# *The Molecular Genetic Investigation of Bipolar Disorder in South Africa*

*Cinda-Lee Cupido (BSc. Med (Hons))*

*A thesis submitted to the Department of Clinical Laboratory Sciences, Faculty of Health Sciences, University of Cape Town, for the degree of Doctor of Philosophy*

*July 2005*

## Declaration

This study was performed from 2000 - 2005 under the supervision of Professor R. Ramesar of the Division of Human Genetics, University of Cape Town.

I hereby certify that this is my own work and has not been presented for a degree at another university.

Signed by candidate

Cinda L. Cupido

July, 2005

# TABLE OF CONTENTS

# ABSTRACT

Bipolar disorder (BPD), previously known as manic depression, is a severe, debilitating, mental illness. Characterized by cyclical swings between mania (elation) and depression (despair), BPD is common, estimated to affect 1 – 1.5 % of the population, and often aggregates in families. Suicide is a common phenotype and endpoint for about 20% of patients diagnosed with BPD. Twin, adoption and family studies provide strong support for a complex genetic aetiology that may involve multiple interacting susceptibility genes and an environmental component. The aim of this study is to identify susceptibility loci for BPD and attempted suicides (sub-phenotype), in South African families, integrating recent developments in statistical, computational and bioinformatics methods.

Based on prior international reports of linkage, ten candidate regions (*1q32, 2q33.3, 4p16.1, 7q11.23, 10q23.31, 12q24, 13q32, 15q12, 16p13.3* and *22q12.1*) were selected and assessed for evidence of linkage in 22 multiplex South African families, under affection status model (ASM) I (BPI and BPII) and ASM II (BPI, BPII and MDE). To investigate the genetic component of BPD, this study included (i) analyses involving single locus and the joint action of two loci, and (ii) the application of model-free/non-parametric linkage methods, using the programs MERLIN, Simwalk2, Allegro and TWOLOC. A modest linkage signal was found on 1q32, and we report evidence for interaction between 1q32 and 10q23.31. Using bioinformatics the candidate genes flagged in region 1q32, included neuron navigator-1 (*NAV*) and adenosine A1 receptor (*ADORA1*), whereas in region 10q23.31, the candidate genes serotonin receptor 7 (*HTR7*) and phospholipase C epsilon 1 (*PLCE1*) were identified.

This study went on to investigate the genetics of suicide attempts, owing to the high rates reported in individuals with an affective disorder. Two different approaches were used. In the association study the frequency of candidate gene (*TPH, DRD4, BDNF, HTR2A and 5-HTT*) polymorphisms were compared between 46 cases of attempted suicides and 160 controls. The only promising finding was the decreased frequencies of the 4/4 genotype and 4-repeat allele in the Caucasian subset, suggesting that the 4-

repeat allele may confer a possible 'protection' against susceptibility to development of an affective disorder and/or attempted suicide. Of the 46 cases, 9 extended families were available for a linkage study. The presence of genetic vulnerabilities, which predispose individuals with mood disorders to suicidal behaviour, were investigated. Single locus analyses were carried out under ASM I (to assess the genetic predisposition to suicide attempts), ASM II (to assess the genetic predisposition to mood disorders) and ASM III (to determine whether inclusion of suicide attempts in the group of mood disorders had any effect on linkage).

Single locus analysis provided statistical evidence in support of linkage between tryptophan hydroxylase (*TPH1*) and brain-derived neurotrophic factor (*BDNF*) in mood disorders. Evidence of linkage to *BDNF*, but not *TPH1* was significant in suicide attempts. It therefore appeared that *BDNF* predisposes susceptible individuals to both mood disorders and suicide attempts, whereas *TPH* predisposes to mood disorders but not to the risk of suicide, suggesting a possible protection factor against suicidal behaviour in this group of mood disorder patients. Two-locus interactions showed evidence of interaction between *BDNF* and *TPH*, *HTR2A* and *TPH*, and *5-HTT* and *TPH*, in the group of mood disorders. The single-locus linkage findings are consistent with international reports implicating *BDNF*, *TPH, HTR2A* and *5-HTT*, in the pathophysiology of mood disorders. However, none of the two-locus interactions described here have been reported before.

Due to the complex nature of BPD, it is clear that researchers cannot rely on classical genetics alone to identify genes for this condition. To this end, the present study integrates classical genetics methodologies, statistical modelling of two-locus interactions and bioinformatics, to identify genes in common pathways, that may be involved in BPD and suicidal behaviour.

# ACKNOWLEDGEMENTS

In presenting this thesis, I would like to convey my sincere gratitude and appreciation to my supervisor **Prof. Raj Ramesar**. His continuing encouragement, valuable advice and enthusiasm as a supervisor made the completion of this project and thesis a stimulating and pleasant learning experience.

I am most grateful to my collaborator **Dr. Heather Cordell** at the Juvenile Diabetes Research Foundation/Wellcome Trust (JDRF/WT) Diabetes and Inflammation Laboratory (DIL) at Cambridge University, United Kingdom. Her assistance in performing the statistical analyses has been invaluable. The hospitality shown to me during my visit to her department and her overall input has been truly appreciated.

I am also grateful to **Dr. Rene Goliath** (UCT, SA), **Prof. Jacquie Greenberg** (UCT, SA), and **Pippin Anderson** (UCT, SA) for all their constructive criticism, support and assistance in writing and structuring this thesis.

I would like to express my sincerest appreciation to **Dr. Ekow Oppon** (UCT, SA), for his assistance regarding the development of the bioinformatics tools. Likewise, I would like to thank **Mrs Amy Pather** (UCT, SA) for her help on the project.

I am also grateful to my **friends and colleagues** in the Division of Human Genetics (UCT, SA) for moral support, technical assistance and friendship.

A special thank you is due to all the families for their participation in the research project and to **Sr. Elize Pietersen** for her dedication and support in recruiting family members.

I am most grateful to the **South African Medical Research Council** for their financial support during the four years of my Ph.D.

Last but not least I would like to acknowledge the support and constant love of my mother **Angeline**, my three sisters **Lizl**, **Tarryn** and **Amber**, my brother-in-law **Lance** and my nephew **Joshua**.

# LIST OF TABLES

# LIST OF FIGURES

# SYMBOLS AND ABBREVIATIONS

| | |
|---|---|
| % | percentage |
| °C | degrees Celsius |
| $\lambda_R$ | risk ratio |
| µg | microgram |
| µl | microlitre |
| µM | micromolar |
| 5-HTT | serotonin transporter |
| A | adenine |
| AA | Alcohol Abuse |
| APA | American Psychiatric Association |
| APM | Affected pedigree member |
| ARP | affected relative pair |
| ASM | affection status model |
| BDNF | brain-derived neurotrophic factor |
| BOD | burden of disease |
| BP | Bipolar |
| bp | base pair (s) |
| BPD | Bipolar disorder |
| BPAD | Bipolar affective disorder |
| C | cytosine |
| cAMP | cyclic adenosine monophosphate |
| CNS | central nervous system |
| CREB | cAMP response element binding protein |
| DALY | disability adjusted life years |
| dATP | α-deoxyadenine-triphosphate |
| dCTP | α-deoxycytosine-triphosphate |
| dGTP | α-deoxyguanosine-triphosphate |
| DNA | deoxyribonucleic acid |
| dNTP | deoxynucleoside triphosphate |
| DRD4 | dopamine receptor 4 |

| | |
|---|---|
| DSM | Diagnostic and Statistical Manual of Mental Disorders |
| dTTP | α-deoxythymine-triphosphate |
| et al. | and others |
| G | guanine |
| GABRA5 | gamma-aminobutyric acid receptor, alpha-5 |
| HCl | hydrochloric acid |
| HRAS1 | v-ha-ras harvey rat sarcoma viral oncogene Homolog |
| HTR2A | serotonin receptor 2A |
| IBD | identical-by-descent |
| IBS | identical-by-state |
| incl. | including |
| INS | insulin |
| kb | kilobase |
| KCl | potassium chloride |
| l | litre |
| LD | linkage disequilibrium |
| lod | log of the odds |
| M | molar |
| MDD | Major depressive disorder |
| MDE | Major depressive episode |
| Mg | magnesium |
| Min | minute (s) |
| ml | millilitre |
| MLS | maximum lod score |
| mM | millimolar |
| MRI | magnetic resonance imaging |
| mRNA | messenger ribonucleic acid |
| NaCl | sodium chloride |
| NaOH | sodium hydroxide |
| ng | nanogram |
| NIMH | National Institute of Mental Health |

| | |
|---|---|
| NPL | non-parametric linkage |
| OOA | Old order Amish |
| PCR | polymerase chain reaction |
| PET | positron emission tomography |
| PFC | prefrontal cortex |
| pH | log $H^+$ ions |
| pmol | picomoles |
| RFLP | Restriction fragment length polymorphism |
| SA | South Africa |
| SCZ | Schizophrenia |
| SNP | single nucleotide polymorphism |
| SSCP | single strand conformational polymorphism |
| T | thymine |
| TBE | Tris-borate/EDTA |
| TCI | temperament and character inventory |
| TDT | transmission disequilibrium test |
| TPH1 | Tryptophan hydroxylase 1 |
| UCT | University of Cape Town |
| UK | United Kingdom |
| US(A) | United States (of America) |
| VNTR | variable number tandem repeat |
| www | world wide web |

# GLOSSARY

**Apoptosis**

Also known as programmed cell death, is the way in which cells that are induced to commit suicide

**Gene**

A sequence of DNA that codes for a particular protein.

**Genotype**

The genetic constitution of an individual either overall or at a specific locus.

**Heredity**

Transmission from generation to generation through the process of reproduction in plants and animals of factors which cause the offspring to resemble their parents.

**Heterogeneity**

Any one of multiple genes may underlie similar phenotypes in different families, such as a wide concert of genes interacting in a common biochemical pathway

**Informed consent**

Consent given by a patient allowing research to be undertaken on his/her biological material.

**Linkage analysis**

A technique, designed to search for co-occurrence of a marker locus and a putative disease gene, using all the available inheritance information from families

**Lod score**

A measure of the likelihood of genetic linkage between loci. The log (base 10) of the odds that the loci are linked (with recombination fraction $\theta$) rather than unlinked. For a Mendelian character, a lod score of +3 is evidence of linkage, one that is –2 is evidence against linkage.

**Microsatellite**

A small run of tandem repeats (usually less than 0.1kb) of a very simple DNA sequence, usually 1-4bp e.g. (CA)n.

**Multifactorial**

Traits are determined by interactions between a gene or genes and the environment, and do not exhibit Mendelian ratios

**Non-parametric analysis**

Is model-independent and therefore tests the deviation of inheritance pattern from random Mendelian segregation under independent assortment

**Oligogenic**

A phenotypic trait produced by two or more genes working together, but distinct from polygenic inheritance in that smaller number of total loci thought to be involved.

**Parametric analysis**

Proposes a specific model to explain the inheritance pattern of a trait-causing gene

**Penetrance**

Is genotype specific and is the probability of being affected given the genotype.

**Polymorphism**

The existence of two or more variants at significant frequencies in the population; any sequence variant present at a frequency of more than 1% in a population.

**Prevalence**

The proportion of individuals in a population having a disease.

**Proband**

The family member whose disease led to ascertainment of family for study.

# PLAN OF THESIS

This dissertation is comprised of five chapters, and the layout is as follows:

**Chapter 1** is a review of the literature, providing a context for BPD, that is a general theoretical background on various aspects of BPD. This is in order that the reader understands the concepts as they are introduced, and is aware of their importance. The work-related chapters 2 to 4 provide the relevant literature pertaining to these specific chapters.

**Chapter 2** provides perspective for the subsequent research, drawing primarily on the BPD research archive in the Division of Human Genetics, at the University of Cape Town South Africa, with some input from international archival resources.

**Chapter 3** and **Chapter 4** present the laboratory-based research and each consists of an introduction, the experimental approach, followed by results and discussion. Chapter 3 describes the investigation of the candidate regions that were chosen for linkage analysis, involving single and two-locus interactions, in bipolar disorder multiplex families. This chapter also describes the development and application of bioinformatics tools to identify candidate genes in linked and other regions. Chapter 4 details a pilot study in the genetics of suicidality in a subset of the cohort outlined in Chapter 2.

**Chapter 5** concludes the dissertation, providing a general discussion of the present and international investigations.

# CHAPTER 1

## REVIEW OF BIPOLAR DISORDER

Based on the Mendelian principles of genetic transmission, the heritability of polymorphisms in DNA has been used in families to map inherited diseases to the genome. The successful application of these techniques to several hereditary diseases including Duchenne Muscular dystrophy (Monaco et al., 1985; Kunkel et al., 1986), cystic fibrosis (Kerem et al., 1989) and Huntington disease (Gusella et al., 1983) has naturally evolved to the use of mapping to even non-Mendelian disorders and indeed to psychiatric disorders. These common disorders do not follow a monogenic pattern of inheritance, but are generally influenced by both genetic and environmental components (Ghosh and Collins, 1996). This thesis outlines an attempt to identify genetic elements predisposing to bipolar disorder (BPD).

The intensity of research dedicated to BPD reflects the degree of scientific and public recognition of its 'burden of disease' in society. This introductory chapter (1) provides an overview of BPD, starting with the long and well recorded history of the disease. Following the historical perspective, the epidemiology of the disease is presented. Detailed descriptions of the clinical aspects of BPD, neurobiology of mood and pharmacological intervention all provide necessary background information on the status of research to date. The understanding of BPD as a multifactorial disorder is presented and the concept of a genetic or hereditary component is introduced. Molecular genetic research and analytical strategies adopted in BPD to date, and as they relate more directly to the proposed research, are described. In light of this historical context, the epidemiological status of the disease, and the current research status, the present study is motivated, and the aims and objectives of the research made explicit.

## 1.1 HISTORY OF BIPOLAR DISORDER

The deeds and accounts of the classical Greek physicians chronicle the recognition of BPDs even in ancient times (reviewed by Angst and Marneros, 2001). In the first century AD, Aretaeus of Cappadocia, cited in Angst and Marneros (2001),

categorically described the inherent connection between mania and melancholia. Notwithstanding the fundamental contributions made by these ancient physicians and subsequent psychiatrists of the 19th century, BPD was first classified as a distinct entity in 1951 by Jean-Pierre Falret (cited in Angst and Marneros 2001). Falret observed BPD, which at the time he termed 'folie circulaire', which is characterised by a continuous cycle of depression and mania with lucid intervals of varying length. Succeeding Falret and with opposing views, Jules Baillarger, in 1854, submitted his judgement 'folie a double forme', that is a disease in which mania and melancholia vacillated, but where the interval was of no consequence (cited in Angst and Marneros 2001).

The German psychiatrist, Emil Kraepelin (1856-1926), upon whose findings contemporary distinctions of psychiatric disease depend, defined manic depressive disorder thus: "As a rule the disease runs its course in isolated attacks more or less sharply defined from each other or from health. Accordingly we distinguish first of all manic states with the essential morbid symptoms of flight of ideas, exalted mood, and pressure of activity, and melancholia or depressive states with sad or anxious moodiness and also sluggishness of thought and action. These two opposed phases of the clinical state have given the disease its name. But besides them we observe also clinical 'mixed forms', in which the phenomena of mania and melancholia are combined with each other" (Angst and Marneros 2001).

Kraepelin's immense contribution to the comprehension, diagnosis and prognosis of manic depressive illness (BPD), qualifies his recognition as the 'father of modern psychiatry' (reviewed by Angst and Marneros, 2001). However, two of Kraepelin's impressions have motivated strong counterarguments from German scholars in psychiatry. Firstly, Kraepelin's unitary concept of BPD included not only attenuated forms, but also much of the sphere of major depressive disorders (Akiskal, 2003). Secondly, in 1899, his binary theory separated manic depressive insanity and dementia praecox (schizophrenia) as distinct disease entities, based on clinical symptoms, course, and prognosis (Crow, 1986; Angst, 2002). However, the more descriptive trend in modern psychiatric classification is in favour of a continuum from BPD to schizophrenia (Crow, 1986; 1990; Angst, 2002), including schizoaffective disorder as an intermediate phenotype (Angst and Marneros, 2001; Benabarre et al., 2001). The vast range of research of BPDs and schizophrenia

internationally, showed similarities in age at onset, lifetime risk, course of illness, world-wide distribution, suicide risk, gender influence and genetic susceptibility (Berrettini, 2000).

In 1966, Perris put forward the nosological distinction between unipolar and BPD, thereby confirming Falret's and Baillarger's concepts of cycling mania and depression in BPD (Perris, 1966; Angst and Marneros, 2001). This binary conceptualisation removed a large portion of depressive disorders from BPD (Akiskal, 2003), and is currently reflected as official nosology in the fourth edition of the Diagnostic and Statistical Manual for mental disorders (DSM-IV, American Psychiatric Association (APA), 1994).

## 1.2 EPIDEMIOLOGY OF BPD

Over recent decades the clinical definition of BPD has been revisited extensively, giving rise to the disparity in reported prevalence even in the same population groups (Greil and Kleindienst, 2003). Although earlier epidemiological studies reported a lifetime prevalence of 0.5-1.5% for BPD (Hwu et al., 1989), more recent population-based investigations indicate a much higher prevalence of this disorder. The reported increase in prevalence is attributed to the improved ascertainment through the development of diagnostic manuals and diagnostic practices (Greil and Kleindienst, 2003). For example (with the exception of the DSM-IV), a patient initially diagnosed as either schizophrenic, unipolar depressive or exhibiting a personality disorder (PD), is now presumed to belong to the BP spectrum (Stoll et al., 1993; Angst and Marneros, 2001; Akiskal and Benazzi, 2003).

The burden of neuropsychiatric disorders, which includes mental and nervous system disorders, was highlighted in the 'Global Burden of Disease Study' (GBD) (Murphy and Lopez, 1997a). In 1990, this cluster of conditions was reported to account for 0.3% of deaths and 10.5% of the burden of disease and injury world-wide. Estimates of the global burden of disease calculated in terms of disability-adjusted life years (DALYs) placed unipolar major depression (UMD) as the fourth most disabling disorder world-wide, preceded by, in descending order, lower respiratory tract infections, diarrhoeal diseases and perinatal disorders. With the inclusion of suicide in the primary tabulations, the burden of UMD was elevated by

nearly 40% (Murphy and Lopez, 1997a). Alcohol use, schizophrenia and BPD were the other neuropsychiatric conditions in the 30 leading contributors to burden of disease. As part of the GBD, projected health scenarios of (future) mortality and disability, predict a rise in the burden of neuropsychiatric disorders, including BPD, from 10.5% in 1990 to 14.7% in 2020 (Murray and Lopez, 1997b).

In South Africa the "National Burden of Disease study" excluded the category of neuropsychiatric conditions per se, but included the categories of mental disorders and nervous system disorders, to make up the category of neuropsychiatric conditions (Bradshaw et al., 2003). Remarkably, this category of neuropsychiatric conditions ranked as the second most burdensome of all medical conditions, after HIV/AIDS, thereby justifying the present research into this group of disorders.

## 1.3 CLINICAL ASPECTS OF BPD

### 1.3.1 Clinical phenotype and diagnostic guidelines

Much of the current understanding of the clinical phenotype of BPD is based on the DSM-IV (APA, 1994). Aimed at presenting a unified and comprehensive nomenclature/nosology for psychiatric illness, more than 300 mental, behavioural and addictive disorders, as well as other psychopathologies, are described and grouped into broad diagnostic categories in the most recent edition of this publication. The DSM-IV uses a multi-axial approach, where the axes are levels of diagnoses that assist the clinician in the planning of treatment and prognostication. Axis I comprises the major groups of clinical disorders, including eating, anxiety, substance-related and mood disorders, amongst others. Of these, the mood disorders comprise those conditions where the predominant feature is a disturbance in mood (affect), such as BPD and major depressive disorder (MDD). Characteristically, individuals affected with BPD (often not recognised by relatives, friends, physicians or even the patients themselves) vacillate between states of major depression, hypomania and frank mania; whereas individuals affected with MDD, experience one or more major depressive episodes (MDE).

According to DSM-IV guidelines, during a MDE the patient experiences a loss of interest or pleasure in his/her normal activities, weight loss or gain, insomnia or hypersomnia, fatigue, psychomotor agitation or retardation, diminished ability to

think or concentrate and a feeling of worthlessness. In order to meet the criteria for diagnosis of a MDE, symptoms must be present nearly everyday for at least two weeks. Furthermore, patients may be plagued by recurrent thoughts of death or suicidal ideation, or they may attempt suicide. A hypomanic episode typically lasts at least four days, during which the patient's symptoms include: elation or irritation, inflated self-esteem or grandiosity, decreased need for sleep, psychomotor agitation, recklessness, impulsivity, incoherent thoughts, distractibility and a flight of ideas. A manic episode, a more extreme case of hypomania, lasts for a minimum period of one week, when the patient experiences all the symptoms of a hypomanic episode as well as difficulty in performing their daily functions.

On account of the different phases that patients with BPD may experience, the DSM-IV has defined two different forms of the condition, BP type I (BPI) and BP type II (BPII). A subject diagnosed with BPI disorder may express one or a combination of the following symptoms:

(i) one manic episode with no history of MDEs,

(ii) currently or most currently in a hypomanic episode, preceded by at least one manic or mixed episode,

(iii) currently or most currently in a manic episode following at least one previous MDE,

(iv) manic or mixed episode, or currently in a mixed episode having had at least one previous MDE, manic episode or mixed episode.

The diagnosis of BPII disorder is based on the presence or history of one or more MDEs and at least one hypomanic episode, without any prior experience of manic or mixed episodes. Often people experience a mixed state, i.e. when the criteria for both a manic and a depressive episode are met everyday for at least one week. DSM-IV criteria also include rapid cycling, a variant of BPD, during which patients experience four or more mood episodes within a 12-month period.

Some individuals endure fluctuating disturbances of mood that involve numerous periods of hypomania and depression, whilst the symptoms do not meet the full criteria for classification of a manic episode or a MDE. If symptoms persist for at least two years, patients are diagnosed with cyclothymic disorder. When a patient

appears to present features of BPD, yet does not meet the "full" criteria for BPD, a diagnosis of BPD "not otherwise specified" (NOS) is given.

Patients with BPD commonly have at least one other (co-occurring) Axis I disorder, such as anxiety and substance abuse, with eating disorders diagnosed to a lesser extent (McElroy et al., 2001; 2004). Co-occurring disorders, such as BPD and substance abuse, or BPD and anxiety disorders, have important clinical implications, because of the prevalence and negative impact on the course, treatment and prognostic response to both disorders (McElroy et al., 2001; Myrick and Brady, 2003; Boylan et al., 2004). Substance use, for example, aggravates bipolar symptoms and functionality, as evidenced by its association with an earlier age of onset of BPD (Feinman and Dunner, 1996), increased rates of psychiatric hospitalisation, delayed remission from acute mania (Goldberg et al., 1999), poor response to lithium among bipolar patients (O'Connell et al., 1991), and higher rates of suicide attempts (Feinman and Dunner, 1996; Potash et al., 2000).

BPD can also present with an Axis II disorder, which includes personality disorders (PDs) such as paranoid PD, borderline PD, obsessive-compulsive PD, and mental retardation. Kraepelin, cited in Cloninger et al. (1998), was the first to identify the existence of enduring/rudimentary personality variants that are essential for the development of psychotic or mood disorder states. In 1999, Henry and colleagues demonstrated the impact of temperament on the clinical evolution of BPD, supporting the existence of a continuum between personality and affective disorders. The three most frequent PDs associated with BPD, as evidenced by the independent studies of Rossi et al. (2001) and Brieger et al. (2003), were obsessive-compulsive PD, borderline PD and avoidant PD.

Variance in personality can arise from four temperament dimensions (harm avoidance, novelty seeking, reward dependence and persistence) and three character dimensions (self-directedness, cooperativeness and self-transcendence) (Cloninger et al., 1993). Based on these dimensions, the Temperament and Character Inventory (TCI), one of a number of rating scales, was developed as a reliable and efficient test for the assessment of personality. TCI traits were found to be related to mood states (Cloninger et al., 1998). More recently, using the TCI, it was reported that patients with BPD scored significantly higher in harm avoidance

and lower in reward dependence, self-directedness and cooperativeness, than a control group (Engstrom et al., 2004). These BPD patients were wearier, less sentimental, more independent, less purposeful and resourceful, less empathetic and had less impulse control than the control subjects. Engstrom and co-workers (2004) noted the possibility that specific personality dimensions may predispose individuals to BPD or alter the natural history of illness and response to treatment. Personality traits are a significant theme of BPD research internationally, and may further contribute to sub-phenotypes within a heterogeneous phenotype.

Based on the number of overlapping diagnostic and syndromic entities observed in families with BPD, a number of clinical practitioners/researchers have contributed to the construction of a bipolar spectrum (Gershon et al., 1982; Weissman et al., 1984; Coryell et al., 1984; Kelsoe 2002; Akiskal, 2003). Ranging in decreasing symptom severity and detriment, the spectrum includes the following: BPI - BPII - recurrent unipolar depression - single episode unipolar depression - cyclothymia - dysthymia - affective personalities. Adding to the complexity of the BPD phenotype, Crow (1986) proposed the concept of a continuum of psychosis with increasing severity and impairment, extending from unipolar, through bipolar affective illness and schizoaffective psychosis, to typical schizophrenia. The observed aggregation of psychotic symptoms, defined as hallucinations or delusions (originally thought to typify schizophrenia), in families with BPD lends support in favour of Crow's continuum of psychosis (Potash et al., 2001; Tsuang et al., 2004). It is currently unresolved as to whether the all-inclusive bipolar spectrum depicts a clinical or genetic continuum (Akiskal, 2003).

DSM-IV does not accurately distinguish between BPI and BPII disorder, except for the difference between mania and hypomania, since age of onset and the duration of episodes are very similar. Furthermore the current diagnostic gold standard, the Structured Clinical Interview for DSM-IV (SCID), classifies the bipolar spectrum conditions as Type II and NOS (Ghaemi et al., 2003) and omits depressive mixed states, MDE and hypomania of less than four days duration, which are commonly found in clinical practice (Akiskal and Benazzi, 2003). Therefore the bipolar spectrum disorders often go unrecognised, or they may be misdiagnosed as symptoms of either impulsive behaviour or alcohol and substance abuse (Lish et al.,

1994). Misdiagnosis or delayed diagnosis can be damaging, as it complicates drug treatment strategies and contributes to suicidal behaviour (Kasper, 2003).

Recognising the need for more descriptive screening instruments or measures for BPD, researchers have sought to develop sufficiently sensitive diagnostic tools to supplement DSM-IV guidelines. The Mood Disorder Questionnaire (MDQ), a brief self-report scale, was developed to screen for a lifetime history of a manic or hypomanic syndrome (Hirschfeld et al., 2000). The MDQ yielded good sensitivity and specificity scores for the bipolar spectrum disorders in a psychiatric outpatient population. Moreover, the MDQ improved the diagnosis of the so-called 'soft' phenotype (Akiskal, 2003), towards the diagnosis of BPII disorder, which frequently stays undetected in psychiatric care (Isometsa et al., 2003). Another instrument, the Bipolar Spectrum Diagnostic Scale (BSDS), aimed at identifying BPII and NOS conditions, was demonstrated to be highly sensitive and specific for bipolar spectrum conditions (Ghaemi et al., 2003). The combined administration of MDQ and BSDS promises to be a powerful screening tool. Well-validated and comprehensive, the Brief Psychiatric Symptom Rating scale-24 item version ($BPRS_{24}$) (cited in Dennehy et al., 2004), assesses symptom changes in psychiatric patients, including manic phase symptoms for patients with BPD. A subset of 10 $BPRS_{24}$ items was incorporated into the Brief Bipolar Disorder Symptom Scale (BDSS), which was developed to measure the critical dimensions and symptoms seen in patients with BPD (Dennehy et al., 2004). In preliminary investigations the BDSS was sensitive to symptom change, suggesting its use as an invaluable supplementary tool in clinical settings for the purposes of assessing symptom severity and treatment response.

While various formal rating scales exist for the assessment of a range of clinical expressions of bipolar spectrum disorders, not all of these tools appear to be well suited for use in everyday clinical practice. The area of diagnosis is therefore an area that is in need of considerable further research. It is likely that once genes are identified in at least a subset of BPDs that the confusing issues of genetic and clinical heterogeneity will begin to be logically defined.

Another reported feature of BPD, is greater neurological dysfunction, predominantly in the region of sequencing of complex motor acts (Negash et al., 2003). It is

possible that these early neurological "soft signs" are stable manifestations of BPD and represent a stable disease process. Furthermore, cognitive abnormalities have been detected in symptomatic BPD patients with trait-related deficits present in verbal recall and sustained attention (Quraishi and Frangou, 2002). The research into personality traits, neurological and cognitive dysfunction has contributed to the understanding of the disease pathology of BPD, however, these neuropsychological aspects are not discussed any further in this thesis.

Along the behavioural dimension, suicide is a serious complication of BPD and a common endpoint for many patients (Barraclough, 1974; Mann, 1998; Jamison, 2000). Over 90% of suicide victims are reported to have suffered a significant psychiatric illness in their lifetime (Isometsa et al., 1995; Harris and Barraclough, 1997; Mann, 1998; Sher et al., 2001), predominantly major depression and BPD (Guze and Robins, 1970; Oquendo et al., 2000; Jamison, 2000; Sher et al., 2001). An estimated 25-50% of patients with BPD admit to at least one suicide attempt in their lifetime (Jamison, 2000; Oquendo et al., 2000; Grunebaum, 2001). Consequently, the mortality rate in BPD is two to three times higher than that of the general population (Muller-Oerlinghausen et al., 2000). Amongst the 20 leading contributors of premature mortality in South Africa, suicide is ranked 10$^{th}$ in males (1.9%), 19$^{th}$ in females (0.7%) and 11$^{th}$ in general (1.4%) (Bradshaw et al., 2003). No definitive epidemiological information is available on the contribution of neuropsychiatric illness to suicide.

Comorbid substance abuse (Tondo et al., 1999), borderline PD, hopelessness and a history of impulsive aggression all increase the risk of suicide attempts and ideation in BPD (Oquendo et al., 2000; Soloff et al., 2000). It is estimated that the lifetime mortality due to suicide is approximately 20% in BPD patients, 15% in individuals suffering from unipolar depression, 18% in alcoholics, 10% in schizophrenics and 5-10% in individuals with some type of PDs. The latter include both borderline and antisocial PDs, characterised by emotional liability, aggression and impulsivity. Furthermore, with the TCI, individuals who scored low in both self-directedness and cooperativeness were at an increased risk of both attempted and completed suicide (Cloninger et al., 1998). Therefore, suicide is not merely a serious consequence of a personal crisis or a psychosocial impediment, but occurs at a significantly elevated incidence in the context of a psychiatric condition,

particularly in the mood disorders. The relationship between suicidality and BPD and specific genetic loci is investigated in Chapter 4 of this current study.

### 1.3.2 Gender and Age

The influences of gender and age on the clinical manifestation of BPD have been reported in the literature, and deserve a mentioning at this stage. According to McGuffin and Katz (1989), in order to give a sufficient account of the complex relationship between the influences of "nature" (genes) and "nurture" (environment) one has to consider gender differences in the aetiology of affective disorders. There are differences in the clinical course of BPD in men and women (Rasgon et al., 2005). Women have higher rates of depression and larger fluctuations in mood than men, and are about two times as likely as men to describe a lifetime history of MDEs (Kessler et al., 1993; Weissman et al., 1993; Bierut et al., 1999; Visscher et al. 2001; Ragson et al., 2005). Furthermore and in line with mood disorders generally, Silberg et al. (1999) reported increased heritability of depression among adolescent girls (aged 8 to 16 years).

Characteristically, BPD manifests at a relatively early age and late onset cases are reported to be rare (Nath et al., 2001; Visscher et al., 2001). Schurhoff et al. (2000) defined early and late onset as before 18 and after 40 years of age, respectively. The most severe manifestations of BPD are generally observed in the early onset group. Commonly, severely affected patients experienced increased psychotic features, increased mixed episodes, greater co-occurrence with panic disorder and inferior prophylactic lithium response. In addition, a reported association exists between early age at onset of BPD and suicidal behaviour and substance use disorders, rapid cycling course (Carter et al., 2003; Ernst and Goldberg, 2004) and a family history of affective disorders (Johnson et al., 2000). Owing to differences in clinical expression and familial risk, it has been implied that early and late onset BPDs may be considered to be distinct sub-forms of BPD (Schurhoff et al., 2000). As probable as the existence of these sub-forms may be, there is still debate in the literature on defining their accurate age at onset thresholds.

## 1.4 NEUROBIOLOGY OF MOOD

The underlying biology of mood disorders, which might aid surer and faster diagnosis, is gradually being exposed. As BPD is a neuropsychiatric condition, involving the central nervous system (CNS), the next two sections of this chapter provide relevant background on the normal functioning of neural circuitry. This is followed by a description of the reported neurological abnormalities underlying the clinical phenotype of BPD, an area that has been extensively studied over the last decade.

### 1.4.1 Neural basis of emotion perception

Emotion perception is based on three processes: (i) identification of the emotional importance of the stimulus; (ii) production of a specific affective condition in response to the stimulus; and (iii) the management of the affective condition and emotional behaviour (Phillips et al., 2003). Findings from recent human lesion (Adolphs et al., 1994; Bechara et al., 1995; Scott et al., 1997) and functional neuroimaging studies (Anderson and Phelps, 2001; Davis and Whalen, 2001; Drevets et al., 2001) suggest that these processes may depend on the functioning of two neural systems: a ventral, and a dorsal system. The ventral system, including the amygdala, insula, ventral striatum, ventral regions of the anterior cingulate gyrus and prefrontal cortex, is responsible for the rapid appraisal of the emotional significance of the stimuli and the subsequent production of the affective state. The dorsal system, including the hippocampus and dorsal regions of the anterior cingulate gyrus and prefrontal cortex, is responsible for executive functions, including selective attention, planning, and laborious rather than automatic regulation of resulting affective states (Scott et al., 1997; Davis and Whalen, 2001; Drevets et al., 2001). The functioning of these two neural systems involves complex signalling networks elicited by neurochemical transmission.

### 1.4.2 Neurochemical transmission

Neurochemical transmission is the process by which an electrical impulse in the presynaptic neuron is converted to a chemical signal (Stahl, 2000). This conversion involves the release of chemical messengers called neurotransmitters at the

synapse between the pre- and postsynaptic neuron. The released neurotransmitter molecules (first messenger) selectively bind to appropriate receptors on the surface of the post-synaptic neuron. This binding action leads to the postsynaptic events of chemical neurotransmission, starting with the production of second messengers, which include cyclic adenosine monophosphate (cAMP) and phosphatidylinositol (PI) (Stahl, 2000). Through a biochemical cascade, these second messengers instruct the postsynaptic neuron to execute numerous actions (for example, to modify its ionic fluxes, to generate or stop neuronal electrical impulses, and to phosphorylate intracellular proteins). These postsynaptic events ultimately extend to the cell nucleus, where they regulate the expression of certain genes. This process initiates a second biochemical cascade, which is dependent on the manner in which specific genes in the previous cycle have been regulated.

Termination of the initial signal occurs when the neurotransmitters detach from the receptor and either diffuse through the extracellular fluid, undergo re-uptake, or are broken down by enzymes (Kandel et al., 2000). Neurotransmitter receptors are also found on the presynaptic neuron (autoreceptors) and function as inhibitors of further neurotransmitter release. Each neurotransmitter has a variable number of receptor subtypes that function in association either with G-proteins, protein kinases, or by the receptor itself in the form of a ligand-gated ion channel (Stahl, 2000).

Neurotransmitters such as the three major monoamines: serotonin, norepinephrine, and dopamine can modify and shape human behaviour (Stahl, 2000). Serotonin, present in both the central nervous and peripheral nervous systems, is formed by the hydroxylation and decarboxylation of tryptophan and then released by serotonergic neurons (Mendes de Oliveira et al., 1998). In turn, serotonin (5-HT) exerts its function upon interaction with specific receptors of which several have been cloned and identified (5HT1 to 5HT7), each with its own subtypes. With one exception, the serotonin receptors are G-protein coupled receptors. Synaptic actions of serotonin are terminated by the re-uptake of serotonin by the serotonin transporter protein (Hoehe et al., 1998), and the destructive enzyme, monoamine oxidase (MAO).

Dopamine and norepinephrine, further classified as catecholamines, are chemical compounds formed from tyrosine (Furlong et al., 1999). These neurotransmitters, apart from their excitatory and inhibitory affects in the peripheral nervous system, are active in the CNS. In the CNS, dopamine receptors effect brain processes that control movement, cognition, emotion, in addition to neuro-endocrine secretion (Grandy et al., 1992; Emilien et al., 1999). The dopamine receptors belong to a superfamily of seven transmembrane domain G-protein coupled receptors (Sherrington et al., 1993). At least six dopamine receptors have been cloned from brain tissue (Emilien et al., 1999; Hiroi et al., 2002). These receptors include the D1 class, divided into D1 and D5 receptor subtypes, which, when activated, stimulate adenylate cyclase, the enzyme that converts ATP to cyclic AMP (Emilien et al., 1999). The D2 class, consisting of the D2, D3 and D4 subtypes, generally inhibit adenylate cyclase activity. Termination of dopamine synaptic activity involves the uptake of dopamine by the dopamine transporter into the presynaptic terminal (Giros et al., 1996). Dopamine is also destroyed enzymatically, by MAO and catechol-o-methyl transferase (COMT).

Dopamine, when transported into norepinephrine-producing neurons, is rapidly transformed into norepinephrine (Kandel et al., 2000). In the CNS, the locus coeruleus of the brain stem is concentrated with norepinephrine-containing cell bodies. Following release into the synaptic cleft, postsynaptic activity involves the excitatory alpha 1 and beta adrenergic receptors and the inhibitory alpha 2 receptor (reviewed in Ressler and Nemeroff, 1999). Norepinephrine activity can be terminated by the norepinephrine transporter, which rapidly returns norepinephrine to the synaptic terminals. The same enzymes that destroy dopamine (MAO and COMT) destroy norepinephrine.

In addition to monoamines, the nervous system makes use of multiple other neurotransmitters, including amino acids of which γ-aminobutyric acid (GABA) is the most important inhibitory neurotransmitter.

As neurochemical transmission is the process that underpins emotion and behaviour, the neurochemical control of any component of this highly regulated matrix may be a potential cause of neuropsychiatric disorders.

## 1.4.3 Cellular phenotype of BPD

In light of the intricacy of brain function, it is not surprising that the aetiology and pathophysiology of BPD is not well understood. However, distinct imbalances in neurochemical control, including reduced glial cell number, excess signal activity, neuropeptide aberrations and monoamine alterations, have been observed in the brains of patients with BPD (reviewed by Vawter et al., 2000). This area of research carries much credence and has guided the selection of candidate loci in the present study.

### *1.4.3.1 The neurotransmitters*

Based largely on pharmacological observations, the monoamine model (Schildkraut, 1965) implicates serotonin, dopamine and norepinephrine in BPD. It is possible that patients with BPD have diminished neurotransmitter transport into the synaptic vesicle and/or presynaptic neuron (Freis, 1954; Schildkraut, 1965). Consequently, the synaptic vesicles are non-functional and a surplus of the respective neurotransmitter exists, resulting in a higher concentration of transmitter in the synaptic cleft. This may also be responsible for the corresponding fluctuation in mood. In their study, Potter and Manji (1994) corroborated the monoamine hypothesis, which generally showed a rise in monoamine turnover occurring in mania and a fall in depression. Serotonin has been implicated in a variety of traits and behaviour such as mood, anxiety, aggression, sleep, appetite and suicidality (Mendes de Oliveira et al., 1998). Imbalances in the dopaminergic system are involved in manic depression, schizophrenia, Parkinson disease and Tourette syndrome (Grandy et al., 1992; Ernilien et al., 1999). Irregularities in norepinephrine circuits have been shown to possibly interfere with concentration, attention, memory, arousal states and sleep regulation (Ressler and Nemeroff, 1999). Further clarification, however, is needed on which of the three monoamines, are predominantly involved in depression and mania.

In addition to the roles of these monoamines, the GABA receptor genes may be involved in the pathogenesis of affective disorders (Petty, 1995). Sanacora and colleagues (1999) reported a significant reduction in cortical GABA concentrations

in the brains of depressed patients, with the idea that diminished amounts of this neurotransmitter could lead to the "uncontrolled" phenotype of BPD.

### *1.4.3.2 Beyond the neurochemical origin*

The monoamine model of BPD (Schildkraut, 1965) fails to explain the mechanism underlying the therapeutic action of antidepressant drugs. Furthermore it does not provide an holistic view of the pathophysiology of depression (Hindmarch, 2002). In addition, it seems as though medications work by modifying synaptic connections and therefore alternative mechanisms, other than the monoamine model, have been investigated to help elucidate the pathophysiology of BPD.

Emerging evidence implicates impairments of structural plasticity and cell survival in the pathophysiology of mood disorders (Duman et al., 1997; Manji et al., 2001; Rohlff and Hollis, 2003). Neuronal plasticity is pivotal to the process by which the brain gathers information and facilitates the execution of the suitable adaptive responses to anticipated situations (Duman, 2002). Disruption of these primary processes can thereby lead to the pathobiology of mood disorders. Recent studies have demonstrated structural changes in the brain in response to stress and in patients with mood disorders (Duman et al., 2000; Duman, 2002; Spedding et al., 2003).

The neuroprotective effects of lithium are mediated through the brain-derived neurotrophic factor (BDNF)/receptor tyrosine kinase (TrkB) signalling pathway (Hashimoto et al., 2002). BDNF, a neurotrophic factor for both the norepinephrine and serotonin neurotransmitter systems (Mamounas et al., 1995), is involved in cell survival and synaptic plasticity (Thoenen, 1995), and is one of the targets of antidepressant treatment (Duman et al., 1997; Duman et al., 1999). Over the past decade, investigation into the pathophysiology of BPD and unipolar depression has focussed on second messenger systems (for example, cAMP and PI) within cells. The cAMP second messenger pathway, including the cAMP response element binding protein (CREB) and G proteins, has been implicated in BPD (Manji et al, 1995; Thome et al., 2000). Furthermore, altered levels of catalytic subunits of cAMP-dependent protein kinase (PKA), a central component of the cAMP signalling cascade, and one of its substrates, RAP1, were reported in the platelets of bipolar

patients (Perez et al., 2000). Abnormalities of the PI system, another second messenger pathway, were found in post-mortem brains of subjects with BPD (see Gould and Manji, 2002 for review). The second messenger system is considered a credible area of research with support, further than that listed above.

Further research points to the disruption of inherent biological rhythms, such as the sleep-wake cycle, hormonal rhythms, temperature regulation and circadian rhythms, as an underlying mechanism in BPD. This is supported for example by the study which showed the secretion of the hormone melatonin, an indolamine derivative of serotonin, which is involved in the timing of circadian rhythm information, as frequently abnormal in patients with BPI (Nurnberger et al., 2000). While this points to the diversity of proposed etiologic mechanisms and the breadth of BPD research, it is not an aspect which is pursued beyond this chapter.

### 1.4.4 Functional and structural brain abnormalities in BPD

Extricating the disease pathology of BPD from the rest of the mood disorders (so that etiologically distinct categories can be formed), as well as defining its link with brain structure and function, is a challenging task. In addition to neuro-anatomical and clinical data, the explosion of brain-imaging technologies in the last decade (such as magnetic resonance imaging (MRI) and positron emission tomography (PET)) is adding considerably to the area of functional neurobiology.

PET images of cerebral blood flow and the rate of glucose metabolism, in bipolar and unipolar depressive patients, identified abnormally decreased activity in the prefrontal cortex (PFC) ventral to the genu of the corpus callosum in the left hemisphere (Drevets et al., 1997). The implicated neuronal region forms part of the multiplex circuitry of emotion and is critically involved in the dysfunctional states of depression and mania. Importantly, the subgenual PFC has been shown to be involved in the regulation of neurotransmitter systems.

Analysis of lesions of the subgenual PFC in humans and experimental animals, implicated this area in the regulation of visceral responses to stimuli incited by stressful and emotional experiences. MRI indicated a corresponding reduction in the mean volume of grey matter of the subgenual PFC in BPD subjects (Drevets et

al., 1997). Furthermore, MRI images of patients with BPD suggested subregion-specific grey matter volume reductions in the PFC (Lopez-Larson et al., 2002).

Information gathered from post-mortem brain studies could supplement and guide neuro-imaging studies in psychiatric patients. Preliminary post-mortem histopathological assessments of subgenual PFC tissue from subjects with MDD and BPD suggest that the abnormal reduction in grey matter volume is associated with a reduction in glia (Drevets et al., 1998). However, no decrement in neuronal number was observed. Recently, Uranova and co-workers (2004) reported a significant reduction in the density of oligodendroglial cells in the PFC of post-mortem human brain specimens of BPD patients.

Despite the notable body of neuro-imaging and neuropathology studies indicating the occurrence of brain changes in BPD, their aetiology and their relation to phenotypic characteristics are yet to be established. Malfunctioning of the subgenual PFC may upset autonomic, neurotransmitter, neuroendocrine, motor and somatosensory responses, which together bring about human emotion (Damasio, 1997). It is through the understanding of the process of neurochemical transmission that psychopharmacology aims to unravel the mechanism of drug action on the brain (Stahl, 2000).

## 1.5 PSYCHOPHARMACOLOGY

The advent of the anti-psychotic drug, chlorpromazine, in 1952, marked the beginning of psychopharmacology (Catalano, 1999). Psychopharmacology heeds the impact of diseases on the CNS, and explains the behavioural effects of psychiatric medicines. While this project does not focus on treatment, but rather on the genetic susceptibility of the disease, it is important to note that an increase in genetic understanding could aid treatment.

The treatment of BPD is variable and reflects the heterogeneity the clinical phenotype of the disease. Treatment modalities in BPD include electroconvulsive therapy, psychotherapy, sleep deprivation, phototherapy and psychopharmacology (reviewed in Compton and Nemeroff, 2000). One of the main aims in the treatment of BPD is mood stabilisation (Griel and Kleindienst, 2003). The desired effects of

mood stabilisers are antimanic, antidepressant (Goodwin et al., 1972), and prophylactic efficacy. It has been proposed that the therapeutic agent should improve at least one phase of illness, without worsening another (Keck and McElroy, 2003).

Amongst the most established mood stabilisers (lithium, valproate and carbamazepine) (Gould et al., 2002), lithium comes the closest to meeting the criteria of a comprehensive mood stabiliser. Since its discovery, lithium is considered the cornerstone of pharmacologic treatment of persons affected with BPD (reviewed in Manji et al., 1999; Lenox and Hahn, 2000). Lithium has been shown to deplete inositol (Berridge et al., 1989), inhibit the enzyme glycogen synthase kinase (GSK3) (Klein and Melton, 1996), and has the neuroprotective effect of rendering cells less susceptible to apoptosis (Hashimoto et al., 2002; Brunello, 2004). In addition, the anticonvulsants, valproate and carbamazepine, are effective antimanic agents in the treatment of patients with BPD (Keck et al., 1998; Keck and McElroy, 2003). In an experiment of explants of rat sensory neurons from dorsal root ganglia cultured in lithium, valproate and carbamazapine, not only was growth cone collapse inhibited, but an increase in growth cone area was observed (Williams et al., 2002). Upon the addition of inositol, these effects were abolished, thereby implicating inositol phosphate signalling in the actions of lithium, valproate and carbamazapine (and also in BPD). Furthermore, the treatment of BPD patients with lithium, valproate or carbamazapine has been shown to be associated with reduced risk of suicidal behaviour (Yerevanian et al., 2003).

Mood stabilisation is also important in the treatment of the disabling depressive phases of BPD, which may sometimes be dominated by mixed episodes (Kasper, 2003). Antidepressants include selective serotonin inhibitors (SSRIs) (for example fluoxetine, paroxetine and setraline) and MAO inhibitors (for example tranylcypromine, imipramine and moclobemide) (reviewed in Compton and Nemeroff, 2000). However, the latter may induce mania and psychosis in vulnerable individuals (Stoll et al., 1994; Preda et al, 2001). The Expert Consensus Guidelines (Francis et al., 1998) prefer bupropion, which is believed to have a lower risk of switching the patient into a manic phase, and SSRIs as first-line treatment for depression in BPD (Frances et al., 1998). The anti-epileptic drug lamotrogine has emerged as a promising new prophylactic treatment for bipolar depression

(Calabrese et al., 1999). In addition to its anti-psychotic properties, olanzapine also has antidepressant effects (Rothschild et al., 1999).

Antidepressant treatment and stress have been shown to target signal transduction and gene expression circuits associated with neuroplasticity in patients with mood disorders (Duman, 2002; Shimizu et al., 2003). Two such targets include CREB and *BDNF*, a major gene regulated by CREB (Duman et al., 1997). Antidepressant administration has been reported to increase the expression of CREB and *BDNF* in rat hippocampus (Nibuya et al., 1996).

Appropriate drug therapy for BPD is complicated, as the usefulness of psychotropic drugs in the treatment of patients has proved to be variable in outcome and side effects. Tohen and co-workers (2002) reported that up to 40% of patients with BPD respond poorly to either lithium or valproate. Thus, monopharmacological (one drug) treatment strategies appear to have different degrees of efficacy (Calabrese et al., 1999; Greil and Kleindienst, 2003), with some displaying notable effects on the depressive phase, the manic phase, psychotic symptoms, or decreasing risk of suicidal behaviour. Instead of treating patients with one mood stabilizer, psychiatrists commonly prescribe combination therapy, combining anti-psychotics (for example olanzapine and risperidone), anti-epileptics (for example lamotrigine and carbamazepine) or anti-depressants (for example imipramine and paroxetine) with mood stabilizers (reviewed by Bowden, 2004). Dependent on the individual patient's manifesting symptoms, disease history and comorbid illness, successful management of BPD needs both acute and chronic (maintenance) treatments of both mania and depression.

## 1.6 BPD: A MULTIFACTORIAL DISORDER

Only a few decades ago psychologists were convinced that human behavioural characteristics and aberrancies could be explained entirely by environmental influences (Sherman et al., 1997). In addition, and classically, psychiatry has focussed its attention on the pathogenicity of the disorder, rather than taking the contributions of heredity into account (Barondes, 1999). In more recent years heredity and the possibility of having a multifactorial pathology (i.e. not a single but many genes with varying effect and interacting with exogenous influences) to the

aetiology of psychiatric conditions has been considered. Several aetiological mechanisms, as discussed in the following sections 1.6.1-1.6.3, are thought to contribute to BPD, owing to the observed bipolar spectrum of clinical phenotypes together with non-Mendelian inheritance patterns in families with the disorder (Lander and Schork, 1994; Craddock and Owen, 1996; Craddock and Jones, 1999).

### 1.6.1 Environmental stimuli

Several studies have implicated psychosocial stressors in the timing, type and outcome of bipolar episodes (Glassner and Haldipur, 1983; Bidzinska, 1984; Miklowitz et al., 1988). These stressors include health, marital and family conflicts, work overload, and emotional and ambitional failures (Bidzinska, 1984). Previously, it was reported that post discharge from hospital, important predictors of the bipolar patient's subsequent clinical course of illness may include: (i) the emotional quality of the family environment (Miklowitz et al., 1988), (ii) the season of the year, (iii) variation in endocrine status, and (iv) drug treatment (Silverstone and Romans-Clarkson, 1989). The onset of mania has also, interestingly, been reported in association with unpleasant life events, in the form of loss or threat (Ambelas, 1979; Ambelas, 1987; Bebbington et al., 1993). This response to stress is immediate and selective, possibly providing a transient shield or buffer which makes the succeeding depression tolerable or even ineffectual (Ambelas, 1987). As observed by Glassner and Haldipur (1983), stressful life events may constitute different etiological components in early and late onset BPDs. The contribution of environmental components to illness is further compounded by the underlying biological predisposition.

### 1.6.2 Evidence of heredity

Genetics offers the ability to dissect the "heritable" aspects/components of disease towards identifying the underlying genes and biology. The measure of heritability describes the relative genetic contribution to any complex trait (Ellsworth and Manolio, 1999a). In the broad sense, heritability is the estimated percentage of the total phenotypic variation owing to genetic influences, relative to environmental factors and measurement error.

Twin studies offer the most compelling evidence for the presence of genetic effects in complex disorders (McGuffin and Katz, 1989; Ghosh and Collins, 1996; Kelsoe, 2003). It has been shown that the concordance rate for BPD, on average, in monozygotic twins is 65%, and 14% in dizygotic twins (McGuffin and Katz, 1989), reflecting an estimated heritability as high as 80%. In these studies, both monozygotic and dizygotic twin subjects were reared together which suggested that sharing all (as opposed to 50%) genes elevates the risk of disease over twofold (Kelsoe, 2003). Twin data have also suggested a genetic component to suicide that may represent a predisposition to the associated psychiatric illness (Roy et al., 1991).

Adoption studies provide another means of separating the hereditary (nature) and environmental factors (nurture) interacting in the aetiology of the disorder (Mendlewicz, 1977; Mitchell et al., 1993; MacKinnon et al., 1997). These investigations have shown that in the presence of genetic effects, the biological parents of an individual with BPD exhibited relatively high BPD rates compared with the adoptive parents.

Another effective measure to describe the genetic contribution to a trait is relative risk, $\lambda_R$ (Risch, 1990a; Ghosh and Collins, 1996). $\lambda_R$ is defined as the recurrence risk for a relative (R) (such as sibling (S) and offspring (O)) of an affected individual, divided by the recurrence risk for the general population (Lander and Schork, 1994). The greater the value of $\lambda_R$, the greater the scale of concordant inheritance for genetic elements is, in affected relative pairs. For BPD, Gershon and colleagues (1982) suggested an increase in the $\lambda_R$ for siblings by a factor of about eight to nine. However, because of BPD's multigenic aetiology, in which many loci add to the inflated $\lambda_R$, it is predicted that the locus-specific $\lambda_R$ would be much less than nine, owing to their minor to moderate effects.

## 1.6.3 Genetic mechanisms involved in BPD

Several genetic mechanisms are known to generate complex patterns of inheritance, including:

*(i) Heterogeneity*

Any one of multiple genes may underlie similar phenotypes in different families, such as a wide concert of genes interacting in a common biochemical pathway (Lander and Schork, 1994; Craddock and Owen, 1996).

*(ii) Epistasis*

Two or more products of disease genes may interact to produce a phenotype (Craddock and Owen, 1996).

*(iii) Incomplete penetrance*

The inheritance of a predisposing allele (genotype) does not result in the manifestation of the phenotype (Lander and Schork, 1994; Craddock and Owen, 1996). Some individuals, however, are phenocopies, that is they do not inherit a predisposing allele but nonetheless express the phenotype due to environmental/random influences (Lander and Schork, 1994), or could simply be due to the effects of other genes.

*(iv) Anticipation*

A clinical phenomenon of increasing severity of disease phenotype observed at an earlier age in successive generations (Lander and Schork, 1994; Craddock and Owen, 1996; MacKinnon et al., 1997).

*(v) Imprinting*

Given the same genetic defect, variation in phenotype between family members may occur due to differential activity of the maternal and paternal copies of a gene (Lander and Schork, 1994; Craddock and Owen, 1996; MacKinnon et al., 1997), the so-called parent-of-origin effect (McMahon et al., 1995; Gershon et al, 1996).

*(vi) Mitochondrial inheritance*

Another plausible genetic mechanism is that of mitochondrial inheritance, which can explain the maternal inheritance pattern of BPD, since mitochondria are transmitted exclusively from a mother (Craddock and Owen, 1996).

*(vii) Variable expressivity*

A variety of related clinical phenotypes can result from the same genes (Kelsoe, 2003).

Some of these genetic mechanisms have been suggested as possible explanations for BPD (Table 1.1).

| Table 1.1: Genetic mechanisms involved in BPD | |
|---|---|
| Genetic mechanism | References |
| Heterogeneity | Hodgkinson et al., 1987 |
| Epistasis | Craddock et al., 1995 |
| Anticipation | McInnes et al., 1993; Nylander et al., 1994 |
| Imprinting | McMahon et al., 1995 |
| Mitochondrial inheritance | McMahon et al., 1995 |

It is possible that two or more of these potential genetic mechanisms interact in a single individual, thereby adding to the complexity of the disease.

## 1.7 MOLECULAR GENETICS OF BPD

### 1.7.1 Models of BPD inheritance

Unlike disorders with Mendelian patterns of inheritance, such as Huntington disease, BPD is unlikely to have a specific mode of inheritance (Barondes, 1999). This is mainly attributed to the spectrum of disorders characteristic of BPD. However, several genetic models for inheritance, as discussed in this section of the dissertation, have been proposed for the complex phenotype of BPD.

In psychiatric disorders, including BPD, an at-risk individual does not inherit the certainty of becoming ill, but instead may be predisposed to illness as a result of environmental and genetic components, termed "liability" (McGuffin and Katz, 1989; Mitchell et al., 1993). Therefore, people whose liability surpasses a certain threshold value, present with the disorder. The simplest form of "liability threshold" model is the single major locus (SML) model, which states that a single gene is responsible for all the genetic variation in liability (McGuffin and Katz, 1989; Mitchell et al, 1993). However, the effect of the major contributing locus can be influenced by reduced penetrance, variable expressivity and phenocopies (Kelsoe, 2003). Furthermore, different major genes may function in each family.

Alternatively, the multifactorial polygenic (MFP) model, assumes that the combined effect of multiple genes and environmental factors contribute to an at-risk individual's liability in an additive fashion (McGuffin and Katz, 1989; Mitchell et al, 1993). Galton first proposed this polygenic or quantitative transmission model (as cited by Kelsoe, 2003). Risch (1990a) described two types of multilocus models: (i) an additive model, which closely approximates genetic heterogeneity, characterised by no interlocus interaction, and (ii) a multiplicative model with locus interaction. The multiplicative or epistatic model serves as a basis for the methodologies implemented in the present study. Such genes that underlie scores on continuous measures, including blood pressure, weight, height, serum cholesterol levels and body-mass index, have been termed quantitative trait loci (QTLs) (Ghosh and Collins, 1996; Craddock and Owen, 1996). Simplified, quantitative traits derive from many genes, each of modest effect (Kelsoe, 2003). Traditionally, psychiatric genetics has focussed its attention on categorical phenotypes (Craddock and Owen, 1996). However, if it can be shown that a suitable continuous measure exists that is biologically associated with the disorder, a QTL approach will prove to be practical and possibly impart valuable information about the disorder itself.

In the multiple threshold model, Reich et al. (1972) extended the quantitative trait model to include qualitatively different phenotypes lying on a spectrum. Assuming that these phenotypes are quantitatively ranked in relationship with each other, as in the case of the bipolar spectrum disorder (Kelsoe, 2002), different thresholds of a latent quantitative trait lead to different disorders.

Yet another and a fourth possible mechanism for the complex inheritance pattern of BPD is the mixed model. This model favours the presence of a single major locus against a multifactorial background, which could include the effects of birth cohort and age of onset (Rice et al., 1987). All these genetic models point to the highly variable approaches adopted, to elucidate the specific genetic factors that may explain the familial transmission of BPD, and other complex diseases.

### 1.7.2 Genes implicated in BPD

A number of susceptibility genes have been implicated in BPD (Table 1.2), either by genome-wide scans or the candidate gene approach. Failure to detect major gene

effects may indicate that individual BPD susceptibility loci represent small to moderate effects, and are required in combination (multiplicative or epistatic inheritance) (Stoltenberg and Burmeister, 2000), and this is what the present study hopes to build on.

Molecular genetic studies have also displayed evidence for overlapping genetic vulnerabilities in BPD and other psychiatric conditions, particularly schizophrenia (Table 1.3). Recently, Mazaide and co-workers (2004) identified a novel susceptibility locus on chromosome 15q26 shared by schizophrenia and BPD. Their findings support the concurrent investigation of schizophrenia and BPD in the same study. Based on studies, such as those listed in Table 1.3, BPD and other psychiatric conditions are likely to be expressions of the same genetic susceptibility, creating a phenotypic spectrum. This implies that current nosology of these disorders will require substantial redefinition during the next decade to model their shared genetic vulnerability, as specific genes and their derived biologies are identified (Berrettini, 2000).

Table 1.2: Summary of candidate loci implicated in BPD

| Locus and Reference | Source Population[a] | Etnicity[b] | SM[c] | Type of analysis[d] | Findings[e] |
|---|---|---|---|---|---|
| 1q32 | | | | | |
| Turecki G et al., 1995 | 10 cases, 10 cons | Bra | CA | CC | BPI was associated with a fragile site on 1q32 |
| Detera-Wadleigh et al., 1999 | 22 fams | Nam+OOA | GS | Non-par. | Lod= 2.67, *P*= 0.00022 at D1S1660-D1S1678 |
| | | | | | |
| 1q42 | | | | | |
| Blackwood et al., 2001 | 1 large fam | Br | CR | Par. | Max lod= 4.5 at translocation at 1q42 |
| Macgregor et al., 2004 | 45 fams | Scot | GS | Par. | lod= 2.63 at D1S103 |
| | | | | | |
| 2p13-16 | | | | | |
| Lui et al., 2003 | 40 fams | Nam+Is | GS | Par. | 2P lod= 3.2 at D2S441 |
| McInnis et al., 2003b | 65 fams | Am | GS | Non-par. | NPL= 2.54 at D2S99 |
| | | | | | |
| 2q | | | | | |
| Zubenko et.al., 2003 | 81 fams | Am | CG | MA | *CREB1* associated with depressive disorders |
| Papiol et al., 2004 | 88 cases, 176 cons | Spanish | CG | CC | Interleukin-1 cluster is associated with BPD |
| | | | | | |
| 3q | | | | | |
| Schosser et al., 2004 | 5 fams | Austrian | CR | Non-par. | NPL score Z all = 4, *P*= 0.000128 at D3S1265 on 3q29 |
| Bailer et al., 2002 | 5 fams | Austrian | GS | Non-par. | NPL score Z all = 3.74, *P*=.0003 at D3S1265 |
| Benedetti et al., 2004 | 185 cases | Italian | CG | Assoc. | *GSK3-beta* associated with age at onset in BPD |
| | | | | | |
| 4p16 | | | | | |
| Blackwood et al., 1996 | 12 families | Scot | GS | Par. | 2P Lod= 4.1 at D4S394 |
| Ewald et al., 1998 | 2 families | Dan | CR | Par. | 2P Lod= 2.0 at D4S394 |
| Detera-Wadleigh et al., 1999 | 22 fams | Nam+OOA | GS | Non-par. | *P*= 0.0022 at D4S2408-D4S2632 |
| Als et al., 2004 | 17 cases, 44 cons | Faroese | | LDM | *P*= 0.0162 at two-marker segment D4S394-D4S2983 |

TABLE 1.2: *Continued*

| Locus and Reference | Source Population[a] | Etnicity[b] | SM[c] | Tye of analysis[d] | Findings[e] |
|---|---|---|---|---|---|
| 4q35 | | | | | |
| Adams et al., 1998 | 1 fam | Aust | GS | Non-par. | NPL= 3.74, *P*=.0003 |
| Willour et al., 2003 | 56 fams | Nam | GS | Non-par. | allele-sharing lod= 2.49 between D4S3335-D4S2390 |
| McInnis et al., 2003b | 65 fams | Am | GS | Non-par. | NPL= 2.8 at D4S1629 |
| Lui et al., 2003 | 40 fams | Nam+Is | GS | Par. | 2P lod= 3.16 at D4S1625 |
| | | | | | |
| 5p15.3 | | | | | |
| Kelsoe et al., 1996 | 25 fams | Nam+OOA+Isl | CR | Par. | Max lod= 2.38 at D5S392 |
| Greenwood et al., 2001 | 50 trios | Nam+OOA | CG | TDT | LD between Dopamine transporter and BPD |
| | | | | | |
| ter5q-q35 | | | | | |
| Coon et al., 1993 | 8 fams | Nam | GS | Par. | Max lod >1 at D5S39, D5S43, D5S62 |
| Edenberg et al., 1997 | 97 fams | Nam | GS | ASP | *P*= 0.008 at D5S1456 |
| Detera-Wadleigh et al., 1999 | 22 fams | Nam+OOA | GS | Non-par. | Lod= 1.7, *P*= 0.0026 in the region of D5S498-D5S408 |
| | | | | | |
| 6pter-p24 | | | | | |
| Ginns et al., 1996 | 5 fams | OOA | GS | Non-par. | *P*= 0.0003 at D6S7 |
| | | | | | |
| 8q24 | | | | | |
| Cichon et al., 2001a | 75 fams | Ger+It+Is | GS | Par. | 2P lod= 3.619 at D8S514 |
| McInnis et al., 2003b | 65 fams | Am | GS | Non-par. | NPL= 3.13 at D8S256 |
| Avramopoulos et al., 2004 | 65 fams | Nam | CR | Par. | lod= 3.32 at D8S256 |
| | | | | | |
| 9q31-34 | | | | | |
| Badenhop et al., 2002 | 13 fams | Br/Irish | GS | Par. | Hlod= 1.74 at D9S1776 |
| Mundo et al., 2003 | 288 trios | Eur | CG | TDT | *GRIN1* confers susceptibility to BPD |
| Lui et al., 2003 | 40 fams | Nam+Is | GS | Par. | 2P lod> 1.9 at D9S938 |

TABLE 1.2: *Continued*

| Locus and Reference | Source Population[a] | Etnicity[b] | SM[c] | Tye of analysis[d] | Findings[e] |
|---|---|---|---|---|---|
| 10p | | | | | |
| Foroud et al., 2000 | 97 fams | Nam | GS | Non-par. | lod= 2.5 at D10S1423 |
| | | | | | |
| 10q21-26 | | | | | |
| Cichon et al., 2001b | 75 fams | Eur | | Non-par. | Max NPL(all) score= 3.12, *P*= 0.0013 at segment D10S1483-D10S217 |
| Lin et al., 2003 | 100 cases, 106 cons | Taiwanese | CG | CC | Association between *5HT7* and BPD |
| | | | | | |
| 11p15 | | | | | |
| Bellivier et al., 1998 | 152 cases, 94 cons | France | CG | CC | Association between *TPH* and BPD |
| Neves-Pereira et al., 2002 | 283 nuclear fams | Eur | CG | Assoc. | |
| Sklar et al., 2002 | 136 T | Nam+Br | CG | TDT | *BDNF* potential risk locus for BPD |
| | | | | | |
| 12p | | | | | |
| Faroane et al., 2004 | 97 fams | Nam | GS | VC | lod> 2.5 near D12S1292 |
| Willeit et al., 2003 | 172 cases, 143 cons | Austrian | CG | CC | Association G-protein coupled signal transduction in the aetiology of affective disorders |
| | | | | | |
| 12q23-q24 | | | | | |
| Dawson et al., 1995 | 45 fams | Br+Ger+Wel | CR | Non-par. | *P*< 0.053 at D12S78 |
| Detera-Wadleigh et al., 1999 | 22 fams | Nam+OOA | GS | Non-par. | Lod= 1.24, *P*= 0.0084 in the region of D12S1343-D12S2070 |
| Shink et al., 2004 | 20 fams | Can | GS | Par. | lod= 3.35 at D12S378 |
| | | | | | |
| 13q12-q13 | | | | | |
| Ginns et al., 1996 | 5 fams | OOA | GS | Non-par. | *P*= 0.0003 at D13S1 |
| McInnis et al., 2003b | 65 fams | Am | GS | Non-par. | NPL= 2.53 at D13S1493 |
| | | | | | |
| 13q32-q34 | | | | | |
| Liu et al. 2001 | 22 fams | Nam | CR | Non-par. | max lod= 3.25, *P*= 0.00546 around D13S779-D13S225 |
| Hattori et al., 2003 | 557 samples | Eur | CG | TDT | Association between *G72/G30* ad BPD |
| Shaw et al., 2003 | 32 fams | Nam | CR | Par. | lod= 2.26 at D13S154 |
| Schumacher et al., 2004 | 300 cases, 300 cons | Ger | CG | CC | *G72* was associated with BPD |

TABLE 1.2: *Continued*

| Locus and Reference | Source Population[a] | Etnicity[b] | SM[c] | Tye of analysis[d] | Findings[e] |
|---|---|---|---|---|---|
| 14q | | | | | |
| Lui et al., 2003 | 40 fams | Nam+Is | GS | ASP | lod= 2.36 at D4S306 |
| Faroane et al., 2004 | 97 fams | Nam | GS | VC | lod= near GATA50C |
| 15q | | | | | |
| Ginns et al., 1996 | 5 fams | OOA | GS | Non-par. | *P*= 0.0003 at D15S45 |
| Papadimitriou et al., 1998 | 48 cases, 50 cons | Greek | CG | CC | Association between *GABRA5* and BPD |
| Maziade et al., 2004 | 21 fams | Can | GS | Par. | Z= 4.59 at D15S122 |
| Faroane et al., 2004 | 97 fams | Nam | GS | VC | lod= 3 near GATA31B |
| | | | | | |
| 16p13 | | | | | |
| Ewald et al., 1995 | 2 fams | Dan | CR | Par. | lod= 2.65 at D16S510 |
| Nyegaard et al., 2002 | 80 cases, 144 cons | Dan | CG | CC | Association between Somatostatin receptor 5 and BPD |
| Itokawa et al., 2003 | 96 fams | Nam | CG | PDT | Association between *GRIN2A* and BPD |
| Maziade et al., 2004 | 21 fams | Can | GS | Par. | Z= 4.10 at D16S1145 |
| | | | | | |
| 17q | | | | | |
| Dick et al., 2003 | 250 fams | Nam | GS | Non-Par. | max lod score=3.63 at D17S928 |
| Lui et al., 2003 | 40 fams | Nam+Is | GS | Par. | 2P lod> 1.9 at D17S921 |
| | | | | | |
| 18p11 | | | | | |
| Berrettini et al., 1994 | 22 fams | Nam+ OOA | CR | ASP | *P*< 0.001 at D18S21 |
| Sjoholt et al., 2004 | 44 cases, 48 cons + 75 T | Nor+Arabian | CG | Assoc. | Association between *IMPA2* and BPD |
| Washizuka et al., 2004 | 189 cases, 222 cons + 105 T | Jp+Nam | CG | Assoc. | association between *NDUFV2* and BPD |
| | | | | | |
| 18q | | | | | |
| McMahon et al., 1997 | 58 fams | Am | CR | Non-Par. | NPL = 2.84, *P*<0.0019 at D18S38 |
| McInnis et al., 2003b | 65 fams | Am | GS | Non-par. | NPL= 2.9 at D18S878 |
| Maziade et al., 2004 | 21 fams | Can | GS | Par. | Z= 4.05 at D18S410 |
| | | | | | |
| 19q12-q13 | | | | | |
| Badenhop et al., 2002 | 13 fams | Br/Irish | GS | Non-par. | NPL= 4.55, P= 0.0002 between D19S414 and D19S220 |

TABLE 1.2: *Continued*

| Locus and Reference | Source Population[a] | Etnicity[b] | SM[c] | Tye of analysis[d] | Findings[e] |
|---|---|---|---|---|---|
| 20p | | | | | |
| Willour et al., 2003 | 56 fams | Nam | GS | Non-par. | allele-sharing lod= 1.82 at D20S162 |
| | | | | | |
| 21q | | | | | |
| Detera-Wadleigh et al., 1996 | 22 fams | Nam | CR | ASP | *P*< 0.01 between D21S270-D21S171 |
| Aita et al., 1999 | 57 fams | Am+Is | CR | Non-par. | 2P heterogeneity lod= 3.35 at D21S1260 |
| | | | | | |
| 22q11 | | | | | |
| Lachman et al., 1997 | 17 fams | Nam+OOA | CR | Par. | Max lod= 2.51 at D22S303 |
| Kelsoe et al., 2001 | 20 fams | Nam | GS | Par. | Max lod= 3.8 at D22S278 |
| Barrett et al., 2003 | 153 nuclear fams + 275 T | Nam+Eur | | TDT | *P*= 0.0019 at P-5 variant |
| Shifman et al., 2004 | 217 cases, 4050 cons | Is | CG | Assoc. | association between *COMT* and BPD |
| | | | | | |
| Xq24 | | | | | |
| Ekholm et al., 2002 | 41 fams | Finnish | CR | Par. | max lod= 2.78 at DXS1047 |
| Lin et al., 2003 | 100 cases, 106 cons | Taiwanese | | CC | Association between *5HT2C* and BPD |

[a]: cons = controls, fams = families, T = trios

[b]: Am = American, Aust = Australian, Bra = Brazilian, Br = British, Can = Canadian, Dan = Danish, Eur = European, Fin = Finish, Ger = German, Isl = Islandic, Is = Israeli, It = Italian, Jp = Japanese, Nam = North American, Nor = Norwegian, OOA = Old Order Amish, Scot = Scottish

[c]: SM = screening method, CG = candidate gene, CR = candidate region, GS = genome scan

[d]: 2P = two point, Assoc = association, CC = case-control, LDM = linkage disequilibrium mapping, Non-par = non-parametric, Par. = parametric, PDT= pedigree disequilibrium test, TDT = transmission disequilibrium test, VC = variance components

[e]: *5HT2C* = serotonin receptor 2C, *5HT7* = serotonin receptor 7, *COMT* = catechol-o-methyltransferase, *CREB1* = cAMP response element-binding protein 1, *GABRA1* = gamma-aminobutyric acid receptor, alpha-1, *GABRA5* = gamma-aminobutyric acid receptor, alpha-5, *GABRA6* = gamma-aminobutyric acid receptor, alpha-6, *GRIN1* = glutamate receptor, ionotropic, n-methyl-d-aspartate, subunit 1,

*GRIN2A* = glutamate receptor, ionotropic, n-methyl-d-aspartate, subunit 2A, *GSK3-beta* = glycogen synthase kinase beta,

*IMPA2* = myo-inositol monophosphatase 2, LD = linkage disequilibrium, *NDUFV2* = NADH-ubiquinone oxidoreductase flavoprotein 2, *TPH* = tryptophan hydroxylase

Table 1.3: Overlap of genetic regions and candidate genes implicated by linkage and association respectively in psychiatric disorders.

| Locus | Phenotype | | | | | | | Reference |
|---|---|---|---|---|---|---|---|---|
| | BPD | SCZ | ADHD | ALC. | NS | AT | OCD | |
| 1q21 | * | * | | | | | | Chandy et al., 1998 |
| 1q32-q41 | | * | | | | | | Hallmayer, 1999 |
| 1q32 | * | | | | | | | Detera-Wadleigh et al., 1999 |
| 1q42 | * | * | | | | | | Macgregor et al., 2004 |
| 1p22.3-1p21 | * | | | | | | | Rice et al., 1997 |
| 1p22-p21 | | | | * | | | | Hallmayer, 1999 |
| | | | | | | | | |
| 2p22.1 | | * | | | | | | Hallmayer, 1999 |
| 2p22.1 | | | | * | | | | Hallmayer, 1999 |
| 2p13-16 | * | | | | | | | Liu et al., 2003 |
| 2q31-q33. | | | | | | * | | Rabionet et al., 2004 |
| 2q33-35 | * | | | | | | | Zubenko et al., 2002 |
| | | | | | | | | |
| 3q13.3 | * | | | | | | | Parsian et al., 1995 |
| 3q13 | | * | | | | | | Maziade et al., 2001 |
| 3q27-28 | * | | | | | | | Liu et al., 2003 |
| 3q29 | * | * | | | | | | Schosser et al., 2004 |
| | | | | | | | | |
| 4q31 | | * | | | | | | Kennedy et al., 1999 |
| 4q31 | * | | | | | | | Liu et al., 2003 |
| 4p16.1 | * | * | | | | | | Als et al., 2004 |
| DRD5 (4p16.1-15.3) | | | * | | | | | Bobb et al., 2004 |
| | | | | | | | | |
| 5p15.31-15.1 | * | | | | | | | Shink et al., 2002 |
| DAT (5p15.3) | | | * | | | | | Bobb et al., 2004 |
| 11.2-q13.3 | | * | | | | | | Sherrington et al., 1993 |
| 5q11 | * | | | | | | | McInnis et al., 2003a |
| 5q23.3-q31.1 | | * | | | | | | Crowe et al., 1999 |
| 5q31.3-q35.1 | * | | | | | | | Crowe et al., 1999 |
| 5q33.3 | | | * | | | | | Arcos-Burgos et al., 2004 |
| | | | | | | | | |
| 6p24-p22 | | * | | | | | | Straub et al., 1995 |
| MOG (6p21.3) | | | | | | | * | Zai et al., 2004 |
| 6p | | * | | | | | | Schwab et al., 2000 |
| 6p | * | | | | | | | Schulze et al., 2004 |
| 6p22-p24 | | * | | | | | | Maziade et al., 2001 |
| | | | | | | | | |
| 8p23.1-p22 | * | | | | | | | Niculescu et al., 2000 |
| 8p23.1-p22 | | * | | | | | | Kaufmann et al., 1998 |
| 8p23 | * | | | | | | | Liu et al., 2003 |
| 8p21-p22 | * | | | | | | | Wildenauer et al., 1999 |
| 8p21-p22 | | * | | | | | | Brzustowicz et al., 1999 |

Table 1.3: *Continued*

| Locus | Phenotype | | | | | | | Reference |
|---|---|---|---|---|---|---|---|---|
| | BPD | SCZ | ADHD | ALC. | NS | AT | OCD | |
| 10p12 | * | | | | | | | McInnis et al., 2003a |
| 10p11-p14 | | * | | | | | | Maziade et al., 2001 |
| 10q | * | | | | | | | Wildenauer et al., 1999 |
| 10q | | * | | | | | | Levinson et al., 1998 |
| 10q24 | * | | | | | | | Liu et al., 2003 |
| | | | | | | | | |
| 11p15 | * | | | | | | | McInnis et al., 2003 |
| DRD4 (11p15.4-15.5) | | | | | * | | | Ebstein et al., 1996 |
| DRD4 (11p15.5) | | | * | | | | | Bobb et al., 2004 |
| DRD4 (11p15.4-15.5) | | | | | | | * | Hemmings et al., 2004 |
| *BDNF* (11p13-15) | | | | | | | * | Hall et al., 2003 |
| *BDNF* (11p13-15) | | | | * | | | | Matsushita et al., 2004 |
| | | | | | | | | |
| 13q32 | | * | | | | | | Blouin et al., 1998 |
| 13q32 | * | | | | | | | Detera-Wadleigh et al., 1999 |
| | | | | | | | | |
| 15q13-q15 | * | | | | | | | Turecki et al., 2001 |
| 15q13-q15 | | * | | | | | | Leonard et al., 1998 |
| 15q11-q13 | | | | | | * | | Cook et al ., 1998 |
| | | | | | | | | |
| 16p13 | * | | | | | | | Ewald et al., 1995 |
| 16p13 | | | | * | | | | Detera-Wadleigh et al. 1999 |
| 16p13 | * | | | | | | | McInnis et al., 2003 |
| | | | | | | | | |
| 17q11-12 | * | | | | | | | Liu et al., 2003 |
| 5-HTT (17q11.1-q12) | | | * | | | | | Bobb et al., 2004 |
| | | | | | | | | |
| 18q21 | * | | | | | | | McMahon et al. 1997 |
| 18q21.1 | | * | | | | | | Van Broeckhoven et al. 1999 |
| 18q12-q21 | * | * | | | | | | Maziade et al., 2004 |
| 22q11 | | * | | | | | | Karayiorgou et al. 1995 |
| 22q11 | * | | * | | | | | Lachman et al. 1997 |
| 22q11 | | | | | | | * | Schindler et al. 2000 |
| | | | | | | | | |
| MAOA (Xp11.23) | * | | | | | | | Lin et al. 2000 |
| Xp11 | | * | | | | | | Wei and Hemmings 2000 |
| MAOA (Xp11.23) | | | | | | | * | Karayiorgou et al. 1999 |

Key: BPD- bipolar disorder; SCZ- schizophrenia; ADHD- attention-deficit hyperactivity disorder; ALC.- alcohol abuse; NS- novelty-seeking; AT- autism; OCD- obsessive compulsive disorder

## 1.8 ANALYTICAL STRATEGIES

Gene identification can involve one of two approaches. The first is a general survey of the entire human genome, i.e. genome-wide scans (Elsworth and Manolio, 1999b). The second is the candidate gene approach, which considers genes based on their location in a previously identified genomic region or because they are suspected to have a role in the physiology of a phenotype/trait (Glatt and Freimer, 2002). The identification of genes via the candidate gene approach has not been fully realised, because the biology underlying many complex diseases, including physical processes and biochemical pathways, is not yet known (Ellsworth and Manolio, 1999b).

In the case of a phenotype with an unknown inheritance pattern, identification of susceptibility genes over a range of genetic models requires the type and size of study sample to be chosen in a manner that will provide optimal statistical power (Craddock and Owen, 1996). Power is a function of the number of genes present and the manner in which these contribute to the disease (Weeks and Lathrop, 1995). Linkage analysis provides the method of measuring the genetic contribution to disease.

### 1.8.1 Linkage analyses

Statistical power to detect linkage in BPD is determined by the locus-specific value of $\lambda_R$ (Risch, 1990b). Moreover, in the case of multiple loci, the ease or difficulty of mapping loci for a given phenotype depends on the value of the largest of the locus-specific $\lambda_R$ values for the contributing loci. In other words, a disorder with a moderate value of $\lambda_R$ may have only loci with individual small effect sizes, which will be quite difficult to map, or it may have one or a few loci accounting for most of the familial effect, and those loci will be relatively easier to map. For example cystic fibrosis has a value for $\lambda_s$ (the risk to a sibling of an affected individual compared to the population prevalence) of about 500, much larger than the $\lambda_s$ values of 15 for type I diabetes, 3.5 for type II diabetes and 8.6 for schizophrenia (Lander and Schork, 1994).

Linkage analysis is a technique, designed to search for co-occurrence of a marker locus and a putative disease gene, using all the available inheritance information from affected subjects and their families (Kruglyak et al., 1996). This technique involves

either parametric analysis, which proposes a specific model to explain the inheritance pattern of a trait-causing gene, or non-parametric analysis, which is model-independent and therefore tests the deviation of inheritance pattern from random Mendelian segregation under independent assortment (Kruglyak et al., 1996; Lander and Schork, 1994; Ott and Hoh, 2000).

#### *1.8.1.1 Parametric linkage analysis*

Traditionally, parametric (model-based) linkage analysis, commonly known as the logarithm of odds (LOD) score method (Morton, 1955) was performed on large, multiply affected, multigenerational pedigrees (Holmans and Craddock, 1997). The LOD score is the logarithmic expression of the likelihood (odds) ratio of the estimated recombination fraction (calculated under the alternative hypothesis of linkage, $\theta<0.5$) and the recombination fraction of 0.5 (calculated under the null hypothesis of no linkage) (McGuffin and Katz, 1989; Mitchell et al., 1993; Nyholt, 2002). Recombination fraction ($\theta$), refers to the number of observed recombination events divided by the total number of informative meioses (McGuffin and Katz, 1989; Ghosh and Collins, 1996). For small distances, the extent of $\theta$ is related to the physical distance between the marker locus and the disease locus. By testing across different values of $\theta$, the maximum LOD score (MLS) is obtained. Traditionally, a LOD score of >3, with odds of 1000 to 1 for linkage, was accepted as firm evidence to declare significant linkage (Morton, 1955; Nyholt, 2000). However, if one selects any two loci in the human genome, those loci are ~50 times (i.e., 46 chromosomes) more likely to reside on different chromosomes (Nyholt, 2000). Even with odds of 1000 to 1 in favour of linkage compared with non-linkage, the reality that non-linkages occur 50 times more often than linkages, it stands to reason that a LOD score of 3 corresponds to odds of ~20 to 1 in favour of linkage. That is, a LOD score of 3 will show to be false in ~1/20 events ($P$=0.5).

With few models for the disease locus, the LOD score method has been successfully applied to diseases with Mendelian patterns of inheritance (Lander and Schork, 1994). These models incorporated parameters such as the number of loci involved, allele frequencies, penetrances of each genotype, and recombination fractions (Olson et al., 1999; Ott and Hoh, 2000). Given good estimates for these parameters and therefore a correct disease model, the LOD score method takes full advantage of the available

pedigree data, thereby rendering it the most powerful linkage analysis technique (Guo and Lange, 2000; Ott and Hoh, 2000). If the assumed model is at least roughly accurate, then mapping of the disease locus becomes relatively straightforward (Guo and Lange, 2000). Using different phenotypic models, LOD score analysis has led to the identification of several genes implicated in BPD (see Table 1.2). Currently, in complex inheritance, a LOD score ≥3 is still necessary although not adequate enough to render a linkage test reliable (Morton, 1998). LOD score methods are sensitive to genetic heterogeneity and can lead to spurious results if the disease model is misspecified (Guo and Lange, 2000).

#### *1.8.1.2 Nonparametric linkage analysis*

In the instance of complex traits, it is not possible to model with certainty all the parameters of familial aggregation (Elston, 1998). Therefore, recent non-parametric (model-free) methods that do not require prior specification of such parameters have become increasingly popular in behavioural genetics (Kruglyak et al, 1996). Non-parametric methods have been used extensively in the identification of susceptibility genes in BPD (Table 1.2), and it is the method of choice in the present study. These include allele-sharing methods, association studies and the transmission disequilibrium test (TDT) (Lander and Schork, 1994; Ghosh and Collins, 1996).

**(i) Allele-sharing methods**, aim to determine whether affected relative pairs share the same chromosomal region identical-by-decent (IBD), i.e. inherited from a common ancestor, more often than expected by chance (Lander and Schork, 1994; Kruglyak and Lander, 1995). With the tendency to be more robust than linkage analysis (Kruglyak and Lander, 1995), allele-sharing methods are often less powerful, when compared to a correctly specified linkage model.

The most common and simplest form of allele-sharing methods use affected sib pairs (ASP), who in a randomly mating population can share 0, 1 or 2 alleles identical by decent (IBD), with corresponding sharing probabilities ($z$) of 0.25 ($z_0$) 0.5 ($z_1$) and 0.25 ($z_2$), respectively (Lander and Schork, 1994). If the marker under investigation is linked to a locus underlying a trait of interest, then affected sibs will tend to share more alleles IBD than expected (Ghosh and Collins, 1996). In some cases the founder alleles may not be informative, i.e. parents are not heterozygous with different genotypes.

Consequently, restriction of the ASP test to such families will undoubtedly result in a loss of information (Nyholt, 2000). In order to deal with data for which IBD status, in the presence of ambiguous marker information, cannot be unequivocally determined for all ASPs, likelihood methods were introduced.

The main approach of current ASP analysis is the likelihood-ratio (LR) or maximum likelihood score (MLS) method of Risch (1990b, 1990c), which tests the null hypothesis $H_0$: $(z_0, z_1, z_2) = (0.25, 0.50, 0.25)$,

$$\text{MLS} = \log_{10} \frac{L(\underline{z})}{L(0.25, 0.50, 0.25)}$$

where the maximum likelihood estimate (MLE) $\underline{z} = (z_0, z_1, z_2)$, the values of $(z_0, z_1, z_2)$ that maximise the likelihood (probability) of the data. For the unrestricted MLS method the only constraint on the MLE is that $z_0+z_1+z_2=1$. The null hypothesis is rejected if the resulting MLS is greater than the test criterion of 3 (with 2 degrees of freedom) which corresponds to a *P* value of 0.001. Not all sharing vectors ($z$) correspond to a plausible genetic model (Holmans, 1993). For example a sharing vector $(z_0, z_1, z_2)=(0.9, 0.05, 0.05)$ would imply that ASPs are not very likely to share disease alleles. This would seem doubtful, as the sibs are both affected, and therefore would have both inherited the same disease allele(s) from their parents. Risch (1990b) demonstrated that the sharing proportions ($z_0$, $z_1$ and $z_2$) for affected siblings can be expressed in terms of relative risks, given by the equations, $z_0 = 0.25/\lambda_s$, $z_1 = 0.5\lambda_o/\lambda_s$ and $z_2 = 0.25\lambda_m/\lambda_s$. The symbol $\lambda_s$ represents the relative risk for a sibling of a proband, where

$$\lambda_s = \frac{\text{Risk for sibs of a proband}}{\text{Population prevalence}}$$

and $\lambda_o$ and $\lambda_m$ are the relative risks for offspring and MZ twins of a proband, respectively. By stating the relative risks in terms of population additive and dominance variances it can be shown that $\lambda_o \leq \lambda_s$ and $\lambda_m \geq \lambda_s$ (Risch, 1990b). Thereby restricting the sharing vector $(z_0, z_1, z_2)$ to realistic genetic models, with the following restrictions imposed on the sharing probabilities:

$$z_0, z_1, z_2 \geq 1,\ z_0 + z_1 + z_2 = 1,\ z_2 + z_0 \geq z_1,\ z_1 \geq 2z_0$$

Holmans (1993) showed that plotting these restrictions graphically in the $(z_0, z_1)$ plane, defines a 'possible triangle', bounded by the lines $z_0 = 0$, $z_1 = 0.5$ and $z_1 = 2z_0$. Maximisation of the MLS is restricted to the set of sharing probabilities defined by $z_1 \leq 0.5$ and $2z_0 \leq z_1$ within the possible triangle, which is more powerful than the

unconstrained method. Restriction to the possible triangle follows a multiplicative model, which describes epistasis (interaction) among loci, has been shown to increase the power of the LR test and allows for dominance variance. Dominance variance occurs when alleles at a locus display dominance-recessive interactions, i.e. when the effect of one allele at a locus masks the effect of the other allele. For the possible triangle restricted test an MLS of 2.3 corresponds to a P value of 0.001, whereas for an unrestricted test an MLS of 3 corresponds to a P value of 0.001.

Other restrictions on the sharing probabilities are possible (Risch, 1990a). By restricting $z_1 = 0.5$ and $z_0 \leq 0.25$, the test follows an additive model and assumes no dominance variance. An additive model assumes that the genetic contribution to the phenotypic variance of a trait results from the additive effects of genes at unlinked loci. This model is a good approximation of genetic heterogeneity, allowing for no interaction between loci (Risch, 1990a), and an MLS of 2.07 corresponds to a P value of 0.001. Another possible restriction is the 'multiplicative model' restriction, which expresses ($z_0$, $z_1$, $z_2$) in terms of a single parameter $p$ such that $z_0=(1- p )^2$, $z_1 =2 p (1- p )$ and $z_2= p^2$ (MLS=2.07, P value of 0.001).

ASP methods make use of nuclear families containing affected sibling pairs, which makes it easier to collect large samples necessary to map susceptibility genes (Holmans and Craddock, 1997). Unlike large, densely affected pedigrees, affected sib pairs are more informative for linkage under oligogenic epistatic models (Risch and Merikangas, 1996). However, ASP methods also have some drawbacks: a relatively large proportion of the inheritance information contained in the pedigree structure is lost, making affected sib pair analysis inefficient (Kruglyak et al., 1996) and they do not produce estimates of recombination fractions (Holmans and Craddock, 1997).

MLS is very flexible and easily extendable to more complex situations. Recently, the MLS method has been extended to; (i) allow the consideration of the joint effects of two loci using ASPs (Cordell et al., 1995), and (ii) affected relative pairs (ARPs) (Cordell et al., 2000).

An alternative approach, the affected pedigree member (APM) method, has been developed to include information from all affected pedigree members (Weeks and Lange, 1988; 1992). This is not a true linkage method as it uses identical-by-state (IBS)

relations, i.e. affected relatives share a common allele at a locus regardless of whether it is inherited from a common ancestor in the pedigree. Excess IBS allele sharing among affected relative pairs and deviation from Mendelian expectation under the hypothesis of non-linkage, are investigated. The APM method, however, lacks a true multipoint formulation, i.e. it only adds together statistics from several marker loci (Weeks and Lange, 1992), and thereby tests for linkage to an extended chromosomal region rather than to a locus (Kruglyak et al., 1996). It therefore cannot be used to confine a specific locus relative to a map of markers.

Linkage mapping is the superior method for localisation of genes of moderate-to-large effect, although some genes of modest effect can also be uncovered (Concannon et al., 1998; Baron, 2001).

**(ii) Association studies**, unconcerned with familial inheritance patterns, are case-control studies examining the co-occurrence of a marker and disease at the population level (Lander and Schork, 1994; Baron, 2001). The gene is implicated in disease susceptibility if the frequency of occurrence of a marker allele is significantly different between unrelated patients and controls, both carefully matched for ethnicity and other factors (Parsian et al., 1991; Lander and Schork, 1994). Such disease-marker association can arise if the marker allele (like single nucleotide polymorphisms, SNPs) at the candidate locus directly influences the risk of disease or if it is in linkage disequilibrium with the actual cause. Association mapping, using a dense genome-wide collection of SNPs (see section 1.8.2.2 Linkage Disequilibrium/HapMap), has been heralded the method of choice in the study of the genetic basis of common complex disorders (e.g. Risch and Merinkangas, 1996). However, population substructure (heterogeneity) presents a serious problem, in that it can induce spurious associations between a candidate marker and a phenotype (Lander and Schork, 1994). The transmission disequilibrium test (family-based methods of association) was proposed in response to this problem (see next section on Linkage disequilibrium). Despite the success and popularity of family-based association tests, however, case-control studies are once again being extensively used with either direct analysis of population substructure (with correction if necessary) through the techniques developed by Pritchard and colleagues (2000), or the use of genomic control developed by Devlin, Roeder and Wasserman (2001). Each of these methodologies involves genotying rather large number of unlinked markers in the cases and controls, particularly for

admixed populations in which there may be slight variations in population substructure between case and control samples.

**(iii) Linkage disequilibrium (LD)** occurs when the marker allele, on the ancestral chromosome, lies close to the trait-causing locus in question, so that these are transmitted together through successive generations (MacKinnon *et al.*, 1997). Therefore LD is most inclined to occur in young, genetically-isolated populations (Lander and Schork, 1994). In finite populations LD can arise from random drift, selection, mutation and non-random mating (Hill and Weir, 1994). Moreover, the recombination fraction between the loci influences the extent of LD, such that LD decays with time in proportion to the recombination fraction (Weeks and Lathrop, 1995). In other words, LD between closely linked loci will decay slowly, whereas LD, as a result of admixture, selection or drift, between unlinked loci will decay very rapidly. Spurious association can occur as an artefact of population admixture (Lander and Schork, 1994), as a result of inappropriate matching of patients and controls (Weeks and Lathrop, 1995), or due to the choice of the candidate gene (Baron, 2001).

When population stratification is an issue, association studies depend on the transmission disequilibrium test (TDT) (Spielman et al., 1993), incorporating family-based control data, to overcome the problem. This test tracks the distorted transfer of disease susceptibility alleles in family triads, i.e. from parents to their affected child. The TDT relies on the assumption that a parent heterozygous for a disease associated allele should transmit that allele or its alternate at a greater frequency to the affected child compared to the unaffected child (Spielman et al., 1993; Lander and Schork, 1994; Sham 1996; Morton and Collins, 1998). Therefore the TDT is specific for association between linked loci and is preferred because of its robustness to population stratification, migration and admixture.

Martin and colleagues (2000) developed the pedigree disequilibrium test (PDT), for analysis of LD in general pedigrees. The PDT uses all potentially informative data, from related nuclear families and discordant sibships from extended pedigrees. Whilst retaining its validity in the presence of population substructure, PDT is more powerful than TDT even in case of misclassification of unaffected individuals.

### *1.8.1.3 Computational analysis*

To provide adequate background to the challenges of analysis of complex disorders, it is necessary, at this stage, to review the range of available computational methods, including those used in the present study.

Given the genotype data obtained from a genome-wide scan, test statistics may be calculated in a single-point (or two-point) or a multipoint manner. When calculating the test statistic in a single-point analysis each marker is analysed independently of all the other markers, whereas a multipoint analysis uses the information from multiple markers when testing for linkage at any particular marker or at any location between the markers. The multipoint approach makes full use of all the available genotyping information, which can increase the power.

A variety of computation programs have been developed for calculating often mathematically intensive linkage statistics, however two main algorithms underlie the calculations. The Elston-Stewart algorithm (Elston and Stewart, 1971) scales linearly with non-founders, individuals whose parents are in the pedigree, and exponentially with the number of loci used in the analysis. This algorithm is most commonly used in parametric linkage analysis, particularly in programs such as LINKAGE/FASTLINK (Lathrop et al., 1984; Cottingham et al., 1993) and VITESSE (O'Connell and Weeks, 1995). Depending on the pedigree size, LINKAGE/FASTLINK and VITESSE programs are best suited to analyse only a handful of loci (Ott and Hoh, 2000). Thus complete multipoint analyses remain a point of contention for general pedigrees (Kruglyak et al., 1996). Numerous families are frequently necessary to gain sufficient power to give significant evidence for linkage (Ellsworth and Manolio, 1999a).

An alternative to the Elston-Stewart algorithm is the Lander-Green algorithm (Lander and Green, 1987), which scales linearly with the number of loci and exponentially with non-founders. Most commonly utilised for non-parametric linkage analysis, the Lander-Green algorithm is equally well suited for parametric linkage analysis (Cox, 2001). Substantial improvements have been made to the algorithm, to extract all the available inheritance information from pedigrees (Kruglyak et al., 1996). The improved Lander-Green algorithm, incorporated into the GENEHUNTER computer program, when applied to parametric linkage analysis, provided rapid multipoint LOD score

calculations, including dozens of highly polymorphic markers. In non-parametric analysis, the same multipoint algorithm provided a new and powerful approach to pedigree analysis. Compared to parametric linkage analysis, loss of power in the non-parametric linkage (NPL) analysis was insignificant. More robust than the ASP and APM methods, the non-parametric linkage (NPL) approach is powerful in the face of an uncertain mode of inheritance. GENEHUNTER utilises two useful scoring functions, $S_{pairs}$ and $S_{all}$ (Whittemore and Halpern, 1994) to measure whether affected individuals share alleles IBD more often than expected under random segregation. $S_{pairs}$ refers to the number of pairs of alleles from distinct affected pedigree members that are IBD. When considering larger sets of affected relatives, such as three or more, the $S_{all}$ function puts extra weight on those individuals sharing the same allele IBD. In general, $S_i$ can be any function to include further information about sharing or lack of sharing, among affected and unaffected individuals. The normalised form of $S_i$ is defined as, $Z = [S - \mu]/\sigma$. Under the null hypothesis, i.e. under the uniform distribution over the possible inheritance vectors, Z has mean ($\mu$) 0 and variance ($\sigma$) 1. To analyse m pedigrees, the scores are combined in the following linear combination, $Z = \sum \gamma_i Z_i$, where $Z_i$ is the normalised score for the *i*th pedigree, and $\gamma_i$ are weighting factors.

Kruglyak et al. (1996), added the constraint $\sum \gamma_i^2 = 1$, so that Z has mean 0 and variance 1 under the null hypothesis of no linkage. Z is referred to as the NPL score, and $NPL_{pairs}$ and $NPL_{all}$ are used to denote the scoring function under investigation. With the concurrent use of multiple markers, in a region or along an entire chromosome, multipoint linkage analysis methods can infer IBD status at every point along the genome, as well as perform information content mapping and reconstruction of haplotypes. In addition GENEHUNTER allows the researcher to carry out LD analysis (by incorporating TDTs) and variance component analysis (for the detection of linkage to quantitative traits). However, GENEHUNTER is still somewhat encumbered by the size of the pedigree under investigation, i.e. only pedigrees of moderate size can be analysed (twice the number of nonfounders, 2N - the number of founders, $F \leq 20$, $\leq 16$ by default) (Kruglyak et al., 1996; Ott and Hoh, 2000). Furthermore, the *P* value (computed by the comparison of the observed NPL score with its complete data distribution) is usually conservative when information on decent is incomplete (Kong and Cox, 1997). This problem is circumvented by the likelihood approach of Kong and Cox (1997) by calculating the test statistic Zlr. Under the null hypothesis $\delta=0$, where $\delta$

is the parameter to be estimated representing the magnitude of deviation from null sharing, Zlr = $\sqrt{2}$ [log likelihood($\delta$) - log likelihood(0)].

When IBD sharing is ambiguous, this allele-sharing model still maximises an exact log likelihood. This allele-sharing model has been incorporated into the program ALLEGRO (Gudbjartsson et al., 2000). Much of the functionality of GENEHUNTER has been included in ALLEGRO, such as the calculation of multipoint parametric lod scores, NPL scores and allele-sharing lod scores, the reconstruction of haplotypes, estimated recombination count between markers and entropy information. ALLEGRO has several advantages over GENEHUNTER, such as the allele-sharing models that it provides, a much shorter execution time and there is no fixed limit on the number of marker loci or the number of non-founders. Unlike GENEHUNTER, that assigns equal weights to large and small families when tabulating the overall score, ALLEGRO allows the user to specify the weighting scheme. This property can also be used to execute conditional or two locus analyses in which the weight of a family relies on the linkage score of the family at a second locus (Cox et al., 1999).

Another program for linkage analysis at the postulated second locus is TWOLOCARP, which implements a likelihood-ratio analysis, based on the two-locus IBD sharing probabilities of affected relative pairs (Cordell et al., 1995; Farrall, 1997; Cordell et al., 2000). This likelihood ratio is a generalisation and extension of the maximum lod score (MLS) method (for a single locus) originally described by Risch (1990c) to overcome the problem of detecting the influences and modelling the action of many disease loci, using all affected relative pairs. TWOLOCARP generates various two-locus models, including a multiplicative (MUL), an additive (ADD) and a general (GEN) model. The MUL model, is an epistatic model (Risch, 1990a), where the two-locus MLS for two unlinked loci is equal to the sum of the individual single-locus MLSs at the two loci; the ADD model, is a good approximation of a heterogeneity (HET) model (Risch, 1990a), where the two loci are considered to be independent causes of the disease; both the MUT and ADD models are embedded in the GEN model, i.e. may be considered as special cases of the GEN model.

Other computer packages for multipoint analysis include, Multipoint Engine for Rapid Likelihood Inference, MERLIN (Abecasis et al., 2002) and SIMWALK2 (Sobel and Lange, 1996; Sobel et al., 2001; Sobel et al., 2002). The algorithms implemented in

MERLIN (Abecasis et al., 2002) use sparse inheritance trees to summarise the gene flow within pedigrees and efficiently calculate exact likelihoods for single markers or for multiple linked markers. MERLIN uses the Whittemore and Halpern (1994) $NPL_{pairs}$ and $NPL_{all}$ statistics to test for allele sharing among affected individuals, by generating a Z score, and its associated p value. The program also uses the Kong and Cox function (delta) to convert this Z score into LODs and calculate p values. MERLIN carries out common single-point and multipoint analyses, including IBD and kinship calculations of pedigree data, non-parametric and variance component linkage analyses, rapid haplotyping, genotype error detection, affected pair linkage analyses and can deal with more markers than other pedigree analysis programs. In addition, simulation studies can be carried out in MERLIN and many of the other software packages commonly used in linkage studies. By simulating data a more pragmatic estimation of the genome-wide significance of a finding is permitted, by taking into consideration the existing information content in the observed data (Wiltshire et al., 2001). The simulate option in MERLIN can generate random chromosomes that is unlinked to any of the traits of interest, the so-called gene-dropping simulations. The authors reported increases in computational speed for MERLIN relative to GENEHUNTER and ALLEGRO, software packages implementing related Lander-Green algorithms.

SIMWALK2 (Sobel and Lange, 1996; Sobel et al., 2001; Sobel et al., 2002), uses the Markov chain Monte Carlo (MCMC) algorithm which is able to analyse large pedigrees (>200 individuals) because it heeds the underlying configurations in proportion to their likelihood. Thus SIMWALK2 calculations are estimates and not exact. The program measures the degree of clustering, among affecteds, of the marker alleles descending from the founders, at each marker locus. The NPL_PAIR statistic (formerly known as STAT D) (Sobel and Lange, 1996, Whittemore and Halpern, 1994) is roughly the sum of conditional kinship coefficients for all affected relative pairs and is similar to the Whittemore and Halpern $S_{pairs}$ statistic as implemented in GENEHUNTER. Both these statistics were designed for traits best modelled by additive inheritance. Simwalk2 performs haplotype, parametric and non-parametric linkage, IBD and mistyping analyses. In instances where the pedigrees under investigation are of varying sizes (small and large), the utility Mega2, Manipulation Engine for Genetic Analysis (Mukhopadhyay et al., 1999) can be used to generate, from data stored in linkage-format files, all the input files required for a combined MERLIN (for small pedigrees) and SIMWALK2 (for large pedigrees) analysis.

There are a large number of computational resources for the linkage mapping of candidate loci, however, the outcome of these strategies are influenced by many factors.

### *1.8.1.4 Factors contributing to the outcome of analyses*

With the availability of various computational resources, the identification of candidate loci, by linkage methods, has seen few loci unambiguously classified. The reasons for this lack of progress are reviewed in this section.

#### 1.8.1.4.1 Misclassification of BPDs

It is not clear which phenotypic denotation best describes the underlying genetics of BPD (Segurado et al., 2003). It is plausible that there exist subtypes of BPD, owing to the presence of heterogeneity. The issue of phenotypic spectra is critical in molecular genetic studies, especially in linkage analysis of relatives that need to be classified as either affected or unaffected. It is therefore possible that diagnostic categories in the initial bipolar linkage studies may have been (i) too inclusive in some instances, thereby presuming that a particular phenotypic spectrum marked a unitary genetic disorder being inherited within families; or (ii) too exclusive of certain diagnoses, which may reduce statistical power (Pauls, 1993; Alda, 2004).

Co-morbidity, often seen with BPD, could lead to the misdiagnoses of these patients. This was illustrated, for example, where patients presenting with both BPD and alcoholism, were diagnosed as alcoholic, instead of BPD, and therefore led to the misclassification as "unaffected" in linkage analysis of BPD (Mitchell et al., 1993).

#### 1.8.1.4.2 Statistical power

Genetic analysis of complex traits requires a strong statistical power, to ensure credible evidence of association (Glatt and Freimer, 2002). The genetic mapping of complex traits involves searching for those chromosomal regions that tend to be shared among affected relatives. This is followed by the calculation of an appropriate linkage statistic, S, and identification of the region in which the deviation, from what one would expect to occur under independent assortment, is significant (Lander and Kruglyak, 1995). The

following guidelines, based on simulation studies, illustrate the Lander and Kruglyak (1995) statistical thresholds for significant linkage:

- a LOD score of 3.3 (pointwise P value of 0.000049) in parametric linkage analysis
- a pointwise probability ($P$ =0.000022) in sib-pair analysis
- a LOD score of 3.6 in non-parametric LOD score analysis
- a $P$ =0.00005-0.00001. in affected-pedigree-member analysis

When significant linkage has been obtained, its credibility has to be confirmed in an independent replication study ($P$<0.01). Modest p-values, and failure to replicate initial findings in recent candidate gene association studies and genome wide linkage scans, has lead to controversy within the field of complex trait analysis (Mein et al., 1998; Lernmark and Ott, 1998; Concannon et al., 1998).

The effects of locus heterogeneity across samples and across families within samples may contribute to several of the conflicting results (Lander and Kruglyak, 1995; Segurado et al., 2003). Prathikanti and McMahon (2001) concluded that ascertainment and methodological differences across linkage studies might contribute to the inconsistent findings. These differences can be seen in the use of multiplex extended pedigrees and parametric linkage analyses by some investigators, whilst others advocate the use of nuclear families with affected sib-pairs and non-parametric linkage approaches (Table 1.3). Further possible explanations for inconsistencies in linkage findings include population heterogeneity, which may lead to incidental sample variation across studies, and statistical outcomes such as inflated linkage results due to multiple comparisons (Lander and Kruglyak, 1995) and appropriate thresholds for significance in replication studies (Lernmark and Ott, 1998). Reduction in power may be due to stringent correction of multiple testing, as these procedures tend to be conservative and therefore impede reproducibility (Bohringer *et al.*, 2003). A requirement for increased power in the disease gene mapping and subsequent replication studies is an adequate sample size (Lander and Kruglyak, 1995; Lernmark and Ott, 1998). With a hypothesized decrease in genetic effect of a locus, there is a near exponential increase in sample size required for high-resolution mapping (Weeks and Lathrop, 1995).

In order to circumvent the statistical problems in human linkage analysis in pedigrees, computer simulation (Monte Carlo) methods can be used to assess the potential for linkage before actual marker typing of family members proceeds (Boehnke, 1986). Furthermore, Monte Carlo procedures can be used to predict genotypes or risks before

gathering the relevant information and to evaluate and compute the significance of information accumulated (Ott, 1989). It is expected that further refinement of the human genome map, in addition to multipoint analysis of closely linked marker loci, will magnify the power of linkage and association tests (Baron, 2001). A more efficient strategy perhaps, involves the combined (rather than separate) search for linkage and association between marker loci and disease.

## 1.8.2 Advances in molecular genetics

Despite the widespread recognition that complex disorders are oligogenic, researchers have been mapping genes one at a time, thereby vastly understating both the number of contributing susceptibility loci and the genetic and molecular complexities underlying complex disorders. Both the power and the resolution of linkage mapping may increase substantially, once genetic approaches include the gene-gene and gene-environment interactions that almost surely distinguish the genetic element of complex disorders (Cox et al., 1999; Cordell et al., 2000). The joint effects of two loci have been considered in insulin-dependent diabetes mellitus (Cordell et al., 2000), non-insulin dependent diabetes mellitus (Cox et al., 1999), obesity (Dong et al., 2005), and only recently have statistical analyses considered potential genomic interactions (epistasis) in BPD. With the application of conditional analyses, McInnis and co-workers (2003a) reported evidence of epistasis between chromosomes 16p13 and 9q21. In a different study, based on the National Institute of Mental Health Genetics Initiative pedigrees (USA), it was shown that loci on chromosomes 6q and 6p interact to increase susceptibility to BPD (Schulze et al., 2004). In addition to improved statistical analyses, i.e. to model potential gene-gene interactions, as in the present study, other promising strategies/approaches have come to the fore, to aid the identification of genetic components involved in complex disorders, and particularly BPD. These advanced strategies are summarised next, in sections 1.8.2.1 – 1.8.2.7, and (with the exception of bioinformatics) are not described further beyond this chapter.

### *1.8.2.1 Endophenotypes*

The incorporation of more phenotypic parameters into the analysis of behavioural and psychiatric traits, might improve the chances of identifying genetic loci involved in the disorder, by improving the statistical power of linkage and association studies (Inoue

and Lupski, 2003). These so-called endophenotypes are biological or psychological features and are generally quantitative in nature. Such quantitative traits could be influenced by fewer genetic and environmental factors than the disease trait itself and therefore possibly more agreeable to identification of susceptibility loci or disease-causing variants (Weeks and Lathrop, 1995). For example, Cloninger et al. (1998) reported an association between both temperament and character personality dimensions and mood disorders. Yoon and colleagues (2001) observed significantly reduced mRNA levels of inositol monophosphatase 1, IMPA1, with an associated increase in basal intracellular calcium concentrations in B lymphocyte cell lines from males with BPI and not in healthy male controls. Both, reduced IMPA 1 activity and elevated basal intracellular calcium concentrations may reflect cellular endophenotypes of bipolar disorder. A recent report has shown an association between genetic risk for BPD and decreased gray matter volume in the right anterior cingulate gyrus and ventral striatum (McDonald et al., 2004).

#### *1.8.2.2 Linkage disequilibrium/HapMap*

Single nucleotide polymorphisms (SNPs) have been acclaimed as the 'markers of choice' for disease mapping (Xiong and Jin, 1999). Highly abundant and accounting for 90% of sequence variation in humans, SNPs are stable genetic markers (Collins et al., 1998). A SNP database has been developed by the international SNP Consortium, in an effort to characterise at least 300 000 SNPs (http//:www.ncbi.nih.nlm.gov/SNP/). With the construction of high-density SNP maps (Wang et al., 1998) conventional candidate gene association studies can be substituted with genome-wide linkage disequilibrium (LD) mapping, using thousands of candidate-SNPs (Risch and Merinkangas, 1996). LD mapping of complex disease traits depends on detecting an association of SNPs, either individually or as haplotypes, with disease-causing mutations (Tishkoff and Verrelli, 2003). According to simulations by Kruglyak (1999), approximately 500 000 SNPs would be required to capture any disease associations in a full genome-wide scan. In addition to marker density, other parameters that affect the success of LD mapping include sample size, population subdivision, demography, and disease incidence (Clark, 2003). Identification of an association relies mainly on the underlying genomic patterns of LD in the population under investigation (Tishkoff and Verrelli, 2003).

More recently, the genome has been depicted as being made up of discrete high LD segments (haplotype blocks) interspersed with very low LD regions (Daly et al., 2001; Gabriel et al., 2002). These haplotype blocks, characterised by low haplotype diversity, may cover as little as one third of a chromosome (Clark, 2003). Johnson and colleagues (2001) determined the common haplotype blocks for nine genes in subjects of European origin and selected a set of 'tagging' SNPs within each block. These haplotype tags not only scanned the common variation of a gene sensitively and comprehensively, but also significantly reduced the number of SNPs from 122 to 34 SNPs (thereby reducing genotyping costs). The National institute of Health (NIH, USA) Haplotype map (HapMap) project is underway to catalogue common haplotype blocks and patterns of LD across the human genome, to assist the mapping of complex disease genes (Couzin, 2002).

#### *1.8.2.3 Pharmacogenomics*

One of the major issues with the management of patients with BPD is patient non-compliance with drug therapy, reported at rates of 51 to 64% (Keck et al., 1998). A further complication is that 30-40% of patients do not respond adequately to the early treatment therapy (Doris et al., 2001; Bradbury, 2003). Moreover, time required in proving the ineffectiveness of a therapy and, accordingly, the identification of non-responders could take up to six weeks. It would, therefore, be highly desirable to have a test that will distinguish the patient as a responder or a non-responder, prior to starting therapy and so eliminate the method of trial and error (Spear et al., 2001; Steimer at al., 2001). The principal, yet un-elucidated, explanation for erroneous drug response remains a patient's individual genetic predisposition (Spear et al., 2001). Hence the origin of the biomedical discipline of pharmacogenetics, which refers to the investigation of inherited variations in the metabolism and activity of xenobiotic agents and drugs (Catalano, 1999; Weber, 2001; Serretti et al., 2002; McCarthy, 2001). The term pharmacogenetics is often used interchangeably and confused with pharmacogenomics, where the latter deals with the impact of a patients' genetic constitution on the efficacy and toxicity of drugs (Peet and Bey, 2001; Schmitz et al., 2001). Pharmacogenetics and pharmacogenomics hold great promise for clinical medicine, where the aim of both is to determine the effect of genetic variation on drug response, with the ultimate goal of finding better and safer therapies for an individual patient, i.e. so-called personalized medicine.

SNPs and their analysis have been synonymous with the concept of pharmacogenomics (Meyerson, 2003). Over the past few years, several groups have reported possible susceptibility genes (see Serretti et al., 2002 for review). Certain studies found an association of the short (s) allele, in the promoter region of the serotonin transporter gene, with a poor response to selective serotonin re-uptake inhibitors (SSRIs). Serretti et al. (2001) reported an association between the A/A and A/C genotypes of the A218C variant in *TPH* and poor response to paroxetine. Variants in the *5-HT2a* receptor (Minov et al., 2001) and G-protein β3 (Zill et al., 2000) have been associated with a response to SSRIs. The near completion of the human genome project will provide a wealth of information on the nature and extent of genetic variation in human drug response and so transform the practice of medicine (Weber, 2001; Gurwitz et al., 2003).

Turecki and co-workers (2001), using a pharmacogenetic approach, mapped susceptibility genes for bipolar disorder. Their genome scan of 31 families ascertained through excellent lithium responders, found evidence for linkage with loci on chromosomes 15q14 and 7q11.2 and interesting findings were observed with markers on chromosomes 6 and 22. These findings highlight the potential of pharmacogenomics, in contributing to the elucidation of the genetic aetiology of BPD.

#### *1.8.2.4 Meta-Analysis*

The need for large sample sizes in the study of complex traits, has led investigators to see existing data in the literature in a new light (Johnson and Todd, 2000). Pooling of former linkage studies of BPD would increase the power of the combined data to detect linkage (Merinkangas and Yu, 2002). Meta-analysis represents one strategy for determining the significance of findings from multiple, independent and related studies (Segurado et al., 2003). In a genome scan meta-analysis (GSMA) of bipolar disorder, data from 18 BPD genome scans were combined, in an effort to identify regions with significant support for linkage (Segurado et al., 2003). Three such regions were identified on chromosomes 9p22.3-21.1, 10q11.21-22.1 and 14q24.1-32.12, using stringent disease models.

### *1.8.2.5 Variance component analysis*

Variance component-based linkage analysis, provide an attractive tool for the mapping of quantitative trait loci influencing complex phenotypes and for accurately estimating their relative effect sizes, using all the available information in extended pedigrees of varying complexity (Blangero et al., 2000). In addition to analysing large pedigrees, the variance component approach is able to analyse multiple loci simultaneously, examine gene-environment interaction, epistasis and calculate multipoint IBD probabilities. These properties have been implemented in the SOLAR (Sequential Oligogenic Linkage Analysis Routines) computer software package (Almasy and Blangero, 1998).

### *1.8.2.6 Expression studies*

In the hope of identifying individual genes and/or functionally related genes that contribute to complex disease, molecular profiling/expression studies, such as high-density array screens and differential display technologies, have started to impact on the understanding of BPD and schizophrenia.

Recently, genome-wide expression analysis, using microarrays, revealed decreased expression of myelin-related and oligodendrocyte-related genes in post-mortem dorsolateral PFC of schizophrenic patients (Hakak et al., 2001). In an independent study, using differential display and quantitative PCR, Tkackev et al. (2003) found that key oligodendrocyte and myelination genes were down-regulated in post-mortem brain samples of BPD and schizophrenia (including transcription factors that co-ordinate these genes), thus lending support for common pathophysiological pathways for these two disorders. Using DNA microarray analysis, Kakiuchi and colleagues (2003) showed down-regulation of expression of genes involved in the endoplasmic reticulum (ER) stress response in lymphoblastoid cells of bipolar patients. Upon further investigations, a polymorphism, -116C→G, in the promoter region of *XBP1* was significantly associated with bipolar disorder in Japanese case-control samples, causing impairment in the XBP1 positive-feedback system.

Bosetti and et al. (2002), using cDNA micro-arrays, examined brain expression of genes in rats fed with lithium. Expression of 50 genes were decreased at 42 days of oral lithium administration, coding for a number of receptors, protein kinases,

transcription and translation factors, markers of energy metabolism, signal transduction and stress response. The finding of lithium's widespread depressive effects proposes that bipolar disorder, in comparison, may be associated with a stimulatory effect on gene expression and brain functional activity. Microarray analysis of human neuroblastoma flat cells exposed to valproate (a mood-stabilizing drug) showed decreased expression of dopamine decarboxylase, dopamine beta-hydroxylase and dihydropyrimidinase-related protein 3 (Jurata et al., 2004).

cDNA array analysis of post-mortem frontal cortex tissue from patients with BPD and controls revealed gene expression differences between the two groups (Bezchlibnyk et al., 2002). This study identified genes encoding the signal transduction proteins, including transforming growth factor-beta 1 (*TGF-beta1*), caspase-8 precursor (*casp-8*) and transducer of erbB2 (*Tob*).

Using oligonucleotide GeneChip microarrays, Niculescu and colleagues (2000) examined gene expression in specific brain regions of rats, treated with methamphetamine. Several novel candidate genes were identified, including signal transduction molecules, transcription factors and metabolic enzymes that may be involved in the pathogenesis of mood disorders and psychosis. Moreover, decreased protein levels of GRK3, were observed in a subset of patient lymphoblastoid cell lines that related to disease severity. Recently, Iwamoto and co-workers (2004) performed oligonucleotide microarray analysis on post-mortem tissue from patients with bipolar disorder and control subjects. The genes encoding receptor and channels (or transporters) were down regulated and those that encode the stress response proteins were up regulated in bipolar disorder.

Yoon and colleagues (2001) examined the mRNA levels of TRPC7 in B lymphoblast cell lines (BLCLs) from BPI patients and healthy controls. Significantly lower TRPC7 mRNA levels were found in BPI patients with high BLCL intracellular $Ca^{2+}$ concentration, thus implicating reduced *TRPC7* gene expression in pathophysiological disturbance of $Ca^{2+}$ homeostasis in a subset of BPI patients.

Expression studies, using human and animal systems, augment efforts to characterise the structure and function of the genetic factors involved in complex disease processes.

### *1.8.2.7 Bioinformatics*

The successful completion of the Human Genome Project, HGP (Guttmacher and Collins, 2003), has led to the harnessing of genomics and proteomics in human disease studies. Simply put, genomics is the study of the interaction of genes, whereas proteomics is the study of all the proteins expressed in a cell or tissue. The HGP and its spin-offs, i.e. genomics and proteomics, are yielding unprecedented quantities of raw sequences, polymorphisms and gene expression data for thousands of genes, including those of the central nervous system in humans (Baron, 2002). Bioinformatics, the application of computer science to molecular biology, comparative genomics, structural genomics, and proteomics, has emerged as a fundamental tool to, (i) cope with the vast amount of genomic data stored in databases (Baron, 2002), and (ii) analyse information through data mining tools (Debes and Urrutia, 2004). Several different databases allowing universal access, with a sophisticated website as a primary means of presenting the data, are available and these include:

1. The Serial Analysis of Gene Expression (SAGE) database, at the National Centre for Biotechnology Information (NCBI), is a public resource that provides a quantitative profile of cellular gene expression. This quantitation does not measure the expression level of a gene, but measures a 'tag' which marks a gene transcript (Lash et al., 2000).
2. The Gene Ontology (GO) site is a database and informatics resource that provides controlled and structured vocabularies (ontologies) to describe properties of gene products, including molecular function, biological process and cellular component (Gene Ontology Consortium, 2004).
3. Databases such as Ensembl (http://www.emsembl.org), the University of California, Santa Cruz (UCSC) Genome browser (http://www.genome.cse.ucsc.edu) and the NCBI (http://www.ncbi.nlm.nih.gov/mapview) store and display annotated sets of genes, their chromosomal locations, and derived proteins from a range of genomes.

Further tools at Ensembl, including BlastView and MartView, provide the user with access to inter-species comparisons at both the genomic and protein sequence levels and a rich series of links to external databases, such as the Genetic Association database (http://geneticassociationdb.nih.gov/cgi-bin/index.cgi), UNIPROT knowledgebase (http://www.expasy.uniprot.org/) and PUBMED

(http://www.expasy.uniprot.org/) amongst others, thereby accommodating a wide range of possible queries (Hammond and Birney, 2004).

In order to gain maximum benefit from the wealth of information deposited in various databases, researchers require powerful computational tools to effectively integrate, clarify and apply it to drive hypothesis-based laboratory research and devise genetically derived benefits for clinical medicine.

## 1.9 THE PRESENT STUDY

Despite the current diagnostic techniques, diagnosis of BPD is complicated and misdiagnosis is commonplace. Nonetheless, BPD has long been recognised and continues to be well researched. Over the last decade intensive efforts have been made to explicate the genetic basis of BPD, in the hope of gleaning valuable insights into the structural and functional pathology, which remains to be fully explained. The genetic findings have, however, been unsatisfactory as investigators struggle with the problems of dealing with phenotypically and genetically heterogeneous conditions, controlling for gene-environment interactions and revealing the contributions of several interacting loci of small individual effect. Motivated by the current understanding of BPD, its impact on the global and South African burden of disease, the present study proposes, to some degree, to address the aforementioned problems, using an integrative approach, in contributing towards a better understanding of the specific neuropathology that underlies BPD. The practical benefits of identifying susceptibility genes for BPD include; (i) making diagnosis surer and faster, (ii) pinpointing biochemical pathways involved in pathogenesis, (iii) the facilitation of rational and more effective treatment strategies and, (iv) the implementation of clinical management strategies to prevent suicide and other morbidities in susceptible individuals.

### 1.9.1 Aims

The aim of this study is to identify susceptibility loci influencing BPD and attempted suicides (sub-phenotype), in South African families, integrating recent developments in statistical, computational and bioinformatics methods.

### 1.9.2 Project Objectives

The specific objectives of this study are to:

i) select candidate loci for genetic investigation

ii) perform linkage analysis, using parametric and non-parametric methods

iii) investigate the association between suicide attempts (sub-phenotype) and candidate loci (SNP)

iv) investigate gene-gene interactions, using statistical and bioinformatics methods

# CHAPTER 2

## THE SOUTH AFRICAN PERSPECTIVE

In the presence of no substantial molecular pathophysiological knowledge, linkage analysis is one of the finest available methods to identify chromosomal regions concealing genes that promote the development of BPD. Successful linkage studies of BPD will require very large samples of informative bipolar families drawn from diverse populations, and/or samples drawn from genetically isolated populations. Such a resource, if made readily accessible to the scientific research community, would not only permit researchers to test various linkage hypotheses (Blehar et al., 1988), but also allow findings of linkage in one family to be investigated speedily in other families. In due course, such an archival resource may favour the determination of (i) the proportion of BPD in which a specific chromosomal locus is involved, and (ii) the various loci contributing to genetic heterogeneity.

In 1988, the National Institute of Mental Health (NIMH, USA) initiated a nation-wide program to establish an archival resource for genetic studies in BPD (Blehar et al., 1988). Known as the NIMH genetics initiative for BPD, four sites were selected for participation, including Indiana University (John Nurnberger, PI), Johns Hopkins University (Raymond DePaulo, PI), the NIMH Intramural research program (Elliot Gershon) and Washington University at St. Louis (Theodore Reich, PI). By 1995, using a common ascertainment and assessment strategy, around 140 families with 1200 individuals had been ascertained and recruited. This national biological resource is complemented by diagnostic information and transformed cell lines from members of multiple families with BPD, providing an inventory or a living "library" of DNA and pedigree data. Based on their informativeness for linkage, 97 families with 540 subjects, enriched for affected relative pairs and parents (where available) were chosen for genotyping. This is one of the largest uniformly ascertained and assessed samples available for genetic studies of BPD. Over the past few years several other groups have recruited large series of families with BPD, for linkage studies (Table 2.1). Families, containing multiple affected individuals, typically had patients diagnosed with BPI and BPII. In some studies, the diagnostic criteria were extended to also include recurrent unipolar depression (RUP), major depression and schizoaffective disorder.

| Table 2.1: International archival resources for genetic studies of families with BPD | | |
|---|---|---|
| Reference | Resource | Diagnoses |
| NIMH Genetics Initiative Bipolar group, 1997 | 97 pedigrees (540 individuals) of North American origin | 232 BPI, 72 BPII, 88 RUP, 32 Schizoaffective bipolar illness |
| McMahon et al., 1997 | 30 pedigrees (259 individuals) of North American origin | 56 BPI, 38 BPII and MDE-R, 38 RUP, 4 Schizoaffective bipolar illness |
| Kelsoe, et al., 2001 | 20 pedigrees (164 individuals) of North American origin | 33 BPI, 15 BPII, 28 MDE-R |
| Cichon et al., 2001 | 75 pedigrees (445 individuals) of German, Israeli and Italian origin respectively | 128 BPI, 40 BPII, 14 32 Schizoaffective bipolar illness, 40 RUP |
| Bennett et al., 2002 | 151 nuclear families (509 individuals) of British and Irish origins | 288 BPI, 12 Schizoaffective bipolar illness, 25 BPII, 8 BNOS, 34 MDD |
| Ekholm et al., 2002 | 41 families (341 individuals) of Finnish origin | 92 BPI, 9 Schizoaffective bipolar illness, 6 BPII/BNOS, 15 MDD |
| Dick et al., 2003 | 250 pedigrees (1152 individuals) of North American origin | 537 BPI/ Schizoaffective bipolar illness, 51 BPII, 98 RUP |

In South Africa (SA), the Medical Research Council (MRC) conducted a survey in 2000, in order to obtain initial estimates of the burden of disease nation-wide (Bradshaw et al., 2003). In this survey the combined categories of mental and nervous system disorders, referred to as neuropsychiatric conditions, had the second highest disability-adjusted life years (DALY) ranking. These conditions could be the cause of mortality due to suicide (part of the injuries category). Furthermore, neuropsychiatric illness is ranked 19$^{th}$ out of the 24 disease categories contributing to premature mortality in SA, making it an essential category of healthcare problems. In 1996, the Division of Human Genetics, at the University of Cape Town (UCT), SA, initiated an investigation into the genetics of familial BPD with the establishment of a BPD archive.

## 2.1 MATERIALS AND METHODS

### 2.1.1 Ascertainment scheme

BPD families were recruited through the systematic screening of local psychiatric hospitals as well as through advocacy groups and advertisements. During the period of 1997 – 1999, three clinically trained interviewers undertook several visits to 4 of the 9 provinces in South Africa (Gauteng, Free State, Eastern Cape and Western Cape). The criteria for ascertainment required a proband and an available first-degree relative, both with a BPI or BPII diagnosis. Interaction with healthcare professionals and advocacy groups resulted in the recruitment of two or more family members with a diagnosis of BPI, BPII, or recurrent MDE. Through further genealogical assessment, additional affected family members were identified and included in the study. All subjects were entered into the research programme after signing the UCT Research Ethics Committee-approved (study approval UCT ERC 081/96) written informed consent forms. In addition, 20 ml of peripheral blood was obtained from each participant and processed to DNA in the molecular genetics laboratory in the Division of Human Genetics.

Individuals from the South African Mixed Ancestry and Caucasian population groups were recruited for the BPD resource. In SA, the Mixed ancestry population is derived from San, Khoi-Khoi, West Africa, Madagascar, Java and Western Europe origins. The South African Caucasian population is of Western European origin, mainly from European and British stock.

### 2.1.2 Diagnostic assessment

Probands and their relatives were interviewed by means of the Structured Clinical Interview for DSM-IV[1] Axis I Disorders (SCID)[2]. Borderline and anti-social PD were the only Axis II personality disorders assessed. Following the interview, a psychiatrist resolved the best-estimate final diagnosis for each subject using information from the SCID interview and all available family and medical records. Individuals who were experiencing an acute manic episode were interviewed once they were stable. No

[1] Diagnostic and Statistical Manual of Mental Disorders – 4th edition.
[2] Patient Edition (February 1996 FINAL.) SCID-I/P (Version 2.0) Modified for the Norvatis Bipolar Disorder Genetic Study Version: January 7.1998. First, M.B. Spitzer, R.L. Gibbon M and Williams, J.B.W.

children under the age of 16 were interviewed unless they had already been diagnosed with BPD by a psychiatrist.

## 2.2 RESULTS

To date, biological material and clinical information from 131 (two and three generation) families, which comprise of 809 individuals, have been archived. The diagnostic categories prevalent in the cohort of 809 individuals included mood, substance use and anxiety disorders (Table 2.2). Table 2.3 illustrates the distribution of these diagnostic categories according to ethnicity and gender; with a greater percentage of females than males recruited for both Caucasian and Mixed ancestry population groups.

Of the total cohort (809), the majority of individuals (n=472) were diagnosed with a mood disorder (Table 2.4). 51.1% of these affected individuals were diagnosed with BPI (age range of 17-87 years) and BPII (age range from 15-88 years), the core phenotypes of this study.

Table 2.2: Diagnostic categories of subjects in BPD archive

| Diagnostic category | Number of individuals |
|---|---|
| Mood disorders | 472 (58.3%) |
| Substance use disorders | 36 (4.5%) |
| Anxiety disorders | 27 (3.3%) |
| Psychotic disorders | 15 (1.9%) |
| Other disorders [1] | 16 (2%) |
| No diagnosis | 240 (29.6%) |
| Interviews not completed[2] | 3 (0.4%) |
| Total **(%)** | **809 (100%)** |

Other disorders, including Attention deficit disorder, Borderline personality disorder, Delusional disorder, Dementia, Migraine

Interviews not completed due to i) interviewee mental state, ii) refused prolonged participation in study iii) lack of collateral information to complete diagnostic process.

| Table 2.3: Ethnic and gender distribution of diagnostic categories | | | | |
|---|---|---|---|---|
| Diagnostic category | Mixed Ancestry | | Caucasian | |
| | M | F | M | F |
| Mood disorders | 69 | 141 | 111 | 151 |
| Substance use disorders | 19 | 2 | 13 | 2 |
| Anxiety disorders | 3 | 7 | 6 | 11 |
| Psychotic disorders | 6 | 5 | 1 | 3 |
| Other disorders | 2 | 3 | 6 | 5 |
| No diagnosis | 41 | 78 | 54 | 67 |
| Interviews not completed | - | - | 3 | - |
| TOTAL | 140 | 236 | 194 | 239 |

| Table 2.4: Summary of mood disorder diagnoses | |
|---|---|
| Diagnosis | Total |
| Bipolar I | 200 (42%) |
| Bipolar II | 43 (9.1%) |
| Cyclothymia | 4 (09%) |
| Bipolar NOS | 5 (1.1%) |
| MDE – recurrent | 113 (24%) |
| MDE – single episode | 92 (%) |
| Dysthymia | 8 (1.7%) |
| DNOS | 7 (1.5%) |
| TOTAL (%) | 472 (100%) |

Table 2.5 illustrates the distribution of these mood disorders according to ethnicity and gender. BPD showed a higher degree of co-occurrence with other Axis I psychiatric disorders (Table 2.6), including anxiety, substance use and psychotic disorders. In some cases, participants (8%) presented with more than two Axis I diagnoses. The Axis II, borderline and anti-social PDs were diagnosed in 1% of participants. Furthermore, with the exception of substance use disorder, women diagnosed with a mood disorder showed greater co-morbidity with all the different psychiatric diagnoses than men.

| Table 2.5: Ethnic and gender distribution of mood disorders | | | | |
|---|---|---|---|---|
| Diagnosis | Mixed ancestry | | Caucasian | |
| | M | F | M | F |
| Bipolar I | 42 | 65 | 48 | 45 |
| Bipolar II | 5 | 8 | 13 | 17 |
| Cyclothymia | 1 | - | 2 | 1 |
| Bipolar NOS | 1 | 2 | 2 | - |
| MDE – recurrent | 10 | 34 | 23 | 46 |
| MDE – single episode | 7 | 25 | 18 | 41 |
| Dysthymia | 2 | 3 | 2 | 1 |
| DNOS | 1 | 4 | 2 | - |
| **Total** | **69** | **141** | **111** | **151** |

| Table 2.6: Comorbid disorders (n=472) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Clinical phenotype | Diagnostic category * | | | | | | | TOTAL |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| Bipolar I | 132 | 19 | 23 | - | 5 | 5 | 15 | 199 |
| Bipolar II | 21 | 3 | 8 | - | 2 | - | 9 | 43 |
| Cyclothymia | 3 | - | 1 | - | - | - | - | 4 |
| Bipolar NOS | 4 | - | - | - | 1 | - | - | 5 |
| MDE – R | 74 | 15 | 9 | 1 | 3 | 1 | 10 | 113 |
| MDE – S | 67 | 10 | 8 | 1 | 1 | 1 | 3 | 91 |
| Dysthymia | 5 | 1 | 2 | - | - | - | - | 8 |
| Depression NOS | 5 | - | 1 | - | 1 | - | - | 7 |
| TOTAL (%) | 311 (66%) | 48 (10%) | 52 (11%) | 2 (0.4%) | 13 (3%) | 7 (1%) | 37 (8%) | 470 (100%) |

* Diagnostic category

1. Primary / Only diagnosis, 2. Co-morbid with Anxiety disorders, 3. Co-morbid with Substance Use disorders,

4. Co-morbid with Psychotic disorders, 5. Co-morbid with Other disorders including Attention deficit disorder, Borderline personality disorder, Delusional disorder, Dementia and Migraine, 6. Co-morbid with Axis II diagnosis specifically borderline + anti-social personality disorders, 7. Co-morbid with >2 Axis I diagnosis

The most prevalent mood disorders in the Mixed ancestry and Caucasian populations were BPD (BPI and BPII) and MDE (recurrent and single) (Table 2.7). More individuals of Mixed Ancestry descent were diagnosed with BPD (n=120), whereas more individuals of Caucasian descent were diagnosed with MDE (n=129).

Table 2.7: Diagnostic distribution of BPD and MDE according to ethnicity

| DIAGNOSIS | ETHNICITY | | |
|---|---|---|---|
| | **Mixed ancestry** | **Caucasian** | **TOTAL** |
| **BPI and II** | **120** | **123** | **243** |
| **MDE recurrent and single** | **76** | **129** | **205** |
| **TOTAL** | **196** | **259** | **448** |

A subset of 73 (9%) individuals, comprising of 48 families (1 - 8 members per family), had a history of suicide attempts (Table 2.8). Of these individuals, 53% were diagnosed as BPI, 13% as BPII, 23% as MDE recurrent and 9% as MDE single depressive episode. Suicide attempts peaked between the ages 31 - 50 years (56%) with the age-group 21 – 30 years having a higher attempt rate (19%) than the 51 – 60 age-groups (11%).

Table 2.8: Number of individuals who attempted suicide according to age groups and diagnostic category.

| AGE | BPI | BPII | MDE - R | MDE - S | Alcohol | TOTAL |
|---|---|---|---|---|---|---|
| 0 – 20 | 0 | 0 | 0 | 1 | 0 | 1 (1%) |
| 21 – 30 | 6 | 1 | 4 | 3 | 0 | 14 (19%) |
| 31 – 40 | 12 | 3 | 6 | 0 | 1 | 22 (30%) |
| 41 – 50 | 12 | 4 | 3 | 0 | 0 | 19 (26%) |
| 51 – 60 | 6 | 0 | 1 | 1 | 0 | 8 (11%) |
| 61 – 70 | 2 | 1 | 3 | 0 | 1 | 7 (10%) |
| 71 – 80 | 1 | 0 | 0 | 1 | 0 | 2 (3%) |
| | 39 (53%) | 9 (12%) | 17 (23%) | 6 (8%) | 2 (3%) | 73 |

## 2.3 DISCUSSION

In any attempt to identify genes underlying disease it is important to have a well-characterised large sample population for linkage analysis (or whatever mode of analysis will be used) and more specifically to aid the identification of multiple genes that exert varying magnitudes of effect on overall susceptibility. In order to locate such susceptibility genes that confer risk to BPD, a sizeable collection of families with subjects diagnosed as having BPD has been assembled in the Division of Human Genetics (UCT) in South Africa. Knowledge at hand indicates that this was the first neuropsychiatric genetics research project undertaken by a Human Genetics laboratory in South Africa. By the year 2000, around 131 families with 809 individuals had been ascertained, with diagnostic information captured in a well-maintained database and biological material archive in the molecular genetics laboratory. The biological resource consists of DNA material, but does not include transformed cell lines. It is possible that the existing resource of 809 relatives might be profitably utilised for parallel statistical and genetic analyses to investigate properties of diagnostic heterogeneity and familial patterns of inheritance, effects of birth cohort, and other such variables.

It is reasonable to keep families and population groups separate when attempting to identify genetic elements in heterogeneous disorders. The Indigenous Black SA population was not recruited because of the unavailability of appropriately translated diagnostic resources. Of the two population groups in this study, a great number was observed in the proportion of subjects affected with BPD (BPI and BPII) in the Mixed Ancestry population compared to the proportion of individuals affected with BPD in the Caucasian population ($P$ = 0.009). Furthermore, there was a lower proportion of subjects affected with MDE (recurrent and single) in the Mixed Ancestry population compared to the proportion of individuals affected with MDE in the Caucasian population ($P$ = 0.009). Therefore it is possible that different susceptibility alleles are present in these two population groups, and supports the separate analysis of the two population groups in subsequent genetic investigations.

Affected family members were diagnosed with BPI, BPII and MDE-R, the so-called BPD spectrum of phenotypes. The most common co-occurring illnesses among individuals

with BPD are anxiety and substance use disorders. More females than males were recruited, which could be attributed to one of two reasons, (1) more females were willing to participate in the study or (2) there is an association between females and bipolar disorder ($P = 0.001$).

At the time of the interview 9% of the recruited individuals admitted to having attempted suicide at least once in their lifetime. Eight years later, this figure is not an accurate representation of the incidence of attempted suicide in the study sample. Furthermore, the diagnoses of individuals could have changed. Therefore the implementation of systematic screening methods is required to update changes in psychiatric status (course of illness) on perhaps a yearly basis for all pedigrees.

In the design of the present study, three data sets (Table 2.9) of the archived families were selected, based on their informativity for genetic analysis and the diagnoses of the subjects (described in subsequent chapters). The research plan involved (ii) the identification of areas of interest, (ii) typing either flanking markers or markers within those areas and (iii) the modelling of two locus interactions.

Table 2.9: Data sets selected for Ph.D investigation

| Data set | Sample | Description of sample |
|---|---|---|
| BPD genetic linkage study | 22 Caucasian pedigrees | Chapter 3 |
| Genetics of Suicide attempts: case-control study | 46 cases, 160 controls | Chapter 4 |
| Genetics of Suicide attempts: linkage analysis of extended families | 9 extended pedigrees | Chapter 4 |

These data sets were chosen for genotyping a subset of chromosomal regions, containing several loci for which there have been reports of suggestive evidence for linkage to BPD and association with suicide attempts. In this report, analytic techniques including non-parametric (model-free) linkage and case-control analyses were performed and are reported separately.

# CHAPTER 3

## GENETICS OF BIPOLAR DISORDER IN SOUTH AFRICAN FAMILIES

The term affective disorders, encompasses several clinical conditions characterised by disturbance of mood (DSM-IV). Estimated to disrupt the lives of up to 15% of the population, affective disorders are perhaps the most destructive group of neuropsychiatric disease in terms of prevalence, mortality, economic cost, and impact on families. Of the affective disorders, the sub-group of BPD can be diagnosed with the most certainty and for which convincing evidence for familial disposition is available. Owing to huge medical, interpersonal, societal and economic costs (Evans and Charney, 2003) current BPD research is underway at an unprecedented rate. This surge of research is aimed at understanding the aetiology of the illness to inform treatment and prevention. In addition to extensive BPD research involving post-mortem (including expression studies), neuro-imaging and biochemical studies (as discussed in Chapter 1), there is great and increasing emphasis on genetics. Genetics is seen as a potentially powerful tool to identify shared genetic elements between individuals bearing similar phenotypic traits. Such shared material is likely to code for biological elements related to the pathophysiology of disease.

During the past 20 years, several genomic regions have been implicated as candidates by parametric and non-parametric linkage analyses (see Risch and Botstein, 1996). Regions on chromosomes 1, 3, 4, 5, 6, 12, 15, 18, 22 and X have all been implicated at one time or another. Occasionally a single gene may have a major effect in some families, however, generally the development of BPD is considered to involve multiple genes and environmental factors (Craddock and Jones, 1999). Due to this multifactorial nature, genetic analysis of BPD is complicated. The possible involvement of multiple genes, which are likely to be heterogeneous and of small individual effect sizes, together with their epistatic interactions add to the complexity of the disorder (Craddock and Jones, 1999). Novel statistical methods that permit the joint analysis of multiple chromosomal regions have successfully been applied to diabetes (Cox et al., 1999; Cordell et al., 2000) and BPD (McInnis et al., 2003a; Schulze et al., 2004). In addition to improved statistical genetic analyses, genomics has emerged as a promising

strategy/approach. Genomics not only focuses on a single gene/locus, but aims to interpret whole genomic information and its impact on biological systems (Guttmacher and Collins, 2002). The incredible advance in the science of genomics resulted in (i) an increase in the amount of data at the molecular and medical levels deposited in public databases; (ii) an increased need to develop tools to mine the data, organise it and create hypothesis-driven laboratory research (Debes and Urrutia, 2004). The advent of bioinformatics (as discussed in Chapter 1) provided researchers with essential tools to help make sense of biological data stored in databases. Moreover, bioinformatics represents a unique way to enhance our understanding of diseases, correlate genotypes to phenotypes, clarify multifactorial inheritance, and identify novel therapeutic targets.

The application of genetic mapping strategies and bioinformatics form the subject of the present investigation, i.e. (i) the application of linkage analysis strategies to candidate regions on chromosomes *1q32, 2q33.3, 4p16.1, 7q11.23, 10q23.31, 12q24, 13q32, 15q12, 16p13.3,* and *22q12.1*, that have been implicated in bipolar disorder (see Table 3.1), and (ii) the development and application of bioinformatics tools to identify and prioritise potential candidate genes in linked and other regions (Table 1.2).

Table 3.1: Candidate regions under investigation

| Candidate region | Reference |
|---|---|
| 1q31-32 | Turecki et al., 1995; Detera-Wadleigh et al., 1999 |
| 2q33.3 | Zubenko et al., 2002; Kelsoe et al., 2001 |
| 4p16.1 | Blackwood et al., 1994; Als et al., 2004 |
| 7q | Detera-Wadleigh et al., 1994; 1997 |
| 10q | Cichon et al., 2001b |
| 12q23-q24 | Dawson et al., 1995; Morisette et al., 1999 |
| 13q32 | Detera-Wadleigh et al., 1999; Liu et al., 2001 |
| 15q12 | Ginns et al., 1996; Craddock and Lendon, 1999 |
| 16p13 | Ewald et al., 1995; Edenberg et al., 1997 |
| 22q11-q13 | Lachman et al., 1997; Kelsoe et al., 2001 |

A brief summary and annotation of these candidate regions (Table 3.1) and the microsatellite markers used for linkage analysis is provided.

i. Chromosome 1q31-32

Interest in this region stems from a cytogenetic study reporting an association between BPI and a fragile site at 1q32 (Turecki et al., 1995). In a genome-wide scan of 22 US pedigrees, suggestive linkage of BPD to 1q31-q32 was reported (Detera-Wadleigh et al., 1999). Using affected sib pair analysis, a multipoint non-parametric lod score of 2.64 at D1S1660-D1S1678 (*P*=0.00022), under a narrow disease definition (including individuals diagnosed with BPI and BPII), was obtained. This finding was supported by parametric linkage analysis, with a lod score of 2.37 at marker GATA124F08, assuming a dominant mode of transmission.

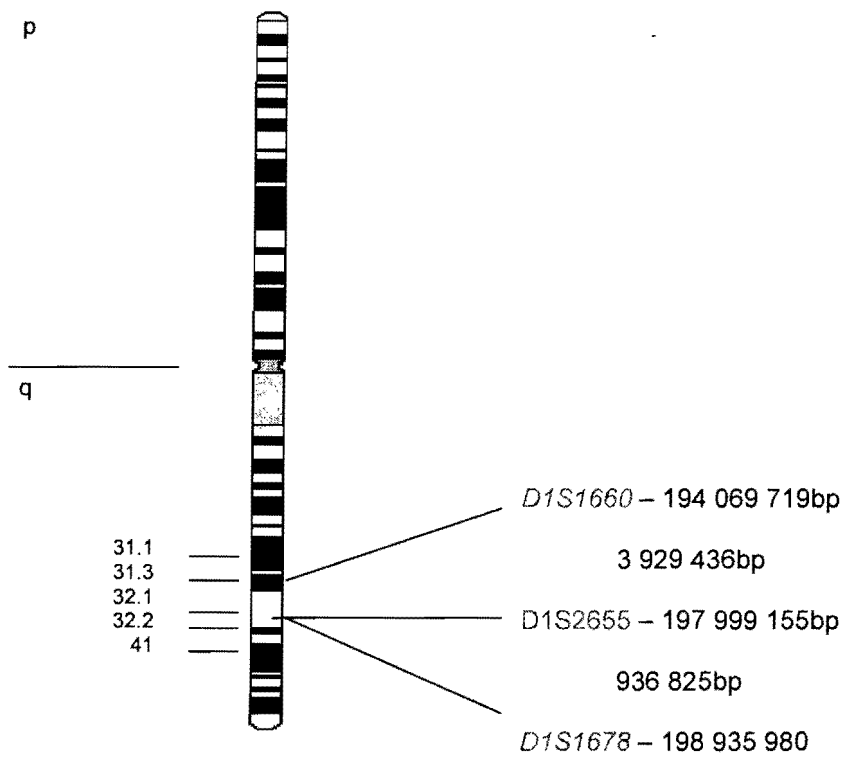


Figure 3.1: Ideogram of chromosome 1q31-q32 and the distances between markers. Key: *Italics* – Markers implicated in literature; Grey - Markers selected for linkage analyses, based on proximity and heterozygosity (>0.60). The orientation (i.e. distance in bp) is from telomere (pter) to centromere.

ii. Chromosome 2q33.3

Zubenko et al. (2002) reported linkage of unipolar mood disorders to a 451kb region of 2q33-34, with maximum non-parametric lod scores of 6.331 and 6.866 at D2S2321 and D2S2208, respectively. Linkage of mood disorders to this region, which contains the *CREB1* gene, was detected only among affected female relative pairs.

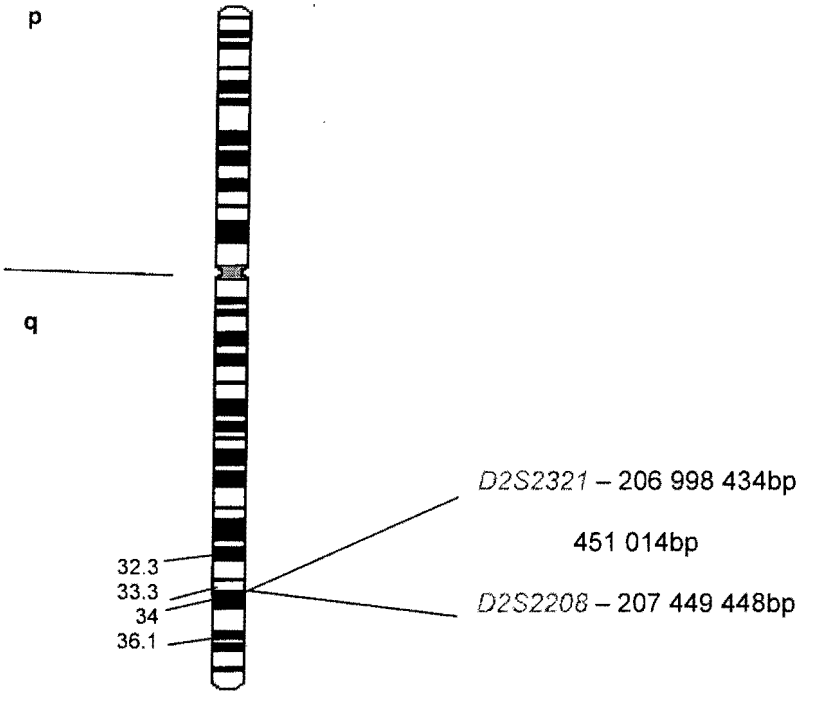


Figure 3.2: Ideogram of chromosome 2q33.3 and the distances between markers. Key: *Italics* – Markers implicated in literature; Grey - Markers selected for linkage analyses, based on proximity and heterozygosity (>0.60). The orientation (i.e. distance in bp) is from telomere (pter) to centromere.

iii. Chromosome 4p16.1

Under a dominant model of inheritance, Blackwood et al. (1994) observed linkage in a large Scottish pedigree, where the marker D4S394 generated a two-point parametric lod score of 4.1. Multipoint analyses with neighbouring markers gave a maximum parametric lod score of 4.8, under a narrow disease phenotype (including BPI, BPII and cases of recurrent and single episode unipolar depression) and a dominant model of transmission. Detera-Wadleigh et al. (1999) reported support for linkage to this region, with a multipoint non-parametric lod score of 1.77, under a broad disease phenotype (including individuals diagnosed with BPI, BPII and MDE). Recently, Als et al. (2004) reported evidence for association between BPD and a two-marker segment of D4S394-D4S2983 (*P*=0.0162) at 4p16.1.

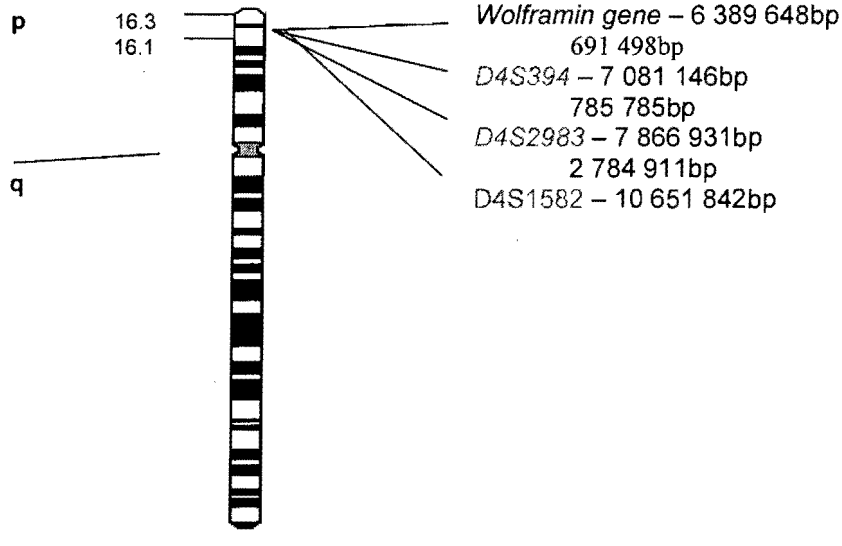


Figure 3.3: Ideogram of chromosome 4p16.1 and the distances between markers. Key: *Italics* – Markers implicated in literature; Grey - Markers selected for linkage analyses, based on proximity and heterozygosity (>0.60). The orientation (i.e. distance in bp) is from telomere (pter) to centromere.

iv. Chromosome 7q

In their analysis of 20 pedigrees, under a dominant model, Detera-Wadleigh et al. (1994) obtained a lod score >3 at D7S78. In the bipolar series of the National Institute of Mental Health (NIMH) Genetics Initiative, chromosome 7q displayed elevated allele sharing (Detera-Wadleigh et al., 1997). Coincident with this location, Detera-Wadleigh et al. (1999) reported evidence for linkage to an area between D7S1799 and D7S501 (lod score of 2.08, *P*=0.00099), under the broad disease phenotype, in 22 US pedigrees.

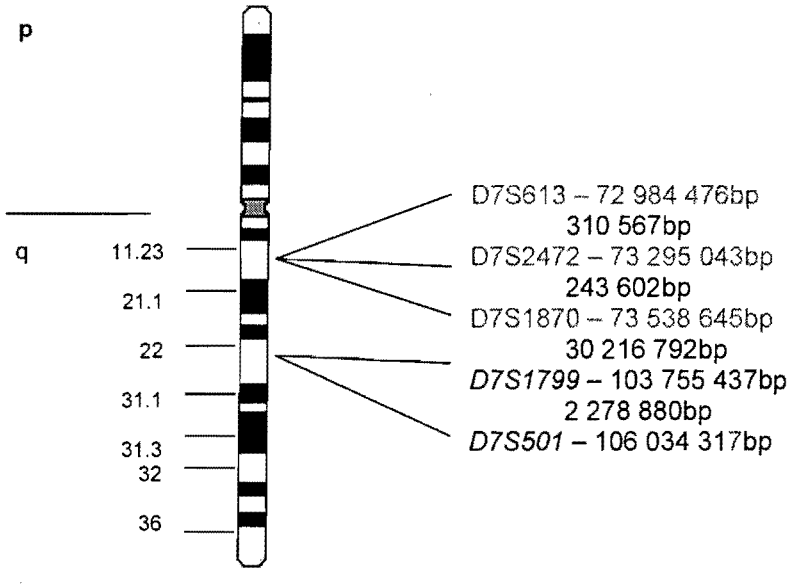


Figure 3.4: Ideogram of chromosome 7q and the distances between markers. Key: *Italics* – Markers implicated in literature; Grey - Markers selected for linkage analyses, based on proximity and heterozygosity (>0.60). The orientation (i.e. distance in bp) is from telomere (pter) to centromere.

v. Chromosome 10q

In their analysis of 75 German pedigrees, Cichon et al. (2001b) obtained a parametric lod score of 2.86 at D10S217 on 10q25-26, under a dominant model and a broad disease definition (including all cases with a diagnosis of BPI, BPII, schizoaffective, bipolar type, recurrent unipolar and unipolar single episode). Kelsoe et al. (2001) reported suggestive linkage of BPD to chromosome 10q with marker D10S1223 (parametric lod score 2.27), under a dominant model of inheritance.

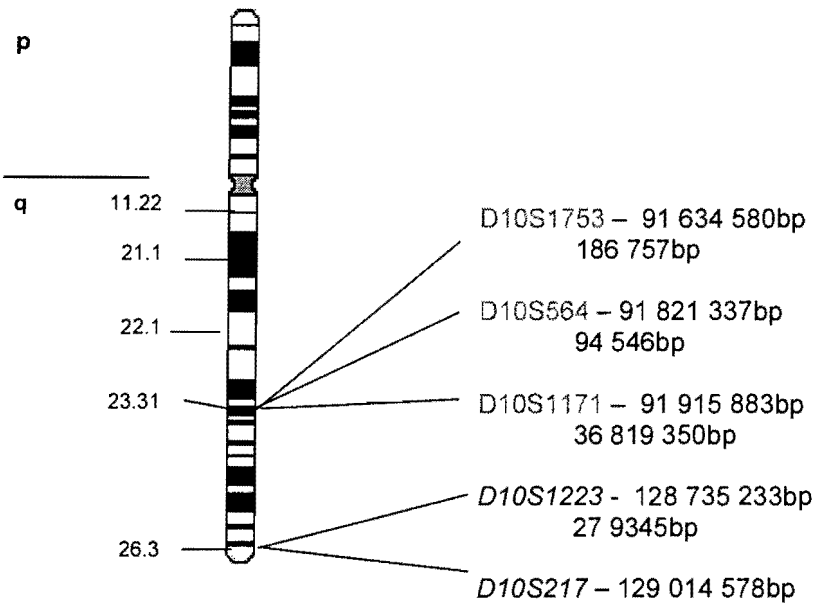


Figure 3.5: Ideogram of chromosome 10q and the distances between markers. Key: *Italics* – Markers implicated in literature; Grey - Markers selected for linkage analyses, based on proximity and heterozygosity (>0.60). The orientation (i.e. distance in bp) is from telomere (pter) to centromere.

vi. Chromosome 12q23-q24

Morisette et al. (1999) reported a non-parametric lod score of 3.92 at D12S82 in a large Canadian family, under a broad disease definition (where a diagnosis of BPI, BPII, schizoaffective, bipolar type, and recurrent major depression are considered as affected). Using non-parametric analyses Dawson et al. (1995) and Detera-Wadleigh et al. (1999) observed modest support for linkage to this region, 12q23-q24, in 44 European families and 22 US pedigrees, respectively. A recent genome-wide scan pointed to a susceptibility locus for BPD on 12q23-q24 with a parametric lod score of 3.35 at D12S378 (Shink et al., 2004).

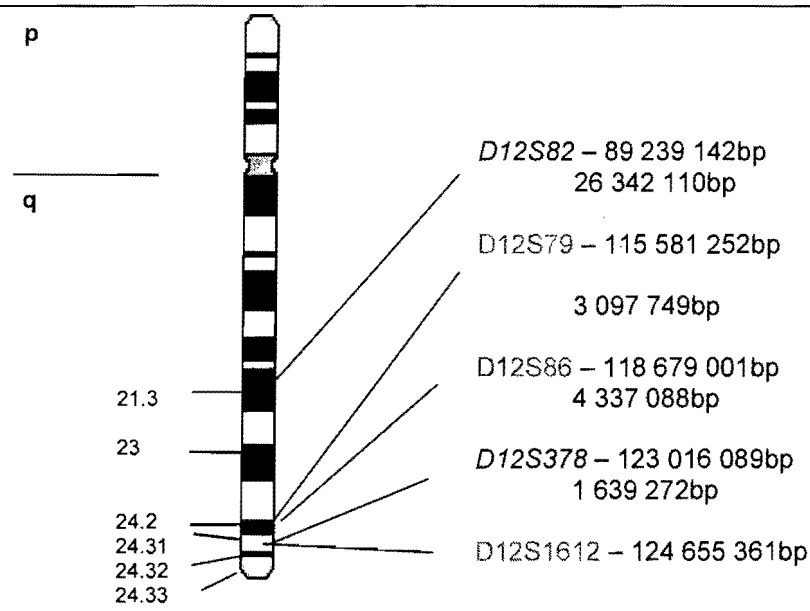


Figure 3.6: Ideogram of chromosome 12q23 and the distances between markers. Key: *Italics* – Markers implicated in literature; Grey - Markers selected for linkage analyses, based on proximity and heterozygosity (>0.60). The orientation (i.e. distance in bp) is from telomere (pter) to centromere.

vii. Chromosome 13q32

In their affected sib pair analysis of 22 US pedigrees, Detera-Wadleigh et al. (1999) obtained the strongest evidence for linkage at chromosome 13q32 that yielded a parametric lod score of 3.5 ($P$=0.000028) at the D13S1252-D13S1271 interval, under a broad disease definition (including individuals diagnosed with BPI, BPII and MDE). By typing nine additional markers in this region, in the same pedigree set, Liu et al. (2001) proved the validity of these findings and suggested a finer localisation of the susceptibility gene(s) on chromosome 13q32.

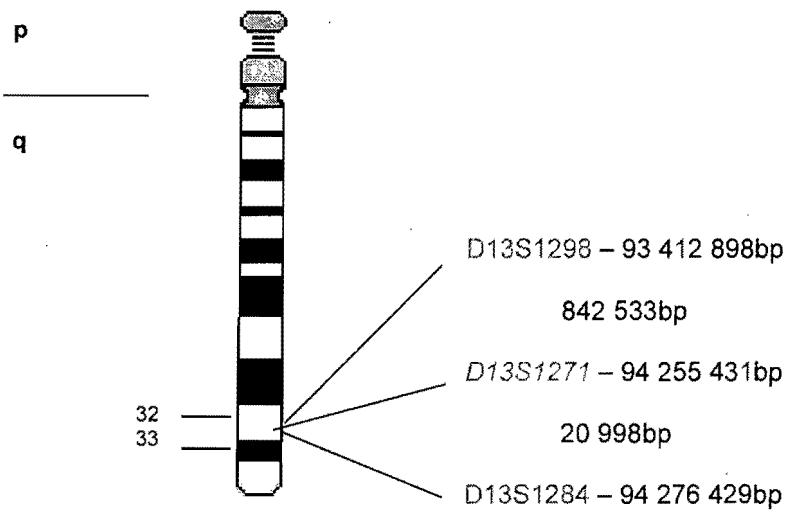


Figure 3.7: Ideogram of chromosome 13q32 and the distances between markers. Key: *Italics* – Markers implicated in literature; Grey - Markers selected for linkage analyses, based on proximity and heterozygosity (>0.60). The orientation (i.e. distance in bp) is from telomere (pter) to centromere.

viii. Chromosome 15q12

Ginns et al. (1996) reported linkage of BPD with excess allele sharing at D15S45 (*P*=0.0003), under the BPI diagnostic category. Detera-Wadleigh and co-workers (1998) found evidence of linkage to D15S695 (*P*=0.017) on 15q21.1. In the linkage data report on the Workshop on Chromosome 15 at the sixth World Congress of Psychiatric Genetics, in Bonn, Germany, the region on chromosome 15q13-q15 was most promising for functional psychoses (Craddock and Lendon, 1999). This region contains the alpha 7 nicotinic receptor and the gamma-aminobutyric acid receptor, alpha-5 (*GABRA5)* genes. Recently, chromosome 15q was identified in a genome-wide scan for regions influencing the age at onset of mania in BPD (Faraone et al., 2004).

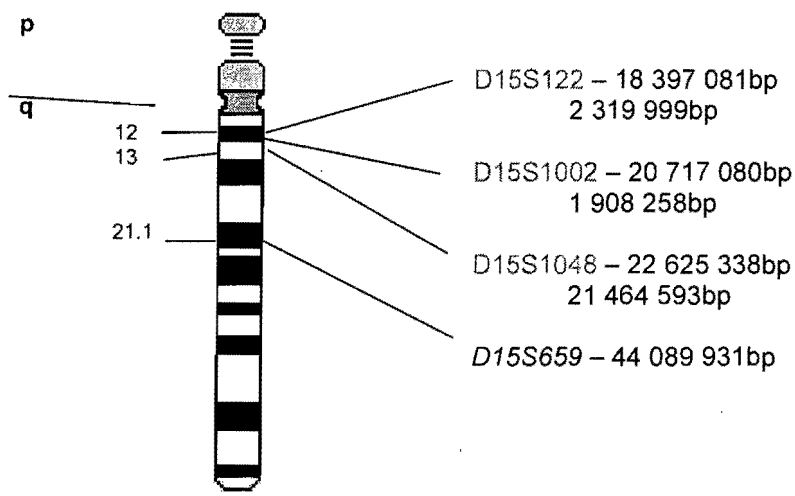


Figure 3.8: Ideogram of chromosome 15q12 and the distances between markers. Key: *Italics* – Markers implicated in literature; Grey - Markers selected for linkage analyses, based on proximity and heterozygosity (>0.60). The orientation (i.e. distance in bp) is from telomere (pter) to centromere.

ix. Chromosome 16p13

Ewald et al. (1995) reported possible evidence for linkage between BPD and D16S510 (parametric lod score 2.52), under a recessive model, in Danish families. In the 97 bipolar pedigrees of the NIMH Genetics Initiative, Edenberg et al. (1997) reported increased IBD sharing at D16S2619 (IBD=0.54, *P*=0.006) on chromosome 16p13.

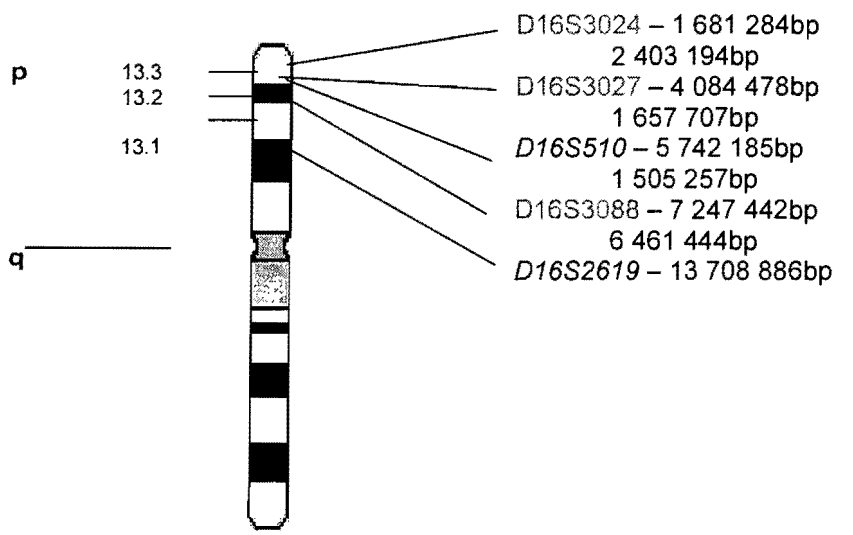


Figure 3.9: Ideogram of chromosome 16p13 and the distances between markers. Key: *Italics* – Markers implicated in literature; Grey - Markers selected for linkage analyses, based on proximity and heterozygosity (>0.60). The orientation (i.e. distance in bp) is from telomere (pter) to centromere.

x. Chromosome 22q11-13

In a study of 17 US BPD families, Lachman et al. (1997), using non-parametric multipoint analyses, reported suggestive evidence for linkage to chromosome 22q11 (*P*=0.017), near the Velo-Cardio-Facial syndrome region. This finding was supported by Detera-Wadleigh et al. (1999), who obtained a non-parametric multipoint lod score of 2.1 (P = 0.00094), under a broad disease definition (including individuals diagnosed with BPI, BPII and MDE) in 22 US pedigrees. The 'right' extension/branch of the Old Order Amish pedigree 110 was selected as the 22$^{nd}$ pedigree, because it seemed to contain a single known ancestor with bipolar disorder (Berrettini et al., 1994). Kelsoe et al. (2001) observed a parametric lod score of 3.84 in the same general region, under a dominant model, in 20 families from the general North American population.

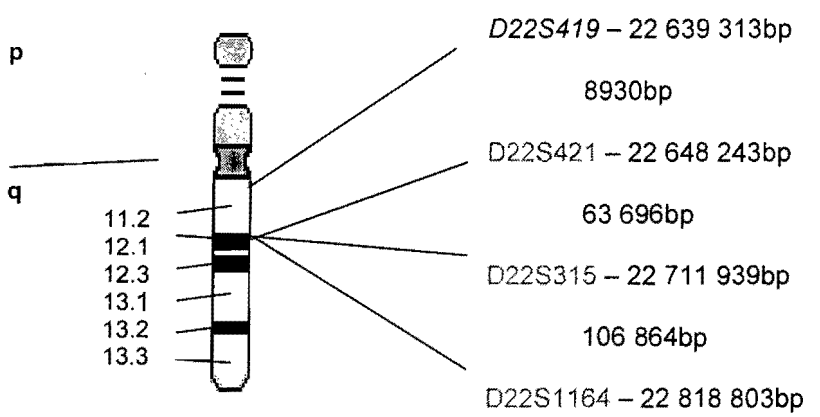


Figure 3.10: Ideogram of chromosome 22q11-q13 and the distances between markers. Key: *Italics* – Markers implicated in literature; Grey - Markers selected for linkage analyses, based on proximity and heterozygosity (>0.60). The orientation (i.e. distance in bp) is from telomere (pter) to centromere.

## 3.1 EXPERIMENTAL APPROACH

### 3.1.1 Subjects

From the large collection of families (n=131), 22 multiply affected pedigrees of Caucasian decent were selected as an initial candidate region screening panel. These families, with 4 to 33 members per family, typically had members diagnosed with BPI, BPII and MDE, with some unaffected relatives to allow inference of identity by descent (IBD) where possible (Table 3.2). A copy of the consent form completed by patients is included as Appendix A. The sample includes 168 individuals with; 31 BPI, 18 BPII, 24 single MDE and 46 recurrent MDE. In order to maximise the utility of the sample for genetic analyses and owing to the uncertainty about the spectrum of disease that is inherited, linkage analysis was carried out using two hierarchical models (Affection status models I and II) of the affected bipolar phenotype. For affection status model I (ASM I), a narrow disease definition was used to describe the affected phenotype, consisting of BPI and BPII. For the second model, ASM II, the affected phenotype was broadened to include those individuals with single and recurrent episodes of major depression.

### 3.1.2 DNA extraction

Peripheral blood (20ml) was collected from subjects into 10ml EDTA-containing plastic tubes. These tubes were labelled with the patient's name, surname and date of birth. Together with the completed consent form, the blood specimens were received by the molecular genetics laboratory of the Division of Human Genetics, UCT, where all the information was captured in a Microsoft Access database. Subsequent storage of blood specimens was at -20°C. The Genomix™ extraction kit (*Talent*, Italy) was used to extract genomic DNA (gDNA) from frozen lymphocytes (Appendix B, section B.1). gDNA samples were quantitatively assessed by spectrophotometry and working dilutions of 200ng/ul were made of the stock concentrations.

| Table 3.2: Summary of diagnoses of 22 Caucasian pedigrees | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| FAMILY | BPI | BPII | MDE - R | MDE - S | AA | Psychosis | Other | No diagnosis | TOTAL |
| 1 | 1 | 1 | 1 | - | - | - | - | 2 | 4 |
| 2 | - | - | - | 5 | - | - | - | - | 5 |
| 3 | 3 | - | 2 | 2 | - | - | - | 1 | 8 |
| 4 | 2 | 1 | 4 | - | 2 | 1 | - | 3 | 13 |
| 5 | 1 | 1 | 3 | - | - | - | 1 | 3 | 9 |
| 6 | 2 | 2 | 1 | 1 | - | - | - | 3 | 9 |
| 7 | 1 | 1 | - | 2 | 1 | - | - | - | 5 |
| 8 | 2 | - | 3 | 1 | 1 | - | 2 | 4 | 13 |
| 9 | 3 | - | 5 | - | - | 1 | - | 2 | 11 |
| 10 | - | - | 4 | 2 | - | - | - | 2 | 8 |
| 11 | 2 | 1 | - | - | - | - | - | 2 | 5 |
| 12 | 3 | 5 | 4 | 1 | - | - | 1 | - | 14 |
| 13 | 1 | 1 | 1 | 3 | - | - | - | - | 6 |
| 14 | 1 | 1 | 1 | - | - | - | - | 1 | 4 |
| 15 | 1 | 1 | 7 | 1 | - | - | - | 3 | 13 |
| 16 | 2 | - | 2 | - | - | - | - | - | 4 |
| 17 | 2 | 1 | 1 | 1 | 1 | - | - | - | 6 |
| 18 | 1 | 1 | - | 2 | - | - | - | 4 | 8 |
| 19 | 1 | 1 | - | 1 | - | - | - | 1 | 4 |
| 20 | 1 | - | 4 | - | - | - | - | 1 | 6 |
| 21 | 2 | 1 | - | - | - | - | 1 | 2 | 6 |
| 22 | 1 | - | 1 | 1 | 1 | - | - | - | 4 |

* Large pedigree was split into pedigrees 9 and 10 for linkage analysis

### 3.1.3 DNA analysis

The microsatellite markers used for genotyping (Table 3.3) consisted of tandem repeated dinucleotides and tetranucleotides. Markers were chosen from the literature and the following databases: the Foundation Jean Dausset - Centre d'Etude du Polymophisme Humain (CEPH) (www.cephb.fr), the National Centre for Biotechnology Information (NCBI) (www.ncbi.nlm.nih.gov/) and the Human Genome Browser (www.genome.ucsc.edu/). Marker heterozygosity ranged from 0.60 to 0.96 and averaged 0.85. The average inter-marker spacing across the candidate locus, in 262 affected relative pairs (ARP), was ~3cM, based on sex-average distances between them (Cooperative Human Linkage Center, CHLC). With few exceptions, multiplex PCR amplification was carried out in 10ul volumes in 96-well plates, containing

| Table 3.3: List of primer sequences for markers genotyped in candidate regions | | | |
|---|---|---|---|
| **Chromosome** | **Markers** | **Forward Primer (5'-3')** | **Reverse primer (5'-3')** |
| 1q32 | D1S1660 | tgctatcctctcaccagtga | gtctgaagttcatgggaacg |
| | D1S2655 | agggtccccaaagagccttc | atggcagcacatcctgcttc |
| | D1S1678 | actttctcacatgaccacagg | cagcgagactctgtcaaaaa |
| 2q33.3 | D2S2321 | ccatgtggtgtttggcag | gagcagaattgcagggc |
| | D2S2208 | ctatttgtaaacatgcggga | agctaagtacctgctcaggaaa |
| 4p16.1 | D4S394 | cccttgagcatcctgacttc | gagtgagcccctgtactcca |
| | D4S2983 | tgtccagttggcaggg | ggtcgcattcattcgc |
| | D4S1582 | atcagggttctccacacaaa | ttggttgaaacttgtggatataaa |
| 7q11.23 | D7S613 | cag cct ggg taa caa aag c | cctccctccctaatccatg |
| | D7S2472 | tct aaa gtc tgc cag gct ac | gcagcgagactccatc |
| | D7S1870 | ttc act cag gaa gtg gc | tggtgatgtgctttactacg |
| 10q23.31 | D10S1753 | ctgctgccaccaacctaa | caagtggagactcgatgaca |
| | D10S564 | tgggaatgtgtctttatcca | gctctaacatagaggccagat |
| | D10S1171 | gggttgaataaacatgcatg | gaatggggatgcagctaaa |
| 12q24 | D12S79 | ttggactgaactgagatgcc | tatgtgcacccagactacca |
| | D12S86 | agctagtctggcatgagcag | ctatcccctgatgatctccc |
| | D12S1612 | tcagcccctgtctcac | gcatccttgatgtccc |
| 13q32 | D13S1298 | gattgttttatattctctgcttttc | cctaccagtgggactttc |
| | D13S1271 | ccgagattgcaccactg | gcggtttttgccattaaag |
| | D13S1284 | gaggtgtccatcaatgc | tgtggctgcaaatgac |
| 15q12 | D15S122 | gataatcatgcccccca | cccagtatctggcacgtag |
| | D15S1002 | gtatcccaaggccataccct | ctcttgctagagacagcagg |
| | D15S1048 | agccgtctttgtgcca | tgcagccactgtggaa |
| 16p13.3 | D16S3024 | acatgctgtgccacct | agctgccagtatatggagga |
| | D16S3027 | atatttggcatctgggg | ccagcatgagttgcttt |
| | D16S3088 | ctctgaatagggtggggatg | aaggaaatctggggtgtacg |
| 22q12.1 | D22S421 | ctgctgcccctaacatatcac | ggccaggagtgtctgaatttta |
| | D22S315 | tgcctattaaactctccactcctta | gcattatgattcatttctcacaga |
| | D22S1164 | atccagacactgcctctc | cattcagaagccatttgc |

200ng gDNA, 10pmol of each fluorescent primer (labelled with either HEX or FAM, purchased from the Synthetic DNA laboratory, UCT, South Africa), reaction buffer containing 10mM Tris-HCl (pH8.3) and 50 mM KCl, 1.0 or 1.5 mM $MgCl_2$, 200uM dATP,dGTP,dCTP, dTTP (*Invitrogen*, UK) and 0.5 units of Biotaq™ (*Bioline*, UK) or Taq DNA polymerase (*Invitrogen*, USA).

All amplifications were carried out on a Perkin Elmer DNA Thermal Cycler. In general, cycling conditions included an initial denaturation step at 95°C for 3 minutes followed by 10 cycles of denaturation at 94°C for 15 seconds, annealing at 55°C (ranging from 60°C to 50°C) for 15 seconds, extension at 72°C for 30 seconds. The remaining 20 cycles included denaturation at 89°C for 15 seconds, annealing at 55°C (dependent on the primer pair in use) for 15 seconds, extension at 72°C for 30 seconds, followed by an extension step at 72°C for 5 minutes.

PCR products were genotyped using the Applied Biosystems (ABI) 377 and 3100 sequencers (Perkin Elmer) (Appendix B section B.3). Allele sizes were called using the Genescan® version 3.1/Genotyper® version 2.1 and Genemapper™ version 3.0 programs, on the ABI 377 and ABI 3100 sequencers respectively. In order to standardise markers run on different sequencing machines, allele sizes generated from two control DNA samples were used. Genotypic data was transferred to Genopedigree™ version 1.0.1(Perkin Elmer), along with the phenotypic data and stored in Genbase™ version 2.0.1. Mendelian inconsistencies were checked manually with Genotyper® version 2.1 software and resolved before linkage analysis. All data for statistical analysis were exported from Genbase™.

## 3.1.4 Statistical analysis

Because the pattern of inheritance of any genes influencing the relative risk for bipolar disorder is unknown (Craddock and McGuffin, 1993), model-free/non-parametric analyses were relied on to assess evidence for linkage due to a single locus and two locus interactions. Model-free methods, based on allele-sharing at marker locations, are not dependent on a specific model proposed for the disease as in the case of parametric LOD score methods. Affected relative pair (ARP) methods were used, as they are more robust than affected sib pair (ASP) methods, to conditions of genetic heterogeneity and multifactorial inheritance (Risch, 1990a-c).

### *3.1.4.1 Single (major)-locus analysis*

Linkage, i.e. the identification of chromosomal positions showing an excess of alleles shared IBD among the affected relative pairs, was assessed under the affected phenotype groupings of ASM I and ASM II. Linkage analyses were performed by using

MERLIN version 1.10.2 (Abecasis et al., 2002; MERLIN website) in conjunction with SimWalk2 version 2.89 (Sobel and Lange, 1996; Sobel et al., 2001; Sobel et al., 2002). To estimate more accurately the relative location of the BPD susceptibility gene and marker loci, multipoint linkage analyses (at three steps/increments equidistant between markers) were carried out, using sex-averaged maps, i.e. the test statistic was evaluated at positions between marker loci (as well as at the marker loci themselves), using allele-sharing information from the markers and their specified locations. The probability of the observed sharing is calculated by using the non-parametric linkage (NPL), Z, score approach (Kruglyak et al., 1996). Pedigrees were analysed using the ALL scoring function of Merlin, which examines all individuals simultaneously and assigns a higher score when more of them share the same allele by decent. Kruglyak and colleagues (1996) have demonstrated that the $NPL_{ALL}$ statistic results in a more powerful test than the $NPL_{PAIRS}$ statistic. However, the ALL scoring function can have skewed distribution in large pedigrees and therefore which statistic is more powerful depends on the true mode of transmission. MERLIN calculates exact non-parametric Z scores to assess the extent of allele-sharing (IBD) among the affected sibling and other relative pairs. One pedigree in the data set was too large to be analysed by MERLIN, and was evaluated by SimWalk2. This software performs Markov Chain Monte Carlo (MCMC) methods to estimate results for the combined sample of relative pairs and larger pedigrees, which could not be analysed by MERLIN. The NPL_ALL statistic of SimWalk2, was used and it is a measure of whether a few founder-alleles are represented in excess in the affecteds. MERLIN analysis was also done after splitting the large pedigree into smaller pedigrees (9 and 10), to compare the results. Mega2 (Manipulation Engine for Genetic Analysis) version 2.5 R3 (Mukhopadhyay et al., 1999) was used to generate from data stored in linkage-format files all the input files required for the MELRIN/SIMWALK2 analysis option.

#### *3.1.4.2 Multi-locus analysis*

Next the joint effects of two loci in BPD were considered. Again model-free methods are most convenient for combining this information into a test for linkage and/or interaction with a second disease locus, as they are not dependent on the specification of a two-locus disease susceptibility model. Two such two-locus analyses, restricted to (locations) which had previously shown evidence of linkage in single-locus analyses,

were carried out. In both analyses, the increased IBD sharing at any given position is considered as an estimation of disease location.

The first two-locus analysis method allows multipoint allele-sharing analysis at a postulated second locus, by weighting families according to their evidence for linkage at the first locus (Cox et al., 1999). Three weighting schemes were used to model positive interactions (such as epistasis) ($weight_{01}$ and $weight_{npl}$) and heterogeneity ($weight_{10}$). $Weight_{01}$ assigns weight 0 to families with 0 or negative linkage scores and weight 1 to families with positive linkage scores; $Weight_{npl}$ assigns weight 0 to families with 0 or negative linkage scores and families with positive linkage scores are assigned the specific linkage score as a weight; $Weight_{10}$ assigns weight 1 to families with 0 or negative linkage scores and weight 0 to families with positive linkage scores. $Weight_{npl}$ is similar to $weight_{10}$, except that more weight is assigned to families with positive linkage scores.

The program Allegro version 1.2c (Gudbjartsson et al., 2000) was used to carry out four allele-sharing model analyses incorporating the specific scoring function ($S_{pairs}$ or $S_{all}$) and family weighting scheme. Gudbjartsson and co-workers (2000) promote the use of the exponential allele-sharing model of Kong and Cox (1997) and cite McPeek (1999) in promoting the Whittemore and Halpern (1994) $S_{pairs}$ scoring function as a concession choice for a scoring function that behaves well over all disease models (see Allegro report which accompanies the Allegro manual). The following allele-sharing models were specified:

1. MODEL mpt exp pairs equal exppairseq.mpt ; this model causes multipoint ('mpt') allele-sharing analysis to be performed using the exponential ('exp') model and the $S_{pairs}$ scoring function with all families weighted equally ('equal'). Using the equal weights scheme, a standard unweighted linkage analysis was performed creating the output file exppairseq.mpt.
2. MODEL mpt exp pairs weights:$weight_{01}$ exppairs01.mpt ; this model causes multipoint ('mpt') allele-sharing analysis to be performed using the exponential ('exp') model and the $S_{pairs}$ scoring function with all families weighted according to scheme weight01.
3. MODEL mpt exp pairs weights:$weight_{10}$ exppairs10.mpt ; this model causes multipoint ('mpt') allele-sharing analysis to be performed using the exponential

('exp') model and the $S_{pairs}$ scoring function with all families weighted according to scheme weight10.

4. MODEL mpt exp pairs weights:$weight_{npl}$ exppairsnpl.mpt; this model causes multipoint ('mpt') allele-sharing analysis to be performed using the exponential ('exp') model and the $S_{pairs}$ scoring function with all families weighted according to scheme weightnpl.

Models 2 - 4 test linkage at the second locus, weighted by the evidence of linkage at the first locus. Zlr statistics are reported for each model at each marker and at three points equidistant between each marker, rather than NPL scores as these tend to be too conservative when inheritance data are incomplete. Under the null hypothesis of no linkage, Zlr has a standard normal distribution, with mean 1 and variance 0.

The difference in Zlr statistics between the equal weights (baseline) model and each of the three weighted models ($diff_{01}$, $diff_{10}$ and $diff_{npl}$) were calculated. The significance associated with these differences in Zlr, when evidence for linkage at a particular location is taken into account using family specific weights, was determined by writing a script to perform permutations in the statistical analysis package R. That is, to see how often performing a weighted analysis gave a result as significant as the result ($diff_{01}$, $diff_{10}$ and $diff_{npl}$) we observed by chance, 1000 permutations were performed using actual data and the R (R Development Core Team, 2004) script *rcoms* shown in Appendix D, section D.1 . In order to correct for using three weighting schemes, the maximum difference (maxdiff) of all three weighting schemes were also calculated.

In the second two-locus method a likelihood-ratio analysis, based on the two-locus IBD sharing probabilities of affected relative pairs (Cordell et al., 1995; Farrall, 1997; Cordell et al., 2000) is performed. This method is a generalisation and extension of the maximum lod score (MLS) method (for a single locus) introduced by Risch (1990c) to the problem of detecting the effects and modelling the action of several disease loci, using all affected relative pairs. The single-locus MLS test statistic, tests for departure from 0.25, 0.5 and 0.25 IBD sharing probabilities: MLS=log10 $L(\underline{z})/L(0.25, 0.5, 0.25)$, where $\underline{z} = (z_0, z_1, z_2)$, the values of $(z_0, z_1, z_2)$ that maximise the likelihood (probability) of the data. The program TWOLOCARP (Cordell et al., 1995; Farrall, 1997; Cordell et al., 2000) was used to calculate two-locus MLS statistics, which model the 3 by 3 matrix of 0, 1, and 2 IBD sharing probabilities at both locus 1 and locus 2, under the null hypothesis that locus 2 is not involved in disease. Prior and posterior probabilities, that

each affected relative pair shares *i* alleles IBD at particular locations on locus 1 and locus 2, were required. The Fortran utility *maketwo.f* (Appendix D section D.2) was used to convert the prior and posterior IBD output files generated by Allegro for each locus into input files, twoprior and twoposterior, required by TWOLOCARP. Various two-locus models are generated, including a multiplicative (MUL), an additive (ADD) and a general (GEN) model. The MUL model, is an epistatic model (Risch, 1990a), where the two-locus MLS for two unlinked loci is equal to the sum of the individual single-locus MLSs at the two loci; the ADD model, is a good approximation of a heterogeneity (HET) model (Risch, 1990a), where the two loci are considered to be independent causes of the disease; both the MUT and ADD models are embedded in the GEN model, i.e. may be considered as special cases of the GEN model. The significance associated with the increased MLS value, when evidence for linkage at a particular location is taken into account, was determined by simulation in MERLIN, using the R script *goodscript* (Appendix D, section D.3). Simulation was carried out by allowing MERLIN to perform gene-dropping, under the assumption of no linkage to disease, at the second (test) locus.

### 3.1.5 Bioinformatics

Literature on the molecular genetics of BPD was reviewed (Table 1.2) and used to compile a list of candidate regions implicated in the aetiology of the disease. Using the genetic markers reported in the literature, the boundaries of these candidate regions were refined using:

i) Genome Database (GDB, www.gdb.org) to find the corresponding D segment number for the reported flanking markers in cases where the marker names were reported in radiation hybrid format, in order to standardise nomenclature.

ii) National Centre Biotechnology Information (NCBI, www.ncbi.nih.gov), to find the nucleotide sequence of the identified flanking markers and their physical distances on the genome.

iii) University of California Santa Cruz (UCSC) BLAT server (http://genome.ucsc.edu) to identify and extract the genes reported in the corresponding intervals in the publications.

In cases where more than one marker was reported in the candidate region of interest (by different investigators), markers were ordered according to base pair positions

along the chromosome and the smallest and largest base pair positions were used to define the boundaries of the interval. In instances where only one marker was reported in the candidate region of the publications, an artificial boundary was established 5cM outward on either side of the reported marker. The resulting long list of genes, extracted from the candidate regions, was processed to remove hypothetical genes as well as repeated genes, henceforth referred to as master list (LIST_MASTER). The processing was executed on a Pentium IV personal computer running on a REDHAT LINUX V8 using a purpose written PERL script clean_tbl.pl.

Three additional short lists were created. The parameters and criteria used to create these additional lists are explained below:

i) LIST_MICROARRAY: This list comprises of genes implicated in Bipolar Disorder (BPD) based on reported micro-array expression studies (Table 3.4).

| Tabe 3.4: Sources of microarray expression studies in BPD |
|---|
| Literature |
| 1. Detera_Wadleigh, 2001<br>2. Niculescu and Kelsoe, 2001<br>3. Kakiuchi et al., 2003<br>4. Tkachev et al., 2003<br>5. Iwamoto et al., 2004 |

ii.) LIST_CNS: This list comprises of genes involved in different neurotransmitter pathways in the Central Nervous System (CNS). The data for this list was obtained from various sources (Table 3.5).

| Table 3.5: Sources of CNS data |
|---|
| Books |
| *Fundamental Neuroscience*, 1999, Academic Press<br>*Essential Psychopharmacology*, 2nd Edition, 2000, Cambridge University Press |
| World wide web |
| http://www.biocarta.com/pathfiles/neurotransmittersPathway.asp<br>http://www.biocarta.com/pathfiles/h_crebPathway.asp<br>http://www.genome.ad.jp/kegg/ |

iii) LIST_SAGE: The genes in this list were obtained from Serial Analysis of Gene Expression (SAGE) data, located at http://www.ncbi.nlm.nih.gov/sage and ftp://ncbi.nlm.nih.gov/pub/sage (Lash et al., 2000). Using the PERL script get_brain_exp_pl., genes were extracted from the human brain SAGE library.

A preliminary tool was designed, using the PERL script get_matches_pl., to match the genes in the individual lists to those in the master list, called LIST_MATCHES, in order to identify (short-list) candidate genes in the reported intervals, implicated in BPD. Next, the genes 'flagged' in the region(s) implicated in the linkage analysis, were investigated by Gene Ontology (Hammond and Birney, 2004) (www.geneontology.org), a resource for molecular function, biological process and cellular component ontologies, for a given gene(s), i.e. to find genes that are related by ontologies.

## 3.2 RESULTS

A total of 32 polymorphic microsatellite markers covering 11 candidate regions (Table 3.1), with an average spacing of ~3 cM between markers, were genotyped in 22 multiplex families. Two affection status models, narrow (ASM I) and broad (ASM II), were deliberately chosen prior to linkage analysis. Non-parametric linkage calculations identified candidate regions with excess IBD sharing.

### 3.2.1 Single-locus analysis

The results of the $NPL_{ALL}$ score analysis, before splitting the larger pedigrees, are depicted in Table 3.6. The table lists all markers and their respective empirical *P* values associated with the $NPL_{ALL}$ statistic of SimWalk2. Under ASM II, the strongest evidence for linkage was produced by chromosome 1q32, generating a *P* value of 0.0526 (NPL=1.2786) at marker D1S1678. After splitting the large pedigree into two smaller pedigrees, analysis in MERLIN, detected greater evidence for linkage to 1q32 under ASM II, yielding a non-parametric *P* value of 0.011 (Zmean=2.28) at marker D1S1678 (Table 3.7). Table 3.7 depicts all markers with the results of the multipoint MERLIN analysis, performed at three increments equidistant between marker loci. Under ASM I, chromosome 12q23-q24 yielded the next strongest support for linkage to BPD, generating a *P* value of 0.0327 (NPL=1.4850) at marker D12S1612. There was no

significant change in the *P* value before and after splitting the large pedigree. Furthermore, none of the other regions showed any evidence of linkage to BPD.

Table 3.6: Linkage data derived through MERLIN/SimWalk2 analysis before splitting the large pedigree

| | | ASM I | ASM II |
|---|---|---|---|
| Chromosomes | Markers | $NPL_{ALL}$ P value | $NPL_{ALL}$ P value |
| 1q32 | D1S1660 | 0.1112 | 0.01030 |
| | D1S2655 | 0.0839 | 0.0753 |
| | D1S1678 | 0.0821 | ***0.0526*** |
| 2q33.2 | D2S2321 | 0.7276 | 0.5937 |
| | D2S2208 | 0.8555 | 0.7693 |
| 4p16.1 | D4S394 | 0.6887 | 0.6566 |
| | D4S2983 | 0.6551 | 0.9759 |
| | D4S1582 | 0.3779 | 0.4494 |
| 7q11.23 | D7S613 | 0.1643 | 0.2831 |
| | D7S2472 | 0.3905 | 0.1992 |
| | D7S1870 | 0.6073 | 0.3704 |
| 10q23.31 | D10S1753 | 0.7216 | 0.3524 |
| | D10S564 | 0.8002 | 0.4348 |
| | D10S1171 | 0.4432 | 0.1154 |
| 12q24 | D12S79 | 0.3032 | 0.1370 |
| | D12S86 | 0.1127 | 0.0950 |
| | D12S1612 | ***0.0327*** | 0.0721 |
| 13q32 | D13S1298 | 0.2336 | 0.5740 |
| | D13S1271 | 0.1745 | 0.3693 |
| | D13S1284 | 0.1841 | 0.2340 |
| 15q12 | D15S122 | 0.3760 | 0.2014 |
| | D15S1002 | 0.3024 | 0.2368 |
| | D15S1048 | 0.2161 | 0.2191 |
| 16p13.3 | D16S3024 | 0.1248 | 0.4460 |
| | D16S3027 | 0.2172 | 0.2439 |
| | D16S3088 | 0.6519 | 0.4538 |
| 22q12.1 | D22S421 | 0.5751 | 0.6933 |
| | D22S315 | 0.5754 | 0.9749 |
| | D22S1164 | 0.2826 | 0.9161 |

Table 3.7: Multipoint linkage data derived through MERLIN analysis after splitting large pedigrees

| | | | ASM I | | ASM II | |
|---|---|---|---|---|---|---|
| Chromosome | Markers | Distance [a] | $P(NPL_{ALL})$ | $P(LOD)$ [b] | $P(NPL_{ALL})$ | $P(LOD)$ [b] |
| 1 | D1S1660 | 0 | 0.08 | 0.04 | 0.04 | 0.03 |
| | | 1.095 | 0.08 | 0.03 | 0.04 | 0.03 |
| | | 2.19 | 0.08 | 0.03 | 0.04 | 0.02 |
| | | 3.285 | 0.08 | 0.04 | 0.03 | 0.02 |
| | D1S2655 | 4.38 | 0.08 | 0.05 | 0.03 | 0.02 |
| | | 4.79 | 0.07 | 0.04 | 0.02 | 0.014 |
| | | 5.2 | 0.07 | 0.04 | 0.02 | 0.011 |
| | | 5.61 | 0.06 | 0.04 | 0.02 | 0.008 |
| | D1S1678 | 6.02 | 0.06 | 0.05 | **0.011** | ***0.006*** |
| 2 | D2S2321 | 0 | 0.8 | 0.7 | 0.7 | 0.7 |
| | | 0.003 | 0.8 | 0.7 | 0.7 | 0.7 |
| | | 0.005 | 0.8 | 0.7 | 0.7 | 0.7 |
| | | 0.008 | 0.8 | 0.7 | 0.8 | 0.7 |
| | D2S2208 | 0.01 | 0.8 | 0.7 | 0.8 | 0.7 |
| 4 | D4S394 | 0 | 0.7 | 0.7 | 0.7 | 0.6 |
| | | 0.37 | 0.7 | 0.7 | 0.8 | 0.7 |
| | | 0.74 | 0.7 | 0.7 | 0.9 | 0.7 |
| | | 1.11 | 0.7 | 0.7 | 0.9 | 0.8 |
| | D4S2983 | 1.48 | 0.7 | 0.6 | 1.0 | 0.8 |
| | | 3.047 | 0.6 | 0.6 | 0.9 | 0.7 |
| | | 4.615 | 0.5 | 0.5 | 0.8 | 0.7 |
| | | 6.183 | 0.5 | 0.5 | 0.7 | 0.7 |
| | D4S1582 | 7.75 | 0.4 | 0.4 | 0.5 | 0.6 |
| 7 | D7S613 | 0 | 0.4 | 0.4 | 0.3 | 0.4 |
| | | 0.333 | 0.4 | 0.4 | 0.3 | 0.3 |
| | | 0.665 | 0.5 | 0.5 | 0.2 | 0.2 |
| | | 0.998 | 0.5 | 0.5 | 0.13 | 0.2 |
| | D7S2472 | 1.330 | 0.5 | 0.5 | 0.08 | 0.12 |
| | | 1.530 | 0.5 | 0.5 | 0.11 | 0.14 |
| | | 1.730 | 0.6 | 0.6 | 0.14 | 0.2 |
| | | 1.930 | 0.6 | 0.6 | 0.2 | 0.2 |
| | D7S1870 | 2.130 | 0.7 | 0.7 | 0.2 | 0.3 |

Table 3.7 : *Continued*

| | | | ASM I | | ASM II | |
|---|---|---|---|---|---|---|
| Chromosomes | Markers | Distance [a] | $P(NPL_{ALL})$ | $P(LOD)^b$ | $P(NPL_{ALL})$ | $P(LOD)^b$ |
| 10 | D10S1753 | 0 | 0.7 | 0.7 | 0.4 | 0.3 |
| | | 0.003 | 0.7 | 0.7 | 0.4 | 0.4 |
| | | 0.005 | 0.7 | 0.7 | 0.4 | 0.4 |
| | | 0.008 | 0.8 | 0.7 | 0.4 | 0.4 |
| | D10S564 | 0.01 | 0.8 | 0.7 | 0.4 | 0.4 |
| | | 1.218 | 0.8 | 0.7 | 0.3 | 0.3 |
| | | 2.425 | 0.7 | 0.7 | 0.3 | 0.2 |
| | | 3.632 | 0.6 | 0.6 | 0.2 | 0.13 |
| | D10S1171 | 4.84 | 0.5 | 0.5 | 0.12 | 0.1 |
| 12 | D12S79 | 0 | 0.4 | 0.4 | 0.13 | 0.13 |
| | | 2.308 | 0.3 | 0.3 | 0.13 | 0.11 |
| | | 4.615 | 0.3 | 0.3 | 0.12 | 0.09 |
| | | 6.923 | 0.2 | 0.2 | 0.11 | 0.09 |
| | D12S86 | 9.23 | 0.2 | 0.2 | 0.10 | 0.09 |
| | | 10.638 | 0.11 | 0.11 | 0.10 | 0.08 |
| | | 12.045 | 0.07 | 0.06 | 0.10 | 0.08 |
| | | 13.453 | 0.04 | 0.04 | 0.09 | 0.09 |
| | D12S1612 | 14.86 | **0.03** | **0.03** | 0.09 | 0.1 |
| 13 | D13S1298 | 0 | 0.3 | 0.3 | 0.5 | 0.5 |
| | | 0.003 | 0.3 | 0.3 | 0.5 | 0.5 |
| | | 0.005 | 0.3 | 0.3 | 0.4 | 0.4 |
| | | 0.008 | 0.2 | 0.3 | 0.4 | 0.4 |
| | D13S1271 | 0.01 | 0.2 | 0.3 | 0.3 | 0.3 |
| | | 0.013 | 0.2 | 0.3 | 0.3 | 0.3 |
| | | 0.015 | 0.2 | 0.3 | 0.3 | 0.2 |
| | | 0.018 | 0.2 | 0.3 | 0.3 | 0.2 |
| | D13S1284 | 0.02 | 0.2 | 0.3 | 0.2 | 0.2 |
| 15 | D15S122 | 0 | 0.4 | 0.4 | 0.2 | 0.14 |
| | | 2.118 | 0.4 | 0.4 | 0.2 | 0.11 |
| | | 4.235 | 0.4 | 0.3 | 0.2 | 0.1 |
| | | 6.353 | 0.3 | 0.3 | 0.2 | 0.12 |
| | D15S1002 | 8.47 | 0.3 | 0.3 | 0.2 | 0.2 |
| | | 9.605 | 0.2 | 0.2 | 0.2 | 0.2 |
| | | 10.74 | 0.2 | 0.2 | 0.2 | 0.2 |
| | | 11.875 | 0.2 | 0.2 | 0.3 | 0.2 |
| | D15S1048 | 13.01 | 0.2 | 0.2 | 0.3 | 0.3 |

Table 3.7 : *Continued*

| | | | ASM I | | ASM II | |
|---|---|---|---|---|---|---|
| Chromosomes | Markers | Distance [a] | $P(NPL_{ALL})$ | $P(LOD)$ [b] | $P(NPL_{ALL})$ | $P(LOD)$ [b] |
| 16 | D16S3024 | 0 | 0.07 | 0.02 | 0.4 | 0.4 |
| | | 0.415 | 0.07 | 0.013 | 0.4 | 0.4 |
| | | 0.83 | 0.08 | 0.02 | 0.3 | 0.3 |
| | | 1.245 | 0.11 | 0.03 | 0.3 | 0.3 |
| | D16S3027 | 1.66 | 0.2 | 0.09 | 0.3 | 0.3 |
| | | 2.763 | 0.2 | 0.14 | 0.4 | 0.3 |
| | | 3.865 | 0.3 | 0.3 | 0.4 | 0.4 |
| | | 4.968 | 0.5 | 0.5 | 0.5 | 0.5 |
| | D16S3088 | 6.07 | 0.7 | 0.6 | 0.6 | 0.6 |
| 22 | D22S421 | 0 | 0.7 | 0.7 | 0.7 | 0.6 |
| | | 0.003 | 0.7 | 0.7 | 0.8 | 0.7 |
| | | 0.005 | 0.7 | 0.7 | 0.9 | 0.7 |
| | | 0.008 | 0.7 | 0.7 | 0.9 | 0.8 |
| | D22S315 | 0.01 | 0.7 | 0.7 | 1 | 0.8 |
| | | 0.013 | 0.6 | 0.6 | 1 | 0.8 |
| | | 0.015 | 0.5 | 0.5 | 1 | 0.8 |
| | | 0.018 | 0.4 | 0.4 | 0.9 | 0.8 |
| | D22S1164 | 0.02 | 0.3 | 0.3 | 0.9 | 0.8 |

[a] Distance in cM, calculated from our data

[b] *P* value for the Kong and Cox allele-sharing lod score (Zlr statistic)

### 3.2.2 Multi-locus analysis

Evidence for linkage was assessed at a second region, 2q33.3, 4p16.1, 7q11.23, 10q23.31, 12q24, 13q32, 15q12, 16p13.3 or 22q12.1, conditioned on the evidence for linkage at 1q32, under ASM II, using 2 two-locus methods. The first method involves weighting a family's contribution to the Zlr statistic at locus 1q32, through the program Allegro, using three weighting schemes (Cox et al., 1999), $weight_{01}$, $weight_{10}$ and $weight_{npl}$. Table 3.8, illustrates the baseline Zlr values observed at the second locus and those weighted by the contribution of 1q32, under ASM II. The most significant difference between baseline and 1q32 weighted Zlr statistics, was observed at region 10, with a difference of 1.0147 ($weight_{npl}$ minus baseline).

| Table 3.8: Zlr statistics of Allegro analysis of two locus interactions, under ASM II, showing positions with increased Zlr scores | | | |
|---|---|---|---|
| Region | Position | Baseline Zlr | 1q32 weighted Zlr |
| 2q33.3 | - | - | - |
| 4p16.1 | 7.750 | 0.2629 | 1.2256 ($weight_{01}$) |
| 7q11.23 | 1.930 | 0.9598 | 1.3738 ($weight_{npl}$) |
| 10q23.31 | 4.840 | 1.7863 | 2.8010 ($weight_{npl}$) |
| 12q24 | 14.860 | 1.2718 | 1.8485 ($weight_{10}$) |
| 13q32 | 0.020 | 0.7637 | 1.6133 ($weight_{npl}$) |
| 15q12 | 8.470 | 1.2212 | 1.3528 ($weight_{npl}$) |
| 16p13.3 | 1.660 | 0.8685 | 2.1287($weight_{10}$) |
| 22q12.1 | - | - | - |

-: Negative contributions to Zlr, no significant increase in 1q32 weighted Zlr

To assess the significance of this difference, 1000 permutations, using the actual data, were performed. Nineteen out of 1000 permutations gave a result greater than 1.0147, with an associated *P* value of 0.019. Next, the permutation test was used to correct for the use of the 3 weighting schemes (by looking at the maximum difference of the three results) and produced a marginally significant result *P* value of 0.05. The second most significant difference, of 1.2606 ($weight_{01}$ minus baseline) was detected at region 16. 31 out of 1000 permutations gave a result greater than 1.0147, with an associated *P* value of 0.031. As with 10q23.31, a similar corrected P value of 0.055 was produced on chromosome 16, when amending for the use of multiple weights.

The second two-locus method involves calculating the MLS statistic, based on the two-locus IBD sharing probabilities of affected relative pairs, using the program TWOLOCARP (Cordell et al., 1995; Farrall, 1997; Cordell et al., 2000). Table 3.9 lists the MLS statistics for the MUT, ADD/HET and GEN models. For all locus pairs, the MLS scores for the MUL model were equal to the MLS results for the single loci (results not shown). Therefore, the multiplicative model does not explain any underlying interaction model that might be present. Overall, the individual GEN model fits better i.e. gave greater MLS scores than the single locus (MUL model) for all locus pairs. To estimate the significance of the difference between single locus (MUL model) MLS scores and GEN model MLS scores, 1000 simulations were carried out using the

Table 3.9: MLS statistics of TWOLOCARP analysis under ASM II showing regions with highest MLS scores for each locus pair under the GEN model

| Locus pair | Position* | MLS | | |
|---|---|---|---|---|
| | | MUL | HET | GEN |
| 1q32 - 2q33.3 | 2.0000 | 0.2841 | 0.3434 | 0.3434 |
| 1q32 - 4p16.1 | 9.0000 | 0.0313 | 0 | 0.8997 |
| 1q32 – 7q11.23 | 1.0000 | 3.0686 | 3.0977 | 3.1086 |
| 1q32 -10q23.31 | 9.0000 | 0.4805 | 0.3713 | 1.3997 |
| 1q32 - 12q24 | 2.0000 | 4.0063 | 4.7598 | 4.7598 |
| 1q32 - 13q32 | 5.0000 | 0.2556 | 0.2150 | 0.3737 |
| 1q32 - 15q12 | 1.0000 | 1.5617 | 2.2958 | 2.2958 |
| 1q32 -16p13.3 | 5.0000 | 1.2728 | 13898 | 1.3979 |
| 1q32 - 22q12.1 | 4.0000 | 0.1496 | 0.0837 | 0.3238 |

* Position (at three steps equidistant between 3 markers) where the highest MLS score was observed for each locus pair

SIMULATE option in MERLIN. Where the primary focus is the detection of a second disease locus (GEN model), out of 1000 simulations none of the MLS statistics proved to be significant. However when evaluating the significance of the statistic GEN-MUL (0.9192) for locus pair 1q32-10q23.31, i.e. to measure whether conditioning on 1q32 improved linkage to 10q23.31, 23 out of 1000 simulations (*P* value 0.023) gave a result greater than 0.9192. This test for interaction is most similar to the weighted analysis performed in Allegro.

Under ASM I, chromosome 12q24 showed weak evidence for linkage to BPD, *P* value of 0.0327. It was decided to carry out the same two-locus methods to assess the evidence for linkage and/or interaction at 1q32, 2q33.3, 4p16.1, 7q11.23, 10q23.31, 13q32, 15q12, 16p13.3 or 22q12.1, conditioned on the evidence for linkage at 12q24, under ASM I. As before, three weighting schemes (Cox et al., 1999) were created for analysis in Allegro, using the individual family contribution to the Zlr statistic at locus 12q24. No significant results were obtained (Zlr<1.96) for any of the locus pairs (Table 3.10).

| Table 3.10: Zlr statistics of Allegro analysis of two locus interactions, under ASM I, showing regions with increased Zlr scores | | | |
|---|---|---|---|
| Region | Position | Baseline Zlr | 12q24 weighted Zlr |
| 1q32 | 5.200 | 1.6627 | 1.8884 ($weight_{01}$) |
| 2q33.3 | 0.007 | -1.0639 | 0.1782 ($weight_{npl}$) |
| 4p16.1 | 7.750 | 0.2629 | 1.2256 ($weight_{01}$) |
| 7q11.23 | 0.665 | 0.8198 | 1.6683 ($weight_{npl}$) |
| 10q23.31 | 4.840 | 0.0386 | 1.1853 ($weight_{npl}$) |
| 13q32 | 0.002 | 0.5470 | 1.7154 ($weight_{npl}$) |
| 15q12 | 10.740 | 0.8453 | 1.5361 ($weight_{npl}$) |
| 16p13.3 | 0.415 | 1.9503 | - |
| 22q12.1 | 0.020 | 0.5809 | 0.8820 ($weight_{npl}$) |

-: No increase in any one of the three 12q24 weighted Zlr

TWOLOCARP (Cordell et al., 1995; Farrall, 1997; Cordell et al., 2000) analyses of locus pairs, under ASM I (Table 3.11), yielded the highest MLS score, 1.5294 (under the GEN model), for locus pair 12q24 and 1q32. As before 1000 simulations were carried out using the SIMULATE option in MERLIN and no significant linkages were detected at the second disease locus, based on the evidence for linkage to the first locus, i.e. 12q24. Furthermore, no evidence for interaction was detected between 12q24 and a second disease locus.

### 3.2.3 Bioinformatics

The number of genes extracted from the list of candidate regions (Table 1.2) implicated in the aetiology of BPD, i.e. the master list, totalled 4713. Three additional short lists were created, LIST_MICROARRAY (with 19 genes), LIST_SAGE (with 487 genes) and LIST_CNS (with 604 genes). The PERL utility get_matches_pl., which matches LIST_MICROARRAY, LIST_SAGE and LIST_CNS with genes in the master list (LIST_MASTER), yielded a short list, LIST_MATCHES, of 'flagged' genes (n=231), i.e. genes present, in LIST_MICROARRAY, LIST_CNS and/or LIST_SAGE; and therefore potential candidate genes involved in BPD. Table 3.12 shows the Gene Ontology information on the candidate genes 'flagged' on chromosomes 1q32 and 10q23.31.

The file containing the various lists and utilities may be downloaded at ftp.proto.hmg.uct.ac.za/CC/LISTS using anonymous login.

Table 3.11: MLS statistics of TWOLOCARP analysis under ASM I showing regions with highest MLS scores for each locus pair under the GEN model

| Locus pair | Position* | MLS | | |
|---|---|---|---|---|
| | | MUL | HET | GEN |
| 12q24 - 1q32 | 7.0000 | 1.0825 | 0.3364 | 1.5294 |
| 12q24 - 2q33.3 | 1.0000 | 0.0253 | 0.2064 | 0.2064 |
| 12q24 - 4p16.1 | 9.0000 | 0.0021 | 0 | 0.1180 |
| 12q24 - 7q11.23 | 1.0000 | 1.4061 | 1.6746 | 1.7193 |
| 12q24 - 10q23.31 | 1.0000 | 0 | 0.0305 | 0.0305 |
| 12q24 - 13q32 | 1.0000 | 1.1906 | 1.3009 | 1.3278 |
| 12q24 - 15q12 | 5.0000 | 0.5706 | 0.3210 | 0.6434 |
| 12q24 -16p13.3 | 3.0000 | 0.5601 | 0.2932 | 0.7109 |
| 12q24 - 22q12.1 | 9.0000 | 0.3174 | 0.0055 | 0.7380 |

* Position (at three steps equidistant between 3 markers) where the highest MLS score was observed for each locus pair

| Table 3.12: Gene ontology data on 'flagged' genes in linked regions 1q32 and 10q23.31 | | | |
|---|---|---|---|
| Gene | Molecular Function | Biological Process | Cellular component |
| *ADORA1*, adenosine receptor A1, 1q32 | A1 adenosine receptor activity, rhodopsin-like receptor activity | Signal transduction, neurogenesis, phagocytosis, inflammatory response, | Integral to plasma membrane |
| *NAV1*, Neuron navigator-1, 1q32 | ATP binding, nucleotide binding | No available information | No available information |
| *HTR7*, serotonin receptor 7, 10q23.31 | Serotonin receptor activity, melanocortin receptor activity | Synaptic transmission, circadian rhythm | Integral to plasma membrane |
| *PLCE1*, phospholipase C epsilon 1, 10q23.31 | Phosphoinositide phospholipase C activity, receptor signalling protein activity | Activation of MAPK, cell proliferation, inositol phosphate-mediated signalling | Plasma membrane, cytosol |

## 3.3 DISCUSSION

The aim of this study was the investigation of the molecular basis of BPD in South African families affected with the condition. The genetics of BPD is complex and is thus likely to be caused by multiple interacting susceptibility genes; therefore non-parametric linkage methods were applied in the identification of those regions that may confer susceptibility to this psychiatric condition. Genotype analyses were conducted on 22 Caucasian families with BPD, using 32 marker loci covering candidate regions on chromosomes 1q32, 2q33.3, 4p16.1, 7q11.23, 10q23.31, 12q24, 13q32, 15q12, 16p13.3, 18p11.2 and 22q12.1. Different analytical methods were employed using the programs MERLIN, Simwalk2, Allegro and TWOLOCARP. Single-locus analyses were carried out in (i) MERLIN, in conjunction with Simwalk2 (MERLIN/Simwalk2) for large pedigrees and (ii) MERLIN, after splitting the one large pedigree. Splitting the large pedigree, gave greater evidence for linkage to 1q32 (*P* value, 0.006, at marker D1S1678), under AMS II, compared to keeping the pedigree together (P value, 0.0526, at marker D1S1678). One possible explanation is that more distant relatives provide less information when a susceptibility allele is common (Badner et al., 1998), thus subsequent analyses were performed using split pedigrees.

Although primary single locus analysis has implicated chromosome 1q32, with marginal evidence for linkage to BPD, under ASM II, it is not considered significant for linkage on the basis of Lander and Kruglyak's (1995) proposed guidelines for a genomic scan. However, previously implicated candidate regions were investigated in the current study. The findings may therefore be considered meaningful on the basis of prior probability. Furthermore, nearly half of the families (41%) contribute positive scores to the total score on 1q32, suggesting that the linkage signal is not driven by a few large families. Under ASM I, chromosome 12q24 also showed marginal evidence for linkage to BPD, but to a lesser extent than the region on 1q32. Broadening of the classification of affected phenotypes to include major depression, may therefore have had a significant effect on the NPL score. None of the other candidate regions, 2q33.3, 4p16.1, 7q11.23, 10q23.31, 13q32, 15q12, 16p13.3 and 22q12.1, showed any evidence of linkage to BPD. However, given the multifactorial nature and expected heterogeneity of BPD, it is likely that contributions by these loci are of such small effect, that linkage is not detected, when examined individually. So,

to increase power to detect linkage to these loci, two-locus analyses were performed. Using model-free methods, the linkage information at 1q32, under ASM II, was incorporated into a test for linkage and/or interaction with a second disease locus, 2q33.3, 4p16.1, 7q11.23, 10q23.31, 12q24, 13q32, 15q12, 16p13.3 or 22q12.1, thereby avoiding indiscriminate application of two-locus analysis. In the first two-locus analysis method of Cox et al. (1999) families were weighted based on their contribution to the Zlr statistic at 1q32, using three weighting schemes. The most promising result was observed at 10q23.31 with $weight_{npl}$, generating a corrected *P* value of 0.05, suggesting an interaction between susceptibility loci on 1q32 and 10q23.31. When modelling interaction between 1q32 and 16p13, using $weight_{01}$, a corrected *P* value of 0.055 was obtained. Based on the TWOLOCARP analysis, the true genetic model is not multiplicative, as the MUL model MLS score was equal to the single locus MLS score. Furthermore, as expected, the individual GEN model better fits the two-locus sharing probabilities than the embedded MUL and HET models, thus indicating that these embedded models do not sufficiently explain the data. In the interest of unbiased reporting, two-locus analyses were also performed on 1q32, 2q33.3, 4p16.1, 7q11.23, 10q23.31, 13q32, 15q12, 16p13.3, and 22q12.1, based on the linkage result obtained at 12q24. Allegro and TWOLOCARP analyses, showed no evidence for linkage to and/or interaction with 1q32, 2q33.3, 4p16.1, 7q11.23, 10q23.31, 13q32, 15q12, 16p13.3, and 22q12.1. These results are all very marginal and could become significant by performing genetic analysis on additional affected family members and by typing the remaining families in our BPD research archive.

Owing to the application of statistical testing on two affection status models, ASM I and ASM II, a cautious interpretation of these findings is warranted. The two models have extensive overlap, and the tests are therefore not entirely independent. Furthermore, the same data are explored using various analytic methods which interrogate possible interactions between two loci, and with no transparent procedure of correcting for these effects (which are common to several linkage analyses in human genetics), false positive results are more likely to occur. The genetic mechanisms of anticipation, genetic imprinting and mitochondrial mutations that are known to generate complex patterns of inheritance have not been focussed on. Therefore, the findings of tentative linkage and interaction should only be regarded as the first steps in understanding the biologic properties of illness and its

significance in South African subpopulations. Nonetheless, the decision to continue with studying these regions might be better based on biological relevance in order to improve power for linkage analysis.

Traditionally, molecular biology research was carried out exclusively at the experimental laboratory bench, but the considerable increase in the amount of data being generated by the genome projects has necessitated the need to integrate computers into this research process. In the present investigation, the development and implementation of bioinformatics tools have been initiated in an effort to narrow down and prioritise the candidate genes, contained in the regions implicated in BPD (Table 1.2), into a short list that may be screened for mutations in affected individuals. With such a short list of candidate genes, hypothesis driven genotype analysis can be performed in the research laboratory. Furthermore, such tools will make it possible; (i) to pinpoint the underlying biochemical pathways most likely to be involved in disease pathology; (ii) to look for possible gene-gene interactions that may explain the phenotype and (iii) to ultimately assign biological meaning to genomic data. The following lists (1-4) were created:

1. a list of genes (LIST_MASTER, n=4672) using standardised HUGO gene symbols (www.gene.ucl.ac.uk/nomenclature) given the input of flanking markers or chromosome bands implicated in BPD (Table 1.2).
2. a list of genes according to the expression profile of human brain (LIST_SAGE, n=487)
3. a list of genes according to micro array expression data on BPD, published in the literature (LIST_MICROARRAY, n=19)
4. a list of genes based on biochemical pathway data (LIST_CNS, n=546) obtained from the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways

A preliminary 'matching' tool was developed that allows the comparison of these lists (2-4) to the list of genes created with regions implicated in BPD (1). Consequently, a ranked list of candidates, flagged according to hits of genes in lists 2, 3 and/or 4 (n=177), was created. This list, called LIST_MATCHES, and all the other lists can be downloaded at ftp.proto.hmg.uct.ac.za/CC/LISTS. Further investigations involve:

i. the interrogation of the Gene Ontology information

(http://www.geneontology.org), which aims to formalise our knowledge about biological processes, molecular functions and cellular components, and return a list of genes related by ontologies

ii. examining and interrogating available tools such as eVoke, the human expression vocabulary at the South African National Bioinformatics Institute (SANBI), creating a list of genes related by eVOC (http://www.sanbi.ac.za/evoc) ontologies,

iii. creating a list of published candidate genes and lists from any other sources that may become available.

Recent developments at Ensembl (http://www.ensembl.org), a comprehensive genome information resource, provide access to a variety of links to external databases, such as protein databases (SWISSPROT and trEMBL), ontology and expression databases (www.geneontology.org and eVoc at SANBI), making data retrieval and data mining sufficiently fast to accommodate a wide range of possible queries. Ensembl provides interactive interfaces such as BlastView, for sequence homology searches and MartView, which permits the user to formulate a query and select the output format, including genes, markers, SNPs, proteins, ontologies, and orthologues. While it is now possible to execute queries on live data and therefore generate more comprehensive lists compared to our original outputs, the comparison of these will still need to be carried out locally.

In 2004, Ogden and co-workers described the use of a convergent functional genomics approach in the identification of genes in BPD. The approach involves the integration of; (i) brain gene expression data from a pharmacogenomic mouse model (involving treatments with methamphetamine and valproate), (ii) human linkage data and (iii) variations in post-mortem brains from patients with bipolar disorder. Similar to our study, the authors came up with a short list of candidate genes for individual analysis in a prioritised manner. *DARPP-32* (dopamine- and cAMP-regulated phosphoprotein of 32kDa) on 17q12, *PENK* (preproenkephalin) on 8q12.1 and *TAC1* (tachykinin 1, substance P) on 7q21 topped their list of candidate genes. More recently and much like the strategy adopted in this study, Hattori et al. (2005), describes a "systems genetics" approach, requiring an intensive literature review and mining of electronic databases for entire pathways, in the compilation of candidate genes.

Individually, the chromosomal regions 1q32 and 10q have been implicated in BPD (Detera-Wadleigh et al., 1999, Cichon et al., 2001b). Based on the evidence of interaction observed between loci on chromosomes 1q32 and 10q23.31, the matching tool described in section 3.1.5, was used to 'flag' candidate genes in these two regions. Located on chromosome 1q32.1, Neuron navigator-1 (*NAV1*), is expressed in the developing brain (Maes et al., 2002). The exact function of this gene is not known in humans, although the *Caenorhabditis elegans* homolog, *unc-53*, is involved in axon guidance. Abnormal neuronal development has been implicated in BPD (Manji et al., 2000). The gene that encodes the adenosine A1 receptor is another interesting gene in this region. *ADORA1* regulates the release of neurotransmitters and, in rodents, antidepressant treatments such as electroconvulsive therapy (Newman et al., 1984), sleep deprivation (Yanik and Radulovacki, 1987) and carbamazepine (Daval et al.,1989) have been shown to increase ADORA1 density. Increased expression of ADORA1 may, therefore, be one of the mechanisms mediating antidepressant effects at the molecular level.

Located on chromosome 10q23.31, the serotonin receptor 7 (*HTR7*) is one of seven classes of receptors that mediate the effects of serotonin in the central nervous system (Barnes and Sharp, 1999). Abnormal functioning of the serotonergic system, including *HTR7*, has been implicated in BPD (Lin et al., 2003). A second candidate gene on 10q23.31, phospholipase C epsilon 1 (*PLCE1*) belongs to the phospholipase family that catalyses the hydrolysis of phosphoinositide 4, 5-biphosphate (PIP2) to diacylglycerol (DAG) and inositol-1, 4, 5-triphosphate (IP3) (Gould and Manji, 2002). Production of second messengers involved in phosphatidylinositol signalling system, i.e. DAG and IP3, depends on a continued source of free myo-inositol for the re-synthesis of precursor phospholipids. Free myo-inositol in turn, is generated by the dephosphorylation of myo-inositol monophosphate by the enzyme myo-inositol monophosphatase (IMPase) (Sjoholt et al., 2004). Known as the inositol depletion hypothesis, inhibition of IMPase enzyme, has been a reputed target of lithium treatment (Berridge et al., 1989).

In order to find biological evidence in support of possible interactions between *NAV1* or *ADORA1* on 1q32 and *HTR7* or *PLCE1* on 10q23.31 the genes were examined by Gene Ontology. Not much is known about the *NAV1* gene (Maes et al., 2002). None of the genes appear to have the same molecular function, although *ADORA1, HTR7*

and *PLCE1* all have receptor activity. Furthermore, the genes are involved in different biological processes although one could argue that not enough is known about these protein products and cannot be ruled out as ontologically candidates. Interestingly, *ADORA1, HTR7* and *PLCE1* are all integral to the plasma membrane; however, further investigation into interactions is required.

In summary, the results suggest that chromosomal regions, 1q32 and 10q23.31 not only contain susceptibility genes, but that their joint effects may play a role in the aetiology of BPD in at least a subset of populations. Further pedigree linkage studies of these regions are needed in delineating genetic vulnerability to BPD. In addition, with the application of bioinformatics, candidate genes on chromosomes 1q32 and 10q23.31 were identified, for genetic analysis in order to understand the underlying pathophysiology of BPD. With the use of bioinformatics, researchers are able to mine the wealth of biological information concealed in the mass of sequence data and to obtain a better understanding of the fundamental biology underlying complex disease and to channel this information to enhance the standard of healthcare.

The recruitment of extended families and "endogamous" populations, together with thoroughly investigated phenotypes need to be strongly encouraged. The present results should be followed up in the populations, in order to "mine" specific genetic aspects of BPD as a homogeneous entity.

As mentioned in Chapter 1 many patients diagnosed with BPD have one time or another in their lifetime attempted suicide, making suicide a serious complication for these individuals. Importantly, identification of these high risk patients is important, in order to prescribe effective treatment, and so reduce morbidity and mortality in patients with BPD. A subset of the patients in our BPD archive has attempted suicide (as described in Chapter 2) and in the next chapter the same methodological approaches, used in this chapter, are applied to try and uncover the underlying genetic susceptibility loci in this "sub-phenotype" of BPD.

# CHAPTER 4

## GENETICS OF SUICIDE ATTEMPTS

Suicidal behaviour is frequently encountered in the course of BPD (see Chapter 1 section 1.3.1). According to Tondo and Baldessarini (2000) suicide attempts in BPD are inclined to be of high lethality, with one completed suicide per three attempts, in contrast with one completed suicide in 18 attempts in the general population.

In 2001, in South Africa the National Injury Mortality Surveillance System (NIMSS, http://www.sahealthinfo.org/violence/nimssannual2001.htm) found that suicide accounted for 46.5% of all non-natural deaths in the 20-34 age ranges across all population groups. In addition, these data show that suicide rates peaked in the Asian population group in the 20-24 year age group with the bulk of cases occurring between years 15-34 (54.4%), while among Caucasians suicide peaked in the 30-34 and 45-49 year age groups (49.7%). Suicide-related deaths peaked in the 25-29 year age group among the Black (~22.4%) and Mixed Ancestry populations (17%). The NIMSS, and the South African National Burden of Disease study (Bradshaw et al., 2003), did not provide information on the influence of neuropsychiatric illness on suicide. However, if one considers suicide to emerge from this group of conditions, it contributes substantially to the health-related socio-economic burden in this country. The study by Lourens (1998) on suicide rates in Cape Town, South Africa, showed that 69% of individuals (N=44) who committed suicide, suffered from a psychiatric condition and, particularly, affective disorders (80%). Because of its size, this pilot study may not be an accurate reflection of the incidence of psychiatric disorders amongst individuals who commit suicide in South Africa, but it may indicate the general trend.

It is apparent that certain vulnerabilities predispose some individuals to suicidal behaviours more than others (Jamieson, 2000). In an effort to explain the occurrence of suicidal behaviour, i.e. the relationship between the underlying biological predisposition to suicide and the precipitants/stressors that trigger it, Mann (1999) developed the stress-diathesis model. This model states the risk for suicidal acts is determined by a stressor (including psychiatric illness and adverse life events) and a

diathesis (predisposition/vulnerability). Suicide attempters compared to non-attempters, with the same psychiatric condition and a similar number of adverse life events, experience more subjective depression and hopelessness and report fewer reasons for living (Mann et al., 1999; 2002). The overwhelming sense of hopelessness and greater number of suicidal ideations can be explained by the presence of vulnerability for such feelings in the face of disease. It is possible that the pressure of greater lifetime aggressivity and impulsivity presents a second vulnerability component in suicidal persons. Examination of the literature highlights a point of contention, with regards to the importance of psychotic symptoms in suicidal behaviour in persons with mood disorders; where some studies found no support for increased suicidality in relation to psychosis and affective disorder, whilst other studies support the association of psychotic features (such as the symptom of delusional thinking) in major depression and suicide (see Sher et al., 2001 for review). Recently, Warman and co-workers (2004) reported increased suicide ideation in a heterogeneous sample. In addition to the well-known risk factors – depression, hopelessness, substance abuse, and poor problem-solving – this increase in suicide ideation was associated with psychosis. Furthermore, in a comparison of multiple versus single suicide attempts, Forman and co-workers (2004) reported greater severity of psycopathology (including psychosis), suicidality (including ideation, intent and lethality) and interpersonal deficits in the group of multiple suicide attempters compared to the single suicide atternpters. The interaction between these various risk factors, diverse as they are in nature and their extent of complexity, still remains to be elucidated.

It has been reported that the family environment and genetic factors influence the threshold for suicidal behaviour in affective disorders (Linkowski et al., 1985; Brent et al., 1996; Sher et al., 2001). The high rates of attempted suicide in individuals diagnosed with affective disorders (Jamison, 2000; Oquendo et al., 2000; Grunebaum, 2001) stress the importance of identifying clinical and genetic characteristics that might aid in the recognition of those patients at high risk of suicidal behaviour. Because of their role in complex behaviour, initial genetic studies have regarded genes of the serotonergic and dopaminergic systems as primary candidates for investigation in suicidal behaviour (see Roy, 2001 for review). Table 4.1 lists a few of the genetic studies on suicidal behaviour.

Dysfunction of the serotonergic neurotransmitter system has been implicated in a variety of traits and behaviours such as mood, anxiety, impulsive aggression, sleep and appetite (Mendes de Oliveira et al., 1998; Serretti et al., 2001). Moreover, alterations in serotonergic neurons may play a role in the pathophysiology of depression (Owens and Nemerhoff, 1994). Many molecular genetic studies have implicated the serotonin transporter (*5-HTT*) and the serotonin receptor 2A (*5-HT2A*) in affective disorders and in suicidal behaviour (see Anguelova et al., 2003a; 2003b for systematic review). Tryptophan hydroxylase (*TPH*1), the rate-limiting enzyme in the biosynthesis of serotonin is a major candidate in the pathogenesis of BPD (Bellivier et al., 1998) and suicidality (Nielsen et al., 1994; Nielsen et al., 1998).

In the central nervous system (CNS), dopamine affects brain processes that control movement, cognition, emotion and neuro-endocrine secretion (Grandy et al., 1992; Emilien et al., 1999). Several studies implicate dopaminergic neurotransmission in BPD, because of the ability of; (i) amphetamine, a dopamine agonist, to precipitate mania (Silverstone and Romans-Clarkson, 1989) and, (ii) dopaminergic drugs to enhance the antidepressant activity in patients with treatment-resistant depressive disorders (Nierenberg et al., 1998).

Not many studies have investigated the dopaminergic system in suicidal behaviour. Even though decreased binding to D1 and D2 receptors in the PFC has been reported, studies have not determined whether there are alterations in dopamine or its major metabolite, homovanillic acid (HVA), in either the PFC or brainstem (Mann, 1998). Reasons for further evaluation of the dopaminergic system in suicidal behaviour include: (i) the observation of reduced cerebrospinal fluid HVA in suicide attempters in the context of major depression (Traskman et al., 1981), and (ii) findings connecting diminished dopaminergic function in major depression (Kapur and Mann, 1992). Genes of the dopaminergic system, including those coding for the dopamine transporter (*DAT*) and the dopamine receptors, as well as the *COMT* enzyme, which catalyses the methylation and hence degradation of catecholamines, are plausible candidates for suicidal behaviour.

In an attempt to further elucidate the pathophysiology of bipolar disorder, investigators have increasingly focussed on intracellular second messenger systems, in particular, the cAMP pathway (Duman et al., 1997). The postsynaptic response to

dopamine is mediated by cAMP (Crocker, 1994) and one of the downstream targets of the cAMP pathway, brain-derived neurotrophic factor (*BDNF*), has been suggested as a candidate molecule in the pathophysiology of psychiatric disorders. Duman et al. (1997) described the association of decreased levels of BDNF with major depressive disorder. Increased expression of *BDNF* was detected in hippocampal regions in subjects treated with antidepressant medications at the time of death, compared with untreated subjects (Chen et al., 2001). Furthermore, by typing SNPs in the gene, two independent research groups reported an association between the *BDNF* gene and BPD (Neves-Pereire et al., 2002; Sklar et al., 2002). More recently, Dwivedi and co-workers (2003) showed reduced gene expression of brain-derived neurotrophic factor (*BDNF*) in postmortem brain of suicide subjects.

Table 4.1: Genetic studies on suicidal behaviour

| Study | Type of study | Sample (origin) | Gene (polymorphism) | Findings |
|---|---|---|---|---|
| Nielsen et al., 1994 | Association study | 56 impulsive and 14 non-impulsive, alcoholic, violent offenders and 20 healthy controls (Finland) | *TPH*1 (A799C) | Association of *TPH*1 with a history of suicide attempts |
| New et al., 2001 | Association study | 145 personality disorder (PD) subjects (USA) | *HTR1B* (G816C) | Association between the G-allele and a history of suicide attempts in Caucasian subjects with PDs |
| Du et al., 1999 | Association study | 24 depressed suicide victims and 31 controls (Hungary) | *5-HTT* (5-HTTLPR), *HTR2A* (T102C and $His_{452}Tyr$) | The *5-HTTLPR* long-allele may be associated with suicide in depressed subjects |
| Du et al., 2000 | Association study | 120 cases with major depression and 131 controls (Canada) | *HTR2A* (T102C) | 102C allele associated with suicidal ideation in major depressive disorder |
| Preuss et al., 2001 | Association study | 163 patients with alcohol dependence and 117 healthy controls (South Germany) | *5-HTT* (5-HTTLPR) | Association between the *5-HTTLPR* short-allele and suicide attempts in alcohol-dependent subjects |
| Zubenko et al., 2004 | Linkage study | 81 families with recurrent, early-onset, major depression (USA) | *Genome-wide scan* | Evidence of suicide risk loci on 2p, 6q, 8p and Xq |
| Jernej et al., 2004 | Association study | 192 suicide victims and 377 controls without a history of neuropsychiatric disorders (Croatia) | *TPH1 (218AC) and 5-HTT (VNTR-2)* | Combined 5-HTT and TPH1 genotypes were associated with increased risk for suicidal behaviour |

The present investigation uses the candidate gene approach, focussing on genes with reported functional relevance for suicide. Brief summaries and annotations of these candidate genes, with their respective markers examined in association and linkage studies, are given below.

*i. Tryptophan hydroxylase (TPH1)*

Tryptophan hydroxylase (*TPH*1), the rate-limiting enzyme in the biosynthesis of serotonin is a major candidate in the pathogenesis of bipolar disorder (Bellivier et al., 1998) and suicidality (Nielsen et al., 1994; Nielsen et al., 1998). The human *TPH*1 gene has been localised to chromosome 11p14-p15.3. A significant association has been reported between the A218C polymorphism (A-allele) in intron 7 and bipolar disorder (Bellivier et al., 1998). The A218C polymorphism is in tight linkage disequilibrium with the polymorphism A779C in intron 7, which showed evidence of association to suicidality (Nielsen et al., 1998).

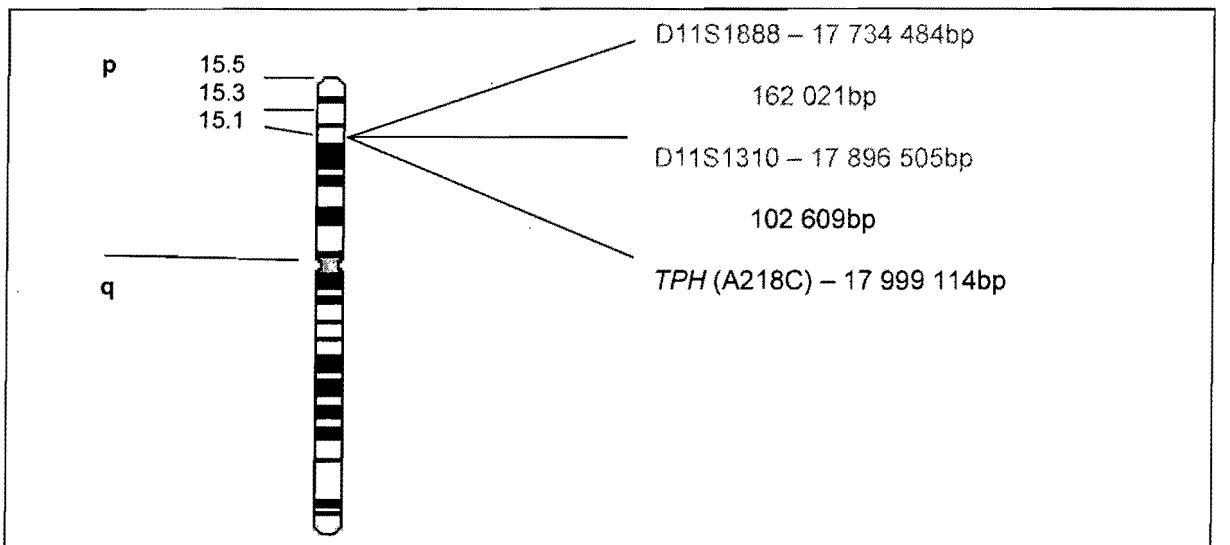


Figure 4.1: Ideogram of chromosome 11 with A218C polymorphism in *TPH1*; Key: Grey – microsatellite markers used for linkage analysis and the physical distances between markers. The orientation (i.e. distance in bp) is from telomere (pter) to centromere.

*ii. $D_4$ dopamine receptor (DRD4)*

The DRD4 gene, mapped to chromosome 11p15.5, contains a hypervariable region in the third cytoplasmic loop, comprised of 2-10 imperfect 48 base-pair variable number tandem repeats (VNTRs) (Van Tol et al., 1992). In an investigation of suicide attempters and normal controls, Persson and co-workers (1999) found no association between the 48bp VNTR of *DRD4* and suicide attempts. Interestingly, however, is the reported association between this polymorphism and novelty seeking and risk-taking behaviours (Benjamin et al., 1995).

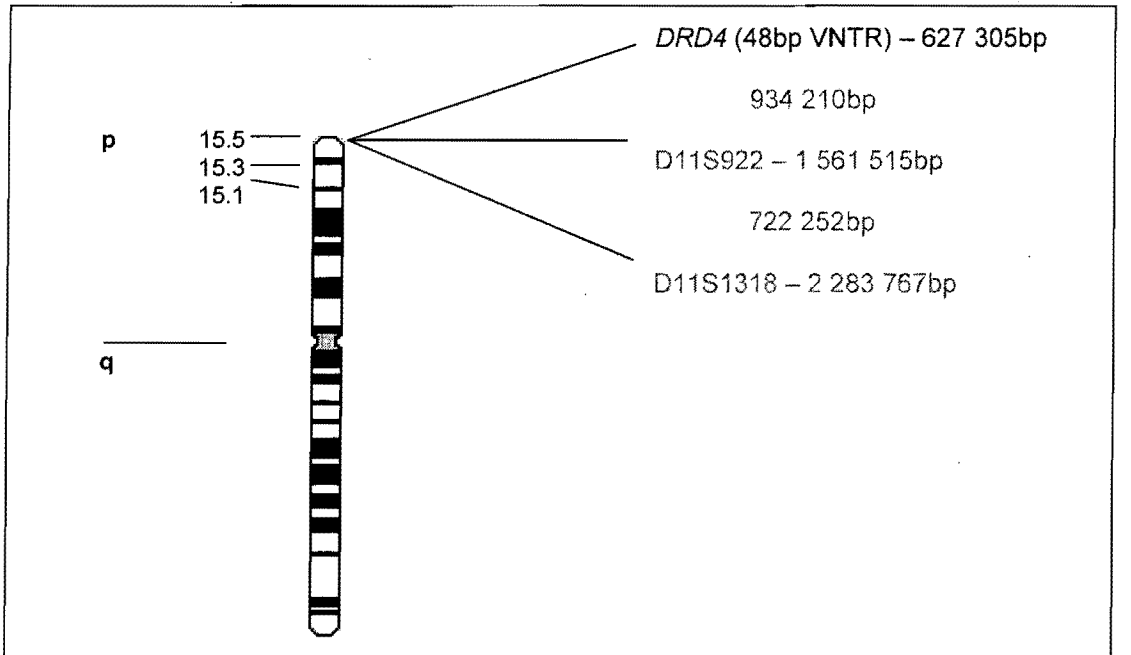


Figure 4.2: Ideogram of chromosome 11 with 48bp-VNTR in *DRD4*;
Key: Grey - the microsatellite markers used for linkage analysis and the distances between markers. The orientation (i.e. distance in bp) is from telomere (pter) to centromere.

*iii. BDNF*

The gene encoding *BDNF* maps to chromosome 11p13-14 (Hanson et al., 1992). It contains a missense SNP at nucleotide 196 (G/A) (Proschel et al., 1992) producing an amino acid substitution, val66met (dbSNP reference number rs6265). The val66met (V66M) polymorphism of *BDNF* has been associated with bipolar disorder (Neves-Pereira et al., 2002).

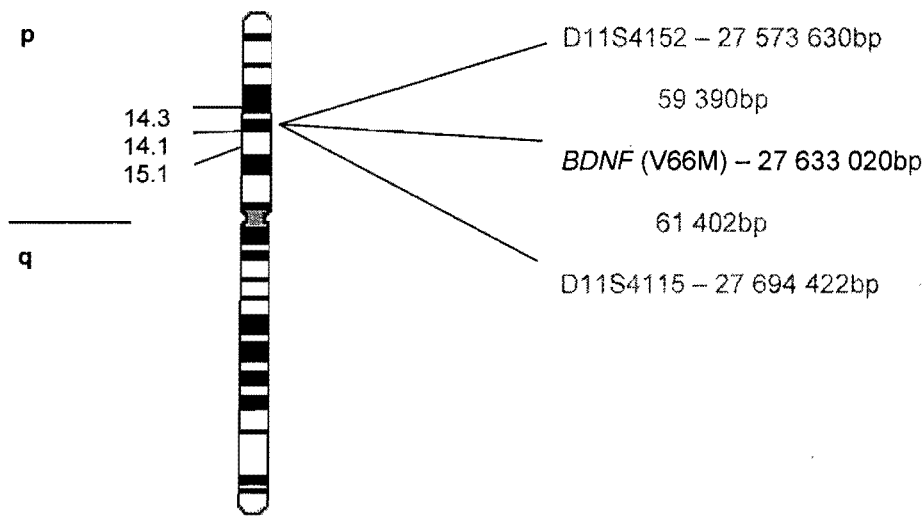


Figure 4.3: Ideogram of chromosome 11 with V66M polymorphism in *BDNF*; Key: Grey - the microsatellite markers used for linkage analysis and the distances between markers. The orientation (i.e. distance in bp) is from telomere (pter) to centromere.

*iv. Serotonin type 2A receptor (5-$HT_{2A}$)*

The human *5-$HT_{2A}$* gene codes for a post-synaptic receptor and is located on chromosome 13q14-q21. A silent mutation at position 102 (102 T/C) has been reported to be associated with bipolar affective disorder (Vincent et al., 1999). Du et al. (2000) and Arias et al. (2001) described an association between *5-$HT_{2A}$* -C allele and suicidality in major depression.

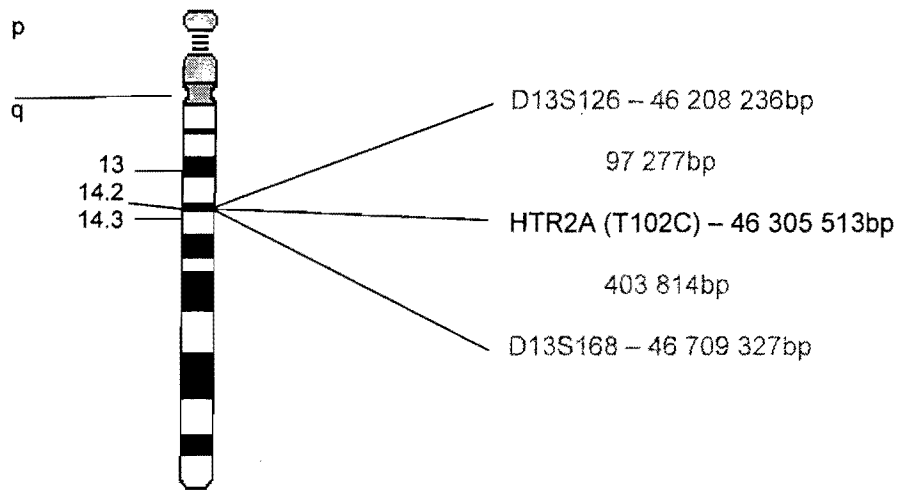


Figure 4.4: Ideogram of chromosome 13 with T102C polymorphism in *HTR2A*; Key: Grey – the microsatellite markers used for linkage analysis and the distances between markers. The orientation (i.e. distance in bp) is from telomere (pter) to centromere.

*v. Serotonin transporter (5-HTT)*

The human 5-HTT gene was mapped to chromosome 17q11.1-q12. A 44 base pair deletion/insertion (44bp ins/del) polymorphism in the promoter region results in two allelic variants, a long (l) variant and a short (s) variant with different transcriptional activities (Lesch et al., 1997). Hauser et al. (2003) described an association between the lower activity (s) allele and bipolar disorder. Du and co-workers (1999) reported an association between this *5-HTT* polymorphism and depressed suicide victims.

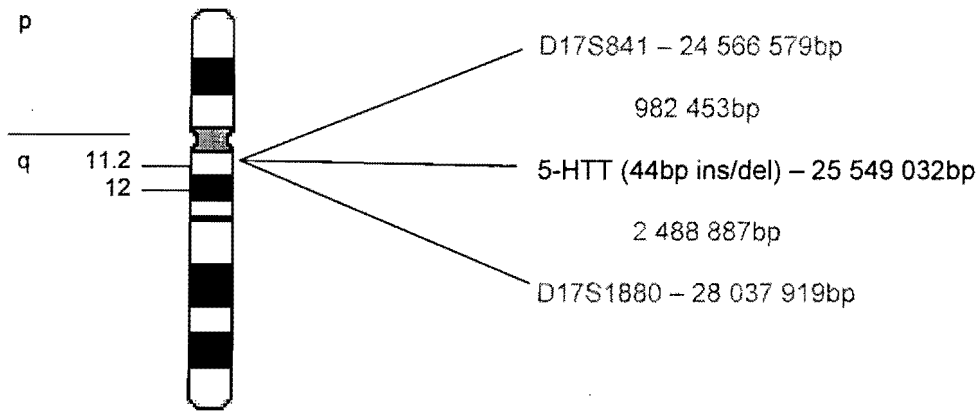


Figure 4.5: Ideogram of chromosome 17 with 44bp ins/del in *5-HTT*;
Key: Grey - the microsatellite markers used for linkage analysis and the distances between markers. The orientation (i.e. distance in bp) is from telomere (pter) to centromere.

***The Current Study***

From the initial cohort of recruited individuals (n=809, described in Chapter 2), a subset of subjects, diagnosed with an affective disorder and/or substance abuse, had a history of suicide attempts. This subset comprised of 73 individuals (between 1 and 8 members from a family), who admitted to attempting suicide at least once in their lifetime (ranging from 1 – 30 attempts). The present investigation was carried out in two stages; (i) an association study, comparing the frequency of the individual candidate gene polymorphisms in 46 unrelated cases (patients), diagnosed with an affective disorder and who have attempted suicide, and in healthy controls, and (ii) a non-parametric linkage study using microsatellite polymorphic markers, where extended families were available. As depicted in Table 4.1, large sample sets are advocated in the genetic dissection of suicidal behaviour. However, the relatively small dataset of the current study is meant to reflect its pilot nature, in order to ascertain the possibility of future large scale investigations.

# 4.1 EXPERIMENTAL APPROACH

## 4.1.1 Subjects

### *4.1.1.1 Cases vs controls*

Of the 73 individuals who admitted to attempting suicide, a total of 46 unrelated individuals from the Mixed Ancestry and Caucasian ethnic groups were included in a case-control (association) study. The cases comprised of patients suffering from BPI, BPII and MDE (single and recurrent) (Table 4.2).

Table 4.2: Summarised diagnostic data of subjects with a history of suicide attempts who were included in this study

| **Clinical diagnosis** | **Caucasian** | **Mixed Ancestry** |
|---|---|---|
| BPI | 19 | 12 |
| BPII | 3 | 2 |
| MDE-S | 2 | 0 |
| MDE-R | 3 | 5 |
| TOTAL | 27 | 19 |

Control samples from the South African subject matched background population (100 Caucasian and 60 Mixed Ancestry) were incorporated into the study. The control subjects were collected from amongst individuals referred to the human genetics laboratory for genetic research unrelated to psychiatric illness. Only research subjects who have formally consented to their DNA samples being used outside of their specific "research project" were utilised. Table 4.3 describes the gender, age, and ethnic characteristics of the cases and the control subjects.

### *4.1.1.2 Extended families*

Of the 46 individuals who attempted suicide, 12 (3 of Mixed Ancestry and 9 of Caucasian origin) had one or more relatives who also attempted suicide, i.e. a family history of suicide attempts. Because of the observed association between some markers and suicidality, genotypes were generated only in those families. Table 4.4

summarises the diagnoses of these nine families, where suicide attempts occurred within the context of an affective disorder.

Table 4.3: Age and gender distribution of cases and control samples

A. Caucasian data-set

| Variable | Cases **(n = 27)** | | Controls **(n = 100)** | |
|---|---|---|---|---|
| | Mean | SEM | Mean | SEM |
| Gender | 78 % F | | 69 % F [a] | |
| Age | 46.11<br>(23 – 70) | 2.62 | 51.16 [b]<br>(9 - 95) | 1.68 |

B. Mixed Ancestry data-set

| Variable | Cases **(n = 19)** | | Controls **(n = 60)** | |
|---|---|---|---|---|
| | Mean | SEM | Mean | SEM |
| Gender | 79 % F | | 71 % F [c] | |
| Age | 33.32<br>(23 – 45) | 1.65 | 48.10 [d]<br>(13 – 77) | 2.84 |

Data on gender and age was available for only 96 [a] and 97 [b] Caucasian control samples.

Data on gender and age was available for only 55 [c] and 40 [d] Mixed Ancestry control samples.

Table 4.4: Summary of the diagnoses in subjects from nine extended families (suicide attempters are shown under the respective diagnoses and in bold)

| FAMILY | BPI | BPII | MDE - R | MDE - S | AA | Psychosis | Other | No diagnosis | TOTAL |
|---|---|---|---|---|---|---|---|---|---|
| 1 | **2** | **1** | 4 | - | 1 | - | - | 2 | 10 |
| 2 | **1**+1 | 2 | **1** | 1 | - | - | - | 4 | 10 |
| 3 | - | - | **6**+2 | **2** | - | 1 | - | 4 | 15 |
| 4 | **1** | - | **1**+1 | - | - | - | - | 1 | 4 |
| 5 | 3 | **2**+3 | **1**+3 | 1 | - | - | 1 | - | 14 |
| 6 | **1** | 1 | **1**+6 | 1 | - | - | - | 3 | 13 |
| 7 | **2** | 1 | 1 | 1 | **1** | - | - | - | 6 |
| 8 | **1** | - | - | - | **1** | - | - | 1 | 3 |
| 9 | **1** | - | 1 | **1** | 1 | - | - | - | 4 |

Linkage analysis was performed for the mood disorder phenotype with and without the inclusion of suicide attempts, using three affection status models (ASMs). In an

attempt to assess whether suicide attempts made a significant contribution to the linkage statistic, ASM I consisted of those affected relatives who have attempted suicide, and ASM II included all affected relatives diagnosed with a mood disorder (excluding those who have attempted suicide). ASM III is ASM I and ASM II combined, in order to determine whether suicide attempts influenced the linkage information at the candidate genes.

## 4.1.2 DNA extraction

As described in Appendix B, section B.1.

## 4.1.3 Genotyping

### *4.1.3.1 Association study*

PCR was performed on genomic DNA for each of the five polymorphisms, shown in Table 4.5, using standard protocols (except for incorporation of 7-deaza dGTP for *DRD4*) and the primers indicated in Table 4.5. Where appropriate, the PCR products were digested with restriction enzymes, following the manufacturers' suggested protocol (Table 4.5). *TPH* and *5-HTT* DNA fragments were electrophoresed on 2.5% agarose gels and visualised by ethidium bromide staining. Digested PCR products were separated on 6% polyacrylamide gels (Appendix C, section C.4) at 220 V for 30min using the flat bed discontinuous buffer system, *Multiphor II* Electrophoresis Unit. Digested products were visualised by silver staining the gels after electrophoresis. In order to confirm the sequence of the restriction enzyme site, samples which exhibited homozygous and heterozygous alleles upon restriction enzyme digestion were selected for sequence analysis. The PCR products were first resolved on a 2% (w/v) agarose gel after which the DNA fragments were excised with a clean, sharp scalpel. DNA fragments were purified from the agarose gel slices using the *QIAquick* Gel extraction protocol (*Qiagen*, UK) (Appendix B, section B.5). Cleaned fragments were sequenced in both directions, using the reverse and forward primers of the original PCR reaction. Sequencing reactions were performed using the *BigDye* terminator cycle sequencing kit (*Applied Biosystems*, USA). Sequence products were purified with Sephadex G-50 fine columns (Princeton separations, USA) (Appendix B, section B.5) and separated by automated gel electrophoresis using an *ABI PRISM* 377 DNA sequencer (*Applied Biosystems*, USA). These samples were used as reference samples on subsequent gels.

### *4.1.3.2 Linkage study*

In addition to the genetic markers (Table 4.5) for each polymorphism, microsatellite markers (Table 4.6) were chosen from Ensembl (www.ensembl.org) and the Human Genome Browser (www.genome.ucsc.edu/) databases, for genotype analysis of the nine extended families. These microsatellite markers, chosen because of their close proximity to the candidate genes, consisted of tandem repeated dinucleotides and tetranucleotides. For three of the five candidate genes, *BDNF*, *HTR2A* and *5-HTT*, one marker on either side of the candidate gene was selected for genotype analysis. Appropriate flanking markers at a reasonable distance for the remaining two candidate genes, *TPH*1 and *DRD4*, did not exist, necessitating the use of markers on only one side of these genes. Marker heterozygosity ranged from 0.65 to 0.88 (average 0.740). The genomic region across the candidate locus of interest ranged from 0.12 - 3.5cM. Target sequences were amplified in 177 affected relative pairs (ARP), in a 10ul multiplex reaction solution, in 96-well plates, containing 200ng gDNA, 10pmol of each fluorescent primer labelled with either HEX or FAM (Synthetic DNA Laboratory, University of Cape Town, South Africa), reaction buffer containing 10mM Tris-HCl (pH8.3) and 50 mM KCl, 1.0 or 1.5 mM $MgCl_2$, 200uM dATP, dGTP, dCTP, dTTP (*Invitrogen*, UK) and 0.1 units of GoTaq DNA polymerase.

After an initial denaturation of the DNA templates for 3 min, 10 cycles were performed, each consisting of 94°C for 15 seconds, 55°C for 15 seconds and 72°C for 30 seconds. The remaining 20 cycles consisted of 89°C for 15 seconds, 50°C for 15 seconds and 72°C for 30 seconds. A final extension step involved incubating 72°C for 5 minutes. PCR products were typed and processed on the Applied Biosystems (ABI) 3100 sequencer (Perkin Elmer) (Appendix B, section B.3.2).

After an initial denaturation of the DNA templates for 3 min, 10 cycles were performed, each consisting of 94°C for 15 seconds, 55°C for 15 seconds and 72°C for 30 seconds. The remaining 20 cycles consisted of 89°C for 15 seconds, 50°C for 15 seconds and 72°C for 30 seconds. A final extension step involved incubating 72°C for 5 minutes. PCR products were typed and processed on the Applied Biosystems (ABI) 3100 sequencer (Perkin Elmer) (Appendix B section B.3.2).

| Table 4.5: The candidate genes and the specific polymorphisms genotyped in this study | | | | | | | |
|---|---|---|---|---|---|---|---|
| Chromosome | Gene Symbol | Polymorphism | Position | Restriction enzyme | Forward primer (5'-3') | Reverse primer (5'-3') | **Reference** |
| 11p15.3-15.5 | *TPH1* | A218C | intron 7 | BfaI | ttcagatcccttctataccccagag | caccactcgatgcaacatttg | Nielsen et al.,1992 |
| 11p15.5 | *DRD4* | 48 bp VNTR | exon 3 | N/A | gcgactacgtggtctactcg | ggtctgcggtggagtctg | Ebstein et al., 1996 |
| 11p13 | *BDNF* | Val66met | Codon66 | BbrP I | caagaggcttgacatcattggctg | ccgacatgtccactgcagtctttt | dbSNP reference number rs6265 |
| 13q14-q21 | *5HT2A* | T34C | exon 1 | MspI | tctgctacaagttctggctt | ctgcagctttttctctaggg | Warren et al., 1993 |
| 17q11.1-q12 | *5-HTT* | 44 bp ins/del | 5' regulatory region | N/A | ggcgttgccgctctgaatgc | gagggactgagctggacaaccac | Matsushita et al., 2001 |

| Table 4.6: List of primer sequences for microsatellite markers genotyped around candidate genes | | | |
|---|---|---|---|
| Chromosome (candidate gene) | Markers | Forward Primer (5'-3') | Reverse Primer (5'-3') |
| 11p15.3-15.5 (*TPH1*) | D11S1888 | ccccagtaccctgtataggc | cacttgtgtgtttgtatcgagtca |
| | D11S1310 | taagaaaaagcacctccaag | ctctgaacttccagtaggt |
| 11p15.5 (*DRD4*) | D11S922 | ggggcatctttggcta | tccggtttggttcagg |
| | D11S1318 | cccgtatggcaacagg | tgtgcatgtncatgagtg |
| 11p13 (*BDNF*) | D11S4115 | tggcatgtaaatntaagagactcac | ctgctacctcagaagtatctcaa |
| | D11S4152 | ccttagcagtagagaataacacaca | tcagccagccagattg |
| 13q14-q21 (*HTR2A*) | D13S126 | tcaccagtaaaatgctattgg | gtgattttcaaatttgctctg |
| | D13S168 | gcctagcccagtggtg | tgcttgtgcctatgttcttg |
| 17q11.1-q12 (*5-HTT*) | D17S841 | tggactttcttacatggcag | aggttagtagtctatgtcacagcg |
| | D17S1880 | agggattgcttgagcc | tgacagaatttgaacactttg |

## 4.1.4 Statistical analysis

### *4.2.4.1 Case-control study*

Comparisons of genotype and allele frequencies, of the respective gene polymorphisms, between the cases and controls were carried out using Pearson's chi-square ($\chi^2$) analysis with two-tailed P value. All statistical calculations were performed using the computer software program Epi-Info, version 6 (Centers for Disease Control [CDC], Atlanta, USA).

### *4.2.4.2 Family-based study*

Model-free single-locus and two-locus linkage analyses are detailed in Chapter 3, section 3.1.4. Briefly, single-locus analyses were carried out under ASM I, II and III, at three increments equidistant between markers (multipoint analysis), using the program MERLIN version 1.10.2 (Abecasis et al., 2002). The ALL scoring function of MERLIN was used to analyse nine families, each of which were adequately sized to be analysed by the software. Any locus showing evidence of linkage, i.e. an excess of IBD-allele sharing among ARPs at a single location, was used to perform two-locus analyses, i.e. modelling the joint effects of two loci. The two-locus analysis methods described in section 3.1.4.2, i.e. the weighted analysis of Cox et al. (1999) using Allegro version 1.2c (Gudbjartsson et al., 2000) and the likelihood-ratio analysis of Cordell et al. (1995, 2000) and Farrall et al. (1997) using TWOLOCARPARP, were performed on the present data set. In the case of the weighted analysis, where the Zlr statistic of the baseline model was improved by the weighted models, 1000 permutations were carried out in the statistical analysis package R, to determine the significance of this improvement. Alternatively, in the case of the likelihood-ratio analysis, where the MLS statistic of the GEN (two-locus) model has improved that of the single locus model, 1000 simulations were carried out in MERLIN (as described previously in Chapter 3 section 3.1.4.2), to determine the significance of conditioning on another locus.

## 4.2 RESULTS

### 4.2.1 Association analysis

In the total sample of unrelated cases, diagnosed with an affective disorder and with a history of suicide attempts (n=46), 59% were of Caucasian descent and 41% were of Mixed Ancestry. Table 4.2 lists the clinical characteristics of the cases under investigation, with BPI as the most common diagnosis in both population groups. The mean age of the Caucasian cases was 46 years, compared to 33 years in the Mixed Ancestry cases (Table 4.3). The majority of cases and controls were females, in both the Caucasian and Mixed Ancestry population groups. Table 4.7 shows the genotype and allele frequencies of the five polymorphisms under investigation, in cases and controls. The genotype distributions in all the groups were in Hardy-Weinberg equilibrium.

Significant differences were observed in the genotypic distributions of the *DRD4* locus between the Caucasian and the Mixed Ancestry cases ($\chi^2$ = 6.20, df = 2, $P$ = 0.045), thereby verifying the need for separate analysis of the two population groups. Significant differences were observed in the genotypic and allelic distribution of the *DRD4* 48 bp VNTR polymorphism between Caucasian cases and control subjects (genotype, $\chi^2$ = 5.976, df = 2, $P$ = 0.05; allele frequency, $\chi^2$ = 8.16, df = 1, $P$ = 0.004; odds ratio, 2.42). Following correction for multiple testing (Bonferroni correction, $p<0.01$), the allele distribution was marginally significant in the Caucasian cases. The 4/4-repeat genotype and the 4-repeat allele were significantly less frequently observed in individuals, with affective disorders who had attempted suicide, than in the background controls. A similar effect in this gene was not observed for the Mixed Ancestry group. No significant differences were observed between cases and controls (in either the genotype distribution or the allele frequencies) for the remaining polymorphisms, in both population groups.

### 4.2.2 Family-based study (allele-sharing)

A total of 15 genetic markers, covering 5 candidate genes, were genotyped in 9 multiplex families, under three affection status models (ASM I, II and III). Non-parametric linkage methods were used to identify excess IBD allele-sharing among affected relative pairs.

Table 4.7: Genotype and allele distributions for the candidate gene loci.

| **Gene** | **Ethnic group** | **Disease status** | **Genotype frequency** | | | | **Allele frequency** | | |
|---|---|---|---|---|---|---|---|---|---|
| *DRD4* | | | **4/4 (%)** | **4/7 (%)** | **4/other (%)** | **other (%)** | **4 (%)** | **7 (%)** | **other (%)** |
| | Caucasian | Cases | 12 (44) | 0 | 10 (37) | 5 (19) | 34(63) | 0 | 20 (37) |
| | | Controls | 68 (68) | 0 | 25 (25) | 7 (7) | 161(81) | 0 | 39 (19) |
| | Mixed Ancestry | Cases | 14 (78) | 1 (5) | 2 (12) | 1 (5) | 31(86) | 1 (3) | 4 (11) |
| | | Controls | 46 (77) | 0 | 8 (13) | 6 (10) | 100(83) | 1 (1) | 19 (16) |

Table 4.7: *Continued*

| **Gene** | **Ethnic Group** | **Disease status** | **Genotype frequency** | | | **Allele frequency** | |
|---|---|---|---|---|---|---|---|
| *TPH1* | | | A/A (%) | A/C (%) | C/C (%) | A (%) | C(%) |
| | Caucasian | Cases | 3 (11) | 16 (59) | 8 (30) | 22 (41) | 32 (59) |
| | | Controls | 17 (17) | 52 (52) | 31 (31) | 86 (43) | 114 (57) |
| | Mixed Ancestry | Cases | 0 | 9 (47) | 10 (53) | 9 (24) | 29 (76) |
| | | Controls | 5 (9) | 30 (48) | 25 (43) | 40 (33) | 80 (67) |
| *BDNF* | | | G/G (%) | G/A (%) | A/A (%) | G (%) | A (%) |
| | Caucasian | Cases | 17 (63) | 7 (26) | 3 (11) | 41 (76) | 13 (24) |
| | | Controls | 62 (62) | 31 (31) | 7 (7) | 155 (77) | 45 (23) |
| | Mixed Ancestry | Cases | 14 (74) | 4 (21) | 1 (5) | 32 (84) | 6 (16) |
| | | Controls | 32 (53) | 26 (43) | 2 (4) | 90 (75) | 30 (25) |

Table 4.7: *Continued*

| **Gene** | **Ethnic Group** | **Disease status** | **Genotype frequency** | | | **Allele frequency** | |
|---|---|---|---|---|---|---|---|
| *5HT2A* | | | **T/T (%)** | **T/C (%)** | **C/C (%)** | **T (%)** | **C (%)** |
| | Caucasian | Cases | 6 (22) | 9 (33) | 12 (45) | 21(39) | 33 (61) |
| | | Controls | 19 (19) | 45 (45) | 36 (36) | 83(42) | 117(58) |
| | Mixed Ancestry | Cases | 4 (21) | 8 (42) | 7 (37) | 16(42) | 22 (58) |
| | | Controls | 7 (12) | 34 (57) | 19 (31) | 48(40) | 72 (60) |
| *5-HTT* | | | **L/L (%)** | **L/S (%)** | **S/S (%)** | **L (%)** | **S (%)** |
| | Caucasian | Cases | 10 (37) | 12 (44) | 5 (19) | 32 (59) | 22 (41) |
| | | Controls | 30 (30) | 51 (51) | 19 (19) | 111 (56) | 89 (44) |
| | Mixed Ancestry | Cases | 9 (47) | 9 (47) | 1 (6) | 27 (71) | 11 (29) |
| | | Controls | 27 (47) | 24 (42) | 6 (11) | 78(68) | 36 (32) |

#### *4.2.2.1 Single-locus analysis*

In single locus analyses the best evidence of linkage was found for *TPH*1 and *BDNF*, under all three ASMs. Table 4.8 lists the results of the multipoint MERLIN analysis, carried out at three steps equidistant between marker loci. In an effort to correct for testing linkage at five loci, we considered *P* values less than 0.01 as significant. *BDNF* yielded significant evidence of linkage under ASM I (at V66M, Lod=1.32, *P* value 0.007) and ASM II (at D11S4152, Lod=1.25, *P* value 0.008). Adding the suicide attempters (ASM I) to individuals with mood disorders (ASM II), similar evidence of linkage was detected to *BDNF*, under ASM III (at D11S4152, Lod=1.26, *P* value 0.007). *TPH*1 showed evidence of linkage under ASM II (at D11S1888, Lod=1.33, *P* value 0.007) and not under ASM I. Adding the suicide attempters (ASM I) with individuals with mood disorders (ASM II), did not make a significant contribution to linkage at *TPH*1 (ASM III, Lod=1.21, *P* value 0.009). Subsequent two-locus interactions were carried out under both ASM I and II.

Table 4.8: Multipoint linkage data derived through MERLIN analysis ($P < 0.01$)

| **Chromosome** | | | **ASM I** | | **ASM II** | | **ASM III** | |
|---|---|---|---|---|---|---|---|---|
| **(Gene)** | **Markers** | **Distance** [a] | ***P*(NPL$_{ALL}$)** | ***P*(LOD)** [b] | ***P*(NPL$_{ALL}$)** | ***P*(LOD)** [b] | ***P*(NPL$_{ALL}$)** | ***P*(LOD)** [b] |
| 11 (*TPH1*) | D11S1888 | 0.000 | 1.41 | 0.03 | 0.0004 | **0.007** | 1.97 | **0.009** |
| | | 0.041 | 1.33 | 0.03 | 0.0004 | 0.007 | 1.92 | 0.010 |
| | | 0.81 | 1.26 | 0.04 | 0.0004 | 0.007 | 1.88 | 0.011 |
| | | 0.122 | 1.18 | 0.04 | 0.0004 | 0.007 | 1.84 | 0.012 |
| | D11S1310 | 0.162 | 1.10 | 0.06 | 0.0004 | 0.007 | 1.80 | 0.013 |
| | | 0.188 | 1.10 | 0.06 | 0.0007 | 0.008 | 1.70 | 0.02 |
| | | 0.213 | 1.10 | 0.06 | 0.0012 | 0.009 | 1.60 | 0.02 |
| | | 0.238 | 1.10 | 0.06 | 0.002 | 0.010 | 1.50 | 0.02 |
| | A218C | 0.264 | 1.10 | 0.06 | 0.003 | 0.012 | 1.40 | 0.02 |
| 11 (*DRD4*) | 48bp VNTR | 0.000 | -0.61 | 0.7 | 0.5 | 0.4 | -0.61 | 0.7 |
| | | 0.233 | -0.74 | 0.8 | 0.5 | 0.5 | -0.74 | 0.8 |
| | | 0.466 | -0.75 | 0.8 | 0.5 | 0.4 | -0.75 | 0.8 |
| | | 0.698 | -0.64 | 0.7 | 0.4 | 0.3 | -0.64 | 0.7 |
| | D11S922 | 0.931 | -0.41 | 0.7 | 0.3 | 0.2 | -0.41 | 0.7 |
| | | 1.111 | -0.45 | 0.7 | 0.3 | 0.2 | -0.45 | 0.7 |
| | | 1.292 | -0.53 | 0.7 | 0.3 | 0.3 | -0.53 | 0.7 |
| | | 1.472 | -0.66 | 0.8 | 0.4 | 0.3 | -0.66 | 0.8 |
| | D11S1318 | 1.653 | -0.82 | 0.8 | 0.4 | 0.4 | -0.88 | 0.8 |

Table 4.8: *Continued*

| **Chromosome** | | | **ASM I** | | **ASM II** | | **ASM III** | |
|---|---|---|---|---|---|---|---|---|
| **(Gene)** | **Markers** | **Distance** [a] | ***P*(NPL$_{ALL}$)** | ***P*(LOD)** [b] | ***P*(NPL$_{ALL}$)** | ***P*(LOD)** [b] | ***P*(NPL$_{ALL}$)** | ***P*(LOD)** [b] |
| 11 (*BDNF*) | D11S4115 | 0.000 | 2.29 | 0.014 | 0.0002 | **0.008** | 2.09 | **0.007** |
| | | 0.015 | 2.41 | 0.011 | 0.0003 | 0.012 | 2.03 | 0.009 |
| | | 0.030 | 2.53 | 0.009 | 0.0005 | 0.02 | 1.97 | 0.010 |
| | | 0.044 | 2.65 | 0.008 | 0.0010 | 0.03 | 1.91 | 0.012 |
| | V66M | 0.059 | 2.77 | **0.007** | 0.002 | 0.05 | 1.85 | 0.014 |
| | | 0.064 | 2.78 | 0.006 | 0.0012 | 0.04 | 1.88 | 0.013 |
| | | 0.068 | 2.79 | 0.006 | 0.0009 | 0.03 | 1.90 | 0.012 |
| | | 0.072 | 2.80 | 0.006 | 0.0009 | 0.02 | 1.91 | 0.011 |
| | D11S4152 | 0.077 | 2.81 | 0.006 | 0.0007 | 0.02 | 1.93 | 0.011 |
| 13 (*HTR2A*) | D13S126 | 0.000 | -0.65 | 0.7 | 0.5 | 0.4 | -0.53 | 0.6 |
| | | 0.025 | -0.65 | 0.7 | 0.5 | 0.4 | -0.52 | 0.6 |
| | | 0.049 | -0.65 | 0.7 | 0.5 | 0.4 | -0.52 | 0.6 |
| | | 0.073 | -0.65 | 0.7 | 0.4 | 0.4 | -0.51 | 0.6 |
| | T102C | 0.098 | -0.66 | 0.7 | 0.4 | 0.3 | -0.50 | 0.6 |
| | | 0.198 | -0.58 | 0.7 | 0.4 | 0.2 | -0.41 | 0.6 |
| | | 0.298 | -0.40 | 0.7 | 0.4 | 0.2 | -0.25 | 0.6 |
| | | 0.390 | -0.13 | 0.6 | 0.3 | 0.12 | -0.02 | 0.5 |
| | D13S168 | 0.498 | 0.24 | 0.4 | 0.3 | 0.12 | 0.28 | 0.4 |

Table 4.8: *Continued*

| **Chromosome** | | | **ASM I** | | **ASM II** | | **ASM III** | |
|---|---|---|---|---|---|---|---|---|
| **(Gene)** | **Markers** | **Distance** [a] | ***P*(NPL$_{ALL}$)** | ***P*(LOD)** [b] | ***P*(NPL$_{ALL}$)** | ***P*(LOD)** [b] | ***P*(NPL$_{ALL}$)** | ***P*(LOD)** [b] |
| 17 (*5-HTT*) | D17S841 | 0.000 | 0.18 | 0.4 | 0.2 | 0.3 | 1.37 | 0.10 |
| | | 0.246 | 0.10 | 0.5 | 0.3 | 0.3 | 1.22 | 0.11 |
| | | 0.491 | 0.01 | 0.5 | 0.4 | 0.4 | 1.08 | 0.14 |
| | | 0.737 | -0.08 | 0.5 | 0.5 | 0.5 | 0.94 | 0.2 |
| | 44bp ins/del | 0.982 | -0.16 | 0.6 | 0.6 | 0.7 | 0.80 | 0.2 |
| | | 1.593 | -0.19 | 0.6 | 0.4 | 0.4 | 0.94 | 0.2 |
| | | 2.204 | -0.21 | 0.6 | 0.3 | 0.2 | 1.07 | 0.13 |
| | | 2.816 | -0.23 | 0.6 | 0.14 | 0.2 | 1.20 | 0.11 |
| | D17S1880 | 3.427 | -0.26 | 0.6 | 0.06 | 0.11 | 1.32 | 0.09 |

[a] Distance in cM, calculated from our data

[b] *P* value for the Kong and Cox allele-sharing lod score (Zlr statistic)

### *4.2.2.2 Two-locus analysis*

*TPH*1 (under ASM II) and *BDNF* (under both ASM I and II) were used as the conditioning loci in assessing evidence for linkage at a second candidate gene (the test locus), i.e. in the first round of two-locus interactions, *DRD4*, *BDNF*, *HTR2A* and *5-HTT* were conditioned on *TPH*1 and in the second *TPH*1, *DRD4*, *HTR2A* and *5-HTT* were conditioned on *BDNF*.

The first method involved modelling positive interactions (using $weight_{01}$ and $weight_{npl}$) and heterogeneity (using $weight_{01}$) (Cox et al., 1999), through the program Allegro, based on the family contributions to the Zlr statistic at *TPH*1 (Table 4.9) and *BDNF* (Table 4.10). Table 4.9 shows the baseline Zlr values detected at the second/test locus and those weighted by the contribution of *TPH*1, under ASM II (mood disorders only). *TPH*1 was not used as a conditioning locus in two-locus interactions under ASM I (suicide attempters only), as it was not significant in single-locus analyses. The significance associated with the increased Zlr, was determined by performing 1000 permutations. None of the differences between baseline and *TPH*1 weighted Zlr statistics at any of the test loci were significant. Table 4.10 shows the baseline Zlr values detected at the second/test locus and those weighted by the contribution of *BDNF*, under ASM I (Table 4.10A) and II (Table 4.10B). The significance of the differences between baseline and *BDNF* weighted Zlr scores, i.e. the increased Zlr, was assessed by permutation. Based on the evidence of linkage at *BDNF*, using family-specific weights, we found no evidence of linkage and/or interaction between *BDNF* and any of the test loci under both ASM I and II.

Table 4.9: *TPH*1 weighted Zlr statistics of two locus interactions, under ASM II (showing positions with increased Zlr scores)

| Gene (test locus) | Position | Baseline Zlr | *TPH*1 weighted Zlr |
|---|---|---|---|
| *DRD4* | 0.940 | 0.9295 | 1.9648 ($weight_{npl}$) |
| *BDNF* | 0.000 | 2.4429 | 2.5482 ($weight_{npl}$) |
| *HTR2A* | - | - | - |
| *5-HTT* | 3.520 | 1.3808 | 1.5313 ($weight_{npl}$) |

-: No increase in *TPH*1 weighted Zlr

Table 4.10: *BDNF* weighted Zlr statistics of two locus interactions (showing positions with increased Zlr scores), under ASM I and ASM II

A. ASM I

| Gene | Position | Baseline Zlr | *BDNF* weighted Zlr |
|---|---|---|---|
| *TPH1* | 0.000 | 1.5862 | 1.7498 ($weight_{01}$) |
| *DRD4* | 0.940 | -0.4852 | 0.1142 ($weight_{01}$) |
| *HTR2A* | 0.500 | 0.1279 | 0.7058 ($weight_{npl}$) |
| *5-HTT* | 0.000 | 0.3523 | 1.0068 ($weight_{10}$) |

B. ASM II

| Gene | Position | Baseline Zlr | *BDNF* weighted Zlr |
|---|---|---|---|
| *TPH1* | 0.000 | - | - |
| *DRD4* | 0.940 | 0.9295 | 1.3480 ($weight_{npl}$) |
| *HTR2A* | 0.500 | 1.9965 | 1.3056 ($weight_{npl}$) |
| *5-HTT* | 0.000 | 1.4810 | 1.8367 ($weight_{npl}$) |

-: No increase in *TPH1* weighted Zlr

The second two-locus method assesses two-locus IBD sharing probabilities of affected relative pairs, by calculating MLS scores with the program TWOLOCARP (Cordell et al., 1995; Farrall, 1997; Cordell at al., 2000). Table 4.11 shows the MLS scores for the MUL, HET and GEN models for *TPH1* locus pairs (under ASM II) and *BDNF* locus pairs (under ASM I and II). For all locus pairs, the MLS scores for the MUL model (defined as the sum of the two single-locus MLSs) were equal to the MLS results for the single loci (results not shown). Therefore, the multiplicative model does not explain any underlying interaction model that might be present.

Table 4.11: MLS statistics of two locus interactions, showing regions with highest MLS scores for each locus pair, under the GEN model

A. ASM I

| Two-Locus pair | Position* | MLS | | |
|---|---|---|---|---|
| | | MUL | HET | GEN |
| *BDNF-TPH1* | 1.0000 | 0.8088 | 1.5512 | 1.9081 |
| *BDNF-DRD4* | 9.0000 | 0.0226 | 0 | 0.2107 |
| *BDNF-HTR2A* | 9.0000 | 0.227 | 0 | 0.1295 |
| *BDNF-5-HTT* | 1.0000 | 0.4826 | 1.6884 | 1.6890 |

B. ASM II

| Two-Locus pair | Position* | MLS | | |
|---|---|---|---|---|
| | | MUL | HET | GEN |
| *TPH1-DRD4* | 5.0000 | 0.8806 | 0.0 | 1.9344 |
| *TPH1-BDNF* | 1.0000 | 2.1558 | 0.4922 | 3.6565 |
| *TPH1-HTR2A* | 9.0000 | 0.8510 | 0.0287 | 2.4541 |
| *TPH1-5-HTT* | 9.0000 | 1.4335 | 0.3812 | 3.0840 |
| *BDNF-TPH1* | 5.0000 | 3.2462 | 1.5827 | 4.7470 |
| *BDNF-DRD4* | 5.0000 | 0.8806 | 0.1885 | 1.8920 |
| *BDNF-HTR2A* | 9.0000 | 0.8510 | 0.2955 | 1.3519 |
| *BDNF-5-HTT* | 9.0000 | 1.4335 | 0.7645 | 1.8993 |

Position (at three steps equidistant between 3 markers) where the highest MLS score was observed for each locus pair

The GEN model gave a greater MLS score than the single locus (MUL model) for all the *TPH*1 locus pairs (under ASM II, Table 4.11B). In order to assess whether this conditioning on *TPH*1 (under the GEN model) is significantly better than the single locus, 1000 simulations were carried out, using the SIMULATE option in MERLIN. The only significant MLS score under the single locus model, was obtained at *BDNF* (*P* value of 0.043), which is consistent with the MERLIN single-locus analysis. Under the GEN model, locus pairs *TPH*1-*BDNF* (*P* value of 0.028) and *TPH*1-5-HTT (*P* value of 0.043) provided significant evidence of interaction. Furthermore, conditioning on *TPH*1

improved linkage at *HTR2A* (*P* value of 0.033). There was no evidence of linkage and/or interaction to *DRD4*.

Similar to the *TPH1* locus pairs, greater MLS scores were observed under the GEN model, compared to the single locus (MUL) model for all *BDNF* locus pairs (under both ASM I and II). In order to assess whether conditioning on *BDNF* (under the GEN model) significantly improves linkage to the single locus, simulations were carried out in MERLIN and, as before, no evidence of interaction was detected under ASM I. Under ASM II, *TPH1* gave the only significant single-locus MLS score (*P* value of 0.016), which is consistent with the MERLIN single-locus analysis. Conditioning on *BDNF* provided evidence of interaction between *BDNF* and *TPH1* (*P* value of 0.014, under the GEN model). None of the other locus pairs proved significant for linkage or interaction.

## 4.3 DISCUSSION

Bipolar disorder includes a component of impulsivity that determines an individual's risk for impulsive acts (Swann et al., 2003). It is widely accepted that suicide attempts (associated with violent and impulsive traits), broadly defined as self-destructive behaviour with the purpose of ending one's life, are more prevalent in mood disorders than in the general population (Guze and Robins, 1970; Oquendo et al., 2000; Jamison, 2000; Sher et al., 2001). Suicidal behaviour is dependent on inherent genetic susceptibility, which is influenced by environmental factors.

Upon closer examination of the diagnostic classifications of the 809 individuals archived in our research laboratory, a subset of 73 individuals diagnosed with an affective disorder also presented with a history of suicide attempts. At the time of the interview, the question of suicide attempts was not of primary concern. It is likely that further directed interrogation of this "phenotype" could have shown a greater incidence of suicide attempts amongst the subjects. It was decided to carry out a pilot study on the genetics of suicide attempts, on 46 of the 73 cases (i.e. unrelated cases suffering from an affective mood disorder) and controls in a test for association between suicide attempts and selected polymorphisms in five candidate genes. The genotypic and allelic

distributions were compared between cases and controls, gathered from the Caucasian (n=100) and Mixed Ancestry (n=60) population groups. In order to eliminate the possible confounding effect of population non-stratification, the Caucasian and Mixed Ancestry population groups were analysed separately.

The greater number of female cases, in both the Caucasian and the Mixed Ancestry population groups can be attributed to one or a combination of three possible reasons, (i) females are more likely to acknowledge attempting suicide, (ii) more females were willing to participate in the study, and (iii) females are at a higher risk than males for suicide attempts (Chen and Dilsaver, 1996; Wunderlich et al., 2001).

An interesting finding of this study was a significant difference in the frequency of the 4/4-repeat genotype in the *DRD4* gene, between Caucasian cases and control subjects (44% vs. 68%). Furthermore, the 4-repeat allele is significantly less frequent in the cases than controls (63% vs. 81%). This allele may therefore confer a possible 'protection' against susceptibility to development of an affective mood disorder and/or attempted suicide. *DRD4* receptors are abundant in the frontal brain, a region recognized to play a vital role in human personality, in particular irritability and impulsivity (Benjamin et al., 1996; Ebstein et al., 1996). Several reports indicated an association between the length of *DRD4* alleles and risk-taking behaviours (for review, see Kluger et al., 2002). Suicide and suicidal behaviour might be considered as an extension of novelty-seeking behaviour with aggression, violence and depression being perhaps significant in the act of suicide (Apter et al., 1995). However, there is limited data on the association of suicidal behaviour and BPD with the *DRD4* exon III polymorphism. The possibility that functional differences in the affinity and intracellular signalling system of the *DRD4,* exon III protein variants may play a role in the aetiology of suicide attempts in affective mood disorders is as yet unclear and warrants further investigation.

In a world-wide population sample (including individuals of African origin) the *DRD4* 4-repeat and 7-repeat alleles were the most common, occurring in 65.1% and 19.2% of the population, respectively (Ding et al., 2002). Interestingly, the 7-repeat allele appears to be very rare in the South African Caucasian and Mixed Ancestry populations included

in the present study. The *DRD4* association was not observed in the Mixed Ancestry subject group and may possibly be due to the small sample size and/or an extensive history of admixture in this ethnic group. The results show no association between suicide attempts and the remaining alleles under investigation.

In an attempt to understand the genetic diathesis (predisposition) component that underlies suicidal behaviour, we used a family-based design and performed a linkage study on the extended family members of some of the cases. Only nine Caucasian cases had additional family members who have also attempted suicide. Additional microsatellite markers (two per candidate gene) were typed in these families and analysed using model-free methods. Analyses were carried out under three affection status models (ASMs). Primary single-locus analyses, in MERLIN, provided the best evidence for linkage to *TPH*1 and *BDNF*, under all three ASMs. The linkage evidence does not comply with the Lander and Kruglyak (1995) proposed guidelines for significant linkage. Their guidelines, however, are for genomic scans and not for candidate genes/regions. Based on prior probability implicating candidate regions in suicidal behaviour, the findings of the current study may therefore prove to be significant.

Excess IBD allele-sharing (NPL score) at *BDNF* was significant under ASM I and ASM II. The group of mood disorders and combined suicide attempts (ASM III) gave the same significant result for *BDNF* as for ASM I and II, suggesting that whatever the change is involved in *BDNF*, it has the same predisposing effect to suicide attempts and mood disorders. We detected evidence of linkage to *TPH*1 in persons with mood disorders (ASM II), but not in suicide attempters (ASM I). It therefore appears that *TPH*1 predisposes to mood disorders, but somehow confers protection against suicide attempts.

In order to increase the power to detect linkage to *DRD4*, *HTR2A* and *5-HTT*, two-locus interactions were carried out under ASM I and II. In the first instance linkage information at *TPH*1 was incorporated into a test for linkage and /or interaction with *DRD4*, *BDNF*, *HTR2A* and *5-HTT*, followed by incorporation of linkage information at *BDNF* into a test for linkage and /or interaction with *TPH*1, *DRD4*, *HTR2A* and *5-HTT*. Using family weights (Cox et al., 1999), none of the loci showed any evidence of interaction with

either *TPH*1 or *BDNF*, which could be ascribed to the choice of weights, i.e. none of the weights specified in our analyses were appropriate in explaining the underlying genetic model. However, two-locus modelling in TWOLOCARP (Cordell et al., 1995; Farrall, 1997, Cordell et al., 2000) allows more general interaction terms between the test locus and the conditioning locus. Modest evidence was observed in support of linkage to *HTR2A* conditioned on *TPH*1 and *5-HTT* conditioned on *TPH*1, in persons diagnosed with a mood disorder. However the most promising result from TWOLOCARP analyses was the improvement of the single-locus result for *TPH*1, by conditioning on *BDNF* (GEN model) and vice versa. The GEN model better explained the underlying genetic model, than the embedded MUL and HET models.

Sklar and co-workers (2002) identified *BDNF* as a potential risk locus in bipolar disorder in a large family based association study of 76 candidate genes. However, their study did not investigate the combined effects of loci. In the face of genetic heterogeneity, few studies have investigated the joint effects of candidate loci in suicidal behaviour and mood disorders. Jernej and co-workers (2004) reported an association between the combined intronic variants of *TPH*1 and *5-HTT* and suicidal behaviour. In an independent study, Serretti and co-workers (2004) reported evidence of interaction between *5-HTT* and *TPH*1 variants in response to selective serotonin re-uptake inhibitors in persons with mood disorders.

The statistical evidence obtained from linkage and two-locus interactions in the current study is upheld by biological relevance. In 1976, Asberg and colleagues provided the first indication that the serotonin neurotransmission system has a function in the pathogenesis of suicidal behaviour. It is the function and growth of serotonin neurons in the brain, i.e. serotonin system development that is promoted by BDNF (Mamounas et al., 1995). Interacting with trk B receptor tyrosine kinase (trk B), BDNF (one of the most important neurotrophins) is involved in cell survival and synaptic plasticity (Thoenen, 1995). Stress is reported to induce depression and decrease the expression of *BDNF* in the hippocampus (Smith et al., 1995), whereas antidepressants enhance BDNF mRNA levels in the brain via HTR2A and β-adrenoceptor subtypes, inhibiting stress -induced decreases in *BDNF* mRNA levels (Alleva and Santucci, 2001). Recently, Dwivedi and co-workers (2003) detected reduced expression of *BDNF* and trk B in post-mortem brain

in suicide subjects, thereby linking BDNF and trk B to the underlying pathophysiology of suicidal behaviour. Furthermore, infusion of BDNF in the adult rat brain has been shown to augment serotonin synthesis by directly enhancing *TPH*1 mRNA levels (Siuciak et al., 1998). *TPH*1 is the rate-limiting enzyme in the biosynthesis of serotonin (Jernej et al., 2004), which in turn binds to postsynaptic receptors, such as HTR2A which activates the downstream effector, phospholipase C (see Serretti and Artioli, 2004 for review). Following the release of serotonin into the synapse, levels of available serotonin and serotonin turnover are regulated by the serotonin transporter (5-HTT), which transports serotonin back to the serotonergic neurons (Du et al., 2001).

Based on the biological evidence, the findings of the present study suggest alterations in *BDNF* expression in both mood disorders and suicide attempts. In mood disorders the changes in *BDNF* expression leads to the attenuation of *TPH*1 expression. It is possible that this attenuation decreases serotonin synthesis and therefore impacts on the binding activity of HTR2A and the re-uptake of serotonin by the serotonin transporter. The exact mechanism underpinning the biological interactions between these candidate genes in mood disorders and suicidal behaviour is unclear. The recent discovery of *TPH*2, a new brain-specific *TPH* isoform is worthy of mention at this stage (Walther et al., 2003). *TPH*2, located on chromosome 12q21, may play a much more important role in serotonin synthesis in the brain than *TPH*1. This might explain why *TPH*1 has no effect on risk for suicide, but perhaps *TPH*2 does.

Future work includes replication of these findings in large data sets, in conjunction with mutation analyses of *BDNF*, *TPH*1 and *TPH*2, *HTR2A* and *5-HTT* in affected families to ascertain their impact on suicidal behaviour in affective mood disorders, and ultimately on public health. Of great interest are the upstream components of *BDNF*, which might influence altered expression of *BDNF* in mood disorders and suicidal behaviour. Antidepressant treatment activates intracellular signal transduction cascades, one of which is the cAMP-CREB cascade (reviewed by Duman et al., 2000). Chronic administration of antidepressants augments $G_s$-protein coupling to adenylate cyclase, particularly increasing levels of cAMP-dependent protein kinase A and CREB. Phosphorylation of CREB involves $Ca^{2+}$-dependent protein kinases, which in turn can be activated by the phosphatidylinositol pathway or by glutamate ionotropic receptors.

CREB, a transcription factor, induces *BDNF*, one gene target of the cAMP-CREB cascade and antidepressant treatment. Any one of the components of the cAMP-CREB cascade, preceding *BDNF* is a potential candidate gene in mood disorders and suicidal behaviour.

**Study Limitations**

*Association Study*. It is difficult to predict the sample size needed to illustrate a genetic association, as it relies on the degree of association, linkage disequilibrium, the accuracy of phenotypic data and heterogeneity of allele frequencies. Nevertheless, the limited number of cases analysed suggests that caution should be exercised with regard to interpreting our findings.

It remains possible that the 4-repeat allele may be in linkage disequilibrium with another functional variant that contributes to the phenotype, and therefore it is suggested that three or more polymorphisms, covering the gene, should be analysed instead of only one. The diagnostic heterogeneity of the patient group and the use of non-accurately matched controls are further limitations of the study.

Our investigation of the genetics of BPD and suicide attempts is ongoing, i.e. the search for and genotype analysis of more SNPs covering the genes of interest continues. In addition the extension of the size of the patient group requires a follow-up interview with previously recruited individuals to assess any change in their diagnosis including suicidal behaviour. This is currently underway as part of a study feedback and data updating process.

*Linkage study*. The small number of families included in linkage analysis warrants a cautious interpretation of the present findings. It is difficult to correct for testing under three affection status models, and using various analytical methods. Furthermore, we did not assess the temperamental variables of impulsivity and aggression among individuals with mood disorders, as these are also associated with suicidal acts.

In summary, the preliminary findings provided evidence for an association between *DRD4* of the dopaminergic pathway and suicide attempts in subjects diagnosed with an

affective mood disorders. This association was not supported by linkage analysis in nine extended families, suggesting that the initial result was either a false-positive or *DRD4* is in linkage disequilibrium with another locus. Linkage analysis provided statistical evidence in support of linkage to *TPH*1 in the group of mood disorders and not suicide attempts, suggesting a protective effect against suicidal behaviour. Alterations in *BDNF* appear to confer risk to mood disorders and suicidal behaviour.

Linkage analysis (family-based studies) has proven to be successful, given that the underlying genetic model was such that the assumptions made under the various analytical methods improved the power of detecting linkage and interaction. Although the sample size was small to begin with, the findings of this study are cause for optimism as biological evidence places *BDNF*, *TPH*1, *HTR2A* and *5-HTT* in the same biochemical pathway, thereby supporting evidence of interaction between them.

# CHAPTER 5

## GENERAL DISCUSSION

The aim of the present research was to identify susceptibility loci in BPD, by incorporating both classical and advanced linkage methodologies. Geneticists working on complex disorders need to abandon dogmatic uniformity in ideas, gambling our achievement on a single methodological approach. It is clear from what is already known in biology that there is considerable interaction between various gene products, which is equally true for those protein products of genes that influence complex disorders, and specifically BPD.

The present study derives from an archival resource for genetic studies in BPD, comprising of biological material and clinical information from 131 families. This resource is comparable to that of international large-scale formal BPD studies in the USA (Blehar et al., 1988), Canada (Shink et al., 2004), Austria (Bailer et al., 2002), United Kingdom-Ireland (Bennett et al., 2002; Lambert et al., 2005), Finland (Ekholm et al., 2002), and Germany (Cichon et al., 2001). However, unlike these huge international enterprises, research initiatives (in psychiatry) in South Africa do not enjoy high priority status in terms of the country's health care expenditure. Few of the eight academic departments of psychiatry in this country have a firm commitment to research, possibly due to the clinical service load. This status quo is exacerbated by few, mostly short-term and non-sustainable, national research fellowships. In spite of these challenges, investigations of the genetics of psychiatric disorders (the present research included) are positive developments. More recently, investigations have focussed on schizophrenia (Abecasis et al., 2004), obsessive-compulsive disorders (Niehaus et al., 2001) and Alzheimer's disease (Heckmann et al., 2004). In addition, there is a recent surge of cross disciplinary interest in developing sustainable research along specific "signature themes" here at the University of Cape Town. In this regard, the invitation by the research directorate to develop the theme of "Brains and Behaviour" bodes well for the BPD research. This project has been recognised as a cornerstone of the Brains and

Behaviour signature theme, highlighting the value of pathology and genetics as instruments in the neurosciences.

As illustrated in Chapter 2 of this thesis, genealogical studies of BPD reveal the phenotypic spectrum, which hinges on the classically characterised BPI or BPII syndromes. This heterogeneity introduces the layered complexity, ranging from the need to define clear phenotypes (and the difficulty of assigning affected/unaffected status to subjects), to the challenge of genetic mapping against the wide range of candidate loci already discussed.

Several genomic regions have been implicated in the aetiology of BPD. However, owing to issues of study design and methods of analysis, many studies have been marked by the limited success of reproducible findings. In an attempt to address these issues, the present study relied on a combination of statistical, computational and bioinformatics methods to improve the chances of identifying susceptibility loci in BPD, in a subset of 22 multiplex South African families. These families were chosen because of the level of confidence in phenotypic designation and the availability of the most number of affected individuals per pedigree. Non-parametric linkage analysis found a modest linkage signal for a susceptibility locus on chromosome 1q32. In light of the multifactorial nature of BPD, the analysis was extended to two-locus models, which implied interaction between regions at 1q32 and 10q23.32. It is important, however, to interpret these findings with caution until there is definitive biological proof of genetic variants influencing the phenotype.

Attempting to locate plausible candidate genes in the regions identified, bioinformatics revealed the candidate genes *NAV* and *ADORA1* in region 1q32, and *HTR7* and *PLCE1* in region 10q23.31. The only information on possible interactions assembled from Gene Ontology was that ADORA1, HTR7 and PLCE1 all have receptor activity and are integral to the plasma membrane. Mutation analysis of these candidate genes and their association with illness are suggested as the next phase of research. Individually, the chromosomal regions 1q32 and 10q have been implicated in BPD (Detera-Wadleigh et al., 1999; Cichon et al., 2001b), however, no other study has reported evidence for statistical interaction (or oth erwise) between susceptibility loci on 1q32 and 10q. The

strategy adopted in the present study is based on the convergent functional genomics approach of Niculescu and co-workers (2000). Their approach, integrating gene expression data from a mouse model with human linkage data, cross validates findings, thereby prioritising candidate genes for individual analysis. In 2004, Ogden and co-workers presented an expanded convergent functional genomics approach, integrating many more independent converging lines of evidence, as a way of dissecting the genetic code of bipolar mood disorders. More recently and much like the strategy adopted in this study, Hattori et al. (2005), describes a "systems genetics" approach, requiring an intensive literature review and mining of electronic databases for entire pathways, in the compilation of candidate genes. Mutation analysis of the candidate genes identified by Niculescu, Ogden, and Hattori, and the modelling of gene-gene interactions between the same candidate genes, might forge optimal strategies for the mapping and identification of susceptibility genes in BPD (as well as other complex disorders).

An intrinsic limitation in the genetic analysis of BPD is the variable phenotype, which all form part of a spectrum of disorders, with variable genetic involvement. In order to address this limitation, we used the parameter of suicide attempts in an effort to enrich for what might be classified as a "sub-phenotype" of BPD, thereby enriching for a smaller subset of candidate genes. Findings from the "pilot" association study (46 cases vs 160 controls) showed decreased frequencies in suicide attempters (from the Caucasian subpopulation) of the 4-repeat allele, of *DRD4*, suggesting that the 4-repeat allele may confer a possible 'protection' against susceptibility to development of an affective disorder and/or attempted suicide. Because nine of the 46 cases were members of families, these were included in linkage analysis. The outcome of single-locus analyses was in support of linkage between suicide attempts (ASM I) and BDNF, and therefore it seems that *BDNF* might be a susceptibility locus in suicide attempts. Two-locus modelling showed no evidence of interaction between BDNF and other loci under ASM I. It is worth noting that under ASM II, which includes individuals with mood disorders, two-locus models showed evidence of statistical interaction between *BDNF* and *TPH1*, *HTR2A* and *TPH1*, and *5-HTT* and *TPH1*, in the broader spectrum of mood disorders. The data suggest that *BDNF* predisposes vulnerable individuals to both mood disorders and suicide attempts, whereas *TPH1* predisposes to mood disorders but

not to the risk of suicide. Perhaps this suggests a protective role for *TPH1* against suicidal behaviour in this group of mood disorder patients. The positive linkage findings of *BDNF* (Dwivedi et al., 2003), *TPH1* (Siuciak et al., 1998), *HTR2A* (reviewed in Serretti and Artioli, 2004) and *5-HTT* (Du et al., 2001), individually, have been implicated in the pathophysiology of suicidal behaviour, by other authors. However, no two-locus interactions have been reported before. Importantly, biological context places each of these candidates in the same biochemical pathway, lending credence to possible interactions between these candidate genes.

In each of the studies carried out, the multiple-testing cost of modelling gene-gene interactions is a major issue. To limit this cost, loci that achieved at least marginal significance in the single locus analyses were used, in all subsequent two-locus models. In addition, permutations and simulations were used to estimate empirical significance of individual (pair-wise) results. Of the genetic mechanisms commonly known to contribute to the complex patterns of inheritance, attention was focussed on epistatic effects only, anticipation, genetic imprinting and mitochondrial mutations were not considered. As far as addressing variability of phenotype is concerned, a refined sub-phenotype of suicide attempts were investigated, however, endophenotypes, which may have improved the chances of detecting linkage and association, were not examined. While the need for accuracy of demographic and phenotypic data is recognised in sample size and study design in the type of research reported here, one should not understate the value of classical genetic analysis, taking into account IBD information reflecting shared genomic regions, at least within pedigrees.

At present the study of geographically and genetically isolated populations provides the best chance of mapping susceptibility loci in complex disorders. Examples of such communities that have grown from a restricted gene pool with fewer segregating genes that predispose to BPD, include the North American Old Order Amish (Ginns et al., 1996), isolated populations of the Faroe Islands (Als et al., 2004), and the Saguenay-Lac-St-Jean area of Quebec in Canada (Morissette et al., 1999). The anthropological history of South Africa provides researchers with valuable 'genetic isolates', such as the European-derived Afrikaner and indigenous African Xhosa populations. These relatively homogeneous gene pools have been investigated in psychiatric genetics, for example

Afrikaners with schizophrenia (Abecasis et al., 2004); Afrikaners with obsessive-compulsive disorder (Hemmings et al., 2004), and a large Xhosa pedigree with early-onset Alzheimer's disease (Heckman et al., 2004). Given the integrative approach and the evidence of statistical interactions of this current study, the next suggested step in contributing to our understanding is to extend genealogies in stratified populations, for example the Afrikaner and the Nguni-speaking indigenous African populations.

Importantly, the genetic research presented here has already created a heightened awareness of the role of genetics in psychiatric disorders and emphasizes the need to maintain and grow large scale research enterprises in BPD, in South Africa. There is also greater understanding of the need to probe genealogical information on bipolar-associated disorders. Although the research in South Africa lags behind the major international centres with established BPD programmes, the attempts to engage with psychiatrists to develop awareness of genetics and biological heterogeneity in psychiatry is gaining favour. The expanding research programme will create the very necessary link between international (and even local) research findings and translation for local clinical consumption

Future work suggested in this study includes the validation of the findings of the association and linkage studies reported here, such as genetic testing on a larger scale, in an increased number of families and candidate gene screening for mutations. Furthermore, systematic screening methods need to be put in place, in order to update changes in psychiatric status (on a yearly basis) for families already in the BPD archive. There is no question that the vast international efforts aimed at the genetic dissection of BPD has already benefited psychiatry, even though the inventory of predisposing genes have not been formed. The demand for phenotypically "homogeneous" samples has led to a sobering of the approach to collect vast amounts of biological material, and the current project will be extended in this regard.

In conclusion, we set out to connect the dots between interacting genes using a multi-faceted approach, bringing together an exploration of a phenotypic spectrum (including suicidality) with statistical analyses, bioinformatics, and data on the relevant biochemical and neurophysiological systems. Strategies such as this, building on past and

incorporating present methodologies, are expected to improve diagnosis, prevention, and treatment of BPD, and indeed can be adopted for other complex disorders.

# REFERENCES AND APPENDICES

# REFERENCES


Abecasis GR, Burt RA, Hall D, Bochum S, Doheny KF, Lundy SL, Torrington M, Roos JL, Gogos JA, Karayiorgou M (2004) Genomewide scan in families with schizophrenia from the founder population of Afrikaners reveals evidence for linkage and uniparental disomy on chromosome 1. Am J Hum Genet 74:403-417

Abecasis GR, Cherny SS, Cookson WO, Cardon LR (2002) Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet 30:97-101

Adams LJ, Mitchell PB, Fielder SL, Rosso A, Donald JA, Schofield PR (1998) A susceptibility locus for bipolar affective disorder on chromosome 4q35. Am J Hum Genet 62:1084-1091

Adolphs R, Tranel D, Damasio H, Damasio A (1994) Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. Nature 372:669-672

Aita VM, Liu J, Knowles JA, Terwilliger JD, Baltazar R, Grunn A, Loth JE, Kanyas K, Lerer B, Endicott J, Wang Z, Penchaszadeh G, Gilliam TC, Baron M (1999) A comprehensive linkage analysis of chromosome 21q22 supports prior evidence for a putative bipolar affective disorder locus. Am J Hum Genet 64:210-217

Akiskal HS (2003) Validating 'hard' and 'soft' phenotypes within the bipolar spectrum: continuity or discontinuity? J Affect Disord 73:1-5

Akiskal HS, Benazzi F (2003) Family history validation of the bipolar nature of depressive mixed states. J Affect Disord 73:113-122

Alleva E, Santucci D (2001) Psychosocial vs. "physical" stress situations in rodents and humans: role of neurotrophins. Physiol Behav 73:313-320

Almasy L, Blangero J (1998) Multipoint quantitative-trait linkage analysis in general pedigrees. Am J Hum Genet 62:1198-1211

Als TD, Dahl HA, Flint TJ, Wang AG, Vang M, Mors O, Kruse TA, Ewald H (2004) Possible evidence for a common risk locus for bipolar affective disorder and schizophrenia on chromosome 4p16 in patients from the Faroe Islands. Mol Psychiatry 9:93-98

Ambelas A (1979) Psychologically stressful events in the precipitation of manic episodes. Br J Psychiatry 135:15-21

Ambelas A (1987) Life events and mania. A special relationship? Br J Psychiatry 150:235-240

Anderson AK, Phelps EA (2001) Lesions of the human amygdala impair enhanced perception of emotionally salient events. Nature 411:305-309

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

Angst J. 2002. Historical aspects of the dichotomy between manic-depressive disorders and schizophrenia. Schizophrenia Research 1;57(1):5-13.

Apter A, Gothelf D, Orbach I, Weizman R, Ratzoni G, Har-Even D, Tyano S (1995) Correlation of suicidal and violent behavior in different diagnostic categories in hospitalized adolescent patients. J Am Acad Child Adolesc Psychiatry 34:912-918

Benedetti F, Bernasconi A, Lorenzi C, Pontiggia A, Serretti A, Colombo C, Smeraldi E (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 355:37-40

Benjamin J, Li L, Patterson C, Greenberg BD, Murphy DL, Hamer DH (1996) Population and familial association between the D4 dopamine receptor gene and measures of Novelty Seeking. Nat Genet 12:81-84

Bennett P, Segurado R, Jones I, Bort S, McCandless F, Lambert D, Heron J, Comerford C, Middle F, Corvin A, Pelios G, Kirov G, Larsen B, Mulcahy T, Williams N, O'Connell R, O'Mahony E, Payne A, Owen M, Holmans P, Craddock N, Gill M (2002) The Wellcome trust UK-Irish bipolar affective disorder sibling-pair genome screen: first stage report. Mol Psychiatry 7:189-200

Berrettini WH (1999) On the future of genetic research in bipolar and schizophrenic syndromes. Neuropsychopharmacology 21:1-2

Berrettini WH (2000) Are schizophrenic and bipolar disorders related? A review of family and molecular studies. Biol Psychiatry 48:531-538

Berrettini WH, Ferraro TN, Goldin LR, Weeks DE, Detera-Wadleigh S, Nurnberger JI, Jr., Gershon ES (1994) Chromosome 18 DNA markers and manic-depressive illness: evidence for a susceptibility gene. Proc Natl Acad Sci U S A 91:5918-5921

Berridge MJ, Downes CP, Hanley MR (1989) Neural and developmental actions of lithium: a unifying hypothesis. Cell 59:411-419

Bezchlibnyk Y, Young LT (2002) The neurobiology of bipolar disorder: focus on signal transduction pathways and the regulation of gene expression. Can J Psychiatry 47:135-148

Bidzinska EJ (1984) Stress factors in affective diseases. Br J Psychiatry 144:161-166

Blackwood DH, Fordyce A, Walker MT, St Clair DM, Porteous DJ, Muir WJ (2001) Schizophrenia and affective disorders--cosegregation with a translocation at chromosome 1q42 that directly disrupts brain-expressed genes: clinical and P300 findings in a family. Am J Hum Genet 69:428-433

Blackwood DH, He L, Morris SW, McLean A, Whitton C, Thomson M, Walker MT, Woodburn K, Sharp CM, Wright AF, Shibasaki Y, St Clair DM, Porteous DJ, Muir WJ (1996) A locus for bipolar affective disorder on chromosome 4p. Nat Genet 12:427-430

Blangero J, Williams JT, Almasy L. 2001. Variance component methods for detecting complex trait loci. Adv Genet. 2001;42:151-81. Review.

Blehar MC, Weissman MM, Gershon ES, Hirschfeld RM (1988) Family and genetic studies of affective disorders. Arch Gen Psychiatry 45:289-292

Blouin JL, Dombroski BA, Nath SK, Lasseter VK, Wolyniec PS, Nestadt G, Thornquist M, et al. (1998) Schizophrenia susceptibility loci on chromosomes 13q32 and 8p21. Nat Genet 20:70-73

Bobb AJ, Castellanos FX, Addington AM, Rapoport JL (2004) Molecular genetic studies of ADHD: 1991 to 2004. Am J Med Genet

Boehnke M (1986) Estimating the power of a proposed linkage study: a practical computer simulation approach. Am J Hum Genet 39:513-527

Bohringer S, Hardt C, Miterski B, Steland A, Epplen JT (2003) Multilocus statistics to uncover epistasis and heterogeneity in complex diseases: revisiting a set of multiple sclerosis data. Eur J Hum Genet 11:573-584

Bosetti F, Seemann R, Bell JM, Zahorchak R, Friedman E, Rapoport SI, Manickam P (2002) Analysis of

Clark LN, Levy G, Tang MX, Mejia-Santana H, Ciappa A, Tycko B, Cote LJ, Louis ED, Mayeux R, Marder K (2003) The Saitohin 'Q7R' polymorphism and tau haplotype in multi-ethnic Alzheimer disease and Parkinson's disease cohorts. Neurosci Lett 347:17-20

Cloninger CR, Bayon C, Svrakic DM (1998) Measurement of temperament and character in mood disorders: a model of fundamental states as personality types. J Affect Disord 51:21-32

Cloninger CR, Svrakic DM, Przybeck TR (1993) A psychobiological model of temperament and character. Arch Gen Psychiatry 50:975-990

Collins FS, Patrinos A, Jordan E, Chakravarti A, Gesteland R, Walters L (1998) New goals for the U.S. Human Genome Project: 1998-2003. Science 282:682-689

Compton MT, Nemeroff CB (2000) The treatment of bipolar depression. J Clin Psychiatry 61 Suppl 9:57-67

Concannon P, Gogolin-Ewens KJ, Hinds DA, Wapelhorst B, Morrison VA, Stirling B, Mitra M, Farmer J, Williams SR, Cox NJ, Bell GI, Risch N, Spielman RS (1998) A second-generation screen of the human genome for susceptibility to insulin-dependent diabetes mellitus. Nat Genet 19:292-296

Cook EH, Jr., Courchesne RY, Cox NJ, Lord C, Gonen D, Guter SJ, Lincoln A, Nix K, Haas R, Leventhal BL, Courchesne E (1998) Linkage-disequilibrium mapping of autistic disorder, with 15q11-13 markers. Am J Hum Genet 62:1077-1083

Coon H, Jensen S, Hoff M, Holik J, Plaetke R, Reimherr F, Wender P, Leppert M, Byerley W (1993) A genome-wide search for genes predisposing to manic-depression, assuming autosomal dominant inheritance. Am J Hum Genet 52:1234-1249

Cordell HJ, Todd JA, Bennett ST, Kawaguchi Y, Farrall M (1995) Two-locus maximum lod score analysis of a multifactorial trait: joint consideration of IDDM2 and IDDM4 with IDDM1 in type 1 diabetes. Am J Hum Genet 57:920-934

Cordell HJ, Wedig GC, Jacobs KB, Elston RC (2000) Multilocus linkage tests based on affected relative pairs. Am J Hum Genet 66:1273-1286

Coryell W, Endicott J, Reich T, Andreasen N, Keller M (1984) A family study of bipolar II disorder. Br J Psychiatry 145:49-54

Cottingham RW, Jr., Idury RM, Schaffer AA (1993) Faster sequential genetic linkage computations. Am J Hum Genet 53:252-263

Couzin J (2002) Genomics. New mapping project splits the community. Science 296:1391-1393

Cox NJ (2001) Computational issues in mapping variation affecting susceptibility to complex disorders: the chicken and the egg. Theor Popul Biol 60:221-225

Cox NJ, Frigge M, Nicolae DL, Concannon P, Hanis CL, Bell GI, Kong A (1999) Loci on chromosomes 2 (NIDDM1) and 15 interact to increase susceptibility to diabetes in Mexican Americans. Nat Genet 21:213-215

Craddock N, Jones I (1999) Genetics of bipolar disorder. J Med Genet 36:585-594

Craddock N, Khodel V, Van Eerdewegh P, Reich T (1995) Mathematical limits of multilocus models: the genetic transmission of bipolar disorder. Am J Hum Genet 57:690-702

Craddock N, Lendon C (1999) Chromosome Workshop: chromosomes 11, 14, and 15. Am J Med Genet 88:244-254

Craddock N, McGuffin P (1993) Approaches to the genetics of affective disorders. Ann Med 25:317-322

American Psychiatric Association., 1994.

Dick DM, Foroud T, Flury L, Bowman ES, Miller MJ, Rau NL, Moe PR, et al. (2003) Genomewide linkage analyses of bipolar disorder: a new sample of 250 pedigrees from the National Institute of Mental Health Genetics Initiative. Am J Hum Genet 73:107-114

Ding YC, Chi HC, Grady DL, Morishima A, Kidd JR, Kidd KK, Flodman P, Spence MA, Schuck S, Swanson JM, Zhang YP, Moyzis RK (2002) Evidence of positive selection acting at the human dopamine receptor D4 gene locus. Proc Natl Acad Sci U S A 99:309-314

Dong C, Li WD, Li D, Price RA (2005) Interaction between obesity-susceptibility loci in chromosome regions 2p25-p24 and 13q13-q21. Eur J Hum Genet 13:102-108

Drevets WC, Gautier C, Price JC, Kupfer DJ, Kinahan PE, Grace AA, Price JL, Mathis CA (2001) Amphetamine-induced dopamine release in human ventral striatum correlates with euphoria. Biol Psychiatry 49:81-96

Drevets WC, Ongur D, Price JL (1998) Neuroimaging abnormalities in the subgenual prefrontal cortex: implications for the pathophysiology of familial mood disorders. Mol Psychiatry 3:220-226, 190-221

Drevets WC, Price JL, Simpson JR, Jr., Todd RD, Reich T, Vannier M, Raichle ME (1997) Subgenual prefrontal cortex abnormalities in mood disorders. Nature 386:824-827

Du L, Bakish D, Lapierre YD, Ravindran AV, Hrdina PD (2000) Association of polymorphism of serotonin 2A receptor gene with suicidal ideation in major depressive disorder. Am J Med Genet 96:56-60

Du L, Faludi G, Palkovits M, Bakish D, Hrdina PD (2001) Serotonergic genes and suicidality. Crisis 22:54-60

Du L, Faludi G, Palkovits M, Demeter E, Bakish D, Lapierre YD, Sotonyi P, Hrdina PD (1999) Frequency of long allele in serotonin transporter gene is increased in depressed suicide victims. Biol Psychiatry 46:196-201

Duman RS (2002) Pathophysiology of depression: the concept of synaptic plasticity. Eur Psychiatry 17 Suppl 3:306-310

Duman RS, Heninger GR, Nestler EJ (1997) A molecular and cellular theory of depression. Arch Gen Psychiatry 54:597-606

Duman RS, Malberg J, Nakagawa S, D'Sa C (2000) Neuronal plasticity and survival in mood disorders. Biol Psychiatry 48:732-739

Duman RS, Malberg J, Thome J (1999) Neural plasticity to stress and antidepressant treatment. Biol Psychiatry 46:1181-1191

Dwivedi Y, Rizavi HS, Conley RR, Roberts RC, Tamminga CA, Pandey GN (2003) Altered gene expression of brain-derived neurotrophic factor and receptor tyrosine kinase B in postmortem brain of suicide subjects. Arch Gen Psychiatry 60:804-815

Eaves LJ (1994) Effect of genetic architecture on the power of human linkage studies to resolve the contribution of quantitative trait loci. Heredity 72 ( Pt 2):175-192

Ebstein RP, Novick O, Umansky R, Priel B, Osher Y, Blaine D, Bennett ER, Nemanov L, Katz M, Belmaker RH (1996) Dopamine D4 receptor (D4DR) exon III polymorphism associated with the human personality trait of Novelty Seeking. Nat Genet 12:78-80

Edenberg HJ, Foroud T, Conneally PM, Sorbel JJ, Carr K, Crose C, Willig C, Zhao J, Miller M, Bowman E, Mayeda A, Rau NL, Smiley C, Rice JP, Goate A, Reich T, Stine OC, McMahon F, DePaulo JR,

Freis ED (1954) Mental depression in hypertensive patients treated for long periods with large doses of reserpine. N Engl J Med 251:1006-1008

Furlong RA, Ho L, Rubinsztein JS, Walsh C, Paykel ES, Rubinsztein DC (1999) Analysis of the monoamine oxidase A (MAOA) gene in bipolar affective disorder by association studies, meta-analyses, and sequencing of the promoter. Am J Med Genet 88:398-406

Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ, Altshuler D (2002) The structure of haplotype blocks in the human genome. Science 296:2225-2229

Gershon ES, Badner JA, Detera-Wadleigh SD, Ferraro TN, Berrettini WH (1996) Maternal inheritance and chromosome 18 allele sharing in unilineal bipolar illness pedigrees. Am J Med Genet 67:202-207

Gershon ES, Hamovit J, Guroff JJ, Dibble E, Leckman JF, Sceery W, Targum SD, Nurnberger JI, Jr., Goldin LR, Bunney WE, Jr. (1982) A family study of schizoaffective, bipolar I, bipolar II, unipolar, and normal control probands. Arch Gen Psychiatry 39:1157-1167

Ghaemi SN, Miller CJ, Berv DA, Klugman J, Rosenquist KJ, Pies RW. 2003. Sensitivity and specificity of a new bipolar spectrum diagnostic scale. Journal of Affective Disorders. In Press, Corrected Proof, Available online 19 September 2003.

Ghosh S, Collins FS (1996) The geneticist's approach to complex disease. Annu Rev Med 47:333-353

Ginns EI, Ott J, Egeland JA, Allen CR, Fann CS, Pauls DL, Weissenbachoff J, Carulli JP, Falls KM, Keith TP, Paul SM (1996) A genome-wide search for chromosomal loci linked to bipolar affective disorder in the Old Order Amish. Nat Genet 12:431-435

Giros B, Jaber M, Jones SR, Wightman RM, Caron MG (1996) Hyperlocomotion and indifference to cocaine and amphetamine in mice lacking the dopamine transporter. Nature 379:606-612

Glassner B, Haldipur CV (1983) Life events and early and late onset of bipolar disorder. Am J Psychiatry 140:215-217

Glatt CE, Freimer NB (2002) Association analysis of candidate genes for neuropsychiatric disease: the perpetual campaign. Trends Genet 18:307-312

Goldberg JF, Garno JL, Leon AC, Kocsis JH, Portera L (1999) A history of substance abuse complicates remission from acute mania in bipolar disorder. J Clin Psychiatry 60:733-740

Goodwin FK, Murphy DL, Dunner DL, Bunney WE, Jr. (1972) Lithium response in unipolar versus bipolar depression. Am J Psychiatry 129:44-47

Gould TD, Manji HK (2002) Signaling networks in the pathophysiology and treatment of mood disorders. J Psychosom Res 53:687-697

Grandy DK, Allen LJ, Zhang Y, Magenis RE, Civelli O (1992) Chromosomal localization of three human D5 dopamine receptor genes. Genomics 13:968-973

Greenwood TA, Alexander M, Keck PE, McElroy S, Sadovnick AD, Remick RA, Kelsoe JR (2001) Evidence for linkage disequilibrium between the dopamine transporter and bipolar disorder. Am J Med Genet 105:145-151

Greil W, Kleindienst N (2003) Concepts in the treatment of bipolar disorder. Acta Psychiatr Scand Suppl:41-46

Grunebaum MF. 2001. Outcomes of suicidal behaviours. Clinical Neuroscience Research 1: 394-404.

Hill WG, Weir BS (1994) Maximum-likelihood estimation of gene location by linkage disequilibrium. Am J Hum Genet 54:705-714

Hindmarch I (2002) Beyond the monoamine hypothesis: mechanisms, molecules and methods. Eur Psychiatry 17 Suppl 3:294-299

Hiroi N, Martin AB, Grande C, Alberti I, Rivera A, Moratalla R (2002) Molecular dissection of dopamine receptor signaling. J Chem Neuroanat 23:237-242

Hirschfeld RM, Williams JB, Spitzer RL, Calabrese JR, Flynn L, Keck PE, Jr., Lewis L, McElroy SL, Post RM, Rapport DJ, Russell JM, Sachs GS, Zajecka J (2000) Development and validation of a screening instrument for bipolar spectrum disorder: the Mood Disorder Questionnaire. Am J Psychiatry 157:1873-1875

Hodgkinson S, Sherrington R, Gurling H, Marchbanks R, Reeders S, Mallet J, McInnis M, Petursson H, Brynjolfsson J (1987) Molecular genetic evidence for heterogeneity in manic depression. Nature 325:805-806

Hoehe MR, Wendel B, Grunewald I, Chiaroni P, Levy N, Morris-Rosendahl D, Macher JP, Sander T, Crocq MA (1998) Serotonin transporter (5-HTT) gene polymorphisms are not associated with susceptibility to mood disorders. Am J Med Genet 81:1-3

Holmans P (1993) Asymptotic properties of affected-sib-pair linkage analysis. Am J Hum Genet 52:362-374

Holmans P, Craddock N (1997) Efficient strategies for genome scanning using maximum-likelihood affected-sib-pair analysis. Am J Hum Genet 60:657-666

Hwu HG, Yeh EK, Chang LY (1989) Prevalence of psychiatric disorders in Taiwan defined by the Chinese Diagnostic Interview Schedule. Acta Psychiatr Scand 79:136-147

Inoue K, Lupski JR (2003) Genetics and genomics of behavioral and psychiatric disorders. Curr Opin Genet Dev 13:303-309

Isometsa E, Henriksson M, Marttunen M, Heikkinen M, Aro H, Kuoppasalmi K, Lonnqvist J (1995) Mental disorders in young and middle aged men who commit suicide. Bmj 310:1366-1367

Isometsa E, Suominen K, Mantere O, Valtonen H, Leppamaki S, Pippingskold M, Arvilommi P (2003) The mood disorder questionnaire improves recognition of bipolar disorder in psychiatric care. BMC Psychiatry 3:8

Itokawa M, Yamada K, Iwayama-Shigeno Y, Ishitsuka Y, Detera-Wadleigh S, Yoshikawa T (2003) Genetic analysis of a functional GRIN2A promoter (GT)n repeat in bipolar disorder pedigrees in humans. Neurosci Lett 345:53-56

Iwamoto K, Kakiuchi C, Bundo M, Ikeda K, Kato T (2004) Molecular characterization of bipolar disorder by comparing gene expression profiles of postmortem brains of major mental disorders. Mol Psychiatry 9:406-416

Jamison KR (2000) Suicide and bipolar disorder. J Clin Psychiatry 61 Suppl 9:47-51

Jernej B, Stefulj J, Hranilovic D, Balija M, Skavic J, Kubat M (2004) Intronic polymorphism of tryptophan hydroxylase and serotonin transporter: indication for combined effect in predisposition to suicide. J Neural Transm 111:733-738

Johnson GC, Esposito L, Barratt BJ, Smith AN, Heward J, Di Genova G, Ueda H, Cordell HJ, Eaves IA, Dudbridge F, Twells RC, Payne F, Hughes W, Nutland S, Stevens H, Carr P, Tuomilehto-Wolf E, Tuomilehto J, Gough SC, Clayton DG, Todd JA (2001) Haplotype tagging for the identification of common disease genes. Nat Genet 29:233-237

H (2001) A genome survey indicates a possible susceptibility locus for bipolar disorder on chromosome 22. Proc Natl Acad Sci U S A 98:585-590

Kennedy JL, Basile VS, Macciardi FM (1999) Chromosome 4 Workshop Summary: Sixth World Congress on Psychiatric Genetics, Bonn, Germany, October 6-10, 1998. Am J Med Genet 88:224-228

Kerem B, Rommens JM, Buchanan JA, Markiewicz D, Cox TK, Chakravarti A, Buchwald M, Tsui LC (1989) Identification of the cystic fibrosis gene: genetic analysis. Science 245:1073-1080

Kessler RC, McGonagle KA, Swartz M, Blazer DG, Nelson CB (1993) Sex and depression in the National Comorbidity Survey. I: Lifetime prevalence, chronicity and recurrence. J Affect Disord 29:85-96

Klein PS, Melton DA (1996) A molecular mechanism for the effect of lithium on development. Proc Natl Acad Sci U S A 93:8455-8459

Kluger AN, Siegfried Z, Ebstein RP (2002) A meta-analysis of the association between DRD4 polymorphism and novelty seeking. Mol Psychiatry 7:712-717

Kong A, Cox NJ (1997) Allele-sharing models: LOD scores and accurate linkage tests. Am J Hum Genet 61:1179-1188

Kruglyak L (1999) Prospects for whole-genome linkage disequilibrium mapping of common disease genes. Nat Genet 22:139-144

Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES (1996) Parametric and nonparametric linkage analysis: a unified multipoint approach. Am J Hum Genet 58:1347-1363

Kruglyak L, Lander ES (1995) High-resolution genetic mapping of complex traits. Am J Hum Genet 56:1212-1223

Kunkel LM (1986) Analysis of deletions in DNA from patients with Becker and Duchenne muscular dystrophy. Nature 322:73-77

Lachman HM, Kelsoe JR, Remick RA, Sadovnick AD, Rapaport MH, Lin M, Pazur BA, Roe AM, Saito T, Papolos DF (1997) Linkage studies suggest a possible locus for bipolar disorder near the velo-cardio-facial syndrome region on chromosome 22. Am J Med Genet 74:121-128

Lambert D, Middle F, Hamshere ML, Segurado R, Raybould R, Corvin A, Green E, O'Mahony E, Nikolov I, Mulcahy T, Haque S, Bort S, Bennett P, Norton N, Owen MJ, Kirov G, Lendon C, Jones L, Jones I, Holmans P, Gill M, Craddock N (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

Lander E, Kruglyak L (1995) Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. Nat Genet 11:241-247

Lander ES, Green P (1987) Construction of multilocus genetic linkage maps in humans. Proc Natl Acad Sci U S A 84:2363-2367

Lander ES, Schork NJ (1994) Genetic dissection of complex traits. Science 265:2037-2048

Lash AE, Tolstoshev CM, Wagner L, Schuler GD, Strausberg RL, Riggins GJ, Altschul SF (2000) SAGEmap: a public gene expression resource. Genome Res 10:1051-1060

Lathrop GM, Lalouel JM, Julier C, Ott J (1984) Strategies for multilocus linkage analysis in humans. Proc Natl Acad Sci U S A 81:3443-3446

Law AJ, Weickert CS, Hyde TM, Kleinman JE, Harrison PJ (2004) Reduced spinophilin but not microtubule-associated protein 2 expression in the hippocampal formation in schizophrenia and

MacKinnon DF, Jamison KR, DePaulo JR (1997) Genetics of manic depressive illness. Annu Rev Neurosci 20:355-373

Maes T, Barcelo A, Buesa C (2002) Neuron navigator: a human gene family with homology to unc-53, a cell guidance gene from Caenorhabditis elegans. Genomics 80:21-30

Mamounas LA, Blue ME, Siuciak JA, Altar CA (1995) Brain-derived neurotrophic factor promotes the survival and sprouting of serotonergic axons in rat brain. J Neurosci 15:7929-7939

Manji HK, Bebchuk JM, Moore GJ, Glitz D, Hasanat KA, Chen G (1999) Modulation of CNS signal transduction pathways and gene expression by mood-stabilizing agents: therapeutic implications. J Clin Psychiatry 60 Suppl 2:27-39; discussion 40-21, 113-116

Manji HK, Chen G, Shimon H, Hsiao JK, Potter WZ, Belmaker RH (1995) Guanine nucleotide-binding proteins in bipolar affective disorder. Effects of long-term lithium treatment. Arch Gen Psychiatry 52:135-144

Manji HK, Drevets WC, Charney DS (2001) The cellular neurobiology of depression. Nat Med 7:541-547

Manji HK, Moore GJ, Rajkowska G, Chen G (2000) Neuroplasticity and cellular resilience in mood disorders. Mol Psychiatry 5:578-593

Mann JJ (1998) The neurobiology of suicide. Nat Med 4:25-30

Mann JJ (2002) A current perspective of suicide and attempted suicide. Ann Intern Med 136:302-311

Mann JJ, Waternaux C, Haas GL, Malone KM (1999) Toward a clinical model of suicidal behavior in psychiatric patients. Am J Psychiatry 156:181-189

Maris RW (2002) Suicide. Lancet 360:319-326

Martin ER, Monks SA, Warren LL, Kaplan NL (2000) A test for linkage and association in general pedigrees: the pedigree disequilibrium test. Am J Hum Genet 67:146-154

Matsushita S, Kimura M, Miyakawa T, Yoshino A, Murayama M, Masaki T, Higuchi S (2004) Association study of brain-derived neurotrophic factor gene polymorphism and alcoholism. Alcohol Clin Exp Res 28:1609-1612

Matsushita S, Yoshino A, Murayama M, Kimura M, Muramatsu T, Higuchi S (2001) Association study of serotonin transporter gene regulatory region polymorphism and alcoholism. Am J Med Genet 105:446-450

Maziade M, Roy MA, Chagnon YC, Cliche D, Fournier JP, Montgrain N, Dion C, Lavallee JC, Garneau Y, Gingras N, Nicole L, Pires A, Ponton AM, Potvin A, Wallot H, Merette C (2004) Shared and specific susceptibility loci for schizophrenia and bipolar disorder: a dense genome scan in Eastern Quebec families. Mol Psychiatry

Maziade M, Roy MA, Rouillard E, Bissonnette L, Fournier JP, Roy A, Garneau Y, Montgrain N, Potvin A, Cliche D, Dion C, Wallot H, Fournier A, Nicole L, Lavallee JC, Merette C (2001) A search for specific and common susceptibility loci for schizophrenia and bipolar disorder: a linkage study in 13 target chromosomes. Mol Psychiatry 6:684-693

McCarthy A (2001) Pharmacogenetics. Bmj 322:1007-1008

McDonald C, Bullmore ET, Sham PC, Chitnis X, Wickham H, Bramon E, Murray RM (2004) Association of genetic risks for schizophrenia and bipolar disorder with specific and generic brain structural endophenotypes. Arch Gen Psychiatry 61:974-984

Miklowitz DJ, Goldstein MJ, Nuechterlein KH, Snyder KS, Mintz J (1988) Family factors and the course of bipolar affective disorder. Arch Gen Psychiatry 45:225-231

Minov C, Baghai TC, Schule C, Zwanzger P, Schwarz MJ, Zill P, Rupprecht R, Bondy B (2001) Serotonin-2A-receptor and -transporter polymorphisms: lack of association in patients with major depression. Neurosci Lett 303:119-122

Mitchell P, Mackinnon A, Waters B. 1993. The genetics of bipolar disorder.
Austr NZ J Psych. 1993 Dec;27(4):560-80. Review.

Monaco AP, Bertelson CJ, Middlesworth W, Colletti CA, Aldridge J, Fischbeck KH, Bartlett R, Pericak-Vance MA, Roses AD, Kunkel LM (1985) Detection of deletions spanning the Duchenne muscular dystrophy locus using a tightly linked DNA segment. Nature 316:842-845

Morissette J, Villeneuve A, Bordeleau L, Rochette D, Laberge C, Gagne B, Laprise C, Bouchard G, Plante M, Gobeil L, Shink E, Weissenbach J, Barden N (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 88:567-587

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

Morton NE (1998) Significance levels in complex inheritance. Am J Hum Genet 62:690-697

Morton NE, Collins A (1998) Tests and estimates of allelic association in complex inheritance. Proc Natl Acad Sci U S A 95:11389-11393

Mukhopadhyay N, Almasy L, Schroeder M, Mulvihill WP, Weeks DE. 1999.
Mega2, a data-handling program for facilitating genetic linkage and association analyses. Am J Hum Genet 65:A436

Muller-Oerlinghausen B, Berghofer A, Bauer M (2002) Bipolar disorder. Lancet 359:241-247

Mundo E, Tharmalingham S, Neves-Pereira M, Dalton EJ, Macciardi F, Parikh SV, Bolonna A, Kerwin RW, Arranz MJ, Makoff AJ, Kennedy JL (2003) Evidence that the N-methyl-D-aspartate subunit 1 receptor gene (GRIN1) confers susceptibility to bipolar disorder. Mol Psychiatry 8:241-245

Murray CJ, Lopez AD (1997) Global mortality, disability, and the contribution of risk factors: Global Burden of Disease Study. Lancet 349:1436-1442

Myrick H, Brady KT (2003) The use of divalproex in the treatment of addictive disorders. Psychopharmacol Bull 37 Suppl 2:89-97

Negash A, Kebede D, Alem A, Melaku Z, Deyessa N, Shibire T, Fekadu A, Fekadu D,
Jacobsson L, Kullgren G. 2004. Neurological soft signs in bipolar I disorder patients. Journal of Affective Disorders 80 (2-3): 221-230

Nath J, Sagar R (2001) Late-onset bipolar disorder due to hyperthyroidism. Acta Psychiatr Scand 104:72-73; discussion 74-75

Neves-Pereira M, Mundo E, Muglia P, King N, Macciardi F, Kennedy JL (2002) The brain-derived neurotrophic factor gene confers susceptibility to bipolar disorder: evidence from a family-based association study. Am J Hum Genet 71:651-655

New AS, Gelernter J, Goodman M, Mitropoulou V, Koenigsberg H, Silverman J, Siever LJ (2001) Suicide, impulsive aggression, and HTR1B genotype. Biol Psychiatry 50:62-65

Newman M, Zohar J, Kalian M, Belmaker RH (1984) The effects of chronic lithium and ECT on A1 and A2 adenosine receptor systems in rat brain. Brain Res 291:188-192

Oquendo MA, Waternaux C, Brodsky B, Parsons B, Haas GL, Malone KM, Mann JJ (2000) Suicidal behavior in bipolar mood disorder: clinical characteristics of attempters and nonattempters. J Affect Disord 59:107-117

Ott J (1989) Computer-simulation methods in human linkage analysis. Proc Natl Acad Sci U S A 86:4175-4178

Ott J, Hoh J (2000) Statistical approaches to gene mapping. Am J Hum Genet 67:289-294

Papadimitriou GN, Dikeos DG, Karadima G, Avramopoulos D, Daskalopoulou EG, Vassilopoulos D, Stefanis CN (1998) Association between the GABA(A) receptor alpha5 subunit gene locus (GABRA5) and bipolar affective disorder. Am J Med Genet 81:73-80

Papiol S, Rosa A, Gutierrez B, Martin B, Salgado P, Catalan R, Arias B, Fananas L (2004) Interleukin-1 cluster is associated with genetic risk for schizophrenia and bipolar disorder. J Med Genet 41:219-223

Parsian A, Chakraverty S, Todd RD (1995) Possible association between the dopamine D3 receptor gene and bipolar affective disorder. Am J Med Genet 60:234-237

Parsian A, Todd RD, Devor EJ, O'Malley KL, Suarez BK, Reich T, Cloninger CR (1991) Alcoholism and alleles of the human D2 dopamine receptor locus. Studies of association and linkage. Arch Gen Psychiatry 48:655-663

Pauls DL (1993) Behavioural disorders: lessons in linkage. Nat Genet 3:4-5

Peet NP, Bey P (2001) Pharmacogenomics: challenges and opportunities. Drug Discov Today 6:495-498

Perez J, Tardito D, Mori S, Racagni G, Smeraldi E, Zanardi R (2000) Altered Rap1 endogenous phosphorylation and levels in platelets from patients with bipolar disorder. J Psychiatr Res 34:99-104

Perris C (1966) A study of bipolar (manic-depressive) and unipolar recurrent depressive psychoses. Introduction. Acta Psychiatr Scand Suppl 194:9-14

Persson ML, Geijer T, Wasserman D, Rockah R, Frisch A, Michaelovsky E, Jonsson EG, Apter A, Weizman A (1999) Lack of association between suicide attempt and a polymorphism at the dopamine receptor D4 locus. Psychiatr Genet 9:97-100

Petty F (1995) GABA and mood disorders: a brief review and hypothesis. J Affect Disord 34:275-281

Phillips ML, Drevets WC, Rauch SL, Lane R (2003) Neurobiology of emotion perception I: The neural basis of normal emotion perception. Biol Psychiatry 54:504-514

Potash JB, Kane HS, Chiu YF, Simpson SG, MacKinnon DF, McInnis MG, McMahon FJ, DePaulo JR, Jr. (2000) Attempted suicide and alcoholism in bipolar disorder: clinical and familial relationships. Am J Psychiatry 157:2048-2050

Potash JB, Willour VL, Chiu YF, Simpson SG, MacKinnon DF, Pearlson GD, DePaulo JR, Jr., McInnis MG (2001) The familial aggregation of psychotic symptoms in bipolar disorder pedigrees. Am J Psychiatry 158:1258-1264

Potter WZ, Manji HK (1994) Catecholamines in depression: an update. Clin Chem 40: 279-287

Prathikanti S, McMahon FJ (2001) Genome scans for susceptibility genes in bipolar affective disorder. Ann Med 33:257-262

Preda A, MacLean RW, Mazure CM, Bowers MB, Jr. (2001) Antidepressant-associated mania and psychosis resulting in psychiatric admissions. J Clin Psychiatry 62:30-33

Sanacora G, Mason GF, Rothman DL, Behar KL, Hyder F, Petroff OA, Berman RM, Charney DS, Krystal JH. 1999. Reduced cortical gamma-aminobutyric acid levels in depressed patients determined by proton magnetic resonance spectroscopy. Archives of General Psychiatry 56(11):1043-7

Schildkraut JJ (1965) The catecholamine hypothesis of affective disorders: a review of supporting evidence. Am J Psychiatry 122:509-522

Schindler KM, Richter MA, Kennedy JL, Pato MT, Pato CN (2000) Association between homozygosity at the COMT gene locus and obsessive compulsive disorder. Am J Med Genet 96:721-724

Schmitz G, Aslanidis C, Lackner KJ (2001) Pharmacogenomics: implications for laboratory medicine. Clin Chim Acta 308:43-53

Schosser A, Fuchs K, Leisch F, Bailer U, Meszaros K, Lenzinger E, Willinger U, Strobl R, Heiden A, Gebhardt C, Kasper S, Sieghart W, Hornik K, Aschauer HN (2004) Possible linkage of schizophrenia and bipolar affective disorder to chromosome 3q29; a follow-up. J Psychiatr Res 38:357-364

Schulze TG, Buervenich S, Badner JA, Steele CJ, Detera-Wadleigh SD, Dick D, Foroud T, Cox NJ, MacKinnon DF, Potash JB, Berrettini WH, Byerley W, Coryell W, DePaulo JR, Jr., Gershon ES, Kelsoe JR, McInnis MG, Murphy DL, Reich T, Scheftner W, Nurnberger JI, Jr., McMahon FJ (2004) Loci on chromosomes 6q and 6p interact to increase susceptibility to bipolar affective disorder in the national institute of mental health genetics initiative pedigrees. Biol Psychiatry 56:18-23

Schumacher J, Jamra RA, Freudenberg J, Becker T, Ohlraun S, Otte AC, Tullius M, Kovalenko S, Bogaert AV, Maier W, Rietschel M, Propping P, Nothen MM, Cichon S (2004) Examination of G72 and D-amino-acid oxidase as genetic risk factors for schizophrenia and bipolar affective disorder. Mol Psychiatry 9:203-207

Schurhoff F, Bellivier F, Jouvent R, Mouren-Simeoni MC, Bouvard M, Allilaire JF, Leboyer M (2000) Early and late onset bipolar disorders: two different forms of manic-depressive illness? J Affect Disord 58:215-221

Schwab SG, Hallmayer J, Albus M, Lerer B, Eckstein GN, Borrmann M, Segman RH, Hanses C, Freymann J, Yakir A, Trixler M, Falkai P, Rietschel M, Maier W, Wildenauer DB (2000) A genome-wide autosomal screen for schizophrenia susceptibility loci in 71 families with affected siblings: support for loci on chromosome 10p and 6. Mol Psychiatry 5:638-649

Scott SK, Young AW, Calder AJ, Hellawell DJ, Aggleton JP, Johnson M (1997) Impaired auditory recognition of fear and anger following bilateral amygdala lesions. Nature 385:254-257

Segurado R, Detera-Wadleigh SD, Levinson DF, Lewis CM, Gill M, Nurnberger JI, Jr., Craddock N, et al. (2003) Genome scan meta-analysis of schizophrenia and bipolar disorder, part III: Bipolar disorder. Am J Hum Genet 73:49-62

Serretti A, Artioli P (2004) From molecular biology to pharmacogenetics: a review of the literature on antidepressant treatment and suggestions of possible candidate genes. Psychopharmacology (Berl) 174:490-503

Serretti A, Cusin C, Rossini D, Artioli P, Dotoli D, Zanardi R (2004) Further evidence of a combined effect of SERTPR and TPH on SSRIs response in mood disorders. Am J Med Genet 129B:36-40

Serretti A, Lilli R, Mandelli L, Lorenzi C, Smeraldi E (2001a) Serotonin transporter gene associated with lithium prophylaxis in mood disorders. Pharmacogenomics J 1:71-77

Smith MA, Makino S, Kvetnansky R, Post RM (1995) Stress and glucocorticoids affect the expression of brain-derived neurotrophic factor and neurotrophin-3 mRNAs in the hippocampus. J Neurosci 15:1768-1777

Sobel E, Lange K (1996) Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker-sharing statistics. Am J Hum Genet 58:1323-1337

Sobel E, Papp JC, Lange K (2002) Detection and integration of genotyping errors in statistical genetics. Am J Hum Genet 70:496-508

Sobel E, Sengul H, Weeks DE (2001) Multipoint estimation of identity-by-descent probabilities at arbitrary positions among marker loci on general pedigrees. Hum Hered 52:121-131

Soloff PH, Lynch KG, Kelly TM, Malone KM, Mann JJ (2000) Characteristics of suicide attempts of patients with major depressive episode and borderline personality disorder: a comparative study. Am J Psychiatry 157:601-608

Spear BB, Heath-Chiozzi M, Huff J (2001) Clinical application of pharmacogenetics. Trends Mol Med 7:201-204

Spedding M, Neau I, Harsing L (2003) Brain plasticity and pathology in psychiatric disease: sites of action for potential therapy. Curr Opin Pharmacol 3:33-40

Spielman RS, McGinnis RE, Ewens WJ (1993) Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). Am J Hum Genet 52:506-516

Stahl SM, (Eds), 2000. Essential psychopharmacology, second edition.
Cambridge University press

Steimer W, Muller B, Leucht S, Kissling W (2001) Pharmacogenetics: a new diagnostic tool in the management of antidepressive drug therapy. Clin Chim Acta 308:33-41

Stoll AL, Mayer PV, Kolbrener M, Goldstein E, Suplit B, Lucier J, Cohen BM, Tohen M (1994) Antidepressant-associated mania: a controlled comparison with spontaneous mania. Am J Psychiatry 151:1642-1645

Stoltenberg SF, Burmeister M (2000) Recent progress in psychiatric genetics-some hope but no hype. Hum Mol Genet 9:927-935

Straub RE, MacLean CJ, O'Neill FA, Burke J, Murphy B, Duke F, Shinkwin R, Webb BT, Zhang J, Walsh D, et al. (1995) A potential vulnerability locus for schizophrenia on chromosome 6p24-22: evidence for genetic heterogeneity. Nat Genet 11:287-293

Swann AC, Bowden CL, Morris D, Calabrese JR, Petty F, Small J, Dilsaver SC, Davis JM (1997) Depression during mania. Treatment response to lithium or divalproex. Arch Gen Psychiatry 54:37-42

Swann AC, Pazzaglia P, Nicholls A, Dougherty DM, Moeller FG (2003) Impulsivity and phase of illness in bipolar disorder. J Affect Disord 73:105-111

Thoenen H (1995) Neurotrophins and neuronal plasticity. Science 270:593-598

Thome J, Sakai N, Shin K, Steffen C, Zhang YJ, Impey S, Storm D, Duman RS (2000) cAMP response element-mediated gene transcription is upregulated by chronic antidepressant treatment. J Neurosci 20:4030-4036

Tishkoff SA, Verrelli BC (2003) Role of evolutionary history on haplotype block structure in the human genome: implications for disease mapping. Curr Opin Genet Dev 13:569-575

Wang DG, Fan JB, Siao CJ, Berno A, Young P, Sapolsky R, Ghandour G, et al. (1998) Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. Science 280:1077-1082

Warman DM, Forman EM, Henriques GR, Brown GK, Beck AT (2004) Suicidality and psychosis: beyond depression and hopelessness. Suicide Life Threat Behav 34:77-86

Warren JT, Jr., Peacock ML, Rodriguez LC, Fink JK (1993) An MspI polymorphism in the hyman serotonin receptor gene (HTR2): detection by DGGE and RFLP analysis. Hum Mol Genet 2:338

Washizuka S, Iwamoto K, Kazuno AA, Kakiuchi C, Mori K, Kametani M, Yamada K, Kunugi H, Tajima O, Akiyama T, Nanko S, Yoshikawa T, Kato T (2004) Association of mitochondrial complex I subunit gene NDUFV2 at 18p11 with bipolar disorder in Japanese and the National Institute of Mental Health pedigrees. Biol Psychiatry 56:483-489

Weber WW (2001) The legacy of pharmacogenetics and potential applications. Mutat Res 479:1-18

Weeks DE, Lange K (1988) The affected-pedigree-member method of linkage analysis. Am J Hum Genet 42:315-326

Weeks DE, Lange K (1992) A multilocus extension of the affected-pedigree-member method of linkage analysis. Am J Hum Genet 50:859-868

Weeks DE, Lathrop GM (1995) Polygenic disease: methods for mapping complex disease traits. Trends Genet 11:513-519

Wei J, Hemmings GP (2000) Searching for a locus for schizophrenia within chromosome Xp11. Am J Med Genet 96:4-7

Weissman MM, Bland R, Joyce PR, Newman S, Wells JE, Wittchen HU (1993) Sex differences in rates of depression: cross-national perspectives. J Affect Disord 29:77-84

Weissman MM, Gershon ES, Kidd KK, Prusoff BA, Leckman JF, Dibble E, Hamovit J, Thompson WD, Pauls DL, Guroff JJ (1984) Psychiatric disorders in the relatives of probands with affective disorders. The Yale University--National Institute of Mental Health Collaborative Study. Arch Gen Psychiatry 41:13-21

Whittemore AS, Halpern J (1994) A class of tests for linkage using affected pedigree members. Biometrics 50:118-127

Wildenauer DB, Schwab SG (1999) Chromosomes 8 and 10 workshop. Am J Med Genet 88:239-243

Willeit M, Praschak-Rieder N, Zill P, Neumeister A, Ackenheil M, Kasper S, Bondy B (2003) C825T polymorphism in the G protein beta3-subunit gene is associated with seasonal affective disorder. Biol Psychiatry 54:682-686

Williams RS, Cheng L, Mudge AW, Harwood AJ (2002) A common mechanism of action for three mood-stabilizing drugs. Nature 417:292-295

Willour VL, Zandi PP, Huo Y, Diggs TL, Chellis JL, MacKinnon DF, Simpson SG, McMahon FJ, Potash JB, Gershon ES, Reich T, Foroud T, Nurnberger JI, Jr., DePaulo JR, Jr., McInnis MG (2003) Genome scan of the fifty-six bipolar pedigrees from the NIMH genetics initiative replication sample: chromosomes 4, 7, 9, 18, 19, 20, and 21. Am J Med Genet 121B:21-27

Wiltshire S, Cardon LR, McCarthy MI (2002) Evaluating the results of genomewide linkage scans of complex traits by locus counting. Am J Hum Genet 71:1175-1182

Wunderlich U, Bronisch T, Wittchen HU, Carter R (2001) Gender differences in adolescents and young adults with suicidal behaviour. Acta Psychiatr Scand 104:332-339

# APPENDICES

**APPENDIX A**: Request form for molecular studies

**APPENDIX B**: Techniques

**APPENDIX C**: Media, Buffers and Solutions

**APPENDIX D**: Scripts

A

# REQUEST FOR MOLECULAR STUDIES (DNA)

**PROF R. RAMESAR**
**Molecular Laboratory**
**Division of Human Genetics**
**1st Floor, Anatomy Building**
**UCT Medical School, Observatory 7925**
**SOUTH AFRICA**

Tel: 27-21- 406 6297 Fax: 27-21- 4477703

***Please fill in all the information requested***

Surname: _____ First Name(s): _____

New Family: Yes ☐ No X (If no, please fill in family name) Family name: _____

Medical Aid: _____ Medical Aid No: _____

Sex: M ☐ F ☐ Date of Birth: Year: _____ Month: _____ Day: _____

Number of children: _____

Ethnic Origin : (please indicate ancestry of both your mother and father ) _____

Contact Address: _____ Town: _____ Fax: Tel: _____

Referring Doctor/Sister: _____ Town: _____ Fax: Tel: _____

Hospital or Address: _____ Town: _____ Fax: Tel: _____

Reason for Referral (Clinical diagnosis):

Affected ☐ At Risk ☐ Carrier ☐ Spouse ☐ Query ☐ Unaffected ☐

| | | | | | |
|---|---|---|---|---|---|
| Becker Muscular Dys. | ☐ | Duchenne Muscular Dys | ☐ | Colonic Carcinoma | ☐ |
| Fragile-X Syndrome | ☐ | Bipolar Disorder | ☐ | Huntington Disease | ☐ |
| Retinitis Pigmentosa | ☐ | Spinocerebellar Ataxia | ☐ | Waardenberg Syndrome | ☐ |

Additional disorders (apparent or previously treated): _____

Additional family history _____

Clinical Details:

Physical disability ☐ Mental retardation ☐ Deafness ☐ Impaired vision ☐ Night blindness ☐

Other: _____

Have samples from this patient been sent to a DNA lab before? (DELETE WHERE NOT APPLICABLE) YES / NO / Don't Know

If Yes, where: _____

***For Laboratory use only:***

*DNA number:* _____ *Vol.Blood:* _____ *(ml)* *Other:* _____

*Date Received: Year:* _____ *Month:* _____ *Day:* _____ *Computer Index No:* _____

# CONSENT FOR DNA ANALYSIS AND STORAGE

1. I, _____ , request that an attempt be made using genetic material to assess the probability that: I / my child / my unborn child (DELETE WHERE NOT APPLICABLE) might have inherited a disease-causing mutation in the gene for: _____

2. I understand that the genetic material for analysis is to be obtained from: blood cells/skin sample/other (specify) (DELETE WHERE NOT APPLICABLE) :

3. I request that no portion of the sample be stored for later use. ☐ (MARK IF APPLICABLE )
   Or
   I request that a portion of the sample be stored indefinitely for (DELETE WHERE NOT APPLICABLE):
   - ( a ) possible re-analysis
   - ( b ) analysis for the benefit of members of my immediate family
   - ( c ) research purposes, subject to the approval of the University of Cape Town Research Ethics Committee, provided that any information from such research will remain confidential.

4. The results of the analysis carried out on this sample of stored biological material will be made known to me, via my doctor, in accordance with the relevant protocol, if and when available.
   In addition, I authorise that they may be made known to: (DELETE WHERE NOT APPLICABLE) :
   other doctors involved in my care
   the following family members: _____
   other: _____

5. I authorise / do not authorise my doctor(s) (DELETE WHERE NOT APPLICABLE) to provide relevant clinical details to the Division of Human Genetics, UCT.

6. I have been informed that:
   - ( a ) there are risks and benefits associated with genetic analysis and storage of biological material and these have been explained to me.
   - ( b ) the analysis procedure is specific to the genetic condition mentioned above and cannot determine the complete genetic makeup of an individual.
   - ( c ) the genetics laboratory is under an obligation to respect medical confidentiality .
   - ( d ) genetic analysis may not be informative for some families or family members.
   - ( e ) even under the best conditions, current technology of this type is not perfect and could lead to incorrect results.
   - ( f ) where biological material is used for research purposes, there may be no direct benefit to me.

7. I understand that I may withdraw my consent for any aspect of the above at any time without this affecting my future medical care.

8. **ALL OF THE ABOVE HAS BEEN EXPLAINED TO ME IN A LANGUAGE THAT I UNDERSTAND AND MY QUESTIONS ANSWERED BY:**

_____ DATE: _____

**Patient signature** _____ **Witnessed consent** _____

**NOTE - PLEASE INSERT A FAMILY PEDIGREE DRAWING ON THE REVERSE OF THIS FORM**

## APPENDIX B

**Techniques**

B.1 DNA isolation from lymphocytes

B.2 DNA oligonucleotides

B.3 Genotype analysis

B.4 Restriction endonuclease digests

B.5 DNA sequencing

USA), dextran blue (Genescan®, PE Applied Biosystems, Foster City, USA) and ROX-500 size standard (Genescan®, PE Applied Biosystems, Foster City, USA). Samples were heat denatured at 95°C for 2min and placed on ice before loading on 12cm (well-to-read) ABI plates.

### B.3.2 3100 Genetic Analyzer

To 1µl of sample (in 96 well plates) add 8µl of loading buffer, containing 0.5µl of ROX-500 size standard (Genescan®, PE Applied Biosystems, Foster City, USA) and 8.5 µl of Hi-Di formamide (Genescan®, PE Applied Biosystems, Foster City, USA). Samples were heat denatured at 95°C for 2min and placed on ice before loading on the 3100 Genetic Analyzer.

## B.4 Restriction Endonuclease Digestion

Restriction enzymes were purchased from:

- Promega, UK
- New England Biolabs, UK
- Roche, Germany

Conditions described by the manufacturers' were followed at all times.

## B.5 DNA Sequencing

- The PCR product was first separated on a 2% agarose gel, after which the DNA fragment was excised from the agarose gel with a clean, sharp scalpel.
- The DNA was purified in *QIAquick* spin columns (*Qiagen*, UK) according to the manufacturer's *QIAquick* Gel extraction protocol.
- Purified PCR products were sequenced using the *BigDye* terminator cycle sequencing kit (*Applied Biosystems*, USA), according to the manufacturer's protocol.
- *Centri-Sep* Spin Columns (*Princeton separations*, USA) were used to remove unincorporated nucleotides.
- The sequencing reactions were separated by automated gel electrophoresis using an *ABI PRISM* 377 DNA sequencer (*Applied Biosystems*, USA).

*Centri-Sep spin column clean-out*

- To hydrate the column, one has to proceed as follows:
- Gently tap the column to insure that the dry gel has settled in the bottom of the spin column
- Remove the column cap and reconstitute the column by adding 0.80ml of reagent grade water or buffer. Leave the column end stopper in place so the column can stand up by itself. Replace the column cap and hydrate the gel by shaking and inverting the column or vortexing briefly. It is important to hydrate all of the dry gel
- Allow at least 30 min of room temperature hydration time before using the columns. Reconstituted columns may be stored refrigerated at 4°C for several days. Longer storage may be accomplished in 10mM sodium azide

## APPENDIX C

**Media, Buffers and Solutions**

C.1 10X TBE

C.2 Agarose loading buffer

C.3 Formamide loading dye

C.4 6% denaturing polyacrylamide gels

C.5 Silver staining

# APPENDIX D

## Scripts

```
      DOUBLE PRECISION posn(500), pri0(2000), pri1(2000),
     +pri2(2000), post0(2000,500), post1(2000,500), post2(2000,500),
     +dumposn, fpri0(2000), fpri1(2000), fpri2(2000),
     +fpost0(2000), fpost1(2000), fpost2(2000), f00, f01, f02,
     +f10, f11, f12, f20, f21, f22, p00, p01, p02, p10, p11, p12,
     +p20, p21, p22


      INTEGER i, j, npairs, nincrs, fam(2000), id1(2000), id2(2000),
     + pfam(2000), pid1(2000), pid2(2000), k,  ffam(2000), fid1(2000),
     + fid2(2000),
     + pffam(2000), pfid1(2000), pfid2(2000)

      open(3, file='maketwo.dat', status='old')
      open(4, file = 'firstprior.mpt', status = 'old')
      open(5, file= 'firstposterior.mpt', status= 'old')
      open(6, file = 'secondprior.mpt', status = 'old')
      open(7, file= 'secondposterior.mpt', status= 'old')

      open(14, file='twoposterior.dat', status='unknown')
      open(15, file='twoprior.dat', status='unknown')


      read(3,*) npairs
      read(3,*) nincrs
      read(4,*)
      read(5,*)
      read(6,*)
      read(7,*)


#     assume only one position at first locus, possibly many at second
#     read in data for second locus

      do 100 i=1,npairs

        read(6,*) fam(i), id1(i), id2(i), pri0(i), pri1(i), pri2(i)
        do 200 j=1,nincrs

          read(7,*) pfam(i), posn(j), pid1(i), pid2(i), post0(i,j),
     +post1(i,j), post2(i,j)
```

```
     +((fid1(i).eq.id2(j)).and.(fid2(i).eq.id1(j)))) then

             f00= fpri0(i)*pri0(j)
             f01= fpri0(i)*pri1(j)
             f02= fpri0(i)*pri2(j)
             f10= fpri1(i)*pri0(j)
             f11= fpri1(i)*pri1(j)
             f12= fpri1(i)*pri2(j)
             f20= fpri2(i)*pri0(j)
             f21= fpri2(i)*pri1(j)
             f22= fpri2(i)*pri2(j)
             do 500 k=1,nincrs

             p00= fpost0(i)*post0(j,k)
             p01= fpost0(i)*post1(j,k)
             p02= fpost0(i)*post2(j,k)
             p10= fpost1(i)*post0(j,k)
             p11= fpost1(i)*post1(j,k)
             p12= fpost1(i)*post2(j,k)
             p20= fpost2(i)*post0(j,k)
             p21= fpost2(i)*post1(j,k)
             p22= fpost2(i)*post2(j,k)

                write(15,999), posn(k), ffam(i), fid1(i), fid2(i),
     +f00, f01, f02, f10, f11, f12, f20, f21, f22

             write(14,999), posn(k), ffam(i), fid1(i), fid2(i),
     +p00, p01, p02, p10, p11, p12, p20, p21, p22

500            continue

             end if

400             continue
300             continue

999     format(f8.3, i12, i12, i12, 9f12.9)

550            continue
             stop
             end
```

```
      write(*,*) ped, id, pa, ma, sex,
     +aff, (al1(i), al2(i), i=1, nloci)

 100  continue
      stop
      end
```

### D.3.3 extractline41.f

```
#  This was part D2 goodscript
#  To extract line 41 of the twomls.out from twolocarp

DOUBLE PRECISION  posn, sing, mul, add, gen
      INTEGER i

      do 100 i=1,40

     read(*,*)

100  continue

     read(*,*) posn, sing, mul, add, gen
     write(*,*) posn, sing, mul, add, gen
      stop
     end
```