STUDY OF TWO BIPOLAR SUSCEPTIBILITY GENES: SLYNAR AND IGF1

Ana Catarina Parente Pereira

University College London

PhD

2009

I, Ana Catarina Parente Pereira 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.

AKNOWLEDGMENTS

ABSTRACT

Linkage studies have implicated the 12q22-24 region in susceptibility to bipolar disorder. In this region alleles at the "Slynar" and Insulin Like Growth Factor 1 (IGF1) genes showed association with bipolar disorder. The Slynar gene is contained within a region of 278 kb on chromosome 12q24 and expresses the sequence AY070435 in the human brain. AY070435 has no known function. A Macaque brain expressed cDNA which is highly homologous to human AY070435 has been cloned and sequenced. To further characterise the human Slynar gene and expressed mRNA transcript studies were carried out to identify Slynar in the mouse and in human neuroblastoma cell lines. Exhaustive efforts were taken to find a mouse homologue but these proved negative. Slynar shared no homology, or partial homology with any other gene in the human genome. The other 12q24 bipolar susceptibility gene IGF1 is highly expressed in the human brain and a well known for its neuromodulatory functions. IGF1 protein has been shown to have an antidepressant and anxiolytic-like effect in the mouse brain. On a genome wide association study (GWAS) with the UCL case control sample, IGF1 was found to be associated to disease with 5 SNPs showing association within the gene. In order to further implicate IGF1 and find the aetiological base pair changes responsible for disease, IGF1 was sequenced. New three new non database SNPs, three previously characterised polymorphisms and a CA repeat were found and genotyped in an extended UCL sample of 1,000 cases and 1,000 controls. One of the novel SNPs and the CA repeat, both located in the promoter region, were associated with bipolar disorder. Haplotype analysis of the GWAS and new markers data confirmed association to bipolar disorder.

TABLE OF CONTENTS

TABLE OF TABLES

TABLE OF FIGURES

1 Introduction

1.1 *Bipolar Disorder*

Bipolar disorder is a lifelong mood disorder characterized by recurrent manic or hypomanic and depressive episodes. It can interfere with cognition and behavior, severely impacting relationships with family, friends, and employers. Social and marital conflict is common, including divorce, sexually transmitted diseases, and unwanted pregnancies, as well as job loss and financial difficulties (Miller 2006).

Age of onset of bipolar disorder ranges from childhood to the late stages of life, but most commonly develops from late 20s to 30s in older studies and from late teens to early 20s in more recent reports. Bipolar disorder has been reported in children as young as 6 years old. Early onset correlates with a greater recurrence of affective disorder in families. Affected relatives have similar ages of onset and symptoms to that of the proband. (Hays *et al.* 1998; O'Mahony *et al.* 2002; Emilien *et al.* 2007).

1.1.1 Historical Background

The origin of the concept of bipolar disorders has its roots in the work and views of the Greek the physicians of the classical period. In fact, the first systematic clinical descriptions of melancholia and mania were by Hippocrates (460-337 BC). However, it was Aretaeus of Cappadocia (150-200 AD) who perceived a relationship between melancholia and mania, recognizing that the two states could occur in the same individual (Angst *et al.* 2001).

Aetiological theories for both melancholia and mania were developed within the theoretical framework of humoralism, established by physicians in the Classical period and remaining the dominant tradition in Europe throughout the Middle Ages. Simplistically stated, excess blood and yellow bile could lead to mania and excess black bile to melancholy. The treatments of blood letting and diet to balance the humors were thus prescribed in accordance with the model, for instance by Robert Burton in his Anatomy of Melancholy in 1621 (Angst *et al.* 2001).

Although various writers continued to note the association of melancholia and mania during the 17th and 18th Centuries, it was not until the middle of 19th Century that a French psychiatrist identified a condition that he termed "folie circulaire". He described a single condition constituting a continuous cycle of melancholy and mania with periods of wellness in between. It was possible that Falret was influenced by the views of German psychiatrist Griesinger, who in 1845 had used the image of a circle to describe the relation of mania and melancholy (Angst *et al.* 2001).

In 1854, the same year that Falret published his theory in full, Jules Baillarger presented his own unitary concept of mania and melancholia, which he referred to as "folie a double forme". Baillarger's views differed from those of Falret, as he recognised that the symptoms alternated, but accorded no importance to the intervening periods of wellness. In contrast, these intervals were central to Falret's concept of a circular disorder (Angst *et al.* 2001).

It was Kraepelin in 1899 that coined the term "manic depressive insanity". This concept unified the affective disorders, including both the circular and depressive types. A broader category was thus created. More importantly, Kraepelin also differentiated manic depressive insanity from dementia praecox. This Kraepelinian dichotomy has had a fundamental impact on the psychiatric community being called the "father of modern psychiatry" (Angst *et al.* 2001).

Kraepelin's unified view of the affective disorders faced considerable opposition, particularly in Germany. Karl Kleist (1953) challenged Kraepelin's concept of manic-depressive disorders, believing that people who suffered from both manic and depressive episodes were suffering from a different illness to those that experienced only one type of affective state. It is to Kleist that we owe current nomenclature of unipolar and bipolar disorder. Kleist's ideas underwent further refinement by another of his colleagues, Karl Leonhard (Angst *et al.* 2001).

Mania is the defining feature of bipolar disorder. Unipolar mania is classified as type of bipolar disorder and unipolar disorder now refers to depression. The work of Angst, Perris and Winkour in the 1960's brought a genetic epidemiological approach to classification. Angst (1959-1963) and Perris (1963-1966) confirmed that unipolar mania was strongly related in genetic terms to "bipolar disorder" and should be included in that category (Angst *et al.* 2001).

In 1881 Mendel used the term hypomania to describe a milder form of mania. This diagnosis category is widely used. (Angst *et al.* 2001).

Accompanying this division of bipolar disorder by severity, investigation has taken place into the relationship between bipolar disorder and schizoaffective disorders. It was Kahlbaum who first suggested that schizoaffective disorder was a separate group (1863). Subsequently Kraepelin was exercised by the problem of defining such cases and he thought that they would fit into one of his categories of manic depressive insanity or dementia praecox (1920). Schizoaffective psychosis was a term first used by Kasanin in 1933. Research since the 1960s has indicated a strong relationship between schizoaffective disorders and bipolar affective disorders. Cadoret (1974) separated schizoaffective disorders into schizoaffective bipolar and schizoaffective unipolar. Based on data that relatives of bipolar probands are at elevated risk of schizoaffective bipolar disorder, Marneros (1999) has argued that schizoaffective bipolar should be included in the bipolar group (Angst *et al.* 2001). In addition, evidence from genetic epidemiology has been gradually accumulating over the past two decades. Family data is coherent with the dichotomous view, that schizophrenia and bipolar affective disorder are different entities (Craddock *et al.* 2005). Family studies have shown some degree of familial coaggregation between schizophrenia and bipolar illness and between schizo-affective disorders and both bipolar disorder and schizophrenia (Craddock *et al.* 2005). That can be explained by the fact that schizoaffective disorders are mainly bipolar or schizophrenia and not a true in between state. Whole-genome linkage studies of schizophrenia (Berrettini 2003; Craddock *et al.* 2005) and bipolar disorder generally implicated different chromosomal regions with very few overlapping shared syndroms. When common susceptibility genes do exist it is likely to be different mutations in the gene that causes different syndroms.

1.1.2 Bipolar Disorder Phenomenology and Classification

1.1.2.1 <u>Phenomenology</u>

Bipolar disorder can cause dramatic swings in mood from overly high and/or irritable to sad and hopeless, and then back again, often with periods of normal mood in between. The intensity of signs and symptoms varies from person to person. The periods of highs and lows are called episodes of mania and depression when they meet criteria for each respectively, and their frequency varies greatly (Emilien *et al.* 2007).

The most common symptoms exhibited during manic episodes include persistently elevated, expansive and/or irritable mood over a identified period of time. The mania is accompanied by feelings of high self-esteem, over optimism, grandiosity, pressure of speech, racing thoughts, distractibility, increased energy with decreased need of sleep, increased sexual drive and overactivity or agitation. There is always loss of inhibitions sometimes leading to reckless involvement in pleasurable, hazardous or embarrassing activities that may have serious marital, social, financial or judicial consequences. Mania symptoms may cause impairment in social and occupational functioning (Kupka *et al.* 2007). The symptoms of a hypomanic episode are similar to those of a mania episode, but they are milder than mania and last for shorter periods. These episodes do not cause a marked functional impairment or require hospitalization (Emilien *et al.* 2007; Kupka *et al.* 2007).

Depression as part of bipolar illness has symptoms very similar to those seen in unipolar depression. Atypical depressive symptoms such as increased sleep,

increased appetite, rejection sensitivity, leaden paralysis and mood reactivity can be observed in a depressive episode in a bipolar patient and can also occur in a rare depressive disorder (El-Mallakh *et al.* 2006; Emilien *et al.* 2007; Kupka *et al.* 2007). In some cases patient recovery from mania can develop a mixed state of depression and mania. They may experience rapidly alternating moods of sadness, irritability, and dysphoria, which can last a week (Kupka *et al.* 2007).

The most common first mood episode in BP appears to be major depression, however bipolar illness can also present first mania, hypomania or mixed symptoms. The disease has a variable course. There may be years between the first and second mood episodes. In addition, symptom fluctuation occurs in most patients with BP. A report from Judd *et al*, in a 14.2 year follow up of BPI patients, showed that patients were symptomatic 50% of the time, with episodes of depression three times more often than mania. Patients with BPI showed week to week changes both in severity and polarity on average six times per year (El-Mallakh *et al.* 2006; Miller 2006).

1.1.2.2   Classification of Bipolar Disorder

For diagnostic purposes the current classifications used are those in DSM-IV (2000) and ICD-10 (1992) (Kupka *et al.* 2007). The Diagnostic and Statistical Manual of Mental Disorders (DSM), published by the American Psychiatric Association, is the standard classification of mental disorders used by mental health professionals in the United States. It describes mood disorders in three parts: (a) mood episodes; (b) mood disorders; (c) characteristics of the most recent episode or the longitudinal course of recurrent illness. The International

Classification of Diseases was developed by the World Health Organization (WHO). The ICD has become the world wide standard diagnostic classification for all diseases.

In the DSM-IV classification bipolar illness can be classified into four subtypes: (a) Bipolar Disorder I; (b) Bipolar Disorder II; (c) cyclothymic disorder; (d) Bipolar Disorder not otherwise specified. To be classified as a Bipolar Disorder I patient, the individual needs to have at least one episode of mania and previous episodes of depression. Most patients with bipolar disorder 1 will experience subsequent manic or depressive episodes. They can also experience hypomanic and mixed episodes, as well as a lower threshold of mood changes between episodes. Patients with bipolar disorder 2 have a history of only hypomania and major depressive episodes. Patients with cyclothymic disorder sometimes only experience manic mild mood changes that do not reach bipolar 1 or bipolar 2 levels. By definition they have symptoms depression and elevated mood for at least 2 years, with no symptom-free period greater than 2 months (Miller 2006).

1.1.3 Diagnosis and Treatment

1.1.3.1 <u>Diagnosis</u>

Diagnostic assessment of bipolar disorder has several aspects: a life time diagnosis according to DSM-IV or/and ICD-10; a diagnosis of the current mood episode or a state of interepisodic remission, rating the severity of the current mood disturbance, and describing the longitudinal course of illness (Kupka *et al.* 2007).

A diagnosis of bipolar disorder is commonly obtained by applying one of the semi-structured interviews by trained clinicians. Many other assessment scales have been developed to use in conjunction with the standard methods of diagnosis (Miller 2006; Kupka *et al.* 2007).

Assessment of bipolar disorder requires a good history of episodic variation over the patients' entire life time. Numerous clinical studies have shown that about 40% of persons who eventually develop bipolar symptoms are initially diagnosed with unipolar depression. There are also studies that show that it takes an average of a decade from the time a patient seeks help from mental health professionals for bipolar disorder and to be confirmed diagnosis. This high percentage of uncertainty can be explained by several factors. Around one half of patients do not know they are experiencing manic episodes. In fact, patients mainly seek help when suffering from depressive episodes. This leads to a general underestimate of the extent of manic symptoms leading to an underdiagnosis of bipolar disorder. One approach to overcome this problem is to access the patient's behavior patterns from family and friends (El-Mallakh *et al.* 2006; Miller 2006).

1.1.3.2   Treatment

Treatment of bipolar disorder requires use of antimanic and antidepressant drugs. About 50% of bipolar disorder patients respond to lithium and lamotrigine. Often patients feel better they may abandon their medication, which may result in a relapse.

Mood stabilization with lithium is a key treatment in all phases of the illness. It has been used in some patients for 30 to 40 years and can be very successful. The second most important drug is lamotrigine and has been shown repeatedly to be better than valproic acid. Carbamezine and other anti-epileptic drugs have also been widely used but with less success (Miller 2006; Emilien *et al.* 2007).

Antipsychotics drugs such as haloperidol and chlorpromazine are also widely used in clinical treatment of mania. Haloperidol and chlorpromazine seem to be good sedatives for manic patients but are not truly antimanic. A recent meta-analysis of atypical antipsychotics - including aripiprazole, olanzapine, quetiapine, resperidone and ziprasidone - concluded that they were all superior to a placebo, but only small differences in efficacy were observed between them. Benzodiazepines are widely used as an adjunct treatment for behavioral disturbance, although they carry the risk of disinhibition and dependence. ECT is sometimes used as an effective treatment for both mania and depression but nowadays is rarely used for mania (Perlis *et al.* 2006; Emilien *et al.* 2007).

On the whole, antidepressants are an effective treatment for BP depressive episodes. There are now three main classes of anti-depressant: tricyclic antidepressants (TCA), monoamine oxidase inhibitors (MAOI), and selective serotonin reuptake inhibitors (SSRI), plus other miscellaneous antidepressants. Efficacy is similar for different antidepressants, although there are notable differences in the side effect profiles. SSRIs are generally used as first line agents for depression because of their higher tolerability and safety in overdose (Malhi *et al.* 2003; El-Mallakh *et al.* 2006).

Psychotherapy can be used in the management of mild depression Psychotherapeutic support may also improve functioning between episodes and reduce distress (Miller 2006).

1.1.3.3 Hospitalization

A severe episode of depression or mania, during which an individual may become a danger to himself/herself or others, requires inpatient hospitalization. Certain patients admit themselves to hospital voluntarily while others need compulsory.

In the first stage of hospitalization, the first priority is to maintain safety and ensure adequate food and liquid intake.

When the patient's condition improves, attempts can be to teach the patient to recognize the symptoms that led to the hospitalization with the hope of preventing future hospitalization. At this point, families should be involved and also learn about the signs and symptoms.

As the time for discharge approaches, the patient begins to spend more time away from the hospital, during the day or even overnight. This time is used to reorganize life out of hospital, to start getting re-involved with family and other activities. It also serves as a time to assess readiness for discharge. Discharge planning should include the patient and family reviewing their perceptions of the onset of the episode, the reasons for hospitalization, and agreeing to a plan for dealing with symptoms should they recur. Liaison with community support services should be made at this time.

The frequency and intensity of follow up care depends on the patient's individual needs. An individual may go for outpatient therapy to a psychiatrist, a community mental health centre or a family doctor. The patient may also attend group sessions to learn how to deal with the disease and to increase self confidence.

1.1.4 Impact of Bipolar on Society

It was believed that no more than 1% of the general population has bipolar disorder. However, new data provided converging evidence for a higher prevalence of up to 5%, when including severe unipolar depression (Akiskal *et al.* 2000).

An epidemiological study of a sample of 112 persons with an ICD-10 diagnosis of bipolar disorder showed that 69.9% of the patients reported a recurrent episodic illness; 25% had a chronic course without clear remissions and 5.4% had a single episode of mania. Moreover, assessed on a lifetime basis, suicidal ideas were common (78.6%) and levels of drug and alcohol abuse/dependence were high (32.1%). The majority (84.8%) had had at least one contact with inpatient, outpatient or emergency services in the previous year (Morgan *et al.* 2005).

The World Health Organization has determined that bipolar disorder is the eighth most important public health problem worldwide (Emilien *et al.* 2007). The Global Burden of Disease Study used Disability Adjusted Life Years (DALYs) to measure disease burden for many conditions (Murray *et al.* 1997). DALYs are

a composite measure of the sum of life years lost to premature death and years lived with disability adjusted for severity. Out of the 107 disorders considered, bipolar disorder ranked in 22nd. According to this study, bipolar disorder accounted for the loss of 1.43×10^7 DALYs in 1990. It has been estimated that the total cost of bipolar disorder in the UK is two billion pounds per year (Das Gupta *et al.* 2002).

In terms of the individual, the disease involves prolonged periods of illness and is likely to recur and worsen. Although virtually all patients recover from acute episodes, only around 20% have prolonged periods of occupational and social stability. This is reflected in a number of measures of reduced social functioning. Bipolar sufferers have higher rates of unemployment, increased use of health services and increased marital dysfunction (Sajatovic 2005). Furthermore, in some bipolar disorder patients, diminished neurocognitive function may exist across all clinical stages, most notably in the neuropsychological domain of executive function, verbal learning, memory and attention (Watson *et al.* 2006).

1.2 Neurobiology of Bipolar Disorder

Numerous theories that attempt to explain bipolar disorder aetiology have evolved in the last few decades. Evidence from family, twin and adoption studies have strongly suggested heritable, genetic factors for bipolar disorder. Biochemical and pharmacological studies led to various hypotheses implicating neurotransmitters, enzymes and neuropeptides, as well as the endocrine and immune systems. Neuroanatomical and neuroimaging techniques have permitted the study of brain neurology and functional effects associated with mood symptoms and disease. However, prior to molecular genetic studies very little was known about the causes of bipolar disorder. Genetic investigation followed by neurobiology has begun to make great advances in aetiological understanding.

1.2.1 Neurotransmitter systems

Many investigators assumed that monoamine transmitters, such as norepinephrine, dopamine and serotonin were implicated in the aetiology of bipolar disorder. More recently, the glutamatergic, cholinergic and GABAergic neurotransmitter systems have been implicated (Newberg *et al.* 2008).

1.2.1.1 <u>Noradrenergic System</u>

The noradrenergic system is involved in both modulating brain function and producing the body's response to emotions. Data has shown that this system is altered in bipolar affective disorders (Bunney *et al.* 1965; Coppen 1969; Coppen 1969). Increased noradrenergic function has been consistently observed in patients with mania. Moreover, the mood stabilizers lithium, valproate and

carbamazepine have been shown to exert effects on the noradrenergic system and/or on its downstream molecular targets, via distinct mechanisms (Nutt 2002; Berton *et al.* 2006).

1.2.1.2   Dopaminergic System

The midbrain dopamine system regulates motor activity, motivation and reward pathways. In addition, the mesolimbic dopamine system plays a critical role in goal-directed behaviour. All of these functions are disrupted in both poles of the disorder (Soares *et al.* 2007; Newberg *et al.* 2008).

Pharmacologic studies show that some medications, such as antipsychotic drugs act on dopamine receptors (Diehl *et al.* 1992; Kapur *et al.* 1992; Gao *et al.* 2005). However, data does not support primary dopaminergic alterations in bipolar patients, even if this system is strongly implicated in the process of hypomania and mania. Rather, the primary abnormality in bipolar disorder seems to be a genetically determined dysregulation of multiple neurotransmitter systems (Nutt 2006; Brugue *et al.* 2007; Dunlop *et al.* 2007; Gershon *et al.* 2007).

1.2.1.3   Serotonergic System

Serotonin has a wide range of effects, including cardiovascular regulation, intestinal motility and modulation of respiration, thermoregulation, circadian rhythm entrainment, sleep-wake cycle, appetite, aggression, mood, sexual behaviour, sensorimotor reactivity, pain sensitivity and learning. Dysfunction of the serotonergic system is thought to be involved in a variety of psychiatric disorders, including mood disorders (Soares *et al.* 2007). Substantial evidence for

the role of serotonin in patients with bipolar disorder comes from alterations in the serotonergic system cause by SSRI medication (Nutt 2002; Berton *et al.* 2006). In addition, data from many different studies, including imaging studies, suggest that abnormalities of the serotonergic system are present in depression and in bipolar disorder (Craddock *et al.* 2001; Higgs *et al.* 2006; Lesch *et al.* 2006; Cannon *et al.* 2007).

1.2.1.4   Cholinergic System

The cholinergic system uses acetylcholine as its neurotransmitter. Acetylcholine (ACh) has functions both in the peripheral nervous system (PNS) and in the central nervous system (CNS). In the central nervous system, ACh acts as a neuromodulator, acting on brain plasticity, CNS excitability and in the reward system (Soares *et al.* 2007).

The cholinergic-aminergic balance hypothesis, supported by pharmacological data, proposes that an increased ratio of cholinergic to adrenergic activity underlies the pathophysiology of depression, whereas the reverse occurs in mania (Davidson 1972; Janowsky *et al.* 1972). In addition, multiple studies have shown that chronic *in vivo* lithium treatment increases ACh synthesis and its release in the rat brain (Jope 1979). In behavioural studies, chronic lithium at clinically relevant doses is reported to enhance a number of cholinergically mediated reponses, including catalepsy and hypothermia (Lerer *et al.* 1985).

1.2.1.5   GABAergic System

γ-aminobutyric acid (GABA) is the major inhibitory nerotransmitter in the CNS, diminuishing the activity of its many target neurons, and thought to be involved in many disorders (Soares *et al.* 2007; Newberg *et al.* 2008).

Not all GABA studies have given consistent results. However, recent postmortem and clinical studies gave a new insight of the potential role of GABA in bipolar disorder (Dean 2004). A study has found that there is an increase in GABA(A) receptors containing alpha5 subunit in the hippocampus from subjects with bipolar I disorder (Dean *et al.* 2005). In addition, abnormalities of GABA synthetic enzymes and GABAergic system related molecules have been found in many studies with bipolar patients (Brambilla *et al.* 2003). There is also an hypothesis that suggests that valproate increases the availability of GABA in synapses. Various studies have shown that GABA does increase in rodent brain after valproate administration. In fact, it seems that valproate increases GABA release and interacts with its transporters. Also, antidepressants, mood stabilizers, electroconvulsive therapy, and GABA agonists have been shown to reverse the depression-like behaviour in animal models and to be effective in unipolar and bipolar patients by increasing brain GABAergic activity (Brambilla *et al.* 2003).

1.2.1.6   Glutamergic System

Glutamate is the major excitatory neurotransmitter in the CNS, and it is known to play a role in regulating the threshold for excitation of most other neurotransmitter systems (Soares *et al.* 2007; Newberg *et al.* 2008).

The strongest direct evidence for a role for glutamate in mood disorders comes from pharmacological and imaging data. Magnetic resonance spectroscopy (MRS) studies in bipolar disorder patients have shown abnormalities of glutamine/glutamate basal levels in brain regions such as prefrontal and anterior cingulate cortices and hippocampus (Yildiz-Yesiloglu *et al.* 2006). In addition, a second study reported that the mean occipital cortex glutamate levels were increased in subjects with depression compared to controls, suggesting that an excess of glutamate in synapses contributes to neuronal atrophy and loss (Sanacora *et al.* 2004). Moreover, drugs used for bipolar disorder treatments such as lithium, lamotrigine and Riluzole, have been shown to act by decreasing glutamate levels, using different pathways (Newberg *et al.* 2008).

AMPA-type glutamate receptors are ionotropic transmembrane receptors for glutamate that mediates fast synaptic transmission in the CNS. AMPA receptors are the most commonly found receptor in the nervous system. They are involved in plasticity and they seem to be a target of mood stabilizers. Chronic lithium and valproate treatment have both been shown to downregulate synaptic expression of an AMPA receptor subunit (GluR1) in the hippocampus, *in vitro* and *in vivo* (Du *et al.* 2004). Interestingly, the antidepressant imipramine has the opposite effect of upregulating the AMPA synaptic strengh in the hippocampus (Du *et al.* 2003). Furthermore, recent studies indicate that AMPA receptor antagonists attenuate several 'manic-like' behaviours in rodents produced by amphetamine administration (Ossowska *et al.* 2004).

The NMDA receptor is an ionotropic glutamate receptor. Activation of NMDA receptors results in the opening of an ion channel that is nonselective to cations. This allows flow of Na^+ and small amounts of Ca^{2+} ions into the cell and K^+ out of the cell. Calcium flux through NMDARs is thought to play a critical role in synaptic plasticity, a cellular mechanism for learning and memory (Newberg *et al.* 2008). Valproate seems to block synaptic responses mediated by NMDA receptors (Basselin *et al.* 2008). Also, ketamine, an NMDA receptors antagonist seems to have a antidepressant action (Engin *et al.* 2009; Phelps *et al.* 2009).

1.2.2 Neuroendocrine Systems

The relationship between endocrine dysfunction and pathological mood states was one of the earliest considerations of biological psychiatry. Of all the endocrine systems hypothesized to be linked to affective disorder, the hypothalamic-pituitary-adrenal (HPA) was one of the most consistent findings. The HPA axis consists of the hypothalamus, pituitary and adrenal glands, various hormones and releasing factors, and it regulates body's acute response to stress (Soares *et al.* 2007). Increased HPA activity has been associated with mixed manic states, depression and less consistently, with classical manic states (Manji *et al.* 2003).

1.2.3 Signalling Networks

Lately, research on pathophysiology and treatment of mood disorders has moved from a focus on neurotransmitters to intracellular cascades. Following neurotransmitter release and binding at the post-synaptic membrane, a second

messenger signalling cascade occurs that ultimately induces a cellular response. This is an extremely complex pathway and dysfunctions on the second messenger mechanisms have also been implicated in the etiology of bipolar disorder.

1.2.3.1 Gαs/cAMP-generating signalling pathway

Gα is a subunit of the major heterotrimeric G protein complex, which is also constituted by the γ and β subunits. These trimeric complexes are loosely associated with G-protein-coupled-receptors (GPCR). Upon receptor stimulation, receptors and G proteins undergo conformational changes that lead to the stimulation of the effector proteins, such as adenylyl cyclase (AC). Cyclic AMP (cAMP) is a second messenger generated by AC, under cell stimulation. When the levels of intracellular cAMP increase, cAMP binds to the cAMP-dependent-protein kinase (PKC) regulatory subunits, leaving the catalytic subunits free to phosphorylate its substrates (Soares *et al.* 2007).

In studies of Gα in peripheral cells and post-mortem brain tissue, Gα predominant species were consistently elevated (Young *et al.* 1991; Young *et al.* 1994). In addition, the cAMP/PKA system has been studied in post-mortem brain of bipolar disorder patients. Data shows that levels of PKA regulatory subunits were significantly lower in cytosolic fractions of frontal, temporal, occipital and parietal cortices, and in the cerebellum and thalamus of bipolar disorder subjects when compared to controls (Chang *et al.* 2003). In addition, lithium, carbamazepine and haloperidol seem to affect the G/cAMP system (Jakobsen *et al.* 1998; Karege *et al.* 1999; Chen *et al.* 2000; Karege *et al.* 2000).

1.2.3.2 Phosphoinositide/Protein Kinase C Signalling Pathway

Protein kinase C (PKC) exists as a family of closely related subspecies, that have a heterogeneous distribution in the brain, and together with other kinases, appear to play a crucial role in the regulation of synaptic plasticity and various forms of learning and memory. PKC is one of the major intracellular mediators of signals generated by neurotransmitter receptors (Manji *et al.* 2003). Allison and Stewart implicated this system in the aetiology of mood disorders when they observed that lithium reduces brain levels of inositol (Allison *et al.* 1971). More recent postmortem and animal data supports a role for PI/PKC signalling system in the pathophysiology of bipolar disorder (Newberg *et al.* 2008). Furthermore, whole-genome association and lithium studies have implicated genes from this pathway in lithium treatment and genetic liability to bipolar disorder (McQuillin *et al.* 2007; Baum *et al.* 2008).

Inositol depletion has been suggested to be a molecular mechanism of the therapeutic effect of lithium (Berridge 1985). Furthermore, some genes of these pathway seem to be downregulated by lithium such as PKC (Bitran *et al.* 1995; Manji *et al.* 1996; Wang *et al.* 2001) and myristoylated, alanin rich C-kinase substrate (MARCKS), a major PKC substrate and a protein involved in regulating long term neuroplastic events (Wang *et al.* 2001). In addition, valproate seems to produce a similar effect to lithium acting by depleting inositol (Shaltiel *et al.* 2004). Nonetheless, these drugs seem to exert they effects through different mechanisms (Newberg *et al.* 2008).

1.2.3.3 GSK-3β Signalling Pathway

Glycogen synthase kinase-3 (GSK-3) plays a critical role in the survival of neurons and it is involved in many pathways implicated in mental illness, such as the dopamine signalling pathway (Beaulieu *et al.* 2005). In addition, new data on DISC1, a previously implicated gene in both schizophrenia and bipolar disorder, shows that DISC1 mediates neural progenitor proliferation by regulating the β-catenin/GSK3β signalling pathway (Mao *et al.* 2009). The results showed that DISC1 stabilises β-catenin by inhibiting GSK3β trough direct action, similarly to lithium action (Figure 1.1) (Harwood 2005; Beaulieu *et al.* 2008; Mao *et al.* 2009).
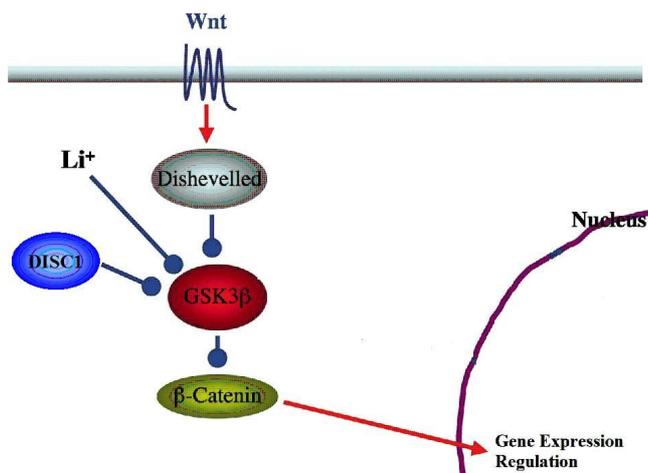
Figure 1.1 GSK3β signalling pathway. The figure shows that GSK3β inhibits β-catenin activity by phosphorylation. When β-catenin is active will act in the expression regulation of other genes and ultimately results in normal progenitor cell proliferation. GSK3β function seems to be regulated by direct interaction with both Lithium and DISC1 (Harwood 2005; Beaulieu *et al.* 2008; Mao *et al.* 2009). Figure adapted from (Coyle 2007).

Recent animal behavioural data have shown that manipulation of the GSK-3 signalling cascade produces both antidepressant and anti-manic effects (Leng *et al.* 2008). It has been shown that inhibition of GSK3β has decreased amphetamine-induced hyperactivity in mice, in a model of mania (Beaulieu *et al.*

2004) and the overexpression of β-catenin in mouse brain mimics the effects of lithium in the amphetamine mania model (Gould *et al.* 2007).

1.2.3.4   <u>Neurotrophic Signalling Cascades</u>

Neurotrophic factors are a family of regulatory factors that mediate the differentiation and survival of neurons, as well as the modulation of synaptic transmission and synaptic plasticity. Increasing evidence suggests that neurotrophic factors inhibit cell death cascades by activating the ERK/MAPK signalling pathway and upregulating the expression of anti-apoptotic proteins such as BCL-2 (Newberg *et al.* 2008). Many of these survival pathways converge at the level of mitochondrial function, and recent findings provide evidence for neurotrophic signalling-mediated mitochondrial dysfunction in bipolar disorder (Kato 2001).

Dysregulation of calcium homeostasis is an essential component of the pathophysiology of classic mitochondriopathies, and impaired regulation of calcium cascades has been found to be the most reproducible biological measure of intracellular alterations described in research on bipolar disorder. Data suggests that the mitochondrial endoplasma reticulum calcium regulation system contributes to the calcium variations seen in bipolar disorder (Kato *et al.* 2003). In addition, a number of microarray studies on postmortem human brain found that mRNAs coding for mitochondrial proteins were decreased in bipolar disorder (Iwamoto *et al.* 2005). All together, there is evidence of a role for

mitochondrial calcium sequestration in the modulation of synaptic plasticity, and as a possible cause of bipolar and related affective disorders.

Pharmacological studies of valproate and lithium showed that the use of therapeutically relevant concentrations of the drugs activate secondary messenger cascades such as the ERK/MAPK survival pathway, which may mediate the antimanic effect of mood disorder (Bielecka *et al.* 2008). Chronic treatment of rats with therapeutic doses of valproate and lithium doubled Bcl-2 levels in the cingulate cortex, dentate gyrus and striatum (Bielecka *et al.* 2008).

1.2.4 Histone Deacetylase

Valproate acts directly on histone deacetylase (HDAC) (Gottlicher *et al.* 2001). HDACs are enzymes that regulate the acetylation/deacetylation of histones, thereby affecting gene expression. Loss of histone acetylases (HATs) and HDACs regulation has been shown to be involved in neuronal dysfunction and degeneration (Selvi *et al.* 2009). Valproate is a HDAC inhibitor, and likely regulates gene expression through its effects on HDAC.

1.2.5 Neuroanatomy Studies

Neuroimaging studies of bipolar disorder have tried to give new insights. Structural imaging suggests abnormalities in prefrontal cortical areas, striatum and amygdala that appear to exist early in the course of illness and, potentially, be present before illness onset (Nugent *et al.* 2006; Terry *et al.* 2009). Other

changes in midline cerebellum, lateral ventricles and other prefrontal regions appear to develop with repeated episodes (Strakowski *et al.* 2005). Spectroscopic studies identify abnormalities of membrane and second messenger metabolism, as well as bioenergetics, in striatum and prefrontal cortex as well, further establishing striatal-thalamic-prefrontal circuits as an integral part of expression of bipolar disorder (Strakowski *et al.* 2005). Finally, functional imaging studies show activation differences between bipolar disorder and controls in these same anterior limbic regions (Robinson *et al.* 2008). Together, neuroimaging studies support a model of bipolar disorder that involves dysfuntions within sub-cortical-prefrontal networks and the associated limbic modulating regions in the amygdale as well as in the midline cerebellum (Strakowski *et al.* 2005).

1.3 Genetics and Bipolar Disorder

1.3.1 Family Studies

One characteristic of a genetic disorder is that it aggregates, or clusters within families. In family studies a proband is the index case ascertained through a systematic catchment procedure and the risk of disease in the relatives of the affected proband subject is investigated. The first and second degree relatives are screened not only for bipolar disorder but also for other psychiatric illnesses. The family morbid risk can therefore be calculated in the families of probands and the families of controls and then compared to the risk in the general population (Ashley-Kosh 2006).

One problem associated with this method is that it is difficult to know if the relative risk observed is due to genetic or family environmental factors. Therefore the observed familial clustering needs to be investigated using adoption studies where genes and environment are better separated (Ashley-Kosh 2006).

Early family studies made no differentiation between bipolar disorder and depression and did not routinely assess psychiatric comorbity (Althoff *et al.* 2005). Therefore little can be inferred about familial risk of bipolar disorder from these studies. Nowadays, many studies from independent groups and different geographical regions have shown a convincing familiality in bipolar disorder. One example is a meta-analysis of eight family studies, that yielded an odds ratio

of 7 for bipolar disorder in first-degree relatives of bipolar disorder probands (Craddock *et al.* 1999).

Investigators have questioned whether bipolar disorder 1 and bipolar disorder 2 represent genetically distinct entities. It was found that the risk of bipolar disorder 1 was similarly elevated in relatives of bipolar disorder 1 probands and bipolar disorder 2 probands. However, the risk of bipolar disorder 2 was much greater in relatives of bipolar disorder 2 probands than bipolar disorder 1 probands (Heun *et al.* 1993). This finding argues for at least some commonality in susceptibility to bipolar disorder 1 and bipolar disorder 2.

There has also been interest in the effect of gender (of proband or relative) and the effect of the nature of relationship (i.e. sibling, parent or offspring) on the risk of bipolar disorder. No clear sex-related effect on risk has been shown and no clear consensus was reached about the nature of the relationship. However, reliable data is only available for risk in first-degree relatives (Craddock *et al.* 1999).

Family studies in bipolar disorder raise a few problems such as the difficult phenotypic definition and possible birth cohort effects and variable age of onset. Other problems relate to the conduct of the studies, e.g. lack of direct interviews (Faraone *et al.* 1990). Nevertheless, studies do exist that indicate familiality in BP.

1.3.2 Segregation Analysis

Segregation analysis is a genetic method that is used to examine the recurrence of an illness within families to determine if the patterns are indicative of an autosomal dominant, X-linked, autossomal recessive, polygenic or mixed genetic and environmental model. Modern methods use likelihood statistics to fit a particular inheritance model to the observed data. By comparing the likelihood of several models, one can determine which model provides the best fit to the data. Segregation analysis does not prove that a particular inheritance model is correct but will determine if the data is consistent with that inheritance model (Ashley-Kosh 2006). All segregation analysis must be carried using strict rules of ascertainment of proband and secondary cases in families.

Segregation analyses on systematically ascertained pedigrees produced results consistent with single gene models, in which a single gene plays a major role in determining disease susceptibility, as well as multigenic multifactorial models. However, these results need to be viewed with caution due to the limited power of the studies in the face of the known extensive degree of genetic heterogeneity found in bipolar disorder (Craddock 2007).

The data from classical genetic epidemiological studies is consistent with models of inheritance that include major gene effects as well as epistatic and polygenic transmission is the favoured transmission in less than 50% of studies (Gurling *et al.* 1991). Nevertheless, the multifactorial polygenic model, where multiple genes interact with each other and with multiple environmental factors to influence susceptibility to disease, is often quoted (Craddock 2007). However

linkage studies using large multiply affective families suggest that major genes have a role on bipolar disorder (Curtis *et al.* 2003).

Other genetic/molecular mechanisms, such as mitochondrial inheritance, dynamic mutation, anticipation and epigenetic effects, have been suggested to induce complex patterns of inheritance in BP (McInnis *et al.* 1993; McMahon *et al.* 1995; Kornberg *et al.* 2000; Craddock 2007). Moreover, many linkage studies have reported cosegregation of X-linked markers with bipolar disorder. An association study of the linked region implicated the G-Protein coupled receptor 50 (GPR50) gene for the X chromosome (Thomson *et al.* 2005).

1.3.3 Twin Studies

Twin and adoption studies are used to differentiate the role of genetics and environment in the aetiological variance of a trait. The premise of twin studies is the comparison of the disease concordance in monozygotic twins (MZ) and dizygotic twins (DZ). Since MZ twins share 100% of their DNA and DZ twins only share 50%, a greater disease concordance in MZ compared with DZ is consistent with the involvement of genetics. The ratio of MZ concordance to DZ concordance is estimated. The ratio is reported as either 'proband-wise' or 'pair-wise'. The proband-wise ratio is defined as the proportion of affected twin partners of probands. In other words, it is the probability that an affected proband's twin will have the disease. The pair-wise method expresses concordance as the proportion of all twin pairs that are concordant (Ashley-Kosh 2006), regardless of whether a twin proband from the same pair was ascertained twice.

An advantage of this method is that generally twins share the same family environment, at least during childhood. However, MZ twins share more similar intrauterine and extrauterine environments than DZ twins. This may lead to a higher concordance among MZ twins. Later in life twins will often not share the same environment, which will increase differences between twins. Misclassification of MZ twins can also occur and affect twins' concordance. However with DNA screening very high certainty can be achieved (Ashley-Kosh 2006).

Twin studies of bipolar disorder have found consistently higher concordance for MZ twins than DZ twins (Rafaelsen 1981). The estimated MZ twins' concordance for narrowly defined bipolar disorder was found to be 50% from pooled data of six available studies (Craddock *et al.* 1999). Recent studies have given very high heritability estimates for bipolar disorder. Thirty MZ and 37 DZ twins from the Maudsley twin register were studied and additive genetic variance was estimated at 85% using a narrow definition of affection (McGuffin *et al.* 2003). Meanwhile, a population based twin study of 26 twin pairs showed a heritability estimate of 93% (Kieseppa *et al.* 2004). Non-shared rather than family environment was implicated in explaining the remaining non genetic variance in these studies.

Discordant MZ twins have been studied in bipolar disorder in order to identify environmental risk or protective factors (Allen 1976). Investigators have generally measured various biological features in the affected and unaffected twin and normal controls. For example, affected twins were found to be

significantly impaired on some measures of visual-spatial functioning and verbal memory compared with the unaffected twin (Gourovitch *et al.* 1999). Furthermore, the right hippocampus was identified as smaller and less asymmetric in the affected co-twin compared to the unaffected co-twin (Noga *et al.* 2001). However, both these studies found significant differences between the unaffected co-twin and normal controls. This raises the spectre that while the BP phenotype is not expressed in the unaffected co-twin the genotype is being expressed sub-clinically - in these cases in terms of brain morphology and neuropsychological function. These 'hidden' effects are referred to as endophenotypes. Discordant MZ twins have also been used to identify risk factors at a molecular level by investigating gene expression differences. These differences have also been used to try and identify aetiological genes (Kakiuchi *et al.* 2003).

1.3.4 Adoption Studies

Adoption studies can also be used to examine the evidence for genetic versus common familial environmental factors. In this approach, cases are ascertained and then the frequency of the disease in biological parents is compared to the frequency in adoptive parents. Higher frequencies in biological parents indicate a strong genetic influence, while higher frequencies in adoptive parents argues in favour of common familial environmental factors (Ashley-Kosh 2006).

Adoption studies of bipolar disorder are rare and only two are reported. Biological parents of the cases are difficult to find. In addition, sometimes children are placed in similar environments to the families of the biological

parents, which may make it less certain that the method can distinguish between genetics and environmental factors. However, when these problems are solved, this approach can provide critical information regarding the etiology of complex traits (Ashley-Kosh 2006).

Two bipolar disorder adoption studies have been published, indicating the strong genetic component that this disease has. A higher risk of BP was found in the biological parents of BP adoptees than in their adoptive parents (Mendlewicz *et al.* 1977). A subsequent study supported this finding, although only ten BP probands were included (Wender *et al.* 1986).

1.4 Mapping Bipolar Disorder

1.4.1 Genetic and Phenotypic Heterogeneity

Locus heterogeneity exists when two or more genes act independently to cause an identical trait. Thus, clinically identical forms of the same disease phenotype can be caused by different genetic etiologies. This is an important consideration in genetic studies because heterogeneity can confound linkage analyses. With more genetic diseases there may also be locus heterogeneity and it is certain that a common disease like bipolar disorder will be heterogenic. Cases of pleiotropism have also been described. Pleiotropism causes phenotypic heterogeneity and means that several disease phenotypes can result from the same mutations (Pericak-Vance 1998).

Heterogeneity causes difficulties in the search for genes involved in the etiology of complex diseases. In linkage studies, locus heterogeneity has been one of the greatest problems in identifying chromosomal regions of importance within families. If homogeneity had been assumed and linkage analysis carried out on a collection of families containing a mixture of two autosomal forms, the results could have been misleading and linkage might have been overlooked. Locus heterogeneity also affects association studies, since the power of association studies decreases as soon as there is more than one susceptibility locus affecting a disease. Generally, to overcome heterogeneity problems, more families are needed to detect a significant result (Pericak-Vance 1998).

1.4.2 Power Analysis

Before undertaken linkage analysis, it is critical to know whether the available pedigree information is sufficient to allow detection of the genes undertaking the trait of interest. Similarly, for allelic and haplotypic association analysis on adequate sample size is essential. Power calculations show that sample sizes of 600 can detect allelic association between a concordant marker allele when only a case allele frequency differs by 4 to 5% in cases compared to controls, assuming complete linkage disequilibrium (LD). Estimates can be obtained prior to collection of samples to develop sampling strategies that will ensure adequate power to detect genetic effects. The interpretation of power studies is based on fundamental statistical concepts such as having complete LD between a marker and a disease allele. The less the LD the greater the sample size is needed (Haines *et al.* 2006).

1.4.3 Genetic Markers

For any disease gene mapping study it is essential to choose informative genetic markers that will be used to test for linkage or association. Markers used for genotyping are restriction fragment length polymorphisms (RFLPs), variable number of tandem repeats (VNTRs), short tandem repeats (STRs) or microsatellite repeats, and single nucleotide polymorphisms (SNPs).

RFLPs were introduced in 1978 and in 1982 became the first modern genotyping markers to be used in a successful linkage study, of Huntington disease (Gusella *et al.* 1983). RFLPs are based on a single base pair change that creates or

eliminates a cleavage site for a restriction enzyme. RFLPs are inherited as simple Mendelian codominant markers, which can be useful in family studies. However they have low heterozygosity (usually < 0.4), which is a severe disadvantage for use in family linkage studies.

The main disadvantage was RFLPs was greatly improved by the identification of VNTRs in 1987 (Nakamura *et al.* 1987). These new markers are made up of specific sets of consensus repetitive sequences that can vary between hundreds of base pairs in length. VNTRs and minisatellites are highly polymorphic and have a high heterozygosity rate.

STRs are widely and evenly distributed in the genome and are relatively easy to score. The number of repeat motifs may vary, however the most useful ones consist of a repeated sequence motif of two (dinucleotide), three (trinucleotide) and four (tetranucleotide) (Litt *et al.* 1989; Weber *et al.* 1989).

SNPs consist of a polymorphism at a single base pair location. In the past SNPs were analysed with restriction enzymes to produce RFLPs. SNPs occur very frequently in the human genome (Wang *et al.* 1998). This high frequency of SNPs made them very useful for association studies, where the power of association is strongest over small distances (Hacia *et al.* 1996). One of their drawbacks is the low heterozygosity, however they can be clustered into haplotypes, which can significantly increase their information content. More than 6 million SNPs have been identified all over the genome due to the effort of the SNP Consortium, the HapMap Project and the Human Genome Sequencing

Project. These can be visualised using internet genome browsers such as NCBI (http://www.ncbi.nlm.nih.gov/), UCSC (genome.ucsc.edu/) and Ensembl (www.ensembl.org/).

1.4.4 Linkage Analysis

Linkage analysis is one of the many methods that can be used to map genes. In fact, the application of molecular biology techniques in combination with statistical linkage analysis has led to rapid localization of genetic loci for many disorders.

In linkage analysis, cosegregation of two or more loci, one of which can be a disease allele, is examined in a family to determine if the two loci demarked by differing alleles segregate independently following Mendelian ratio (alleles observed and compared against random segregation – 50%). Disease alleles will segregate with a marker allele if they are close on the same chromosome.

The measure of genetic distance is the recombination fraction, which is the probability that a recombination will occur in a defined distance along a chromosome in relation to the genetic marker. Recombination occurs when homologous chromosomes cross over. A non-recombinant offspring is one in which the parental type remains intact. Multiple crossing-overs may occur between two loci if they are in opposite ends of large chromosomes. However, recombinants between closely linked marker alleles will rarely suffer crossing over events maintaining that region more conserved. The recombination fraction is referred as ɵ. ɵ ranges from 0 to 0.5:

$\theta = 0$ when two genes are completely linked

$\theta = 0.5$ when two genes segregate independently

1.4.4.1   Parametric Linkage Analysis

Haldane and Smith, in 1947, developed a method for linkage analysis. They proposed the probability ratio test for linkage. This ratio is the probability of the data at some given value of θ divided by the probability of non linkage ($\theta = 0.5$) (Haldane *et al.* 1947). Morton *et al* further developed Haldane and Smith methodology and first defined logarithm of the odds (LOD) score as the $\log_{10}$ of the odds for linkage (Morton 1955).

The parametric LOD score is a likelihood-based statistical measure which computes the likelihood of the observed markers in a family sample showing linkage as opposed to the null hypothesis of non linkage. It can only be inferred from observed phenotypes of family members and for that requires the assumption of certain parameters in a genetic model. That is the reason why these studies are named parametric or model-based linkage analysis. There are four major advantages of model-based linkage analysis: (a) statistically, it is a more powerful approach than any nonparametric method if the genetic model assumed is approximately correct; (b) it uses every family member's phenotypic and genotypic information; (c) It provides an estimate of the recombination fraction between marker and disease locus; (d) It provides a statistical test for linkage and for genetic (locus) heterogeneity.

The genetic model used in parametric linkage analysis includes (a) the mode of inheritance of the trait; (b) the trait and marker allele frequencies; (c) the penetrance values for each possible disease genotype, that is, the probabilities of expressing the disease phenotype given the genotype; (d) sex specificity of recombination fractions (Terwilliger *et al.* 1993).

The model is mostly used assuming a mutation rate of zero and using a sex-averaged recombination fraction. When the genetic model is unknown LOD scores may still be calculated. However, analysis still has to assume that the penetrance and recombination fraction have been estimated correctly. If they have not, this can lead to false negatives. Though, LOD score has been successfully used in many complex traits. It has been used to find single-gene effects in subset of families and also has been applied in conjunction with non-parametric models in more complex linkage signals.

1.4.4.1.1 Two-Point Analysis

The likelihood (*L*) of observing the cosegregation of disease allele and a marker allele is calculated with variable values of θ (that range between 0 and 0.5) and compared to the likelihood when $\theta = 0.5$, which is the null hypothesis (H_0). H_0 assumes no linkage and the alternate hypothesis (H_A) assumes that the disease and the marker locus are linked. θ is the recombination fraction, *R* the number of recombinants, and *NR* the number of non recombinants. The total number of offspring of that family (*N*) will be *R+NR*.

$L = \theta^R (1-\theta)^{NR}$

Log$_{10}$ of the ratio of these likelihoods is then determined for each value of ɵ within the range, and each of the resulting numbers is referred to as the LOD score, *z(x)*, where *x* represents a particular value of ɵ within the range of recombination fractions (ɵ =0 to ɵ =0.5):

z(x) = log$_{10}$ [*L*(pedigree given ɵ =x)/*L*(pedigree given ɵ =0.5)]

The value *z(x)* is referred to as a two point LOD score, since it involves linkage between two loci. It is the value of the log$_{10}$ likelihood at a specific value of ɵ. To demonstrate linkage, there must be evidence of cosegregation that allows rejection of the H$_0$. LOD scores of 3 or more (1000:1 odds in favour of linkage) are indicative of linkage. Values under 3 and above 2 are useful for replication. A LOD of 2 on the X chromosome is considered sufficient to show linkage because of the high prior probability of observing ahead father to sun transmission.

Since diseases rarely accord to perfect Mendelian ratios, LOD score analysis has undergone many revisions in an attempt to improve its ability to model traits whose inheritance shows such deviation. Morton realised that locus heterogeneity could reduce the power of the LOD score method to detect linkage and developed a test for linkage heterogeneity, the K-test (Morton 1956). Other tests of heterogeneity have subsequently been developed, including the Admixture or A test (Smith *et al.* 1963) and the B test (Risch 1988). It can be readily appreciated that the phenomenon of incomplete penetrance could have a great effect on likelihood calculations. Hence, a penetrance function was introduced in an attempt to account for this (Ott 1974).

1.4.4.2 <u>Non-Parametric Models</u>

When the underlying genetic model cannot be specified with any confidence, as happens with majority of complex traits, parametric LOD score analysis looses much of its power, and results can be misleading. The solution to this problem was to develop methods that would rely less on the genetic model specifications. These analytical procedures offer a robust approach when model parameters are less certain, but they all have the weakness of being unable to detect locus heterogeneity. The term non parametric linkage analysis is used to describe methods that do not take into account the recombination fractions and the penetrance.

Methods for non parametric linkage analysis can be distinguished based on four key features: (a) which members of the family are included in the analysis; (b) how phenotypic similarity and estimates of trait locus genetic effect are incorporated; (c) how genotypic similarity at the marker locus is measured and estimated (by identity by state or identity by descent); (d) whether the analysis includes identity by state estimates at single or multiple markers. The greatest variability in study design occurs in deciding which members of the family are included in the analysis and how many markers can be analysed.

Model-free analysis can be performed on affected sib-pairs or more extended pedigrees. Sib-pair analysis, where no recombination fraction is specified, requires the genotypes of two affected siblings with or without their parents. For any given locus the probabilities that the siblings share 0, 1 or 2 alleles identical by descent, on score are 0.25, 0.5 and 0.25 respectively. Evidence that these

probabilities do not hold implies that the marker allele is linked to the disease locus. In extended pedigrees other family members such as cousins and affected individuals across several generations can be used. These additional affective relatives add more useful information to the sample, but do not contribute to within family information since they must be cleared as independent sibling pairs. The Affected Pedigree Member Method (Weeks *et al.* 1988; Weeks *et al.* 1992) and the Nonparametric Linkage Score (NPL) method, as implemented in the program GENEHUNTER, can be used to look for excess allele sharing in extended pedigrees.

1.4.4.3   Two-Locus Linkage

Linkage analysis has traditionally considered loci individually in relation to the disease. However, complex inheritance may arise through gene interaction. Attempts have been made to perform two-locus linkage analysis. The first report of a two-locus linkage analysis was applied to Multiple Sclerosis families and provided evidence for two unlinked loci (Tienari *et al.* 1994). More recently, a systemic two-dimensional genome scan of essential hypertension was performed and was able to identify novel epistatic loci (Bell *et al.* 2006). In BP, two-locus admixture linkage analysis was employed, with evidence that it increased the power to detect linkage when loci on chromosome 21q and 11p were both taken into account and assumed to be acting independently (Smyth *et al.* 1997).

1.4.5 Allelic and Haplotypic Association Studies

Association studies look for a significantly increased or decreased frequency of a marker allele, genotype or haplotypes, with a disease trait than would be expected by chance if there was no association between marker(s) and phenotype.

Allelic association can occur between linked or unlinked loci. The term linkage disequilibrium often is used to refer to allelic association between linked loci. Conceptually, this is the same as standard linkage analysis, however the recombination distances being analysed are very small, generally smaller than 1cM, and the recombination events can only be inferred based on the level of sharing the same allele.

There are several events that may be responsible for allelic association, namely mutation, migration, selection and genetic drift. After an initial event generates an allelic association, the association begins to decay. The rate of decay is related to the recombination rate and is higher in unlinked loci than linked loci. In general, the closest the marker is to the disease, the longer the marker allele-disease association will persist through the generations. Variations within and between populations may occur, especially due to phenomenon such as regional selection pressure or recent population admixture. Population stratification may be a problem that affects case-control association studies, and must be kept in mind when selecting the individuals that will be used in the population sample for the study, since they should be closely matched for ancestry.

1.4.5.1 <u>Measures of Allelic Association</u>

There are many methods to calculate linkage disequilibrium and all measure deviation from independent assortment of alleles at different loci. The three measures mentioned are disequilibrium coefficient, correlation coefficient (r^2) and Lewontin's D'.

The disequilibrium coefficient (D) directly measures the deviation of the haplotype frequency from the product of the allele frequencies. The coefficient is zero when there is no association.

$D = \Pr(AB)-\Pr(A)\Pr(B)$

where P(A) is frequency of allele A and P(B) the frequency of allele B and P(AB) the frequency of AB haplotype. This method can give results that are difficult to interpret since its maximum value depends on the allelic frequencies. So alternative standardized methods where developed, being the most commonly used the correlation coefficient and Lewontin's D'. The correlation coefficient (r^2) is based on D being proportional to the covariance of the allele counts at two loci:

$r^2=D/\{P(A)[1-\Pr(A)]\Pr(B)[1-\Pr(B)]\}$

This is a standard statistical measure with well known properties. When r^2 is zero there is no allelic association.

Lewontin's D' is the disequilibrium coefficient standardized by its maximum value:

D'= D/min{Pr(A)[1-Pr(B)], Pr(B)[1-Pr(A)] } if D>0

D'=D/min {-Pr(A)Pr(B),-[1-Pr(A)][1-Pr(B)]} if D<0

Again, when D' equals to zero there is no association.

Both correlation coefficient and Lewontin's D' values fall in between -1 and 1, being the maximum values independent of allelic frequencies. However the measures themselves are dependent on the allele frequencies and their evolutionary histories.

1.4.5.2 Tests for Association

There are two types of association studies, dependent on the type of sample used. The case-control tests use unrelated individuals who are affected (cases) and unaffected (controls). Case control studies compare allele or genotype frequencies in the cases to the frequency in a set of matched controls. Family-based tests of association use affected individuals and their relatives. Allele frequencies in affected individuals are compared to family-based controls, typically parental controls or unaffected siblings.

1.4.5.2.1 Case-control Test

The case-control test is a standard test for the differences between allele or genotype frequencies in the two samples. For balanced data where the same number of cases and controls is compared, the test used is:

$T_{cc} = [\sum^m (a_i - u_i)]^2 / \sum^m (a_i - u_i)^2$, being $i=1$

Where m is the number of marker alleles and a_i and u_i are the number of times the ith allele is found in the cases and controls, respectively. The statistics can be compared to a chi-square test with $m-1$ degrees of freedom. Following a 2x2 contingency (Table 1.1), the chi-squared calculation for the case-control test, with one degree of freedom, would be:

$T_{cc} = [(ad-bc)^2(a+b+c+d)]/[(a+b)(c+d)(b+d)(a+c)]$

Table 1.1 2x2 Contingency table for case/control analysis

Marker	Cases	Controls	*Total*
Present	A	B	a+b
Absent	C	D	c+d
Total	a+c	b+d	a+b+c+d

Odds ratio (OR) can also be used to measure association between genetic marker and disease in case-control studies. The OR is calculated using the following formula:

$OR = (a/b)/(b/d) = ad/bc$

Confidence intervals (CI) which specify the range of true values for the OR can be derived by estimating the variance and are calculated from the following equation:

$CI = (OR)\exp z[\pm(var(\ln OR)]^{1/2}$

$var(\ln OR) = (1/a)+(1/b)+(1/c)+(1/d)$

where, for a 95% confidence limit or a significance level of 0.05, $z=1.96$.

1.4.5.3 Haplotype Analysis

Haplotype analysis looks for association between disease and a particular combination of alleles at adjacent loci on the same chromosome. The strongest associations are found when a particular allele or haplotype has a low frequency in the population and is increased in cases. Thus, there is potential for a stronger association when examining haplotypes, because many haplotypes will have a low population frequency. In fact, Daly *et al* observed in their association studies in the 5q31 region, that a creation of a comprehensive haplotypic map facilitates mapping approaches based on haplotypes within LD blocks (Daly *et al.* 2001).

However, haplotypic analysis has some difficulties. Haplotypes are not generally observed directly. Instead, they are typically inferred from the individual's genotypes. Some haplotypes can be predicted with some certainty, but when the individual is heterozygous at more than a locus, the haplotype assignment is ambiguous. To overcome this problem the EM algorithm was developed to estimate haplotypes frequencies (Ott 1977). Since then, many association tests have been created to test for haplotypic association in case-control and family studies. The one adopted in the UCL research group was GENECOUNTING (Curtis *et al.* 2006). In addition, haplotype analysis is very sensitive to genotyping error (Moskvina *et al.* 2006). This may lead to a highly inflated significance value for haplotypes in the absence of any single marker associations. Various tests can be preformed to try to identify genotyping error, including testing for Hardy Weinberg Equilibrium (HWE). HWE can also be used to check for aberrant haplotypic association due to sample mismatches.

1.5 Linkage and Association Results in Bipolar Disorder

1.5.1 Linkage

Many linkage studies have been carried out to find susceptibility regions to bipolar disorder (Table 1.2). Two meta-analyses of bipolar disorder genome scans have been conducted. The first study examined seven published genome scans for bipolar disorder, and found the strongest evidence for susceptibility loci on 13q and 22q (Badner *et al.* 2002). A more recent meta-analysis (Segurado *et al.* 2003) of thirteen genome scans of bipolar disorder, including the UCL bipolar linkage study (Curtis *et al.* 2003), was carried out using non parametric rank order statistics to test the randomness of the distribution of positive lods across the genome. The rank order of positive lods showed that several loci, linked to bipolar disorder on chromosomes 1p32.1, 1q31-q32, 11p13 -p13.3 and 21q21.3-qter were non random with p values of less than 0.05. The study also provided some support for regions on chromosome 9p22.3-21.1, 10q11.21-22.1, 14q24.1-32.12 and regions of chromosome 18.

However, many other chromosomal regions have been implicated in bipolar disorder linkage studies (Hayden *et al.* 2006). In the UCL whole genome linkage study where large families containing many cases of bipolar and unipolar disorder were used, linkage was found on chromosomes 1q32, 11p15, 12q24, 21q21 and Xq26-28 (Curtis *et al.* 2003). An NIMH Genetics Initiative showed significant linkage signals at 10q25, 10p12, 16q24, 16p13 and 16p12 (Cheng *et al.* 2006). Linkage at loci on 9q31.1-34, 6q23-24 and 2q33-36 were reported in a

Swedish study on an isolated population (Venken *et al.* 2005). A linkage study in Quebec implicated the 12q24 region (Shink *et al.* 2005).

In addition, several consortia have combined data to strengthen evidence for valid linkages. The Wellcome Trust UK-Irish bipolar affective disorder sibling pair genome scan study gave further support on regions 6q16-21, 4q12-21, 9p21, 10p14-21 and 18q22 (Lambert *et al.* 2005). A study of four European family samples of bipolar disorder provided more evidence on chromosome 4q31 and 6q24 (Schumacher *et al.* 2005). One of the largest studies, where data from 11 bipolar disorder genome wide linkage scan studies involving 5179 individuals from 1067 families was used, established significant linkage on chromosome 6q and 8q (McQueen *et al.* 2005).

Table 1.2 Summary of the main studies done for linkage analysis on bipolar disorder.

Chromosomal region	Study
1p32.1	(Segurado *et al.* 2003)
1q31-32	(Curtis *et al.* 2003; Segurado *et al.* 2003)
2q33-36	(Venken *et al.* 2005)
4p	(Le Hellard *et al.* 2007)
4q	(Lambert *et al.* 2005; Schumacher *et al.* 2005)
4q21	(Cassidy *et al.* 2007)
5q33	(Herzberg *et al.* 2006; Jasinska *et al.* 2009)
6q	(Lambert *et al.* 2005; McQueen *et al.* 2005; Schumacher *et al.* 2005; Venken *et al.* 2005)
7q36	(Cassidy *et al.* 2007)
8q	(McQueen *et al.* 2005)
8p23.1	(Walss-Bass *et al.* 2006)
8q13.3	(Walss-Bass *et al.* 2006)
9p22-21	(Segurado *et al.* 2003; Lambert *et al.* 2005; Cassidy *et al.* 2007)
9q	(Venken *et al.* 2005)
10p	(Lambert *et al.* 2005; Cheng *et al.* 2006)
10p14	(Joo *et al.* 2009)
10q	(Segurado *et al.* 2003, Venken, 2008 #390; Cheng *et al.* 2006)
11p	(Curtis *et al.* 2003; Segurado *et al.* 2003)
12q23-24	(Dawson *et al.* 1995; Ewald *et al.* 1998; Morissette *et al.* 1999; Degn *et al.* 2001; Ewald *et al.* 2002; Abkevich *et al.* 2003; Curtis *et al.* 2003; Shink *et al.* 2005; Cassidy *et al.* 2007)
13q	(Badner *et al.* 2002, Wigg, 2009 #396)
13q13-q14	(Maziade *et al.* 2009)

14q24	(Cassidy *et al.* 2007)
15q26	(Vazza *et al.* 2007; McAuley *et al.* 2009)
16q24	(Cheng *et al.* 2006)
16p	(Ross *et al.* 2008)
16p13-12	(Cheng *et al.* 2006; Cassidy *et al.* 2007, Jones, 2007 #391; Merette *et al.* 2008)
18q	(Segurado *et al.* 2003; Lambert *et al.* 2005)
20	(Ross *et al.* 2008)
21q21	(Curtis *et al.* 2003; Segurado *et al.* 2003)
22q	(Badner *et al.* 2002)
22q11	(Savitz *et al.* 2007)
Xq26-28	(Curtis *et al.* 2003, Wigg, 2009 #396)

1.5.2 Association

Fine mapping of linkage regions has given rise to the detection of possible genes implicated on bipolar disorder aetiology. There are many studies implicating many different genes in bipolar disorder. To have an overall idea of the genes so far implicated, data from many association studies have been compiled on Table 1.3. These data prove locus heterogeneity and point to the need for large sample sizes.

Table 1.3 Compilation of different studies implicating genes in bipolar disorder. This table was created using the Genetic Association Database (http://geneticassociationdb.nih.gov). The p-values mentioned were selected for being the most positive ones.

Gene Symbol	Gene Name	Chromosome	P-values	Publication
ALOX12	Arachidonate 12-lipoxygenase	17 p13.1	0.010	(Fridman *et al.* 2003)
ANK3	Ankyrin 3	10q21		(Ferreira *et al.* 2008; Schulze *et al.* 2009)
ARNTL	Aryl Hydrocarbon Receptor Nuclear Translocator-like	11p15	0.025	(Nievergelt *et al.* 2006)
BCR	Breakpoint cluster region	22q11	0.0054	(Hashimoto *et al.* 2005)
BDNF	Brain-derived neurotrophic fac	11p13	0.000394	(Neves-Pereira *et al.* 2002; Rybakowski *et al.* 2003; Lohoff *et al.* 2005; Green *et al.* 2006; Kremeyer *et al.* 2006; Muller *et al.* 2006; Okada *et al.* 2006; Vincze *et*

				al. 2008; Petryshen *et al.* 2009; Xu *et al.* 2009)
CACNA1C	calcium channel, voltage-dependent, L type, alpha 1C subunit	12p13.3		(Ferreira *et al.* 2008)
CLOCK	Clock homolog (mouse)	4q12	0.026	(Benedetti *et al.* 2003)
COMT	Catechol-O-methyltransferase	22q11.21-q11.23	0.002	(Kirov *et al.* 1998; Mynett-Johnson *et al.* 1998; Papolos *et al.* 1998; Burdick *et al.* 2007; Zhang *et al.* 2009)
DAOA	D-amino acid oxidase activator	13q33.2	0.01	(Williams *et al.* 2006; Prata *et al.* 2008; Bass *et al.* 2009)
DGKH	Diacylglycerol kinase, eta	13q14.11	1.5×10^{-8}	(Baum *et al.* 2008)
DISC1	Disrupted In Schizophrenia 1	1q42.1	0.00026	(Hodgkinson *et al.* 2004; Thomson *et al.* 2005; Schosser *et al.* 2009)
DRD1	Dopamine receptor D1	5q35.1	0.016	(Severino *et al.* 2005; Dmitrzak-Weglarz *et al.* 2006; Del Zompo *et al.* 2007)
DRD2	Dopamine receptor D2	11q23	0.029	(Perez de Castro *et al.* 1995; Massat *et al.* 2002)
DRD4	Dopamine receptor D4	11p15.5	0.001	(Muglia *et al.* 2002; Lopez Leon *et al.* 2005)
DTNBP1	Dystrobrevin binding protein 1	6p22.3	0.014	(Fallin *et al.* 2005; Breen *et al.* 2006; Joo *et al.* 2007; Gaysina *et al.* 2008)
GABRA1	Gamma-aminobutyric acid (GABA) 1	5q34-q35	0.0008	(Horiuchi *et al.* 2004)
GABRA3	Gamma-aminobutyric acid (GABA) 3	Xq28	0.0004	(Massat *et al.* 2002; Craddock *et al.* 2008)
GABRA4	Gamma-aminobutyric acid (GABA) 4	4p14-q12	-	(Craddock *et al.* 2008)
GABRA5	Gamma-aminobutyric acid (GABA) 5	15q11.2-q12	0.024	(Papadimitriou *et al.* 1998; Otani *et al.* 2005; Craddock *et al.* 2008)
GSK3B	Glycogen synthase kinase 3 bet	3q13.3	0.0047	(Benedetti *et al.* 2004; Szczepankiewicz *et al.* 2006)
GRP50	G receptor protein 50	Xq27	0.007	(Thomson *et al.* 2005)
HTR2A	5-hydroxytryptamine (serotonin) 2A receptor	13q14-q21	0.007	(Chee *et al.* 2001; Bonnier *et al.*

				2002; Ranade *et al.* 2003)
HTR3A	5-hydroxytryptamine (serotonin) 3A receptor	11q23.1	0.00016	(Niesler *et al.* 2001)
HTR4	5-hydroxytryptamine (serotonin) 4 receptor	5q31-q33	0.002	(Ohtsuki *et al.* 2002)
MAOA	Monoamine oxidase A	Xp11.4-p11.3	0.037	(Lim *et al.* 1995; Lin *et al.* 2000; Preisig *et al.* 2000; Gutierrez *et al.* 2004; Muller *et al.* 2007)
MLC1	Megalencephalic leukoencephalopathy with subcortical cysts 1	22q13.33	0.002	(Verma *et al.* 2005; Selch *et al.* 2007)
MTHFD1	5,10-methylenetetrahydrofolate dehydrogenase	14q24	0.0054	(Kempisty *et al.* 2007)
MTR	Methionine synthase	1q43	0.0172	(Kempisty *et al.* 2007)
NAPG	N-ethylmaleimide-sensitive fusion (NSF)-attachment protein	18p11	0.027	(Weller *et al.* 2006; Li *et al.* 2009)
NCAM1	Neural cell adhesion molecule 1	11q23.1	0.004	(Arai *et al.* 2004)
NDUFV2	Nicotinamide adenine dinucleotide (NADH)ubiquinone oxidoreductase	18p11	0.009	(Washizuka *et al.* 2004; Zhang *et al.* 2009)
NOS3	Nitric oxide synthase 3	7q36	0.021	(Reif *et al.* 2006)
OTX2	Orthodenticle homeobox 2	14q21-q22	0.003	(Sabunciyan *et al.* 2007)
PER3	Period homolog 3 (Drosophila)	1p36.23	0.008	(Nievergelt *et al.* 2006)
S100B	Glial cell-derived neurotrophic factor	21q22	0.009	(Roche *et al.* 2007)
SLC6A4	Serotonin transporter	17q11.1-q12	0.001	(Mynett-Johnson *et al.* 2000; Mellerup *et al.* 2001; Bellivier *et al.* 2002; Serretti *et al.* 2004; Lasky-Su *et al.* 2005; Masoliver *et al.* 2006)
SYNGR1	Synaptogyrin 1	22q13.1	0.00007	(Verma *et al.* 2005)
TAAR6	Trace amine associated receptor 6	6q23.2	0.00002	(Pae *et al.* 2008)
TPH2	Tryptophan hydroxylase 2	12q21.1	0.004	(Van Den Bogaert *et al.* 2006; Harvey *et al.* 2007; Lin *et al.* 2007; Lopez *et al.* 2007; Grigoroiu-Serbanescu *et al.* 2008)
TRPM2	Transient receptor potential protein	21q22.3	0.008	(McQuillin *et al.* 2006; Xu *et al.*

	melastatin type 2			2006; Xu *et al.* 2009)
VMAT1	Vesicular monoamine transporter 1	8p21.3	0.003	(Lohoff *et al.* 2006)
WFS1	Wolframin	4p16	0.03	(Koido *et al.* 2005)
XBP1	X-box binding protein 1	22q12.1	0.01	(Masui *et al.* 2006)

Table 1.3 shows that there are some genes that have been consistently implicated in bipolar disorder, for example BDNF (Neves-Pereira *et al.* 2002; Rybakowski *et al.* 2003; Lohoff *et al.* 2005; Muller *et al.* 2006; Okada *et al.* 2006) and COMT (Kirov *et al.* 1998; Mynett-Johnson *et al.* 1998; Papolos *et al.* 1998). Members of the GABA receptors family have also shown to be associated to bipolar disorder (GABRA1, P= 0.0008; GABRA5, P=0.024) (Papadimitriou *et al.* 1998; Horiuchi *et al.* 2004; Otani *et al.* 2005).

Interestingly, genes encoding proteins belonging to pathways of the monoamine transmitters, dopamine and serotonin have not shown consistently positive results. However, the serotonin transporter (Mynett-Johnson *et al.* 2000; Mellerup *et al.* 2001; Bellivier *et al.* 2002; Serretti *et al.* 2004; Lasky-Su *et al.* 2005; Masoliver *et al.* 2006), Tryptophan hydroxylase 2 (Van Den Bogaert *et al.* 2006; Harvey *et al.* 2007; Lopez *et al.* 2007) and the serotonin receptors 2A (Chee *et al.* 2001; Bonnier *et al.* 2002; Ranade *et al.* 2003), 3A (Niesler *et al.* 2001) and 4 (Ohtsuki *et al.* 2002), all belonging to the serotonin pathway have been implicated on bipolar disorder. In fact, recently has been shown that missense mutations in Tryptophan hydroxylase 2 have functional properties (McKinney *et al.* 2009). In addition, the Dopamine receptor D1 (Severino *et al.* 2005; Dmitrzak-Weglarz *et al.* 2006), D2 (Perez de Castro *et al.* 1995) and D4 (Lopez Leon *et al.* 2005) might also be associated with bipolar disorder.

It is also of notice that the DISC1, the gene famous for its implication in schizophrenia it has been associated to bipolar disorder (Hodgkinson *et al.* 2004; Thomson *et al.* 2005; Hennah *et al.* 2008). It is possible that DISC1 mutations cause both schizophrenia and bipolar disorder but no actual mutations have been found.

1.5.2.1   Association studies on Chromosome 12

Linkage disequilibrium mapping was carried out on the linkage region of 12q23-24 in a UK Caucasian case-control sample of 347 cases and 374 controls (Glaser *et al.* 2005). Two SNPs (rs3847953 and rs933399) and an insertion/deletion with putative functional relevance (which are in high LD with each other and with the microsatellite marker) showed significant or nearly significant association with bipolar disorder after Bonferroni-correction (reaching nominal p values from $p = 0.002$ to $p = 0.005$). In addition, a significant signal ($p = 0.0016$) was identified for one microsatellite marker. Haplotypic analysis gave evidence of the involvement of CUX2 and the hypothetical protein FLJ32356.

Association studies on the chromosomal region 12q24.31 in a population from Saguenay-Lac-St-Jean (Quebec), revealed significant allelic associations between the bipolar phenotype and markers NBG6 (P = 0.008) and NBG12 (P < 10^{-3}). The authors looked for candidate genes in the NBG12 area and found KIAA1595 and FLJ22471, both hypothetical proteins (Shink *et al.* 2005). Further fine mapping on chromosome 12 highlighted the purinergic receptor P2X, ligand-gated ion channel, 7 (P2RX7) as a candidate gene. However P2RX7 has been delivering conflicting data. A French Canadian population sample was

studied and the presence of two susceptibility loci, the P2RX7 and CaMKK2 genes was detected (Barden *et al.* 2006). The strongest association with bipolar disorder was with the non-synonymous SNP P2RX7-E13A (rs2230912) (p=0.000708). However a study done in a large UK case-control sample, neither rs2230912 nor any of 8 other SNPs genotyped across P2RX7 were found to be associated with bipolar disorder (Green *et al.* 2009).

In a UCL and Danish collaboration, a region of 298 kb on chromosomal region 12q24 was found to be associated with BP. Fine mapping of the region implicated an unknown gene, Slynar (Kalsi *et al.* 2006). Slynar gene is discussed in more detail in section 1.6.1.

In the 12q21 region, Tryptophan hydroxylase (TPH2) has be found to be associated to bipolar disorder, with the strongest *p*Value being 0.004 (Table 1.3). TPH2 is the rate-limiting enzyme in the serotonin (5-HT) biosynthetic pathway responsible for the regulation of serotonin levels (Van Den Bogaert *et al.* 2006; Harvey *et al.* 2007; Lopez *et al.* 2007).

1.5.3 Genome Wide Association Studies (GWAS)

The principle of GWAS is similar to locus specific tests of association. However instead of the starting point being a specific gene or chromosomal region, the variants that may confer risk to disease are analysed in the whole genome (Hennah *et al.* 2008; Porteous 2008). Advances which made GWAS possible where: the development of the HapMap resource that facilitates the design and genotyping of SNPs for the association studies; the development of microarrays

enables to up to 1 million SNPs to be genotyped in a single DNA sample; the existence of large clinical sample collections has helped overcome the problem of heterogeneity (WTCCC 2007; Baum *et al.* 2008; Ferreira *et al.* 2008; Sklar *et al.* 2008).

The Wellcome Trust Case Control Consortium (WTCCC) (WTCCC 2007) used a sample of 2000 individuals for each of 7 major diseases and a shared set of 3000 controls. The group used the GeneChip 500k Mapping Array Set (Affimetrix chip) that comprises 500,568 SNPs. The strongest signal in bipolar disorder was with the marker rs420259 at chromosome 16p12 (P=6.3 X 10^{-8}). At this locus there are very interesting genes such as PALB2 (Partner and Localizer of BRCA2) which is involved in the stability of key nuclear structures; NDUFAB1 (NADH dehydrogenase (ubiquinone) 1, alpha/beta subcomplex, 1) which encodes a subunit of complex I of the mitochondrial respiratory chain; and DCTN5 (dynactin 5) which encodes a protein involved in cellular transport and that interacts with DISC1. Association was also detected with the SNP rs1526805 with p=2,2 X10^{-7}, which is close to KCNC2, a Shaw-related voltage-gated potassium channel. Other highly ranked genes include GABRB1 (GABA A Receptor Beta 1) encoding a ligand-gated ion channel (rs7680321 P=6.2 X10^{-5}), GRM7 (Glutamate Receptor Metabotropic 7, rs1485171 P=9.7 X 10-5) and SYN3 (Synapsin III, rs11089599, P=7.2 X10^{-5}).

Baum *et al* carried out a GWAS in two case control samples of European origin (461 bipolar disorder 1 cases and 563 matching controls from NIMH and 772 bipolar disorder 1 cases and 876 control from a German sample) (Baum *et al.*

2008). They pooled the DNA and ran the samples in the Illumina-Hap550 chip. A total of 88 SNPs covering 80 genes were replicated in both samples. The strongest association came from a marker on the first intron of DGKH (Diacylglycerol kinase eta, $P=1.5 \times 10^{-8}$). This gene encodes a protein involved in the lithium-sensitive phosphatidyl inositol pathway. Other replicated signals were in the gene NXN (combined $P=0.0003$, which encodes a protein neuroredoxin; VGCNL1, a voltage-gated ion channel highly expressed in the brain (combined $P= 8.1 \times 10^{-5}$), DFNB3 gene that encodes the neuronally expressed protein whirlin (combined $P=0.0001$) and SORC2, which encode a VPS10 domain containing receptor, expressed in the developing brain (combined P from 0.0005 to 1.4×10^{-5}). The authors also argue that the results detected markers that had a weak but reproducible association. That none of the replicated genes confer a large population risk of disease. And that these results are compatible with either heterogeneity or a polygenic threshold model.

Sklar *et al* performed a genome wide scan in 1461 patients with bipolar disorder 1 and 2008 controls from the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD) and the UCL collection (Sklar *et al.* 2008). They used the GeneChip Human Mapping 500K Array Set from Affimetrix and successfully genotyped 372,193 SNPs. The strongest single SNPs where in MYO5B (Myosin 5B, $p=1.66 \times 10^{-7}$) and TSPAN8 (tetraspanin-8, $p=6.11 \times 10^{-7}$). Haplotypic analysis supported the single markers results highlighting again MYO5B and TSPAN8 and also EGFR (Epidermal Growth Factor Receptor, $p= 8.36 \times 10^{-8}$). Replications were carried out on two samples, the NIMH trio family sample (n=409 trios) and the Edinburgh case control sample (365 cases

and 351 controls). The three strongest markers were not successfully replicated. However, replicated pValues below 0.05 in a 53Kb region of 18q22 were observed.. The SNPs were approximately 146 Kb from CDH7, a brain expressed cell to cell adhesion protein. Similarly, rs10491113 was found near TMEM132E, an uncharacterized transmembrane protein. Also, two SNPs near MBP (Myelin Basic Protein) and GRM3 (Glutamate Receptor Metabotropic 3) showed positive allelic association.

A analysis of Sklar $et\ al$, WTCCC and the Ferreira GWAS results showed that a SNP was consistently associated to disease in all studies, with $p>0.05$. The SNP is located on the third intron of CACNA1C, the alpha subunit of the L-type voltage-gated calcium channel. Within this gene there are other associated SNPs in both WTCCC and Sklar $et\ al$. Additionally, Sklar $et\ al$ also compared their results to Baum study. The strongest result in the Baum study, for the DGKH, was not positive on the STEP-BD/UCL study. However, Baum's fifth most positive marker, near the gene DFNB31 even not being represented on Sklar's platform, there were other markers for the same gene that showed to be positive with p of approximately 0.0003. Interestingly, the WTCCC study also showed positive association with markers at this locus with $p < 10^{-4}$. The authors argue that the results obtained show that the disease common variant hypothesis does not fit for bipolar disorder. However, if common variants do exist for the disease they may be of very small effect and thus require very large samples to be reliably detected.

Ferreira *et al* used the same data and platform as the WTCCC and STEP-BD/UCL and genotyped more 1,098 individuals with bipolar disorder and 1,267 from the ED-DUB-STEP2 dataset, which included additional samples from the STEP-BD study as well as bipolar cases and controls from the University of Edinburgh and Trinity College Dublin (Ferreira *et al.* 2008). The combined results of the new sample with the STEP-BD/UCL and WTCCC data identified many susceptibility loci for bipolar disorder. Using the genotyped data and imputation, 1.8 million variants were tested in 4,387 cases and 6,209 controls. This study identified a region of strong association (rs10994336, $p = 9.1 \times 10^{-9}$) in ANK3 (ankyrin G). Ferreira *et al* also found further support for the previously reported CACNA1C (alpha 1C subunit of the L-type voltage-gated calcium channel; combined $P = 7.0 \times 10^{-8}$, rs1006737). To conclude, the authors suggest that ion channelopathies may be involved in the pathogenesis of bipolar disorder.

A meta analysis of data from the Baum *et al* and WTCCC GWAS data was preformed (Baum *et al.* 2008). Markers in genes like ZIP3 (combined $P=5 \times 10^{-6}$), a zinc transporter with roles in synaptic transmission and JAM3 (junctional adhesion molecule 3, $p=1 \times 10^{-6}$) a cell to cell adhesion molecule, show consistently positive results in each of the studies and in a combined analysis. The DGKH gene also seems to show consistent association. The most significant marker of Baum's study (rs1012053 $p=1.5 \times 10^{-8}$) showed little evidence of association in the WTCCC study. However, in the latter there are six markers within 2Kb of DGKH that show evidence of association with P-values as low as

0.0006. The authors justify this result has reflect of allelic heterogeneity at the same risk locus.

GWAS has allowed to powerful studies of association across the genome with no prior hypotheses. The results so far point to heterogeneity as being the main complication of genetic studies of bipolar disorder. Ferreira *et al* tested for polygenic transmission in the largest sample so far and could find no evidence for interaction between any associated genes (Ferreira *et al.* 2008). This suggests that bipolar disorder may be caused by multiple rare variations that have an addictive effect on the phenotype rather than a polygenic effect.

It seems that the complexity of the genetics of bipolar disorder, which involves trait, locus and allelic heterogeneity will require many new samples before consensus is revealed. Researchers are using GWAS data to do pathway analysis (Askland *et al.* 2008; Torkamani *et al.* 2008) and/or to converge functional genomics with association data (Le-Niculescu *et al.* 2008). Le-Niculescu *et al* proposes the use of convergent functional genomics in human and animal models to highlight genes that may not have been highlighted in a genetic study (Le-Niculescu *et al.* 2008). They have integrated data from WTCCC, Baum and Sklar GWAS data with human post mortem brain gene expression data, blood gene expression data as well as relevant animal model data. Genes were prioritized according to a score of multiple independent lines of evidence. The top four candidate genes were ARNTL, a transcription factor involved in the circadian rhythms; BDNF; ALDH1A1 that has been implicated in brain development; and KLF12, a transcription factor that represses a zinc finger transcription factor. The

direct pathway analysis using the WTCCC data (Torkamani *et al.* 2008) showed that the enriched signalling pathways were: heparin sulphate and heparin metabolism (*p*=0.01), a pathway involved in the inactivation of dopamine by sulfation; Cytoskeleton remodelling-Alpha-1A adrenergic receptor-dependent inhibition of PI3K (*p*=0.01); niacin-HDL metabolism (*p*=0.03); and glutamate regulation of dopamine D1A receptor signalling (*p*=0.04). Another study has done a pathway-based analysis using the NIMH and WTCCC GWAS data (Askland *et al.* 2008). Their findings suggest involvement of ion channel structural and regulatory genes, including voltage-gated ion channels and the broader ion channel group that comprises both voltage and ligand-gated channels.

1.5.4 Copy Number Variations

Recent studies have brought to attention to copy number variations (CNVs) as a source of genetic variation and a possible factor contributing to disease (Stankiewicz *et al.* 2002; Iafrate *et al.* 2004; Sebat *et al.* 2004). CNVs are segments of DNA ranging from one kilobase to several megabases, for which copy-number differences occur due to duplication, insertional transposition or deletion events (Feuk *et al.* 2006; Redon *et al.* 2006). CNVs can disrupt a gene's function by occuring within the gene or in regulatory regions. CNVs may also generate novel fusion proteins, create new imprinting patterns and create differential allelic expression (Lupski *et al.* 2005).

Rare CNVs have been associated with mental retardation, autism and schizophrenia (de Vries *et al.* 2005; Stefansson *et al.* 2008; Weiss *et al.* 2008). In

addition, recent data has been published implicating CNVs in BP. A study done on 46 individuals in a three-generation Old Order Amish pedigree, using the Illumina HumanHap550 BeadChip, has identified 50 CNV regions that ranged in size from 12 to 885 kb (Yang *et al.* 2009). 19 of these CNVs regions were well characterized regions and were compared with expression data and 11 (58%) were associated with expression changes of genes within, partially within or near these CNV regions in fibroblasts or lymphoblastoid cell lines at a nominal P value <0.05. A set a set four CNVs, located at 6q27, 9q21.11, 12p13.31 and 15q11, were more present in cases than controls. The variants were affecting the expression of neuronal genes within or near the CNVs. A second study on a small sample of 35 control, 35 schizophrenia and 35 BP patients was carried out using the bacterial artificial chromosome (BAC) array comparative genome hybridization (aCGH) (Wilson *et al.* 2006). The CNV analysis showed aberrations in BAC clones containing regions 1p34.3, 14q23.3 and 22q12.3. These three clones were found to contain brain-expressed genes GLUR7, GRM7 AKAP5 and CACNG2. These genes are involved in glutamate signalling on bipolar and schizophrenia patients.

1.6 Mouse as an animal model

`

Genes cannot be manipulated in humans, therefore scientists turn to other organisms to investigate gene function. Animals such as the fruit fly and the nematode worm have been used in the past. However, over the past century, the mouse has developed into the premier animal model system for genetic research.

Analysis of the mammalian lineage indicate that mice and humans diverged 75 million years ago, which is a recent evolutionary event (Madsen *et al.* 2001; Murphy *et al.* 2001). Both groups share similar body plans, organ systems and mechanisms of physiological regulation. Mice and humans genomes are also arranged in a similar pattern. Their genomes share roughly the same size, contain the same number of genes and a large proportion of the mouse and human genomes are syntenic, which means that they possess chromosomal regions with the same order of genes. In addition, approximately 99% of mouse genes have human counterparts. Conversely, mouse orthologs can be identified for 99% of human genes (Tecott 2003).

The existence of a cortex and other evolutionary adaptations in humans make the human brain and behaviour much more complex than in mice. However, the brains of vertebrates share a common structural organization and the brain structure is well conserved among mammals. This means that even though there are differences, the extensive neuroatomical and genetic homologies give rise to a wide variety of behavioural processes that are conserved between species (Tecott 2003), such as hunger, fear, aggression, sleep, circadian rhythms, classical and operand conditioning and sexual behaviour. Also, there are

behavioural responses to drugs that are common among humans and mice. Looking at the genetic perturbations and behaviour, there are mutated genes that have similar effects in both species. An example is a behavioural disturbance caused by the monoamine oxidase A gene, where both species become more aggressive and display impulsive sexual behaviours towards females (Brunner *et al.* 1993; Cases *et al.* 1995). In addition, the hypothalamic neuropeptide orexin gene when mutated in both humans and mice cause sudden episodes of inactivity. This syndrome is denominated narcolepsy (Chemelli *et al.* 1999; Nishino *et al.* 2000; Peyron *et al.* 2000). We must remember that not all mutations will result in similar phenotypes in the two species, which is a consequence of the evolutionary differences between them. Despite these discrepancies, mice still are a valuable tool on the study of genes and behaviour.

One of the key events that turned the mice into a successful tool in science was the development and application of new molecular technologies. Strategies, such as transgenic and gene targeting, are used to generate lines of mice with enhanced, reduced or altered gene expression. This allows precise engineering of genetic alterations to study the neuronal basis of behaviour (Tecott 2003; Holmes *et al.* 2004). Also the capacity of create homozygous populations of mice through inbreeding techniques permitted minimization of the extent to which genetic factors contribute to variability in response to experimental manipulations. This is an advantage since the strains possess a unique set of alleles and these results in distinct biological properties such as size, cancer susceptibility or behavioural traits (Tecott 2003).

In addition to all the advantages mentioned before, there are also economical and practical issues that make mice popular in labs. Mice are small, easy to maintain in the laboratory and (compared to most mammals) have a short breeding cycle (about 2 months). They can produce 4 to 15 offspring per litter and approximately one litter every month (Wolfensohn *et al.* 2003).

All of these reasons make mice very useful in the study of genes, and more specifically in the study of behavioural disorders. Mouse models are beginning to be developed for bipolar disorder. Since bipolar disorder is a complex disease with heterogeneous type of behaviours, an endophenotype based research that focuses specific behaviours in transgenic or normal mice have been carried out. There are models that focus on features of depression and others that have been used to assess manic behaviours (Cryan *et al.* 2007). An example of a mouse mania model was developed by manipulating hippocampal GluR1 and GluR2 receptors resulting in manic-like behaviour (Du *et al.* 2008). There is also a study with lamotrigine where researchers study its antidepressant-like effect in the mouse and also the involvement of the noradrenergic system (Kaster *et al.* 2007). Kato *et al* used the mouse as a model to study the effect of mitocondrial DNA mutations in the mouse brain. They produced a mouse with mitochondrial DNA deletions that presented bipolar-like behaviours (Kato 2006).

1.7 Slynar and IGF1

1.7.1 Slynar

A maximum lod score of 2.11 to bipolar disorder was found on chromosome 12q23-q24.1 (Dawson *et al.* 1995). Later a lod score of 3.67 and sib pair linkage p value of 0.00003 at 12q23-24 was published (Ewald *et al.* 1998; Ewald *et al.* 2002). A haplotype sharing study of 12q24 in a genetic isolate also confirmed this region being linked to bipolar disorder (Degn *et al.* 2001). In Canada, (Morissette *et al.* 1999; Shink *et al.* 2005) a kindred showed linkage to bipolar affective disorder at 12q23-12q24 with a lod of 3.35. Multiply affected families with just unipolar affective disorder have also shown very good evidence for linkage at C12q22-24 with a multipoint heterogeneity LOD score of 6.1 (p=0.0000007) (Abkevich *et al.* 2003). A genome wide scan in a sample of 60 affected sibling pairs (ASP) in Irish families gave suggestive evidence for linkage at 12q24. The p values for the results of the multipoint linkage analysis were p=0.0146 and with the pairwise linkage analysis gave a p=0.043892 (Cassidy *et al.* 2007).

The UCL linkage study of 12q22-24 gave a lod of 2.7 (Curtis *et al.* 2003). The region was investigated by fine mapping to find the exact position of a bipolar susceptibility locus on 12q24 in two case-control association samples originating from Denmark and the UK. In a sample of 681 bipolar cases and 550 controls statistically significant allelic association with bipolar disorder was found for twelve markers (Kalsi *et al.* 2006). The region covered in represented in Figure 1.2. The pattern of allelic association was similar in the Danish and UK samples.

In the Danish sample the most significant single marker association was with the microsatellite 307CA2 (p = 0.0014). The most significant single marker association in the UK sample was with D12SDK1 (p = 0.0002). Tests of haplotype association in the UK sample with the markers 29818insT and 307CA1 gave a permutation test pValue of 0.0013. A three marker haplotype was also significantly associated with bipolar disorder in the UK sample with permutation test p value of 0.0079. The results validate the earlier linkage studies and strongly implicate a 300kb region on 12q24 that contributes to the aetiology of affective disorders.

Chromosome Region 12q24
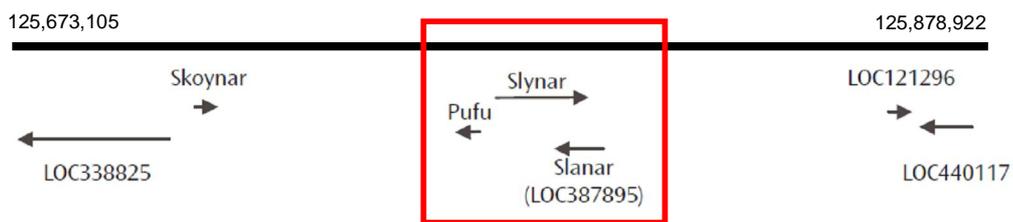
Figure 1.2 Schematic of the 2Mb region on chromosome 12q24 that surrounds Slynar gene. The genes Skoynar, Slynar and LOC121296 are transcribed from the sense strand and LOC338825, Pufu, Slanar and LOC440117 are transcribed from the antisense strand. Fine mapping in the UCL and Danish samples had implicated a 300 Kb region (indicated by the red square) in which Slynar was the most promising candidate gene (Kalsi *et al.* 2006). The genetic distance was obtained from UCSC genome browser, version May 2004. Figure adapted from (Kalsi *et al.* 2006).

Within the 300 Kb region, using RT-PCR on brain mRNA several CNS expressed sequences in this region were identified (Figure 1.2). The identified genes were Slynar, Pufu and Slanar. The most promising candidate gene in this 300 kb region is denoted Slynar in the Aceview database (Unigene cluster Hs.369455) or AY070435 in the UCSC genome browser, since Slynar bears homology to a macaque brain-expressed cDNA and to a chimpanzee genomic sequence (Kalsi *et al.* 2006). The region is not well characterised in terms of

transcription, alternative splicing and protein translation and no known function has been attributed to this gene. So far, four alternative splice variants have been found in mRNA samples from GeneBank (Figure 1.3): AY070435 or Slynar, which is the hypothetical protein, and the cDNA was found in a brain sample (Guo *et al.* 2001); AK091521 is a foetal brain cDNA (Ota *et al.* 2004); BC039096 is a IMAGE clone found in testis (Strausberg *et al.* 2002); CR615184 is a clone found in human foetal brain (Li *et al.* 2004). Slynar has been sequenced to identify aetiological base pair changes, however up to now no changes were found to be associated with disease. On the reverse strand of Slynar another brain expressed cDNA was detected. This gene is denominated Slanar or LOC387895 by ACEVIEW and BC069215 by UCSC browser (Figure 1.2). However, due to the stronger homology to primates, Slynar was chosen for functional studies in this region and not Slanar.

The study of early onset forms of a disorder has proved a successful strategy in the mapping of genetic traits. It has been found that the risk to first degree relatives is elevated in the families of bipolar probands with early onset. An analysis of the UCL bipolar sample data was done grouping people with an early onset, <21 yrs, using OPCRIT and SADS-L data (n=313). Allelic association was tested by comparing this group against controls (n=644) for 30 markers between 1634tet and D12SDK1. Chi squared and CLUMP tests were used where appropriate. Haplotype analysis using GENECOUNTING was performed. As previously reported seven 12q24 markers were associated with bipolar disorder the full UCL sample (Kalsi *et al.* 2006). Increased evidence for allelic association was observed in the early onset group. A total of ten markers were found to be

associated, generally at increased levels of significance. Haplotype analysis also yielded increased evidence for association with pValues ranging from <0.0001 to 0.02. Reference to HapMap data indicates that a relatively common haplotype containing the minor (A) allele of the SNP rs7133178 may harbour aetiological changes. This SNP is in the intron of Slynar however is in the exon of BC069215 (LOC387895), the gene that is located on the negative strand.
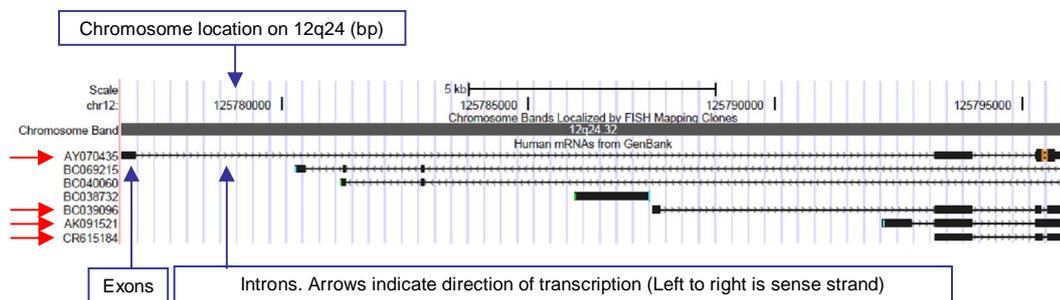
Figure 1.3 Alignment of Slynar (AY070435) and its alternative splicing transcripts with the genomic DNA (figure taken from UCSC genomic database, assembly March 2006). On the top, the continuous grey bar represents the chromosome band demarking the relative location. Below are represented all the human mRNAs found to be expressed from this region. The sequence is represented by small black boxes and thinner line with arrows. The black boxes correspond to exons and the thinner line represents the introns. The arrows show the direction of the gene's transcription. The red arrows (on the left of the legend) indicate four transcripts correspondent to Slynar and the alternative splicing products: AY070435, BC039096, AK091521 and CR615184. The remaining transcripts do not refer to Slynar, as they are transcribed from the antisense strand. In exon 3 and 4 of AY070435 sequence, the orange bands correspond to a difference in the sequence between the genome and the mRNA found.

1.7.2 IGF1

Insulin Growth Factor 1 (IGF1) is a 70 amino acid peptide that shares homology with insulin (Rotwein 1991). IGF1 actions are mediated by a cell surface receptor called the Type 1 IGF Receptor (IGF1R). IGF1R is widely expressed in the brain, being concentrated in neuron rich structures such as hippocampus and the olfactory bulb (Aberg *et al.* 2006). These receptors are receptor tyrosine kinases. After IGF1 binds IGF1R, it induces autophosphorylation of the insulin

receptor substrate (IRS) adaptor proteins. These IRS proteins bind to tyrosine phosphate docking sites on the activated receptors, undergo phosphorylation themselves, and then recruit additional signalling proteins to transduce IGF1 actions (Bondy *et al.* 2004). Additionally, actions of IGF1 are modulated by a group of cystine-rich IGF-binding proteins (IGFBPs). Six different IGFBPs have been characterized to date, and they prolong IGF1 half life by impending proteolysis and modulate the peptide's interaction with IGF1R (Duan 2002).

Apart from regulating somatic growth and metabolic processes, IGF1 has many functions in the brain. The best established activity of IGF1 is its crucial role in brain development (Russo *et al.* 2005). In the adult, IGF1 is a pleiotropic peptide involved in numerous processes to maintain brain homeostasis. At a cellular/molecular level, IGF1 acts in basic energy regulatory loops, cell sorting and cell to cell communication, among others, making this peptide essential for normal brain function (Trejo *et al.* 2004). At a system level, IGF1 acts on neuronal specific traits such as synaptic plasticity and neurotransmission. IGF1 also has an extensive neuroprotective activity (Fernandez *et al.* 2007). IGF1 is expressed in the brain, especially during brain development and in response to injury (Bondy *et al.* 1991; Lee *et al.* 1993). IGF1 can also be transported to the brain through the blood brain barrier (Reinhardt *et al.* 1994). It has been suggested that IGFBP, which is abundantly expressed along the blood-brain barrier, may help IGF1 be transported into the brain (Lee *et al.* 1993; Reinhardt *et al.* 1994).

Indirect evidence suggests that stress and depression inhibit neurogenesis in the adult hippocampus, which leads to the notion that hippocampus cell proliferation

and neurogenesis may be a target for the treatment of depression (Kempermann *et al.* 2003; Santarelli *et al.* 2003). IGF1 is present in the adult brain in the hippocampus and the olfactory bulb (Kar *et al.* 1993). Repeated systemic or central administration of IGF1 has been shown to increase neurogenesis in the dentate gyrus of the hippocampus (Aberg *et al.* 2000). It has also been shown that increase of neurogenesis by exercise may be mediated by IGF1 (Trejo *et al.* 2001). In addition, it was reported that chronic administration of the antidepressants venlafaxine and fluoxetine increased the protein levels of IGF1 in the hippocampus (Khawaja *et al.* 2004). In a recent paper, central administration of IGF1 in mice has shown to have an antidepressant-like effect (Hoshaw *et al.* 2005). A recent review has linked antipsychotic drug mechanisms, the insulin signalling pathway and possible abnormalities in this pathway with schizophrenia (Girgis *et al.* 2008).

1.7.2.1   IGF1 GWAS results

A GWAS study was carried out on 506 UCL bipolar cases and 510 controls, together with an American STEP-BD sample (Sklar *et al.* 2008).

In this study, 10 out of 37 SNPs covering the IGF1 region were found to be associated with BP (Table 1. Table 8.7). Five of the associated SNPs (rs5742688, rs2072592, rs12423791, rs4764698, rs5742615) are located within the gene in the intronic regions with p values ranging from 0.0057 to 0.000037. The other five markers are located upstream IGF1, in the putative promoter region of IGF1. These markers span a region of approximately 3000 bp. However, it is known that promoter control regions can be as far as 14.5 kb (Ho *et al.* 2006).

Assessing Hardy-Weinberg equilibrium (HWE) is often employed as an important initial step for genotype data quality checking in genetics studies (Nielsen *et al.* 1999). All the UCL data was checked for HWE. The only SNP that was out of HWE was rs5742615. This is the most significantly associated SNP (p=0.000037). Analysis of all data was done in the presence and absence of this SNP, and it had no effect on pValues.

Table 1.4 Tests of association of SNPs in IGF1 region in the UCL GWAS data. UCSC March 2004 assembly positions. * SNP out of Hardy-Weinberg Equilibrium.

Marker	Marker Location	Alleles and Observed Allele Frequencies						χ2	P value
rs5742688		G			A				
Controls	101314993	13	0.013		1001	0.987		7.637	**0.0057**
Cases		31	0.031		977	0.969			
rs2072592		G			A				
Controls	101316099	992	0.973		28	0.027		11.152	**0.00080**
Cases		1004	0.992		8	0.008			
rs12423791		C			G				
Controls	101361295	989	0.973		27	0.027		13.334	**0.00030**
Cases		998	0.994		6	0.006			
rs4764698		G			C				
Controls	101362527	987	0.973		27	0.027		13.375	**0.00030**
Cases		998	0.994		6	0.006			
rs5742615*		G			T				
Controls	101373268	1002	0.969		31	0.031		17.006	**0.000037**
Cases		1002	0.994		6	0.006			
rs35765		C			A				
Controls	101384163	85	0.083		935	0.917		5.257	**0.0219**
Cases		115	0.114		897	0.886			
rs2607983		A			G				
Controls	101402171	114	0.112		904	0.888		7.153	**0.0075**
Cases		154	0.152		858	0.848			
rs12309723		A			C				
Controls	101425980	975	0.956		45	0.044		3.969	**0.0464**
Cases		984	0.972		28	0.028			
rs703542		G			C				
Controls	101443545	135	0.133		881	0.867		4.975	**0.0257**
Cases		170	0.168		840	0.832			
rs703548		G			A				
Controls	101447140	139	0.137		879	0.863		4.04	**0.0444**
Cases		170	0.169		838	0.831			

2 Aims of the Project

The objective of the project was to study two susceptibility genes for bipolar disorder that are located in the linkage to bipolar disorder region 12q23-24.

My aim was to characterize Slynar and to attempt to find out what function it has in the human brain. In order to achieve that, studies were done in a mouse model and in human cell lines. The first step of this project was to find a homologous RNA transcript of Slynar in the mouse. To detect the gene and confirm its sequence, techniques such as RT-PCR, qPCR, cloning and sequencing were used. Once confirmed Slynar expression in the mouse, attempts to obtain a full cDNA clone of Slynar through RACE and cDNA library screening were done. In addition, a northern blot was done to evaluate where Slynar was expressed in the mouse, how many alternative transcripts can be found and what are the sizes of the different transcripts. Similarly to what it was described for the mouse project, studies to detect and characterise Slynar in the human were carried out in neuroblastoma SH-SY5Y cells, differentiated and non-differentiated into mature neurons, and in human brain RNA. In addition to detection of the gene by RT-PCR and qPCR and detection of a Slynar cDNA clone by RACE and cDNA library screening, knock out assays of Slynar in SH-SY5Y cells were also done to try to analyse the effect of the gene silencing in the cells.

As showed in section 1.6.2 of this thesis, through the analysis of GWAS data IGF1 was implicated in bipolar disorder. My aim was to sequence IGF1 to find possible aetiological base pair changes and to understand how these changes may affect the gene's function and its pathways. Any SNPs or polymorphic

microsatellites found by sequencing were genotyped using Kaspar, Melting Curve analysis or Polyacrilamide gels methodologies. Association analysis was carried out using Chi square tests, with the exeption of the microsatellites that were analysed using the statistical software CLUMP. All haplotypic analysis was carried out using GENECOUNTING and whole genome significance analysis was done by COMBASSOC.

3 Methodology

3.1.1 DNA Extraction

3.1.1.1 DNA Extraction from Whole Blood Cells

Genomic DNA was extracted from whole blood samples using a standard cell lysis, proteinase K digestion, phenol/chloroform ethanol precipitation method. It is a two day method. In the first day the leukocytes are isolated and then lysed to release the genomic material, followed by overnight digestion of proteins with proteinase K. The second day involves the isolation and extraction of genomic DNA. The DNA was extracted from fresh blood if it was practical to do so. Otherwise the blood was stored at -80^0C.

On day one…

If the blood was stored at $-80\,^o$C, it was allowed to gently thaw out in a water bath at $30\text{-}35^o$C. This was done to prevent the cells from lysing from shock and the release of damaging enzymes such as DNase.

The fully thawed blood was then transferred to 50ml centrifuge tubes, which were then topped to 50ml with 1x lysis buffer. The tube was inverted several times to ensure mixing. This allowed the lysis buffer to break up the whole red blood cells (erythrocytes) whilst leaving the DNA containing white blood cells (leukocytes) intact. The lysate was then spun in a balanced centrifuge at 3000rpm (2000g) at 4^oC for 15 minutes to pellet and isolate the white blood cells. The pellet was re-suspended in 50ml of 1x lysis buffer to remove any remaining red blood cells and again centrifuged at 3000 rpm for 15 minutes.

The cleaned white blood cell pellet was re-suspended in 500µl of 10% sodium dodecyl sulfate (SDS), which would break open the cells by disrupting the lipid membrane, hence releasing its contents including the genomic DNA into the solution. In addition, 10ml proteinase K buffer was added to provide stable optimum conditions for the proteinase K enzyme. To the lysate, 50µl of 20mg/ml proteinase K was mixed in. The presence of the enzyme would break down complex protein allowing for their efficient removal and also to deactivate any damaging enzymes such as DNAase. This lysate was incubated in a water bath at 55°C (just below the optimum temperature of proteinase K at 65 ^{o}C) on a shaking platform overnight, to allow for the reaction to occur as far as possible.

On day two...

5ml of buffered phenol along with 5ml of chloroform (1:24 Isoamyl alcohol: chloroform) was mixed with the cell lysate. The phenol-chloroform mixture is used to dissolve the lipids and agglomerate the proteins. After centrifugation, layers do appear. The nucleic acids will be on the top aqueous layer while agglomerated protein would be lower down.

The mixture was then centrifuged at room temperature for 15 min at 3000rpm which allow the mixture to separate out into the organic solvent (phenol-chloroform) layer at the bottom (containing lipids), the protein interphase, and on top the aqueous layer containing the nucleic acid. This aqueous layer was transferred into a fresh labelled 50ml centrifuge tube for DNA precipitation. The organic solvents are disposed of safely according to UCL's health and safety disposal of hazardous solvents protocol.

It is difficult to extract the aqueous layer up to the interphase without disturbing the layer itself. So we used 1 gram of PVPP (Polyvinylpolypyrrolidine) mixed with 5ml of TE to improve the separation of layers. Once centrifuged with the organic mixture, PVPP will hold down the protein interphase along with the organic solvents allowing one to remove the maximum amount of the aqueous layer without contamination.

From this aqueous layer, DNA was precipitated by adding 1500µl of 3M Sodium Acetate. 30ml of absolute ethanol are added to the tube and then the tube was inverted gently until the DNA precipitated out of 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, in order not to interfere with the Polymerase Chain Reaction (PCR). The DNA clump was then transferred into a labelled 1.5ml micro screw tube containing 500µl of TE. The DNA samples were stored away from light at room temperature for about a week. This was to allow the DNA to dissolve into solution before quantification.

3.1.1.2   DNA Extraction from Saliva Samples

The Oragene DNA Self-Collection Kit (Oragene) 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 released when the lid is tightened immediately after the saliva collection.

Next step was the incubation of the container with the sample in a water bath at 50°C for 1 hour. 160µL of Oragene DNA Purifier it was added to 4 mL of sample and this was incubated in ice for 10 minutes. Then, the samples were centrifuged at high speed for 10 minutes at room temperature. The pellet was discarded and an equal volume of room temperature absolute ethanol was added. The samples were well mixed 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 centrifuged at high speed for 10 minutes at room temperature. The supernatant was discarded and the DNA pellet was left to dry. The DNA was rehydrated in 500µL of TE and stored at room temperature, protected from light, for a week. This allowed the DNA to dissolve completely.

3.1.2 DNA Quantification

3.1.2.1   Quantification of Saliva and Whole Blood DNA Samples

All DNA samples were quantified with Picogreen (Molecular Probes), in a Light Cycler480 (Roche).

The DNA samples were quantified by transferring 2µl of each sample into a fresh labelled 2ml microcentrifuge tube containing 78µl of TE. These tubes were thoroughly mixed and 2µl of each of these samples were transferred to 96 well plate, containing 18µl of TE. Along side the samples to be quantified was a series of wells containing known concentrations of DNA derived from calf thymus. This was diluted to produce samples at 0, 1, 10, 20, 50, 75 and 100ng/µl. This was done in order to produce standard curve from which the concentration of the samples could be derived from.

The DNA samples were mixed with 20µl of Picogreen (150µl Picogreen dissolved in 30ml of TE) which is a fluorescent dye which specifically binds to double stranded DNA. The fluorescence is directly proportional to the quantity of DNA present. From the results, some of the DNA was diluted into a concentration of 25ng/µl for direct use. The rest was labelled as stock with its known concentration. All DNA is stored below 5^0C in a dark place to maintain its quality.

3.1.2.2 Quantification of cDNA and plasmid DNA

For the quantification of DNA samples, such as cDNA or plasmid DNA that resulted from bacterial purification, was used a NanoDrop® ND-1000 Spectrophotometer (Thermo Scientific). This spectrophotometer measures DNA samples up to 3200 ng/mL without dilution and with high accuracy and reproducibility. It uses a spectrum range from 220 to 750 nm. A 1 µl sample was pipetted onto the end of a fiber optic cable (the receiving fiber). A second fiber optic cable (the source fiber) was then brought into contact with the liquid

sample causing the liquid to bridge the gap between the fiber optic ends. The gap was controlled to both 1mm and 0.2 mm paths. A pulsed xenon flash lamp provides the light source and a spectrometer utilizing a linear CCD array was used to analyze the light after passing through the sample. The data was then analyzed by the ND-1000 PC software.

3.1.3 RNA Extraction

3.1.3.1 <u>RNA Extraction from Tissue</u>

Different body tissues (brain, cerebellum, liver, lungs, heart and testis) were extracted from adult mice. The animals were killed under the Schedule 1, appropriate methods of human killing, which came to effect in 1997. The sacrifice was done by exposure to carbon dioxide in a rising concentration, followed by the dislocation of the neck, to ensure that the killing process was completed (Wolfensohn *et al.* 2003).

1 ml of TRIZOL Reagent (Invitrogen) was added to each 50 to 100 mg of mouse tissue. The tissue was homogenized with the homogenizer DIAX 900 (Heidolph), and samples were incubated at room temperature for 5 minutes. After adding 200 µL of chloroform to each sample, the tubes were shaken vigorously for 15 seconds and stored at room temperature for 15 minutes. The samples were spun down in a microcentrifuge (centrifuge 5417R, Eppendorf) at 12,000 g for 15 minutes at 4° C to phase separate the sample. To precipitate the RNA, the aqueous phase was transferred to a fresh tube and 500 µL of isopropanol was added. The samples were stored at room temperature for 10 minutes and

centrifuged at 12,000 g for 10 min at 4° C. The supernatant was removed and the pellet was washed by vortexing in 1mL 75% ethanol. This was centrifuged at 7,500 g for 5 min and 4° C. In the end, the pellet was air-dried for 3 to 5 minutes and ressuspended in 100 µL of RNAse free water.

3.1.3.2   RNA Extraction from Cells

Cells were grown in a 75 cm^2 T-Flask following the normal growth routine. When the 100% confluency was reached, 1 mL per 10 cm^2 (area based on the culture dish) of TRIZOL Reagent (Invitrogen) was added to lyse the cells. The subsequent procedure is similar to what it is done with homogeneized tissue samples (see section 3.1.3.1). A chloroform and isopropanol RNA extraction was done, followed by rehydration of RNA in RNAse free water.

3.1.3.3   Total and PolyA$^+$ RNA Purification

To purify total RNA the RNeasy® Mini Kit (Qiagen) was used. This technology combines the selective binding properties of a silica-gel-based membrane with the speed of microspin technology. The protocol followed was the RNeasy® Mini Protocol for RNA cleanup from the RNeasy® Mini Handbook (June 2001). These samples were subjected to DNaseI treatment, as described in the same handbook.

Two different kits were used for the purification of poly A$^+$ RNA. The Oligotex® mRNA Kit (Qiagen) is based on spin columns purification and the mRNA Isolation Kit (Roche) uses magnetic separation. In the Oligotex® mRNA Kit spin columns are used which have an affinity reagent, oligotex suspension, for the

detection, isolation and purification of nucleic acids containing a poly A^+ sequence. The protocol followed was the Oligotex mRNA Spin-Column Protocol from the Oligotex® Handbook (May 2002). In the mRNA Isolation Kit (Roche), a biotin-labeled oligo(dT)$_{20}$ probe and streptavidin magnetic particles are added to total RNA. The biotin probe will bind to the poly A^+ sequence which has high affinity to bind streptavidin particles. With a magnetic separator streptavidin magnetic particles are captured and with them the mRNA. The protocol used was from the kit's instruction manual (Roche, June 2005).

With the Oligotex® mRNA Kit the poly A^+ RNA samples were cleaner, however with mRNA Isolation Kit the obtained yields were higher.

3.1.4 RNA Quantification and Quality Assessment

The Agilent 2100 bioanalyzer (Agilent Technologies) is an analytical instrument based on lab-on-a-chip technology. The system uses a miniature glass chip that contains a network of interconnected channels and reservoirs. The channels are filled with a gel matrix, composed by a sieving polymer and a fluorescence dye, and there are wells for buffer and the sample. Sample components are electrophoretically separated and detected by the emission of fluorescence. Analysis is done on gel images and electropherograms (Figure 3.1). These allow total RNA and mRNA quantification and accurate quality evaluation.
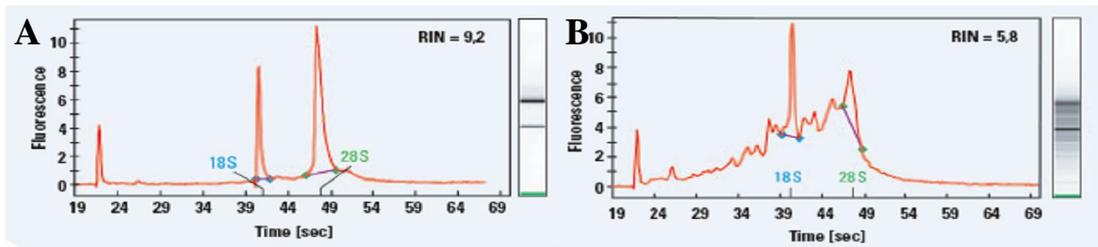
Figure 3.1 An example of a total RNA electropherogram analysis. A: Intact total RNA sample; B: Partially degraded total RNA sample (Reagent Kit Guide; Edition 2003, Agilent Technologies).

To prepare the chip and to analyse the RNA samples on the 2100 bioanalzer I used the RNA 6000 Nano Assay (Agilent Technologies) and the Reagent Kit Guide (Edition 2003) protocol was followed. The RNA Integrity Number (RIN) is calculated to give a better quantitative evaluation of the quality of the RNA. Values obtained were within the range expected for good quality RNA. However, these results were lost due to a software upgrade done to the system.

When a simple quantification of the RNA sample was required, the Nanodrop system was used (3.1.2.2).

3.1.5 cDNA Synthesis

For the conversion of total RNA or mRNA into cDNA anchored oligo dT primers (Promega) were used, or random nonamers primers (Promega), or a mix of both.

For the primer annealing reaction, the following were added to a 0.5 mL microcentrifuge tube: (1) mRNA between 50 ng to 1 μg; (2) 1μL of the primers; (3) RNase free water (Ambion) up to a total volume of 11μl. The reaction

mixture was incubated at 70° C for 10 min and immediately put on ice to allow primers to anneal with mRNA template. For the extension process the following were added to the tube: 10X First Strand Reaction, 0.1 M DTT, 50 mM dNTP mix and 10 U of SuperScript™ III Reverse Trancriptase, in a total volume of 20 µl (All reagents from Invitrogen). These reaction mixtures were incubated at 42° C for 1.5 hours for the extension process. In the end, cDNA can be used immediately or stored at -20° C.

As a negative control in subsequent PCR reactions, Reverse Transcriptase (RT) was not added to some of the RNA samples. This means that if any amplification occurs during PCR it was due to gDNA contamination.

3.1.6 Primer Design

Primers were generally 15-25 nucleotides long for specific targeting and amplification and must span the specific region of interest.

Primers were designed using the Primer3 software, developed by MIT Whitehead Institute (Rozen *et al.* 2000). It is free software, which is available on the internet (http://frodo.wi.mit.edu/). Primer 3 uses algorithms to pick up primer sequences, taking in consideration the following parameters:

    (a) sequence specificity;

    (b) avoidance of repetitive sequences;

    (c) forward and reverse primers with similar melting temperatures (Tm);

(d) pairing of primers with a low probability of forming loops (which would decrease amplification efficiency).

For sequencing reactions, M13 tails (Table 3.1) were added to the 5' end of the forward and reverse primers. These M13 sequences worked as sequencing priming binding sites during subsequent sequencing amplification reaction. This makes the whole sequencing procedure easier, since in all the sequencing amplification reactions only one pair of M13 primers were used, regardless of the region to be sequenced. After adding the M13 tail to the primers 5' end, the entire primer sequence was checked to see if secondary structures such as a hairpin loop would inhibit the amplification. This was checked by using the OligoCalc: Oligonucleotide Properties Calculator (Kibbe 2007) (http://www.basic.northwestern.edu/biotools/oligocalc.html). For genotyping reactions a similar procedure is followed, but the M13 tail is only added to one of the primers. The addition of the M13 tail to the forward or reverse primer will depend on the secondary structure formation, which was checked using OligoCalc.

Table 3.1 M13 tails sequences.

M13	M13F CACGACGTTGTAAAACGAC	M13R GGATAACAATTTCACACAGG

After being designed, the primers were ordered from Sigma Genosys and diluted in TE buffer at a concentration of 100 μM.

3.1.7 Polymerase Chain Reaction (PCR)

3.1.7.1   <u>Mastermix for General Optimization and Amplification</u>

Newly designed primers need to be optimised to find the optimal master mix concentrations and PCR conditions to produce the best results and a successful amplification of the targeted region. Optimisation was carried out on a few DNA samples that have been excluded from the study and PCR products were visualised on agarose or polyacrylamide gels.

For optimisation of primers, four common master mixes conditions were used to amplify the target region (Table 3.2). These master mixes can be manipulated to produce the optimal conditions by, for example, altering the magnesium concentration or by the presence or absence of betaine. The magnesium ion concentration variation influences the specificity of the PCR reaction and also acts as a co-enzyme to the Taq polymerase. The presence of 5M betaine can be valuable, since betaine helps to reduce the melting point and it is useful in regions of high GC content. The optimisation conditions that provided the best results were then used to amplify the entire sample. Where the samples are suboptimal further manipulation of the PCR conditions were performed.

All the master mix components were from Bioline with the exception of the primers and the betaine that were from Sigma.

Table 3.2 Volumes (μl) of reagents in a single 12μl reaction and the common conditions used for primer optimisation and amplification.

	Condition A	Condition B	Condition C	Condition D
	2.0mM MgCl$_2$ + Betaine	2.0mM MgCl$_2$ + No Betaine	2.5mM MgCl$_2$ + Betaine	2.5mM MgCl$_2$ + No Betaine
10X Buffer	1.2	1.2	1.2	1.2
Betaine 5M	2.4	0	2.4	0
50mM MgCl$_2$	0.48	0.48	0.6	0.6
25mM dNTP	0.1	0.1	0.1	0.1
F primer 100pmol/μl	0.036	0.036	0.036	0.036
R primer 100pmol/μl	0.036	0.036	0.036	0.036
M13 800 100pmol/μl				
M13 700 100pmol/μl	0.036	0.036	0.036	0.036
Taq 1U/μl	0.19	0.19	0.19	0.19
DNA 25ng/μl	2.0	2.0	2.0	2.0
Water	5.52	7.92	5.39	7.79
Total Volume (μl)	12.0	12.0	12.0	12.0

3.1.7.2   <u>PCR Cycling Conditions</u>

Once the primers are added to the master mix containing all the required reagents for DNA amplification (such as dNTPs, Taq, magnesium and buffer and also the genomic template) the PCR can begin. The PCR was run on the PCR thermal cycler MWG-HT Primus 96.

The PCR runs in a three step cycle of:

1. Denaturation: which typically occurs at 93-95^{0}C

2. Reannealing: where the primers bind specifically to their complementary sequence. The annealing temperature usually depends on the melting temperature (T$_m$) of the expected duplex. This tends to be approximately 5^{0}C below the expected T$_m$.

3. DNA synthesis: extension of the complementary strand initiated by the annealed primer which occurs at 70-75^{0}C.

This runs for approximately 25-35 cycles and is finished by a final extended hold at 72^{0}C to complete all synthesis and extensions of the targeted region. After about 25 cycles approximately 2^{26} copies of the targeted region would be made.

Usually, one of the three standard PCR programs was used to amplify a specific region successfully. They are Standard 55^{0}C, Standard 60^{0}C and Touch Down. These programs can be manipulated to adjust the annealing temperature and by adding or removing one or two cycles to obtain a clean specific amplification of the target region. The cycles are set out in Table 3.3.

Table 3.3 PCR cycling conditions.

Standard 55^{0}C	**Standard 60^{0}C**	**Touch Down**
Lid Heated to 105^{0}C	Lid Heated to 105^{0}C	Lid Heated to 105^{0}C
Products are denatured at 94^{0}C for 5 minutes	Products are denatured at 94^{0}C for 5 minutes	Products are denatured at 94^{0}C for 5 minutes
35 cycles of:- 94^{0}C – 30 seconds 55^{0}C – 30 seconds 72^{0}C – 30 seconds	35 cycles of:- 94^{0}C – 30 seconds 60^{0}C – 30 seconds 72^{0}C – 30 seconds	3 cycles of:- 94^{0}C – 30 seconds 63^{0}C – 30 seconds 72^{0}C – 30 seconds
Hold at 72^{0}C – 10 minutes Store at 4^{0}C	Hold at 72^{0}C – 10 minutes Store at 4^{0}C	3 cycles of 94^{0}C – 30 seconds 60^{0}C – 30 seconds 72^{0}C – 30 seconds
		3 cycles of 94^{0}C – 30 seconds 57^{0}C – 30 seconds 72^{0}C – 30 seconds
		3 cycles of 94^{0}C – 30 seconds 54^{0}C – 30 seconds 72^{0}C – 30 seconds
		3 cycles of 94^{0}C – 30 seconds

51^0C – 30 seconds
72^0C – 30 seconds

20 cycles of
94^0C – 30 seconds
48^0C – 30 seconds
72^0C – 30 seconds

Hold at 72^0C – 10 minutes
Store at 4^0C

3.1.8 Electrophoresis

Gel electrophoresis was used to separate DNA fragments according to size. The nucleic acid on the gel always runs from the cathode (negative pole) to the anode (positive pole), because DNA carries a negative charge due to the phosphodiester backbone. As the DNA migrates through the gel matrix the fragments separate out according to size with the small fragments migrating faster and further than bigger sized amplimers. The more concentrated the gel matrix, and the further the product has to migrate, the finer the resolution of fragment separation becomes.

3.1.8.1 Agarose Gels

Low resolution agarose gels were used to check on the amplification of PCR products. The different sized fragments were separated by electrophoresis. A 1 to 2% agarose gel was created by dissolving agarose powder (Sigma) in 1x TBE (the electrolyte used). A molten solution was created by heating the solution in a microwave until the agarose powder was fully dissolved. Ethidium bromide (Sigma) was then added and mixed with the cooling molten agarose to make a final concentration of 0.5µg/ml. The molten agarose was poured into a sealed

casting plate and a comb was inserted to cast wells within the agarose. The gel was then allowed to cool to set.

When the gel was set, the comb was removed to form the wells and the gel was immersed in the electrolyte (1x TBE). The samples were loaded into the wells with a molecular ladder running along side (such as Hyperladder IV, for PCR products sized between 100 and 1000 bp or Hyperladder I for amplimers that can as long as 10000bp, Bioline). The molecular ladder allows the size of the amplimer to be estimated and also can be used as rough quantification method. The voltage was held at a set value (the higher the voltage the faster the migration) for the length of time needed to the product migrate far enough through the gel to be properly resolved by eye. As the product migrates through the gel the ethidium bromide intercalates with the DNA. This allows the visualisation of the DNA bands under the Ultra Violet light.

3.1.8.2   Polyacrylamide Gels

Polyacrylamide gels were used to resolve DNA fragments to less than one base difference. The gels were prepared using ammonium persulfate (APS, Sigma) to polymerise the acrylamide. 800µl of 10% APS was added to every 100ml of buffered SequaGel XR solution (National Diagnostics). The gel apparatus consisted of two glass plates separated using two 2mm thick spacers. When genotyping was carried out the plates were 25 cm length. If the gel was being set up for a sequencing run, the plates used were 44 cm length. The gel was then poured and allowed to set for two hours. Once the gel was set, the surplus

polyacrylamide was cleaned and a shark tooth comb was inserted to form the wells that allowed sample loading.

The gel was loaded vertically into the Li-COR 4200 DNA sequencer (Li-COR Biosciences). Both on the top and on the bottom of the gel there are two containers that hold 500ml of electrolyte (1x TBE). This allows a current to be applied through the gel for electrophoresis to occur. A scanning laser at the bottom of the gel continually scans for both 700nm and 800nm wavelengths. The signal was then converted into a graphical image to view.

The samples were first mixed with an equal volume of 0.001% fuschin (w/v, fuschin dissolved in formaldehyde) loading buffer (Sigma) and denatured at 95^0C for approximately three minutes, to allow the DNA strands to become single stranded. The gel was preheated to 45^0C before loading the gel, to help prevent the formation of secondary structure and hence allow the linear DNA run in proportion to their length. When running a genotyping reaction the molecular weight marker MicroSTEP20a (Microzone Limited) was used. The molecular weight marker was also mixed with loading buffer like the samples, and its similarly denatured. The size of the molecular weight markers were as follows: 70, 90, 100, 120, 140, 160, 180, 200, 220, 240, 260, 280, 300, 310, 330, 340, 350, 360, 380 and 400 base pairs.

3.1.9 PCR Product Cleaning Method

Purification of the PCR products was done using the microCLEAN system (Microzone Limited). microCLEAN is a DNA clean-up reagent that is used for all DNA including PCR products. It removes the enzymes and reagents from the PCR reaction. The microCLEAN protocol (version 2006) was used.

An equal volume of microCLEAN was added to DNA sample. The samples were mixed by vortexing and left for 5 minutes at room temperature.

If the DNA was in tubes, the samples were spun down at high speed for 5 minutes. The supernatant was removed and the tube was centrifuged briefly to remove drags. The pellet was ressuspended in the volume of ultra pure water (Sigma). The samples were left for at least 5 minutes at room temperature to allow DNA to rehydrate.

When the samples were in 96 well-plates, the mix microCLEAN-DNA was spun at 2000 to 4000 g for 40 minutes. The supernatant was removed by inverting the plate onto tissue paper in the centrifuge holder and centrifuged at low speed (<40 g) for 30 seconds. The pellet was ressuspended in the appropriate volume of ultra pure water (Sigma), and left to rehydrate for 5 minutes.

3.1.10   Sequencing

Sequencing was used for mutation detection and also to confirm clones sequences.

To screen for mutations on IGF1 that could be associated to disease, DNA samples from our UCL case/control sample were selected based on bipolar patients that had inherited alleles and haplotypes that had previously shown association to the disorder. This was carried out using the Big Dye Sequencing Method.

Plasmid DNA sequences originated from cloning were screened by PCR before being sequenced to ensure it contained the expected ligand. These samples were sequenced using both Sanger-Coulson and the Big Dye Sequencing Method.

3.1.10.1 Choice of Samples to Sequence

UCL sample

U.K. National Health Service (NHS) multicentre and Local Research Ethics Committee approval was obtained. Each volunteer was given an information sheet and description of the study. All volunteers signed an approved consent form. The cases were recruited from London NHS psychiatric services and from volunteers who were members of the Manic Depression Fellowship, a support organisation for sufferers of BP. The control subjects were recruited from London branches of the National Blood Service, from local NHS family doctor clinics and from university student/staff volunteers.

It is crucially important for association studies that the case and control sample are derived from the same population. Ancestry criteria for participation in the study had been set out at the very start of the study. Volunteers had to be of British or Irish ancestry. This was defined as having at least three of four grandparents of such ancestry. Volunteers with a grandparent of non-western European ancestry were excluded. All volunteers were screened with an ancestry questionnaire.

All volunteers were interviewed using the SADS-L (Endicott *et al.* 1978). This information was supplemented by material from case-notes where appropriate. A second psychiatrist reviewed the diagnoses. All the bipolar disorder cases were rated with the 90 item OPCRIT checklist (McGuffin *et al.* 1991). Information regarding family history of mental disorder and drug treatment response was also collected. Phenotypic information was collated in a Microsoft Access database.

620 'supernormal' controls were recruited for the study. Control volunteers had no personal history of mental disorder or any family history of bipolar disorder, schizophrenia or alcoholism. It has been argued that it is more efficient to use such controls (Morton *et al.* 1998). A further 480 DNA control samples were added to our UCL sample. These were ECACC Human Random Control (HRC) DNAs from a control population of randomly selected, non-related normal UK Caucasian blood donors. The DNA was extracted from lymphoblastoid cell lines derived by EBV transformation of peripheral blood lymphocytes.

The UCL case/control sample used in this study had a total of 937 cases and 941 controls. The statistical power of the UCL sample was calculated using an internet based software (http://www.dssresearch.com). A Two-Tail Test using percentage values was done. If the allele frequency in cases was 50% and in controls was 57% with a significance of 0.05 the UCL case/control had a statistical power of 86.1%. This means that the sample can detect and replicate 7% allele frequency differences 86.1% of the time.

3.1.10.2 Chain Termination Sequencing Method (Sanger-Coulson)

M13 tailed PCR products for mutation detection or plasmid clones containing M13 primer sequences (Section 3.1.6) were sequenced simultaneously in both directions using IRD 700/800 labelled M13 forward and reverse oligonucleotides. The products were visualised using a Li-COR 4200 DNA Sequencer.

The M13F/R primer during the PCR reaction binds to the specific template and allows extension and synthesis of the complementary strand by the Sequitherm DNA polymerase (Microzone). The advantage of using M13F IRD700 and M13R IRD800 is that both strands can be sequenced at the same time and used as a crosscheck for one another to assess whether novel bands were artefacts or real polymorphisms.

In sequencing, the usual dNTPs are used along with a small amount of dideoxynucleotide (like ddATP) which is incorporated efficiently into a growing polynucleotide strand but prevents further synthesis. This happens due to

dideoxynucleotide lacking a hydroxyl group at the 3 prime position of the sugar component and this prevents further incorporation into the polynucleotide. Therefore chain termination occurs wherever a dideoxynucleotide is incorporated. As a result the reaction can run with each individual ddNTP and will result with terminations at every base of the sequence showing exactly where each base belongs in a sequence.

For the reaction mix, 3µl of the cleaned DNA template was mixed with 6.8µl of the sequencing master mix. 2µl of this was mixed with 2µl of each ddNTP as shown in Figure 3.2. The sequencing reaction was then PCR amplified by the Sequitherm program, which has a longer extension time to allow the full extension prior to incorporation of the chain terminating nucleotide. The cycling conditions are in Table 3.4. This was carried out on a 96 well plate and enables one to sequence 16 individuals at a time on a single polyacrylamide gel. All the reagents Reagents used were from the SequiTherm EXCEL™ II (Microzone Limited) kit with the exception of the M13F 700 and M13R 800 that came from MWG.

Table 3.4 Sequencing cycling conditions for Sanger-Coulson chain termination method.

Sequitherm PCR cycle:

Lid Heated to 105^0C
Products are denatured at 94^0C for 5 minutes

30 cycles of:-
92^0C – 30 seconds
60^0C – 30 seconds
70^0C – 1 minute
End

Store at 4^0C.
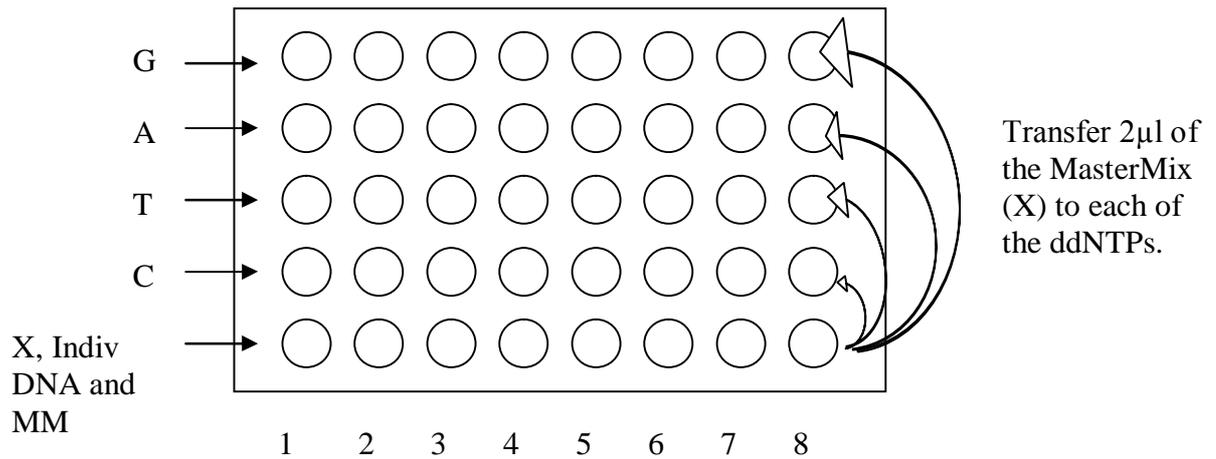
Figure 3.2 Layout of the sequencing reaction, showing how the Master Mix and DNA are mixed with each ddNTP separately.

As mentioned before, after the sequencing PCR an equal volume of Fuchsin loading buffer was added to the PCR products and the samples were ran in a polyacrilamide gel that had 44 cm length and 2 mm thick.

3.1.10.3 Big Dye Terminator Method

The Big Dye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) is based on the Sanger method explained in the previous section. However the dye is not on the M13 forward and reverse primers but on the dideoxynucleotide terminators. This means that each of the four dideoxynucleotide will have a different wavelengths emission. This kit has a special formulation that gives increased robustness, more even peaks and longer read lengths, which will be reflected on better and more accurate results.

As before, the region of interest was amplified with the M13 tailed primers, and cleaned with the MicroCLEAN system (section 3.1.9) and were ressuspended in 5 µL of ultra pure water (Sigma).

The cleaned samples were amplified with the BigDye Terminator sequencing program (Table 3.5). This program was optimized to work with the BigDye Terminator v3.1 master mix (Table 3.6). At this point, two plates with mastermix are prepared, one with the forward primer and a second with the reverse primers.

Table 3.5 BigDye Sequencing PCR cycles.

BigDye Terminator kit PCR cycle:

Lid Heated to 105^0C
Products are denatured at 94^0C for 5 minutes

25 cycles of:
96^0C – 10 seconds
50^0C – 5 seconds
60^0C – 4 minutes
End

Store at 4^0C.

Table 3.6 MasterMix conditions for the Bigdye sequencing PCR.

Reagent	Volume per well (µL)
Terminator Ready Reaction Mix	1
BigDye Sequencing Buffer 5X	1.5
10 µM M13 primer F or R	0.32
Template	1
Water	6.18
Total Volume	10

Once the amplification reaction was completed, the samples were cleaned again. For that an ethanol precipitation was preformed. 2.5 µL of 125 mM EDTA and 30µL of 100% ethanol were added to each well, and the plates were left at room temperature for 10 mins for DNA precipitation. Then, they were spun for 60 min at 3870 rpm. The plates were turned upside down on the centrifuge holder, spun for 1 min at 1000 rpm and left at room temperature for 15 min.

The plates were then covered in a foil seal, they were sent to the Center of Comparative Genomics, in the Department of Biology, UCL. The samples were run on the 3730*xl* DNA Analyzer (ABI). The Applied Biosystems 3730*xl* DNA Analyzer is an automated, high throughput, 96-capillary electrophoresis system used for analyzing fluorescently labeled DNA fragments. First, purified dried samples are ressuspended in injection solution. The injections solution used is formamide. Then the samples plates were placed on the machine. The run cycle is represented on Figure 3.1.
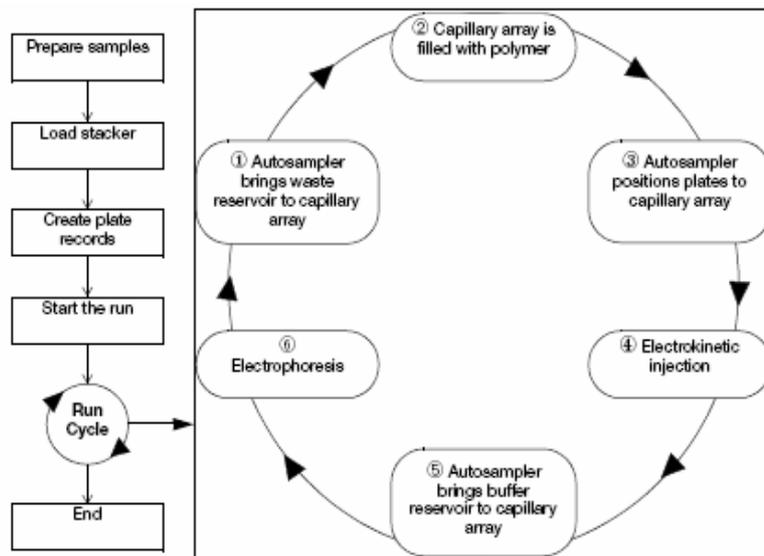
Figure 3.3 Representation of the Sequencing Machine Applied Biosystems 3730xl DNA Analyzer run cycle (Applied Biosystems 3730xl DNA Analyzer Protocol; Edition 2006).

The results were analysed using the Staden Package Pregap4 and Gap4 (Staden 1996). Pregap4 provides a graphical user interface to set up the processing required to prepare trace data for assembly or analysis. It also automates these processes. The possible processes which can be set up and automated include trace format conversion, quality analysis, vector clipping, contaminant screening,

repeat searching and mutation detection. Gap4 performs sequence assembly, contig ordering based on read pair data, contig joining based on sequence comparisons, assembly checking, repeat searching, experiment suggestion, read pair analysis and contig editing. It has graphical views of contigs, templates, readings and traces which all scroll in register. Contig editor searches and experiment suggestion routines use confidence values to calculate the confidence of the consensus sequence and hence identify only places requiring visual trace inspection or extra data. With this software sequences were screened for mutations. If a possible mutation was detected, DNA databases, such as UCSC and NCBI, were used to check if it was a new mutation or an already identified one. In addition, mutations in the promoter regions were examined using the Transcription Element Search System (TESS) software (Schug *et al.* 1997). TESS is a web tool for predicting transcription factor binding sites in DNA sequences. It can identify binding sites using site or consensus strings and positional weight matrices from the TRANSFAC, JASPAR, IMD, and their own CBIL-GibbsMat database. Interesting mutations present in the 32 cases more often than in the 32 controls were genotyped in the UCL case/control sample.

3.1.11   Genotyping

3.1.11.1 <u>KASPar Method</u>

Some SNPs were genotyped using the KBiosciences competitive allele specific PCR system (KASPar). This assay is a modified version of the Amplifluor (Myakishev *et al.* 2001) genotyping method from Millipore. This method involves the allele specific amplification of SNP alleles using two tailed locus

specific oligonucleotides and a standard locus specific reverse primer (Figure 3.4 Figure 3.5). In addition, two tailed oligonucleotides labelled with different fluorescent dyes are used in the reaction. In the early cycles of amplification the allele specific oligonucleotides compete with one another for template. The best fitting (matching) oligonucleotide binds the template with high affinity and thereby creates more of this allele's template for subsequent amplification. In the later cycles the appropriate tailed fluorescent oligonucleotide is incorporated. The end point fluorescence is then detected using a fluorescent plate reader. There are four different possible outcomes each with different fluorescent intensities: low intensity for both fluorescent dyes indicating that no amplification has occurred; high intensity for one fluorescent dye and low intensity for the other indicating a homozygote template for the appropriate SNP allele; the opposite scenario; and finally moderate intensities for both fluorescent dyes indicating a heterozygous template for the SNP.
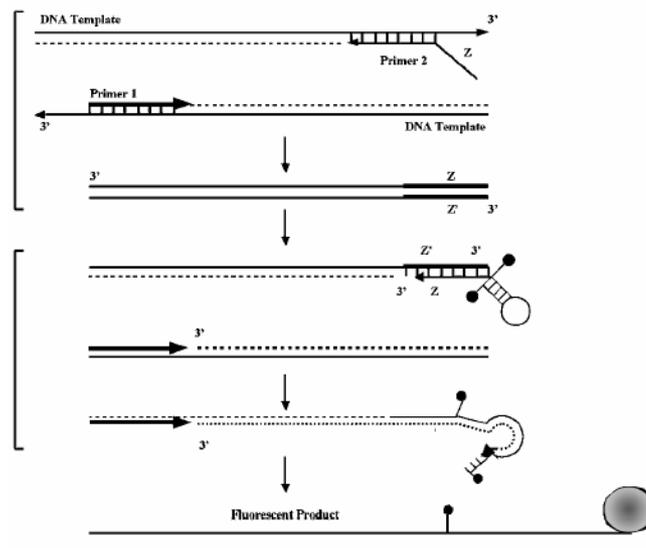
Figure 3.4 KASPar method for genotyping a two allele SNP using allele-specific primers coupled with two Universal Amplifluor primers (Bengra *et al.* 2002).
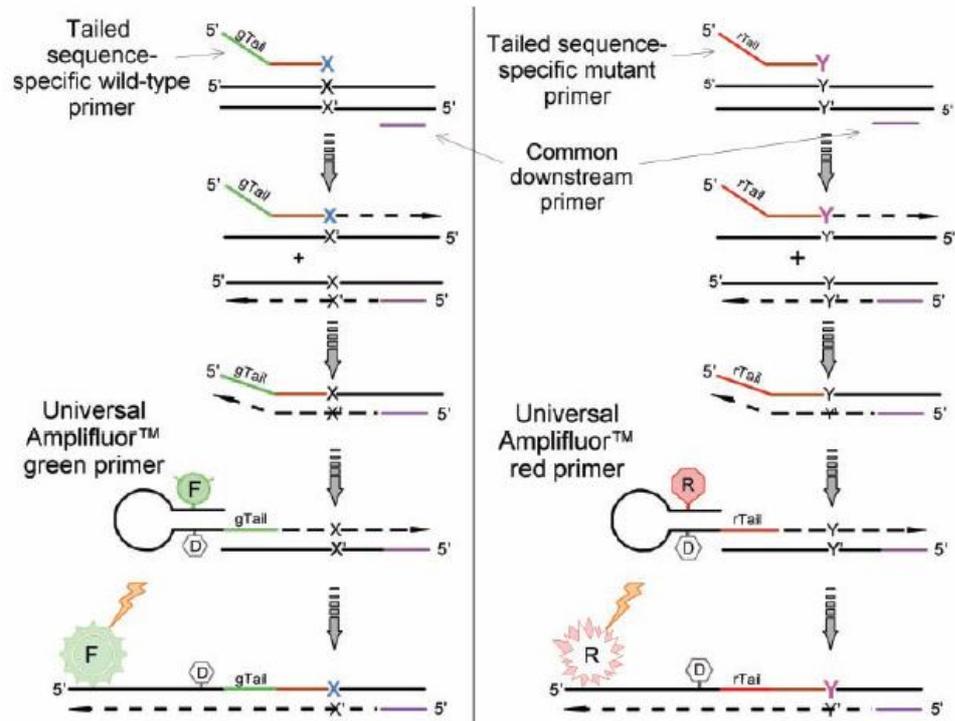
Figure 3.5 An alternative diagrammatic representation of two-allele SNP detection using allele-specific primers coupled with two Universal Amplifluor primer (Bengra *et al.* 2002).

For more common SNPs the assay was sent to KBiosciences for SNP assay development. The SNPs assays were optimized at KBiosciences and were ran in our laboratory in the Light Cycler 480 (Roche). The Light Cycler software allows mutation detection and clusters the different genotypes (Figure 3.6). For more rare SNPs, the primers were designed using the Primer Picker (KBiosciences) and ordered from Sigma Genosys. These assays were optimized in our lab and also run in the Light Cycler480 (Roche).
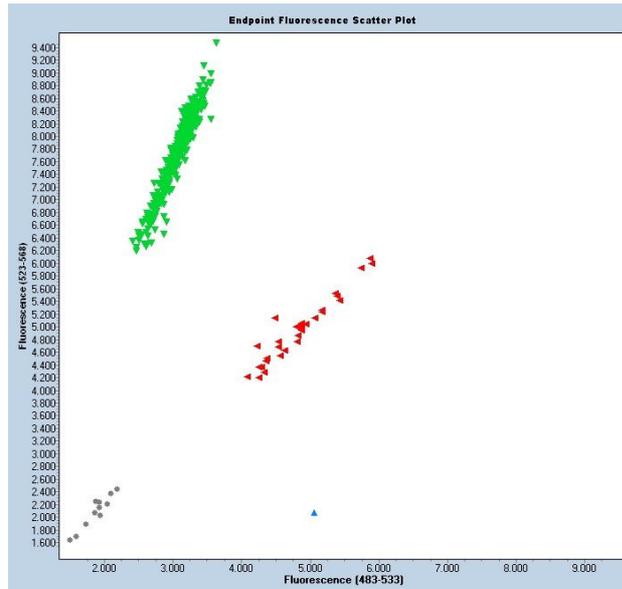
Figure 3.6 Shows of the clusters created by the LightCycler mutation detection software based on the genotyping data of an IGF1 SNP.

3.1.11.2 High Resolution Melting Curve Method

High-resolution melting has proven to be a highly sensitive method for SNPs studies. The PCR is performed in the presence of a double-strand DNA binding dye. Samples are then denatured and the nucleic acid melting is tracked by monitoring the fluorescence of the samples across a defined temperature range, generating melting profiles that can be used to identify the presence of sequence variation within the amplicon. Single point mutations, small insertions, and deletions can all be identified using high-resolution melting.

To genotype the samples using this method the LightCycler 480 High Resolution Melting Master (Roche) was used. The protocol followed is stated on the kit's handbook (February 2008). The samples were run in the LightCycler 480 (Roche) and the data analysed on the LightCycler 480 Gene Scanning software (Roche). The software analysed the different profiles of the melting curves.

Heterozygous and homozygous samples will present different melting profiles (Figure 3.7). The LightCycler 480 software analyses these differences by first normalising the data and then temperature shifting the curves such that differences in the melting temperatures between both homozygous and the heterozygous can be noticed. In the end, the software will group the samples according to the three possible genotypes (Figure 3.8).
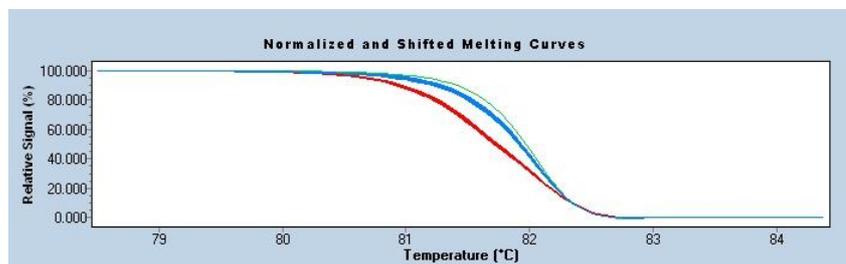
Figure 3.7 Representation of the melting curves profiles of an IGF1 SNP. Result obtained using the HRM method in the LightCycler 480. In green and blue we have the homozygous samples and in red the heterozygous samples.
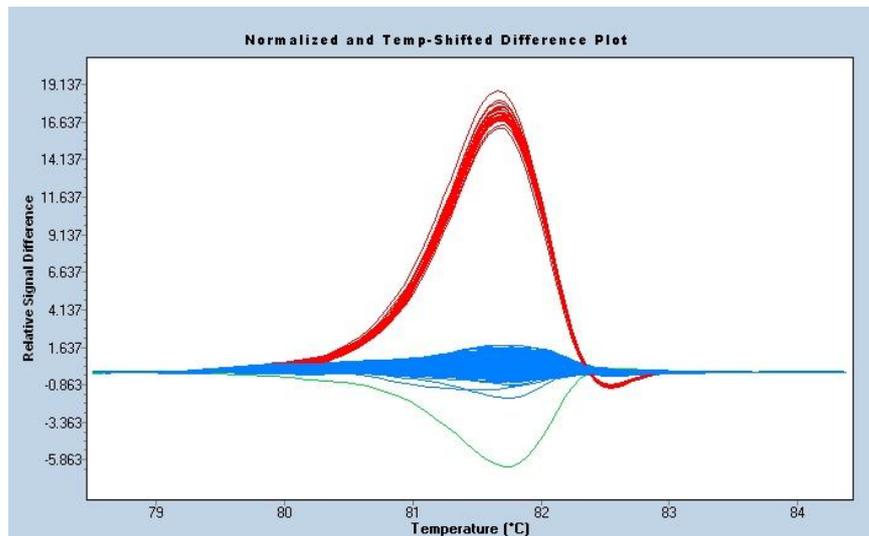
Figure 3.8 Representation of the difference plot of the genotyping data of an IGF1 SNP. The plot groups the samples according to their melting temperature. In green and blue there are represented the homozygous samples and in red the heterozygous samples.

3.1.11.3 <u>Microsatellites</u>

The microsatellite detected by sequencing was amplified in our bipolar sample and matched control sample, along with an M13 forward primer (700 or 800 wavelengths). The products were then pooled and mixed with an equal volume of loading buffer (fuchsin). The sample was loaded on to a vertical polyacrylamide gel. Also, eight molecular weight (MW) markers were run along side the samples. The MW were evenly spaced to enable differentiation between the polymorphic marker alleles and also to check the uniformity of the run to be able to call the alleles correctly and give the correct size. The gel was analysed using a Li-COR 4200 (Li-COR Biosciences), a dual laser scanning laser.

In addition to analysing the entire case control sample, a minimum of 17% of all the samples are repeated in separate microtitre plates as a cross check, to ensure the alleles were called correctly and the common allele size was kept constant between gels to prevent any gel shifts. This helped to improve genotyping accuracy and reliability.

SAGA-GT is a software program (developed by Li-COR) which is used to efficiently genotype gel images produced by the Li-COR. It has the ability to load up the sample identifications, detect the lanes which they are in, and to identify the location of each marker with the aid of the molecular weights. The program detects the uniformity of the gel with aid of the spaced molecular weights and produces a "desmile line" to allow correct calling of allele sizes. Although SAGA-GT is able to automatically call the genotypes, each individual

genotype was checked by eye and corrected if necessary by two independent researchers.

Both the entire sample of cases and controls were genotyped with the aid of SAGA-GT (Figure 3.9). The genotypes were then read by a second individual blind to diagnosis. Any discrepancies between callers resulted in the individual in being re-amplified for the marker and re-genotyped, until the discrepancy was resolved.
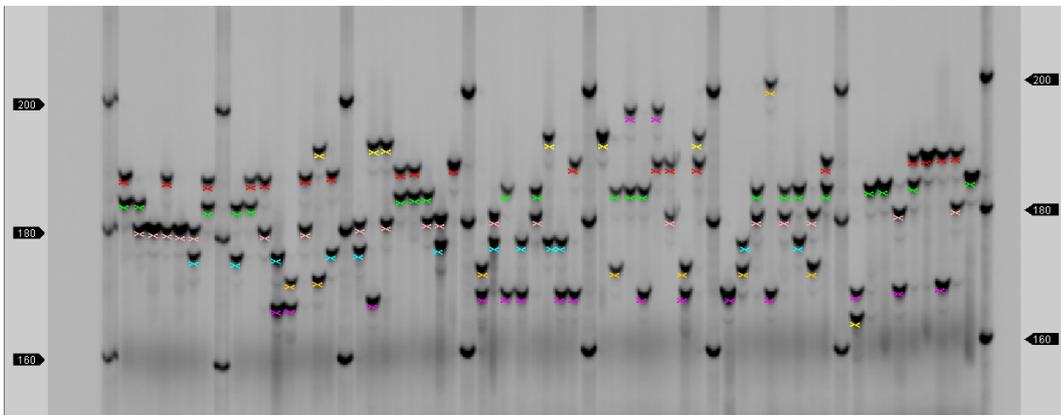
Figure 3.9 Shows a tetranucleotide repeat marker that has been analysed using SAGA-GT software. Allele size calls for all the genotypes are shown using coloured "X" symbols.

The genotyping data was then extracted for each marker from SAGA-GT and entered into a database in Microsoft Access. This database contains a list of all individual identification codes and has a program that checks for conflicts from the data of individuals that have been genotyped more than once. Conflicts were rechecked on SAGA-GT to determine the reason for the conflict and amended or discarded.

3.1.11.4 <u>Data analysis</u>

For all the markers a repeat 17% of samples in each microtitre plate were reduplicated to detect error and confirm the reproducibility of genotypes.

The data was analysed to confirm Hardy-Weinberg equilibrium (HWE). Markers with lack of HWE were rejected and repeated. Genotypic and allelic associations for SNPs were tested using χ^2 tests. Microsatellite data was analyzed for allelic association with bipolar disorder using CLUMP which employs an empirical Monte Carlo test of significance and which does not require correction for multiple alleles (Sham *et al.* 1996). Subtests of the CLUMP programme are: T1 - Pearson's χ^2 statistic of the 'raw' contingency table; T2 - the χ^2 statistic of a table with rare alleles grouped together to prevent small expected cell counts; T3 - the largest of the χ^2 statistics of 2×2 tables each of which compares one allele against the rest grouped together; T4 - the largest of the χ^2 statistics of all possible 2×2 tables comparing any combination of alleles against the rest.

Tests of haplotypic association were carried out using GENECOUNTING, which computes the maximum likelihood estimates of haplotypic frequencies from phase-unknown data, and the empirical significance was obtained using permutation tests (Zhao *et al.* 2000; Zhao *et al.* 2002; Curtis *et al.* 2006). Since running 44 markers on GENECOUNTING for all cases and controls requires a lot of processing power and is time consuming, a preliminary analysis of the haplotypes was done using SCANASSOC. SCANASSOC is part of GENECOUNTING's software. SCANASSOC computes the maximum likelihood estimates of haplotypic frequencies from phase-unknown data similarly to GENECOUNTING, however does not perform the 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. 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 (Broad Institute) (Petryshen *et al.* 2003).

To combine all the markers to implicate association at the level of the whole gene rather than just for individual markers or haplotypes, it was used COMBASSOC. This software combines the p values from all single locus 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).

3.1.12   Cloning

DNA cloning is a process where there is the transfer of a DNA fragment of interest from one organism to a self-replicating genetic element such as a bacterial plasmid. The DNA of interest can then be propagated in a foreign host cell. This technology has been used since the 1970s, and it has become a common practice in molecular biology labs.

Bacterial plasmids are self-replicating extra-chromosomal circular DNA molecules, and are often used as they can generate multiple copies of the same gene. To clone a gene, a DNA fragment containing the gene of interest is isolated

from chromosomal DNA using restriction enzymes, or from PCR reactions. Then the fragment is ligated into a plasmid that has complementary ends to the ones of the fragment and transferred into suitable host cells. In these cells, the plasmid will replicate and will increase the number of copies of the gene of interest.

3.1.12.1 Cloning Process

The regions of interest were amplified by standard PCR and the PCR products were purified by the microCLEAN system. Then were cloned into the pDrive Cloning Vector (Qiagen) using the Qiagen Cloning Kit (Qiagen).
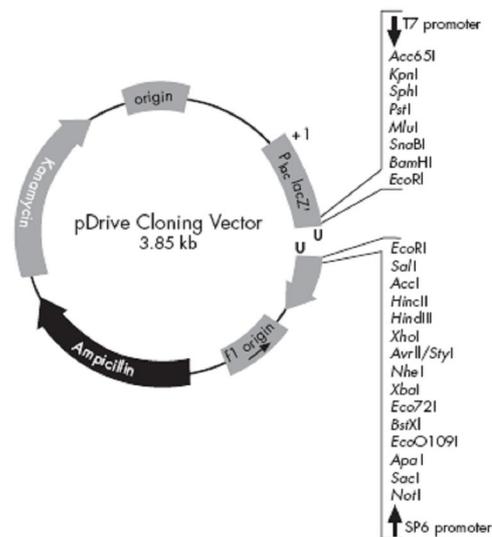
Figure 3.10 Schematic representation of the pDrive Cloning Vector (Qiagen) used in the cloning experiments (Cloning Qiagen Kit Handbook ; April 2001).

The pDrive Cloning Vector (Figure 3.10) contains antibiotic markers (kanamycin and ampicillin). These confer host cell resistance to the antibiotics, if the cells where transfected with the vector. Also the vector has the LacZ α peptide, for α-

complementation, disrupted by the MCS (Multiple Cloning Site). If the ligation of the insert to the MCS was successful, the bacterial colony will be white, if not, the colony will be blue, when grown on the appropriate media. This technique allows for the quick and easy detection of successful ligation. The molecular mechanism for blue white screening is based on the Lac operon, and β-galactosidase protein expression. The chemical required for this screen is X-gal, which functions as indicator, and Isopropyl β-D-1-thiogalactopyranoside (IPTG), which functions as the inducer of the Lac operon. The hydrolysis of colourless X-gal by the β-galactosidase causes the characteristic blue color in the colonies.

The ligation kit works based on the fact that *Taq* and other non-proofreading Polymerases generate a single A overhanging at each end of PCR products. The pDrive Cloning Vector has U overhangs at each end which hybridize to PCR products A overhang with high specificity.

For the ligation, a ligation-reaction mixture with: 1X Ligation Master Mix, 50 ng of pDrive Cloning Vector DNA, 1 to 4 µl of PCR product and distiled water, all in a final volume of 10 µl, was prepared and incubated at 4 to 16°C, for 30 mins. Finished the incubation, we proceed with the tranformation of the host cells. The followed protocol is on the Cloning Qiagen Kit Handbook (April 2001).

3.1.12.2 Transformation

For the transformation reaction was done with MAX Efficiency® DH5α™ Competent Cells. These chemically competent cells contain the φ80d*lac*ZΔM15 marker for α-complementation of the β-galactosidase and are not ampicillin

resistant. These characteristics were essential to get only bacteria transformed with the vector and to select the ones that have the insert.

Very briefly, the competent cells were thawed on wet ice. Then, the ligation mixture was added to 100 µl of cells and incubated for 30 mins on ice. Cells were heat chocked for 45 seconds in a 42°C water bath, and placed on ice for more 2 mins. After, 0.9mL of SOC medium were added to cells at room temperature, and they were incubated for 1 h, at 37°C and 225 rpm. In the end, 25 µL or 100 µL of cells were spread into LB plates containing 100 µg/ml ampicillin (Sigma), 50 µM IPTG (Sigma) and 80 µg/ml X-gal (Sigma). Also 50 pg of pUC19 were used to transform 100 µL of competent cells. This control was essential to determine the transformation efficiency.

The transformation steps were followed as stated on the Max Efficiency® DH5α™ Competent Cells protocol (December 2001).

3.1.12.3 Colony Selection

Once there was growth of the colonies on the plates, the selection was done. Since the insert disrupts the plasmid Lac Z sequence white colonies were expected. The selected colonies were collected with a tip and were dipped into water in a PCR plate and in a falcon tube with 2 mL YT medium with 100 µg/mL of ampicillin. In the PCR plate we checked which clones had the right insert. Meanwhile, the cells in the falcon tube were grown overnight at 37° C, at 200 rpm. From this grown clones the plasmid was extracted and purified.

3.1.12.4 <u>Plasmid Extraction and Purification</u>

For plasmid extraction and purification was used the QIAprep® Spin Miniprep Kit (Qiagen). QIAprep miniprep procedure is based on alkaline lysis of bacterial cells followed by adsorption of DNA onto a silica membrane in the presence of high salts. The four main steps are the preparation and cleaning of the bacterial lysate, the adsorption of DNA onto the silica membrane in the spin column, the washing of the samples and the elution of the DNA. The protocol followed was on the QIAprep® Miniprep Handbook (June 2005).

The yields of purified plasmid were estimated running the samples in a 1% agarose gel with the ladder that corresponds to the size of the DNA, or the samples were quantified using the Nanodrop system.

3.1.13   Northern Blot

The Northern blot is a gel-transfer hybridization technique for detecting and determining the size of specific RNA molecules. The RNA is separated by electrophoresis in an agarose gel under denaturing conditions. Then is transferred to a membrane, crosslinked and hybridized with a specific radioactive or nonisotopic probe. The blot can be probed using labeled DNA, in vitro transcribed RNA and oligonucleotides (Sambrook *et al.* 1989).

3.1.13.1 Formaldehyde/Formamide denaturing agarose gel

The methodology used was based on the methodology described in Maniatis (Sambrook *et al.* 1989) and the Hybond NX leaflet (Amersham Biosciences , 2003).

To separate the RNA by size a 1.6% (W/V) agarose gel with 0.7 M formaldehyde and 1X MOPS was used. Also was included 0.01 µg/mL of ethidium bromide to visualize the RNA after the electrophoresis.

2µg of RNA sample was added to a buffer containing 2.2M formaldehyde, 50% (V/V) formamide and 0.5X MOPS. The samples were denatured at 55°C for 15 min and mixed to 1X loading buffer. As a ladder was used 0.5-10 Kb RNA ladder (Invitrogen). RNA samples were separated at 120 V until the front of the loading buffer was at two thirds of the gel. 1X MOPS buffer was used as electrophoresis buffer. If the work was with total RNA, the RNA in the gel was visualized with UV light to check its integrity. In total RNA was possible to observe two bands corresponding to the 28S and 18S ribosomal RNA. When a smear is absent and the ratio of 28S and 18S RNA is around 2:1 means that the RNA is in good condition.

All of this work was done in the fume hood since formaldehyde vapors are toxic.

3.1.13.2 RNA Transfer to the Nylon Membrane

The transfer was done trough capillary blotting, where buffer is drawn from a reservoir and passes trough the gel into a stack of absorbent paper. The nucleic acid is eluted from the gel by moving stream of buffer and is deposit on a

membrane (Sambrook *et al.* 1989). For this work was used the nylon membrane Hybond-NX (Amersham Biosciences). Nylon membranes have the advantage of standing up well in several rounds of hybridization, however this can also lead to a higher background.

Before the assembling of the transfer apparatus the gel was washed twice with 10x SSC for 15 minutes with gentle agitation. Since formaldehyde vapors are very toxic these washes will remove, or decrease, the levels of formaldehyde in the gel.

To prepare the capillary blotting system, first a tray was half-filled with 10X SSC (transfer buffer). Afterwards it was assembled as shown in Figure 3.11. Air bubbles must be avoided in the capillary system. They were removed by rolling a clean glass rode over the surface. The gel must also be surrounded by Saran Wrap to prevent the transfer buffer being absorbed directly into the paper towels. In the end, a weight was applied to the top of the apparatus to ensure a tight connection between the layers of material used in the transfer system.

After an overnight transfer, when dismantling the apparatus, the membrane was marked to allow identification of the wells, and kept with the sheets of filter paper embedded in transfer buffer.
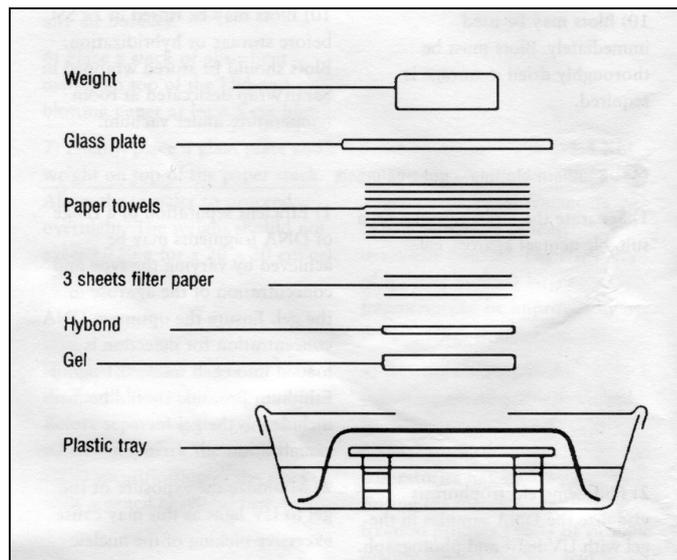
Figure 3.11 Diagrammatic representation of a capillary blotting apparatus (Hybond-NX Amersham Biosciences Handbook).

3.1.13.3 Fixation of RNA to Membrane

After the transfer, the RNA must be fixed to the nylon membrane. That was done with the Stratalinker® UV Crosslinker 2400 (Stratagene). The UV light permits the formation of a covalent bond between the amino group of the nylon membrane and the uracil of RNA. The crosslinking was done at an UV exposure of 70,000 microjoules/cm^2. Blots can be used immediately or stored wrapped in Saran wrap under vacuum at -20°C.

3.1.13.4 Pre-Made Membranes

Commercial Northern blots were also used:

(a) Mouse 12 major tissues polyA$^+$ RNA Northern Blot (Origene): contained RNA samples from brain, heart, kidney, liver, lung, muscle, skin, small intestine, spleen, stomach, testis and thymus;

(b) Mouse Brain Northern Blot II (Panomics): included poly A$^+$ from frontal cortex, back cortex, hippocampus, olfactory bulb, thalamus, striatum, cerebellum and whole brain.

Both blots had 2µg of polyA$^+$ RNA in each lane and were normalized by the mRNA amounts.

3.1.13.5 <u>Riboprobe Synthesis</u>

RNA probes show major advantages when compared with DNA probes. RNA probes are highly sensitive and form more stables complexes with both RNA and DNA samples. Furthermore, they can be synthesized with high efficiency and can be separated from the DNA template just using DNase enzyme. If RNase digestion is needed to remove non specific bounded RNA probes, this won't affect the RNA:RNA or DNA:RNA complexes (Sambrook *et al.* 1989).

To synthesize the RNA probe the Strip-EZ™ RNA T7 Kit (Ambion) was used. This kit allows to synthesis of a probe that is stable under conditions used for hybridization and washing, but which can be cleaved by a degradation buffer. Degradation of the probe happens by cleavage at a modified nucleotide that is incorporated during probe synthesis. This allows probe removal without affecting the Northern blot membrane.

The probe template was a sequence cloned into the pDrive vector. In the vector, the cloned sequence was flanked by two phage RNA polymerases promoter sequences, the T7 and the SP6. The insert orientation was previously checked by

sequencing, which showed that was aligned with the T7 phage RNA promoter. The recombinant plasmid was then linearized with a restriction enzyme downstream of the insert to be transcribed. The restriction enzyme used was Hind III (Promega), overnight at 37°C. Initiation of the transcription is one of the limiting steps of the in vitro transcription reactions. Also RNA polymerases are very processive witch can generate long heterogeneous RNA templates from templates within non linearised vectors. Thus, even a small amount of circular plasmid will decrease the probes synthesis efficiency. Therefore, linearization efficiency was checked by electrophoresis in a 1% agarose gel with 0.5 µg/mL of ethidium bromide. Following linearization, template was cleaned from restriction enzymes and salts. One twentieth volume of 0.5 M EDTA, one tenth volume of 3 M sodium acetate and two volumes of ethanol were added to the restriction reaction mixture, and incubated for 15 minutes at -20°C. Then DNA was pelleted for 15 min in a microcentrifuge at top speed. The supernatant was removed and the pellet was re-spun for a few seconds. The residual fluid was removed and the pellet was diluted in water in a concentration of 1µg/µL.

For RNA probe synthesis was followed the Strip-EZ™ RNA Kit instruction manual (Version 401a). The probe was synthesized using the T7 phage RNA polymerase, which binds its promoter sequence with high affinity. The radiolabeled nucleotides [α-^{32}P]UTP (Amersham) were incorporated in the probe. After transcription reaction, the probe was DNase I treated to remove DNA template. To remove any free nucleotides that might increase the background, probes were cleaned with the Micro Bio-Spin® P-30 Tris Chromatography Columns (Bio Rad). These columns were packed with a special

grade Bio-Gel® P polyacrylamide P-30 gel matrix, which allow non-interactive size separation of particles.

To check if the probes were properly labelled, the levels of radioactivity were checked using a Geiger counter.

3.1.13.6 Hybridisation

If the membrane was previously prepared and stored, or, if it was a commercial membrane, it was hydrated in 100 mL of 4X (w/v) SSC for 10 minutes, shaking. If not, pre-hybridisation would be done immediately.

The hybridisation temperature depends on the type of probe used. For RNA probes, the pre and hybridisation temperature used was 68°C, since the high temperature will decrease background and will prevent non specific binding. The hybridization buffer used was the UltraHyb™ (Ambion). The buffer was always placed in the oven at 68°C, 30 minutes before being used. For pre-hybridisation, the membrane was placed in a hybridisation bottle with 10 mL of hybridisation buffer, at 68°C for 30 minutes to 1 hour.

When the probe was ready, it was denatured by boiling for 5 minutes and placed in ice for 2 minutes. The probe was added to the pre-warmed hybridisation buffer. Then, the previous pre-hybridisation buffer was replaced by the new hybridisation buffer with probe, and incubated overnight at 68°C.

The washing steps were very important since the probes that didn't bind were removed. The washes were done from high stringent to less stringent conditions. Carefully, the hybridisation buffer with the RNA probe was removed from hybridisation bottle. The membrane was washed twice with 2X (w/v) SSC; 0,1% (v/v) SDS and twice with 0.25X (w/v) SSC; 0,1% (v/v) SDS, at 68°C for 15 minutes each. Before exposure to X-ray film, the levels of radioactivity were checked with the Geiger counter. If the levels were higher than 20 counts per minute (cpm), the membrane would be washed again with the second washing solution, until the cpm levels were in the expected range.

3.1.13.7 Exposure and Development of the Film

After the washes the membrane was wrapped in Saran wrap and exposed Kodak BioMax MS X-Ray film (Kodak). All the following steps were done in a dark room to avoid damaging the film. Membrane and film were kept in a cassette film holder at -80°C. This cassette not only prevents light to damage the film but also it has an emulsion that intensifies the exposure when kept at -80°C. The time of exposure could vary depending on how strong was the radioactive signal. Followed by exposure, X-ray film was developed in the Compact X4 Automatic X-ray Film Processor (Xograph).

3.1.13.8 Probe Removal

Northern membranes can be probed more than once. For that, probes that are binding the mRNA need to be removed. The riboprobes used were synthesized with the Strip-EZ™ RNA T7 Kit (Ambion). When the RNA probe was being

transcribed, modified nucleotides were incorporated. These modified nucleotides were cleaved by a special buffer (Degradation buffer). This allows the probe removal without the use of high temperature methods that can affect the RNA quality. The Strip-EZ™ RNA Kit instruction manual (Version 401a) was followed. If after the first round the membrane still radioactive, the procedure was repeated.

3.1.14   Quantitative Real Time PCR (qRT-PCR)

Quantitative polymerase chain reaction (qPCR) is a modification of the PCR used to rapidly measure the quantity of DNA or cDNA present in a sample. One of the most sensitive qPCR methods is the real time PCR (RT-PCR).

The advantage of RT-PCR is the ability to monitor the progress of the PCR as it occurs, precisely measuring the amount of amplicon at each and every cycle throughout the PCR process. For that, there is the detection and quantification of a fluorescent reporter. The signal increases in direct proportion to the amount of PCR product in a reaction (Figure 3.12). By plotting fluorescence against the cycle number an amplification plot is generated that represents the accumulation of product over the duration of the entire PCR reaction. There is a high correlation between the signal and the amount of product present initially.
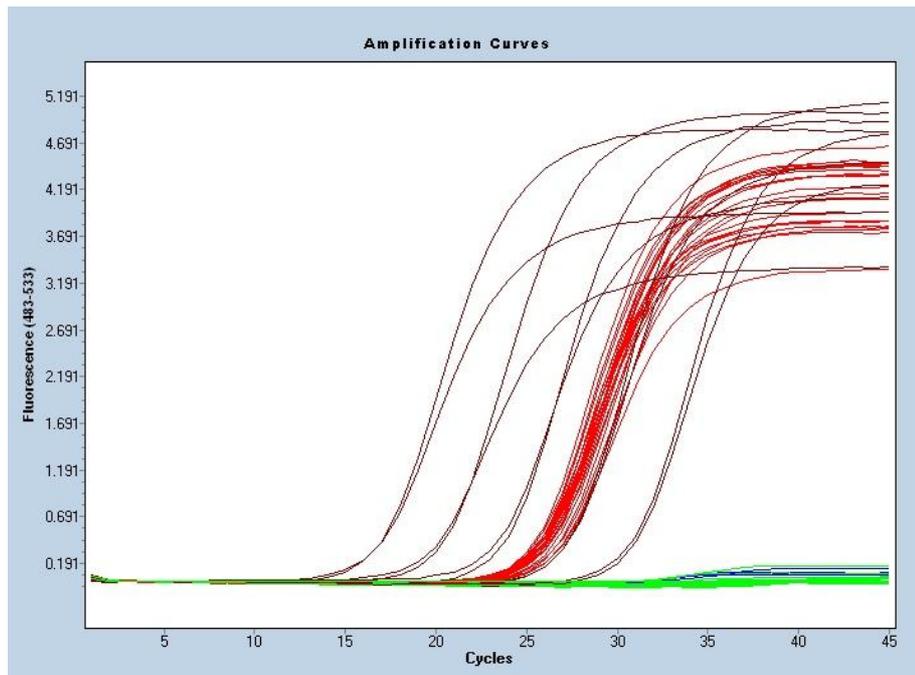
Figure 3.12 Amplification curves of the MAPK1 gene. In red we have the unknown samples, in green the non template controls and in dark red the standard curves.

In the present work the qRT-PCR was carried out with the LightCycler® 480 (Roche).

3.1.14.1 Detection Systems Used

One detection method used was the Universal Probe Library (Roche). This system uses hydrolysis probes (Figure 3.13). The hydrolysis probe relies on the 5'-3' exonuclease activity of *Taq* polymerase, which degrades a hybridized non-extendible DNA probe during the extension step of the PCR. This probe is designed to hybridize to a region within the amplicon and is duel labelled with a reporter dye and a quenching dye. The close proximity of the quencher suppresses the fluorescence of the reporter dye. Once the exonuclease activity of *Taq* polymerase degrades the probe, the fluorescence of the reporter increases at a rate that is proportional to the amount of template present.
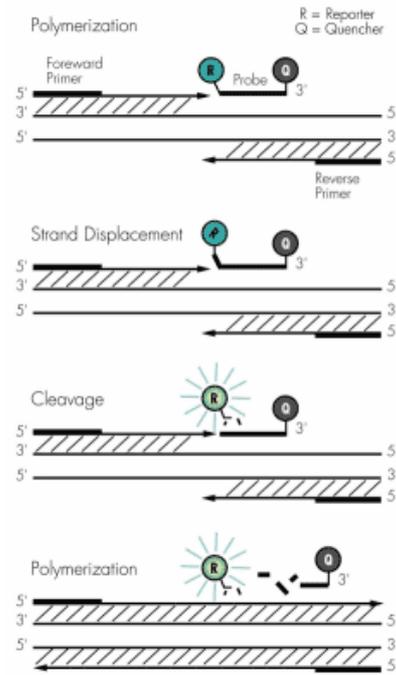
Figure 3.13 Diagrammatic representation of the hydrolysis probes detection method (http://realtime-pcr.blogspot.com).

The other method was the QuantiFast™ SYBR® Green PCR (Qiagen). SYBR Green provides the simplest and most economical format for detecting and quantitating PCR products in real-time reactions. SYBR Green binds double-stranded DNA, and upon excitation emits light (Figure 3.14). Thus, as a PCR product accumulates, fluorescence increases.
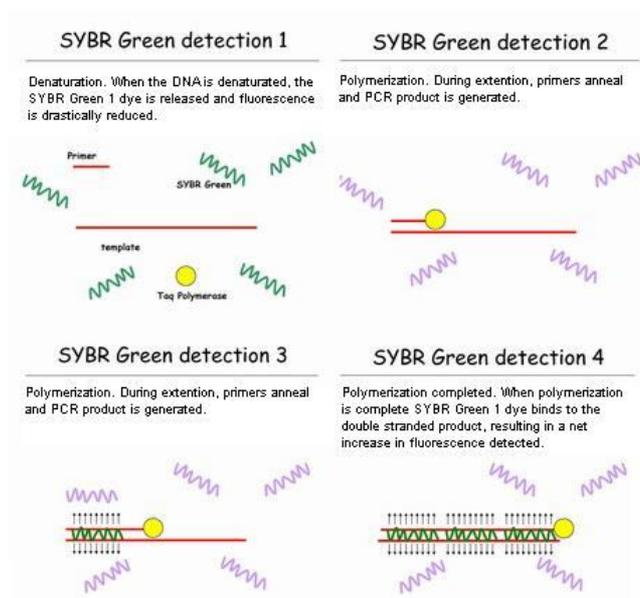
Figure 3.14 Diagrammatic representation of the SYBR Green detection method (http://www.cogentech.it/Pcr/methodology.php).

The advantages of SYBR Green are that it is inexpensive, easy to use, and sensitive. Since the dye binds double-stranded DNA, there is no need to design a probe for any particular target being analyzed. The disadvantage is that SYBR Green will bind to any double-stranded DNA in the reaction, including primer-dimers and other non-specific reaction products, which results in an overestimation of the target concentration. Detection by SYBR Green requires extensive optimization. Since the dye cannot distinguish between specific and non-specific product accumulated during PCR, follow up assays are needed to validate results. For single PCR product reactions with well designed primers, SYBR Green can work extremely well, with spurious non-specific background only showing up in very late cycles.

3.1.14.2 <u>Experimental Design</u>

Primers design is one of the first steps of the experimental design. The PCR primers should be 18 to 24 bp of length. Should not be complementary of each other, especially at the 3' side and should not contain internal secondary structures. Primers have to contain 40% to 60% of G/C and also a balanced distribution of G/C and A/T domains. The melting temperature (Tm) must allow the primers to annealing to the template.

It is also important to choose the detection system. In this work both hydrolysis probes and SYBR green were used. For the latter, QuantiFast™ SYBR® Green PCR kit (Qiagen) was used. The QuantiFast SYBR Green PCR Kit provides rapid real-time quantification of DNA and cDNA targets. The fluorescent dye SYBR Green I in the master mix enables the analysis of many different targets without having to synthesize target-specific labeled probes. High specificity and sensitivity in PCR are achieved by the use of the hot-start enzyme, HotStarTaq *Plus* DNA Polymerase, together with a specialized fast PCR buffer. The buffer also contains ROX dye, which allows fluorescence normalization on certain cyclers. For the used Light Cycler 480 (Roche) this normalization dye is not needed since it has a homogeneous temperature distribution and also a homogeneous illumination and imaging due to a long object-image distance. It also has a large field lens to efficiently collect rays also from lateral wells.

The hydrolysis probes were from the Universal Probe Library (UPL, Roche). The combination of the Roche online assay design software and the 165 pre-validated real-time PCR probes allows quantifying virtually any transcript in the

transcriptome of a large number of organisms. The almost universal coverage of the Universal Probe Library is due to their length of only 8-9 nucleotides. To maintain the specificity, Tm, and assay compatibility that hybridizing qPCR probes require the duplex-stabilizing DNA analogue LNA (Locked Nucleic Acid) was included in the sequence of each probe. The Universal Probe Library probes are fully compatible with the hydrolysis probe real-time PCR detection format. They are labelled 5'-terminal with fluorescein (FAM) and 3'- proximal with a dark quencher dye.

The standard curve is essential for the quantification analysis. Thus, some important considerations need to be taken before starting the qRT-PCR run. To be used as a standard a sequence must be as homologous to the target as possible (same amplicon length; G/C content), to ensure similar amplification efficiency. The standard sequences can be obtained from PCR (linear plasmid DNA; purified PCR products; reference genomic DNA) or from a reverse transcription reaction (*in vitro* transcribed RNA; reference cDNA from total RNA or mRNA). To make a reliable standard curve a minimum of 3 points must be used. However, it is more accurate to have at least 5 points. This curve points correspond to serial dilutions of the standard sequence. One important optimization step is to run a wide range of concentrations that can go from 10^1 to 10^9 copies per μL, in order to make sure that the unknown sample fits into the standard curve.

Not only is the standard curve essential, but the number of replicates is also important. In all assays we used a minimum of 3 replicates of both unknown

samples and all the standard points. The higher the number of replicates used, the more confident we can be of our results. The more similar the replicates are found to be means then the more efficient the reaction was likely to have been. The replicates points will affect the standard curve slope, and that will affect efficiency.

The controls are important in any experimental assay. In qRT-PCR both negative and positive controls must be used. The non template control (NTC) will indicate if the amplification products are specific or due to a contamination or primer dimer. The positive control is a control of the PCR runs and conditions. This will give us confidence on our results when a negative result is obtained.

3.1.14.3 qRT-PCR Normalization

Although real-time RT-PCR is widely used to quantify biologically relevant changes in mRNA levels, there remain a number of problems associated with its use. These include the inherent variability of RNA, variability of extraction protocols that may co-purify inhibitors, and different reverse transcription and PCR efficiencies (Bustin *et al.* 2004). Several strategies have been proposed to normalize real-time RT-PCR data. These range from ensuring that a similar sample size is chosen to using an internal housekeeping or reference gene. These approaches are not mutually exclusive (Huggett *et al.* 2005).

One way of reducing experimental error is by ensuring a similar sample size is obtained, by sampling similar tissue volume or weight. This can be difficult, but will avoid misleading results (Huggett *et al.* 2005).

In the work for this PhD a normalization step using different reference genes was used. For qPCR experiments done in the mouse, the housekeeping genes were: cyclophilin; beta-actin (ACTB); glyceraldehyde-3-phosphate dehydrogenase (GAPD); hypoxanthine-guanine phosphoribosyltransferase (HPRT) and glucuronidase beta (Gus B). For studies done in human cell line samples the house keeping genes chosen were ACTB, HPRT and phosphoglycerate kinase 1 (PGK1).

Normalizing to a reference gene is a simple and popular method for internally controlling for error in real-time RT-PCR. This strategy targets RNAs encoded by genes, which have been collectively called housekeeping genes and benefits from the fact that all the steps required to obtain the final PCR measurement are controlled for. The procedure is simplified as both the gene of interest and the reference gene is measured using real-time RT-PCR. Reference genes can also control for different input RNA amounts used in the reverse transcription step; however, because this can vary with reverse transcriptase type it must be validated (Huggett *et al.* 2005). The most commonly used reference genes include b-actin, glyceraldehyde-3-phosphate dehydrogenase (GAPDH), hypoxanthine-guanine phosphoribosyl transferase (HPRT) and 18S ribosomal RNA. They are historical carry overs and were used as references for many years in Northern blots, RNase protection assays and conventional RT-PCR assays. However, the advent of real-time PCR placed the emphasis on quantitative change, and resulted in a re-evaluation of the use of these "classic" reference genes. Nevertheless, studies continue to use arbitrarily chosen reference genes

for this purpose. There are countless examples of published work that have used a particular reference gene for normalization without any mention of a validation process. More recently, there have been a number of reports that demonstrate that the "classic" reference genes can vary extensively and are unsuitable for normalization purposes due to large measurement error (Dheda *et al.* 2004; Huggett *et al.* 2005).

This occurs when the reference gene is regulated by the experimental conditions. One way of solving this problem is the use of multiple reference genes rather than relying on a single RNA transcript. This is a robust method for providing accurate normalization and is consequently favorable if fine measurements are to be made.

For the normalization analysis there are a number of programs based on the excel platform that allow the assessment of multiple reference genes. One example is geNorm (Vandesompele *et al.* 2002) that allows the most appropriate reference gene to be chosen by using the geometric mean of the expression of the candidate cDNA (http://medgen.ugent.be/~jvdesomp/genorm/). This software was used for the normalization and validation of the qRT-PCR results.

3.1.14.4 Running Programs

When hydrolysis probes were used the master mix was prepared as stated on Table 3.7. This master mix and the 2 µL of sample were added to each well of the 96 well plate.

Table 3.7 Master Mix used for the hydrolysis probes detection method.

Component	Final concentration
Water, PCR grade	-
Universal ProbeLibrary Probe (10 μM)	100nM
Forward Primer (20 μM)	200nM
Reverse Primer (20 μM)	200nM
LightCycler® 480 Probes Master (2x)	1x
Total Volume	**23 μL**

The program for the hydrolysis probes run is described on Table 3.8. The LightCycler must be setup for the detection format, block type and reactions volume. For this specific run, Mono Color Hydrolysis Probe was used as the detection format. The 96 well plate was used for the block type and 25 µL was used as the reaction volume.

Table 3.8 Running program for the qPCR with hydrolysis probes.

Target (°C)	Acquisition Mode	Hold (hh:mm:ss)	Ramp Rate (°C/s)	Acquisitions (per °C)
Pre-Incubation		1 Cycle; Analysis Mode: None		
95	None	00:10:00	4.4	-
Amplification		45 Cycles; Analysis Mode: Quantification		
95	None	00:00:10	4.4	-
60	None	00:00:20	2.0	-
72	Single	00:00:20	4.4	-
Cooling		1 Cycle; Analysis Mode: None		
40	None	00:00:10	2.0	-

For QuantiFast SYBR Green (Qiagen) experiments, a master mix in a total volume of 23µL was prepared and then 2µL of sample added to the well. The master mix was prepared as described in Table 3.9.

Table 3.9 Master Mix conditions for the qPCR run with SYBR green.

Components	Final Concentration
Water, PCR grade	-
QuantiFast SYBR Green PCR Master Mix (2x)	1x
Forward Primer (20 µM)	200 nM
Reverse Primer (20 µM)	200 nM
Final Volume	**23 µL**

The run programme was not that different from the one used for the hydrolysis probes. However, at the end of the SYBR green program, there is the melting curve detection. The program set up is shown in Table 3.10. As before the machine was set up for Mono Color Hydrolysis Probe as the detection format, the 96 well plate as block type and 25 µL as the reaction volume.

Table 3.10 Run programming for the qPCR with the SYBR green method.

Target (°C)	Acquisition Mode	Hold (hh:mm:ss)	Ramp Rate (°C/s)	Acquisitions (per °C)
Pre-Incubation		1 Cycle; Analysis Mode: None		
95	None	00:15:00	4.4	-
Amplification		45 Cycles; Analysis Mode: Quantification		
94	None	00:00:15	2.0	-
55	None	00:00:20	2.0	-
72	Single	00:00:20	2.0	-
Melting Curve		1 Cycle; Analysis Mode: Melting Curves		
95	None	00:00:05	4.4	-
65	None	00:01:00	2.2	-
97	Continuous	-	-	5
Cooling		1 Cycle; Analysis Mode: None		
40	None	00:00:10	1.5	-

3.1.14.5 Interpretation of Results

During the early cycles of the PCR reaction, there is little change in the fluorescence signal. As reaction progresses, fluorescence signal begins to increase each cycle. A fluorescence threshold, that is set to decipher amplification from background, can be set above the baseline in the exponential

portion of the plot. The cycle threshold (C_t) is defined as the cycle number at which the fluorescent signal from each amplification reaction crosses the fixed threshold. The baseline must be set up carefully to allow accurate C_t determination.

Quantification is done by comparing C_t values from unknown samples against its standard curve or against another gene that can serve as an internal control. For that, a dilution series of known concentration, that must fit the dynamic range of the sample, is done to establish a standard curve. The C_t values for these known samples are used to generate the standard curve. This curve is useful not only for quantification, but also contains information about performance of the qPCR reaction. The parameters of the curve provide information about amplification efficiency, replicate consistency, and the theoretical detection limit of the reaction. The detection limit was detected by analysis of the correlation coefficient (r^2), y-intercept and the slope of the standard curve. The r^2 is a measure of the reproducibility, and also indicates how well data fits the standard curve. Ideally on a r^2 of 1 should be obtained. The y-intercept corresponds to the theoretical limit of detection of the reaction, or the C_t value expected if the lowest copy number of target denoted on the X-axis were to be amplified. The RT-PCR is theoretically capable of detecting one copy however a number of 10 is commonly specified as the lowest target level that can be reliably quantified. The slope of the log-linear phase reports the efficiency of the amplification reaction. The efficiency (E) is calculated by:

$E=10^{(-1/Slope)}$

Ideally the efficiency of a PCR should be 100%, meaning the amount of template doubles every cycle, E=2.

The LightCycler® 480 System software has the following analysis modules: melting temperature peak calling, absolute quantification, relative quantification and genotyping analysis. For absolute quantification, unknown samples are quantified based on the standard curve. For relative quantification, PCR is done on both gene of interest and another standard gene (e.g. housekeeping gene). The amount of the gene of interest will be expressed relative to this internal standard. The melting curve is the change in fluorescence observed as a double-stranded DNA dissociating into single-stranded DNA, when the temperature of the reaction increases. The fluorescence is plotted against temperature, and then the change in fluorescence over the change in temperature is plotted against temperature. The melting curve analysis is a way to check RT-PCR reactions for primer-dimmer artefacts, contamination and to ensure reaction specificity. Because the melting temperature of nucleic acids is affected by length, G+C content, and the presence of base mismatches, among other factors, products can often be distinguished by their melting characteristics. The melting curve is only possible on reactions where the fluorophore remains associated with the amplicon, such as the SYBR® Green system.

3.1.15   cDNA Library Screening

A cDNA library refers to a complete, or nearly complete, set of the transcriptome of a cell, type of tissue or organism. In an attempt to find the clone or clones of the human Slynar mouse homologue, adult mouse brain cDNA library was

screened. Also an adult human brain cDNA was screened to find human Slynar and its overlapping full gene clone.

In the work described here the RapidScreen™ Arrayed cDNA Library panels (Origene) were used. The cDNA library was in the form of a 96 well plate. At the beginning, a Master Plate that contains DNA for 500,000 cDNA clones was screened. Each well of the master plate corresponded to a Sub-Plate. The Sub-Plates were 96 well plates with around 50 clones per well. These plates were *Escherichia coli* glycerol stocks, which meant that positive wells could be plated into LB/ampicillin agar plates. The bacterial colonies would grow and would be screened in order to find the right clone. These plates were screened using PCR methodology with specific primers. Figure 3.15 summarises the procedure.

Day 1	Screen Master Plate by 96-well PCR and identify positive wells by electrophoresis
Day 2	Order corresponding Subplate(s)
Day 3	Screen Subplate(s) by 96-well PCR and identify positive wells by electrophoresis
Day 4	Plate bacteria from positive wells of Subplate(s) onto selective media plate
Day 5	Screen colonies by 96-well PCR and prepare plasmid DNA from positive colonies

Figure 3.15   Representation of the cDNA library screening work flow (RapidScreen Arrayed cDNA Library Handbook, Edition 2007).

The mouse cDNA library was screened with the mouse specific set of primers using the touch down PCR cycling and the master mix conditions B (Section 3.1.7). If any positive wells were detected, the *E. coli* from the glycerol stocks were plated into LB agar plates containing 100 µg/ml ampicillin. From the agar plate 96 colonies were picked up and identified. The picked up colonies were amplified in a PCR reaction to screen for the presence of our clone of interest. When the clone was detected, bacterial colonies were grown in 2 mL of LB and 100 µg/mL of ampicillin. The clones are in a pCMV6-XL4 vector (Figure 3.16). This vector has an ampicilin resistance gene and two sequencing priming sites. These sites were used for sequencing, by the Big Dye Terminator v3.1 method (3.1.10.2). A similar procedure was done with the human cDNA library. The ideal PCR condition was B for both primers and the PCR program used was the touch down for SS primers for the human Slynar and touch down with extra cycles for BC-1to5 primers for BC069215.
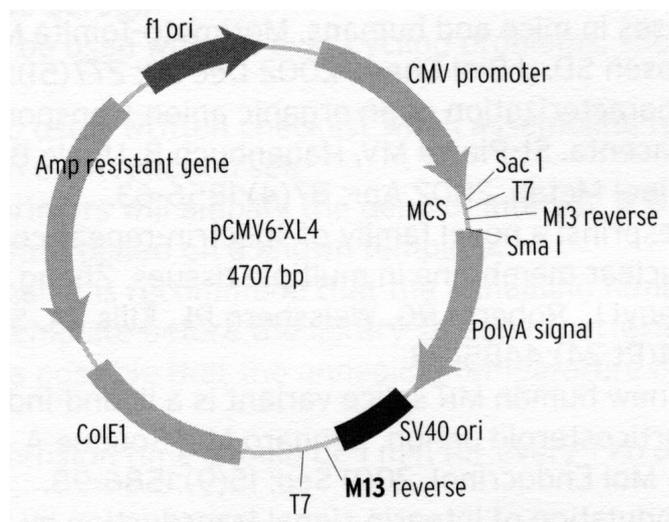
Figure 3.16 Representation of the pCMV6-XL4, the bacterial vector that contains the cDNA sequences (RapidScreen Arrayed cDNA Library Handbook, Edition 2007).

3.1.16　　Rapid Amplification of cDNA Ends (RACE)

3.1.16.1 Principles of RACE

Rapid Amplification of cDNA Ends (RACE) is a procedure for amplification of nucleic acid sequences from a messenger RNA template between a defined internal site and unknown sequences at either the 3' or the 5'-end of the mRNA.

3' RACE takes advantage of the natural poly(A) tail in mRNA as a generic priming site for PCR amplification. In this procedure, mRNAs are converted into cDNA using reverse transcriptase (RT) and an oligo-dT adapter primer. Specific cDNA is then directly amplified by PCR using a gene-specific primer (GSP) that anneals to a region of known exon sequences and an adapter primer that targets the poly(A) tail region. This permits the capture of unknown 3'-mRNA sequences that lie between the exon and the poly(A) tail.

5' RACE, or "anchored" PCR, is a technique that facilitates the isolation and characterization of 5' ends from low-copy messages. First strand cDNA synthesis is primed using a gene-specific antisense oligonucleotide (GSP1). This permits cDNA conversion of specific mRNA, or related families of mRNAs. Following cDNA synthesis, the first strand product is purified from unincorporated dNTPs and GSP1. Then, TdT (Terminal deoxynucleotidyl transferase) is used to add homopolymeric tails to the 3' ends of the cDNA. In the original protocol, tailed cDNA is then amplified by PCR using a mixture of three primers: a nested gene-specific primer (GSP2), which anneals 3' to GSP1; and a combination of a complementary homopolymer-containing anchor primer and corresponding adapter primer which permit amplification from the homopolymeric tail. This

allows amplification of unknown sequences between the GSP2 and the 5'-end of the mRNA.

3.1.16.2 <u>5'RACE</u>

In this work 5' RACE was used to isolate and characterize the 5' prime end of specific transcripts. The 5' RACE System for Rapid Amplification of cDNA Ends (Invitrogen) was used.

The main steps of the 5' RACE System are: (a) synthesis of first strand cDNA, (b) purification of first strand products, (c) homopolymeric tailing, (d) preparation of target cDNA for subsequent amplification by PCR. Control RNA, DNA, and primers are provided for monitoring system performance.

To summarize the procedure (Figure 3.17), first strand cDNA was synthesized from total or poly(A)+ RNA using a gene-specific primer (GSP1) and the reverse transcriptase enzyme SuperScript™ III (Invitrogen). After first strand cDNA synthesis, the original mRNA template was removed by treatment with the RNase Mix (mixture of RNase H, which is specific for RNA:DNA heteroduplex molecules, and RNase T1). Unincorporated dNTPs, GSP1, and proteins were separated from cDNA using a S.N.A.P. Column. A homopolymeric tail was then added to the 3'-end of the cDNA using TdT and dCTP. Since the tailing reaction was performed in a PCR-compatible buffer, the entire contents of the reaction were directly amplified by PCR without intermediate organic extractions, ethanol precipitations, or dilutions. PCR amplification was accomplished using Elongase® Enzyme Mix (Invitrogen), which consists of a mixture of *Taq* and

Pyrococcus species GB-D thermostable DNA polymerases, a nested, gene-specific primer (GSP2) that anneals to a site located within the cDNA molecule, and a novel deoxyinosine-containing anchor primer provided with the RACE kit. Following amplification, 5' RACE products were cloned into the pDrive Cloning Vector from the Qiagen Cloning Kit (Qiagen) and sequenced by the Big Dye Terminator V3.1 method (3.1.10.3).
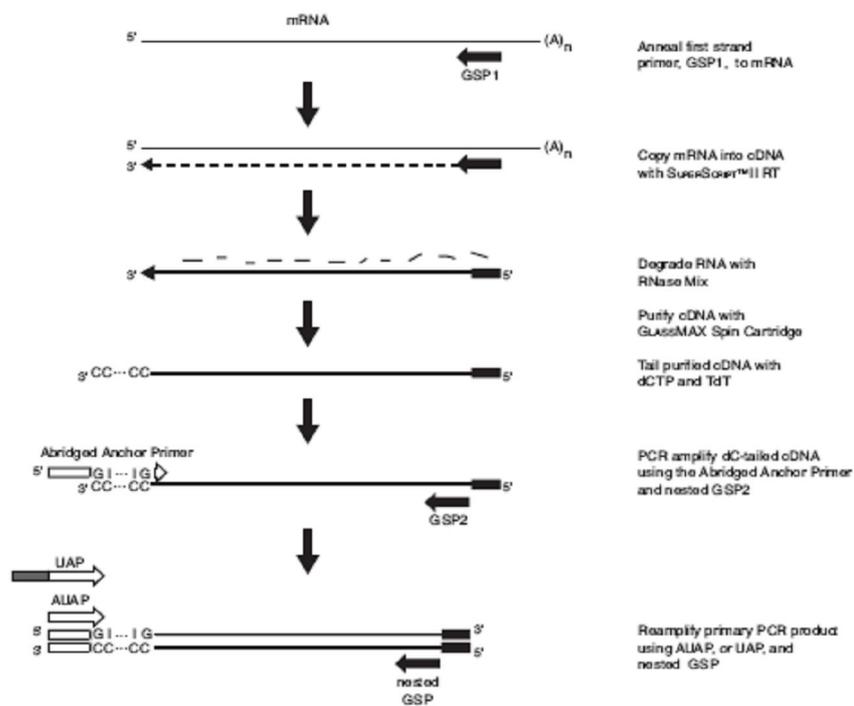
Figure 3.17 Diagrammatic representation of the 5'RACE work flow (5' RACE System for Rapid Amplification of cDNA Ends Handbook; Version 2.0).

The protocol used is stated on the Invitrogen user manual, 5' RACE System for Rapid Amplification of cDNA Ends, Version 2.0 (2004).

3.1.16.3 3' RACE

To run the 3' RACE experiments we used the 3' RACE System for Rapid Amplification of cDNA Ends (Invitrogen).

The procedure is summarized on Figure 3.18.The first step is to anneal a specific adapter primer to the RNA that will bind to the oligo-dT tail. Then the extension of the new DNA strand is done, using the Superscript II Reverse Polymerase (Invitrogen). The RNA template is degraded incubating the DNA:RNA complex with RNase H. The newly synthesized cDNA is PCR amplified using a specific primer designed for our gene and a second primer that binds to the adapter primer sequence used on the reverse transcription step. A second PCR amplification is done with the same primer for the adaptor primer sequence and a second gene specific primer. At this point, the products obtained will be cloned and sequenced.
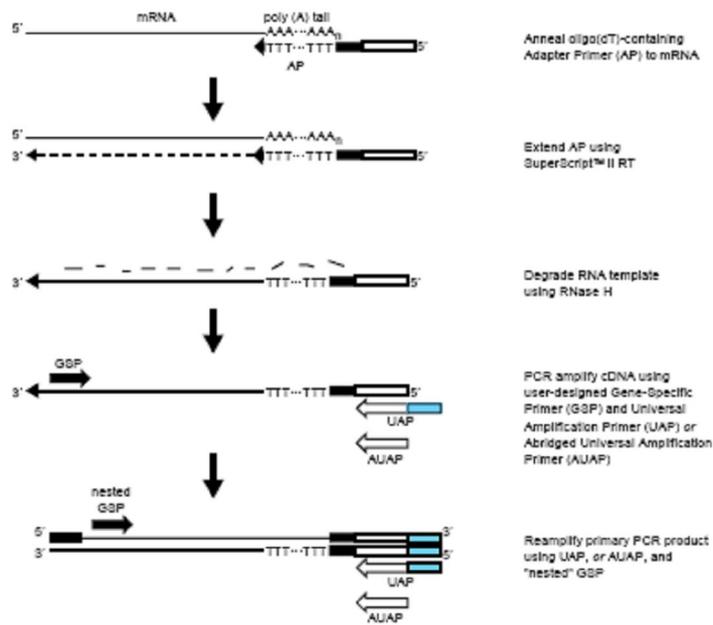
Figure 3.18 Summary of the 3'RACE experiment (3' RACE System for Rapid Amplification of cDNA Ends; Version E).

The protocol used is stated on the Invitrogen user manual, 3' RACE System for Rapid Amplification of cDNA Ends, Version E (2004).

3.1.17 Animal Cell Culture

3.1.17.1 Cell Lines

For this work HeLa and SH-SY5Y neuroblastoma cell lines, both from ECACC, were used.

HeLa cells were the first continuously cultured human malignant cell line, derived from the cervical adenocarcinoma of Henrietta Lacks (one of several pseudonyms). These cells have an epithelial morphology and they have an adherent growth.

SH-SY5Y is a thrice cloned sub line of the neuroblastoma cell line SK-N-SH which was established in 1970 from a metastic bone tumour. The cells came from the brain of a 4 year old girl with a neuroblastoma. These cells grow as a mixture of floating and adherent cells. The cells grow as clusters of neuroblastic cells with multiple, short, fine cell processes (neurites). Cells will aggregate, form clumps and float. SH-SY5Y cells have a 48h doubling time.

3.1.17.2 Growth Conditions

Both cells lines were grown in 75 cm^2 T-Flasks (Nunc) in 15 mL of Dulbecco Modified Eagle's Medium (DMEM, Sigma) with 1% Pencillin-Streptomycin 10x (Sigma) and 10% Fetal Bovine Serum (FBS, Gibco). The cells were kept at 37°C with 5% CO_2.

3.1.17.3 Harvesting of Cells

When cells reached 80 to 90% confluency, they were harvested and subcultured. The medium was removed and cells were rinsed with 10 mL of 1x PBS (Sigma) in cell culture water (Sigma). For SH-SY5Y cell line, the removed medium was centrifuged to recover floating cells. To harvest adherent cells, 5 mL of 1x Trypsin-EDTA Solution (Sigma) was added. They were incubated for a minute at 37°C with 5% CO_2 to let cells to detach. After, 10 mL of fresh medium with serum was added to deactivate the trypsin. Cells were harvested and centrifuged, and cell pellet was ressuspended in 1mL of fresh media. Cells were subcultured in a 1:5 to 1:10 ratio.

3.1.17.4 Cryogenic Storage of Animal Cell Lines

To freeze cells, a freezing medium must be prepared. Freezing medium is complete growth medium supplemented with 5% (v/v) DMSO (Sigma). DMSO is a cryoprotective agent that is used to minimize damage to cells during freezing process. Cells were harvested using the standard protocol that was routinely used for subculturing. However, instead of being resuspended in growth media, cells are ressuspended in freezing media in a final cell concentration of 2 to 5 million viable cells per mL. 1 mL of cell suspension is then added to appropriate cryogenic vials (Nunc). For a good recovery of cells the freezing step must be done at a slow cooling rate. For that, it was used a freezer container, Nalgene®, Mr.Frosty (Sigma). The freezer container is made of polycarbonate, it has a blue high-density polyethylene closure, also a white high-density polyethylene vial holder and foam insert. It provides the critical, repeatable, 1 °C/min cooling rate that is required for successful cryopreservation of cells. For the freezing step, the container was filled with isopropyl alcohol and kept overnight at -80°C. Cells can be stored at -80°C for 6 months or in liquid nitrogen or in its vapor phase up to 2 years.

To recover the culture, the vial must be thawed at room temperature for 1 minute and, after, in water bath at 37°C. To avoid contamination is better to keep the cap and the o-ring away from the water. It is important to thaw cells quickly to avoid cell membrane damage. The content of the vial was centrifuged and cells were ressuspended in 1mL of growth media. Cells were seeded in a 25 mL T-Flask and grown in 7.5 mL of full growth media, at 37°C with 5% CO_2.

3.1.17.5 <u>SH-SY5Y Neuroblastoma Cells Differentiation</u>

Cells were seeded at an initial density of 10^4 cells/ cm^2 in culture dishes. These culture dishes were previously coated with 1 mL/cm^2 of 0.01% Poly-L-Lysine (Sigma). The plates were left to stand for 1 h at 37 °C with the Poly-L-Lysine solution, followed by a PBS washing step.

All trans Retinoic Acid (Sigma) was added the day after plating at a final concentration of 10 µM in DMEM with 15% FBS. The cells were kept under these conditions for 5 days, changing medium every 2 days. After the 5 days, cells were washed 3 times with DMEM and incubated with 50ng/mL of *rh*BDNF (Sigma) for another 5 days. After these time cells were checked on the microscope to check if they acquired neuronal characteristics.

This method was developed based on Encinas *et al* 2000 and Gimenez-Cassina *et al* 2006 publications (Encinas *et al.* 2000; Gimenez-Cassina *et al.* 2006).

3.1.1 siRNA

RNA interference (RNAi) is an evolutionary conserved endogenous biological pathway that has been used as an experimental tool to analyse the function of mammalian genes. A double-stranded RNA (dsRNA) is used to bind to and promote the degradation of target RNAs, resulting in knockdown of the expression of specific genes. RNAi can be induced in mammalian cells by the introduction of synthetic double-stranded small interfering RNAs (siRNAs) 21–23 base pairs (bp) in length or by plasmid and viral vector systems that express

double-stranded short hairpin RNAs (shRNAs) that are subsequently processed to siRNAs by the cellular machinery (Leung *et al.* 2005).

The interfering method chosen was siRNA. These are small RNA molecules around 21 to 23 bp long, that present a 19 bp RNA duplex with a 2 nucleotide overhang on the 3' end. This method has the advantage that can elude the cell's stress response decreasing cellular death and interfering less on cells homeostasis. The disadvantage could be the fact that provides a transient knock down. However, the aim of our studies was to do a simple knock down and analyse its effect on cells gene expression, and not to create a stable knock out cell line.

For the siRNA assays the cells used were neuroblastoma SH-SY5Y cells and the gene to be silenced was Slynar. For this gene four different siRNAs were created to increase the probabilities of knocking down the cells. In addition, MAPK1 siRNA was used as a positive control and a random oligomere, that won't silence any gene, was used to check the effect of the siRNA assay on cells. All the siRNAs were designed by Qiagen. The transfection reagent used was Lipofectamine 2000 (Invitrogen) and the kit's protocol was followed trough out the experiments (11 July 2006 version). To check if the knock down assay was working the expression level of the genes was quantified by qRT-PCR (3.1.14).

3.1.1.1 Day 1: Seeding the cells

On day one, cells were seeded at an initial density of 10^4 cells/ cm^2 in 6 well plate culture dishes. These culture dishes were previously coated with 1 mL/cm^2 of 0.01% Poly-L-Lysine (3.1.17.5).

The cells were grown in a DMEM media supplemented with 10% FBS, but with no antibiotics. Using media with antibiotics during transfection causes cell death. The cells were left overnight at 37°C with 5% CO_2.

3.1.1.2 Day 2: Transfection

The siRNAs are prepared according to the manufacturer's instructions (Qiagen) and were transfected into the cells using Lipofectamine 2000 (Invitrogen). Lipofectamine 2000 is a cationic liposome based reagent that provides high transfection efficiency and high levels of transgene expression in a range of mammalian cell types.

In one tube, 100 pmol of siRNA was diluted in 250 μL of DMEM with no serum or antibiotics. In a second tube, 5 μL of Lipofectamine 2000 was diluted in 250 μL of DMEM, again with no antibiotics or FBS. The two tubes were gently mixed and kept for 5 minutes at room temperature. After the incubation, the siRNA was combined with the Lipofectamine 2000 and the mix was incubated 20 minutes at room temperature. Finally, the full contents of the tube were added to the well with cells and media, and the cells were incubated for 24h at 37°C with 5% CO_2.

3.1.1.3 Day 3: RNA extraction and quantification by qPCR

After the cells were incubated for 24h with the siRNA-Lipofectamine 2000 complexes, the RNA was extracted using the method described on section (3.1.3.2) and cDNA was prepared as stated on section (3.1.5). The newly prepared cDNA was then quantified on the LightCycler 480 (Roche). The MAPK1 was quantified using SYBR Green Assay from Roche and the primers are from Qiagen. Both Slynar and the housekeeping genes, ACTB, PGK1 and HPRT were quantified by hydrolysis probe method, using the Roche UPL system. Data was analysed using the LightCycler 480 absolute quantification software (Roche).

4 Slynar Results

4.1 Slynar Studies in the Mouse

4.1.1 Slynar Mouse Homologue

As it was said in Section 1.7.1, not much is known about Slynar. However Slynar's region has been implicated in bipolar disorder, as it was shown by fine mapping studies at UCL and in Aarhus (Denmark) (Kalsi *et al.* 2006). The aim of this study was to characterize Slynar gene and try to understand its expression and function. Hence, we would like to use the mouse as a model for our studies.

Slynar gene was never identified on the mouse genome. The initial goal was to find a Slynar mouse homologue, to enable the gene's characterization. Analysis of the UCSC genome browser alignment between the human and the mouse genome (Figure 4.1) showed that the mouse genome on chromosome 5, shows a better alignment with the human Slynar gene in exons 2, 3 and 4. This specific chromosome 5 region in the mouse has no genes annotated on the browser. Figure 4.1 shows uniquely the alignment between human and mouse genomes, however does not give homology levels.

Since there was complete alignment between human region of Slynar and the mouse in the region of exon 2, 3 and 4 (Figure 4.1), experiments were developed in order to detect a possible Slynar mouse homologue. The final goal was to carry out expression studies and possibly create a bipolar mouse model of Slynar.

Single line: Smaller gaps mostly due to deletions in the mouse or insertions in the human

Box: Aligning region between the human and the mouse genome

Double line: complex gaps that involve substantial sequence in both species.
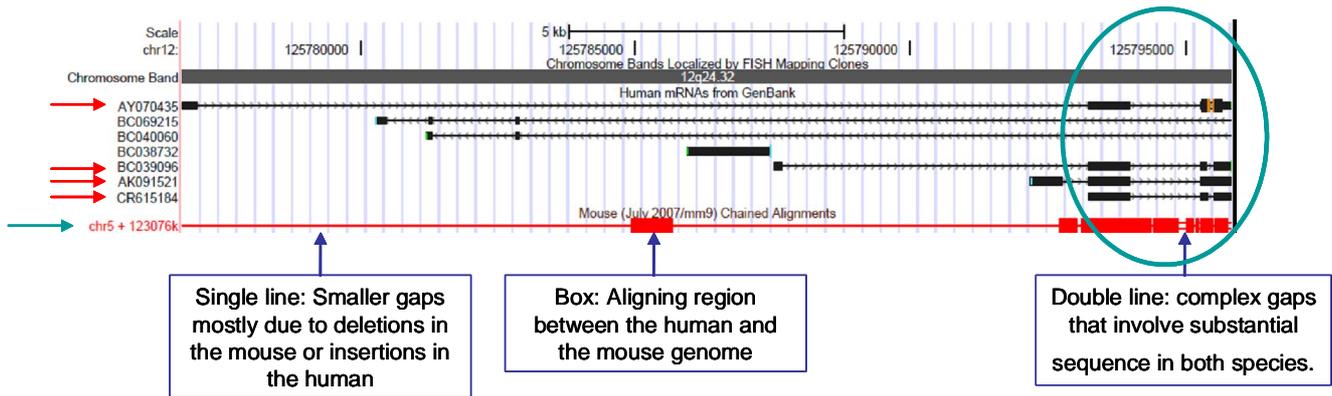
Figure 4.1 Alignment of the human Slynar sequences with the genomic DNA and with the mouse genome (figure taken from UCSC genomic database, assembly March 2006). On the top, the continuous grey bar represents the chromosome band demarking the relative location. Below are represented all the human mRNAs found to be expressed from this region. The sequence is represented by small black boxes and thinner line with arrows. The black boxes correspond to exons and the thinner line represents the introns. The arrows show the direction of the gene's transcription. The red arrows (on the left of the legend) indicate four transcripts correspondent to Slynar and the alternative splicing products: AY070435, BC039096, AK091521 and CR615184. The remaining transcripts do not refer to Slynar, as they are transcribed from the antisense strand. In exon 3 and 4 of AY070435 sequence, the orange bands correspond to a difference in the sequence between the genome and the mRNA found. On the bottom, indicated with a light blue arrow (left of legend) is the alignment of the mouse genome with the human genome. The sequence may contain a box, a single line or a double line. The box represents perfect alignment between the human and the mouse genome. The single line indicates that small gaps may exist between the two sequences. Double lines indicate that complex gaps in long portions of the sequence do exist. The blue circle surrounds the region where the mouse genome better aligns with the human genome. This region covers exon 2, 3 and 4 of Slynar (AY070435).

4.1.2 Finding Slynar Homologue Transcripts

Using the conserved regions between mouse and human as templates, primers for the mouse homologue were designed.   These primers were covering the most conserved region between human Slynar exon 3 and exon 4 (Figure 4.1). Primers sequences can be find in Table 8.1.

To try to amplify a possible transcript from mouse brain RNA, a RT-PCR reaction was set up. The first step was to produce cDNA from mouse brain total RNA. For that, different reactions were done using different primers: reaction 1 –

RNA with a mix of random primers and oligo dT primers; reaction 2 – random primers; reaction 3 – oligo dT primers. The different mixes of primers were used to test which pair would produce the best quality cDNA. Also a negative control was included, where no reverse transcriptase was added, to check for genomic DNA contamination in the RNA sample (Lane 4, 6, 8 from Figure 4.2). The cDNA was then used in a PCR reaction, set up with primers for the conserved region.

The RT-PCR resulted in the amplification of a 220 bp sequence (Figure 4.2). The amplification only occurred on lanes 1, 3, 5 and 7, showing that there was no gDNA contamination. The amplicon has the same size in both cDNA and genomic DNA, which indicates that the intron was not spliced out. Since a product was amplified from cDNA produced with oligo-dT primers, this indicates that it had a poly-A tail.
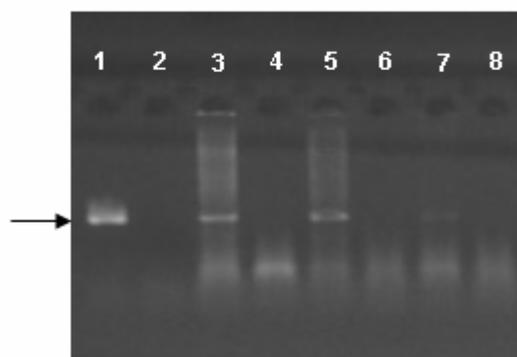
Figure 4.2 Mouse brain cDNA PCR products from the amplification with specific mouse primers for the region between end of exon 3 and the middle of exon 4: (1) gDNA; (2) water; (3) cDNA produced with a mix of random and oligo dT primers; (4) No RT control for random and oligo dT primers produced cDNA; (5) cDNA produced with random primers; (6) control for random primers produced cDNA; (7) cDNA produced with oligo-dT primers; (8) control for oligo-dT primers produced cDNA; arrow, 220 bp PCR product.

cDNA from testis was also amplified using the same specific mouse primers giving rise to an amplicon of the same size as the one in the brain (data not shown). This suggests that the gene is expressed in this tissue. This is not surprising since Slynar's expression has been detected in human testis.

The PCR product amplified with the mouse specific set of primers was cloned as stated in section 3.1.12. A total of 96 clones were screened by PCR, using the same primers for region exon 3 to 4 (Table 8.1), to confirm which of these clones contained the expected sequence. Among the clones that came positive for the presence of the sequence, 16 were plasmid purified and sequenced using the LiCor sequencer. The sequences were submitted for BLAT (UCSC genome browser) and BLAST (NCBI) searches confirming that it was the correct sequence and that was localized to the correct mouse chromosome 5 region. When comparing the clone sequences to the genome sequence they were 98 to 100% homologous. The small differences found between the mouse genomic sequence and the sequences of a small number of clones may be due to errors during the sequencing procedure. The sequence of a representative clone can be found in section 8.1.1.

4.1.3 Slynar's Alternative Transcription

The previous results suggested that transcription was occurring from this specific conserved mouse region. However it was not known if there was alternative transcription happening as for the human Slynar homologue. Additionally, it was also unknown what the size of these transcripts were and where they were expressed. Hence, mouse tissue mRNA blot and a mouse brain tissue blot were

screened using a specific riboprobe consisting of the end of exon 2, an intron and the middle of exon 4 (Section 8.1.5). This probe was produced using the cloned amplicon as a template.
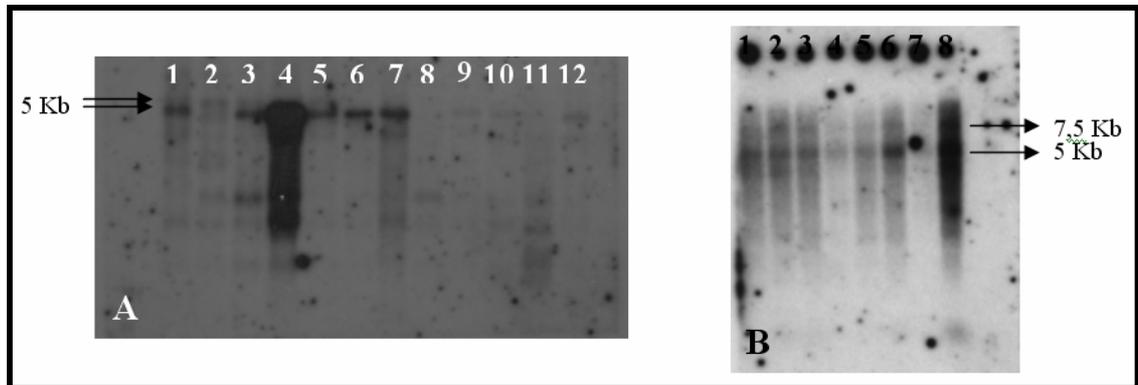
Figure 4.3 Northern blots carried out with mRNA of different mouse tissues. Northern Blot A is a Mouse Tissue mRNA blot: **(1) brain; (2) heart; (3) kidney; (4) liver; (5) lung; (6) muscle; (7) skin; (8) small intestine; (9) spleen; (10) stomach; (11) testis; (12) thymus. Northern blot B is a mouse brain mRNA blot - (1) frontal cortex, (2) back cortex, (3) hippocampus, (4) olfactory bulb, (5) thalamus, (6) striatum, (7) cerebellum, (8) whole brain. Probe labelled with ^{32}P-αUTP.**

In both Northern blots a band of 4.5 to 5 kb was detected. This sequence seemed to be expressed in several tissues (Figure 4.3 A), and more highly expressed in the liver (Figure 4.3 A lane 4). Furthermore, a larger band was detected in the heart and kidney (Figure 4.3 A, lanes 2 and 3). In the mouse brain blot a 7,460 bp band was detected in many different brain tissues (Figure 4.3 B). Some smaller bands such as the ones seen in lanes 3 and 8 can be found (Figure 4.3 A). However, these bands were not considered since they were not present in the brain tissue and could be products of degradation.

These results suggested that this gene was alternatively spliced in the mouse. Moreover, since the transcript was found to be expressed in different tissues, it may indicate that Slynar is a complex multifunctional gene.

Even though these results look promising, the data needs to be carefully analysed. The transcript sizes obtained in the northern blots are in a range of 4.5 to 7.5 Kb long. The predicted transcript size in the human form of Slynar is 1.4 Kb long. This difference can be an indication that the northern blot sequences may be non specific and may not represent a Slynar mouse homologous transcript.

4.1.4 The search for the detected transcripts – getting the cDNA clone

To learn more about Slynar and its function, it was essential to get a full clone representative of the detected transcripts to proceed with functional and expression studies. The first step was to screen an arrayed mouse brain cDNA Library panel (OriGene) using the set of primers used in 4.1.2. First, a master plate (96 well plate) containing 500,000 cDNA clones was screened. Each one of the positive wells corresponds to a sub-plate. This sub-plate is an E-coli glycerol stock and contains 50 clones per well. Positive clones detected by PCR on the sub-plate are then plated on LB/ampicillin agar and the desired clone can be extracted and sequenced.

This master plate and a replica were screened several times and with different primers, however no clones were detected. Since Slynar's expression levels in

the human brain are low we can assume that the same would be expected in the mouse. So, there was a chance of the transcript being missed and not included in the cDNA library. The sequence of all the primers used to screen the cDNA library can be seen in Table 8.2.

To solve this problem a 5' RACE was done. For that, RNA from mouse brain and liver was extracted. From the Northern blot results the liver sample seems to have higher levels of expression than the brain, thus the liver RNA was used to increase the chances of detecting a mouse Slynar clone. A positive and a negative control were also included in the experiment. For the cDNA synthesis step the mouse Slynar specific reverse primer was used (Table 8.3). In this way we would have only specific Slynar cDNA.

After the RACE amplification steps no clone was detected. The RACE was repeated using different brain and liver samples, all freshly extracted. Also the specific primers for Slynar were changed, but Slynar was still undetectable. Once again this could be due to very low levels of Slynar transcription.

4.1.5 Slynar and Lithium

Lithium is the most widely prescribed and effective mood-stabilizing drug used for the treatment of bipolar affective disorder. McQuillin *et al*, in the UCL, studied brain mRNA from 10 mice after treatment with lithium and compared them with 10 untreated controls, to understand how lithium produces changes in the brain. They used Affymetrix MOE430E 2.0 microarrays after 2 weeks of lithium treatment to detect expression changes and found 121 genes with

significant changes (McQuillin *et al.* 2007). Based on these results, and since Slynar seems to be involved in bipolar disorder, research was carried out to check if Slynar transcripts in the mouse were changed by lithium, by qPCR. Pre-existing RNA sample extracted from 10 brains from mice treated with lithium and 10 from control were used.

These experiments were being run in parallel with the experiments from Section 4.1.4, and in accordance to the RACE and cDNA library results, Slynar was not detected by qPCR. Negative and positive controls were included in the assay which showed that the technique was working. The Roche probe sets and primers were changed (Table 8.4), to rule out a specificity problem, but this did not provide any explanation for the results.

4.1.6 Discussion / Conclusion

An unknown transcript, thought to be Slynar's mouse homologue, was detected from mouse brain cDNA. This transcript was again detected in two different northern blots. In fact, the Northern blots seemed to show the existence of alternative transcripts. However, these results failed to replicate on qPCR. Also, all attempts to detect a representative cDNA clone from 5'RACE or cDNA library screening failed.

One possible explanation for the negative results is the fact that Slynar has very low levels of expression in the human. Thus its detection in the human would be difficult and may explain why it is also so difficult to detect in the mouse. If the

Northern blot results were correct, Slynar is a very long transcript in the mouse, which can make it difficult to amplify and detect in techniques such as RACE and cDNA libraries.

On the other hand, there is always the possibility that non-specific products were being detected in the Northern blots and in the RT-PCR experiments. The template used for the riboprobe synthesis, was obtained from the RT-PCR and it was cloned and sequenced to confirm that it was representative of the correct conserved region. Some signs of possible non-specific amplification may have been observed from the northern blot results. The obtained sequences are longer than would be expected, since the largest predicted mRNA sequence in the human in 1.48 Kb long. In addition, when the human and mouse sequence were compared, some conservation was found but the homology was not strong. It may be that a gene is being transcribed in the chromosome 5 mouse region but it may not be a Slynar homologue.

As it was stated in Section 1.6, approximately 99% of mouse genes have human counterparts and mouse orthologs can be identified for 99% of human genes (Tecott 2003). There is a possibility that Slynar belongs to the small group of genes that do not have a mouse counterpart, suggesting that it is a gene specific to higher mammals. This could explain why Slynar could not be detected in the mouse.

This led to the conclusion that cell line and CNS human experiments were likely to be more productive, since all experimental processes in the mouse seemed to be exhausted and no conclusive results were obtained.

4.2 Slynar Studies in the Human

4.2.1 Slynar in the human

Everything that is known about Slynar is from human cell experiments. The alternative splice variants that have been detected were derived from foetal and adult human brain tissue, testis, pooled germ cell tumours and kidney tumours and are listed in the UCSC genome browser. However, very little is understood. The region is not well characterised in terms of transcription, alternative splicing and protein translation and no known function has been attributed to the gene. The aim of the research was to try to clarify Slynar's function, using HeLa and neuroblastoma SH-SY5Y cell lines as models. Understanding the gene's function would help one to build an understanding of how it may cause bipolar disorder and how its dysfunction affects cells and molecular pathways.

4.2.2 Detecting Slynar in Cell Lines

As was mentioned before (Section 1.7.1), Slynar's alternative transcripts were found in different human tissues. Slynar transcription variants that were detected in human mRNAs from GeneBank are listed in the UCSC genome browser (Figure 4.4). Slynar or AY070435 is the predicted gene and CR615184, BC035096 and AK091521 are possible alternative transcripts. To detect the different Slynar variants a pair of primers was designed (Table 8.1). Looking at the alignment of the primer sequences to Slynar sequence (Figure 4.4), it is possible to see that the forward primer is located in the end of exon 3 and the reverse primer is in the end of exon 4. The aim was to detect as many transcripts as possible in the two cell lines models, HeLa and SH-SY5Y.

164

Human Slynar Exon 3 to 4 primers. The primers are the thin vertical lines. The double horizontal lines correspond to the sequence in between the primers
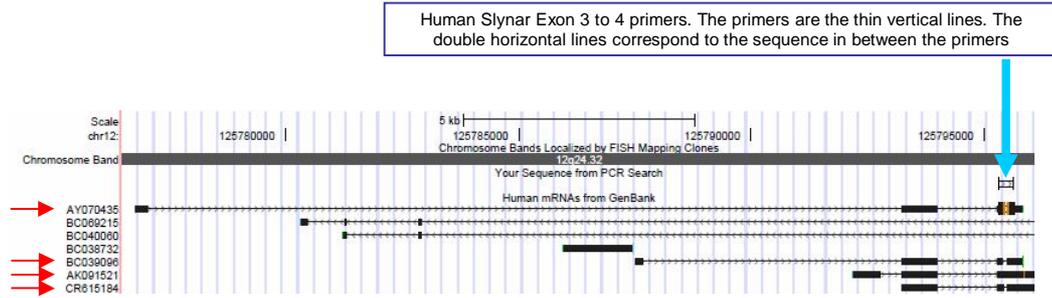
Figure 4.4 Alignment of Slynar (AY070435), its alternative splicing transcripts and the primers Human Slynar Exon 3 to 4 forward and reverse, with the genomic DNA (figure taken from UCSC genomic database, assembly March 2006). On the top, the continuous grey bar represents the chromosome band demarking the relative location. The blue arrow shows the primers used to amplify the region between end of exon 3 and exon 4. The primers are represented by two very thin black vertical lines. The two horizontal lines in between the forward and reverse primer is the sequence amplified by these pair of primers. The region below represents all the human mRNAs known to be expressed. The mRNA sequence is represented by small black boxes and thinner line with arrows. The black boxes correspond to exons and the thinner line represents the introns. The arrows show the direction of the gene's transcription. The red arrows (on the left of the legend) indicate four transcripts correspondent to Slynar and the alternative splicing products: AY070435, BC039096, AK091521 and CR615184. The remaining transcripts do not refer to Slynar, as they are transcribed from the antisense strand. In exon 3 and 4 of AY070435 sequence, the orange bands correspond to a difference in the sequence between the genome and the mRNA found.

For the RT-PCR experiments, RNA was extracted from HeLa and SH-SY5Y neuroblastoma cells. These were run alongside a sample of human brain RNA, acquired from Ambion. cDNA was synthesised. A reaction with and without reverse transcriptase was run to check for genome DNA contamination.
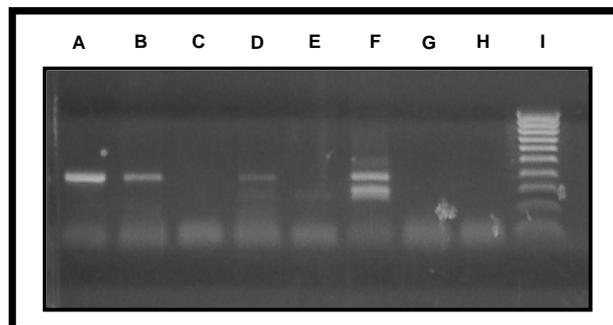
Figure 4.5 Human brain cDNA PCR products from the amplification with specific primers for the region between exon 3 and exon 4: (A) gDNA; (B) SH-SY5Y cDNA; (C) SH-SY5Y cDNA negative control; (D) HeLa cDNA; (E) HeLa cDNA negative control; (F) Human Brain cDNA; (G) Human Brain cDNA negative control; (H) water; (I) Molecular weight marker HyperLadder IV (Bioline)

165

The results of the RT-PCR experiment are shown in Figure 4.5. Three transcripts were detected in human brain RNA (Figure 4.5 F). The transcripts were approximately 400 bp, 300 bp and 250 bp long. Moreover, there was some level of expression in the RNA from HeLa and SH-SY5Y cells (Figure 4.5 B and D). In both cell lines only one transcript, of around 400 bp, was detected. This transcript shares the same size of the genomic band. However, the detection of the bands was not due to genomic DNA contamination because all negative controls are band free (Figure 4.5 C, E, G and H).

The three transcripts obtained from human brain RNA were extracted from gels, cloned and sequenced. The sequences obtained were compared to the human genome sequence with BLAT from the UCSC genome browser. AK091521 seems to be a 400 bp band and is a clone obtained from foetal brain. This clone shares the same size as the genomic band since the intron is not spliced between exon 3 and 4. The 300 bp band seems to represent AY070435, which is the predicted cDNA. Finally, the 250 bp may be identified as BC039096, an IMAGE clone obtained from testis.

These results confirm that Slynar is expressed in the human brain. Furthermore, the presence of one transcript in human cell lines mean that some level of transcription occurs in these cell models. However, given that neuroblastoma cells are better representatives of the human brain cells and have the capacity of being differentiated into mature neurons, these cells were used in the subsequent Slynar experiments.

4.2.3 Human Slynar's Clone

To do more experiments on Slynar and try to find its cellular function it was essential to find a full clone of the gene. For that we ran a 3' and 5' RACE on human brain, non-differentiated SH-SY5Y and differentiated SH-SY5Y RNA. Neuroblastoma cells were differentiated using the retinoic acid and BDNF protocol (Encinas *et al.* 2000; Gimenez-Cassina *et al.* 2006). All sequences detected from RACE were cloned and 8 clones of each RACE sequence were sequenced to confirm that they corresponded to the specific Slynar's transcripts.
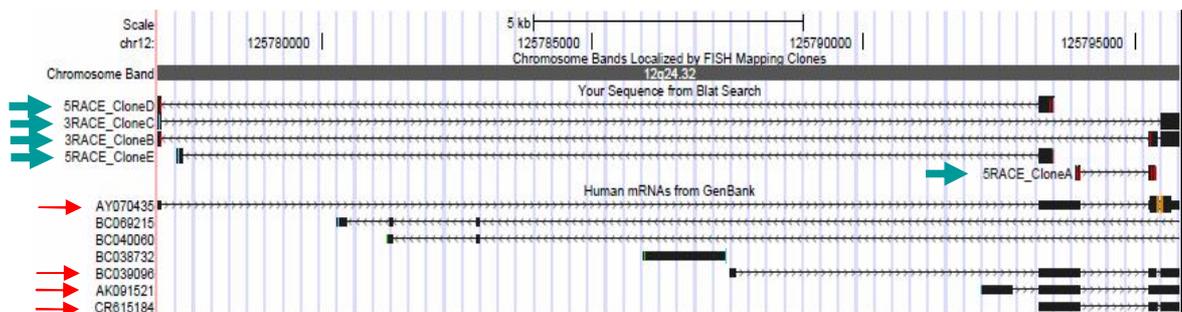
Figure 4.6 Clones obtained from 3' and 5' RACE aligned with the GeneBank mRNA sequences on the UCSC genome browser and the genomic DNA (figure taken from UCSC genomic database, assembly March 2006). On the top, the continuous grey bar represents the chromosome band demarking the relative location. The blue arrows shows the aligned sequences of the clones. The 5' RACE clones, which show the 5' side of Slynar are 5RACE_CloneA, 5RACE_CloneD and 5RACE_CloneE. The 3' RACE clones, that show the 3' end of Slynar, are 3RACE_CloneB and 3RACE_CloneC. The full boxes in the clone sequence are the sequence of the clone that was submitted to BLAT. The thin line represents the genomic sequence that was not represented in the clone, possibily because it is an intronic region. The region below represents all the human mRNAs known to be expressed. The mRNA sequence is represented by small black boxes and thinner line with arrows. The black boxes correspond to exons and the thinner line represents the introns. The arrows show the direction of the gene's transcription. The red arrows (on the left of the legend) indicate four transcripts correspondent to Slynar and the alternative splicing products: AY070435, BC039096, AK091521 and CR615184. The remaining transcripts do not refer to Slynar, as they are transcribed from the antisense strand. In exon 3 and 4 of AY070435 sequence and the clones sequences there are coulored bands than corresponds to differences between the mRNA or clone sequences and the genome. Red means that genome and query sequence have different bases at this position; orange indicates that query sequence has an insertion; blue shows when query sequence extends to beyond the end of alignment.

In all the RACE experiments no Slynar clones were detected from SH-SY5Y whether they were differentiated or non-differentiated. These experiments were repeated at least 3 times and also alternative primers were used to try to detect as many clones as possible. All primers sequences can be found in Table 8.3.

Running 5' and 3' RACE on human brain RNA lead to the detection of three 5'RACE clones: 5RACE_CloneA, 5RACE_CloneD and 5RACE_CloneE, and two 3' RACE clones: 3RACE_CloneB and 3RACE_CloneC (Figure 4.6). The 5'RACE ends are a representation of the beginning of Slynar transcription and the 3'RACE clones indicate where the mRNA stops. However, when these clones were aligned with each other and with the gene's transcripts sequences, against the human genome (Figure 4.6), it was not possible to obtain a full clone from the 5' end to the 3'end. This means that a full transcript of Slynar was not obtained. Even after repeating the RACE experiment with freshly prepared cells and alternative specific primers (sequences of primers in section 8.1.3), there were no improvements. The RACE technique failed to give a full length Slynar clone.

In a further attempt to obtain a full Slynar clone a human brain cDNA library panel (OriGene) was used, using different sets of primers (Table 8.2). Unfortunately, after screening the library several times with more than one set of primers, a Slynar clone was not detected.

Both RACE and cDNA library screening failed to deliver a clone for Slynar. This could be explained by the fact that Slynar was present in low abundance, which

means that it could be hard to detect with these techniques. Also, long sequences are harder to obtain from RACE, and the predicted cDNA sequences on UCSC genome browser (March 2004) ranges from 1300 to 1750 bp. This may explain why it was not possible to obtain a full clone from RACE.

4.2.4 Detection of Slynar by qPCR

After the failure of getting a full clone of Slynar it was decided to run a qPCR experiment to further confirm that Slynar was being expressed in the human brain and SH-SY5Y non differentiated cell line.

For that, primers and a probe were developed for qPCR using the UPL system (Roche). The primers covered the exon 3 to exon 4 region. The primer sequences, probe reference and housekeeping genes used are described in Table 8.4 and Table 8.5. RNA from human brain (Ambion) and RNA from non-differentiated SH-SY5Y cells were used. Then cDNA was synthesised. A RT and no RT control reaction was done to check for genomic DNA contamination. The samples were then cycled in the Light Cycler 480 (Roche).
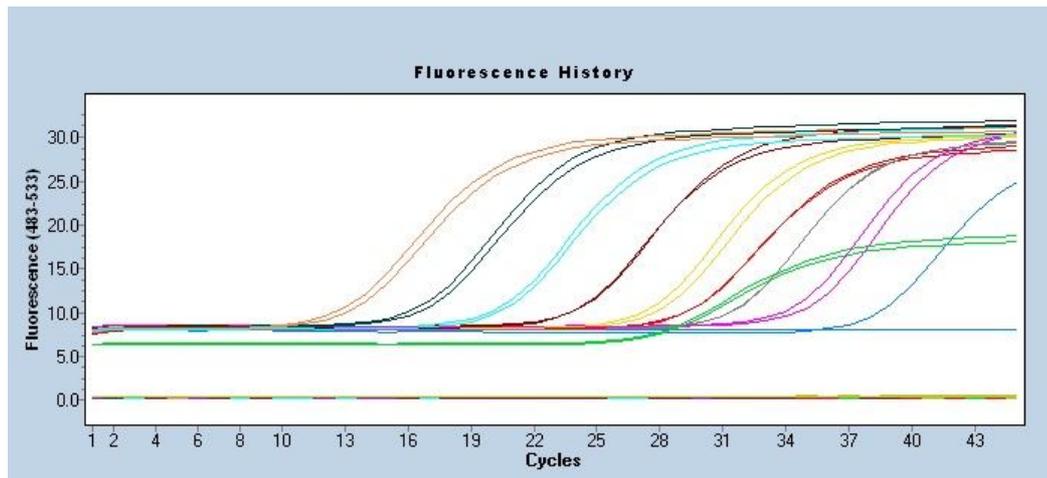
Figure 4.7 qPCR plot. Horizontal lines: correspond to water and no RT control; red colored lines: human brain RNA sample; green coloured lines: SHSY5Y RNA sample; All the other curves: standard curve. Plot acquired from the Roche software.

The plot in Figure 4.7 shows that Slynar was detected by qPCR (Figure 4.7 red and green curves). The qPCR reaction showed once more that Slynar does exist but at a very low level. For reactions with any housekeeping gene, a dilution of 1 in 100 of the cDNA was used. To detect Slynar it was necessary to use 2μL of pure cDNA. As an example, in the experiment exemplified on Figure 4.7, on the same SH-SY5Y RNA sample 2.05 copies/μL of Slynar were detected and 8.58×10^{4} copies/μL of the house keeping gene PGK1.

4.2.5 Silencing Slynar

The fact that we could not obtain a full clone of Slynar made it difficult to do overexpression experiments with the gene in SH-SY5Y neuroblastoma cells. However, since it was known that some transcription of Slynar was detected in SH-SY5Y cells when amplifying the region between exon 3 to 4, it was decided to develop 4 different siRNA oligos for this region. The aim was to knock down

Slynar in neuroblastoma cells and see how it affects gene expression using microarrays. The aim of the microarray experiment was to identify pathways that were being altered by Slynar and elucidate potential functions of the gene in the human brain. It was also going to be carried out imaging experiments on calcium flux changes in SH-SY5Y cells. This experiment was based on the fact that recent GWAS studies have implicated ion channel genes such as CACNA1C in bipolar disorder (Ferreira *et al.* 2008; Sklar *et al.* 2008).
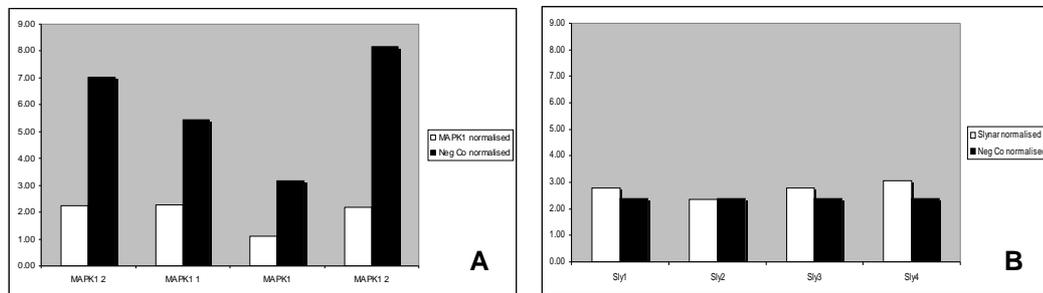
Figure 4.8 Plots representing the siRNA results for MAPK1 positive control and Slynar, when the siRNA concentration used was 50 pmol and the RNA colected 24h post-transfection. (A) MAPK1 silencing results: MAPK1 1 and MAPK 1 2 correspond to two different transfection reagents. 1 is Qiagen reagent and 2 is Lipofectamine2000. The two first MAPK1 1 and MAPK1 2 results were obtained with generation 15 batch of SH-SY5Y cells, and the two second pair of results were obtained using generation 16 batch of cells. (B) Slynar silencing results: Sly 1, Sly 2, Sly 3 and Sly 4 correspond to four different siRNA oligos. For this experiment, all siRNAs were transfected using Lipofectamine2000. The qPCR quantified data was normalized using PGK1 reference gene. White columns: samples knock down, Black columns: samples transfected with a random siRNA that won't silence genes.

Unfortunately, after optimizing the technique and achieving a stable knock down with the positive control siRNA (MAPK1), the same was not achieved for Slynar. Slynar control cells had the same level of expression as the knock down cells and MAPK1 expression was reduced 70 to 80% when using 50 pmol of siRNA and collecting the RNA 24h post-transfection (Figure 4.8). Knock down experiments were attempted at different times (24h, 48h and 72h) with improvements on MAPK1 or Slynar knock down results. In addition different

quantities of siRNA were delivered to cells, ranging from 50 pmol to 250 pmol. The different concentrations of siRNA had no effect on Slynar silencing levels, however 50 pmol showed to be best oligo concentration to reduce MAPK1 expression levels. Finally, two different transfection agents were tested: the siRNA transfection reagent from Qiagen and Lipofectamine 2000 from Invitrogen. The transfection reagent had no effect on Slynar silencing, however it is possible to see that with Lipofectamine2000 (samples MAPK1 2 from Figure 4.8 A) higher levels of silencing were achieved than with the Qiagen reagent (samples MAPK1 1 from Figure 4.8 A) in silencing MAPK1 gene. This indicates that in this experiment Lipofectamin2000 was a more efficient than Qiagen transfection reagent. Looking at MAPK1 results, it is also possible to see that in the two batches of cells there are differences (Figure 4.8 A), being the generation 15 batch more efficiently silenced than the generation 16 batch. The two batches of SH-SY5Y cells have one splitting step of difference to each other. This shows that splitting cells does affect the culture and that ideally replicates must be done from the same batch of cells.

To conclude, after optimizing and consistently silencing the positive control gene, MAPK1, the same was not obtained for Slynar gene. So, no evidence of knock down of Slynar in SH-SY5Y cells was found (Figure 4.8).

4.2.6 Discussion/Conclusion

In the Kalsi *et al* study, Slynar seemed to be the most promising gene in the 300 Kb sequence within 12q24 bipolar disorder susceptibility region (Kalsi *et al.* 2006). To understand the function of the gene at a cellular level and how it might increase risk to bipolar disorder many assays were developed. RT-PCR and q-PCR confirmed that Slynar is expressed in the human brain RNA, with at least 3 different transcripts identified in the UCL (Figure 4.5). Also, a transcript that was not spliced between exon 3 and 4 was identified as being homologous to AK091521 was found in SH-SY5Y cells (Figure 4.5). However it was not possible to obtain a full length clone by RACE and cDNA library screening methods. From RACE we identified five clones, three from 5'RACE and two from 3'RACE (Figure 4.6). However these clones would not align in a way that would deliver a full clone. Hence, it was impossible to overexpress the Slynar gene and investigate its effect in cells. Therefore attempts were made to knock down Slynar in SH-SY5Y cells using the siRNA silencing method. This technique failed to give clear evidence of silencing which meant that was not possible to do further studies on Slynar.

In an attempt to try to predict a possible structure and function for a putative Slynar peptide, the Molecular Modeling Database (MMDB) software from the National Centre for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/) was used. The MMDB contains 3D macromolecular structures, including proteins and polynucleotides. The database includes over 40,000 structures and is linked to the rest of the NCBI databases, including sequences, bibliographic citations, taxonomic classifications, and

sequence and structure neighbours. When submitting the sequences found in human brain RNA and also the RACE clone sequences, no known structure was detected. It is an indication that Slynar possibly does not encode a protein and that it may be a non-protein-coding RNA (npcRNA).

In recent years, npcRNAs have emerged as a major part of the eukaryotic transcriptome and are known to play important roles in cellular metabolism including RNA processing and mRNA stability. Indeed, altered expression of some of these npcRNAs has been associated with cancer, neurodegenerative diseases such as Alzheimer's disease, as well as various types of mental retardation and psychiatric disorders (Xie *et al.* 2008). Its importance in development has also been shown (Amaral *et al.* 2008; Erson *et al.* 2008). Moreover, the progressive maturation and functional plasticity of the nervous system in health and disease seems to involve a dynamic interplay between the transcriptome and the environment. There is increasing evidence that dynamic changes to chromatin, chromosomes and nuclear architecture are regulated by RNA signalling. Although the precise molecular mechanisms are not well understood, they appear to involve the differential recruitment of a hierarchy of generic chromatin modifying complexes and DNA methyltransferases to specific loci by RNAs during differentiation and development. This process seems to implicate npcRNAs and also RNA editing. RNA editing is a primary means by which hardwired genetic information in animals can be altered by environmental signals, especially in the brain, indicating a dynamic RNA-mediated interplay between the transcriptome, the environment and the epigenome. Moreover, RNA-directed regulatory processes may also transfer epigenetic information not

only within cells but also between cells and organ systems, as well as across generations. RNA-based epigenetic mechanisms appear to be essential for neurological development and may be the cause of a spectrum of neurological diseases (Mehler *et al.* 2007; Mattick *et al.* 2009). From a psychiatric disorder point of view, epigenetic mechanisms have been implicated in bipolar disorder. Epigenomic profiling reveals DNA-methylation changes associated with major psychosis (Mill *et al.* 2008). An epigenetic aberration from the normal DNA methylation status of REELIN may confer susceptibility to psychiatric disorders (Tamura *et al.* 2007). In a recent study, genes of the extended dopaminergic (DAergic) system such as membrane-bound catechol-O-methyltransferase (MB-COMT), monoamine oxidase A (MAOA), dopamine transporter 1 (DAT1), tyrosine hydroxylase (TH), dopamine (DA) receptors 1 and 2 (DRD1/2), and related genes (e.g., reelin and BDNF) are cited to illustrate the associations between differential promoter DNA methylations and disease phenotype in schizophrenia and bipolar disorder. This study suggests that epigenetic alterations of the dopaminergic system do exist in major psychiatric disorders (Abdolmaleky *et al.* 2008).

It is possible that there is a long distance effect from the region of chromosome 12 encoding Slynar on a remote gene, which could be the real cause of the positive association with bipolar disorder. It has been found that more than 50% of genes uses tissue-specific promoters and that can be located in set of exons outside the current boundaries of their annotation (Denoeud *et al.* 2007). For some genes the promoters of other neighboring genes are used in specific cells and/or developmental stages. Slynar's region may also contain enhancer and

175

silencer sequences, which could mean that a mutation in this region may affect a gene in a totally different location. Enhancers and silencers are a type of regulatory DNA sequences that may be located 5' or 3' or within an exon or intron of a gene. These regulatory sequences are considered to act via a DNA-loop. The enhancer or silencer and core promoter are brought into close proximity by looping out the intervening DNA (Maston *et al.* 2006). Whereas there are common motifs in core and proximal promoters, enhancers and silencers do not contain many distinctive sequence motifs. Therefore they cannot easily be identified on the basis of their DNA sequence alone (Maston *et al.* 2006; de Vooght *et al.* 2009).

To conclude, we failed to find Slynar's function in this work. However there is a possibility that Slynar is a non protein coding sequence. There is a possibility that Slynar may be a regulatory RNA. It may also be a region rich in regulatory sequences for other genes. In any of these hypotheses aetiological base pair change in the 12q24 Slynar region could ultimately result in the disruption of a gene or a pathway and be indirectly causative of bipolar disorder.

5 IGF1 Results

5.1 IGF1 Sequencing

From the UCL GWAS sample, 32 cases and 32 controls were selected, according to the most representative haplotypes (16 cases that presented rs5742688 allele A and 16 cases presenting the haplotype AAG composed by rs5742688, rs35765 and rs2607983 SNPs).

The chosen samples were sequenced using 29 pairs of primers (Primers sequences in Table 8.6) covering a total of approximately 9000 bp. The primers were designed to cover 1500 bp of the promoter region, the exons, intron/exon junctions and the 3' UTR of IGF1. The sequencing was done using the Big Dye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) and ran on the 3730*xl* DNA Analyzer (ABI). The results were analysed using the Staden Package Pregap4 and Gap4 (Staden 1996).
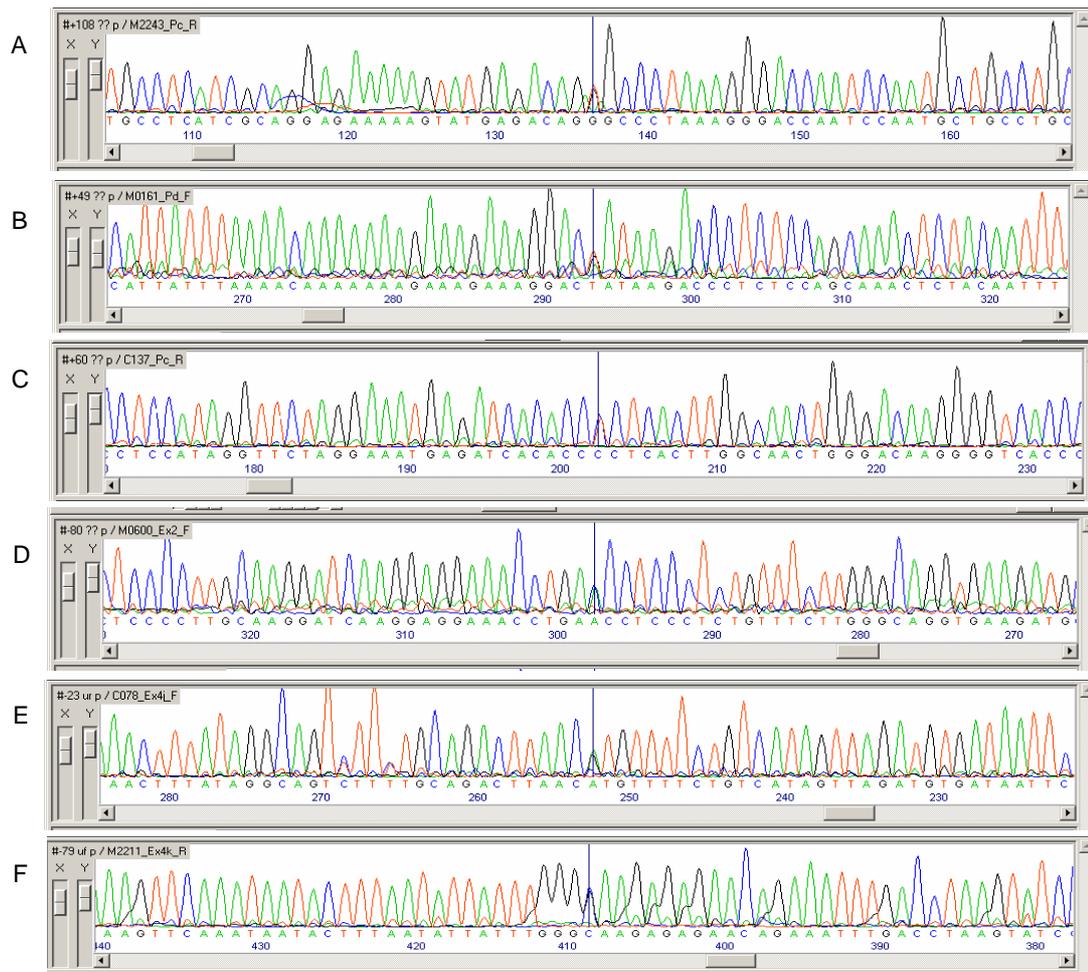
Figure 5.1 IGF1 sequencing traces. In the center of the traces there are the SNPs found in IGF1. (A) Pc, (B) Pd, (C) rs5742612, (D) rs5742620, (E) rs6214, (F) 4k. Figure obtained from the Staden package traces analysis.

Analysing the sequences (Figure 5.1), seven new markers were found: three non database SNPs (4k, Pc and Pd), three database SNPs (rs6214, rs5742612 and rs5742620) and a previously published CA repeat (Rotwein *et al.* 1986). The SNPs Pc, Pd, rs5742612 and the microsatellite are located in the promoter region. The CA repeat is located approximately 750 bp before the start of IGF1. Database SNP rs5742620 is in intron 1, approximately 26 bp from beginning of exon 2. SNPs rs6214 and 4k are located in the 3'untranslated region (UTR).

All the markers found from sequencing, that showed to be more represented in the 32 cases than controls or not present at all in controls, were genotyped in the

full UCL sample of 937 cases and 941 controls. The CA repeat found in the promoter region was genotyped by amplifying the sequence using PCR primers that were directly labelled with a fluorescent probe. The amplified samples were loaded manually and electrophoresed on Li-COR DNA 4200 sequencers (Li-COR Biosciences). The SAGA-GT (Li-COR) program was used to perform the microsatellite analysis. The identified SNPs were genotyped using two different methods. SNPs rs5742620, 4k and rs6214 were genotyped in the UCL lab using the KBiosciences competitive allele specific PCR system (KASPar) and Pd, rs5742612 and Pc were genotyped using High-resolution melting Curve Assay from Roche.

Results of allelic association tests for these markers are shown in Table 5.1 and Table 5.2. The association tests for SNPs were done using a simple Chi^2 test. The CA repeat association test was done using CLUMP. Out of the seven new markers two showed evidence for association (Pd, p=0.006; CA repeat, p=0.0132).

Table 5.1 Tests of association of SNPs in the IGF1 region that were found by sequencing. Association test done using Chi[2]. UCSC March 2004 assembly positions.

Marker	Marker Location	Alleles and Observed Allele Frequencies					χ2	P value
Ex4k		G			C			
Controls	101295526	1786	0.975	46	0.025		0.299	0.583
Cases		1793	0.978	41	0.022			
rs6214		G			A			
Controls	101296036	769	0.414	1089	0.586		0.415	0.519
Cases		785	0.424	1065	0.576			
rs5742620		C			A			
Controls	101372067	51	0.028	1779	0.972		3.724	0.053
Cases		73	0.04	1783	0.96			
Pd		G			T			
Controls	101377223	1805	0.973	51	0.027		7.490	0.006
Cases		1825	0.985	27	0.015			
rs5742612		T			C			
Controls	101377331	1781	0.96	75	0.04		0.753	0.385
Cases		1793	0.965	65	0.035			
Pc		T			G			
Controls	101377396	1785	0.96	75	0.04		0.753	0.385
Cases		1797	0.965	65	0.035			

Table 5.2 CLUMP analysis of the CA repeat. UCSC March 2004 assembly positions. *p*Values from all CLUMP tests in Table 8.8.* *p*Value from CLUMP T3 test.

Marker	Marker Location	Alleles and Observed Allele Frequencies																				χ2	P value
		278		284		298		300		302		304		306		308		310					
CA Repeat																							
Controls	101377593	1	0.001	5	0.003	25	0.014	104	0.060	1150	0.664	325	0.188	83	0.048	37	0.021	2	0.001	9.461	0.0132*
Cases		1	0.001	9	0.005	28	0.016	98	0.056	1113	0.632	352	0.200	128	0.073	31	0.018		0.000		

5.2 IGF1 Haplotypic association

Linkage desiquilibrium (LD) analysis of all the GWAS data and the new markers found by sequencing was done using GENECOUNTING software package (Zhao *et al.* 2000; Zhao *et al.* 2002; Curtis *et al.* 2006). Linkage disequilibrium (LD) between pairs of markers after being calculated was observed using LocusView 2.0 (Petryshen *et al.* 2003) (Figure 5.2)
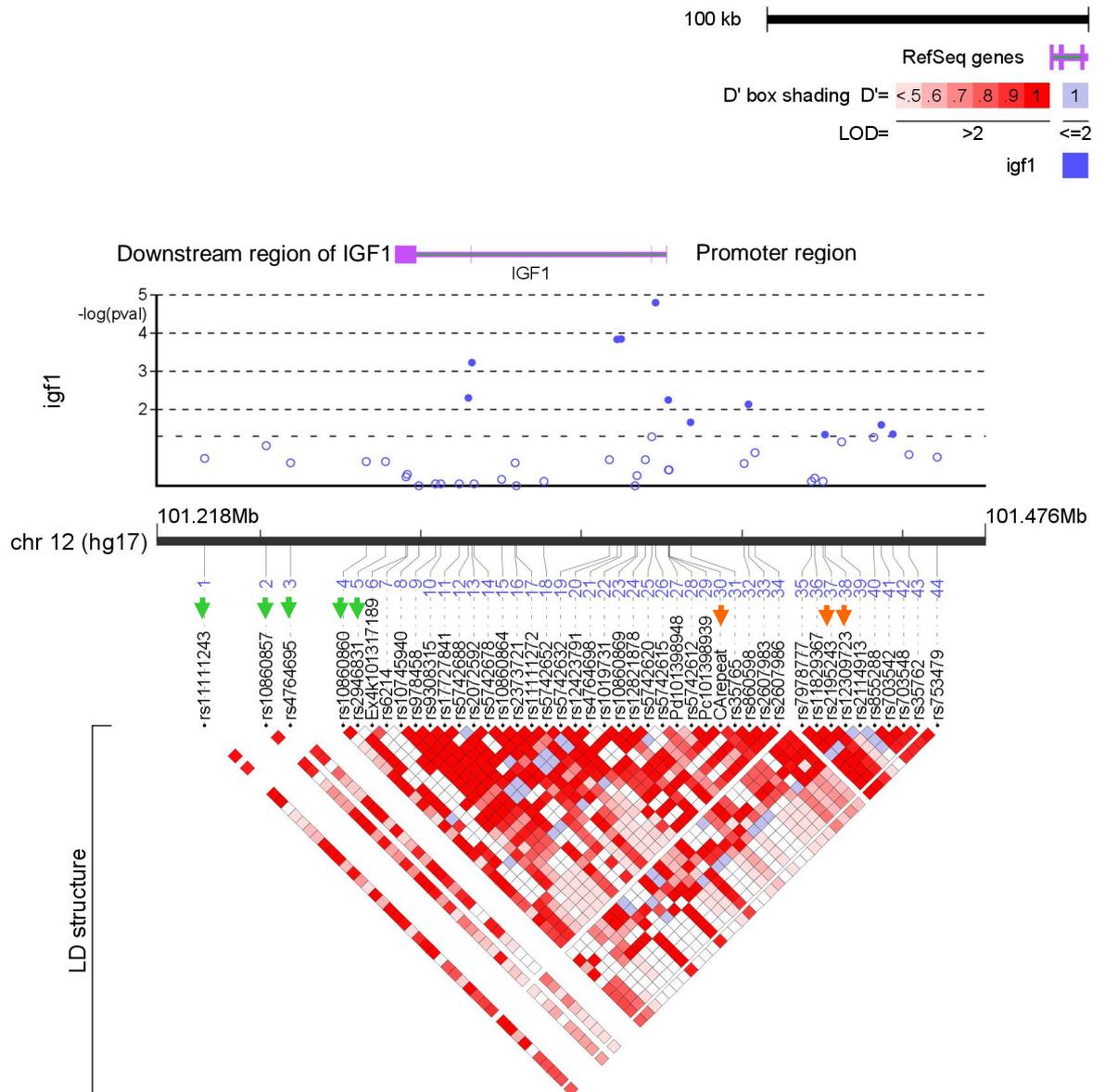
Figure 5.2 LD plot with all markers covering IGF1 region, analysed in the 2000 case/control UCL sample. LD representation done using LocusView 2.0 (Broad Institute). On the top, the purple bar represents IGF1 gene. The vertical bands correspond to the exons and the horizontal band represents the introns. IGF1 is transcribed from the negative strand of the chromosome. Below there is an logarithm plot, plotting IGF1 log values of the 44 markers. The full dots correspond to markers that showed association to bipolar disorder. The empty dots are from markers that did not show association to bipolar disorder. The black band represents the chromosome band demarking the relative location of IGF1. Below, there is a representation of the LD block of the 44 IGF1 markers. Strong red equals a D' of 1. D' <1 is represented with lighter shades of red. The five green arrows are pointing out the SNPs that correspond to HAP A and the three orange arrows are indicating HAP B.

Due to the high number of markers used in the haplotypic analysis of IGF1, a preliminary analysis of the haplotypes was done using SCANASSOC. This

181

software, that is part of GENECOUNTING software package, computes the maximum likelihood estimates of haplotypic frequencies but does not do the permutation testing to calculate the empirical significance. Analysing SCANASSOC's raw data together it was possible to choose the haplotypes that were more represented in cases rather than controls. The chosen haplotypes were inputted into GENECOUNTING to test for the significance of the results.

Haplotypic analysis of the region showed that there are two groups of alleles that show empirical association, as it is possible to analyse in table 5.3 empirical significances, calculated using GENECOUNTING. HAP A includes five markers downstream of the gene (rs11111243, rs10860857, rs4764695, rs10860860 and rs2946831) (Figure 5.2, green arrows indicating markers). These markers, on their own, do not show allelic association as we can see on the dot plot in Figure 5.2. However, the five SNPs together show an empirical haplotypic association of 0.0006 (Table 5.3). In addition, the haplotype comprising the CA repeat, rs2195243 and rs12309723 (located in the promoter region, Figure 5.2, orange arrows indicating markers) gives an empirical pValue of 0.00068 (Table 5.3). Looking at the dot plot from Figure 5.2 is possible to see that the three markers of this haplotype are significant on their own. However, the three markers together have a higher level of significance than on their own. It should be noted that GENECOUNTING computes significances for all possible haplotypes as a global test and does not give significance for individual haplotypes for a given set of markers.

Table 5.3 Haplotypic association with bipolar disorder and IGF1 markers. Results obtained using GENECOUNTING software. The group of alleles of HAP A and HAP B represented in the table were the ones showing a higher presence in cases than controls.

Haplotype name	rs11111243	rs10860857	rs4764695	rs10860860	rs2946831	CA repeat	rs2195243	rs12309723	Controls	Cases	Global Empirical Significance
				Alleles					Frequency		
HAP A	C	C	C	T	A				9.96%	15.85%	0.0006
HAP B						306	G	A	3.07%	6.18%	0.00068018

In addition, a global analysis of the all dataset was done using COMBASSOC (Curtis *et al.* 2008). This indicates the significance of the entire set of markers and haplotypes in the gene and gives a 'gene wide' analysis of association which needs no further correction for multiple markers within the gene. Two data analysis were done: (a) using the 44 markers data; (b) using the 44 markers and the two haplotypes, HAP A and HAP B. COMBASSOC analysis of the entire set of markers gave an empirical significance of 0.009. The joined analyses of the 44 markers, HAP A and HAP B gave an empirical significance of 0.0058.

5.3 Discussion/Conclusion

IGF1 has been implicated in many diseases. There is increasing evidence of IGF1 having a protective effect against metabolic syndrome, endothelial dysfunction, atherosclerotic plaque development or ischemic myocardial damage (Conti *et al.* 2004). There are studies that indicate that high plasma levels of IGF1 may reduce risk of type 2 diabetes and coronary disease (Sandhu 2005). Also, a 2008 review brought together the findings that implicate the IGF1 pathway and IGF1 receptor to the progressive loss of neurogenic capacity of the aging brain and to the pathogenesis of Alzheimer's disease (Puglielli 2008). The data presented here implicates IGF1 in bipolar disorder.

The GWAS study in the UCL sample highlighted 10 SNPs, five within IGF1 and five in the promoter region (Table 1., Figure 5.2). These SNPs showed association with bipolar disorder (Table 1.), the strongest P-value being *P*=0.000037.

Sequencing of IGF1 did not identify any potential aetiological base pair changes. However, new markers were detected and genotyped in an extended UCL sample of 937 cases and 941 controls. Database SNP rs5742620 showed a trend to association (Table 5.2) is located in an intronic region, between exon 1 and 2. Due to the close proximity to exon 2 (around 24 bp from the beginning of exon 2, Figure 5.2), this SNP could change the splicing patterns of this particular intron. To analyse this possibility the sequence of this region, with both SNP alleles, was analysed using the Splice Site Prediction by Neural Network (http://www.fruitfly.org/index.html). The software uses training and tested sets of human and *Drosophila melanogaster* splice sites, for testing splice site predictors other human or Drosophila DNA sequences (Reese *et al.* 1997). The base pair changes due to the presence of the SNP rs5742620 do not seem to affect the predicted splicing site. Nevertheless, there is always the possibility that a change so close to the exon may affect the transcription, for example, by the changing folding structure of the DNA chain in that region decreasing the efficiency of transcription.

The two significantly associated markers found by sequencing were Pd, a non database SNP and a CA repeat. The two markers show association with bipolar disorder, respectively, with significance values of 0.006 and 0.0132 (Table 5.1

and Table 5.3). Both markers are between 750 to 1000 bp from IGF1 (Figure 5.2). It is known that the promoter is a regulatory region of DNA located upstream of a gene and plays an important role in transcriptional regulation. Some 1% of single base-pair substitutions causing human genetic disease occur within gene promoter regions, where they disrupt the normal processes of gene activation and transcriptional initiation and usually decrease or increase the amount of mRNA and thus protein (Cooper 2002). Examples of promoter mutations causing disease include β-thalassemia, Bernard- Soulier syndrome, pyruvate kinase deficiency, familial hypercholesterolemia, and hemophilia (Maston *et al.* 2006). Hence, the importance of the finding of the five GWAS positive markers in IGF1 and the two additional markers found by sequencing. Changes in the promoter region could disrupt expression of IGF1 which could be causative of disease.

The haplotypic analysis of the CA repeat with other two SNPs (Table 5.3) further implicate the promoter region and the microsatellite in bipolar disorder, with a *p*Value of 0.0006. This microsatellite has been published in the past due to its possible involvement in many diseases. It has been shown to be associated with left ventricular hypertrophy (LVH) (Bleumink *et al.* 2005). The absence of the common allele seems to show higher susceptibility to (LVH). Vaessen *et al*, reported that individuals not carrying the commonest allele of the dinucleotide repeat showed lower body weight, lower birth weight and increased relative risk of type 2 diabetes and myocardial infarction (Vaessen *et al.* 2001; Vaessen *et al.* 2002). The CA repeat's common allele has also been implicated in increased levels of fasting serum trygliceride in glucose-tolerant subjects (Nielsen *et al.*

2004). The muscle strength response to strength training was studied in a set of 67 adults and replicated in a bigger sample of 128 individuals. The data suggests that the presence of the IGF1 promoter polymorphism may influence strength response to strength training (Kostek *et al.* 2005; Hand *et al.* 2007). IGF1's promoter dinucleotide has also been linked and associated (p-value = 0.0002) with obesity risk traits in a study of 190 parents and 312 adult offspring (Sun *et al.* 1999). It may be relevant that this microsatellite has been conserved among species, being found in mice, rats, pigs, sheep, cattle and equids (Shimatsu *et al.* 1987; Dickson *et al.* 1991; Kirkpatrick 1992; Caetano *et al.* 1998; Andrade *et al.* 2008).

The COMBASSOC global analysis shows that IGF1 region is associated with bipolar disorder.

The positive results are further confirmed by examination of data from UCL/STEP-BD (Sklar *et al.* 2008), WTCCC GWAS (WTCCC 2007) and the Ferreira data (Ferreira *et al.* 2008). The IGF1 region was covered by SNPs included in the chips and show positive association in all the three samples (Table 5.4). In both WTCCC and Ferreira *et al* samples rs7955640 was the SNP most strongly associated with bipolar disorder (0.0132 and 0.0053, respectively). This SNP is located upstream the IGF1 gene. In the STEP-BD sample, rs7955640 shows no association, but there are three SNPs, all with *p*Values of 0.0301 also upstream IGF1. In the WTCCC sample we can also observe that three SNPs with *p*Values ranging from 0.03 and 0.04 show association downstream of the IGF1 gene.

Table 5.4 Association analysis of IGF1 region in STEP-BD (Ferreira *et al.* 2008), Wellcome Trust (Lopez *et al.* 2007) and Ferreira et al (Ferreira *et al.* 2008) data.

SNPs	Position	WTCCC	STEPBD	Ferreira et al
rs7313075	101154809	>0.05	>0.05	>0.05
rs12426318	101159651	**0.0317**	>0.05	>0.05
rs17439974	101192963	>0.05	>0.05	>0.05
rs10745938	101213386	**0.0362**	>0.05	>0.05
rs4764695	101281243	**0.0426**	>0.05	>0.05
rs5742688	101336656	>0.05	>0.05	>0.05
rs2072592	101337762	>0.05	>0.05	>0.05
rs12423791	101382958	>0.05	>0.05	>0.05
rs4764698	101384190	>0.05	>0.05	>0.05
rs5742615	101394931	>0.05	>0.05	>0.05
rs35765	101405826	>0.05	>0.05	>0.05
rs860598	101422576	>0.05	**0.0309**	>0.05
rs2607983	101423834	>0.05	**0.0310**	>0.05
rs12309723	101447643	>0.05	>0.05	>0.05
rs703542	101465208	>0.05	>0.05	>0.05
rs703548	101468803	>0.05	>0.05	>0.05
rs7955640	101492429	**0.0132**	>0.05	**0.0053**
rs10778179	101495406	>0.05	>0.05	>0.05
rs4609653	101496200	>0.05	>0.05	0.0443
rs1457602	101498898	**0.0468**	>0.05	**0.0082**
rs10745943	101504886	>0.05	>0.05	>0.05
rs2100635	101515117	>0.05	>0.05	**0.0343**
rs7977073	101515601	>0.05	>0.05	**0.0239**
rs17032879	101517463	>0.05	**0.0301**	>0.05

Differences between the results of the different population samples are likely to be due to allelic and locus heterogeneity. The three samples are based on different populations. The STEP-BD (Sklar *et al.* 2008) study uses a population sample from America, the WTCCC (WTCCC 2007) uses individuals that were living within England, Scotland and Wales ('Great Britain') and identified themselves has white Europeans and the UCL sample (Sklar *et al.* 2008) had bipolar disorder subjects who had both parents of English, Irish, Welsh or Scottish descent and three out of four grandparents of the same descent. In addition, there is the possibility that bipolar disorder follows a polygenic model where interaction of modest effects from many genetic variants in each affected individual may be the cause of disease.

In conclusion, the new data seems to suggest the involvement of IGF1 in bipolar disorder.

6 Future Work

6.1 Slynar

Unfortunately, during the time of PhD it was not possible to unravel the truth about Slynar. This project focused on Slynar, however there are other unknown transcripts in the implicated 300 Kb region. This region has been implicated as a susceptibility region for bipolar disorder in two different populations (UCL and Denmark) (Kalsi *et al.* 2006). The next step would be to carry out a project where the surrounding genes are characterized. They could be expressed in the mouse and they too could be studied for expression patterns in different tissues.

The idea that Slynar may be a ncpRNA or even a region rich in critical regulatory sequences for other genes should not be discarded. With the development of new techniques it should become easier to study Slynar and its possible regulatory functions further.

It is also important that more replication studies are done in bipolar disorder studies. It is important to note that the Slynar region is not represented on the chips used for the latest GWAS (WTCCC 2007; Baum *et al.* 2008; Ferreira *et al.* 2008; Sklar *et al.* 2008). So, the lack of results implicating Slynar in the latest GWAS is due to lack coverage and not to a failure of finding association.

Slynar has been previously sequenced by members of our team, however no aetiological base pair changes were found. The gene should now be resequenced, to include the surrounding areas. It is known that regulatory regions can be in

found as far as 14 Kb (Ho *et al.* 2006) and 5' untranslated regions have been shown to contain critical regulatory elements (Zimmermann *et al.* 2005; de Vooght *et al.* 2008). The intronic regions should also not be overlooked. There are examples of mutations in intronic regions that can cause disease. An example is myotonic dystrophy (type 2), which is caused by a CGTG expansion, of approximately 5000 repeats, located in the zinc finger protein 9 (ZNF9) gene (Liquori *et al.* 2001). The three SNP haplotype, located in intron 1 of OCA2 (Oculocutaneous Albinism type 2), is thought to influence eye colour variation (Duffy *et al.* 2007).

6.2 IGF1

Similarly to what was described in Section 6.1 for Slynar, IGF1 association ought to be replicated in further samples to confirm association. IGF1 can also be subjected to more sequencing covering further 5' and 3' non translated areas and introns. It has been discussed on the previous section that the gene surrounding areas and introns may carry important regulatory sequences (Liquori *et al.* 2001; Zimmermann *et al.* 2005; Ho *et al.* 2006; Duffy *et al.* 2007; de Vooght *et al.* 2008). A base pair change in these regions could hold the answer to understanding how IGF1 may cause bipolar disorder.

IGF1 is a very well known gene. However, until now it has never been implicated in bipolar disorder. The next step it to understand how IGF1 may be a risk factor to bipolar disorder. For that, knock out studies with cell lines and animal models should be done. A subsequent expression analysis of other genes implicated in bipolar in the presence and absence of the knock out, could give an

insight in how IGF1 may interact with other genes to cause bipolar disorder. An example of sets of genes to be used in this study could be BDNF and serotonin related genes. BDNF has been implicated in the past on bipolar disorder (Neves-Pereira *et al.* 2002; Sklar *et al.* 2002) and association with the disease has been found on the short allele of serotonin transporter gene (5-HTT) (Collier *et al.* 1996). More recently, studies have shown that the central administration of IGF1 and BDNF in rats lead to a long lasting antidepressant-like effect (Hoshaw *et al.* 2005). The authors discuss that the BDNF and IGF1 antidepressant-like effect may be mediated through serotonergic pathways or receptors. This theory is supported by data indicating that administration of BDNF increases serotonin tissue content and the mRNA of tryptophan hydroxaylase (the limiting step in serotonin production), in the rat brain (Siuciak *et al.* 1996; Siuciak *et al.* 1998). In addition, a more recent study where IGF1 was centrally administered to rats led to the hypothesis that IGF1 initiates a long lasting cascade of neurochemical effects involving increased serotonin levels that results in antidepressant-like behavioural effects (Hoshaw *et al.* 2008). It has been shown that the downstream signalling pathways of IGF1, BDNF and 5-HT demonstrate a high degree of overlap (Mattson *et al.* 2004). Malberg *et al* in a review paper discuss that the functional effects of these three signalling pathways might be a combination of increasing cell proliferation and neurogenic pathways, all with the objective of increasing synaptic strength and plasticity (Malberg *et al.* 2005).

Another interesting study could be to overexpress IGF1 in an animal model. As far as I could find in the literature, attempts have only been made to create a depressive behaviour in animals by blocking IGF1, followed by IGF1

administration to confirm IGF1's antidepressant-like effect. However it would be interesting to see what the effect of having too much IGF1 and to check if this resembled the bipolar phenotype of mania.

Another obvious experiment could be the use of sequences with the positive SNPs that are located on the promoter region and to do an electrophoretic mobility shift assay (EMSA), to observe if they affect the binding of a possible regulatory protein. If it had a real effect, the mutation could be used to create a transgenic mouse and see how the animal behaviour would be affected.

If further replications of association are successful, IGF1 pathways may become a target for new drug treatments of affective disorder. If mutations can be found than predictive testing and studies of treatment outcome can be envisaged. It seems likely that the current phase of genetic research with GWAS is analogous to the days of linkage analysis. At that time the complication was locus heterogeneity and only when many studies were completed could a consensus be reached. For association studies the problem is haplotypic diversity with different disease haplotype frequencies being present even within UK.

7 Bibliography

Abdolmaleky, H. M., C. L. Smith, J. R. Zhou and S. Thiagalingam (2008). "Epigenetic alterations of the dopaminergic system in major psychiatric disorders." Methods Mol Biol **448**: 187-212.

Aberg, M. A., N. D. Aberg, H. Hedbacker, J. Oscarsson and P. S. Eriksson (2000). "Peripheral infusion of IGF-I selectively induces neurogenesis in the adult rat hippocampus." J Neurosci **20**(8): 2896-903.

Aberg, N. D., K. G. Brywe and J. Isgaard (2006). "Aspects of growth hormone and insulin-like growth factor-I related to neuroprotection, regeneration, and functional plasticity in the adult brain." ScientificWorldJournal **6**: 53-80.

Abkevich, V., N. J. Camp, C. H. Hensel, C. D. Neff, D. L. Russell, D. C. Hughes, A. M. Plenk, M. R. Lowry, R. L. Richards, C. Carter, G. C. Frech, S. Stone, K. Rowe, C. A. Chau, K. Cortado, A. Hunt, K. Luce, G. O'Neil, J. Poarch, J. Potter, G. H. Poulsen, H. Saxton, M. Bernat-Sestak, V. Thompson, A. Gutin, M. H. Skolnick, D. Shattuck and L. Cannon-Albright (2003). "Predisposition locus for major depression at chromosome 12q22-12q23.2." Am J Hum Genet **73**(6): 1271-81.

Akiskal, H. S., M. L. Bourgeois, J. Angst, R. Post, H. Moller and R. Hirschfeld (2000). "Re-evaluating the prevalence of and diagnostic composition

within the broad clinical spectrum of bipolar disorders." <u>J Affect Disord</u> **59 Suppl 1**: S5-S30.

Allen, M. G. (1976). "Twin studies of affective illness." <u>Arch Gen Psychiatry</u> **33**(12): 1476-8.

Allison, J. H. and M. A. Stewart (1971). "Reduced brain inositol in lithium-treated rats." <u>Nat New Biol</u> **233**(43): 267-8.

Althoff, R. R., S. V. Faraone, D. C. Rettew, C. P. Morley and J. J. Hudziak (2005). "Family, twin, adoption, and molecular genetic studies of juvenile bipolar disorder." <u>Bipolar Disord</u> **7**(6): 598-609.

Amaral, P. P. and J. S. Mattick (2008). "Noncoding RNA in development." <u>Mamm Genome</u> **19**(7-8): 454-92.

Andrade, P. C., D. A. Grossi, C. C. Paz, M. M. Alencar, L. C. Regitano and D. P. Munari (2008). "Association of an insulin-like growth factor 1 gene microsatellite with phenotypic variation and estimated breeding values of growth traits in Canchim cattle." <u>Anim Genet</u> **39**(5): 480-5.

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

Arai, M., M. Itokawa, K. Yamada, T. Toyota, M. Arai, S. Haga, H. Ujike, I. Sora, K. Ikeda and T. Yoshikawa (2004). "Association of neural cell

adhesion molecule 1 gene polymorphisms with bipolar affective disorder in Japanese individuals." <u>Biol Psychiatry</u> **55**(8): 804-10.

Ashley-Kosh, Ed. (2006). <u>Determining Genetic Component of a Disease</u>. Genetic Analysis of Complex Disease. New Jersey, Wiley.

Askland, K., C. Read and J. Moore (2008). "Pathways-based analyses of whole-genome association study data in bipolar disorder reveal genes mediating ion channel activity and synaptic neurotransmission." <u>Hum Genet</u>.

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

Barden, N., M. Harvey, B. Gagne, E. Shink, M. Tremblay, C. Raymond, M. Labbe, A. Villeneuve, D. Rochette, L. Bordeleau, H. Stadler, F. Holsboer and B. Muller-Myhsok (2006). "Analysis of single nucleotide polymorphisms in genes in the chromosome 12Q24.31 region points to P2RX7 as a susceptibility gene to bipolar affective disorder." <u>Am J Med Genet B Neuropsychiatr Genet</u> **141B**(4): 374-82.

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

Bass, N. J., S. R. Datta, A. McQuillin, V. Puri, K. Choudhury, S. Thirumalai, J. Lawrence, D. Quested, J. Pimm, D. Curtis and H. M. Gurling (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.

Basselin, M., L. Chang, M. Chen, J. M. Bell and S. I. Rapoport (2008). "Chronic administration of valproic acid reduces brain NMDA signaling via arachidonic acid in unanesthetized rats." Neurochem Res **33**(11): 2229-40.

Baum, A. E., N. Akula, M. Cabanero, I. Cardona, W. Corona, B. Klemens, T. G. Schulze, S. Cichon, M. Rietschel, M. M. Nothen, A. Georgi, J. Schumacher, M. Schwarz, R. Abou Jamra, S. Hofels, P. Propping, J. Satagopan, S. D. Detera-Wadleigh, J. Hardy and F. J. McMahon (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.

Baum, A. E., M. Hamshere, E. Green, S. Cichon, M. Rietschel, M. M. Noethen, N. Craddock and F. J. McMahon (2008). "Meta-analysis of two genome-wide association studies of bipolar disorder reveals important points of agreement." Mol Psychiatry **13**(5): 466-7.

Beaulieu, J. M., S. Marion, R. M. Rodriguiz, I. O. Medvedev, T. D. Sotnikova, V. Ghisi, W. C. Wetsel, R. J. Lefkowitz, R. R. Gainetdinov and M. G. Caron (2008). "A beta-arrestin 2 signaling complex mediates lithium action on behavior." Cell **132**(1): 125-36.

Beaulieu, J. M., T. D. Sotnikova, S. Marion, R. J. Lefkowitz, R. R. Gainetdinov and M. G. Caron (2005). "An Akt/beta-arrestin 2/PP2A signaling complex mediates dopaminergic neurotransmission and behavior." Cell **122**(2): 261-73.

Beaulieu, J. M., T. D. Sotnikova, W. D. Yao, L. Kockeritz, J. R. Woodgett, R. R. Gainetdinov and M. G. Caron (2004). "Lithium antagonizes dopamine-dependent behaviors mediated by an AKT/glycogen synthase kinase 3 signaling cascade." Proc Natl Acad Sci U S A **101**(14): 5099-104.

Bell, J. T., C. Wallace, R. Dobson, S. Wiltshire, C. Mein, J. Pembroke, M. Brown, D. Clayton, N. Samani, A. Dominiczak, J. Webster, G. M. Lathrop, J. Connell, P. Munroe, M. Caulfield and M. Farrall (2006). "Two-dimensional genome-scan identifies novel epistatic loci for essential hypertension." Hum Mol Genet **15**(8): 1365-74.

Bellivier, F., M. Leroux, C. Henry, F. Rayah, F. Rouillon, J. L. Laplanche and M. Leboyer (2002). "Serotonin transporter gene polymorphism influences age at onset in patients with bipolar affective disorder." Neurosci Lett **334**(1): 17-20.

Benedetti, F., A. Bernasconi, C. Lorenzi, A. Pontiggia, A. Serretti, C. Colombo and E. Smeraldi (2004). "A single nucleotide polymorphism in glycogen synthase kinase 3-beta promoter gene influences onset of illness in patients affected by bipolar disorder." Neurosci Lett **355**(1-2): 37-40.

Benedetti, F., A. Serretti, C. Colombo, B. Barbini, C. Lorenzi, E. Campori and E. Smeraldi (2003). "Influence of CLOCK gene polymorphism on circadian mood fluctuation and illness recurrence in bipolar depression." Am J Med Genet B Neuropsychiatr Genet **123B**(1): 23-6.

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

Berrettini, W. (2003). "Evidence for shared susceptibility in bipolar disorder and schizophrenia." Am J Med Genet C Semin Med Genet **123C**(1): 59-64.

Berridge, M. J. (1985). "Phosphoinositides and signal transductions." Rev Clin Basic Pharm **5 Suppl**: 5S-13S.

Berton, O. and E. J. Nestler (2006). "New approaches to antidepressant drug discovery: beyond monoamines." Nat Rev Neurosci **7**(2): 137-51.

Bielecka, A. M. and E. Obuchowicz (2008). "Antiapoptotic action of lithium and valproate." Pharmacol Rep **60**(6): 771-82.

Bitran, J. A., H. K. Manji, W. Z. Potter and F. Gusovsky (1995). "Down-regulation of PKC alpha by lithium in vitro." Psychopharmacol Bull **31**(2): 449-52.

Bleumink, G. S., A. F. Schut, M. C. Sturkenboom, J. A. Janssen, J. C. Witteman, C. M. van Duijn, A. Hofman and B. H. Stricker (2005). "A promoter polymorphism of the insulin-like growth factor-I gene is associated with left ventricular hypertrophy." Heart **91**(2): 239-40.

Bondy, C. and E. Chin (1991). "IGF-I mRNA localization in trigeminal and sympathetic nerve target zones during rat embryonic development." Adv Exp Med Biol **293**: 431-7.

Bondy, C. A. and C. M. Cheng (2004). "Signaling by insulin-like growth factor 1 in brain." Eur J Pharmacol **490**(1-3): 25-31.

Bonnier, B., P. Gorwood, M. Hamon, Y. Sarfati, C. Boni and M. C. Hardy-Bayle (2002). "Association of 5-HT(2A) receptor gene polymorphism with major affective disorders: the case of a subgroup of bipolar disorder with low suicide risk." Biol Psychiatry **51**(9): 762-5.

Brambilla, P., J. Perez, F. Barale, G. Schettini and J. C. Soares (2003). "GABAergic dysfunction in mood disorders." Mol Psychiatry **8**(8): 721-37, 715.

Breen, G., D. Prata, S. Osborne, J. Munro, M. Sinclair, T. Li, S. Staddon, D. Dempster, R. Sainz, B. Arroyo, R. W. Kerwin, D. St Clair and D. Collier (2006). "Association of the dysbindin gene with bipolar affective disorder." Am J Psychiatry **163**(9): 1636-8.

Brugue, E. and E. Vieta (2007). "Atypical antipsychotics in bipolar depression: neurobiological basis and clinical implications." Prog Neuropsychopharmacol Biol Psychiatry **31**(1): 275-82.

Brunner, H. G., M. Nelen, X. O. Breakefield, H. H. Ropers and B. A. van Oost (1993). "Abnormal behavior associated with a point mutation in the structural gene for monoamine oxidase A." Science **262**(5133): 578-80.

Bunney, W. E., Jr. and J. M. Davis (1965). "Norepinephrine in depressive reactions. A review." Arch Gen Psychiatry **13**(6): 483-94.

Burdick, K. E., B. Funke, J. F. Goldberg, J. A. Bates, J. Jaeger, R. Kucherlapati and A. K. Malhotra (2007). "COMT genotype increases risk for bipolar I disorder and influences neurocognitive performance." Bipolar Disord **9**(4): 370-6.

Bustin, S. A. and T. Nolan (2004). "Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction." J Biomol Tech **15**(3): 155-66.

Caetano, A. R. and A. T. Bowling (1998). "Characterization of a microsatellite in the promoter region of the IGF1 gene in domestic horses and other equids." Genome **41**(1): 70-3.

Cannon, D. M., M. Ichise, D. Rollis, J. M. Klaver, S. K. Gandhi, D. S. Charney, H. K. Manji and W. C. Drevets (2007). "Elevated serotonin transporter binding in major depressive disorder assessed using positron emission

tomography and [11C]DASB; comparison with bipolar disorder." Biol Psychiatry **62**(8): 870-7.

Cases, O., I. Seif, J. Grimsby, P. Gaspar, K. Chen, S. Pournin, U. Muller, M. Aguet, C. Babinet, J. C. Shih and et al. (1995). "Aggressive behavior and altered amounts of brain serotonin and norepinephrine in mice lacking MAOA." Science **268**(5218): 1763-6.

Cassidy, F., C. Zhao, J. Badger, E. Claffey, S. Dobrin, S. Roche and P. McKeon (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.

Chang, A., P. P. Li and J. J. Warsh (2003). "cAMP-Dependent protein kinase (PKA) subunit mRNA levels in postmortem brain from patients with bipolar affective disorder (BD)." Brain Res Mol Brain Res **116**(1-2): 27-37.

Chee, I. S., S. W. Lee, J. L. Kim, S. K. Wang, Y. O. Shin, S. C. Shin, Y. H. Lee, H. M. Hwang and M. R. Lim (2001). "5-HT2A receptor gene promoter polymorphism -1438A/G and bipolar disorder." Psychiatr Genet **11**(3): 111-4.

Chemelli, R. M., J. T. Willie, C. M. Sinton, J. K. Elmquist, T. Scammell, C. Lee, J. A. Richardson, S. C. Williams, Y. Xiong, Y. Kisanuki, T. E. Fitch, M.

Nakazato, R. E. Hammer, C. B. Saper and M. Yanagisawa (1999). "Narcolepsy in orexin knockout mice: molecular genetics of sleep regulation." Cell **98**(4): 437-51.

Chen, G., M. I. Masana and H. K. Manji (2000). "Lithium regulates PKC-mediated intracellular cross-talk and gene expression in the CNS in vivo." Bipolar Disord **2**(3 Pt 2): 217-36.

Cheng, R., S. H. Juo, J. E. Loth, J. Nee, I. Iossifov, R. Blumenthal, L. Sharpe, K. Kanyas, B. Lerer, B. Lilliston, M. Smith, K. Trautman, T. C. Gilliam, J. Endicott and M. Baron (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.

Collier, D. A., M. J. Arranz, P. Sham, S. Battersby, H. Vallada, P. Gill, K. J. Aitchison, M. Sodhi, T. Li, G. W. Roberts, B. Smith, J. Morton, R. M. Murray, D. Smith and G. Kirov (1996). "The serotonin transporter is a potential susceptibility factor for bipolar affective disorder." Neuroreport **7**(10): 1675-9.

Conti, E., C. Carrozza, E. Capoluongo, M. Volpe, F. Crea, C. Zuppi and F. Andreotti (2004). "Insulin-like growth factor-1 as a vascular protective factor." Circulation **110**(15): 2260-5.

Cooper, D. N. (2002). "Human gene mutation in pathology and evolution." J Inherit Metab Dis **25**(3): 157-82.

Coppen, A. (1969). "Defects in monoamine metabolism and their possible importance in the pathogenesis of depressive syndromes." Psychiatr Neurol Neurochir **72**(2): 173-80.

Coppen, A. J. (1969). "Biochemical aspects of depression." Int Psychiatry Clin **6**(2): 53-81.

Coyle, J. T. (2007). "What can a clock mutation in mice tell us about bipolar disorder?" Proc Natl Acad Sci U S A **104**(15): 6097-8.

Craddock, N., Ed. (2007). Genetics of Bipolar Disorder. Bipolar Disorders Basic Mechanisms and Therapeutic Implications. New York, Informa Healthcare.

Craddock, N., S. Dave and J. Greening (2001). "Association studies of bipolar disorder." Bipolar Disord **3**(6): 284-98.

Craddock, N. and I. Jones (1999). "Genetics of bipolar disorder." J Med Genet **36**(8): 585-94.

Craddock, N., L. Jones, I. R. Jones, G. Kirov, E. K. Green, D. Grozeva, V. Moskvina, I. Nikolov, M. L. Hamshere, D. Vukcevic, S. Caesar, K. Gordon-Smith, C. Fraser, E. Russell, N. Norton, G. Breen, D. St Clair, D.

A. Collier, A. H. Young, I. N. Ferrier, A. Farmer, P. McGuffin, P. A. Holmans, P. Donnelly, M. J. Owen and M. C. O'Donovan (2008). "Strong genetic evidence for a selective influence of GABA(A) receptors on a component of the bipolar disorder phenotype." Mol Psychiatry.

Craddock, N., M. C. O'Donovan and M. J. Owen (2005). "The genetics of schizophrenia and bipolar disorder: dissecting psychosis." J Med Genet **42**(3): 193-204.

Craddock, N. and M. J. Owen (2005). "The beginning of the end for the Kraepelinian dichotomy." Br J Psychiatry **186**: 364-6.

Cryan, J. F. and D. A. Slattery (2007). "Animal models of mood disorders: Recent developments." Curr Opin Psychiatry **20**(1): 1-7.

Curtis, D., G. Kalsi, J. Brynjolfsson, M. McInnis, J. O'Neill, C. Smyth, E. Moloney, P. Murphy, A. McQuillin, H. Petursson and H. Gurling (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 and P. C. Sham (2006). "Program report: GENECOUNTING support programs." Ann Hum Genet **70**(Pt 2): 277-9.

Curtis, D., A. Vine and J. Knight (2008). "A simple method for assessing the strength of evidence for association at the level of the whole gene." Computational Biology and Chemistry: Advances and Applications **1**: 115-120.

Daly, M. J., J. D. Rioux, S. F. Schaffner, T. J. Hudson and E. S. Lander (2001). "High-resolution haplotype structure in the human genome." Nat Genet **29**(2): 229-32.

Das Gupta, R. and J. F. Guest (2002). "Annual cost of bipolar disorder to UK society." Br J Psychiatry **180**: 227-33.

Davidson, J. (1972). "Cholinergic-adrenergic hypothesis of mania and depression." Lancet **2**(7789): 1249.

Dawson, E., E. Parfitt, Q. Roberts, J. Daniels, L. Lim, P. Sham, M. Nothen, P. Propping, M. Lanczik, W. Maier and et al. (1995). "Linkage studies of bipolar disorder in the region of the Darier's disease gene on chromosome 12q23-24.1." Am J Med Genet **60**(2): 94-102.

de Vooght, K. M., R. van Wijk and W. W. van Solinge (2008). "GATA-1 binding sites in exon 1 direct erythroid-specific transcription of PPOX." Gene **409**(1-2): 83-91.

de Vooght, K. M., R. van Wijk and W. W. van Solinge (2009). "Management of Gene Promoter Mutations in Molecular Diagnostics." Clin Chem.

de Vries, B. B., R. Pfundt, M. Leisink, D. A. Koolen, L. E. Vissers, I. M. Janssen, S. Reijmersdal, W. M. Nillesen, E. H. Huys, N. Leeuw, D. Smeets, E. A. Sistermans, T. Feuth, C. M. van Ravenswaaij-Arts, A. G. van Kessel, E. F. Schoenmakers, H. G. Brunner and J. A. Veltman (2005). "Diagnostic genome profiling in mental retardation." Am J Hum Genet **77**(4): 606-16.

Dean, B. (2004). "The neurobiology of bipolar disorder: findings using human postmortem central nervous system tissue." Aust N Z J Psychiatry **38**(3): 135-40.

Dean, B., E. Scarr and M. McLeod (2005). "Changes in hippocampal GABAA receptor subunit composition in bipolar 1 disorder." Brain Res Mol Brain Res **138**(2): 145-55.

Degn, B., M. D. Lundorf, A. Wang, M. Vang, O. Mors, T. A. Kruse and H. Ewald (2001). "Further evidence for a bipolar risk gene on chromosome 12q24 suggested by investigation of haplotype sharing and allelic association in patients from the Faroe Islands." Mol Psychiatry **6**(4): 450-5.

Del Zompo, M., V. De Luca, G. Severino, X. Ni, S. Mulas, D. Congiu, M. P. Piccardi and J. L. Kennedy (2007). "Haplotype association study between DRD1 gene and bipolar type I affective disorder in two samples from Canada and Sardinia." Am J Med Genet B Neuropsychiatr Genet **144B**(2): 237-41.

Denoeud, F., P. Kapranov, C. Ucla, A. Frankish, R. Castelo, J. Drenkow, J. Lagarde, T. Alioto, C. Manzano, J. Chrast, S. Dike, C. Wyss, C. N. Henrichsen, N. Holroyd, M. C. Dickson, R. Taylor, Z. Hance, S. Foissac, R. M. Myers, J. Rogers, T. Hubbard, J. Harrow, R. Guigo, T. R. Gingeras, S. E. Antonarakis and A. Reymond (2007). "Prominent use of distal 5' transcription start sites and discovery of a large number of additional exons in ENCODE regions." Genome Res **17**(6): 746-59.

Dheda, K., J. F. Huggett, S. A. Bustin, M. A. Johnson, G. Rook and A. Zumla (2004). "Validation of housekeeping genes for normalizing RNA expression in real-time PCR." Biotechniques **37**(1): 112-4, 116, 118-9.

Dickson, M. C., J. C. Saunders and R. S. Gilmour (1991). "The ovine insulin-like growth factor-I gene: characterization, expression and identification of a putative promoter." J Mol Endocrinol **6**(1): 17-31.

Diehl, D. J. and S. Gershon (1992). "The role of dopamine in mood disorders." Compr Psychiatry **33**(2): 115-20.

Dmitrzak-Weglarz, M., J. K. Rybakowski, A. Slopien, P. M. Czerski, A. Leszczynska-Rodziewicz, P. Kapelski, M. Kaczmarkiewicz-Fass and J. Hauser (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.

Du, J., T. K. Creson, L. J. Wu, M. Ren, N. A. Gray, C. Falke, Y. Wei, Y. Wang, R. Blumenthal, R. Machado-Vieira, P. Yuan, G. Chen, M. Zhuo and H. K. Manji (2008). "The role of hippocampal GluR1 and GluR2 receptors in manic-like behavior." J Neurosci **28**(1): 68-79.

Du, J., N. A. Gray, C. Falke, P. Yuan, S. Szabo and H. K. Manji (2003). "Structurally dissimilar antimanic agents modulate synaptic plasticity by regulating AMPA glutamate receptor subunit GluR1 synaptic expression." Ann N Y Acad Sci **1003**: 378-80.

Du, J., N. A. Gray, C. A. Falke, W. Chen, P. Yuan, S. T. Szabo, H. Einat and H. K. Manji (2004). "Modulation of synaptic plasticity by antimanic agents: the role of AMPA glutamate receptor subunit 1 synaptic expression." J Neurosci **24**(29): 6578-89.

Duan, C. (2002). "Specifying the cellular responses to IGF signals: roles of IGF-binding proteins." J Endocrinol **175**(1): 41-54.

Duffy, D. L., G. W. Montgomery, W. Chen, Z. Z. Zhao, L. Le, M. R. James, N. K. Hayward, N. G. Martin and R. A. Sturm (2007). "A three-single-nucleotide polymorphism haplotype in intron 1 of OCA2 explains most human eye-color variation." Am J Hum Genet **80**(2): 241-52.

Dunlop, B. W. and C. B. Nemeroff (2007). "The role of dopamine in the pathophysiology of depression." Arch Gen Psychiatry **64**(3): 327-37.

El-Mallakh, R. and S. Ghaemi, Eds. (2006). <u>Bipolar Depression - A Comprehensive Guide</u>. Arlington, American Psychiatric Publishing.

Emilien, G., L. Septien, C. Brisard, E. Corruble and M. Bourin (2007). "Bipolar disorder: how far are we from a rigorous definition and effective management?" <u>Prog Neuropsychopharmacol Biol Psychiatry</u> **31**(5): 975-96.

Encinas, M., M. Iglesias, Y. Liu, H. Wang, A. Muhaisen, V. Cena, C. Gallego and J. X. Comella (2000). "Sequential treatment of SH-SY5Y cells with retinoic acid and brain-derived neurotrophic factor gives rise to fully differentiated, neurotrophic factor-dependent, human neuron-like cells." <u>J Neurochem</u> **75**(3): 991-1003.

Endicott, J. and R. L. Spitzer (1978). "A diagnostic interview: the schedule for affective disorders and schizophrenia." <u>Arch Gen Psychiatry</u> **35**(7): 837-44.

Engin, E., D. Treit and C. T. Dickson (2009). "Anxiolytic- and antidepressant-like properties of ketamine in behavioral and neurophysiological animal models." Psychopharmacology **206**(2):281-9.

Erson, A. E. and E. M. Petty (2008). "MicroRNAs in development and disease." <u>Clin Genet</u> **74**(4): 296-306.

Ewald, H., B. Degn, O. Mors and T. A. Kruse (1998). "Significant linkage between bipolar affective disorder and chromosome 12q24." Psychiatr Genet **8**(3): 131-40.

Ewald, H., T. Flint, T. A. Kruse and O. Mors (2002). "A genome-wide scan shows significant linkage between bipolar disorder and chromosome 12q24.3 and suggestive linkage to chromosomes 1p22-21, 4p16, 6q14-22, 10q26 and 16p13.3." Mol Psychiatry **7**(7): 734-44.

Fallin, M. D., V. K. Lasseter, D. Avramopoulos, K. K. Nicodemus, P. S. Wolyniec, J. A. McGrath, G. Steel, G. Nestadt, K. Y. Liang, R. L. Huganir, D. Valle and A. E. Pulver (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.

Faraone, S. V., W. S. Kremen and M. T. Tsuang (1990). "Genetic transmission of major affective disorders: quantitative models and linkage analyses." Psychol Bull **108**(1): 109-27.

Fernandez, S., A. M. Fernandez, C. Lopez-Lopez and I. Torres-Aleman (2007). "Emerging roles of insulin-like growth factor-I in the adult brain." Growth Horm IGF Res **17**(2): 89-95.

Ferreira, M. A., M. C. O'Donovan, Y. A. Meng, I. R. Jones, D. M. Ruderfer, L. Jones, J. Fan, G. Kirov, R. H. Perlis, E. K. Green, J. W. Smoller, D.

Grozeva, J. Stone, I. Nikolov, K. Chambert, M. L. Hamshere, V. L. Nimgaonkar, V. Moskvina, M. E. Thase, S. Caesar, G. S. Sachs, J. Franklin, K. Gordon-Smith, K. G. Ardlie, S. B. Gabriel, C. Fraser, B. Blumenstiel, M. Defelice, G. Breen, M. Gill, D. W. Morris, A. Elkin, W. J. Muir, K. A. McGhee, R. Williamson, D. J. Macintyre, A. W. Maclean, D. St Clair, M. Robinson, M. Van Beck, A. C. Pereira, R. Kandaswamy, A. McQuillin, D. A. Collier, N. J. Bass, A. H. Young, J. Lawrence, I. Nicol Ferrier, A. Anjorin, A. Farmer, D. Curtis, E. M. Scolnick, P. McGuffin, M. J. Daly, A. P. Corvin, P. A. Holmans, D. H. Blackwood, H. M. Gurling, M. J. Owen, S. M. Purcell, P. Sklar and N. Craddock (2008). "Collaborative genome-wide association analysis supports a role for ANK3 and CACNA1C in bipolar disorder." <u>Nat Genet</u>. **40**(9):1056-8.

Feuk, L., A. R. Carson and S. W. Scherer (2006). "Structural variation in the human genome." <u>Nat Rev Genet</u> **7**(2): 85-97.

Fridman, C., E. P. Ojopi, S. P. Gregorio, E. H. Ikenaga, D. H. Moreno, F. N. Demetrio, P. E. Guimaraes, H. P. Vallada, W. F. Gattaz and E. Dias Neto (2003). "Association of a new polymorphism in ALOX12 gene with bipolar disorder." <u>Eur Arch Psychiatry Clin Neurosci</u> **253**(1): 40-3.

Gao, K., P. Gajwani, O. Elhaj and J. R. Calabrese (2005). "Typical and atypical antipsychotics in bipolar depression." <u>J Clin Psychiatry</u> **66**(11): 1376-85.

Gaysina, D., S. Cohen-Woods, P. C. Chow, L. Martucci, A. Schosser, H. A. Ball, F. Tozzi, J. Perry, P. Muglia, I. W. Craig, P. McGuffin and A. Farmer

(2008). "Association of the dystrobrevin binding protein 1 gene (DTNBP1) in a bipolar case-control study (BACCS)." <u>Am J Med Genet B Neuropsychiatr Genet</u>.

Gershon, A. A., T. Vishne and L. Grunhaus (2007). "Dopamine D2-like receptors and the antidepressant response." <u>Biol Psychiatry</u> **61**(2): 145-53.

Gimenez-Cassina, A., F. Lim and J. Diaz-Nido (2006). "Differentiation of a human neuroblastoma into neuron-like cells increases their susceptibility to transduction by herpesviral vectors." <u>J Neurosci Res</u> **84**(4): 755-67.

Girgis, R. R., J. A. Javitch and J. A. Lieberman (2008). "Antipsychotic drug mechanisms: links between therapeutic effects, metabolic side effects and the insulin signaling pathway." <u>Mol Psychiatry</u>.

Glaser, B., G. Kirov, E. Green, N. Craddock and M. J. Owen (2005). "Linkage disequilibrium mapping of bipolar affective disorder at 12q23-q24 provides evidence for association at CUX2 and FLJ32356." <u>Am J Med Genet B Neuropsychiatr Genet</u> **132B**(1): 38-45.

Gottlicher, M., S. Minucci, P. Zhu, O. H. Kramer, A. Schimpf, S. Giavara, J. P. Sleeman, F. Lo Coco, C. Nervi, P. G. Pelicci and T. Heinzel (2001). "Valproic acid defines a novel class of HDAC inhibitors inducing differentiation of transformed cells." <u>Embo J</u> **20**(24): 6969-78.

Gould, T. D., H. Einat, K. C. O'Donnell, A. M. Picchini, R. J. Schloesser and H. K. Manji (2007). "Beta-catenin overexpression in the mouse brain phenocopies lithium-sensitive behaviors." Neuropsychopharmacology **32**(10): 2173-83.

Gourovitch, M. L., E. F. Torrey, J. M. Gold, C. Randolph, D. R. Weinberger and T. E. Goldberg (1999). "Neuropsychological performance of monozygotic twins discordant for bipolar disorder." Biol Psychiatry **45**(5): 639-46.

Green, E. K., D. Grozeva, R. Raybould, G. Elvidge, S. Macgregor, I. Craig, A. Farmer, P. McGuffin, L. Forty, L. Jones, I. Jones, M. C. O'Donovan, M. J. Owen, G. Kirov and N. Craddock (2009). "P2RX7: A bipolar and unipolar disorder candidate susceptibility gene?" Am J Med Genet B Neuropsychiatr Genet.

Green, E. K., R. Raybould, S. Macgregor, S. Hyde, A. H. Young, M. C. O'Donovan, M. J. Owen, G. Kirov, L. Jones, I. Jones and N. Craddock (2006). "Genetic variation of brain-derived neurotrophic factor (BDNF) in bipolar disorder: case-control study of over 3000 individuals from the UK." Br J Psychiatry **188**: 21-5.

Grigoroiu-Serbanescu, M., C. C. Diaconu, S. Herms, C. Bleotu, J. Vollmer, T. W. Muhleisen, D. Prelipceanu, L. Priebe, R. Mihailescu, M. J. Georgescu, D. Sima, M. Grimberg, M. M. Nothen and S. Cichon (2008).

"Investigation of the tryptophan hydroxylase 2 gene in bipolar I disorder in the Romanian population." Psychiatr Genet **18**(5): 240-7.

Guo, J. H., Q. Zan and L. Yu (2001). GenBank.

Gurling, H. and L. Rifkin (1991). The genetic aspects of affective disorders. Biological Aspects of Affective Disorders. R. Horton and C. Katona. London, Academic Press.

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

Gutierrez, B., B. Arias, C. Gasto, R. Catalan, S. Papiol, L. Pintor and L. Fananas (2004). "Association analysis between a functional polymorphism in the monoamine oxidase A gene promoter and severe mood disorders." Psychiatr Genet **14**(4): 203-8.

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

Haines, J. and M. Pericak-Vance, Eds. (2006). Genetic Analysis of Complex Disease. New Jersey, Wiley-Liss.

Haldane, J. B. and C. A. Smith (1947). "A new estimate of the linkage between the genes for colourblindness and haemophilia in man." Ann Eugen **14**(pt 1): 10-31.

Hand, B. D., M. C. Kostek, R. E. Ferrell, M. J. Delmonico, L. W. Douglass, S. M. Roth, J. M. Hagberg and B. F. Hurley (2007). "Influence of promoter region variants of insulin-like growth factor pathway genes on the strength-training response of muscle phenotypes in older adults." J Appl Physiol **103**(5): 1678-87.

Harvey, M., B. Gagne, M. Labbe and N. Barden (2007). "Polymorphisms in the neuronal isoform of tryptophan hydroxylase 2 are associated with bipolar disorder in French Canadian pedigrees." Psychiatr Genet **17**(1): 17-22.

Harwood, A. J. (2005). "Lithium and bipolar mood disorder: the inositol-depletion hypothesis revisited." Mol Psychiatry **10**(1): 117-26.

Hashimoto, R., T. Okada, T. Kato, A. Kosuga, M. Tatsumi, K. Kamijima and H. Kunugi (2005). "The breakpoint cluster region gene on chromosome 22q11 is associated with bipolar disorder." Biol Psychiatry **57**(10): 1097-102.

Hayden, E. P. and J. I. Nurnberger, Jr. (2006). "Molecular genetics of bipolar disorder." Genes Brain Behav **5**(1): 85-95.

Hays, J. C., K. R. Krishnan, L. K. George and D. G. Blazer (1998). "Age of first onset of bipolar disorder: demographic, family history, and psychosocial correlates." Depress Anxiety **7**(2): 76-82.

Hennah, W., P. Thomson, A. McQuillin, N. Bass, A. Loukola, A. Anjorin, D. Blackwood, D. Curtis, I. J. Deary, S. E. Harris, E. T. Isometsa, J. Lawrence, J. Lonnqvist, W. Muir, A. Palotie, T. Partonen, T. Paunio, E. Pylkko, M. Robinson, P. Soronen, K. Suominen, J. Suvisaari, S. Thirumalai, D. S. Clair, H. Gurling, L. Peltonen and D. Porteous (2008). "DISC1 association, heterogeneity and interplay in schizophrenia and bipolar disorder." Mol Psychiatry.

Herzberg, I., A. Jasinska, J. Garcia, D. Jawaheer, S. Service, B. Kremeyer, C. Duque, M. V. Parra, J. Vega, D. Ortiz, L. Carvajal, G. Polanco, G. J. Restrepo, C. Lopez, C. Palacio, M. Levinson, I. Aldana, C. Mathews, P. Davanzo, J. Molina, E. Fournier, J. Bejarano, M. Ramirez, C. A. Ortiz, X. Araya, C. Sabatti, V. Reus, G. Macaya, G. Bedoya, J. Ospina, N. Freimer and A. Ruiz-Linares (2006). "Convergent linkage evidence from two Latin-American population isolates supports the presence of a susceptibility locus for bipolar disorder in 5q31-34." Hum Mol Genet **15**(21): 3146-53.

Heun, R. and W. Maier (1993). "Bipolar II disorders in six first-degree relatives." Biol Psychiatry **34**(4): 274-6.

Higgs, B. W., M. Elashoff, S. Richman and B. Barci (2006). "An online database for brain disease research." BMC Genomics **7**: 70.

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

Hodgkinson, C. A., D. Goldman, J. Jaeger, S. Persaud, J. M. Kane, R. H. Lipsky and A. K. Malhotra (2004). "Disrupted in schizophrenia 1 (DISC1): association with schizophrenia, schizoaffective disorder, and bipolar disorder." Am J Hum Genet **75**(5): 862-72.

Holmes, A., J. E. Lachowicz and D. R. Sibley (2004). "Phenotypic analysis of dopamine receptor knockout mice; recent insights into the functional specificity of dopamine receptor subtypes." Neuropharmacology **47**(8): 1117-34.

Horiuchi, Y., J. Nakayama, H. Ishiguro, T. Ohtsuki, S. D. Detera-Wadleigh, T. Toyota, K. Yamada, M. Nankai, H. Shibuya, T. Yoshikawa and T. Arinami (2004). "Possible association between a haplotype of the GABA-A receptor alpha 1 subunit gene (GABRA1) and mood disorders." Biol Psychiatry **55**(1): 40-5.

Hoshaw, B. A., T. I. Hill, J. J. Crowley, J. E. Malberg, X. Khawaja, S. Rosenzweig-Lipson, L. E. Schechter and I. Lucki (2008).

"Antidepressant-like behavioral effects of IGF-I produced by enhanced serotonin transmission." Eur J Pharmacol **594**(1-3): 109-16.

Hoshaw, B. A., J. E. Malberg and I. Lucki (2005). "Central administration of IGF-I and BDNF leads to long-lasting antidepressant-like effects." Brain Res **1037**(1-2): 204-8.

Huggett, J., K. Dheda, S. Bustin and A. Zumla (2005). "Real-time RT-PCR normalisation; strategies and considerations." Genes Immun **6**(4): 279-84.

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

Iwamoto, K., M. Bundo and T. Kato (2005). "Altered expression of mitochondria-related genes in postmortem brains of patients with bipolar disorder or schizophrenia, as revealed by large-scale DNA microarray analysis." Hum Mol Genet **14**(2): 241-53.

Jakobsen, S. N. and O. Wiborg (1998). "Selective effects of long-term lithium and carbamazepine administration on G-protein subunit expression in rat brain." Brain Res **780**(1): 46-55.

Janowsky, D. S., M. K. el-Yousef, J. M. Davis and H. J. Sekerke (1972). "A cholinergic-adrenergic hypothesis of mania and depression." Lancet **2**(7778): 632-5.

Jasinska, A. J., S. Service, D. Jawaheer, J. Deyoung, M. Levinson, Z. Zhang, B. Kremeyer, H. Muller, I. Aldana, J. Garcia, G. Restrepo, C. Lopez, C. Palacio, C. Duque, M. Parra, J. Vega, D. Ortiz, G. Bedoya, C. Mathews, P. Davanzo, E. Fournier, J. Bejarano, M. Ramirez, C. Araya Ortiz, X. Araya, J. Molina, C. Sabatti, V. Reus, J. Ospina, G. Macaya, A. Ruiz-Linares and N. B. Freimer (2009). "A narrow and highly significant linkage signal for severe bipolar disorder in the chromosome 5q33 region in Latin American pedigrees." Am J Med Genet B Neuropsychiatr Genet.

Joo, E. J., T. A. Greenwood, N. Schork, R. A. McKinney, A. D. Sadovnick, R. A. Remick, P. E. Keck, S. L. McElroy and J. R. Kelsoe (2009). "Suggestive evidence for linkage of ADHD features in bipolar disorder to chromosome 10p14." Am J Med Genet B Neuropsychiatr Genet.

Joo, E. J., K. Y. Lee, S. H. Jeong, J. S. Chang, Y. M. Ahn, Y. J. Koo and Y. S. Kim (2007). "Dysbindin gene variants are associated with bipolar I disorder in a Korean population." Neurosci Lett 418(3): 272-5.

Jope, R. (1979). "Effects of lithium treatment in vitro and in vivo on acetylcholine metabolism in rat brain." J Neurochem 33(2): 487-95.

Kakiuchi, C., K. Iwamoto, M. Ishiwata, M. Bundo, T. Kasahara, I. Kusumi, T. Tsujita, Y. Okazaki, S. Nanko, H. Kunugi, T. Sasaki and T. Kato (2003). "Impaired feedback regulation of XBP1 as a genetic risk factor for bipolar disorder." Nat Genet 35(2): 171-5.

Kalsi, G., A. McQuillin, B. Degn, M. D. Lundorf, N. J. Bass, J. Lawrence, K. Choudhury, V. Puri, M. Nyegaard, D. Curtis, O. Mors, T. Kruse, S. Kerwin and H. Gurling (2006). "Identification of the Slynar gene (AY070435) and related brain expressed sequences as a candidate gene for susceptibility to affective disorders through allelic and haplotypic association with bipolar disorder on chromosome 12q24." Am J Psychiatry **163**(10): 1767-76.

Kapur, S. and J. J. Mann (1992). "Role of the dopaminergic system in depression." Biol Psychiatry **32**(1): 1-17.

Kar, S., J. G. Chabot and R. Quirion (1993). "Quantitative autoradiographic localization of [125I]insulin-like growth factor I, [125I]insulin-like growth factor II, and [125I]insulin receptor binding sites in developing and adult rat brain." J Comp Neurol **333**(3): 375-97.

Karege, F., C. Buresi, J. Golaz, M. Schwald and A. Malafosse (2000). "Decreased expression of Galphas mRNA and protein levels in lithium-treated bipolar affective disorder." Hum Psychopharmacol **15**(3): 191-197.

Karege, F., J. Golaz, M. Schwald and A. Malafosse (1999). "Lithium and haloperidol treatments differently affect the mononuclear leukocyte Galphas protein levels in bipolar affective disorder." Neuropsychobiology **39**(4): 181-6.

Kaster, M. P., I. Raupp, R. W. Binfare, R. Andreatini and A. L. Rodrigues (2007). "Antidepressant-like effect of lamotrigine in the mouse forced swimming test: evidence for the involvement of the noradrenergic system." Eur J Pharmacol **565**(1-3): 119-24.

Kato, T. (2001). "The other, forgotten genome: mitochondrial DNA and mental disorders." Molecular Psychiatry **6**(6): 625-633.

Kato, T. (2006). "The role of mitochondrial dysfunction in bipolar disorder." Drug News Perspect **19**(10): 597-602.

Kato, T., M. Ishiwata, K. Mori, S. Washizuka, O. Tajima, T. Akiyama and N. Kato (2003). "Mechanisms of altered Ca2+ signalling in transformed lymphoblastoid cells from patients with bipolar disorder." Int J Neuropsychopharmacol **6**(4): 379-89.

Kempermann, G. and G. Kronenberg (2003). "Depressed new neurons adult hippocampal neurogenesis and a cellular plasticity hypothesis of major depression." Biol Psychiatry **54**(5): 499-503.

Kempisty, B., J. Sikora, M. Lianeri, A. Szczepankiewicz, P. Czerski, J. Hauser and P. P. Jagodzinski (2007). "MTHFD 1958G>A and MTR 2756A>G polymorphisms are associated with bipolar disorder and schizophrenia." Psychiatr Genet **17**(3): 177-81.

Khawaja, X., J. Xu, J. J. Liang and J. E. Barrett (2004). "Proteomic analysis of protein changes developing in rat hippocampus after chronic antidepressant treatment: Implications for depressive disorders and future therapies." J Neurosci Res **75**(4): 451-60.

Kibbe, W. A. (2007). "OligoCalc: an online oligonucleotide properties calculator." Nucleic Acids Res **35**(Web Server issue): W43-6.

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

Kirkpatrick, B. W. (1992). "Identification of a conserved microsatellite site in the porcine and bovine insulin-like growth factor-I gene 5' flank." Anim Genet **23**(6): 543-8.

Kirov, G., K. C. Murphy, M. J. Arranz, I. Jones, F. McCandles, H. Kunugi, R. M. Murray, P. McGuffin, D. A. Collier, M. J. Owen and N. Craddock (1998). "Low activity allele of catechol-O-methyltransferase gene associated with rapid cycling bipolar disorder." Mol Psychiatry **3**(4): 342-5.

Koido, K., S. Koks, T. Nikopensius, E. Maron, S. Altmae, E. Heinaste, K. Vabrit, V. Tammekivi, P. Hallast, A. Kurg, J. Shlik, V. Vasar, A. Metspalu and E. Vasar (2005). "Polymorphisms in wolframin (WFS1) gene are possibly related to increased risk for mood disorders." Int J Neuropsychopharmacol **8**(2): 235-44.

Kornberg, J. R., J. L. Brown, A. D. Sadovnick, R. A. Remick, P. E. Keck, Jr., S. L. McElroy, M. H. Rapaport, P. M. Thompson, J. B. Kaul, C. M. Vrabel, S. C. Schommer, T. Wilson, D. Pizzuco, S. Jameson, L. Schibuk and J. R. Kelsoe (2000). "Evaluating the parent-of-origin effect in bipolar affective disorder. Is a more penetrant subtype transmitted paternally?" J Affect Disord **59**(3): 183-92.

Kostek, M. C., M. J. Delmonico, J. B. Reichel, S. M. Roth, L. Douglass, R. E. Ferrell and B. F. Hurley (2005). "Muscle strength response to strength training is influenced by insulin-like growth factor 1 genotype in older adults." J Appl Physiol **98**(6): 2147-54.

Kremeyer, B., I. Herzberg, J. Garcia, E. Kerr, C. Duque, V. Parra, J. Vega, C. Lopez, C. Palacio, G. Bedoya, J. Ospina and A. Ruiz-Linares (2006). "Transmission distortion of BDNF variants to bipolar disorder type I patients from a South American population isolate." Am J Med Genet B Neuropsychiatr Genet **141B**(5): 435-9.

Kupka, R. and W. Nolen, Eds. (2007). Classification of Bipolar Disorders - Implications for Clinical Research. Bipolar Disorders Basic Mechanisms and Therapeutic Implications. New York, Informa Healthcare.

Lambert, D., F. Middle, M. L. Hamshere, R. Segurado, R. Raybould, A. Corvin, E. Green, E. O'Mahony, I. Nikolov, T. Mulcahy, S. Haque, S. Bort, P. Bennett, N. Norton, M. J. Owen, G. Kirov, C. Lendon, L. Jones, I. Jones, P. Holmans, M. Gill and N. Craddock (2005). "Stage 2 of the Wellcome

Trust UK-Irish bipolar affective disorder sibling-pair genome screen: evidence for linkage on chromosomes 6q16-q21, 4q12-q21, 9p21, 10p14-p12 and 18q22." Mol Psychiatry **10**(9): 831-41.

Lasky-Su, J. A., S. V. Faraone, S. J. Glatt and M. T. Tsuang (2005). "Meta-analysis of the association between two polymorphisms in the serotonin transporter gene and affective disorders." Am J Med Genet B Neuropsychiatr Genet **133B**(1): 110-5.

Le-Niculescu, H., S. D. Patel, M. Bhat, R. Kuczenski, S. V. Faraone, M. T. Tsuang, F. J. McMahon, N. J. Schork, J. I. Nurnberger, Jr. and A. B. Niculescu, 3rd (2008). "Convergent functional genomics of genome-wide association data for bipolar disorder: Comprehensive identification of candidate genes, pathways and mechanisms." Am J Med Genet B Neuropsychiatr Genet.

Le Hellard, S., A. J. Lee, S. Underwood, P. A. Thomson, S. W. Morris, H. S. Torrance, S. M. Anderson, R. R. Adams, P. Navarro, A. Christoforou, L. M. Houlihan, S. Detera-Wadleigh, M. J. Owen, P. Asherson, W. J. Muir, D. H. Blackwood, N. R. Wray, D. J. Porteous and K. L. Evans (2007). "Haplotype analysis and a novel allele-sharing method refines a chromosome 4p locus linked to bipolar affective disorder." Biol Psychiatry **61**(6): 797-805.

Lee, W. H. and C. Bondy (1993). "Insulin-like growth factors and cerebral ischemia." Ann N Y Acad Sci **679**: 418-22.

Leng, Y., M. H. Liang, M. Ren, Z. Marinova, P. Leeds and D. M. Chuang (2008). "Synergistic neuroprotective effects of lithium and valproic acid or other histone deacetylase inhibitors in neurons: roles of glycogen synthase kinase-3 inhibition." J Neurosci **28**(10): 2576-88.

Lerer, B. and M. Stanley (1985). "Effect of chronic lithium on cholinergically mediated responses and [3H]QNB binding in rat brain." Brain Res **344**(2): 211-9.

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.

Leung, R. K. and P. A. Whittaker (2005). "RNA interference: from gene silencing to gene-specific therapeutics." Pharmacol Ther **107**(2): 222-39.

Li, W. B., C. Gruber, J. Jessee and D. Polayes (2004). Full-length cDNA libraries and normalization, GenBank.

Li, X., J. Zhang, Y. Wang, J. Ji, F. Yang, C. Wan, P. Wang, G. Feng, K. Lindpaintner, L. He and G. He (2009). "Association study on the NAPG gene and bipolar disorder in the Chinese Han population." Neurosci Lett **457**(3): 159-62.

Lim, L. C., J. Powell, P. Sham, D. Castle, N. Hunt, R. Murray and M. Gill (1995). "Evidence for a genetic association between alleles of monoamine oxidase A gene and bipolar affective disorder." Am J Med Genet **60**(4): 325-31.

Lin, S., S. Jiang, X. Wu, Y. Qian, D. Wang, G. Tang and N. Gu (2000). "Association analysis between mood disorder and monoamine oxidase gene." Am J Med Genet **96**(1): 12-4.

Lin, Y. M., S. C. Chao, T. M. Chen, T. J. Lai, J. S. Chen and H. S. Sun (2007). "Association of functional polymorphisms of the human tryptophan hydroxylase 2 gene with risk for bipolar disorder in Han Chinese." Arch Gen Psychiatry **64**(9): 1015-24.

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

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.

Lohoff, F. W., J. P. Dahl, T. N. Ferraro, S. E. Arnold, J. Gallinat, T. Sander and W. H. Berrettini (2006). "Variations in the vesicular monoamine transporter 1 gene (VMAT1/SLC18A1) are associated with bipolar i disorder." Neuropsychopharmacology **31**(12): 2739-47.

Lohoff, F. W., T. Sander, T. N. Ferraro, J. P. Dahl, J. Gallinat and W. H. Berrettini (2005). "Confirmation of association between the Val66Met polymorphism in the brain-derived neurotrophic factor (BDNF) gene and bipolar I disorder." Am J Med Genet B Neuropsychiatr Genet **139B**(1): 51-3.

Lopez Leon, S., E. A. Croes, F. A. Sayed-Tabatabaei, S. Claes, C. Van Broeckhoven and C. M. van Duijn (2005). "The dopamine D4 receptor gene 48-base-pair-repeat polymorphism and mood disorders: a meta-analysis." Biol Psychiatry **57**(9): 999-1003.

Lopez, V. A., S. Detera-Wadleigh, I. Cardona, L. Kassem and F. J. McMahon (2007). "Nested association between genetic variation in tryptophan hydroxylase II, bipolar affective disorder, and suicide attempts." Biol Psychiatry **61**(2): 181-6.

Lupski, J. R. and P. Stankiewicz (2005). "Genomic disorders: molecular mechanisms for rearrangements and conveyed phenotypes." PLoS Genet **1**(6): e49.

Madsen, O., M. Scally, C. J. Douady, D. J. Kao, R. W. DeBry, R. Adkins, H. M. Amrine, M. J. Stanhope, W. W. de Jong and M. S. Springer (2001). "Parallel adaptive radiations in two major clades of placental mammals." Nature **409**(6820): 610-4.

Malberg, J. E. and J. A. Blendy (2005). "Antidepressant action: to the nucleus and beyond." Trends Pharmacol Sci **26**(12): 631-8.

Malhi, G. S., P. B. Mitchell and S. Salim (2003). "Bipolar depression: management options." CNS Drugs **17**(1): 9-25.

Manji, H. K., Y. Bersudsky, G. Chen, R. H. Belmaker and W. Z. Potter (1996). "Modulation of protein kinase C isozymes and substrates by lithium: the role of myo-inositol." Neuropsychopharmacology **15**(4): 370-81.

Manji, H. K., J. A. Quiroz, J. L. Payne, J. Singh, B. P. Lopes, J. S. Viegas and C. A. Zarate (2003). "The underlying neurobiology of bipolar disorder." World Psychiatry **2**(3): 136-46.

Mao, Y., X. Ge, C. L. Frank, J. M. Madison, A. N. Koehler, M. K. Doud, C. Tassa, E. M. Berry, T. Soda, K. K. Singh, T. Biechele, T. L. Petryshen, R. T. Moon, S. J. Haggarty and L. H. Tsai (2009). "Disrupted in schizophrenia 1 regulates neuronal progenitor proliferation via modulation of GSK3beta/beta-catenin signaling." Cell **136**(6): 1017-31.

Masoliver, E., A. Menoyo, V. Perez, V. Volpini, E. D. Rio, J. Perez, E. Alvarez and M. Baiget (2006). "Serotonin transporter linked promoter (polymorphism) in the serotonin transporter gene may be associated with antidepressant-induced mania in bipolar disorder." Psychiatr Genet **16**(1): 25-9.

Massat, I., D. Souery, J. Del-Favero, L. Oruc, M. M. Noethen, D. Blackwood, M. Thomson, W. Muir, G. N. Papadimitriou, D. G. Dikeos, R. Kaneva, A. Serretti, R. Lilli, E. Smeraldi, M. Jakovljevic, V. Folnegovic, M. Rietschel, V. Milanova, F. Valente, C. Van Broeckhoven and J. Mendlewicz (2002). "Excess of allele1 for alpha3 subunit GABA receptor gene (GABRA3) in bipolar patients: a multicentric association study." Mol Psychiatry **7**(2): 201-7.

Massat, I., D. Souery, J. Del-Favero, S. Van Gestel, A. Serretti, F. Macciardi, E. Smeraldi, R. Kaneva, R. Adolfsson, P. O. Nylander, D. Blackwood, W. Muir, G. N. Papadimitriou, D. Dikeos, L. Oruc, R. H. Segman, S. Ivezic, H. Aschauer, M. Ackenheil, S. Fuchshuber, H. Dam, M. Jakovljevic, L. Peltonen, C. Hilger, F. Hentges, L. Staner, V. Milanova, E. Jazin, B. Lerer, C. Van Broeckhoven and J. Mendlewicz (2002). "Positive association of dopamine D2 receptor polymorphism with bipolar affective disorder in a European Multicenter Association Study of affective disorders." Am J Med Genet **114**(2): 177-85.

Maston, G. A., S. K. Evans and M. R. Green (2006). "Transcriptional regulatory elements in the human genome." Annu Rev Genomics Hum Genet **7**: 29-59.

Masui, T., R. Hashimoto, I. Kusumi, K. Suzuki, T. Tanaka, S. Nakagawa, H. Kunugi and T. Koyama (2006). "A possible association between the -116C/G single nucleotide polymorphism of the XBP1 gene and lithium prophylaxis in bipolar disorder." Int J Neuropsychopharmacol **9**(1): 83-8.

Mattick, J. S., P. P. Amaral, M. E. Dinger, T. R. Mercer and M. F. Mehler (2009). "RNA regulation of epigenetic processes." Bioessays **31**(1): 51-9.

Mattson, M. P., S. Maudsley and B. Martin (2004). "A neural signaling triumvirate that influences ageing and age-related disease: insulin/IGF-1, BDNF and serotonin." Ageing Res Rev **3**(4): 445-64.

Maziade, M., Y. C. Chagnon, M. A. Roy, A. Bureau, A. Fournier and C. Merette (2009). "Chromosome 13q13-q14 locus overlaps mood and psychotic disorders: the relevance for redefining phenotype." Eur J Hum Genet **17**(8): 1034-42.

McAuley, E. Z., I. P. Blair, Z. Liu, J. M. Fullerton, A. Scimone, M. Van Herten, M. R. Evans, K. C. Kirkby, J. A. Donald, P. B. Mitchell and P. R. Schofield (2009). "A genome screen of 35 bipolar affective disorder pedigrees provides significant evidence for a susceptibility locus on chromosome 15q25-26." Mol Psychiatry **14**(5): 492-500.

McGuffin, P., A. Farmer and I. Harvey (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., F. Rijsdijk, M. Andrew, P. Sham, R. Katz and A. Cardno (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, G. A. Chase, S. G. Simpson, C. A. Ross and J. R. DePaulo, Jr. (1993). "Anticipation in bipolar affective disorder." <u>Am J Hum Genet</u> **53**(2): 385-90.

McKinney, J. A., B. Turel, I. Winge, P. M. Knappskog and J. Haavik (2009). "Functional properties of missense variants of human tryptophan hydroxylase 2." <u>Hum Mutat</u>.

McMahon, F. J., O. C. Stine, D. A. Meyers, S. G. Simpson and J. R. DePaulo (1995). "Patterns of maternal transmission in bipolar affective disorder." <u>Am J Hum Genet</u> **56**(6): 1277-86.

McQueen, M. B., B. Devlin, S. V. Faraone, V. L. Nimgaonkar, P. Sklar, J. W. Smoller, R. Abou Jamra, M. Albus, S. A. Bacanu, M. Baron, T. B. Barrett, W. Berrettini, D. Blacker, W. Byerley, S. Cichon, W. Coryell, N. Craddock, M. J. Daly, J. R. Depaulo, H. J. Edenberg, T. Foroud, M. Gill, T. C. Gilliam, M. Hamshere, I. Jones, L. Jones, S. H. Juo, J. R. Kelsoe, D. Lambert, C. Lange, B. Lerer, J. Liu, W. Maier, J. D. Mackinnon, M. G. McInnis, F. J. McMahon, D. L. Murphy, M. M. Nothen, J. I. Nurnberger, C. N. Pato, M. T. Pato, J. B. Potash, P. Propping, A. E. Pulver, J. P. Rice, M. Rietschel, W. Scheftner, J. Schumacher, R. Segurado, K. Van Steen, W. Xie, P. P. Zandi and N. M. Laird (2005). "Combined analysis from eleven linkage studies of bipolar disorder provides strong evidence of susceptibility loci on chromosomes 6q and 8q." <u>Am J Hum Genet</u> **77**(4): 582-95.

McQuillin, A., N. J. Bass, G. Kalsi, J. Lawrence, V. Puri, K. Choudhury, S. D. Detera-Wadleigh, D. Curtis and H. M. Gurling (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.

McQuillin, A., M. Rizig and H. M. Gurling (2007). "A microarray gene expression study of the molecular pharmacology of lithium carbonate on mouse brain mRNA to understand the neurobiology of mood stabilization and treatment of bipolar affective disorder." Pharmacogenet Genomics **17**(8): 605-17.

Mehler, M. F. and J. S. Mattick (2007). "Noncoding RNAs and RNA editing in brain development, functional diversification, and neurological disease." Physiol Rev **87**(3): 799-823.

Mellerup, E., B. Bennike, T. Bolwig, H. Dam, L. Hasholt, M. B. Jorgensen, P. Plenge and S. A. Sorensen (2001). "Platelet serotonin transporters and the transporter gene in control subjects, unipolar patients and bipolar patients." Acta Psychiatr Scand **103**(3): 229-33.

Mendlewicz, J. and J. D. Rainer (1977). "Adoption study supporting genetic transmission in manic--depressive illness." Nature **268**(5618): 327-9.

Merette, C., M. A. Roy, A. Bureau, A. Fournier, C. Emond, D. Cliche, V. Jomphe, Y. C. Chagnon and M. Maziade (2008). "Replication of linkage

with bipolar disorder on chromosome 16p in the Eastern Quebec population." Am J Med Genet B Neuropsychiatr Genet **147B**(6): 737-44.

Mill, J., T. Tang, Z. Kaminsky, T. Khare, S. Yazdanpanah, L. Bouchard, P. Jia, A. Assadzadeh, J. Flanagan, A. Schumacher, S. C. Wang and A. Petronis (2008). "Epigenomic profiling reveals DNA-methylation changes associated with major psychosis." Am J Hum Genet **82**(3): 696-711.

Miller, K. (2006). "Bipolar disorder: etiology, diagnosis, and management." J Am Acad Nurse Pract **18**(8): 368-73.

Morgan, V. A., P. B. Mitchell and A. V. Jablensky (2005). "The epidemiology of bipolar disorder: sociodemographic, disability and service utilization data from the Australian National Study of Low Prevalence (Psychotic) Disorders." Bipolar Disord **7**(4): 326-37.

Morissette, J., A. Villeneuve, L. Bordeleau, D. Rochette, C. Laberge, B. Gagne, C. Laprise, G. Bouchard, M. Plante, L. Gobeil, E. Shink, J. Weissenbach and N. Barden (1999). "Genome-wide search for linkage of bipolar affective disorders in a very large pedigree derived from a homogeneous population in quebec points to a locus of major effect on chromosome 12q23-q24." Am J Med Genet **88**(5): 567-87.

Morton, N. (1955). "Sequential tests for the detection of linkage." American Journal of Human Genetics **7**: 277-318.

Morton, N. E. (1956). "The detection and estimation of linkage between the genes for elliptocytosis and the Rh blood type." Am J Hum Genet **8**(2): 80-96.

Morton, N. E. and A. Collins (1998). "Tests and estimates of allelic association in complex inheritance." Proc Natl Acad Sci U S A **95**(19): 11389-93.

Moskvina, V. and K. M. Schmidt (2006). "Susceptibility of biallelic haplotype and genotype frequencies to genotyping error." Biometrics **62**(4): 1116-23.

Muglia, P., A. Petronis, E. Mundo, S. Lander, T. Cate and J. L. Kennedy (2002). "Dopamine D4 receptor and tyrosine hydroxylase genes in bipolar disorder: evidence for a role of DRD4." Mol Psychiatry **7**(8): 860-6.

Muller, D. J., V. de Luca, T. Sicard, N. King, J. Strauss and J. L. Kennedy (2006). "Brain-derived neurotrophic factor (BDNF) gene and rapid-cycling bipolar disorder: family-based association study." Br J Psychiatry **189**: 317-23.

Muller, D. J., A. Serretti, T. Sicard, S. Tharmalingam, N. King, P. Artioli, L. Mandelli, C. Lorenzi and J. L. Kennedy (2007). "Further evidence of MAO-A gene variants associated with bipolar disorder." Am J Med Genet B Neuropsychiatr Genet **144B**(1): 37-40.

Murphy, W. J., E. Eizirik, W. E. Johnson, Y. P. Zhang, O. A. Ryder and S. J. O'Brien (2001). "Molecular phylogenetics and the origins of placental mammals." <u>Nature</u> **409**(6820): 614-8.

Murray, C. J. and A. D. Lopez (1997). "Global mortality, disability, and the contribution of risk factors: Global Burden of Disease Study." <u>Lancet</u> **349**(9063): 1436-42.

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

Mynett-Johnson, L., C. Kealey, E. Claffey, D. Curtis, L. Bouchier-Hayes, C. Powell and P. McKeon (2000). "Multimarkerhaplotypes within the serotonin transporter gene suggest evidence of an association with bipolar disorder." <u>Am J Med Genet</u> **96**(6): 845-9.

Mynett-Johnson, L. A., V. E. Murphy, E. Claffey, D. C. Shields and P. McKeon (1998). "Preliminary evidence of an association between bipolar disorder in females and the catechol-O-methyltransferase gene." <u>Psychiatr Genet</u> **8**(4): 221-5.

Nakamura, Y., M. Leppert, P. O'Connell, R. Wolff, T. Holm, M. Culver, C. Martin, E. Fujimoto, M. Hoff, E. Kumlin and et al. (1987). "Variable number of tandem repeat (VNTR) markers for human gene mapping." <u>Science</u> **235**(4796): 1616-22.

Neves-Pereira, M., E. Mundo, P. Muglia, N. King, F. Macciardi and J. L. Kennedy (2002). "The brain-derived neurotrophic factor gene confers susceptibility to bipolar disorder: evidence from a family-based association study." Am J Hum Genet **71**(3): 651-5.

Newberg, A. R., L. A. Catapano, C. A. Zarate and H. K. Manji (2008). "Neurobiology of bipolar disorder." Expert Rev Neurother **8**(1): 93-110.

Nielsen, D. and D. Zaykin (1999). "Novel tests for marker-disease association using the Collaborative Study on the Genetics of Alcoholism data." Genet Epidemiol **17 Suppl 1**: S265-70.

Nielsen, E. M., L. Hansen, M. Lajer, K. L. Andersen, S. M. Echwald, S. A. Urhammer, T. Hansen and O. Pedersen (2004). "A common polymorphism in the promoter of the IGF-I gene associates with increased fasting serum triglyceride levels in glucose-tolerant subjects." Clin Biochem **37**(8): 660-5.

Niesler, B., T. Flohr, M. M. Nothen, C. Fischer, M. Rietschel, E. Franzek, M. Albus, P. Propping and G. A. Rappold (2001). "Association between the 5' UTR variant C178T of the serotonin receptor gene HTR3A and bipolar affective disorder." Pharmacogenetics **11**(6): 471-5.

Nievergelt, C. M., D. F. Kripke, T. B. Barrett, E. Burg, R. A. Remick, A. D. Sadovnick, S. L. McElroy, P. E. Keck, Jr., N. J. Schork and J. R. Kelsoe (2006). "Suggestive evidence for association of the circadian genes

PERIOD3 and ARNTL with bipolar disorder." Am J Med Genet B
Neuropsychiatr Genet **141B**(3): 234-41.

Nishino, S., B. Ripley, S. Overeem, G. J. Lammers and E. Mignot (2000).
"Hypocretin (orexin) deficiency in human narcolepsy." Lancet
355(9197): 39-40.

Noga, J. T., K. Vladar and E. F. Torrey (2001). "A volumetric magnetic
resonance imaging study of monozygotic twins discordant for bipolar
disorder." Psychiatry Res **106**(1): 25-34.

Nugent, A. C., M. P. Milham, E. E. Bain, L. Mah, D. M. Cannon, S. Marrett, C.
A. Zarate, D. S. Pine, J. L. Price and W. C. Drevets (2006). "Cortical
abnormalities in bipolar disorder investigated with MRI and voxel-based
morphometry." Neuroimage **30**(2): 485-97.

Nutt, D. J. (2002). "The neuropharmacology of serotonin and noradrenaline in
depression." Int Clin Psychopharmacol **17 Suppl 1**: S1-12.

Nutt, D. J. (2006). "The role of dopamine and norepinephrine in depression and
antidepressant treatment." J Clin Psychiatry **67 Suppl 6**: 3-8.

O'Mahony, E., A. Corvin, R. O'Connell, C. Comerford, B. Larsen, R. Jones, F.
McCandless, G. Kirov, A. G. Cardno, N. Craddock and M. Gill (2002).
"Sibling pairs with affective disorders: resemblance of demographic and
clinical features." Psychol Med **32**(1): 55-61.

Ohtsuki, T., H. Ishiguro, S. D. Detera-Wadleigh, T. Toyota, H. Shimizu, K. Yamada, K. Yoshitsugu, E. Hattori, T. Yoshikawa and T. Arinami (2002). "Association between serotonin 4 receptor gene polymorphisms and bipolar disorder in Japanese case-control samples and the NIMH Genetics Initiative Bipolar Pedigrees." Mol Psychiatry **7**(9): 954-61.

Okada, T., R. Hashimoto, T. Numakawa, Y. Iijima, A. Kosuga, M. Tatsumi, K. Kamijima, T. Kato and H. Kunugi (2006). "A complex polymorphic region in the brain-derived neurotrophic factor (BDNF) gene confers susceptibility to bipolar disorder and affects transcriptional activity." Mol Psychiatry **11**(7): 695-703.

Ossowska, K., M. Pietraszek, J. Wardas and S. Wolfarth (2004). "Potential antipsychotic and extrapyramidal effects of (R,S)-3,4-dicarboxyphenylglycine [(R,S)-3,4-DCPG], a mixed AMPA antagonist/mGluR8 agonist." Pol J Pharmacol **56**(3): 295-304.

Ota, T., Y. Suzuki, T. Nishikawa, T. Otsuki, T. Sugiyama, R. Irie, A. Wakamatsu, K. Hayashi, H. Sato, K. Nagai, K. Kimura, H. Makita, M. Sekine, M. Obayashi, T. Nishi, T. Shibahara, T. Tanaka, S. Ishii, J. Yamamoto, K. Saito, Y. Kawai, Y. Isono, Y. Nakamura, K. Nagahari, K. Murakami, T. Yasuda, T. Iwayanagi, M. Wagatsuma, A. Shiratori, H. Sudo, T. Hosoiri, Y. Kaku, H. Kodaira, H. Kondo, M. Sugawara, M. Takahashi, K. Kanda, T. Yokoi, T. Furuya, E. Kikkawa, Y. Omura, K. Abe, K. Kamihara, N. Katsuta, K. Sato, M. Tanikawa, M. Yamazaki, K. Ninomiya, T. Ishibashi, H. Yamashita, K. Murakawa, K. Fujimori, H.

Tanai, M. Kimata, M. Watanabe, S. Hiraoka, Y. Chiba, S. Ishida, Y. Ono, S. Takiguchi, S. Watanabe, M. Yosida, T. Hotuta, J. Kusano, K. Kanehori, A. Takahashi-Fujii, H. Hara, T. O. Tanase, Y. Nomura, S. Togiya, F. Komai, R. Hara, K. Takeuchi, M. Arita, N. Imose, K. Musashino, H. Yuuki, A. Oshima, N. Sasaki, S. Aotsuka, Y. Yoshikawa, H. Matsunawa, T. Ichihara, N. Shiohata, S. Sano, S. Moriya, H. Momiyama, N. Satoh, S. Takami, Y. Terashima, O. Suzuki, S. Nakagawa, A. Senoh, H. Mizoguchi, Y. Goto, F. Shimizu, H. Wakebe, H. Hishigaki, T. Watanabe, A. Sugiyama, M. Takemoto, B. Kawakami, M. Yamazaki, K. Watanabe, A. Kumagai, S. Itakura, Y. Fukuzumi, Y. Fujimori, M. Komiyama, H. Tashiro, A. Tanigami, T. Fujiwara, T. Ono, K. Yamada, Y. Fujii, K. Ozaki, M. Hirao, Y. Ohmori, A. Kawabata, T. Hikiji, N. Kobatake, H. Inagaki, Y. Ikema, S. Okamoto, R. Okitani, T. Kawakami, S. Noguchi, T. Itoh, K. Shigeta, T. Senba, K. Matsumura, Y. Nakajima, T. Mizuno, M. Morinaga, M. Sasaki, T. Togashi, M. Oyama, H. Hata, M. Watanabe, T. Komatsu, J. Mizushima-Sugano, T. Satoh, Y. Shirai, Y. Takahashi, K. Nakagawa, K. Okumura, T. Nagase, N. Nomura, H. Kikuchi, Y. Masuho, R. Yamashita, K. Nakai, T. Yada, Y. Nakamura, O. Ohara, T. Isogai and S. Sugano (2004). "Complete sequencing and characterization of 21,243 full-length human cDNAs." Nat Genet **36**(1): 40-5.

Otani, K., H. Ujike, Y. Tanaka, Y. Morita, T. Katsu, A. Nomura, N. Uchida, T. Hamamura, Y. Fujiwara and S. Kuroda (2005). "The GABA type A

receptor alpha5 subunit gene is associated with bipolar I disorder." Neurosci Lett **381**(1-2): 108-13.

Ott, J. (1974). "Estimation of the recombination fraction in human pedigrees: efficient computation of the likelihood for human linkage studies." Am J Hum Genet **26**(5): 588-97.

Ott, J. (1977). "Counting methods (EM algorithm) in human pedigree analysis: linkage and segregation analysis." Ann Hum Genet **40**(4): 443-54.

Pae, C. U., H. S. Yu, D. Amann, J. J. Kim, C. U. Lee, S. J. Lee, T. Y. Jun, C. Lee, I. H. Paik, A. A. Patkar and B. Lerer (2008). "Association of the trace amine associated receptor 6 (TAAR6) gene with schizophrenia and bipolar disorder in a Korean case control sample." J Psychiatr Res **42**(1): 35-40.

Papadimitriou, G. N., D. G. Dikeos, G. Karadima, D. Avramopoulos, E. G. Daskalopoulou, D. Vassilopoulos and C. N. Stefanis (1998). "Association between the GABA(A) receptor alpha5 subunit gene locus (GABRA5) and bipolar affective disorder." Am J Med Genet **81**(1): 73-80.

Papolos, D. F., S. Veit, G. L. Faedda, T. Saito and H. M. Lachman (1998). "Ultra-ultra rapid cycling bipolar disorder is associated with the low activity catecholamine-O-methyltransferase allele." Mol Psychiatry **3**(4): 346-9.

Perez de Castro, I., J. Santos, P. Torres, G. Visedo, J. Saiz-Ruiz, C. Llinares and J. Fernandez-Piqueras (1995). "A weak association between TH and DRD2 genes and bipolar affective disorder in a Spanish sample." J Med Genet **32**(2): 131-4.

Pericak-Vance, Ed. (1998). Basic Concepts in Linkage Analysis. Approaches to Genes mapping in Complex Human Diseases. New York, Wiley-Liss.

Perlis, R. H., J. A. Welge, L. A. Vornik, R. M. Hirschfeld and P. E. Keck, Jr. (2006). "Atypical antipsychotics in the treatment of mania: a meta-analysis of randomized, placebo-controlled trials." J Clin Psychiatry **67**(4): 509-16.

Petryshen, T., A. Kirby and M. Ainscow. (2003). "LocusView unpublished software from Broad Institute." from www.broad.mit.edu/mpg/locusview/.

Petryshen, T. L., P. C. Sabeti, K. A. Aldinger, B. Fry, J. B. Fan, S. F. Schaffner, S. G. Waggoner, A. R. Tahl and P. Sklar (2009). "Population genetic study of the brain-derived neurotrophic factor (BDNF) gene." Mol Psychiatry.

Peyron, C., J. Faraco, W. Rogers, B. Ripley, S. Overeem, Y. Charnay, S. Nevsimalova, M. Aldrich, D. Reynolds, R. Albin, R. Li, M. Hungs, M. Pedrazzoli, M. Padigaru, M. Kucherlapati, J. Fan, R. Maki, G. J. Lammers, C. Bouras, R. Kucherlapati, S. Nishino and E. Mignot (2000).

"A mutation in a case of early onset narcolepsy and a generalized absence of hypocretin peptides in human narcoleptic brains." Nat Med **6**(9): 991-7.

Phelps, L. E., N. Brutsche, J. R. Moral, D. A. Luckenbaugh, H. K. Manji and C. A. Zarate, Jr. (2009). "Family history of alcohol dependence and initial antidepressant response to an N-methyl-D-aspartate antagonist." Biol Psychiatry **65**(2): 181-4.

Porteous, D. (2008). "Genetic causality in schizophrenia and bipolar disorder: out with the old and in with the new." Curr Opin Genet Dev **18**(3): 229-34.

Prata, D., G. Breen, S. Osborne, J. Munro, D. St Clair and D. Collier (2008). "Association of DAO and G72(DAOA)/G30 genes with bipolar affective disorder." Am J Med Genet B Neuropsychiatr Genet **147B**(6): 914-7.

Preisig, M., F. Bellivier, B. T. Fenton, P. Baud, A. Berney, P. Courtet, P. Hardy, J. Golaz, M. Leboyer, J. Mallet, M. L. Matthey, D. Mouthon, E. Neidhart, M. Nosten-Bertrand, E. Stadelmann-Dubuis, J. Guimon, F. Ferrero, C. Buresi and A. Malafosse (2000). "Association between bipolar disorder and monoamine oxidase A gene polymorphisms: results of a multicenter study." Am J Psychiatry **157**(6): 948-55.

Puglielli, L. (2008). "Aging of the brain, neurotrophin signaling, and Alzheimer's disease: is IGF1-R the common culprit?" Neurobiol Aging **29**(6): 795-811.

Rafaelsen, O. J. (1981). "[Biological perturbations in affective disorders in man]." Encephale **7**(4 Suppl): 413-9.

Ranade, S. S., H. Mansour, J. Wood, K. V. Chowdari, L. K. Brar, D. J. Kupfer and V. L. Nimgaonkar (2003). "Linkage and association between serotonin 2A receptor gene polymorphisms and bipolar I disorder." Am J Med Genet B Neuropsychiatr Genet **121B**(1): 28-34.

Redon, R., S. Ishikawa, K. R. Fitch, L. Feuk, G. H. Perry, T. D. Andrews, H. Fiegler, M. H. Shapero, A. R. Carson, W. Chen, E. K. Cho, S. Dallaire, J. L. Freeman, J. R. Gonzalez, M. Gratacos, J. Huang, D. Kalaitzopoulos, D. Komura, J. R. MacDonald, C. R. Marshall, R. Mei, L. Montgomery, K. Nishimura, K. Okamura, F. Shen, M. J. Somerville, J. Tchinda, A. Valsesia, C. Woodwark, F. Yang, J. Zhang, T. Zerjal, J. Zhang, L. Armengol, D. F. Conrad, X. Estivill, C. Tyler-Smith, N. P. Carter, H. Aburatani, C. Lee, K. W. Jones, S. W. Scherer and M. E. Hurles (2006). "Global variation in copy number in the human genome." Nature **444**(7118): 444-54.

Reese, M. G., F. H. Eeckman, D. Kulp and D. Haussler (1997). "Improved splice site detection in Genie." J Comput Biol **4**(3): 311-23.

Reif, A., A. Strobel, C. P. Jacob, S. Herterich, C. M. Freitag, T. Topner, R. Mossner, S. Fritzen, A. Schmitt and K. P. Lesch (2006). "A NOS-III haplotype that includes functional polymorphisms is associated with bipolar disorder." Int J Neuropsychopharmacol **9**(1): 13-20.

Reinhardt, R. R. and C. A. Bondy (1994). "Insulin-like growth factors cross the blood-brain barrier." Endocrinology **135**(5): 1753-61.

Risch, N. (1988). "A new statistical test for linkage heterogeneity." Am J Hum Genet **42**(2): 353-64.

Robinson, J. L., E. S. Monkul, D. Tordesillas-Gutierrez, C. Franklin, C. E. Bearden, P. T. Fox and D. C. Glahn (2008). "Fronto-limbic circuitry in euthymic bipolar disorder: evidence for prefrontal hyperactivation." Psychiatry Res **164**(2): 106-13.

Roche, S., F. Cassidy, C. Zhao, J. Badger, E. Claffey, L. Mooney, C. Delaney, S. Dobrin and P. McKeon (2007). "Candidate gene analysis of 21q22: support for S100B as a susceptibility gene for bipolar affective disorder with psychosis." Am J Med Genet B Neuropsychiatr Genet **144B**(8): 1094-6.

Ross, J., W. Berrettini, W. Coryell, E. S. Gershon, J. A. Badner, J. R. Kelsoe, M. G. McInnis, F. J. McMahon, D. L. Murphy, J. I. Nurnberger, Jr., T. Foroud, J. P. Rice, W. B. Scheftner, P. Zandi, H. Edenberg and W. Byerley (2008). "Genome-wide parametric linkage analyses of 644 bipolar pedigrees suggest susceptibility loci at chromosomes 16 and 20." Psychiatr Genet **18**(4): 191-8.

Rotwein, P. (1991). "Structure, evolution, expression and regulation of insulin-like growth factors I and II." Growth Factors **5**(1): 3-18.

Rotwein, P., K. M. Pollock, D. K. Didier and G. G. Krivi (1986). "Organization and sequence of the human insulin-like growth factor I gene. Alternative RNA processing produces two insulin-like growth factor I precursor peptides." J Biol Chem **261**(11): 4828-32.

Rozen, S. and H. Skaletsky (2000). "Primer3 on the WWW for general users and for biologist programmers." Methods Mol Biol **132**: 365-86.

Russo, V. C., P. D. Gluckman, E. L. Feldman and G. A. Werther (2005). "The insulin-like growth factor system and its pleiotropic functions in brain." Endocr Rev **26**(7): 916-43.

Rybakowski, J. K., A. Borkowska, P. M. Czerski, M. Skibinska and J. Hauser (2003). "Polymorphism of the brain-derived neurotrophic factor gene and performance on a cognitive prefrontal test in bipolar patients." Bipolar Disord **5**(6): 468-72.

Sabunciyan, S., R. Yolken, C. M. Ragan, J. B. Potash, V. L. Nimgaonkar, F. Dickerson, I. C. Llenos and S. Weis (2007). "Polymorphisms in the homeobox gene OTX2 may be a risk factor for bipolar disorder." Am J Med Genet B Neuropsychiatr Genet **144B**(8): 1083-6.

Sajatovic, M. (2005). "Bipolar disorder: disease burden." Am J Manag Care **11**(3 Suppl): S80-4.

Sambrook, Fritsh and Maniatis (1989). Molecular Cloning A Laboratory Manual, CSH.

Sanacora, G., R. Gueorguieva, C. N. Epperson, Y. T. Wu, M. Appel, D. L. Rothman, J. H. Krystal and G. F. Mason (2004). "Subtype-specific alterations of gamma-aminobutyric acid and glutamate in patients with major depression." Arch Gen Psychiatry 61(7): 705-13.

Sandhu, M. S. (2005). "Insulin-like growth factor-I and risk of type 2 diabetes and coronary heart disease: molecular epidemiology." Endocr Dev 9: 44-54.

Santarelli, L., M. Saxe, C. Gross, A. Surget, F. Battaglia, S. Dulawa, N. Weisstaub, J. Lee, R. Duman, O. Arancio, C. Belzung and R. Hen (2003). "Requirement of hippocampal neurogenesis for the behavioral effects of antidepressants." Science 301(5634): 805-9.

Savitz, J., L. van der Merwe, M. Solms and R. Ramesar (2007). "A linkage and family-based association analysis of a potential neurocognitive endophenotype of bipolar disorder." Neuromolecular Med 9(2): 101-16.

Schosser, A., D. Gaysina, S. Cohen-Woods, P. C. Chow, L. Martucci, N. Craddock, A. Farmer, A. Korszun, C. Gunasinghe, J. Gray, L. Jones, F. Tozzi, J. Perry, P. Muglia, M. J. Owen, I. W. Craig and P. McGuffin (2009). "Association of DISC1 and TSNAX genes and affective disorders

in the depression case-control (DeCC) and bipolar affective case-control (BACCS) studies." <u>Mol Psychiatry</u>.

Schug, J. and G. Christian Overton (1997) "TESS: Transcription Element Search Software on the *www.cbil.upenn.edu/tess/* <u>Technical Report</u>

Schulze, T. G., S. D. Detera-Wadleigh, N. Akula, A. Gupta, L. Kassem, J. Steele, J. Pearl, J. Strohmaier, R. Breuer, M. Schwarz, P. Propping, M. M. Nothen, S. Cichon, J. Schumacher, M. Rietschel and F. J. McMahon (2009). "Two variants in Ankyrin 3 (ANK3) are independent genetic risk factors for bipolar disorder." <u>Mol Psychiatry</u> **14**(5): 487-91.

Schumacher, J., R. Kaneva, R. A. Jamra, G. O. Diaz, S. Ohlraun, V. Milanova, Y. A. Lee, F. Rivas, F. Mayoral, R. Fuerst, A. Flaquer, C. Windemuth, E. Gay, S. Sanz, M. J. Gonzalez, S. Gil, F. Cabaleiro, F. del Rio, F. Perez, J. Haro, C. Kostov, V. Chorbov, A. Nikolova-Hill, V. Stoyanova, G. Onchev, I. Kremensky, K. Strauch, T. G. Schulze, P. Nurnberg, W. Gaebel, A. Klimke, G. Auburger, T. F. Wienker, L. Kalaydjieva, P. Propping, S. Cichon, A. Jablensky, M. Rietschel and M. M. Nothen (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." <u>Am J Hum Genet</u> **77**(6): 1102-11.

Sebat, J., B. Lakshmi, J. Troge, J. Alexander, J. Young, P. Lundin, S. Maner, H. Massa, M. Walker, M. Chi, N. Navin, R. Lucito, J. Healy, J. Hicks, K. Ye, A. Reiner, T. C. Gilliam, B. Trask, N. Patterson, A. Zetterberg and M. Wigler (2004). "Large-scale copy number polymorphism in the human genome." <u>Science</u> **305**(5683): 525-8.

Segurado, R., S. D. Detera-Wadleigh, D. F. Levinson, C. M. Lewis, M. Gill, J. I. Nurnberger, Jr., N. Craddock, J. R. DePaulo, M. Baron, E. S. Gershon, J. Ekholm, S. Cichon, G. Turecki, S. Claes, J. R. Kelsoe, P. R. Schofield, R. F. Badenhop, J. Morissette, H. Coon, D. Blackwood, L. A. McInnes, T. Foroud, H. J. Edenberg, T. Reich, J. P. Rice, A. Goate, M. G. McInnis, F. J. McMahon, J. A. Badner, L. R. Goldin, P. Bennett, V. L. Willour, P. P. Zandi, J. Liu, C. Gilliam, S. H. Juo, W. H. Berrettini, T. Yoshikawa, L. Peltonen, J. Lonnqvist, M. M. Nothen, J. Schumacher, C. Windemuth, M. Rietschel, P. Propping, W. Maier, M. Alda, P. Grof, G. A. Rouleau, J. Del-Favero, C. Van Broeckhoven, J. Mendlewicz, R. Adolfsson, M. A. Spence, H. Luebbert, L. J. Adams, J. A. Donald, P. B. Mitchell, N. Barden, E. Shink, W. Byerley, W. Muir, P. M. Visscher, S. Macgregor, H. Gurling, G. Kalsi, A. McQuillin, M. A. Escamilla, V. I. Reus, P. Leon, N. B. Freimer, H. Ewald, T. A. Kruse, O. Mors, U. Radhakrishna, J. L. Blouin, S. E. Antonarakis and N. Akarsu (2003). "Genome scan meta-analysis of schizophrenia and bipolar disorder, part III: Bipolar disorder." <u>Am J Hum Genet</u> **73**(1): 49-62.

Selch, S., A. Strobel, J. Haderlein, J. Meyer, C. P. Jacob, A. Schmitt, K. P. Lesch and A. Reif (2007). "MLC1 polymorphisms are specifically associated

with periodic catatonia, a subgroup of chronic schizophrenia." <u>Biol Psychiatry</u> **61**(10): 1211-4.

Selvi, R. B. and T. K. Kundu (2009). "Reversible acetylation of chromatin: implication in regulation of gene expression, disease and therapeutics." <u>Biotechnol J</u> **4**(3): 375-90.

Serretti, A., P. N. Malitas, L. Mandelli, C. Lorenzi, C. Ploia, B. Alevizos, C. Nikolaou, F. Boufidou, G. N. Christodoulou and E. Smeraldi (2004). "Further evidence for a possible association between serotonin transporter gene and lithium prophylaxis in mood disorders." <u>Pharmacogenomics J</u> **4**(4): 267-73.

Severino, G., D. Congiu, C. Serreli, R. De Lisa, C. Chillotti, M. Del Zompo and M. P. Piccardi (2005). "A48G polymorphism in the D1 receptor genes associated with bipolar I disorder." <u>Am J Med Genet B Neuropsychiatr Genet</u> **134B**(1): 37-8.

Shaltiel, G., A. Shamir, J. Shapiro, D. Ding, E. Dalton, M. Bialer, A. J. Harwood, R. H. Belmaker, M. L. Greenberg and G. Agam (2004). "Valproate decreases inositol biosynthesis." <u>Biol Psychiatry</u> **56**(11): 868-74.

Sham, P. C., D. Curtis and C. J. MacLean (1996). "Likelihood ratio tests for linkage and linkage disequilibrium: asymptotic distribution and power." <u>Am J Hum Genet</u> **58**(5): 1093-6.

Shimatsu, A. and P. Rotwein (1987). "Mosaic evolution of the insulin-like growth factors. Organization, sequence, and expression of the rat insulin-like growth factor I gene." J Biol Chem **262**(16): 7894-900.

Shink, E., M. Harvey, M. Tremblay, B. Gagne, P. Belleau, C. Raymond, M. Labbe, M. P. Dube, R. G. Lafreniere and N. Barden (2005). "Analysis of microsatellite markers and single nucleotide polymorphisms in candidate genes for susceptibility to bipolar affective disorder in the chromosome 12Q24.31 region." Am J Med Genet B Neuropsychiatr Genet **135B**(1): 50-8.

Shink, E., J. Morissette, R. Sherrington and N. Barden (2005). "A genome-wide scan points to a susceptibility locus for bipolar disorder on chromosome 12." Mol Psychiatry **10**(6): 545-52.

Siuciak, J. A., C. Boylan, M. Fritsche, C. A. Altar and R. M. Lindsay (1996). "BDNF increases monoaminergic activity in rat brain following intracerebroventricular or intraparenchymal administration." Brain Res **710**(1-2): 11-20.

Siuciak, J. A., M. S. Clark, H. B. Rind, S. R. Whittemore and A. F. Russo (1998). "BDNF induction of tryptophan hydroxylase mRNA levels in the rat brain." J Neurosci Res **52**(2): 149-58.

Sklar, P., S. B. Gabriel, M. G. McInnis, P. Bennett, Y. M. Lim, G. Tsan, S. Schaffner, G. Kirov, I. Jones, M. Owen, N. Craddock, J. R. DePaulo and

E. S. Lander (2002). "Family-based association study of 76 candidate genes in bipolar disorder: BDNF is a potential risk locus. Brain-derived neutrophic factor." Mol Psychiatry **7**(6): 579-93.

Sklar, P., J. W. Smoller, J. Fan, M. A. Ferreira, R. H. Perlis, K. Chambert, V. L. Nimgaonkar, M. B. McQueen, S. V. Faraone, A. Kirby, P. I. de Bakker, M. N. Ogdie, M. E. Thase, G. S. Sachs, K. Todd-Brown, S. B. Gabriel, C. Sougnez, C. Gates, B. Blumenstiel, M. Defelice, K. G. Ardlie, J. Franklin, W. J. Muir, K. A. McGhee, D. J. MacIntyre, A. McLean, M. VanBeck, A. McQuillin, N. J. Bass, M. Robinson, J. Lawrence, A. Anjorin, D. Curtis, E. M. Scolnick, M. J. Daly, D. H. Blackwood, H. M. Gurling and S. M. Purcell (2008). "Whole-genome association study of bipolar disorder." Mol Psychiatry **13**(6): 558-69.

Smith, E. M. and R. C. Juvinall (1963). "Design Refinement of the Linkage Feeder." Arch Phys Med Rehabil **44**: 609-15.

Smyth, C., G. Kalsi, D. Curtis, J. Brynjolfsson, J. O'Neill, L. Rifkin, E. Moloney, P. Murphy, H. Petursson and H. Gurling (1997). "Two-locus admixture linkage analysis of bipolar and unipolar affective disorder supports the presence of susceptibility loci on chromosomes 11p15 and 21q22." Genomics **39**(3): 271-8.

Soares, J. and A. Young (2007). Bipolar Disorders Basic Mechanisms and Therapeutic Implications. New York, Informa Health Care.

Staden, R. (1996). "The Staden sequence analysis package." <u>Mol Biotechnol</u> **5**(3): 233-41.

Stankiewicz, P. and J. R. Lupski (2002). "Molecular-evolutionary mechanisms for genomic disorders." <u>Curr Opin Genet Dev</u> **12**(3): 312-9.

Stefansson, H., D. Rujescu, S. Cichon, O. P. Pietilainen, A. Ingason, S. Steinberg, R. Fossdal, E. Sigurdsson, T. Sigmundsson, J. E. Buizer-Voskamp, T. Hansen, K. D. Jakobsen, P. Muglia, C. Francks, P. M. Matthews, A. Gylfason, B. V. Halldorsson, D. Gudbjartsson, T. E. Thorgeirsson, A. Sigurdsson, A. Jonasdottir, A. Jonasdottir, A. Bjornsson, S. Mattiasdottir, T. Blondal, M. Haraldsson, B. B. Magnusdottir, I. Giegling, H. J. Moller, A. Hartmann, K. V. Shianna, D. Ge, A. C. Need, C. Crombie, G. Fraser, N. Walker, J. Lonnqvist, J. Suvisaari, A. Tuulio-Henriksson, T. Paunio, T. Toulopoulou, E. Bramon, M. Di Forti, R. Murray, M. Ruggeri, E. Vassos, S. Tosato, M. Walshe, T. Li, C. Vasilescu, T. W. Muhleisen, A. G. Wang, H. Ullum, S. Djurovic, I. Melle, J. Olesen, L. A. Kiemeney, B. Franke, C. Sabatti, N. B. Freimer, J. R. Gulcher, U. Thorsteinsdottir, A. Kong, O. A. Andreassen, R. A. Ophoff, A. Georgi, M. Rietschel, T. Werge, H. Petursson, D. B. Goldstein, M. M. Nothen, L. Peltonen, D. A. Collier, D. St Clair and K. Stefansson (2008). "Large recurrent microdeletions associated with schizophrenia." <u>Nature</u> **455**(7210): 232-6.

Strakowski, S. M., M. P. Delbello and C. M. Adler (2005). "The functional neuroanatomy of bipolar disorder: a review of neuroimaging findings." Mol Psychiatry **10**(1): 105-16.

Strausberg, R. L., E. A. Feingold, L. H. Grouse, J. G. Derge, R. D. Klausner, F. S. Collins, L. Wagner, C. M. Shenmen, G. D. Schuler, S. F. Altschul, B. Zeeberg, K. H. Buetow, C. F. Schaefer, N. K. Bhat, R. F. Hopkins, H. Jordan, T. Moore, S. I. Max, J. Wang, F. Hsieh, L. Diatchenko, K. Marusina, A. A. Farmer, G. M. Rubin, L. Hong, M. Stapleton, M. B. Soares, M. F. Bonaldo, T. L. Casavant, T. E. Scheetz, M. J. Brownstein, T. B. Usdin, S. Toshiyuki, P. Carninci, C. Prange, S. S. Raha, N. A. Loquellano, G. J. Peters, R. D. Abramson, S. J. Mullahy, S. A. Bosak, P. J. McEwan, K. J. McKernan, J. A. Malek, P. H. Gunaratne, S. Richards, K. C. Worley, S. Hale, A. M. Garcia, L. J. Gay, S. W. Hulyk, D. K. Villalon, D. M. Muzny, E. J. Sodergren, X. Lu, R. A. Gibbs, J. Fahey, E. Helton, M. Ketteman, A. Madan, S. Rodrigues, A. Sanchez, M. Whiting, A. Madan, A. C. Young, Y. Shevchenko, G. G. Bouffard, R. W. Blakesley, J. W. Touchman, E. D. Green, M. C. Dickson, A. C. Rodriguez, J. Grimwood, J. Schmutz, R. M. Myers, Y. S. Butterfield, M. I. Krzywinski, U. Skalska, D. E. Smailus, A. Schnerch, J. E. Schein, S. J. Jones and M. A. Marra (2002). "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences." Proc Natl Acad Sci U S A **99**(26): 16899-903.

Sun, G., J. Gagnon, Y. C. Chagnon, L. Perusse, J. P. Despres, A. S. Leon, J. H. Wilmore, J. S. Skinner, I. Borecki, D. C. Rao and C. Bouchard (1999).

"Association and linkage between an insulin-like growth factor-1 gene polymorphism and fat free mass in the HERITAGE Family Study." Int J Obes Relat Metab Disord **23**(9): 929-35.

Szczepankiewicz, A., M. Skibinska, J. Hauser, A. Slopien, A. Leszczynska-Rodziewicz, P. Kapelski, M. Dmitrzak-Weglarz, P. M. Czerski and J. K. Rybakowski (2006). "Association analysis of the GSK-3beta T-50C gene polymorphism with schizophrenia and bipolar disorder." Neuropsychobiology **53**(1): 51-6.

Tamura, Y., H. Kunugi, J. Ohashi and H. Hohjoh (2007). "Epigenetic aberration of the human REELIN gene in psychiatric disorders." Mol Psychiatry **12**(6): 519, 593-600.

Tecott, L. H. (2003). "The genes and brains of mice and men." Am J Psychiatry **160**(4): 646-56.

Terry, J., M. Lopez-Larson and J. A. Frazier (2009). "Magnetic resonance imaging studies in early onset bipolar disorder: an updated review." Child Adolesc Psychiatr Clin N Am **18**(2): 421-39, ix-x.

Terwilliger, J. D. and J. Ott (1993). "A novel polylocus method for linkage analysis using the lod-score or affected sib-pair method." Genet Epidemiol **10**(6): 477-82.

Thomson, P. A., N. R. Wray, J. K. Millar, K. L. Evans, S. L. Hellard, A. Condie, W. J. Muir, D. H. Blackwood and D. J. Porteous (2005). "Association between the TRAX/DISC locus and both bipolar disorder and schizophrenia in the Scottish population." Mol Psychiatry **10**(7): 657-68, 616.

Thomson, P. A., N. R. Wray, A. M. Thomson, D. R. Dunbar, M. A. Grassie, A. Condie, M. T. Walker, D. J. Smith, D. J. Pulford, W. Muir, D. H. Blackwood and D. J. Porteous (2005). "Sex-specific association between bipolar affective disorder in women and GPR50, an X-linked orphan G protein-coupled receptor." Mol Psychiatry **10**(5): 470-8.

Tienari, P. J., J. D. Terwilliger, J. Ott, J. Palo and L. Peltonen (1994). "Two-locus linkage analysis in multiple sclerosis (MS)." Genomics **19**(2): 320-5.

Torkamani, A., E. J. Topol and N. J. Schork (2008). "Pathway analysis of seven common diseases assessed by genome-wide association." Genomics **92**(5): 265-72.

Trejo, J. L., E. Carro, C. Lopez-Lopez and I. Torres-Aleman (2004). "Role of serum insulin-like growth factor I in mammalian brain aging." Growth Horm IGF Res **14 Suppl A**: S39-43.

Trejo, J. L., E. Carro and I. Torres-Aleman (2001). "Circulating insulin-like growth factor I mediates exercise-induced increases in the number of new neurons in the adult hippocampus." <u>J Neurosci</u> **21**(5): 1628-34.

Vaessen, N., P. Heutink, J. A. Janssen, J. C. Witteman, L. Testers, A. Hofman, S. W. Lamberts, B. A. Oostra, H. A. Pols and C. M. van Duijn (2001). "A polymorphism in the gene for IGF-I: functional properties and risk for type 2 diabetes and myocardial infarction." <u>Diabetes</u> **50**(3): 637-42.

Vaessen, N., J. A. Janssen, P. Heutink, A. Hofman, S. W. Lamberts, B. A. Oostra, H. A. Pols and C. M. van Duijn (2002). "Association between genetic variation in the gene for insulin-like growth factor-I and low birthweight." <u>Lancet</u> **359**(9311): 1036-7.

Van Den Bogaert, A., K. Sleegers, S. De Zutter, L. Heyrman, K. F. Norrback, R. Adolfsson, C. Van Broeckhoven and J. Del-Favero (2006). "Association of brain-specific tryptophan hydroxylase, TPH2, with unipolar and bipolar disorder in a Northern Swedish, isolated population." <u>Arch Gen Psychiatry</u> **63**(10): 1103-10.

Vandesompele, J., K. De Preter, F. Pattyn, B. Poppe, N. Van Roy, A. De Paepe and F. Speleman (2002). "Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes." <u>Genome Biol</u> **3**(7): 1-11.

Vazza, G., C. Bertolin, E. Scudellaro, A. Vettori, F. Boaretto, S. Rampinelli, G. De Sanctis, G. Perini, P. Peruzzi and M. L. Mostacciuolo (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.

Venken, T., S. Claes, S. Sluijs, A. D. Paterson, C. van Duijn, R. Adolfsson, J. Del-Favero and C. Van Broeckhoven (2005). "Genomewide scan for affective disorder susceptibility Loci in families of a northern Swedish isolated population." Am J Hum Genet **76**(2): 237-48.

Verma, R., S. Kubendran, S. K. Das, S. Jain and S. K. Brahmachari (2005). "SYNGR1 is associated with schizophrenia and bipolar disorder in southern India." J Hum Genet **50**(12): 635-40.

Verma, R., M. Mukerji, D. Grover, B. R. C, S. K. Das, S. Kubendran, S. Jain and S. K. Brahmachari (2005). "MLC1 gene is associated with schizophrenia and bipolar disorder in Southern India." Biol Psychiatry **58**(1): 16-22.

Vincze, I., N. Perroud, C. Buresi, P. Baud, F. Bellivier, B. Etain, C. Fournier, F. Karege, M. L. Matthey, M. Preisig, M. Leboyer and A. Malafosse (2008). "Association between brain-derived neurotrophic factor gene and a severe form of bipolar disorder, but no interaction with the serotonin transporter gene." Bipolar Disord **10**(5): 580-7.

Walss-Bass, C., A. P. Montero, R. Armas, A. Dassori, S. A. Contreras, W. Liu, R. Medina, D. Levinson, M. Pereira, I. Atmella, L. NeSmith, R. Leach, L. Almasy, H. Raventos and M. A. Escamilla (2006). "Linkage disequilibrium analyses in the Costa Rican population suggests discrete gene loci for schizophrenia at 8p23.1 and 8q13.3." Psychiatr Genet **16**(4): 159-68.

Wang, D. G., J. B. Fan, C. J. Siao, A. Berno, P. Young, R. Sapolsky, G. Ghandour, N. Perkins, E. Winchester, J. Spencer, L. Kruglyak, L. Stein, L. Hsie, T. Topaloglou, E. Hubbell, E. Robinson, M. Mittmann, M. S. Morris, N. Shen, D. Kilburn, J. Rioux, C. Nusbaum, S. Rozen, T. J. Hudson, R. Lipshutz, M. Chee and E. S. Lander (1998). "Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome." Science **280**(5366): 1077-82.

Wang, H. Y., G. P. Johnson and E. Friedman (2001). "Lithium treatment inhibits protein kinase C translocation in rat brain cortex." Psychopharmacology (Berl) **158**(1): 80-6.

Wang, L., X. Liu and R. H. Lenox (2001). "Transcriptional down-regulation of MARCKS gene expression in immortalized hippocampal cells by lithium." J Neurochem **79**(4): 816-25.

Washizuka, S., K. Iwamoto, A. A. Kazuno, C. Kakiuchi, K. Mori, M. Kametani, K. Yamada, H. Kunugi, O. Tajima, T. Akiyama, S. Nanko, T. Yoshikawa and T. Kato (2004). "Association of mitochondrial complex I subunit

gene NDUFV2 at 18p11 with bipolar disorder in Japanese and the National Institute of Mental Health pedigrees." <u>Biol Psychiatry</u> **56**(7): 483-9.

Watson, S., J. M. Thompson, J. C. Ritchie, I. Nicol Ferrier and A. H. Young (2006). "Neuropsychological impairment in bipolar disorder: the relationship with glucocorticoid receptor function." <u>Bipolar Disord</u> **8**(1): 85-90.

Weber, J. L. and P. E. May (1989). "Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction." <u>Am J Hum Genet</u> **44**(3): 388-96.

Weeks, D. E. and K. Lange (1988). "The affected-pedigree-member method of linkage analysis." <u>Am J Hum Genet</u> **42**(2): 315-26.

Weeks, D. E. and K. Lange (1992). "A multilocus extension of the affected-pedigree-member method of linkage analysis." <u>Am J Hum Genet</u> **50**(4): 859-68.

Weiss, L. A., Y. Shen, J. M. Korn, D. E. Arking, D. T. Miller, R. Fossdal, E. Saemundsen, H. Stefansson, M. A. Ferreira, T. Green, O. S. Platt, D. M. Ruderfer, C. A. Walsh, D. Altshuler, A. Chakravarti, R. E. Tanzi, K. Stefansson, S. L. Santangelo, J. F. Gusella, P. Sklar, B. L. Wu and M. J. Daly (2008). "Association between microdeletion and microduplication at 16p11.2 and autism." <u>N Engl J Med</u> **358**(7): 667-75.

Weller, A. E., J. P. Dahl, F. W. Lohoff, T. N. Ferraro and W. H. Berrettini (2006). "Analysis of variations in the NAPG gene on chromosome 18p11 in bipolar disorder." Psychiatr Genet **16**(1): 3-8.

Wender, P. H., S. S. Kety, D. Rosenthal, F. Schulsinger, J. Ortmann and I. Lunde (1986). "Psychiatric disorders in the biological and adoptive families of adopted individuals with affective disorders." Arch Gen Psychiatry **43**(10): 923-9.

Williams, N. M., E. K. Green, S. Macgregor, S. Dwyer, N. Norton, H. Williams, R. Raybould, D. Grozeva, M. Hamshere, S. Zammit, L. Jones, A. Cardno, G. Kirov, I. Jones, M. C. O'Donovan, M. J. Owen and N. Craddock (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, V. Chopra, B. L. Melnyk, W. G. Honer and R. A. Holt (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.

Wolfensohn, S. and M. Lloyd (2003). Handbook of Laboratory Animal Management and Welfare, Blackwel Publishing.

WTCCC (2007). "Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls." Nature **447**(7145): 661-78.

Xie, W., W. Ted Brown and R. B. Denman (2008). "Translational regulation by non-protein-coding RNAs: different targets, common themes." Biochem Biophys Res Commun **373**(4): 462-6.

Xu, C., P. P. Li, R. G. Cooke, S. V. Parikh, K. Wang, J. L. Kennedy and J. J. Warsh (2009). "TRPM2 variants and bipolar disorder risk: confirmation in a family-based association study." Bipolar Disord **11**(1): 1-10.

Xu, C., F. Macciardi, P. P. Li, I. S. Yoon, R. G. Cooke, B. Hughes, S. V. Parikh, R. S. McIntyre, J. L. Kennedy and J. J. Warsh (2006). "Association of the putative susceptibility gene, transient receptor potential protein melastatin type 2, with bipolar disorder." Am J Med Genet B Neuropsychiatr Genet **141**(1): 36-43.

Xu, J., Y. Liu, P. Wang, S. Li, Y. Wang, J. Li, D. Zhou, Z. Chen, T. Zhao, T. Wang, H. Xu, Y. Yang, G. Feng, L. He and L. Yu (2009). "Positive association between the brain-derived neurotrophic factor (BDNF) gene and bipolar disorder in the Han Chinese population." Am J Med Genet B Neuropsychiatr Genet.

Yang, S., K. Wang, B. Gregory, W. Berrettini, L. S. Wang, H. Hakonarson and M. Bucan (2009). "Genomic landscape of a three-generation pedigree segregating affective disorder." PLoS ONE **4**(2): e4474.

Yildiz-Yesiloglu, A. and D. P. Ankerst (2006). "Neurochemical alterations of the brain in bipolar disorder and their implications for pathophysiology: a

systematic review of the in vivo proton magnetic resonance spectroscopy findings." Prog Neuropsychopharmacol Biol Psychiatry **30**(6): 969-95.

Young, L. T., P. P. Li, A. Kamble, K. P. Siu and J. J. Warsh (1994). "Mononuclear leukocyte levels of G proteins in depressed patients with bipolar disorder or major depressive disorder." Am J Psychiatry **151**(4): 594-6.

Young, L. T., P. P. Li, S. J. Kish, K. P. Siu and J. J. Warsh (1991). "Postmortem cerebral cortex Gs alpha-subunit levels are elevated in bipolar affective disorder." Brain Res **553**(2): 323-6.

Zhang, J., X. Li, Y. Wang, J. Ji, F. Yang, G. Feng, P. Wan, K. Lindpaintner, L. He and G. He (2009). "Association study on the mitochondrial gene NDUFV2 and bipolar disorder in the Chinese Han population." J Neural Transm **116**(3): 357-61.

Zhang, Z., K. Lindpaintner, R. Che, Z. He, P. Wang, P. Yang, G. Feng, L. He and Y. Shi (2009). "The Val/Met functional polymorphism in COMT confers susceptibility to bipolar disorder: evidence from an association study and a meta-analysis." J Neural Transm.

Zhao, J. H., D. Curtis and P. C. Sham (2000). "Model-free analysis and permutation tests for allelic associations." Hum Hered **50**(2): 133-9.

Zhao, J. H., S. Lissarrague, L. Essioux and P. C. Sham (2002). "GENECOUNTING: haplotype analysis with missing genotypes." Bioinformatics **18**(12): 1694-5.

Zimmermann, N., J. L. Colyer, L. E. Koch and M. E. Rothenberg (2005). "Analysis of the CCR3 promoter reveals a regulatory region in exon 1 that binds GATA-1." BMC Immunol **6**(1): 7.

8 Appendix

8.1 Primers and Probes Sequences for Slynar Experiments

8.1.1 RT-PCR

Primers sequences used for both mouse and human cell lines RT-PCR experiments.

Table 8.1 Slynar primer sequences for RT-PCR

	Forward Primers	Reverse Primers
Slynar Mouse Exon 3 to 4	CTGCCAGATCTGATTTTCTGC	GGATGGGAAAGAACAGAAGG
Slynar Human Exon 3 to 4	ATGGAGCCTGTTGAGCCACAG	GCTGCTAAGGAACTCAGTCTG

The clone sequence of the 220 bp band found in the mouse was:

CGGCCGCGAGCTCGGGCCCACACGTGTGGTCTAGAGCTAGCCTAGGC

TCGAGAAGCTTGTCGACGAATTCAGATTGGATGGGAAAGAACAGAAG

GCAATAGGAATGTCAGTCACTGCCCATGGACTGGGAGGACAGCAAAA

CAAAAATCCAGCAGCTGACATGGTATCTATAAGATGATGTCCAACAC

AACCCTCTGGCTACCCTGGGCCTCATCAGCAGTGAGAAGTCATGGAT

CATGCATTCTCAATGACCTTCCAGAAGACCAGGCAGAGGGCAGAAAA

TCAGATCTGCAGAATCACGAATTCTGGATCCATACGTAACGCGTCTGC

AGCATGCGTGG

8.1.2 cDNA Library Screening

Primers sequences used for Slynar's detection on the cDNA library. The table

shows the primers used for both mouse and human cDNA library screenings.

Table 8.2 Slynar primers used for cDNA library screening

	Forward Primers	Reverse Primers
Slynar Mouse Exon 3 to 4	CTGCCAGATCTGATTTTCTGC	GGATGGGAAAGAACAGAAGG
Slynar Human Exon 3 to 4	ATGGAGCCTGTTGAGCCACAG	GCTGCTAAGGAACTCAGTCTG
Slynar_AK Human Exon 1 to 3	TCTCTGTGCCTGAGTCGTTG	AGGGAGAAGTCGATTCAGCA
Slynar_DA7 Human Exon 1 to 2	CGGTTCAGAAAAGTTGTCAGG	AAGCCAACACCTCTCTCCTT

8.1.3 RACE

The table lists the sequences for the gene specific primers used for 3' and 5'

RACE experiments, for both human and mouse.

Table 8.3 Slynar primers used for RACE

	5' RACE Gene Specific Primers
Slynar Mouse Exon 4	GGATGGGAAAGAACAGAAGG
Slynar Mouse Exon 3a	GCAGAGGGCAGAAAATCAGA
Slynar Mouse Exon 3b	TTCCAGAAGACCAGGCAGAG
Slynar Mouse Exon 2	AAGGTATAATGTGCAGGCTTTTT
Slynar Human Exon 4a	GCTGCTAAGGAACTCAGTCTG
Slynar Human Exon 4b	TTCCCCAACAGTGAAAATTCGTG
Slynar Human Exon 3	AGGGAGAAGTCGATTCAGCA
	3'RACE Gene Specific Primers
Slynar Human Exon 1a	GGAGTTTCCGTGCCAGAG
Slynar Human Exon 1b	CGGTTCAGAAAAGTTGTCAGG
Slynar_AK Human Exon 1a	CACACGGCTACATTTCGGTA
Slynar_AK Human Exon 1b	TCTCTGTGCCTGAGTCGTTG

Sequences from the clones obtained from human 3' and 5' RACE:

Human 3' RACE clone B

TTTTTTTTTTTTTTTTTTAGAATGCAGGATTTTAATTCTAAGGGAAAG
GAAAATTCTCCATGTATCTAATCTAAAAAAAACATGGAAGTCATCAT
GAGTTTGTTTCTGGGGGATGATATCAGATGAAAATGGGCCCTTAGGT
AGTAGGAGGTGGCAGGTCTCAGTTCAGTGTTTACACATGGCTCCACG
CAGAGCTACAGACTGAGTTCCTTAGCAGCTTCCGTAGGAGGTCCTTTG
AAGAAGGGCTGCTCTGCATTCGATATGTAGAAACTTGTGGAGAAGAT
GAGGAAACTGAGGAAAACCACCAATGCAATCCACAATCCAATCCCTT
GACTCTCTGTTTTCCCCAACAGTGAAAATTCGTGTACAGAGTATTCCC
CAATGACTTCTCTCAGGCAATGTAGAAAGGGAGAAGTCGATTCAGCA
TTGACTAGTTGGTACTGTGGCTCAACAGGCTCCATCTCTGAAAGTGTT
GTTTCTTCCTTTTCCGGAGCTGGAACTGCCTGTCCGCGTGGACCCTCTC
CTCTGGCACGGAAACTCC

Human 3' RACE clone B

TGGAGTTTCCGTGCCAGAGGAGCGGGTCCACGCGGACAGGCAGTTCC
AGCTCCGGAGAGTCAAGGGATTGGATTGTGGATTGCATTGGTGGTTTT
CCTCAGTTTCCTCATCTTCTCCACAAGTTTCTACATATCGAATGCAGA
GCAGCCCTTCTTCAAAGGACCTCCTACGGAAGCTGCTAAGGAACTCA
GTCTGTAGCTCTGCGTGGAGCCATGTGTAAACACTGAACTGAGACCT
GCCACCTCCTACTACCTAAGGGCCCATTTTCATCTGATATCATCCCCC
AGAAACAAACTCATGATGACTTCCATGTTTTTTTTAGATTAGATACAT
GGAGAATTTTCCTTTCCCTTAGAATTAAAATCCTGCATTCTAATTTTCA
AAAAAAAAAAAAAAA

Human 5' RACE clone A

CGAGGTTAGGGGTTATGTGTGAGATAGTATGCCAGCCTTTTTTTTTTTT
TTTTTTTTTTTTTTTTGTCCAGCGCGAGCGTTCGTATCGGATCCGAA
TTGTGATTGGCCACGCGGCGACTAGTACGGGGGGGGGGGGGGCACAAA
GGGCTGAACTTGCAGGAGCCCCGCTTCTCAACCTCCATCGCTTTATGA
ATGAATTCACTAGAAAGAAAAGGAAGAATCAACACTTTCAAAGATGG
AGCCTGTTGAGCCACAGTACCAACTAGTCAATGCTGAATCGACTTCTC
CCTTTCTACATTGCCTGAAAAAAGTCATTGGGGAATACTTTGTACTAG
TCGACGCGTGGCCA

Human 5' RACE clone D

CCTCCTCGGGCTCGCTATCTGTTTGTTTTAGATGCAAAGCCAAAAATT
TTTTTTTTTTTTTTTTTTTATTTTTTTCTCTTACTCTAATGTTCTTATCTG
ATCCAGAATTCCTGATTGGTAGATTGCAACAAAATATTCTTGAATTTA
GACATATATATTGAATTTTATAAAGCCACAAATATATGTCTTGAGGAA
TACATGGTAGTTTGACAAGCCAACACCTCTCTCCTTAATCCAAAATTC

AAAAGGTAGTATAATTCTTTCTTTCCCAGCAGCAGAGGAAGTGACTAT
TGCAACCCACATTCTGAGTGTGAAATGAGGCCTCATTCAGTTTCCTAA
TTTCCAAGATTCTTTTTCAGCACCCAGCTGAATAAAGTATGTCCCGGA
GCTGGAACTGCCTGTCCGCGTGGACCCTCTCCTCTGGCACGGAAACTC
CGTGTCTGCAGACTGCCTCAGCCCTGGAGACTTCTGCCGGGAACCCG
GGACACCGAGAGGGGCCCAGCCCTTCCCCATCTCTCTCCTGACCGTTA
CTCCGGACAGGTAGGGTCTGCGCTCCTCTCCCCCCCCCCCCCCGTACT
AGTCTACGCGTGGCCA

Human 5' RACE clone E

CTCCCCAAGGAAAATTTCCGTCGATTTGGCTAAGATTGCAACAAAAT
ATCCTTTAGAATGTTAGACATATATATTGAATTTTATAAAGCCACAAA
AATATGTCTTGAGGAATACATGGTAGTTTGACAAGCCAACACCTCTCT
CCTTAATCCAAAATTCAAAAGGTAGTATAATTCTTTCTTTCCCAGCAG
CAGAGGAAGTGACTATTGCAACCCACATTCTGAGTGTGAAATGAGGC
CTCATTCAGTTTCCTAATTTCCAAGATTCTTTTTCAGCACCCAGCTGAA
TAAAGTATGTCCTGTGCCGGGGAAGGCGCCAAGCACCCGCCCCAGCG
CGCCGCACCTGCCGGGGCCGCGATCTCCCAGCTCATCGAGTCCCTCGC
GCCCGTGGCCCTTTGACCCCCCCCCC

8.1.4 qPCR

The first table shows the sequences for the primers used on qPCR and also the
specific UPL Roche probe for the region.

Table 8.4 Slynar primers and UPL probes used for qPCR

	Forward Primers	Reverse Primers	UPL probes
Slynar Mouse Exon 3 to 4 (a)	aaggcaataggaatgtcagtcac	tgtcagctgctggatttttg	56
Slynar Mouse Exon 3 to 4 (b)	tccatgacttctcactgctga	ggtatctataagatgatgtccaacaca	106
Slynar Human Exon 3 to 4	cctgttgagccacagtacca	caatgacttctctcaggcaatg	7

The second table contains the list of reference genes used on the qPCR
experiments for the mouse and the human cell lines. These set of primers were
obtained from Roche or Qiagen and were ready to use optimised mixes. MAPK1
was not used as a reference gene for qPCR but was a positive control for siRNA.

Table 8.5 List of house keeping genes used in human and mouse qPCR experiments

Genes	Source	Mouse	Human	Function
ACTB	Roche	✓	✓	Reference
HPRT	Roche	✓	✓	Reference
Gus B	Qiagen	✓		Reference
GAPD	Roche	✓		Reference
Cyclophilin	Qiagen	✓		Reference
PGK1	Roche		✓	Reference
MAPK1	Qiagen		✓	siRNA positive control

8.1.5 Northern Blot

The riboprobe used was based on an amplified sequence from Slynar Mouse Exon 3 to 4 (8.1.1). The sequence had to be cloned and sequenced to confirm that was specific for the region of interest.

The riboprobe sequence:

CTGCCAGATCTGATTTTCTGCCCTCTGCCTGGTCTTCTGGAAGGTCATT

GAGAATGCATGATCCATGACTTCTCACTGCTGATGAGGCCCAGGGTA

GCCAGAGAGTTGTGTTGGACATCATCTTATAGATACCATGTCAGCTGC

TGGATTTTTGTTTTGCTGTCCTCCCAGTCCATGGGCAGTGACTGACATT

CCTATTGCCTTCTGTTCTTTCCCATCC

8.1.6 siRNA

The four siRNA oligos sequences are unknown. A sequence of the region of Slynar exon 3 to exon 4 was sent to Qiagen, for the siRNA to be designed. Due to the company terms and conditions, we unaware of which regions of Slynar's sequence were used to design the oligos. Nonetheless, we do know that exon 3 and exon 4 are targets.

Slynar exon 3 and 4 target sequence:

CAGAGAGTCAAGGGATTGGATTGTGGATTGCATTGGTGGTTTTCCTCA
GTTTCCTCATCTTCTCCACAAGTTTCTACATATCGAATGCAGAGCAGC
CCTTCTTCAAAGGACCTCCTACGGAAGCTGCTAAGGAACTCAGTCTGT
AGCTCTGCGTGGAGCCATGTGTAAACACTGAACTGAGACCTGCCACC
TCCTACTACCTAAGGGCCCATTTTCATCTGATATCATCCCCCAGAAAC
AAACTCATGATGACTTCCATGTTTTTTTTAGATTAGATACATGGAGAA
TTTTCCTTTCCCTTAGAATTAAAATCCTGCATTCT

8.2 IGF1 Sequencing primers

The table represents the list of primers used to sequence IGF1. The primers cover a part of the 5' and 3'untraslated regions, exons and intron/exon junctions. To these sequences was also added a M13 tail (Table 3.1 M13 tails sequences.).

Table 8.6 Primers used for IGF1 sequencing

IGF1 region	Forward Primer	Reverse Primer
IGF1_Ex1a	GCAGCACTTAAATAATTGGGTTGGAA	TCATGCCCAGCAGAAAGTTAATCA
IGF1_Ex1b	CGGAAGCCCTGCAGAAGTGG	GAGGGAGAGAGAGAGAAGGCAAATG
IGF1_Ex2	GGCCTAGGATGGCTGCCAGA	CCGAGACACGCTCCATCCAC
IGF1_Ex3	CATCGTCCATAGCGGTGGGA	CCGGGAGACATACTGGCATTCA
IGF1_Ex4a	GACAAGGTTGCTGAATGAATGGC	ACCAAGGTTCAAGCTCTCTATTAGCTC
IGF1_Ex4b	AGGAGGCCAAATTCGGCAAA	CGCAAGTAGAGGGAGTGCAGGA
IGF1_Ex4c	TCTTTGGCTCCAGGCTTCCC	GCACTTCTTTCTACACAACTCGGGC
IGF1_Ex4d	CGCCAGTCCAATTTGCATCA	GCCGAATTTGGCCTCCTCAA
IGF1_Ex4e	TTGGCCAGTTATTTGGATAGCTTCA	GGGAAGCCTGGAGCCAAAGA
IGF1_Ex4f	CCCTTTCACTGGTAGAAATCTCTTTG	GACTGGCGAGTCCAGAGAGGAA
IGF1_Ex4g	CTTTCAACTGGAAACTCTAGTCAAGCA	GGGCGCTTGAGTTGCTGAGA
IGF1_Ex4h	TGCATGAAATAATCAAGCCTGGG	CAAAGAGATTTCTACCAGTGAAAGGG
IGF1_Ex4i	CCATCTTGGGAAGAGGAGTCCA	CATGCCTGCTCAGAAGGGTAGC
IGF1_Ex4j	TTGGTTGCTCCTTTCTATGAAATCTGA	TCCAACATTATTTGAATTGAGCACCTC
IGF1_Ex4k	GAGTGGATTCTGATGGAGAGCTGC	GGAAAGCTGAAAGATGCACTGCC
IGF1_Ex4l	GGATTCTCAAGGGTGGACAGGC	TTGATGCAAACCCTGGAAGTCA
IGF1_Ex4m	GTTGAAAGGTGGTGGTGGCTAGA	TGCAGCTCTCCATCAGAATCCA
IGF1_Ex4n	TCACAGATTGTTAGCCATCTCTTTCA	GCCACCACCACCTTTCAACTTT
IGF1_Ex4o	TTTATGGTCTTTGCAAGGGAGGG	TCTCCCTTCACCCAGACATCTCA
IGF1_Ex4p	GAGTTTCAGCTTGGTCAGCCCTC	GCAGAACCTGTTTGGCTCTCCTC
IGF1_Ex4q	CTGAGGCGGGCAAATCACAA	AACCAATTCCTATCTGGAACAATGCTT
IGF1_Ex4r	CAGCTCCGGTTATTAGGAGAAACTCTG	GAGGGCTGACCAAGCTGAAACTC
IGF1_Ex4s	TGCAGTGTGTTTAGCAGCGGG	TGGATCTTAAACATGATCCTTCTCTCC
IGF1_In3a	CATCATCATCTAGCTCCAGCAGGC	CAGGTGACCCAGCGCCTCTT
IGF1_In3b	GGGAATCTGGGAACTTCTATGACACA	GCCAAAGACACATCCAGGAGGG
IGF1_Pa	TCTCCTGCGATGAGGCAAAGA	CCACCACTCCTGGGAAACCA
IGF1_Pb	CCCTTGTCCCAGTTGCCAAG	AAACCAAAGGGAAATAGGTACAAACTG
IGF1_Pc	TGTGTTAGTGACAGGGTTCGCAGA	CCCTTCTCCCAGAGTGGTGGG
IGF1_Pd	TCTCTCTCTCCCTCTTCTGGCA	TCTTTGCCTCATCGCAGGAGAA

8.3 IGF1 Association Table

On chapter 5, I include the table of the positive GWAS SNPs (Table 1.) and the SNPs found by sequencing IGF1 (Table 5.1). This additional table contains all the 43 SNPs (from GWAS and sequencing) used for allelic and haplotypic associations and COMBASSOC analysis. In bold are highlighted $P<0.05$.

Table 8.7 List of markers used for IGF1 association tests. Includes markers form GWAS and from sequencing.

Marker	Marker Location	Alleles and Observed AlleleFrequencies						χ2	P value
rs11111243		C			G				
Controls	101232831		691	0.68		325	0.32	1.713	0.191
Cases			714	0.707		296	0.293		
rs10860857		C			A				
Controls	101252032		604	0.608		390	0.392	2.899	0.089
Cases			642	0.645		354	0.355		
rs4764695		T			C				
Controls	101259580		496	0.486		524	0.514	1.331	0.249
Cases			517	0.512		493	0.488		
rs10860860		T			A				
Controls	101283300		692	0.678		328	0.322	1.385	0.239
Cases			711	0.703		301	0.297		
rs2946831		A			C				
Controls	101289247		52	0.051		968	0.949	1.419	0.234
Cases			64	0.063		948	0.937		
Ex4k*		G			C				
Controls	101295526		1786	0.975		46	0.025	0.299	0.583
Cases			1793	0.978		41	0.022		
rs6214*		G			A				
Controls	101296036		769	0.414		1089	0.586	0.415	0.519
Cases			785	0.424		1065	0.576		
rs10745940		C			T				
Controls	101299663		281	0.275		739	0.725	0.008	0.930
Cases			280	0.277		730	0.723		
rs978458		G			A				
Controls	101304706		256	0.251		764	0.749	0.011	0.918
Cases			256	0.253		756	0.747		
rs9308315		T			A				
Controls	101306360		250	0.247		762	0.753	0.016	0.899
Cases			248	0.249		746	0.751		
rs17727841									
Controls	101312097		179	0.175		841	0.825	0.039	0.843
Cases			181	0.179		831	0.821		
rs5742688		G			A				
Controls	101314993		13	0.013		1001	0.987	7.637	0.006
Cases			31	0.031		977	0.969		
rs2072592		G			A				
Controls	101316099		992	0.973		28	0.027	11.152	0.001
Cases			1004	0.992		8	0.008		
rs5742678		C			G				
Controls	101316799		256	0.251		764	0.749	0.011	0.918
Cases			256	0.253		756	0.747		
rs10860864		C			T				
Controls	101325329		175	0.173		835	0.827	0.184	0.668
Cases			182	0.181		826	0.819		
rs2373721		G			C				
Controls	101329512		228	0.224		792	0.776	1.308	0.253
Cases			247	0.245		761	0.755		
rs11111272		G			C				
Controls	101329908		282	0.278		734	0.722	0.014	0.907

Cases				281	0.28	723	0.72	
rs5742652		C				T		
Controls	101338533		23	0.023	995	0.977	0.098	0.754
Cases			25	0.025	987	0.975		
rs5742632		A				G		
Controls	101358941		763	0.751	253	0.249	1.577	0.209
Cases			784	0.775	228	0.225		
rs12423791		C				G		
Controls	101361295		989	0.973	27	0.027	13.334	3.000E-04
Cases			998	0.994	6	0.006		
rs4764698		G				C		
Controls	101362527		987	0.973	27	0.027	13.375	3.000E-04
Cases			998	0.994	6	0.006		
rs1019731		G				T		
Controls	101366892		856	0.844	158	0.156	0.008	0.929
Cases			849	0.846	155	0.154		
rs10860869		A				T		
Controls	101367519		726	0.716	288	0.284	0.375	0.540
Cases			734	0.728	274	0.272		
rs12821878		G				A		
Controls	101370134		745	0.738	265	0.262	1.585	0.208
Cases			768	0.762	240	0.238		
rs5742620*		C				A		
Controls	101372067		51	0.028	1779	0.972	3.724	0.053
Cases			73	0.04	1783	0.96		
rs5742615**		G				T		
Controls	101373268		1002	0.969	31	0.031	17.006	3.726E-05
Cases			1002	0.994	6	0.006		
Pd*		G				T		
Controls	101377223		1805	0.973	51	0.027	7.490	0.006
Cases			1825	0.985	27	0.015		
rs5742612*		T				C		
Controls	101377331		1781	0.96	75	0.04	0.753	0.385
Cases			1793	0.965	65	0.035		
Pc*		T				G		
Controls	101377396		1785	0.96	75	0.04	0.753	0.385
Cases			1797	0.965	65	0.035		
rs35765		C				A		
Controls	101384163		85	0.083	935	0.917	5.257	0.022
Cases			115	0.114	897	0.886		
rs860598		G				C		
Controls	101400913		145	0.148	837	0.852	1.292	0.256
Cases			163	0.166	817	0.834		
rs2607983		A				G		
Controls	101402171		114	0.112	904	0.888	7.153	0.008
Cases			154	0.152	858	0.848		
rs2607986		T				G		
Controls	101404245		136	0.136	862	0.864	2.228	0.136
Cases			160	0.16	840	0.84		
rs7978777		C				A		
Controls	101421787		20	0.02	986	0.98	0.110	0.740
Cases			22	0.022	978	0.978		
rs11829367		G				A		
Controls	101422834		20	0.02	1000	0.98	0.239	0.625

Cases			23	0.023		989	0.977		
rs2195243		G			C				
Controls	101425453		227	0.223		793	0.777	0.065	0.799
Cases			230	0.227		782	0.773		
rs12309723		A			C				
Controls	101425980		975	0.956		45	0.044	3.969	0.046
Cases			984	0.972		28	0.028		
rs2114913		A			G				
Controls	101431171		983	0.964		37	0.036	3.253	0.071
Cases			989	0.977		23	0.023		
rs855288		C			G				
Controls	101441098		140	0.137		880	0.863	3.710	0.054
Cases			170	0.168		842	0.832		
rs703542		G			C				
Controls	101443545		135	0.133		881	0.867	4.975	0.026
Cases			170	0.168		840	0.832		
rs703548		G			A				
Controls	101447140		139	0.137		879	0.863	4.040	0.044
Cases			170	0.169		838	0.831		
rs35762		G			A				
Controls	101452114		154	0.152		862	0.848	2.020	0.155
Cases			177	0.175		835	0.825		
rs753479		A			G				
Controls	101461029		154	0.151		866	0.849	1.803	0.179
Cases			175	0.173		837	0.827		

UCSC March 2004 assembly positions

* Newly found SNPs by sequencing

** SNP out of Hardy-Weinberg

8.4 CLUMP test results for CA the repeat

Table 8.8 CLUMP p-values for the four CLUMP tests.

CLUMP test	X2	Pvalue
T1	15.161	0.041
T2	12.24	0.053
T3	9.567	0.013
T4	10.698	0.021