Glowinski, et al. (2011). "A 3p26-3p25 genetic linkage finding for DSM-IV major depression in heavy smoking families." Am J Psychiatry **168**(8): 848-52.
Pericak-Vance, J. L. H. a. M. (2006). Genetic analysis of complex disease. New Jersey, John Wiley & Sons, Inc.
Perlis, R. H., J. Huang, et al. (2010). "Genome-wide association study of suicide attempts in mood disorder patients." Am J Psychiatry **167**(12): 1499-507.
Pin, J. P., T. Galvez, et al. (2003). "Evolution, structure, and activation mechanism of family 3/C G-protein-coupled receptors." Pharmacol Ther **98**(3): 325-54.
Pinheiro, P. S. and C. Mulle (2008). "Presynaptic glutamate receptors: physiological functions and mechanisms of action." Nat Rev Neurosci **9**(6): 423-36.
Pini, S., V. de Queiroz, et al. (2005). "Prevalence and burden of bipolar disorders in European countries." Eur Neuropsychopharmacol **15**(4): 425-34.
Pitychoutis, P. M., A. Zisaki, et al. (2010). "Pharmacogenetic insights into depression and antidepressant response: does sex matter?" Curr Pharm Des **16**(20): 2214-23.
Post, R. M., G. S. Leverich, et al. (2010). "Early-onset bipolar disorder and treatment delay are risk factors for poor outcome in adulthood." J Clin Psychiatry **71**(7): 864-72.
Prata, D., G. Breen, et al. (2008). "Association of DAO and G72(DAOA)/G30 genes with bipolar affective disorder." Am J Med Genet B Neuropsychiatr Genet **147B**(6): 914-7.
Price, A. L., N. J. Patterson, et al. (2006). "Principal components analysis corrects for stratification in genome-wide association studies." Nat Genet **38**(8): 904-9.

Price, J. (1968). The genetics of depressive behaviour. Recent developments in affective disorders. A. C. a. A. Walk, British Journal of Psychiatry (Special Publisher 2)**:** 37-54.
Prickett, T. D., X. Wei, et al. (2011). "Exon capture analysis of G protein-coupled receptors identifies activating mutations in GRM3 in melanoma." Nat Genet **43**(11): 1119-26.
Priebe, L., F. A. Degenhardt, et al. (2012). "Genome-wide survey implicates the influence of copy number variants (CNVs) in the development of early-onset bipolar disorder." Mol Psychiatry **17**(4): 421-32.
Pritchard, J. K. and N. J. Cox (2002). "The allelic architecture of human disease genes: common disease-common variant...or not?" Hum Mol Genet **11**(20): 2417-23.
Pulver, A. E., V. K. Lasseter, et al. (1995). "Schizophrenia: a genome scan targets chromosomes 3p and 8p as potential sites of susceptibility genes." Am J Med Genet **60**(3): 252-60.
Purcell, S., S. S. Cherny, et al. (2003). "Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits." Bioinformatics **19**(1): 149-50.
Purcell, S., B. Neale, et al. (2007). "PLINK: a tool set for whole-genome association and population-based linkage analyses." Am J Hum Genet **81**(3): 559-75.
Rao, J. S., M. Kellom, et al. (2012). "Dysregulated glutamate and dopamine transporters in postmortem frontal cortex from bipolar and schizophrenic patients." J Affect Disord **136**(1-2): 63-71.
Raymont, V., D. Bettany, et al. (2003). "The Maudsley bipolar disorder project. Clinical characteristics of bipolar disorder I in a Catchment area treatment sample." Eur Psychiatry **18**(1): 13-7.
Regier, D. A., M. E. Farmer, et al. (1990). "Comorbidity of mental disorders with alcohol and other drug abuse. Results from the Epidemiologic Catchment Area (ECA) Study." Jama **264**(19): 2511-8.
Rieder, R. O., L. S. Mann, et al. (1983). "Computed tomographic scans in patients with schizophrenia, schizoaffective, and bipolar affective disorder." Arch Gen Psychiatry **40**(7): 735-9.
Risch, N. and K. Merikangas (1996). "The future of genetic studies of complex human diseases." Science **273**(5281): 1516-7.
Sajikumar, S., S. Navakkode, et al. (2005). "Synaptic tagging and cross-tagging: the role of protein kinase Mzeta in maintaining long-term potentiation but not long-term depression." J Neurosci **25**(24): 5750-6.
Sanacora, G., C. A. Zarate, et al. (2008). "Targeting the glutamatergic system to develop novel, improved therapeutics for mood disorders." Nat Rev Drug Discov **7**(5): 426-37.
Sanderson, J. L. and M. L. Dell'Acqua (2011). "AKAP signaling complexes in regulation of excitatory synaptic plasticity." Neuroscientist **17**(3): 321-36.
Sandi, C. (2004). "Stress, cognitive impairment and cell adhesion molecules." Nat Rev Neurosci **5**(12): 917-30.
Sartorius, L. J., G. Nagappan, et al. (2006). "Alternative splicing of human metabotropic glutamate receptor 3." J Neurochem **96**(4): 1139-48.
Sartorius, L. J., D. R. Weinberger, et al. (2008). "Expression of a GRM3 splice variant is increased in the dorsolateral prefrontal cortex of individuals carrying a schizophrenia risk SNP." Neuropsychopharmacology **33**(11): 2626-34.
Saus, E., A. Brunet, et al. (2010). "Comprehensive copy number variant (CNV) analysis of neuronal pathways genes in psychiatric disorders identifies rare variants within patients." J Psychiatr Res **44**(14): 971-8.

Savitz, J. B. and W. C. Drevets (2012). "Neuroreceptor imaging in depression." Neurobiol Dis (Epub).
Scherag, A., C. Dina, et al. (2010). "Two new Loci for body-weight regulation identified in a joint analysis of genome-wide association studies for early-onset extreme obesity in French and german study groups." PLoS Genet **6**(4): e1000916.
Schneck, C. D., D. J. Miklowitz, et al. (2008). "The prospective course of rapid-cycling bipolar disorder: findings from the STEP-BD." Am J Psychiatry **165**(3): 370-7; quiz 410.
Schneider, M. R., M. P. DelBello, et al. (2012). "Neuroprogression in bipolar disorder." Bipolar Disord **14**(4): 356-74.
Schoepp, D. D., R. A. Wright, et al. (2003). "LY354740, an mGlu2/3 receptor agonist as a novel approach to treat anxiety/stress." Stress **6**(3): 189-97.
Schulze, T. G. and F. J. McMahon (2003). "Genetic linkage and association studies in bipolar affective disorder: a time for optimism." Am J Med Genet C Semin Med Genet **123C**(1): 36-47.
Schumacher, J., R. A. Jamra, et al. (2004). "Examination of G72 and D-amino-acid oxidase as genetic risk factors for schizophrenia and bipolar affective disorder." Mol Psychiatry **9**(2): 203-7.
Schumacher, J., R. Kaneva, et al. (2005). "Genomewide scan and fine-mapping linkage studies in four European samples with bipolar affective disorder suggest a new susceptibility locus on chromosome 1p35-p36 and provides further evidence of loci on chromosome 4q31 and 6q24." Am J Hum Genet **77**(6): 1102-11.
Scott, L. J., P. Muglia, et al. (2009). "Genome-wide association and meta-analysis of bipolar disorder in individuals of European ancestry." Proc Natl Acad Sci U S A **106**(18): 7501-6.
Sebat, J., B. Lakshmi, et al. (2007). "Strong association of de novo copy number mutations with autism." Science **316**(5823): 445-9.
Sebat, J., B. Lakshmi, et al. (2004). "Large-scale copy number polymorphism in the human genome." Science **305**(5683): 525-8.
Segurado, R., S. D. Detera-Wadleigh, et al. (2003). "Genome scan meta-analysis of schizophrenia and bipolar disorder, part III: Bipolar disorder." Am J Hum Genet **73**(1): 49-62.
Seifuddin, F., P. B. Mahon, et al. (2012). "Meta-analysis of genetic association studies on bipolar disorder." Am J Med Genet B Neuropsychiatr Genet **159**(5): 508-18.
Selvaraj, S., D. Arnone, et al. (2012). "Grey matter differences in bipolar disorder: a meta-analysis of voxel-based morphometry studies." Bipolar Disord **14**(2): 135-45.
Serretti, A. and L. Mandelli (2008). "The genetics of bipolar disorder: genome 'hot regions,' genes, new potential candidates and future directions." Mol Psychiatry **13**(8): 742-71.
Severino, G., D. Congiu, et al. (2005). "A48G polymorphism in the D1 receptor genes associated with bipolar I disorder." Am J Med Genet B Neuropsychiatr Genet **134B**(1): 37-8.
Shaltiel, G., S. Maeng, et al. (2008). "Evidence for the involvement of the kainate receptor subunit GluR6 (GRIK2) in mediating behavioral displays related to behavioral symptoms of mania." Mol Psychiatry **13**(9): 858-72.
Sham, P. C. and D. Curtis (1995). "Monte Carlo tests for associations between disease and alleles at highly polymorphic loci." Ann Hum Genet **59**(Pt 1): 97-105.
Shi, J., J. B. Potash, et al. (2011). "Genome-wide association study of recurrent early-onset major depressive disorder." Mol Psychiatry **16**(2): 193-201.

Shibata, H., A. Tani, et al. (2009). "Association study of polymorphisms in the group III metabotropic glutamate receptor genes, GRM4 and GRM7, with schizophrenia." Psychiatry Res **167**(1-2): 88-96.
Shifman, S., M. Bronstein, et al. (2004). "COMT: a common susceptibility gene in bipolar disorder and schizophrenia." Am J Med Genet B Neuropsychiatr Genet **128B**(1): 61-4.
Shigemoto, R., A. Kinoshita, et al. (1997). "Differential presynaptic localization of metabotropic glutamate receptor subtypes in the rat hippocampus." J Neurosci **17**(19): 7503-22.
Shigemoto, R., A. Kulik, et al. (1996). "Target-cell-specific concentration of a metabotropic glutamate receptor in the presynaptic active zone." Nature **381**(6582): 523-5.
Shih, R. A., P. L. Belmonte, et al. (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): 260-83.
Shim, S. S., M. D. Hammonds, et al. (2012). "Effects of 4-weeks of treatment with lithium and olanzapine on long-term potentiation in hippocampal area CA1." Neurosci Lett **524**(1): 5-9.
Shyn, S. I., J. Shi, et al. (2011). "Novel loci for major depression identified by genome-wide association study of Sequenced Treatment Alternatives to Relieve Depression and meta-analysis of three studies." Mol Psychiatry **16**(2): 202-15.
Signori, E., C. Bagni, et al. (2001). "A somatic mutation in the 5'UTR of BRCA1 gene in sporadic breast cancer causes down-modulation of translation efficiency." Oncogene **20**(33): 4596-600.
Simonyi, A., L. A. Miller, et al. (2000). "Region-specific decline in the expression of metabotropic glutamate receptor 7 mRNA in rat brain during aging." Brain Res Mol Brain Res **82**(1-2): 101-6.
Sinclair, D., J. M. Fullerton, et al. (2012a). "Glucocorticoid receptor 1B and 1C mRNA transcript alterations in schizophrenia and bipolar disorder, and their possible regulation by GR gene variants." PLoS One **7**(3): e31720.
Sinclair, D., M. J. Webster, et al. (2012b). "Glucocorticoid receptor mRNA and protein isoform alterations in the orbitofrontal cortex in schizophrenia and bipolar disorder." BMC Psychiatry **12**(1): 84.
Singerman, B. (1981). "DSM-III: historical antecedents and presents significance." J Clin Psychiatry **42**(11): 409-10.
Singh, I. S., J. R. He, et al. (2002). "A high affinity HSF-1 binding site in the 5'-untranslated region of the murine tumor necrosis factor-alpha gene is a transcriptional repressor." J Biol Chem **277**(7): 4981-8.
Sklar, P., S. Ripke, et al. (2011). "Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4." Nat Genet **43**(10): 977-83.
Sklar, P., J. W. Smoller, et al. (2008). "Whole-genome association study of bipolar disorder." Mol Psychiatry **13**(6): 558-69.
Smith, E. N., C. S. Bloss, et al. (2009). "Genome-wide association study of bipolar disorder in European American and African American individuals." Mol Psychiatry **14**(8): 755-63.
Smith, E. N., D. L. Koller, et al. (2011). "Genome-wide association of bipolar disorder suggests an enrichment of replicable associations in regions near genes." PLoS Genet **7**(6): e1002134.
Smoller, J. W. and C. T. Finn (2003). "Family, twin, and adoption studies of bipolar disorder." Am J Med Genet C Semin Med Genet **123C**(1): 48-58.

Spitzer, R. L., Endicott J (1977). The Schedule for Affective Disorders and Schizophrenia, Lifetime Version. New York, New York State Psychiatric Institute.
Spitzer, R. L., J. Endicott, et al. (1975). "Research diagnostic criteria." Psychopharmacol Bull **11**(3): 22-5.
Spitzer, R. L., J. Endicott, et al. (1978). "Research diagnostic criteria: rationale and reliability." Arch Gen Psychiatry **35**(6): 773-82.
Spooren, W. P., F. Gasparini, et al. (2001). "Novel allosteric antagonists shed light on mglu(5) receptors and CNS disorders." Trends Pharmacol Sci **22**(7): 331-7.
St Clair, D., D. Blackwood, et al. (1990). "Association within a family of a balanced autosomal translocation with major mental illness." Lancet **336**(8706): 13-6.
Staden, R. (1996). "The Staden sequence analysis package." Mol Biotechnol **5**(3): 233-41.
Stine, O. C., F. J. McMahon, et al. (1997). "Initial genome screen for bipolar disorder in the NIMH genetics initiative pedigrees: chromosomes 2, 11, 13, 14, and X." Am J Med Genet **74**(3): 263-9.
Straub, R. E. and D. R. Weinberger (2006). "Schizophrenia genes - famine to feast." Biol Psychiatry **60**(2): 81-3.
Suarez BK, H. C., Van Eerdewegh P (1994). Genetic Approaches to Mental Disorders. Washington DC, American Psychiatric Press.
Sullivan, P. F., M. J. Daly, et al. (2012). "Genetic architectures of psychiatric disorders: the emerging picture and its implications." Nat Rev Genet **13**(8): 537-51.
Sullivan, P. F., E. J. de Geus, et al. (2009). "Genome-wide association for major depressive disorder: a possible role for the presynaptic protein piccolo." Mol Psychiatry **14**(4): 359-75.
Sun, Y., L. Zhang, et al. (2001). "Serial analysis of gene expression in the frontal cortex of patients with bipolar disorder." Br J Psychiatry Suppl **41**: s137-41.
Suzuki, G., N. Tsukamoto, et al. (2007). "In vitro pharmacological characterization of novel isoxazolopyridone derivatives as allosteric metabotropic glutamate receptor 7 antagonists." J Pharmacol Exp Ther **323**(1): 147-56.
Swanson, C. J., M. Bures, et al. (2005). "Metabotropic glutamate receptors as novel targets for anxiety and stress disorders." Nat Rev Drug Discov **4**(2): 131-44.
Takaki, H., R. Kikuta, et al. (2004). "Positive associations of polymorphisms in the metabotropic glutamate receptor type 8 gene (GRM8) with schizophrenia." Am J Med Genet B Neuropsychiatr Genet **128B**(1): 6-14.
Tang, B., T. Thornton-Wells, et al. (2011). "Comparative linkage meta-analysis reveals regionally-distinct, disparate genetic architectures: application to bipolar disorder and schizophrenia." PLoS One **6**(4): e19073.
Targum, S. D. (1988). Genetic issues in treatment. Affective disorder and the family: Assessment and treatment. G. L. H. a. I. D. G. J.F.Clarkin. New York, Guilford**:** 196-212.
Tondo, L. and R. J. Baldessarini (1998). "Rapid cycling in women and men with bipolar manic-depressive disorders." Am J Psychiatry **155**(10): 1434-6.
Trabzuni, D., S. Wray, et al. (2012). "MAPT expression and splicing is differentially regulated by brain region: relation to genotype and implication for tauopathies." Hum Mol Genet **21**(18): 4094-103.
Tsuang MT, F. S. (1990). The genetics of mood disorders., Baltimore: Johns Hopkins University Press.
Tsunoka, T., T. Kishi, et al. (2009). "Association analysis of group II metabotropic glutamate receptor genes (GRM2 and GRM3) with mood disorders and fluvoxamine response

in a Japanese population." Prog Neuropsychopharmacol Biol Psychiatry **33**(5): 875-9.
Turecki, G., P. Grof, et al. (2001). "Mapping susceptibility genes for bipolar disorder: a pharmacogenetic approach based on excellent response to lithium." Mol Psychiatry **6**(5): 570-8.
Vadasz, C., M. Saito, et al. (2007). "Glutamate receptor metabotropic 7 is cis-regulated in the mouse brain and modulates alcohol drinking." Genomics **90**(6): 690-702.
van der Schot, A. C., R. Vonk, et al. (2009). "Influence of genes and environment on brain volumes in twin pairs concordant and discordant for bipolar disorder." Arch Gen Psychiatry **66**(2): 142-51.
Vazza, G., C. Bertolin, et al. (2007). "Genome-wide scan supports the existence of a susceptibility locus for schizophrenia and bipolar disorder on chromosome 15q26." Mol Psychiatry **12**(1): 87-93.
Viguera, A. C., R. J. Baldessarini, et al. (2001). "Response to lithium maintenance treatment in bipolar disorders: comparison of women and men." Bipolar Disord **3**(5): 245-52.
Vine, A. E., A. McQuillin, et al. (2009). "No evidence for excess runs of homozygosity in bipolar disorder." Psychiatr Genet **19**(4): 165-70.
Vossen, R. H., E. Aten, et al. (2009). "High-resolution melting analysis (HRMA): more than just sequence variant screening." Hum Mutat **30**(6): 860-6.
Walsh, T., J. M. McClellan, et al. (2008). "Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia." Science **320**(5875): 539-43.
Wang, K. S., X. F. Liu, et al. (2010). "A genome-wide meta-analysis identifies novel loci associated with schizophrenia and bipolar disorder." Schizophr Res **124**(1-3): 192-9.
Weber, J. L. and P. E. May (1989). "Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction." Am J Hum Genet **44**(3): 388-96.
WHO (1967). Manual of the International Statistical Classification of Diseases, Injuries and Causes of Death. Geneva, World Health Organization.
WHO (1993). The tenth revision of the international statistical classification of diseases and related health problems (ICD-10): Diagnostic criteria for research purpose. Geneva, Switzerland, World Health Organization.
Wieronska, J. M., K. Klak, et al. (2007). "Citalopram influences mGlu7, but not mGlu4 receptors' expression in the rat brain hippocampus and cortex." Brain Res **1184**: 88-95.
Wieronska, J. M., B. Legutko, et al. (2008). "Olfactory bulbectomy and amitriptyline treatment influences mGlu receptors expression in the mouse brain hippocampus." Pharmacol Rep **60**(6): 844-55.
Wigg, K., Y. Feng, et al. (2009). "Genome scan in sibling pairs with juvenile-onset mood disorders: Evidence for linkage to 13q and Xq." Am J Med Genet B Neuropsychiatr Genet **150B**(5): 638-46.
Williams, N. M., E. K. Green, et al. (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): 366-73.
Wilson, G. M., S. Flibotte, et al. (2006). "DNA copy-number analysis in bipolar disorder and schizophrenia reveals aberrations in genes involved in glutamate signaling." Hum Mol Genet **15**(5): 743-9.
Wittchen, H. U., F. Jacobi, et al. (2011). "The size and burden of mental disorders and other disorders of the brain in Europe 2010." Eur Neuropsychopharmacol **21**(9): 655-79.
Wittwer, C. T. (2009). "High-resolution DNA melting analysis: advancements and limitations." Hum Mutat **30**(6): 857-9.

WTCCC (2007). "Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls." Nature **447**(7145): 661-78.
Wu, S. C. and R. J. Solaro (2007). "Protein kinase C zeta. A novel regulator of both phosphorylation and de-phosphorylation of cardiac sarcomeric proteins." J Biol Chem **282**(42): 30691-8.
Xu, B., J. L. Roos, et al. (2008). "Strong association of de novo copy number mutations with sporadic schizophrenia." Nat Genet **40**(7): 880-5.
Xu, H., R. Cheng, et al. (2011). "Fine mapping of candidate regions for bipolar disorder provides strong evidence for susceptibility loci on chromosomes 7q." Am J Med Genet B Neuropsychiatr Genet **156**(2): 168-76.
Yosifova, A., T. Mushiroda, et al. (2011). "Genome-wide association study on bipolar disorder in the Bulgarian population." Genes Brain Behav **10**(7): 789-97.
Yutzy, S. H., C. R. Woofter, et al. (2012). "The increasing frequency of mania and bipolar disorder: causes and potential negative impacts." J Nerv Ment Dis **200**(5): 380-7.
Zarate, C., Jr., R. Machado-Vieira, et al. (2010). "Glutamatergic modulators: the future of treating mood disorders?" Harv Rev Psychiatry **18**(5): 293-303.
Zhang, D., L. Cheng, et al. (2009). "Singleton deletions throughout the genome increase risk of bipolar disorder." Mol Psychiatry **14**(4): 376-80.
Zhang, X., J. Zhu, et al. (2007). "Dishevelled promotes axon differentiation by regulating atypical protein kinase C." Nat Cell Biol **9**(7): 743-54.
Zhao, J. H., S. Lissarrague, et al. (2002). "GENECOUNTING: haplotype analysis with missing genotypes." Bioinformatics **18**(12): 1694-5.
Zhou, R., P. Yuan, et al. (2009). "Evidence for selective microRNAs and their effectors as common long-term targets for the actions of mood stabilizers." Neuropsychopharmacology **34**(6): 1395-405.
Zivanovic, O. and A. Nedic (2012). "Kraepelin's concept of manic-depressive insanity: one hundred years later." J Affect Disord **137**(1-3): 15-24.
Zollner, S. and J. K. Pritchard (2007). "Overcoming the winner's curse: estimating penetrance parameters from case-control data." Am J Hum Genet **80**(4): 605-15.

Appendices

APPENDIX I

Buffers and solutions

DNA Extraction from blood samples

10 X Lysis buffer

100 mM NaCl, 100 mM EDTA; 5.84 g of NaCl and 37.22 g EDTA was added to one litre of MilliQ water and autoclaved.

Proteinase K enzyme (20 mg/ml) (Sigma-Aldrich, UK)

Proteinase K buffer

50mM Tris.HCl (pH8.0), 50mM EDTA, 100mM NaCl; 25ml 2M Tris, 100ml 0.5M EDTA, and 25ml 4M NaCl were added and made up to one litre using MilliQ water.

10% SDS

10g of SDS was dissolved in 100 ml of autoclaved water.

Polyvinylpolypyrrolidine (PVPP) (Sigma-Aldrich, UK)

Phenol (Sigma-Aldrich, UK)

1:24 Isoamylalcohol:chloroform solution

One part of Isoamylalcohol (Sigma-Aldrich, UK) was mixed with 24 parts of chloroform (Sigma-Aldrich, UK) and mixed well.

3 M Sodium Acetate

Absolute ethanol (Sigma-Aldrich, UK)

Low EDTA TE

10 mM Tris, 0.1 mM EDTA

Purification of PCR products

MicroCLEAN

0.5 M NaCl, 1 mM Tris HCl pH 8.0, 0.1 mM EDTA, 20 % w/v PEG8000, 1.75 mM $MgCl_2$. All the contents were added to 50 ml of PCR water and the solution was heated gently for the contents to dissolve. The solution was filter sterilized using a 0.45 µM filter.

PAGE for microsatellite genotyping

SequaGel XR (EC-842)(National Diagnostics, UK)

Ammonium persulfate 10 %(w/v)

Ammonium persulfate (APS) acts as a catalyst for the copolymerization of acrylamide and bisacrylamide gels by generating free radicals during an oxido-recuction reaction with TEMED. Fresh 10 % APS solution was prepared by adding 0.1 g APS to 1 ml of deionized water in an eppendorf tube.

10X TBE electrophoresis buffer

Cell culture

Complete growth medium

Dulbecco's modified Eagle's medium (DMEM, Sigma-Aldrich, UK) supplemented with 10% FBS (Sigma-Aldrich, UK) and 1% Penicllin – Streptomycin (Sigma-Aldrich, UK).

Trypsin-EDTA solution (Sigma-Aldrich, UK)

1 X Phosphate buffered saline (PBS) (Sigma-Aldrich, UK)

0.4% Trypan blue solution (Sigma-Aldrich, UK)

Trypan blue is a vital stain that interacts with the cells only when the cell membrane is damaged as the chromophore is negatively charged. Non-viable cells absorb the dye and appear blue and asymetrical whereas viable cells are seen as bright round cells.

Freezing medium

DMEM supplemented with 20% FBS, 10% DMSO and 1% PenStrep

EMSA

Cytoplasmic Lysis buffer

10mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) pH 7.6, 1 mM EDTA (ethylene diamine tetraacetic acid), 0.1 mM EGTA (ethylene glycol tetraacetic acid), 10 nM KCl, 1 mM dithiothreitol (DTT), 1 mM vanadate, 1 tablet of 'complete' protease inhibitor (Roche Diagnostics Ltd, UK).

Nuclear Lysis buffer

20mM HEPES pH 7.6, 0.2mM EDTA, 0.1mM EGTA, 25% glycerol, 0.42M NaCl, 1mM DTT, 1mM vanadate and 1/2 tablet 'complete' protease inhibitor.

Nonidet P-40 (NP-40, Sigma-Aldrich, UK)

1X PBS

The cytoplasmic and the nuclear lysis buffers were filter sterilized using 0.2 µm membrane filter and stored at 4^{o}C.

Polyacrylamide gel for EMSA

40% Acrylamide/Bis-acrylamide solution (Mix Ratio 19:1) (Sigma-Aldrich, UK)

10X TBE

TEMED (Tetramethylethylenediamine)

10 % APS (w/v)

Binding buffer

100 mM Tris, 500 mM NaCl, 10 mM DTT, (pH 7.5), 2.5 mM DTT, 50 ng/µl of Poly (dI.dC) (Sigma-Aldrich, UK) 5 mM $MgCl_2$, and 0.05 % NP-40

Transient transfection

Lipofectamine 2000 (Invitrogen Life technologies, UK)

DMEM supplemented with 10% FBS and with and without Penstrep

Opti-MEM® I Reduced Serum Medium

Glycerol stocks

80% glycerol

In a glass bottle, 8 ml of 100% glycerol (Sigma-Aldrich, UK) was added to 2 ml of MilliQ water and the mixture was autoclaved on a liquid cycle (121^{o}C for 30 minutes).

Appendix II

Primer sequences for screening GRM7 and GRM3

Table A2.1 Primer sequences used for screening GRM7

Primer ID	Primer sequence
GRM7_ExPa_F	CAAAGAAAGGCATTCCCAATCAATG
GRM7_ExPa_R	TGCTGACTGCTCCCAAACCC
GRM7_ExPb_F	GCGTCTTTAAGTCAAGAATCAGGCTC
GRM7_ExPb_R	TGCTGCTCGCTCTCTCCAACA
GRM7_Ex1a_F	CTGGGCTTTCCCGGAGGAG
GRM7_Ex1a_R	GGCTACCATGATGGAGACCGAA
GRM7_Ex1b_F	GGAAGCGATGCTCTACGCCC
GRM7_Ex1b_R	TGTCAAGCCGGATGGAGTGG
GRM7_Ex2_F	TCCCATGCTAACCTCCTTCCC
GRM7_Ex2_R	CAGGTGTACTTCACAGGACTCTGGG
GRM7_Ex3_F	TCCTCGCTTAGCAACAGCATCA
GRM7_Ex3_R	CAAGCACAGACTTCGCCGCT
GRM7_Ex4_F	TTGCTCTTGTTTCTGAAGATGC
GRM7_Ex4_R	TTTGAAACAACCCCACATGA
GRM7_Ex5_F	TTTCGGTCGACACACAAGAGGA
GRM7_Ex5_R	CCTGGCTGGAACTGAACAAGAGG
GRM7_Ex6_F	TCGAAGCAGTGTTTTCTTTAAGC
GRM7_Ex6_R	CCAAACCATAACCCCCAAAT
GRM7_Ex7_F	GGGGACCTATTAAAGGGGATA
GRM7_Ex7_R	CCACAACAAATTGCATGCTC
GRM7_Ex8a_F	TCTTGCAAGCTGAGGGTAAA
GRM7_Ex8a_R	GATGCCCGTCAAAAGAACAT
GRM7_Ex8c_F	GGGATCATTGCCACCATCTTTG
GRM7_Ex8c_R	ATCCAATGGGCTTGGCTTCG
GRM7_Ex8d_F	GCCCAACATCACAACTGGCAA
GRM7_Ex8d_R	TCCAGTGGAATGAACGTGGCTC
GRM7_Ex9_F	GCATGTGCATGGATTGACCC
GRM7_Ex9_R	TGGCATTCACCAAGGCAACC
GRM7_Ex10_F	TGTTTGGGCTGGTTCCCTCA
GRM7_Ex10_R	GGCAGCTGCTGGCACTAAGC
GRM7_Ex11_12_F	CCAAGGTCAGTTAATCTCCAAAA
GRM7_Ex11_12_R	GGGACTCCCAGTCCAGTTT
GRM7_Ex13_F	GAAACTGGCCACCAGGGTCA
GRM7_Ex13_R	GGATGTTGCTGGCATGTGATGA
GRM7_Ex14_F	TGCAGGAAGAGTAGCATTCATGGG
GRM7_Ex14_R	GGCTTTAAGGCATCATTGCTCCA

Primer ID	Primer sequence
grm7_15a_F1	CTTAGAGTCAGAGGAATC
grm7_15a_R1	GCTCTCAGTGGTCCATTC
GRM7_Ex15c_F	TCACTGACATCAGCACTGCCAA
GRM7_Ex15c_R	GGCATGCATTCCAAAGATCAAGA
GRM7_Ex15d_F	CCACTGCACATCATGTTTTT
GRM7_Ex15d_R	TCTTGGACAATATCACAGAATTGAA
GRM7_Ex15e_F	CGTTAATCTTGCTGCTTATGTGCCA
GRM7_Ex15e_R	CCTCTTCCCTTCCCACTTGCTG
GRM7_Ex3a_F	ATGGTCCTGTCTTTGCGTTC
GRM7_Ex3a_R	TTTAGTTTTATTCCACTTATTCCACTG
GRM7_Ex3b_F	CTTGGACGGCCATTAGGACT
GRM7_Ex3b_R	CCACAATAAAGAAACTTGAACAGG
GRM7_Ex3c_F	AAAAATGTTATTTGTATGGCTCAAAA
GRM7_Ex3c_R	CATCTTACTTCCCTGCTTCTTG
GRM7_Ex3d_F	TTTTCACAATCATGCTCTGCT
GRM7_Ex3d_R	TTCTTTAAAATATAAACACCCAACAA
GRM7_Ex3e_F	ACGACCAATTTGGCATTTTTA
GRM7_Ex3e_R	AAGAAACATCTAAAATCAGCAACG
GRM7_Ex3f_F	TTCAAATAATAAATTAGCCTTTTGTTC
GRM7_Ex3f_R	AAGATACAAAGTACCAAAACTGAATGC
GRM7_Ex3g_F	GCAACCTTGTTTAAGGCACT
GRM7_Ex3g_R	CAGGAAAAGATATTACCAAACAAAA
grm7_9a_F1	GCATCAGGAAGCTCTAGAAG
grm7_9a_R1	GATGCCGTTATCATGGTAGTG
GRM7_Ex9b_F	TTCCTCTAGCATGTGAATGTTGA
GRM7_Ex9b_R	AAATCCCTTCTGCCTTAAACAA
GRM7_Ex9c_F	TTGCATTTAGAATGAGGCAGA
GRM7_Ex9c_R	ACACTGAATGCCACAGCAAG
GRM7_Ex9d_F	ATTGTGGGCTCCAGTTTGAC
GRM7_Ex9d_R	ACGCCCAGCCATAAATTCTT
GRM7_Ex9e_F	TTCCTGAAGAGGCAAAAAGC
GRM7_Ex9e_R	TGCAAATTTGACAAGACAACC
GRM7_Ex9f_F	GGCAACAAGAGCAAAACTCC
GRM7_Ex9f_R	CACACAAAGCATTCCGCATA
GRM7_Pd_F	CCAGATGGGGAAACTAAAACC
GRM7_Pd_R	GAAAGAGGTGATGGCAGGAA
GRM7_Pe_F	TTATCCCTGAATGGCTCTGG
GRM7_Pe_R	TGGCAATGTAACGAAGAAACC
GRM7_Pf_F	TCATTCTTTATCTTTCTCTATCTCCA
GRM7_Pf_R	TGCCCCTCTTACTTGTTTCAA
GRM7_Pg_F	CGAGACATGTGCAGGTGAAA
GRM7_Pg_R	GTACCTGCTCCTTCCTGCAA
GRM7_Ph_F	TCCAGAAACTGTCTATAGCGACTAA

Primer ID	Primer sequence
GRM7_Ph_R	AGCACCTCAGGCTGTCTTTG
grm7_nPa_F2	GACAGTGCTATAAGCAAGATTTG
GRM7_isonew_nPa_R1	CCAGATCCCTTTCACTGGTC
grm7_nPb_F1	AGCACAGTGGCTGTCCGG
grm7_nPb_R1	GCTCAGTCGCAATTCAAC
GRM7_isonew_nex4_F	GGTTGGCAAGAAAGATTGGA
GRM7_isonew_nex4_R	AATGCTCTTTTGGCCAGGAT
GRM7_isonew_nex5_F	AAGGATGCTTAAGGACCAACTG
GRM7_isonew_nex5_R	CCACAAAGCTAAGAGAGAAGTGG
GRM7_isonew_nex7_F	AGTTACCTTTCCTGCTCCTCTC
GRM7_isonew_nex7_R	AGTGGAAAGCGATTCCTCAA
GRM7_isonew_nex8_F	GCCTAAACTGTGTCCACTAGGC
GRM7_isonew_nex8_R	CCACTTGCTTTATGGGGAAG

Table A2.2 Primer sequences used for screening GRM3

Primer ID	Primer sequence
GRM3_ExPa_F	TGGCGTAACGTGGGAGAGGA
GRM3_ExPa_R	CCTTTCAACGTATTCTTTCTGGAATGC
GRM3_ExPb_F1	GCATTCCAGAAAGAATACGTTGAAAGG
GRM3_ExPb_R1	GGAACCTTAAATCGGTGACG
GRM3_Ex1a_F	CGTCACCGATTTAAGGTTCC
GRM3_Ex1a_R	CTAGAGCGCTTGGCTTTCAG
GRM3_Ex1b_F	GGCTCGCAGTGTGCAGTTGA
GRM3_Ex1b_R	AGCTCTTCCCAGCCGAGTCC
GRM3_Ex1c_F	ATTAGATGCGACGGCTTCAG
GRM3_Ex1c_R	TCCTCCTCTGGGACCCTTAC
GRM3_Ex2a_F	GAGTGCCTGGTGCTGTGGTG
GRM3_Ex2a_R	CCAATGCATAGGTATCCCTTGAACA
GRM3_Ex2c_F	TCAACGCCTGGAAGCCATGT
GRM3_Ex2c_R	CCCAGTGAAACCCAGCCTCC
GRM3_Ex3a_F	CCCTTTCCTGAAGCACACAC
GRM3_Ex3a_R	GACTTGCGGATGTTGGAG
GRM3_Ex3b_F	GCCATGGCTGAGATCTTGCG
GRM3_Ex3b_R	CGCCTGTGGTTGCGTTTGTT
GRM3_Ex3c_F	GTTCGACCGCTACTTCCAG
GRM3_Ex3c_R	GCCATGTCGTGTACCAGC
GRM3_Ex4a_F	TGGCACCCTTTGATACTCAGTAATGG
GRM3_Ex4a_R	CCTCAGGAAGGTCATAGCATCCA
GRM3_Ex4b_F	TCACTGGGCAGAAACCTTATCGC
GRM3_Ex4b_R	CATCGAAGATGCGGGCAATG
GRM3_Ex4c_F	CCCTTGGTCAAAGCATCGGG
GRM3_Ex4c_R	GGCCAACCAGATGATGCACG
GRM3_Ex4d_F	ACCCTTGCAGAGAAGCGGGA
GRM3_Ex4d_R	CAGGCTCATCATGGCATTTGAA
GRM3_Ex5c_F	GGCAAGCATTCAGAAGCACCC
GRM3_Ex5c_R	CCCAGCACTTAGCACAATACCTGAC
GRM3_Ex6a_F	TCCCTCACCTCCTTCCCACC
GRM3_Ex6a_R	CCCTGTCACCAATGCTCAGCTC
GRM3_Ex6d_F	GAGCTGAGCATTGGTGACAGGG
GRM3_Ex6d_R	CTGGGCAATTCTCGGCTTCC

APPENDIX III

Primers for genotyping microsatellite and SNP markers

Table A3.1 Primer sequences for genotyping novel microsatellite markers at chromosome 1p36 generated for fine mapping of a bipolar disorder susceptibility locus

Marker	Type	Primer ID	Primer sequence
PRKCZ$(GT)_{17}$	Dinucleotide repeat	PRKCZ$(GT)_{17}$F	AAAATACGCCCATTCACTCG
		PRKCZ$(GT)_{17}$R	GGGCAGGTTTCAAAGCA
PRKCZ$(CA)_{13}$	Dinucleotide repeat	PRKCZ$(CA)_{13}$F	AGCAAGAGCACACACACAGC
		PRKCZ$(CA)_{13}$R	TGGCTTTGTCTGAGGAA
PRKCZ$(AC)_{24}$	Dinucleotide repeat	PRKCZ$(AC)_{24}$F	GGAGGCACAACACGGAC
		PRKCZ$(AC)_{24}$R	CCAAGATTTGGGCTTTGTGT

Table A3.2 Primers for genotyping GRM7 SNPs using Kaspar and HRM assay

Primer ID	Primer sequence
rs114774914_F1	GAAGGTGACCAAGTTCATGCTCAGGATGCCAGTACAGGTTGCA
rs114774914_F2	GAAGGTCGGAGTCAACGGATTCAGGATGCCAGTACAGGTTGCT
rs114774914_R1	CAGGAGGAGACGCCGCTCATAA
rs56173829_F1	GAAGGTGACCAAGTTCATGCTAGGTACGACATCAGATGGCAAAGT
rs56173829_F2	GAAGGTCGGAGTCAACGGATTAGGTACGACATCAGATGGCAAAGA
rs56173829_R2	CTAGGTTGCAAGGTTTTGAAATTTTCTGTA
rs17726576_F1	GAAGGTGACCAAGTTCATGCTGTTTGTTTTCGAATGCCTTGTTTTC
rs17726576_F2	GAAGGTCGGAGTCAACGGATTCACTGTTTGTTTTCGAATGCCTTGTTTTT
rs17726576_R1	ATTCCACCGTCTGAGAGAATA
rs342034_F1	GAAGGTGACCAAGTTCATGCTCCCCGGAGGGAGCGCAC
rs342034_F2	GAAGGTCGGAGTCAACGGATTGCCCCGGAGGGAGCGCAT
rs342034_R2	AACATCCTGAGGCTCTTCCAGGTA
3_6901914_F1	GAAGGTGACCAAGTTCATGCTCCATCACCTCTTTCCTGCTCTTT
3_6901914_F2	GAAGGTCGGAGTCAACGGATTCCATCACCTCTTTCCTGCTCTTA
3_6901914_R2	CTGGTGCACCTGGTGGCAGTTT
GRM7_Ex3f_7313045_F1	GAAGGTGACCAAGTTCATGCTAATACCTAAACCAGGAAAAGATATTACCA
GRM7_Ex3f_7313045_F2	GAAGGTCGGAGTCAACGGATTACCTAAACCAGGAAAAGATATTACCG
GRM7_Ex3f_7313045_R2	GCAAATATGTTCCAAGTAATATTGTCTATA
GRM7_rs35106713_Ex2_F1	GAAGGTGACCAAGTTCATGCTGAAGTCATAGCGCCGGTCATCA
GRM7_rs35106713_Ex2_F2	GAAGGTCGGAGTCAACGGATTAAGTCATAGCGCCGGTCATCG
GRM7_rs35106713_Ex2_R2	GCATCAACGGCACCCGAGCTAA
rs140139253_F1	GAAGGTGACCAAGTTCATGCTGAGAACTCTACAAGCCACAGACG
rs140139253_F2	GAAGGTCGGAGTCAACGGATTAAGAGAACTCTACAAGCCACAGACA
rs140139253_R2	CTCTGATCAAGCTGGCTCTCTTCAA
rs138571076_F1	GAAGGTGACCAAGTTCATGCTGTTGTTTAATTTCCAAGTGTTTGGAGAT
rs138571076_F2	GAAGGTCGGAGTCAACGGATTGTTGTTTAATTTCCAAGTGTTTGGAGAC
rs138571076_R4	CAAAGTAGAAATCAATGACAGAAAGATAAA
rs192193072_F1	GAAGGTGACCAAGTTCATGCTATATTTGTAGAAAAGTCCTAATGGCCG
rs192193072_F2	GAAGGTCGGAGTCAACGGATTAAATATTTGTAGAAAAGTCCTAATGGCCA

Primer ID	Primer sequence
rs192193072_R2	CCCCTGGGTCTTGTGCCCCTT
rs721774_F1	GAAGGTGACCAAGTTCATGCTGAATGTGCCAGCAGTGCCCC
rs721774_F2	GAAGGTCGGAGTCAACGGATTGAATGTGCCAGCAGTGCCCT
rs721774_R1	CCCTAAACAGAGTTAAATGAAAAGGTCATT
nPb_7467774_F1	GAAGGTGACCAAGTTCATGCTCTTATTGTTTCTTCACTACACAAATG
nPb_7467774_F2	GAAGGTCGGAGTCAACGGATTCTCTTATTGTTTCTTCACTACACAAATA
nPb_7467774_R2	CCAAATTGTTTCCTACGGCA
rs1965222_F1	GAAGGTGACCAAGTTCATGCTGATGTTAGAAGGATGGGATAGAGACT
rs1965222_F2	GAAGGTCGGAGTCAACGGATTATGTTAGAAGGATGGGATAGAGACC
rs1965222_R1	CTGGAAGGACTTTCCATGCCAACAT
GRM7_9c_7698252_F1	GAAGGTGACCAAGTTCATGCTACTTGTCCTAGGACTTCTGTCAC
GRM7_9c_7698252_F2	GAAGGTCGGAGTCAACGGATTAGACTTGTCCTAGGACTTCTGTCAT
GRM7_9c_7698252_R2	CAACATTGTGGGCTCCAGTTTGACTT
GRM7_15e_7758348_F1	GAAGGTGACCAAGTTCATGCTGGTTTTTAATATGCCTTTCAGAAATGGC
GRM7_15e_7758348_F2	GAAGGTCGGAGTCAACGGATTGGGTTTTTAATATGCCTTTCAGAAATGGT
GRM7_15e_7758348_R2	CAAGTAACATGCTTTAGCTCACGACTA
Pa_6877167_GRM7	Pre-designed assay from Kbiosciences, UK
rs2229902	Pre-designed assay from Kbiosciences, UK
rs140995942_F_HRM	GTTAGGAGTTGGTGACCAG
rs140995942_R_HRM	TGGAGTTTTGCTCTTGTTGC

Table A3.3 Primers for genotyping GRM3 SNPs using KASPar assay

Primer ID	Primer sequence
rs166677_F1	GAAGGTGACCAAGTTCATGCTGGAATGCAATAGCATTGACCACG
rs166677_F2	GAAGGTCGGAGTCAACGGATTCTGGAATGCAATAGCATTGACCACA
rs166677_R2	AAACACAAAGGAAGAGGAATTTTGCTTGTA
rs17161026_F1	GAAGGTGACCAAGTTCATGCTCAACTTTCAAGTAGGAATACTTTCCAC
rs17161026_F2	GAAGGTCGGAGTCAACGGATTCCAACTTTCAAGTAGGAATACTTTCCAT
rs17161026_R2	GGCGATACAACGTGTTCAATTTCCAAAAT
rs2073549_F1	GAAGGTGACCAAGTTCATGCTCGTCTAGGTACCCTGGCTCA
rs2073549_F2	GAAGGTCGGAGTCAACGGATTCGTCTAGGTACCCTGGCTCT
rs2073549_R2	CTCTTATAAGGGTATATCTGCAGAGTCTT
rs148754219_F1	GAAGGTGACCAAGTTCATGCTCATGGTTGGTCCTCCTCATCC
rs148754219_F2	GAAGGTCGGAGTCAACGGATTCTCATGGTTGGTCCTCCTCATCT
rs148754219_R2	GGACAAAGCCAGTAAGCTACCTCTT
rs184681725_F1	GAAGGTGACCAAGTTCATGCTGTGAGAGCCAGGGAGGCC
rs184681725_F2	GAAGGTCGGAGTCAACGGATTGTGTGAGAGCCAGGGAGGCA
rs184681725_R1	CAGTGTGCAGTTGAGTCGCGAGTA
rs139639092_F1	GAAGGTGACCAAGTTCATGCTATTGCATTGTTGATATAATCCGTTGATGAT
rs139639092_F2	GAAGGTCGGAGTCAACGGATTGCATTGTTGATATAATCCGTTGATGAC
rs139639092_R2	GTGTTGGGGGAACAAAAGGAAAATATCAT
rs2228595_F1	GAAGGTGACCAAGTTCATGCTGGCCACCCAGGTGAAGGAG
rs2228595_F2	GAAGGTCGGAGTCAACGGATTCTGGCCACCCAGGTGAAGGAA
rs2228595_R1	GACGACTCGCGGGAGCTCATT

Appendix IV

Primers used for CNV validation and cloning

Table A4.1 Primers for validating CNVs in GRM7 using RNase P CNR assay

Primer ID	Sequence	Roche UPL probe
GRM7_CNV_p19_F1	GCATAGCTGACAGAAGGAAGC	19
GRM7_CNV_p19_R1	TGTCCTGCAAGACGTATGCT	
GRM7_CNV_p19_F2	AGACAGAGTCCCACTCTCTTGG	19
GRM7_CNV_p19_R2	TGTGGTGAGCCAAAATTGC	
GRM7_CNV_p22_F1	AGTCGGCAGCCTTGTTAAAG	22
GRM7_CNV_p22_R1	TGTACCACCTTGGTTTCACACT	
GRM7_CNV_p38_F1	CGCTTATTTATCACCCGTTCC	38
GRM7_CNV_p38_R1	TTGAAGAGGCATCTACCTGGA	

Table A4.2 Primers used for sequence verification of pJet1.2 and pGL3 vector constructs

Primer ID	Primer sequence
pJET1.2_F	CGACTCACTATAGGGAGAGCGGC
pJET1.2_R	AAGAACATCGATTTTCCATGGCAG
GLprimer2	CTTTATGTTTTTGGCGTCTTCCA
RVprimer3	CTAGCAAAATAGGCTGTCCC
RVprimer4	GACGATAGTCATGCCCCGCG

Table A4.3 Primers used for site-directed mutagenesis in the construction of reporter vector constructs

Primer ID	Primer sequence
GRM3_rs148754219A_F_SDM	GCTACCTCTTTTGTGTCAGATGAGGAGGACCAACC
GRM3_ rs148754219A _R_SDM	GGTTGGTCCTCCTCATCTGACACAAAAGAGGTAGC
GRM3_SDM_5BPDEL_F	GTTGTCGGTGCGAGCCATGGAAGACGCC
GRM3_SDM_5BPDEL_R	GGCGTCTTCCATGGCTCGCACCGACAAC
GRM3_SDM_8BPDEL_F	GAGTTGTCGGTGCGAGTGGAAGACGCCAAAAA
GRM3_SDM_8BPDEL_R	TTTTTGGCGTCTTCCACTCGCACCGACAACTC
GRM3_SDM_3BPDEL_F	GAGTTGTCGGTGCGAGTGGAAGACGCCAAAAA
GRM3_SDM_3BPDEL_R	TTTTTGGCGTCTTCCACTCGCACCGACAACTC
PGL3PROM_SDM_a/g_F	CTGTTGGTAAAGCCACCGTGGAAGACGCCAAAAAC
PGL3PROM_SDM_a/g_R	GTTTTTGGCGTCTTCCACGGTGGCTTTACCAACAG

Appendix V

Association analysis of GRM7 and GRM3 in UCL1 bipolar GWAS data

Table A5.1 Association of GRM7 SNPs in the UCL1 sample (SNP data from UCL1 bipolar GWAS (Sklar et al. 2008))

SNP ID	BP	A1	F_A	F_U	A2	CHISQ	P	OR
rs1400286	6861715	A	0.0414	0.048	C	0.5522	0.4574	0.855
rs17046124	6861866	C	0.0129	0.006	T	2.733	0.09827	2.203
rs17288442	6866379	A	0.297	0.3323	C	3.129	0.0769	0.8467
rs17046139	6869166	T	0.06434	0.07059	C	0.3273	0.5672	0.9054
rs342045	6885895	C	0.1232	0.1363	T	0.8029	0.3702	0.8903
rs342040	6891874	C	0.1185	0.133	T	0.9953	0.3184	0.8761
rs3749380	6903297	T	0.3833	0.3961	C	0.3631	0.5468	0.9476
rs340659	6906622	C	0.06814	0.04804	T	3.864	0.04932	1.449
rs340657	6907600	G	0.443	0.4314	A	0.29	0.5902	1.048
rs456835	6911400	T	0.414	0.419	C	0.05218	0.8193	0.9799
rs17046231	6913374	T	0.05882	0.06961	C	1.021	0.3122	0.8354
rs9870018	6927311	T	0.2523	0.2682	C	0.6859	0.4075	0.9209
rs3846161	6939834	G	0.2151	0.2348	A	1.172	0.279	0.8931
rs11915789	6950603	A	0.02091	0.02357	T	0.1642	0.6853	0.885
rs17046322	6952739	A	0.04604	0.03543	G	1.505	0.2199	1.314
rs10490857	6959982	G	0.1411	0.1441	A	0.03801	0.8454	0.976
SNP_A-4194743	6962448	C	0.02033	0.01287	T	1.768	0.1836	1.592
rs2116711	6968207	T	0.09982	0.1048	C	0.1403	0.7079	0.9473
rs17694650	7001308	T	0.1112	0.1125	C	0.00929	0.9232	0.9867
rs4686101	7006880	C	0.4575	0.4555	T	0.008035	0.9286	1.008
rs4686102	7006960	T	0.02354	0.02705	C	0.2581	0.6114	0.867
rs17751439	7010587	G	0.01845	0.01775	A	0.01438	0.9045	1.04
rs781393	7011534	C	0.2519	0.245	T	0.1316	0.7168	1.038
rs1356268	7021095	G	0.04638	0.04314	A	0.1289	0.7195	1.079
rs6801970	7026273	C	0.295	0.2922	G	0.02105	0.8846	1.014
rs576913	7026696	A	0.1388	0.1598	G	1.834	0.1756	0.8473
rs6808554	7028319	A	0.08656	0.101	G	1.291	0.2559	0.8436
rs2069062	7034549	C	0.1934	0.2078	G	0.6876	0.407	0.9137
rs1499161	7045455	C	0.4724	0.4941	A	0.9957	0.3183	0.9166
SNP_A-1937262	7046445	C	0.1971	0.2126	G	0.7795	0.3773	0.9089
rs1532544	7052752	A	0.1679	0.1755	G	0.2032	0.6521	0.9482
rs6781223	7065496	C	0.464	0.4863	T	1.044	0.307	0.9146

SNP ID	BP	A1	F_A	F_U	A2	CHISQ	P	OR
rs11711367	7077524	T	0.4216	0.4436	A	1.018	0.3129	0.9145
rs6771606	7085713	G	0.2072	0.1831	A	1.94	0.1637	1.166
rs17234886	7092021	C	0.2177	0.2366	G	1.059	0.3034	0.898
rs9820417	7104786	G	0.2303	0.2672	A	3.645	0.05622	0.8206
rs17824866	7104822	C	0.2255	0.2475	T	1.392	0.2381	0.8853
rs6803027	7110070	A	0.05515	0.05906	G	0.1492	0.6993	0.93
rs6777970	7110108	A	0.05525	0.06004	G	0.2221	0.6375	0.9155
rs6772333	7116696	G	0.2677	0.2959	A	2.04	0.1532	0.8701
rs9311976	7122638	A	0.2421	0.2292	G	0.4824	0.4873	1.074
rs9812630	7124872	A	0.4006	0.4069	G	0.08701	0.768	0.9741
rs9875041	7125230	C	0.4061	0.4157	A	0.194	0.6596	0.9614
rs9856068	7125486	A	0.05268	0.05588	C	0.105	0.746	0.9395
rs6785425	7125901	C	0.05	0.05413	T	0.1815	0.6701	0.9196
rs13072518	7131501	T	0.4476	0.4588	G	0.2671	0.6053	0.9558
rs9826341	7132442	C	0.414	0.4239	G	0.2091	0.6475	0.9603
SNP_A-2134151	7138976	G	0.05147	0.05392	A	0.0634	0.8012	0.9521
rs1909397	7139094	A	0.4372	0.4518	T	0.4535	0.5007	0.9425
rs9837834	7142227	G	0.3814	0.39	A	0.1622	0.6872	0.9646
rs9837989	7142307	C	0.3801	0.3907	T	0.2441	0.6213	0.9565
rs6764411	7143551	C	0.3824	0.3927	A	0.2378	0.6258	0.9573
rs9876724	7152640	A	0.3665	0.4008	G	2.616	0.1058	0.8649
rs6777701	7153339	C	0.4456	0.4702	T	1.28	0.2578	0.9054
rs6778030	7153652	C	0.461	0.4862	T	1.335	0.248	0.9038
rs11928865	7155702	A	0.2532	0.2701	T	0.7786	0.3776	0.9162
rs6443090	7168009	G	0.1875	0.1873	C	0.000208	0.9885	1.002
rs9814809	7170741	C	0.2546	0.2761	G	1.237	0.266	0.8956
rs1878164	7173002	T	0.4328	0.4559	C	1.137	0.2863	0.9107
rs4441639	7179682	T	0.2569	0.2805	C	1.484	0.2232	0.8869
rs3828472	7188006	G	0.2994	0.3051	T	0.07996	0.7774	0.9735
rs3749448	7188116	T	0.1957	0.2006	C	0.07762	0.7806	0.9699
rs7650218	7191923	A	0.1951	0.1944	G	0.001767	0.9665	1.005
rs9814881	7198602	G	0.1895	0.1905	A	0.003252	0.9545	0.9936
rs12497688	7201316	C	0.1859	0.1887	T	0.02806	0.867	0.9814
rs17235018	7212512	G	0.3208	0.3327	C	0.3386	0.5606	0.9473
rs17288561	7212607	T	0.3686	0.3703	C	0.007056	0.9331	0.9924
rs11131064	7214945	T	0.3787	0.3477	C	2.175	0.1403	1.143
rs1400166	7218175	T	0.3056	0.321	C	0.5695	0.4505	0.9309
rs2875257	7221090	A	0.3088	0.3353	G	1.691	0.1935	0.8858
rs10510353	7221165	C	0.3051	0.3343	T	2.061	0.1511	0.8744
rs11708019	7229619	G	0.3327	0.3755	A	4.214	0.04009	0.8293
rs1963265	7233935	T	0.3404	0.3089	G	2.363	0.1242	1.155
rs1508724	7241745	A	0.3346	0.2745	G	8.947	0.002779	1.329
rs9823996	7244509	G	0.329	0.2863	C	4.515	0.03359	1.223

SNP ID	BP	A1	F_A	F_U	A2	CHISQ	P	OR
rs11710946	7246241	A	0.3603	0.4216	G	8.309	0.003945	0.7728
rs6769814	7251433	G	0.3401	0.2917	A	5.669	0.01726	1.251
rs10510354	7252190	C	0.3631	0.4216	G	7.568	0.005941	0.7821
rs13070476	7259313	C	0.2197	0.254	T	3.376	0.06616	0.8273
rs951557	7269035	T	0.443	0.4451	C	0.009256	0.9234	0.9916
rs2136152	7282402	T	0.1562	0.1745	G	1.274	0.2591	0.876
rs4686119	7282739	T	0.1645	0.1804	C	0.9302	0.3348	0.8947
rs908465	7283715	T	0.1633	0.1812	G	1.177	0.2781	0.8818
rs1605705	7290044	G	0.1562	0.1725	A	1.019	0.3127	0.888
rs4095095	7309658	T	0.3139	0.3139	C	1.83E-06	0.9989	1
rs6443099	7320514	C	0.4006	0.4059	T	0.0621	0.8032	0.9781
rs1876614	7321289	A	0.1596	0.1729	G	0.6701	0.413	0.9085
SNP_A-4250891	7321712	G	0.009242	0.00789	A	0.1125	0.7373	1.173
rs12637466	7321909	T	0.4292	0.4127	C	0.5867	0.4437	1.07
rs7635212	7328063	G	0.3803	0.3931	A	0.3659	0.5453	0.9473
rs6443100	7333382	T	0.3759	0.3947	G	0.7814	0.3767	0.9238
rs7632044	7335947	C	0.1621	0.1722	G	0.3912	0.5317	0.9295
rs2133450	7336452	C	0.4524	0.42	A	2.211	0.137	1.141
rs2291867	7340164	G	0.3696	0.3894	A	0.8638	0.3527	0.9194
SNP_A-1971589	7341024	G	0.4594	0.4222	A	2.938	0.08652	1.163
SNP_A-1971590	7341994	A	0.3759	0.399	G	1.184	0.2765	0.9072
rs17047073	7349612	A	0.381	0.3996	G	0.7635	0.3822	0.9248
rs9849147	7350358	A	0.3704	0.3902	T	0.8753	0.3495	0.9194
rs10510356	7350877	C	0.3824	0.3988	G	0.5994	0.4388	0.9331
rs17697853	7366195	A	0.4375	0.402	G	2.729	0.09854	1.157
rs9990013	7366352	A	0.4384	0.4006	C	3.086	0.07897	1.168
rs7621537	7372535	C	0.3419	0.3745	G	2.435	0.1187	0.8677
rs17047149	7380288	G	0.005515	0.01282	A	3.127	0.07698	0.427
rs10510364	7381271	T	0.1155	0.1258	C	0.5008	0.4792	0.9079
rs6810141	7383890	A	0.4314	0.4266	G	0.04832	0.826	1.02
rs1508720	7384636	A	0.1206	0.122	C	0.009945	0.9206	0.9868
rs1499199	7386444	C	0.4072	0.3951	A	0.3193	0.572	1.052
rs11131069	7386487	G	0.1654	0.1798	C	0.7608	0.3831	0.904
rs9836538	7387400	G	0.3969	0.3842	C	0.3552	0.5512	1.055
rs9837019	7387710	T	0.3971	0.3843	C	0.3592	0.549	1.055
rs17047183	7389204	T	0.09444	0.1069	C	0.8951	0.3441	0.8717
rs989126	7389526	C	0.3137	0.3089	A	0.05339	0.8173	1.022
rs7628504	7390577	G	0.2143	0.1962	A	1.038	0.3082	1.117
rs17047199	7392719	T	0.09375	0.1088	A	1.317	0.2512	0.8472
rs7622749	7397946	T	0.1232	0.1375	C	0.9586	0.3275	0.8809
rs7340751	7414354	G	0.4705	0.4706	T	6.19E-06	0.998	0.9998
rs2324122	7414789	G	0.4162	0.4081	A	0.142	0.7063	1.034
rs10510366	7415452	T	0.1447	0.1591	C	0.8463	0.3576	0.894

SNP ID	BP	A1	F_A	F_U	A2	CHISQ	P	OR
rs1499079	7433728	A	0.05463	0.04635	C	0.7458	0.3878	1.189
rs1499204	7437037	C	0.4553	0.4802	G	1.301	0.2541	0.9047
rs12489041	7437599	G	0.4594	0.4882	A	1.752	0.1857	0.8908
rs712767	7443688	T	0.3854	0.3703	C	0.5061	0.4768	1.066
rs1066658	7445630	G	0.384	0.3696	A	0.4651	0.4953	1.063
rs6775424	7448339	A	0.5138	0.4833	G	1.953	0.1623	1.13
rs712775	7458710	G	0.2486	0.248	A	0.000951	0.9754	1.003
rs712777	7463958	C	0.2518	0.2436	T	0.1915	0.6616	1.045
rs11717750	7464013	T	0.5	0.4617	C	3.089	0.07882	1.166
rs17655560	7466155	G	0.4559	0.4863	C	1.952	0.1624	0.8851
rs712779	7476597	T	0.1575	0.1412	C	1.096	0.2951	1.137
rs712782	7481801	G	0.1593	0.1402	A	1.505	0.2199	1.162
rs17717959	7481854	T	0.5084	0.4743	C	2.42	0.1198	1.146
rs712785	7485714	C	0.1507	0.1343	G	1.16	0.2815	1.144
rs779701	7518772	G	0.2902	0.3078	A	0.7798	0.3772	0.9193
rs779699	7519647	C	0.2745	0.2967	G	1.264	0.2608	0.897
rs752300	7520084	T	0.2677	0.296	C	2.062	0.151	0.8694
rs779694	7521741	G	0.2878	0.3084	C	1.068	0.3014	0.9061
rs1083801	7523498	G	0.2831	0.2588	A	1.568	0.2105	1.131
rs779706	7524042	G	0.2895	0.3098	C	1.033	0.3096	0.9079
rs779705	7524103	G	0.2831	0.3078	A	1.551	0.213	0.8878
rs779733	7527372	A	0.2405	0.2454	C	0.06846	0.7936	0.9733
rs9870680	7529555	A	0.4596	0.4569	G	0.01541	0.9012	1.011
rs12494654	7533393	T	0.4329	0.4382	C	0.06086	0.8051	0.9785
rs10222587	7533618	T	0.1271	0.1257	C	0.008426	0.9269	1.012
rs3804945	7539974	T	0.3505	0.3517	C	0.003312	0.9541	0.9947
rs11131078	7548067	T	0.3217	0.3127	C	0.1945	0.6592	1.042
rs756084	7561149	C	0.454	0.4283	A	1.404	0.236	1.11
rs1106486	7562364	C	0.3336	0.35	T	0.6247	0.4293	0.9299
rs1121606	7564242	G	0.4677	0.4454	A	1.038	0.3083	1.094
rs9819987	7567664	G	0.4678	0.4461	T	1.004	0.3164	1.092
rs3804928	7576514	A	0.203	0.2269	G	1.789	0.1811	0.8675
rs779749	7576541	T	0.3633	0.3865	C	1.185	0.2764	0.9059
rs779742	7583058	G	0.4733	0.4539	A	0.7939	0.3729	1.081
rs17664833	7583314	C	0.09835	0.102	G	0.07639	0.7823	0.9607
rs779741	7583602	C	0.4133	0.4333	A	0.8655	0.3522	0.9211
rs1351938	7598417	A	0.4458	0.4637	G	0.6844	0.4081	0.9301
rs17665113	7598811	C	0.1094	0.1061	T	0.059	0.8081	1.035
rs1143740	7602077	C	0.3722	0.3833	T	0.2755	0.5996	0.9539
rs1531939	7617025	C	0.2865	0.2958	G	0.2169	0.6414	0.9559
rs1485174	7620828	T	0.2831	0.277	C	0.09623	0.7564	1.031
rs1485172	7621158	A	0.4288	0.4283	G	0.000636	0.9799	1.002
rs9826424	7621512	A	0.2887	0.2959	G	0.1282	0.7204	0.9662

SNP ID	BP	A1	F_A	F_U	A2	CHISQ	P	OR
rs3804906	7621654	A	0.2877	0.2922	C	0.05117	0.821	0.9785
rs3804904	7621977	A	0.286	0.2947	C	0.1937	0.6599	0.9586
rs3792460	7622270	C	0.284	0.2879	T	0.03846	0.8445	0.981
rs11716647	7624410	A	0.4256	0.4224	G	0.02143	0.8836	1.013
rs11717471	7624469	G	0.2831	0.2804	A	0.01891	0.8906	1.013
rs4143516	7624941	C	0.2848	0.2824	T	0.01413	0.9054	1.012
rs1872400	7624974	A	0.427	0.4245	G	0.01318	0.9086	1.01
rs3792457	7631222	A	0.08333	0.0789	G	0.1381	0.7102	1.061
rs3804886	7636612	C	0.1939	0.2108	G	0.927	0.3356	0.9008
rs3804883	7639708	G	0.2969	0.2971	A	8.52E-05	0.9926	0.9991
rs11713266	7645217	G	0.1498	0.1451	T	0.09318	0.7602	1.038
rs1485171	7651185	A	0.1596	0.1542	C	0.1144	0.7352	1.041
rs17047734	7662963	A	0.07721	0.06287	G	1.655	0.1983	1.247
rs1485167	7664334	C	0.1765	0.1939	T	1.059	0.3035	0.8909
rs6799329	7666694	T	0.4118	0.399	C	0.3547	0.5515	1.054
rs3792452	7666784	A	0.1753	0.1931	G	1.117	0.2906	0.8879
rs17723289	7666994	G	0.1685	0.1781	C	0.3408	0.5594	0.9349
rs10510370	7667672	T	0.1682	0.1775	A	0.3155	0.5743	0.9373
rs3804867	7667909	T	0.1682	0.1775	C	0.3155	0.5743	0.9373
rs9860274	7673242	T	0.3833	0.3863	G	0.02005	0.8874	0.9874
rs3804859	7678340	C	0.1471	0.1368	A	0.4526	0.5011	1.088
rs3804857	7678617	C	0.1876	0.1947	T	0.1681	0.6818	0.9554
rs1872397	7680410	T	0.1472	0.1396	C	0.2409	0.6236	1.064
rs1872394	7680566	A	0.3318	0.334	G	0.01132	0.9153	0.9902
rs9819314	7681695	T	0.3309	0.332	G	0.003091	0.9557	0.9949
rs11918634	7683434	C	0.2924	0.2853	T	0.1259	0.7227	1.035
rs17673467	7685663	C	0.4426	0.4526	T	0.2103	0.6465	0.9605
rs4686145	7686392	A	0.06342	0.07549	C	1.19	0.2753	0.8293
rs4686146	7688564	G	0.3315	0.3471	T	0.5688	0.4507	0.9329
rs4686148	7688662	A	0.3093	0.3248	C	0.5844	0.4446	0.9307
rs3804850	7688702	T	0.2243	0.2264	C	0.01344	0.9077	0.988
rs3864076	7708008	G	0.5046	0.4652	A	3.24	0.07185	1.171
rs11706732	7710584	A	0.431	0.4095	T	0.9762	0.3231	1.092
rs2324209	7710841	G	0.06066	0.04804	C	1.626	0.2022	1.28
rs9870241	7710922	A	0.05985	0.04536	G	2.199	0.1381	1.34
SNP_A-2274307	7712944	T	0.06019	0.04519	C	2.353	0.1251	1.353
rs3804843	7715397	T	0.05545	0.04331	G	1.641	0.2002	1.297
rs9860395	7718570	C	0.06461	0.0507	T	1.839	0.1751	1.293
rs6782528	7721032	A	0.43	0.4479	T	0.6851	0.4078	0.9298
rs162802	7721613	G	0.5018	0.4931	A	0.1594	0.6897	1.035
rs162785	7727101	T	0.4788	0.4676	C	0.2641	0.6073	1.046
rs329044	7748565	G	0.03676	0.02353	T	3.133	0.07674	1.584
rs9860560	7748736	C	0.4393	0.4422	A	0.01697	0.8964	0.9886

SNP ID	BP	A1	F_A	F_U	A2	CHISQ	P	OR
rs1504047	7748843	G	0.4393	0.4422	T	0.01697	0.8964	0.9886
rs332938	7753230	A	0.4384	0.4402	G	0.006749	0.9345	0.9928
rs1155966	7757788	G	0.04926	0.05894	A	0.9606	0.327	0.8272
rs1857697	7758107	C	0.4476	0.4549	T	0.113	0.7367	0.971
rs1352411	7799831	A	0.07486	0.1023	G	4.697	0.03021	0.7102
rs1027527	7805160	G	0.1527	0.16	A	0.2093	0.6473	0.9462
rs6443124	7813180	T	0.1866	0.1853	A	0.005758	0.9395	1.009
rs6443125	7816412	A	0.1847	0.1853	G	0.001062	0.974	0.9963
rs9841316	7820659	T	0.3408	0.3274	G	0.4196	0.5171	1.062

BP, Position in base pairs of SNP based on NCBI build 37 (hg19); A1, Allele 1;F_A, Frequency of A1 in cases; F_U, Frequency of A1 in controls; A2, Allele 2; CHISQ, Basic allelic test chi-square (1df); P, Asymptotic p-value for this test; OR, Estimated odds ratio (for A1, i.e. A2 is reference).

Table A5.2 Association of GRM3 SNPs in the UCL1 sample (SNP data from UCL1 bipolar GWAS (Sklar et al. 2008))

SNP	BP	A1	F_A	F_U	A2	CHISQ	*P*	OR
rs1405881	85897433	C	0.10	0.11	T	0.13	0.72	0.95
rs1405880	85897732	C	0.10	0.10	G	0.00	0.96	0.99
rs274637	85898289	C	0.39	0.34	T	5.95	**0.01**	1.25
rs274621	85917450	C	0.38	0.33	T	7.63	**0.01**	1.29
rs802437	85930852	A	0.41	0.36	G	5.13	**0.02**	1.23
rs802439	85931065	A	0.29	0.31	G	1.01	0.32	0.91
rs802441	85931754	T	0.29	0.31	C	1.17	0.28	0.90
rs802459	85952689	T	0.40	0.35	C	5.47	**0.02**	1.24
rs802467	85973261	G	0.40	0.42	A	0.83	0.36	0.92
rs2189812	85993650	C	0.32	0.28	G	5.22	**0.02**	1.25
rs10238436	86014795	C	0.22	0.24	G	2.05	0.15	0.86
rs16888204	86014885	G	0.03	0.04	A	1.59	0.21	0.74
rs2237553	86019184	T	0.25	0.29	C	3.60	0.06	0.83
rs12704286	86019592	A	0.25	0.29	G	3.93	**0.05**	0.82
rs2237554	86023547	G	0.25	0.29	T	3.81	0.05	0.82
SNP_A-2126810	86031330	C	0.07	0.08	A	1.49	0.22	0.81
rs2888551	86032662	G	0.03	0.03	A	0.00	1.00	1.00
rs723631	86046243	C	0.24	0.29	G	6.18	**0.01**	0.78
rs17160978	86049959	T	0.04	0.04	C	0.00	0.97	1.01
rs17697445	86050097	A	0.03	0.03	G	0.00	0.97	1.01
rs2158786	86050669	A	0.24	0.29	G	6.29	**0.01**	0.78
rs16888210	86050734	T	0.04	0.04	C	0.00	0.97	1.01
rs2237559	86051267	A	0.04	0.04	G	0.00	0.97	1.01
rs2299221	86063065	T	0.04	0.04	A	0.56	0.45	0.85
rs12704289	86066792	C	0.07	0.08	T	0.98	0.32	0.85
rs7804907	86072859	T	0.24	0.28	C	5.61	**0.02**	0.79
rs2237563	86087819	G	0.26	0.19	A	16.94	**3.85×10^{-5}**	1.56
rs7806785	86090967	G	0.24	0.28	C	5.41	**0.02**	0.79
rs2299224	86092054	G	0.04	0.04	A	0.28	0.60	0.89
rs17697609	86097034	G	0.04	0.04	A	0.42	0.52	0.86
rs17161024	86105481	A	0.12	0.11	G	0.17	0.68	1.06
rs10952891	86111902	T	0.12	0.11	C	0.07	0.79	1.04
rs7788115	86116982	A	0.15	0.17	T	1.34	0.25	0.87
rs6955452	86117235	G	0.06	0.05	C	2.15	0.14	1.33
rs6955917	86117414	A	0.07	0.05	G	2.48	0.12	1.35
rs2299231	86134276	G	0.13	0.10	A	4.40	**0.04**	1.35
rs10236047	86157288	G	0.12	0.10	A	3.54	0.06	1.31
rs10258008	86180097	A	0.12	0.09	C	4.19	**0.04**	1.34

SNP	BP	A1	F_A	F_U	A2	CHISQ	*P*	OR
rs7781178	86180995	G	0.28	0.27	A	0.09	0.76	1.03
rs11974622	86181028	A	0.12	0.10	G	3.95	**0.05**	1.33
rs2519713	86189318	C	0.41	0.47	A	5.93	**0.01**	0.80

CHR, Chromosome; SNP, SNP ID; BP, Position in base pairs of SNP based on NCBI build 36 (hg18); A1, Allele 1;F_A, Frequency of this allele in cases; F_U, Frequency of this allele in controls; A2, Allele 2; CHISQ, Basic allelic test chi-square (1df); P, Asymptotic p-value for this test; OR, Estimated odds ratio (for A1, i.e. A2 is reference).

Appendix VI

Power calculation

A graphical representation of power calculation for different odds ratios and minor allele frequencies for UCL2 sample. Genetic power calculator (http://pngu.mgh.harvard.edu/~purcell/gpc/) was used for generating the table A6.1 and the graph (Figure A6.1).

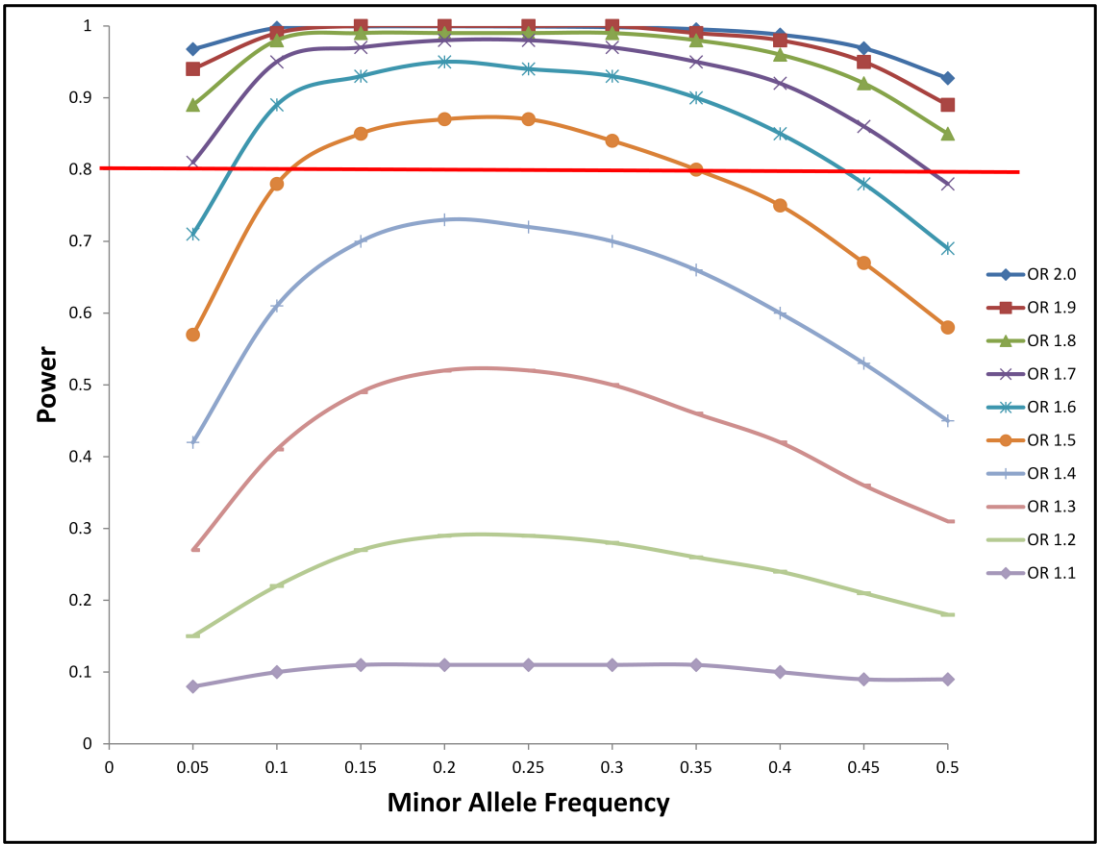


Figure A6.1 Power calculation for UCL2 replication sample

Table A6.1 Power calculation for UCL2 replication sample

Odds ratio	Minor allele frequency 0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5
2	**0.9676**	**0.9973**	**0.9992**	**0.9994**	**0.9991**	**0.9982**	**0.9954**	**0.9878**	**0.9688**	**0.9272**
1.9	**0.94**	**0.99**	**1**	**1**	**1**	**1**	**0.99**	**0.98**	**0.95**	**0.89**
1.8	**0.89**	**0.98**	**0.99**	**0.99**	**0.99**	**0.99**	**0.98**	**0.96**	**0.92**	**0.85**
1.7	**0.81**	**0.95**	**0.97**	**0.98**	**0.98**	**0.97**	**0.95**	**0.92**	**0.86**	0.78
1.6	0.71	**0.89**	**0.93**	**0.95**	**0.94**	**0.93**	**0.9**	**0.85**	0.78	0.69
1.5	0.57	0.78	**0.85**	**0.87**	**0.87**	**0.84**	**0.8**	0.75	0.67	0.58
1.4	0.42	0.61	0.7	0.73	0.72	0.7	0.66	0.6	0.53	0.45
1.3	0.27	0.41	0.49	0.52	0.52	0.5	0.46	0.42	0.36	0.31
1.2	0.15	0.22	0.27	0.29	0.29	0.28	0.26	0.24	0.21	0.18
1.1	0.08	0.1	0.11	0.11	0.11	0.11	0.11	0.1	0.09	0.09

POWER

Appendix VII

miRNA binding site prediction for rs56173829 using RegRNA

Input sequences for GRM7 3' UTR

>Wild type GRM7
GGCTGACCTGTCTTATTACGTATGTACTTCTAGGTTGCAAGGTTTTGAAA
TTTTCTGTACAGTTTGTGAGGACCTTTGC**A**CTTTGCCATCTGATGTCGTAC
CTCGGTTCACTGTTTGTTTTCGAATGCCTTGTTTTCATAGAGCCCTATTCT
CTCAGACGGTGGAATATTTGGAAAAATTTTAAAACAATTAAAATTTTAAA
GCAATCTTGGCAGACTAAAACAAGTACATCTGTACATGACTGTATAATTA
CGATTATAGTACCACTGCACATCATGTTTTTTTTTTTAAGACAAAAAA

>Mutant GRM7
GGCTGACCTGTCTTATTACGTATGTACTTCTAGGTTGCAAGGTTTTGAAA
TTTTCTGTACAGTTTGTGAGGACCTTTGC**T**CTTTGCCATCTGATGTCGTAC
CTCGGTTCACTGTTTGTTTTCGAATGCCTTGTTTTCATAGAGCCCTATTCT
CTCAGACGGTGGAATATTTGGAAAAATTTTAAAACAATTAAAATTTTAAA
GCAATCTTGGCAGACTAAAACAAGTACATCTGTACATGACTGTATAATTA
CGATTATAGTACCACTGCACATCATGTTTTTTTTTTTAAGACAAAAAA

Table A7.1 mi-RNAs predicted to bind to GRM7 3' UTR sequence containing wild-type rs56173829-A allele

miRNA ID	Location	Len	Hybridization	MFE	Score
hsa-miR-106a	51~85	35	miRNA: 3' gauGGACGUG----ACAUU--------CGUGAAAa 5' :|||:|| |||:| ||||||| -Target:5' tttTCTGTACAGTTTGTGAGGACCTTTGCACTTTg 3'	-13.7	152
hsa-miR-106b	53~85	33	miRNA: 3' uaGACGUG----ACAGU--------CGUGAAAu 5' |||:|| ||| | ||||||| Target:5' ttCTGTACAGTTTGTGAGGACCTTTGCACTTTg 3'	-14.84	147
hsa-miR-130a	58~83	26	miRNA: 3' uacggGAAAAUU----GUAACGUGAc 5' :|| |:| | ||||||| Target:5' tacagTTTGTGAGGACCTTTGCACTt 3'	-8.31	147
hsa-miR-130b	62~83	22	miRNA: 3' uacgggaaagUAGUAACGUGAc 5' | | ||||||| Target:5' gtttgtgaggACCTTTGCACTt 3'	-8.31	144
hsa-miR-17	51~85	35	miRNA: 3' gauGGACGUG----ACAUU--------CGUGAAAc 5' :|||:|| |||:| ||||||| Target:5' tttTCTGTACAGTTTGTGAGGACCTTTGCACTTTg 3'	-17.7	152
hsa-miR-182	63~90	28	miRNA: 3' ucACACUC---AAGAUG-GUAACGGUuu 5' |||||| |:|:| | ||||| Target:5' ttTGTGAGGACCTTTGCACTTTGCCAtc 3'	-16.7	146
hsa-miR-19a	59~82	24	miRNA: 3' agUCAAAACGUAUCU--AAACGUGu 5' || ||||:: :|| ||||||| Target:5' acAG-TTTGTGAGGACCTTTGCACt 3'	-14.5	162
hsa-miR-19b	59~82	24	miRNA: 3' agUCAAAACGUACCU--AAACGUGu 5' || ||||:: ||| ||||||| Target:5' acAG-TTTGTGAGGACCTTTGCACt 3'	-16.9	166
	59~82	24	miRNA: 3' agUCAAAACGUACCU--AAACGUGu 5' || ||||:: ||| ||||||| Target:5' acAG-TTTGTGAGGACCTTTGCACt 3'	-16.9	166

MFE, minimum free energy

Table A7.1 mi-RNAs predicted to bind to GRM7 3' UTR sequence containing wild-type rs56173829-A allele contd.

miRNA ID	Location	Len	Hybridization	MFE	Score
hsa-miR-20a	51~85	35	miRNA: 3' gauGGACGUG----AUAUU--------CGUGAAAu 5' :|||:|| |:|:| ||||||| Target:5' tttTCTGTACAGTTTGTGAGGACCTTTGCACTTTg 3'	-14.6	148
hsa-miR-20b	51~85	35	miRNA: 3' gauGGACGUG----AUACU--------CGUGAAAc 5' :|||:|| |:||| ||||||| Target:5' tttTCTGTACAGTTTGTGAGGACCTTTGCACTTTg 3'	-18.1	152
hsa-miR-301a	61~83	23	miRNA: 3' cgAAACUGUUAUGAUAACGUGAc 5' |||| |: || ||||||| Target:5' agTTTGTGAGGACCTTTGCACTt 3'	-10.3	161
hsa-miR-301b	61~83	23	miRNA: 3' cgAAACUGUUAUAGUAACGUGAc 5' |||| |: | | ||||||| Target:5' agTTTGTGAGGACCTTTGCACTt 3'	-9.1	161
hsa-miR-3666	63~83	21	miRNA: 3' agccguagaUGUGAACGUGAc 5' || :||||||| Target:5' tttgtgaggACCTTTGCACTt 3'	-9.72	148
hsa-miR-3678-3p	59~81	23	miRNA: 3' ggcCAGGCA-UGUUUGAGACGUc 5' ||::|| | :| ||:|||| Target:5' acaGTTTGTGAGGACCTTTGCAc 3'	-13.1	143

Table A7.2 mi-RNAs predicted to bind to GRM7 3' UTR sequence containing mutant rs56173829-T allele

miRNA ID	Location	Len	Hybridization	MFE	Score
hsa-miR-1276	66~85	20	miRNA: 3' acagagguguccCGAGAAAu 5' ||||||| Target:5' gtgaggacctttGCTCTTTg 3'	-9.5	140
hsa-miR-182	63~90	28	miRNA: 3' ucACACUC---AAGAUG-GUAACGGUuu 5' |||||| |:|:| | ||||| Target:5' ttTGTGAGGACCTTTGCTCTTTGCCAtc 3'	-16.3	146
hsa-miR-3683	74~100	27	miRNA: 3' acuAUGAUGAA-GGU----UACAGCGu 5' |:|| :|| ||| |||||: Target:5' cttTGCTCTTTGCCATCTGATGTCGTa 3'	-13.8	141
hsa-miR-489	75~99	25	miRNA: 3' cgACGGCAUA---UACACUACAGUg 5' |||: | | || |||||||: Target:5' ttTGCTCTTTGCCATCTGATGTCGt 3'	-13.4	144

Appendix VIII

Publication