The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.
Elucidation of Bipolar Disorder: A Convergent Approach using Genetics and Imaging

Name: Shareefa Dalvie (DLVSHA006)

Project Supervisor: Prof. Rajkumar S. Ramesar

Project Co-supervisors: Prof. Dan J. Stein and Dr Neil Horn

Division of Human Genetics
Presented for M.Sc. (Med)
Faculty of Health Sciences
University of Cape Town
6 May 2011
Acknowledgements

Thank you to my Creator, for guiding me through this challenging task

The completion of this thesis is largely due to the following individuals:

Prof. Rajkumar Ramesar for providing me the opportunity to be part of the Human Genetics lab and for supervision

Prof. Dan J. Stein and Dr Neil Horn for supervision

Sr Gameda Benefeld for obtaining the blood samples, conducting SCID interviews and for always being available to help

Dr Collet Dandara for help with editing and guidance regarding stats analysis

Dr Fleur Howells and Dr Leesha Singh for providing brain imaging analysis data

Mrs Kholiwe Dubula for assistance with DNA isolation

The Division of Human Genetics. A special thank you to Ms Kusha Kalideen, Ms Lindiwe Lamola, and Ms Zeinsonisaa Latief

My parents and siblings for always supporting me

National Research Foundation (NRF) and the University of Cape Town (UCT) for funding
Plagiarism Declaration

1. I know that plagiarism is wrong. Plagiarism is to use another's work and to pretend that it is one's own.

2. I have used the *American Journal of Medical Genetics* convention for citation and referencing. Each significant contribution to, and quotation in, this thesis from the work, or works, of other people has been attributed, and has been cited and referenced.

3. This thesis is my own work using my own words.

4. I have not allowed, and will not allow, anyone to copy my work with the intention of passing it off as his or her own work.

Name: Shareefa Dalvie

Signature:

Date: 6 May 2011
# Table of Contents

Acknowledgements 2

Plagiarism Declaration 3

List of Figures 8

List of Tables 9

Abbreviations 10

Abstract 14

Chapter 1: Introduction 16

1.1 Features and Characteristics of Bipolar Disorder 16
   1.1.1 DSM-IV criteria for BPDI 18

1.2 Kraepelinian Dichotomy: BPD and Schizophrenia 18

1.3 Techniques Used to Determine the Aetiology of BPD 20
   1.3.1 Mouse Models 20
   1.3.2 Human Gene Expression Data 23
   1.3.3 Human Blood Evidence 24
   1.3.4 Genetics and Family Studies 25
   1.3.5 Brain Imaging 29

1.4 Convergent Functional Genomics 32

1.5 Candidate Genes 33
   1.5.1 KLF12 33
   1.5.2 ARNTL 33
   1.5.3 BDNF 35
   1.5.4 ALDH1a1 36
1.5.5 *A2PB1* 37
1.5.6 *MBP* 38
1.5.7 *AK3L1* 39
1.5.8 *GSK3β* 39
1.5.9 *ANK3* 40
1.5.10 *CACNA1C* 41

1.6 Aims and Objectives 42

**Chapter 2: Materials and Methods** 44

2.1 BPD Cohort Description 44
2.1.1 Cohort Description for Case-Control Analysis 45
2.1.2 Cohort Description for Brain Imaging Analysis 45

2.2 DNA Isolation and Quality Determination 46

2.3 Gene and SNP Selection, Primer Design and Synthesis 47
2.3.1. Gene and SNP selection 47
2.3.2 External Primer Design 48

2.4 Polymerase Chain Reaction (PCR) 49

2.5 SNaPshot PCR 51
2.5.1 SNaPshot PCR Internal Primer Design 53
2.5.2 Sap and ExoI Clean-up of PCR Products 55
2.5.3 SNaPshot PCR Conditions 56
2.5.4 Post- SNaPshot SAP Treatment 57
2.5.5 Capillary Electrophoresis on ABI PRISM 58
2.5.6 Analysis of Results 58

2.6 Genotype Validation 58
2.6.1 PCR of DNA Samples for Cycle Sequencing 59
2.6.2 Cycle Sequencing 59
2.6.3 Ethanol Precipitation 59
2.6.4 Capillary Electrophoresis on ABI Prism 60
2.6.5 Analysis of Results

2.7 Statistical Analysis
   2.7.1 Genotype Frequencies, Hardy-Weinberg Equilibrium Calculations and Logistic Regression with R
   2.7.2 MQLS (adjusting for relatedness) and Pedstats

2.8 Neuroimaging of Subset of Cohort
   2.8.1 Statistical Analysis of Neuroimaging Results

Chapter 3: Results

3.1 BPD Cohort Description
   3.1.1 Cohort Description for Case-Control Analysis
   3.1.2 Cohort Description for Brain Imaging Analysis

3.2 DNA Quality Determination

3.3 Gene and SNP Selection, Primer Design and Synthesis

3.4 Polymerase Chain Reaction (PCR)
   3.4.1 PCR Optimisation
   3.4.2 PCR Conditions for Candidate Genes

3.5 SNaPshot PCR
   3.5.1 Sap and ExoI Clean-up of PCR Products
   3.5.2 Capillary Electrophoresis on ABI PRISM
   3.5.3 Analysis of Results

3.6 Genotype Validation
   3.6.1 PCR of DNA Samples for Cycle Sequencing
   3.6.2 Cycle Sequencing
   3.6.3 Analysis of Results

3.7 Statistical Analysis for Case Control Association
   3.7.1 Genotype Frequencies, Hardy Weinberg Equilibrium Calculations and Logistic Regression with R
3.7.2 $M_{QLS}$ (adjusting for relatedness) and Pedstats

3.8 Statistical Analysis of Neuro-imaging

Chapter 4: Discussion

4.1 Future Directions

4.2 Conclusions

Websites Referenced

References

Appendices

Appendix 1: Patient Consent Form

Appendix 2: Gentra Puregene™ DNA Isolation Protocol

Appendix 3: Reagents, Buffers and Solutions

Appendix 4: GeneScan -120 Liz size standard

Appendix 5: Ethanol Precipitation

Appendix 6: Annealing Temperatures ($T_a$) for Temperature Gradient Reactions

Appendix 7: Additional Temperature Gradient Gel Pictures

Appendix 8: Psychosis and relapse in bipolar disorder are related to GRM3, DAOA, and GRIN2B genotype. Dalvie et al. [2010]
List of Figures

Figure 1: An illustration of the SNaPshot PCR principle 52
Figure 2: Ethnic group distribution of BPD archive 64
Figure 3: Distribution of most prevalent phenotypes of BPD cohort 65
Figure 4: Number of males and females in case and control groups 66
Figure 5: Number of Caucasian and Mixed Ancestry individuals 66
Figure 6: Integrity gel of genomic DNA 67
Figure 7: KLF12 rs9600160 (180bp) temperature gradient 69
Figure 8: Group A PCR Products for three primer pairs 71
Figure 9: Group B PCR Products for two primer pairs 71
Figure 10: Group C PCR Products for two primer pairs 72
Figure 11: Group D PCR Products for two primer pairs 72
Figure 12: Group E PCR Products for two primer pairs 73
Figure 13: SNaPshot results for Group A-B 74
Figure 14: SNaPshot results for Group C-D-E 74
Figure 15: PCR amplification for cycle sequencing reaction 75
Figure 16: Sequencing results for BDNF rs6265 76
Figure 17: Sequencing results for ANK3 rs9804190 76
Figure 18: Genotype counts of cases and controls 78
Figure 19: Genotype counts of cases and controls 79
List of Tables

Table 1: Diagnostic criteria of mood episodes characteristic of BPD 19
Table 2: Genes and pathways implicated in human expression studies 24
Table 3: Summary of recent BPD Genome Wide Association Studies 28
Table 4: Candidate genes from CFG and GWAS and respective SNPs 43
Table 5: Selected candidate genes, respective SNPs and CFG scores 48
Table 6: External primer sequences for amplification of candidate gene regions 50
Table 7: PCR cycling conditions 51
Table 8: Dyes assigned to each of the ddNTPs 52
Table 9: SNaPshot internal primer sequences 54
Table 10: Conditions for Sap and ExoI purification 55
Table 11: Sap and ExoI DNA purification 56
Table 12: The components of the SNaPshot PCR reaction 57
Table 13: Conditions for SNaPshot PCR 57
Table 14: Components of the cycle sequencing reaction 60
Table 15: Cycle sequencing PCR conditions 60
Table 16: Optimal annealing temperature determined for each primer pair 69
Table 17: \( R \) logistic regression output for case-control data 80
Table 18: Pedigree structure from Pedstats output 81
Table 19: \( M_{QLS} \) Output for Case Control Association 82
Table 20: Kruskal Wallis and ANOVA values for each of the SNPs and imaging measures 83
Abbreviations

°C     Degrees Celsius
µl     Microlitre
µM     Micromolar
A2BP1  Ataxin-2 Binding Protein 1
ADHD   Attention Deficit Hyperactivity Disorder
AEI    Allelic Expression Imbalance
AK3L1  Adenylate Kinase 3-Like 1
ALDH1A1 Aldehyde Dehydrogenase 1 Family, Member A1
ALDH1A2 Aldehyde Dehydrogenase 1 Family, Member A2
AMP    Adenosine Monophosphate
ANK2   Ankyrin-2
ANK3   Ankyrin-3
ANOVA  Analysis of Variance
AOO    Age of Onset
ARNTL  Aryl Hydrocarbon Nuclear Translocator–Like
BDNF   Brain Derived Neurotrophic Factor
bp     Base pairs
BPD    Bipolar Disorder
BPDI   Bipolar Disorder type I
BPDII  Bipolar Disorder type II
BPDNOS BPD not otherwise specified
BRCA2  Breast Cancer 2
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CACNA1A</td>
<td>Calcium Channel, Voltage-Dependent, L type, Alpha 1A Subunit</td>
</tr>
<tr>
<td>CACNA1C</td>
<td>Calcium Channel, Voltage-Dependent, L type, Alpha 1C Subunit</td>
</tr>
<tr>
<td>CANB1</td>
<td>Calcium Channel, Voltage-Dependent, Beta-1</td>
</tr>
<tr>
<td>CACNB2</td>
<td>Calcium channel, voltage dependent, beta 2 subunit</td>
</tr>
<tr>
<td>CFG</td>
<td>Convergent Functional Genomics</td>
</tr>
<tr>
<td>CNS</td>
<td>Central Nervous System</td>
</tr>
<tr>
<td>Cry</td>
<td>Cryptochrome</td>
</tr>
<tr>
<td>DAOA</td>
<td>D-amino acid Oxidase Activator</td>
</tr>
<tr>
<td>DCTN5</td>
<td>Dynactin 5</td>
</tr>
<tr>
<td>DGKH</td>
<td>Diacylglycerol Kinase</td>
</tr>
<tr>
<td>dH₂O</td>
<td>Distilled Water</td>
</tr>
<tr>
<td>dNTP</td>
<td>Deoxynucleotide Triphosphates</td>
</tr>
<tr>
<td>ddNTP</td>
<td>Dideoxynucleotide Triphosphates</td>
</tr>
<tr>
<td>DSM-IV</td>
<td>Diagnostic and Statistical Manual of Mental Disorders-IV</td>
</tr>
<tr>
<td>ExoI</td>
<td>Exonuclease I</td>
</tr>
<tr>
<td>EtBr</td>
<td>Ethidium Bromide</td>
</tr>
<tr>
<td>GRIN2B</td>
<td>Ionotropic Glutamate Receptor, N-methyl D-aspartate 2B</td>
</tr>
<tr>
<td>GRM3</td>
<td>Glutamate Receptor, Metabotropic 3</td>
</tr>
<tr>
<td>GSK3β</td>
<td>Glycogen Synthase Kinase 3β</td>
</tr>
<tr>
<td>GTP</td>
<td>Guanosine Triphosphate</td>
</tr>
<tr>
<td>GWAS</td>
<td>Genome-Wide Association Study</td>
</tr>
<tr>
<td>HCV</td>
<td>Hippocampal Volume</td>
</tr>
<tr>
<td>HWE</td>
<td>Hardy Weinberg Equilibrium</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>ICV</td>
<td>Intracranial Volume</td>
</tr>
<tr>
<td>KLF</td>
<td>Kruppel-like Factor</td>
</tr>
<tr>
<td><em>KLF12</em></td>
<td><em>Kruppel-like Factor 12</em></td>
</tr>
<tr>
<td>MBP</td>
<td>Myelin Basic Protein</td>
</tr>
<tr>
<td>MDD</td>
<td>Major Depressive Disorder</td>
</tr>
<tr>
<td>MDE</td>
<td>Major Depressive Episode</td>
</tr>
<tr>
<td>Met</td>
<td>Methionine</td>
</tr>
<tr>
<td>mm</td>
<td>Millimetres</td>
</tr>
<tr>
<td>MRI</td>
<td>Magnetic Resonance Imaging</td>
</tr>
<tr>
<td>MWM</td>
<td>Molecular Weight Marker</td>
</tr>
<tr>
<td>NCBI</td>
<td>National Centre for Biotechnology Information</td>
</tr>
<tr>
<td><em>NDUFAB1</em></td>
<td><em>NADH dehydrogenase (ubiquinone) 1, alpha/beta subcomplex, 1</em></td>
</tr>
<tr>
<td>ng</td>
<td>Nanogram</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear Magnetic Resonance</td>
</tr>
<tr>
<td><em>PALB2</em></td>
<td><em>Partner and Localizer of BRCA2</em></td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase Chain Reaction</td>
</tr>
<tr>
<td>PD</td>
<td>Personality Disorder</td>
</tr>
<tr>
<td><em>Per</em></td>
<td><em>Period</em></td>
</tr>
<tr>
<td><em>PER3</em></td>
<td><em>Period Homolog 3</em></td>
</tr>
<tr>
<td>PTSD</td>
<td>Post traumatic Stress Disorder</td>
</tr>
<tr>
<td><em>RARβ</em></td>
<td><em>Retinoic Acid Receptor β</em></td>
</tr>
<tr>
<td>RBC</td>
<td>Red Blood Cell</td>
</tr>
<tr>
<td><em>RXRβ</em></td>
<td><em>Retinoid X Receptor β</em></td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>SABP</td>
<td>Schizoaffective Disorder, Bipolar Type</td>
</tr>
<tr>
<td>SAP</td>
<td>Shrimp Alkaline Phosphatase</td>
</tr>
<tr>
<td>SCA</td>
<td>Spinocerebellar Ataxia</td>
</tr>
<tr>
<td>SCID</td>
<td>Structured Clinical Interview for the DSM-IV</td>
</tr>
<tr>
<td>SCZ</td>
<td>Schizophrenia</td>
</tr>
<tr>
<td>SNP</td>
<td>Single Nucleotide Polymorphism</td>
</tr>
<tr>
<td>Ta</td>
<td>Annealing Temperature</td>
</tr>
<tr>
<td>TBE</td>
<td>Tris Borate EDTA</td>
</tr>
<tr>
<td>UCT</td>
<td>University of Cape Town</td>
</tr>
<tr>
<td>Val</td>
<td>Valine</td>
</tr>
<tr>
<td>VBM</td>
<td>Voxel Based Morphometry</td>
</tr>
</tbody>
</table>
Abstract

Introduction: Bipolar Disorder (BPD) is a severe, neuropsychiatric disorder which affects approximately 1% of the population. Due to the high heritability of BPD, genetic variations amongst affected individuals have been investigated. Recently, researchers have developed a method termed Convergent Functional Genomics (CFG) which aims to identify a prioritized candidate gene list, based on the results of previous studies, underlying the pathology of BPD and other psychiatric illnesses. In addition, neuro-imaging studies have revealed that BPD is associated with structural changes in the brain and that certain brain regional volumes are heritable.

Aims and Objectives: The aims were to determine whether variants, within ten selected candidate genes, have an association with BPD and whether any relationship exists between these variants and brain imaging volumes in subjects with BPD. The objectives were to i) select a list of BPD candidate genes, ii) identify a cohort of individuals from the BPD registry, iii) genotype the candidate genes using a polymerase chain reaction (PCR) based technology, iv) analyse the genotyping data with the appropriate statistical methods, and v) obtain brain imaging data and perform the appropriate statistical analysis

Materials and Methods: Candidate genes were selected based on the most significant findings of previous publications. Multiplex SNaPshot™ PCR was used to genotype 224 case and 202 control samples. Cases comprised of individuals in a South African cohort
of Mixed Ancestry and Caucasians, affected with BPD Type I. Magnetic resonance imaging (MRI) data was collected for a subset of the case group (n=22).

**Results:** Preliminary results suggest associations between SNPs in the candidate genes, *ankyrin3 (ANK3)*, *ataxin-2 binding protein1 (A2BP1)* and *myelin basic protein (MBP)*, with BPD. No significant associations were found between genotypes and brain imaging data.

**Conclusions:** These results suggest that BPD may have a common aetiology with neurodegenerative disorders. Polymorphisms within *ANK3, A2BP1* and *MBP* should be further investigated to determine the specific effects of these variants on the expression of the respective genes.
Chapter 1: Introduction

Bipolar Disorder (BPD) is a severe psychiatric illness which is characterized by extremes in mood and affects approximately 1% of the population. Much research has been undertaken to determine the underlying causes of this highly heterogeneous disorder. However, due to its high heritability, genetic factors have been thought to be a major contributor to the underlying aetiology. Recently, researchers have developed a Convergent Functional Genomics (CFG) approach which aims to derive a prioritized candidate gene list, based on the results of a wide range of previous studies, including brain imaging and animal models. The 113 CFG prioritized genes that have been identified comprise those involved in a wide range of central nervous system (CNS) function including neuronal maintenance, circadian rhythm regulation and transcriptional regulation [Le-Niculescu et al., 2008]. In addition, several other genes have been implicated through hypothesis-free genome-wide association studies (GWAS) [Stratton, 2007; Baum et al., 2008; Ferreira et al., 2008; Sklar et al., 2008; Zhang et al., 2008; Hattori et al., 2009; Scott et al., 2009; Smith et al., 2009]. Many of these genes overlap with those identified by the CFG approach and are ideal candidates to interrogate in studies investigating the aetiology of BPD.

1.1 Features and Characteristics of Bipolar Disorder

BPD is a severe, debilitating mental disorder affecting approximately 1% of the population [Merikangas et al., 2007]. This disorder is characterized by mood episodes ranging from an elevated mood to severe depression [Ogden et al., 2004]; onset usually occurs during early adulthood, and affects both males and females at equal frequencies.
[Miller, 2006]. The affected individual may experience a decline in quality of life, productivity, and longevity [Thase and Denko, 2008] and psychotic features such as delusions and hallucinations may also be a feature [American Psychiatric Association. Task Force on DSM-IV, 1994]. Individuals with BPD have greater difficulty in sustaining interpersonal relationships and have trouble maintaining employment compared to those without mood disorders: a total of 54% of affected individuals were dismissed from their jobs compared to 29% of a matched unaffected group [Calabrese et al., 2003]. The same study showed that persons suffering with the disorder have difficulty abiding by the law; and have a significantly higher chance of being arrested for crimes. Violent behaviour is often observed during manic episodes [American Psychiatric Association. Task Force on DSM-IV, 1994]. In 1991, the total economic burden of BPD in the USA was estimated at US $45 billion. This cost is attributed to medication, hospitalisation, misdiagnosis, and the estimated financial burden due to productivity loss [Hirschfeld and Vornik, 2005]. Between 10% and 20% of patients successfully commit suicide [Müller-Oerlinghausen et al., 2002; Bauer and Pfennig, 2005]. Currently, the diagnosis of BPD is based exclusively on clinical symptoms as the molecular pathology has yet to be discovered [Craddock and Sklar, 2009].

Based on different symptomatic criteria, there are several types of BPD distinguishable in the population: i) BPD type I (BPDI), ii) BPD type II (BPDII), iii) Cyclothymia, and iv) “BPD not otherwise specified” (BPDNOS) [American Psychiatric Association. Task Force on DSM-IV, 1994]. For the purposes of this study, a more detailed description of BPDI, the most severe type, is described below.
1.1.1 DSM-IV criteria for BPD

According to DSM-IV (Diagnostic and Statistical Manual of Mental Disorders-IV) criteria [American Psychiatric Association. Task Force on DSM-IV, 1994], BPD is a recurrent mood disorder, characterized by the presence of one or more manic or mixed episodes (see Table 1 for diagnostic criteria of each mood episode). The occurrence of one or more major depressive episodes may also be a feature. The mania associated with BPD should not be due to Substance-Induced Mood Disorder or Schizoaffective Disorder. During severe manic episodes violent or abusive behavior may be experienced. On average, untreated BPD patients experience four episodes each decade. There are no differences in BPD prevalence across gender and ethnicity. However, it has been noted that males are more likely to experience a manic episode before a depressive episode and vice-versa for females. Most individuals return to an improved or functional state between episodes [American Psychiatric Association. Task Force on DSM-IV, 1994].

1.2 Kraeplinian Dichotomy: BPD and Schizophrenia

Less than a hundred years ago, Emil Kraepelin classified Schizophrenia (SCZ) and BPD as being two separate disease entities, which he termed dementia praecox and manic-depression, respectively. In other words, each disorder has a different underlying cause [Craddock and Owen, 2005; Greene, 2007]. However, these two disease entities share common clinical symptoms [Maier et al., 2005] and there exists an intermediate phenotype, termed Schizoaffective Disorder [Lichtenstein et al., 2009]. Schizoaffective Disorder is characterized by features of both SCZ and BPD [Craddock et al., 2009]. Family history research has shown the occurrence of SCZ and BPD in the same family [Wildenauer et al., 1999]. Relatives of BPD patients have an increased risk for SCZ, including adopted children whose biological parents have SCZ [Lichtenstein et al., 2009].
This suggests a genetic/biological overlap between the two groups of disorders [Maier et al., 2005]. There is a greater amount of genetic information available pertaining to the aetiology of SCZ than there is for BPD [McGuffin et al., 2003].

**Table 1: Diagnostic criteria of mood episodes characteristic of BPD**

<table>
<thead>
<tr>
<th>Episode</th>
<th>DSM-IV Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Major Depressive Episode (MDE)</strong></td>
<td>a) Five or more of the following symptoms</td>
</tr>
<tr>
<td></td>
<td>• Depressed or irritable (children and adolescents) mood</td>
</tr>
<tr>
<td></td>
<td>• Decreased interest or pleasure in almost all activities</td>
</tr>
<tr>
<td></td>
<td>• Significant changes in body weight</td>
</tr>
<tr>
<td></td>
<td>• Insomnia or hypersomnia</td>
</tr>
<tr>
<td></td>
<td>• Psychomotor agitation</td>
</tr>
<tr>
<td></td>
<td>• Fatigue and decrease in energy</td>
</tr>
<tr>
<td></td>
<td>• Feelings of worthlessness and excessive guilt</td>
</tr>
<tr>
<td></td>
<td>• Diminished concentration and increased indecisiveness</td>
</tr>
<tr>
<td></td>
<td>• Increased thoughts of death and suicidal ideation</td>
</tr>
<tr>
<td></td>
<td>b) Symptoms do not meet criteria for mixed episode</td>
</tr>
<tr>
<td></td>
<td>c) Symptoms cause functional impairment</td>
</tr>
<tr>
<td></td>
<td>d) Symptoms not a result of physiological effect of substance</td>
</tr>
<tr>
<td></td>
<td>e) Symptoms not as a result of bereavement and are present for longer than two months</td>
</tr>
<tr>
<td><strong>Manic Episode</strong></td>
<td>a) Period of abnormally elevated or irritable mood lasting at least one week</td>
</tr>
<tr>
<td></td>
<td>b) Three or more of the following symptoms are present to a significant degree</td>
</tr>
<tr>
<td></td>
<td>• Inflated self-esteem or grandiosity</td>
</tr>
<tr>
<td></td>
<td>• Decreased need for sleep</td>
</tr>
<tr>
<td></td>
<td>• Talkative</td>
</tr>
<tr>
<td></td>
<td>• Racing thoughts</td>
</tr>
<tr>
<td></td>
<td>• Easily distractable</td>
</tr>
<tr>
<td></td>
<td>• Increased goal-directed activity or psychomotor agitation</td>
</tr>
<tr>
<td></td>
<td>• Excessive participation in pleasurable activities with potential for painful consequences</td>
</tr>
<tr>
<td></td>
<td>c) Symptoms do not meet criteria for mixed episode</td>
</tr>
<tr>
<td></td>
<td>d) Symptoms cause functional impairment</td>
</tr>
<tr>
<td></td>
<td>e) Symptoms not a result of physiological effect of substance</td>
</tr>
</tbody>
</table>
### Table 1: Diagnostic criteria of mood episodes characteristic of BPD continued

| Mixed Episode | a) Criteria met for both Major Depressive Episode and Manic Episode nearly every day during a one week period  
| b) Symptoms cause functional impairment  
| c) Symptoms not a result of physiological effect of substance |

| Hypomanic Episode | a) Period of abnormally elevated or irritable mood lasting at least four days  
| b) Three or more of the following symptoms are present to a significant degree  
• Inflated self-esteem or grandiosity  
• Decreased need for sleep  
• Talkative  
• Racing thoughts  
• Easily distractible  
• Increased goal-directed activity or psychomotor agitation  
• Excessive participation in pleasurable activities with potential for painful consequences  
| c) Change in functioning  
| d) Disturbance in mood observable by others  
| e) Episode not severe enough to caused marked functional impairment  
| f) Symptoms not a result of physiological effect of substance |

[American Psychiatric Association (DSM-IV), 2005]

### 1.3 Techniques Used to Determine the Aetiology of BPD

#### 1.3.1 Mouse Models

Mouse models have proven to be a reliable resource for geneticists. With the completion of the mouse genome sequence and the relative abundance of different mouse strains, the phenotypic effects of certain mutations are relatively easily observable [Seong et al., 2002]. Human disorders are modeled in animals via pharmacological or genetic...
manipulation [Le-Niculescu et al., 2008]. In general, an animal model needs to meet the following three criteria in order for it to be considered effective:

(a) **Face Validity** - the model should have similar behavioural patterns as human disease.

(b) **Construct Validity** - the model should have the same or similar underlying pathological characteristics as the human disease.

(c) **Predictive Validity** - Symptoms of the animal model that are characteristic of human disease should improve with administration of medication [Jornada et al., 2009].

An animal model of BPD would help in the understanding of the aetiology as well as provide a means to test treatment options. However, due to the cyclical nature of BPD, the innovation of an animal model, which meets all three criteria, has proven to be elusive. There are models which mimic certain characteristics of the disorder, such as depression and mania, but none are able to provide a mechanism by which mood states are alternated [Machado-Vieira et al., 2004]. Each of the existing BPD models can be categorized as either pharmacological, nutritional, environmental, or genetic [Kato et al., 2007]. An example of each category is discussed below:

i) **Pharmacological Model: Ouabain** - It has been shown that sodium-potassium-ATPase function is significantly reduced in individuals suffering from psychiatric disorders and that affected individuals appear to have increased intracellular levels of sodium and calcium [el-Mallakh et al., 1995], which suggests dysfunction of these ion transporters. Ouabain is an inhibitor
of sodium-potassium-ATPase and has been shown to induce a manic-like state in mice by increasing locomotor activity [el-Mallakh et al., 1995; Freitas et al., 2010].

ii) **Nutritional Model: Homocysteine** - Serum levels of homocysteine, a neurotoxin, has been found to be elevated in BPD patients. This is due to the effects of diet, inadequate vitamin B consumption, a decrease in physical exercise and an increase in smoking. In a mouse model, mild cognitive impairments were observed after chronic exposure to homocysteine [Levine et al., 2005].

iii) **Environmental Model: Learned Helplessness (LH)** - This mouse model is used to represent depression. This model, which is developed by treating the animal with electroshocks, displays symptoms of human depression and is unable to cope under difficult situations. The symptoms of this model are alleviated with the administration of antidepressants [Nakatani et al., 2004].

iv) **Genetic Model: Glucocorticoid Overexpression** - Glucocorticoid receptors have several functions including modulation of stress response and cognition. Overexpression of glucocorticoid receptors in the forebrain have shown an increase in anxiety-like and depressive symptoms in mice [Wei et al., 2004].
1.3.2 Human Gene Expression Data
Due to the nature of BPD, the brain would be the preferable tissue for gene expression studies. However, this may only be obtained from postmortem tissue which may present with a few disadvantages; for example, gene expression may be significantly altered due to death, tissue preservation methods [Matigian et al., 2007] or medication [Nakatani et al., 2006].

Nevertheless, Nakatani et al. [2006] performed a genome-wide expression study on seven postmortem BPD brains and seven control brains. They found 108 genes were differentially expressed in BPD brains and that these genes only had a minimal fold change. This is in keeping with the hypothesis that psychiatric illnesses are a manifestation of subtle changes in gene expression. Pathway analysis indicated that most of the genes could be classified as either forming part of the processes of cellular growth and proliferation, nervous system development, or cell death. In another expression study, also involving postmortem brains, pathways that were over-represented included the G-protein coupled receptor signaling pathway, stimulus response pathways, intracellular transport, and the ubiquitin cycle [Ryan et al., 2006]. Table 2 lists some of the expression studies that have been performed and the major pathways that have been implicated in each. There appears to be no consensus of results amongst these expression studies. This could be due to different tissue “quality” in each of the studies which in turn has an effect on tissue pH levels. For example, energy metabolism and mitochondrial genes are typically down-regulated in low pH conditions [Nakatani et al., 2006].
Table 2: Genes and pathways implicated in human expression studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Genes/Pathways implicated</th>
<th>Up-Downregulated</th>
<th>Tissue type</th>
<th>Specific Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tkachev et al. [2003]</td>
<td>Oligodendrocyte-related</td>
<td>↓</td>
<td>brain</td>
<td>Prefrontal Cortex</td>
</tr>
<tr>
<td></td>
<td>Myelin-related</td>
<td>↓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iwamoto et al. [2004]</td>
<td>Channels, receptors, transporters</td>
<td>↓</td>
<td>brain</td>
<td>Prefrontal Cortex</td>
</tr>
<tr>
<td></td>
<td>Stress response</td>
<td>↑</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iwamoto et al. [2005]</td>
<td>Mitochondria-related</td>
<td>↓</td>
<td>brain</td>
<td>Prefrontal Cortex</td>
</tr>
<tr>
<td>Nakatani et al. [2006]</td>
<td>Cellular growth, proliferation</td>
<td>↑↓</td>
<td>brain</td>
<td>Prefrontal Cortex</td>
</tr>
<tr>
<td></td>
<td>Nervous system development</td>
<td>↑↓</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cell death</td>
<td>↑</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ryan et al. [2006]</td>
<td>G-protein coupled receptor signaling</td>
<td>↑</td>
<td>brain</td>
<td>Prefrontal and Orbitofrontal Cortex</td>
</tr>
<tr>
<td></td>
<td>Response to stimulus</td>
<td>↑</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Intracellular transport</td>
<td>↓</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ubiquitin cycle</td>
<td>↓</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1.3.3 Human Blood Evidence
Currently, there are no available blood tests that can be used to determine the presence of a mood disorder in a patient. Diagnosis is strictly limited to clinical symptoms and patient self-report [Craddock and Sklar, 2009]. This presents with certain disadvantages to psychiatrists in obtaining an accurate diagnosis which is necessary for effective treatment and management. Therefore, discovery of appropriate peripheral tissue biomarkers for BPD would allow for greater accuracy in patient diagnosis. Unfortunately,
these types of studies have proven to be inconclusive and a definitive or other blood biomarker has yet to be discovered.

1.3.4 Genetics and Family Studies

Studies have shown that BPD has heritability estimates ranging from 59% to 85% [Lichtenstein et al., 2009; McGuffin et al., 2003]. First degree relatives and half-siblings of affected individuals are at a significantly increased risk (5-10%) of developing the disorder [Müller-Oerlinghausen et al., 2002; Lichtenstein et al., 2009]. Adoption studies have revealed that the biological relatives of a proband have a 31% higher risk of developing BPD than the relatives of control individuals. In a BPD twin study, it was shown that monozygotic and dizygotic twins have a 67% and 19% concordance, respectively [McGuffin et al., 2003]. Therefore, genetics is thought to hold a key to the understanding of BPD. However, efforts at elucidating the genetic cause of this disorder have been hampered by genetic heterogeneity, environmental factors and phenotypic variability that characterise BPD [Le-Niculescu et al., 2008]. As a result, researchers have adopted different approaches to determine the factors underlying the pathophysiology of BPD. One of these approaches involves the examination of intermediate phenotypes which are associated with the disorder under investigation but are easier to investigate genetically [Gottesman and Gould, 2003]. It is thought that discovering the genetic components of the more homogenous endophenotypes will aid in the understanding of genetically complex disorders like BPD. A particular trait needs to meet the following criteria to be considered as an endophenotype [Gottesman and Gould, 2003]:
• Is associated with a disorder in the population

• Is heritable

• Is evident in individuals whether the disorder is active or not

• Co-occurs with disease within families

• Manifests in non-affected family members of affected individuals at a higher frequency than in the general population

1.3.4.1 **Linkage and Association Studies**
The candidate approach investigates genes which have been reported to be linked to disease in previous research, or focuses on loci based on their purported role in the biology of the disorder. Linkage analysis involves the genotyping of genetic markers which are significantly associated with a disease phenotype. These approaches are more cost-effective than GWAS but require a hypothesis regarding the relationship of the gene to disease [Hirschhorn and Daly, 2005]. Regions on chromosomes 6, 7, 8, 10, 11, 12, 13, 15, 16, 17, and 22 have previously been reported to be linked to the BPDI phenotype [Ogden et al., 2004; Neves-Pereira et al., 2002; Sklar et al., 2002; Hattori et al., 2003; Chen et al., 2005; Schumacher et al., 2004; Fallin et al., 2005; Martucci et al., 2006; Avramopoulos et al., 2007].
1.3.4.2 Genome-Wide Association Studies
Over the last several years, researchers interested in mapping the genetic basis of several complex disorders have invested their resources in GWAS which do not require an *a priori* hypothesis. GWAS utilize high-throughput genotyping methods, on very large cohorts of subjects/patients and controls, to determine whether genetic variants, such as single nucleotide polymorphisms (SNPs) spread across the genome, have an association with disease status compared to equally large numbers of controls [Hirschhorn and Daly, 2005]. To date, eight GWAS have been performed on BPD cohorts; these are summarized in Table 3.

In 2007, The Wellcome Trust Case-Control Consortium (WTCCC) carried out a GWAS for seven common diseases, one of which was BPD [Stratton, 2007]. Their strongest signal of association for BPD was in the region of chromosome 16p12, which includes genes such as *PALB2* (*partner and localiser of BRCA2*), *NDUFAB1* (*NADH dehydrogenase (ubiquinone) 1, alpha/beta subcomplex, 1*), and *DCTN5* (*dynactin 5*) [Stratton, 2007]. The subsequent GWAS for BPD implicated genes such as *DGKH* (*diacylglycerol kinase*), *ANK3* (*ankyrin 3*), and *CACNA1C* (*Calcium Channel, Voltage-Dependent, L type, Alpha 1C Subunit*) [Baum et al., 2008; Ferreira et al., 2008]. However, the results from these studies have proven difficult to replicate in independent samples.
## Table 3: Summary of recent BPD Genome-Wide Association Studies

<table>
<thead>
<tr>
<th>Study</th>
<th>No. of Samples</th>
<th>Phenotype</th>
<th>Ethnicity</th>
<th>Type of Polymorphism</th>
<th>Regions/Genes Implicated</th>
<th>Genotyping Platform</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stratton, [2007]</td>
<td>2000 cases 3000 controls</td>
<td>BPD</td>
<td>Caucasian</td>
<td>SNPs</td>
<td>16p12</td>
<td>Affymetrix GeneChip 500K</td>
</tr>
<tr>
<td>Baum et al. [2008]</td>
<td>1233 cases 1439 controls</td>
<td>BPDI</td>
<td>Caucasian</td>
<td>SNPs</td>
<td>DGKH (13q14.11)</td>
<td>Illumina HumanHap550</td>
</tr>
<tr>
<td>Sklar et al. [2008]</td>
<td>1461 cases 2008 controls</td>
<td>BPDI</td>
<td>Caucasian</td>
<td>SNPs</td>
<td>MYO5B (18q21)</td>
<td>Affymetrix GeneChip 500K</td>
</tr>
<tr>
<td>Ferreira et al. [2008]</td>
<td>4387 cases 6209 controls</td>
<td>BPDI, BPDII, SABP*</td>
<td>Caucasian</td>
<td>SNPs</td>
<td>ANK3 (10q21)</td>
<td>Afymetrix GeneChip 500K</td>
</tr>
<tr>
<td>Zhang et al. [2008]</td>
<td>1001 cases 1033 controls</td>
<td>BPDI, SABP</td>
<td>Caucasian</td>
<td>SNPs and CNVs</td>
<td>DEC1 (9q32), IRF2 (4q34.1), VEPH1 (3q24), EXT11 (1p36.1), PCDH15 (10q21.1), MYT1L (2p25.3), CSMD1 (8p23.2), DYNC1I1 (7q21.3), PCSK5 (9q21.3), OLFM3 (1p22), PLXNA2 (1q32.2), KCNMB2 (3q26.2), INTS9 (8p21.1), AUTS2 (3q25), NUP107 (12q15)</td>
<td>Affymetrix SNP 6.0</td>
</tr>
<tr>
<td>Hattori et al. [2009]</td>
<td>1st stage: 107 cases 107 controls 2nd stage: 395 cases 409 controls</td>
<td>BPDI, BPDII</td>
<td>Japanese</td>
<td>SNPs</td>
<td>DEC1 (9q32), IRF2 (4q34.1), VEPH1 (3q24), EXTL1 (1p36.1), PCDH15 (10q21.1), MYT1L (2p25.3), CSMD1 (8p23.2), DYNC1I1 (7q21.3), PCSK5 (9q21.3), OLFM3 (1p22), PLXNA2 (1q32.2), KCNMB2 (3q26.2), INTS9 (8p21.1), AUTS2 (3q25), NUP107 (12q15)</td>
<td>Illumina Beadarray</td>
</tr>
<tr>
<td>Scott et al. [2009]</td>
<td>2076 cases 1676 controls</td>
<td>BPDI</td>
<td>Caucasian</td>
<td>SNPs</td>
<td>1p31.1, 3p21 (ITIH1, GNL3, NEK4, ITIH3), 5q15 (MCTP1),</td>
<td>Illumina HumanHap550</td>
</tr>
<tr>
<td>Smith et al. [2009]</td>
<td>1001 cases 1033 controls 345 cases 670 controls</td>
<td>BPDI, SABP</td>
<td>African American</td>
<td>SNPs</td>
<td>Xq27.1, NAP5</td>
<td>Affymetrix SNP 6.0</td>
</tr>
</tbody>
</table>

*Schizoaffective Disorder, Bipolar Type
1.3.5 Brain Imaging
Research within psychiatry has investigated the links between disease states, genotypes, psychological processes and brain structure and function. This effort has been facilitated by the development of recent neuro-imaging techniques [Dager et al., 2008]. Over the past 30 years, neuro-imaging studies have revealed that BPD is associated with structural changes in the brain [Vita et al., 2009]. However, different regions of the brain have been implicated and replication studies have proven to be inconclusive. Also, the causal reason, age of onset, and course of these brain abnormalities in affected individuals are still largely unknown [Vita et al., 2009]. Magnetic resonance imaging (structural–voxel based morphometry (VBM)) is a commonly used imaging technique that is discussed below in greater detail.

1.3.5.1 Magnetic Resonance Imaging (MRI)
This technique is able to provide a high resolution, three-dimensional view of soft tissues. The main advantage of this technique is its relative safety as it is non-invasive and does not involve any ionizing radiation [Leondes, 2005]. The MRI image is obtained by the use of a tomograph. The main constituent of this instrument is a large cylindrical coil which functions as a magnet. There are additional coils which form part of an electric circuit, whereby the electrical current within the magnet is able to produce a steady magnetic field. The subject is placed inside the cylindrical coil and exposed to the magnetic field [Hennel and Kryst-Widzgowska, 1997].
MRI relies on nuclear magnetic resonance (NMR), whereby protons and neutrons that constitute the nucleus of a molecule, have what is known as an angular momentum or spin [Leondes, 2005]. MRI involves the interaction between the nuclear spin of a proton within the nucleus of a hydrogen atom and the applied magnetic field. The human body consists mainly of lipids (~33%) and water (~70%) [Sheng and Huggins, 1979] hence, an abundance of hydrogen atoms in the brain [Leondes, 2005]. These “protons” once placed in a magnetic field, absorb radio waves (electromagnetic energy) at a certain frequency, and align themselves to the magnetic field by releasing the absorbed energy [Sajjad and Alam, 2007].

A meta-analysis of MRI studies for first episode BPD patients indicated that intracranial as well as white matter volumes were significantly lower in affected individuals compared to controls [Vita et al., 2009]. The authors propose that the apparent deficiency in white matter may indicate irregular neural development of the brain. Reduction in grey matter volumes has also often been found to be associated with BPD. Sassi et al. [2002] found that lithium treatment increased grey matter volumes compared to untreated patients. However, changes in grey matter volume could not be replicated in other studies [Vita et al., 2009; McDonald et al., 2005; Scherk et al., 2008]. These discrepancies in results could be due to the effects of medication and small sample sizes [van der Schot et al., 2009], suggesting a need for additional studies. Several studies have examined the relationship between hippocampal volumes and BPD [Blumberg et al., 2003; Rimol et al., 2010]. Findings from these studies suggest that individuals affected with BPD have reduced hippocampal volumes compared to controls. Similar to grey
matter volumes, treatment with lithium appeared to increase the hippocampal volumes of affected individuals compared to unmedicated individuals [Yucel et al., 2008].

### 1.3.5.2 Imaging Genetics

Certain brain volumes have been shown to be heritable [Goldman et al., 2008]. For example, twin studies have shown that frontal lobe structures and the hippocampus have heritabilities of 90-95% and 40-69%, respectively [Peper et al., 2007]. These findings indicate that genes may underlie the variation in brain structures seen across individuals.

van der Schot et al. [2009] examined the effects of genes and the environment on brain volumes in twin pairs concordant and discordant for BPD. They concluded that a decrease in white matter is associated with a genetic risk for developing BPD. This finding was also found in a previous twin study [Kieseppä et al., 2003]. Similarly, another twin study found that hippocampal volumes were smaller in the affected twin when compared to the unaffected co-twin [Noga et al., 2001]. It is thought that genes underlying BPD may be the cause for the observed decreases in white matter and hippocampal volumes. However, the biological mechanisms underlying these associations are still unknown.

Brain structural volumes and function have been investigated as endophenotypes of psychiatric illnesses [McDonald et al., 2004; Cannon and Keller, 2006; Glahn et al., 2007]. As endophenotypes are easier to investigate genetically [Gottesman and Gould, 2003], a
field which has received some attention of late is that of imaging genetics. Imaging genetics is a type of genetic association study in which genotypes are correlated with brain structure or function. This type of study is favoured as it gives a direct indication of the effect of underlying genetics on brain structure and function [de Geus et al., 2008] which may aid in the understanding of the genetics underlying complex psychiatric illnesses.

1.4 Convergent Functional Genomics
Due to the high heritability and familial pattern of inheritance of BPD, researchers have commonly focused on the genetics of this disorder. Many of these studies, as discussed previously, have focused on single gene approaches or expression studies which have, in turn, generated results implicating numerous genes. However, most of these findings could not be replicated particularly in those studies involving differing ethnic groups. With this abundance of data available it has become difficult to determine which finding would most likely be significant for follow-up studies. Therefore, researchers have focused their attention on a more comprehensive integrated approach to determine the genetic components of this and other psychiatric disorders. Le Niculescu et al. [2008] devised a gene prioritization method which they have termed CFG. CFG integrates data generated from GWAS together with results obtained from independent genomic investigations such as association or linkage studies, animal models and gene expression data. This approach aims to provide a means of prioritizing candidates for future hypothesis-driven studies [Le-Niculescu et al., 2008]. Using this approach, Le-Niculescu et al. [2008] generated an “optimal” candidate gene list for BPD.
1.5 Candidate Genes
All of the candidate genes and respective SNPs that have been investigated in the current study are described below and summarised in Table 4.

1.5.1 KLF12
This gene has not been functionally implicated in any psychiatric disorder but has been one of the genes that obtained the highest CFG score out of a total of 113 genes [Loculescu et al., 2008]. KLF12 is a member of a family of 16 kruppel-like factors (KLFs) that are DNA binding transcriptional regulators. Through their role as gene regulators, KLFs are involved in numerous cellular processes such as proliferation, apoptosis and development [Zhang et al., 2005]. KLF12 is located on chromosome 13q22.1 and is specifically involved in the repression of the AP-2 alpha gene [Roth et al., 2000]. AP-2 genes have been implicated in the pathophysiology of cancers as they play a role in tumour progression [Roth et al., 2000]. With this, KLF12 is also thought to play a role in the development of cancers. A recent study implicated KLF12 in gastric cancers, and proposed this gene as a target for therapeutic intervention [Nakamura et al., 2009]. Baum et al. [2008] found that an intronic SNP (rs9600160) within KLF12 was associated with BPD (p< 0.05).

1.5.2 ARNTL
ARNTL (also known as BMAL1), is located on chromosome 11p15.2 and encodes the aryl hydrocarbon nuclear translocator-like protein which forms a heterodimer with the protein CLOCK and is involved in circadian rhythm regulation [Wolting and McGlade, 1998; Hirota and Fukada, 2004]. The regulation of the circadian rhythm is accomplished
by the binding of the ARNTL/CLOCK complex to the E-box sequences of circadian genes such as *Period (Per)* and *Cryptochrome (Cry)*, thus promoting their transcription. Due to the cyclical nature of BPD, impairment in circadian rhythm patterns has often been implicated in the aetiology [Mansour et al., 2005]. Individuals undergoing a manic phase generally require less sleep or experience periods of insomnia, and those experiencing a depressive episode often have hypersomnia [Shi et al., 2008]. Sleep-deprived mice exhibit manic-like symptoms [Benedetti et al., 2008].

*ARNTL* has also been found to be up-regulated in post-mortem brains of BPD-I patients [Nakatani et al., 2006]. In a comprehensive study of BPD, Mansour et al. [2006] used separate study designs to identify predisposing regions of the genome. In the sub-cohort study, involving a case-parent trio study design, they found a positive association between two SNPs in *ARNTL* (rs7107287 and rs895682) and BPD-I. Both of these SNPs are located in intronic regions of the gene. The case-control study design also revealed an association between *ARNTL* and BPD-I patients. However, the SNP with the significant finding was rs1481892 and not the SNPs previously associated in the a case-parent trio study design [Mansour et al., 2006]. Haplotype analysis revealed an association between polymorphisms within *ARNTL* and another circadian gene, *period homolog 3 (PER3)*, with BPD. However, after correction for multiple testing this association was no longer significant [Nievergelt et al., 2006].
1.5.3 BDNF

BDNF is located on chromosome 11p13, is 70 kb in size and has 11 exons [Sheikh et al., 2010]. This gene has often been suggested in the aetiology of psychiatric disorders, as the protein it encodes is involved in the growth and maintenance of neurons within the CNS. Various extraneous factors such as stress and certain types of drugs affect the expression of this gene [Duman, 2002]. Interestingly, although this family of neurotrophins appear to be ubiquitously expressed in vertebrate nervous systems, similar sequences have not been detected in either of the C. elegans or D. melanogaster genomes suggesting that particular neurotrophic factors have functions in higher order neural activities such as learning, memory and behavior [Chao, 2000]. Several studies have investigated whether variants within BDNF increase risk for developing BPD. Of particular interest is the Val66Met (rs6265) functional polymorphism due to a G/A substitution [McHughen et al., 2010], in the prodomain of the protein [Bath and Lee, 2006]. Although, the prodomain is cleaved from the mature protein, it is thought to be involved in the trafficking of BDNF. The Val66Met polymorphism has also been found on to be associated with BPD in several independent studies [Neves-Pereira et al., 2002; Sklar et al., 2002; Xu et al., 2010]. Most of these studies have found that BPD patients have a higher frequency of the valine (Val) allele. Subsequent studies did not show an association between the Val66Met polymorphism and the BPD phenotype [Green et al., 2006; Vinberg et al., 2009]. However, Green et al. [2006] noted that the Val allele was significantly over-transmitted in individuals who presented with rapid cycling BPD. The Val66Met polymorphism has been shown to have an association with the human brain motor system and short-term learning in healthy individuals [McHughen et al., 2010]. The methionine (Met) allele has been associated with poorer episodic memory and
hippocampal activation [Egan et al., 2003]. Also, individuals affected with BPD, possessing a Met allele, show a reduction in their temporal lobe grey matter volumes compared to those without a Met allele [McIntosh et al., 2007]. Similarly, Met carriers in a control population group, are more likely to have smaller hippocampal formations, [Pezawas et al., 2004; Bueller et al., 2006] as are individuals affected with major depression [Frodl et al., 2007].

1.5.4 ALDH1a1

ALDH1A1, on chromosome 9q21.13, encodes an enzyme which is part of the aldehyde dehydrogenase family of proteins which consists of 16 distinct enzymes [Duester, 2001]. Aldehyde dehydrogenases are involved in the conversion of vitamin A (retinol) to its active form, retinoic acid, which in turn functions as a regulator of gene expression and plays an important role in the modulation of apoptosis [Palha and Goodman, 2006]. Retinoic acid has been postulated to play a role in the development of SCZ [Goodman, 1998; LaMantia, 1999]. This is based on the observation that certain abnormalities that are characteristic of SCZ are also a feature of retinoid dysfunction. Some common characteristics include mental deficits and congenital abnormalities such as craniofacial and digital malformations. It has also been observed that several loci that have been previously linked to SCZ contain genes that are involved in the retinoic acid cascade. These include the genes retinoid X receptor β (RXRβ) on chromosome 6p21.3 and the retinoic acid receptor β (RARβ) on chromosome 3p24.3 [Goodman, 1998]. Therefore, genes involved in the retinoic acid pathway are often investigated in relation to SCZ [Wan et al., 2009]. An association has previously been found between ALDH1A2
(aldehyde dehydrogenase type 2) and SCZ in a Chinese population group [Wan et al., 2009]. As BPD and SCZ are thought to have some genetic overlap [Maier et al., 2005], it would be reasonable to consider \textit{ALDH1A1} as a possible candidate gene for BPD.

\subsection*{1.5.5 \textit{A2BP1}}

\textit{A2BP1} or \textit{ataxin-2 binding protein 1}, predominantly expressed in both muscle and brain, encodes a protein which binds to the c-terminus of ataxin 2. A polyglutamine expansion in the gene \textit{ataxin-2} has been implicated the aetiology of spinocerebellar ataxia (SCA), a severe hereditary neurodegenerative disorder [Shibata et al., 2000]. The exact function of the \textit{A2BP1} gene is unknown [Ralser et al., 2005], however, it is thought to play a role in the splicing of downstream targets [Lee et al., 2009]. A translocation which results in a deletion in exon1 of the \textit{A2BP1} gene was previously observed in an individual with autism. This deletion resulted in decreased \textit{A2BP1} expression in the affected individual compared to controls [Martin et al., 2007]. Autism is a neuro-developmental disorder which shares some behavioural features with BPD and SCZ. Based on results from GWAS, it has been suggested, that these three psychiatric disorders, although distinct, share some genetic overlap [Carroll and Owen, 2009]. Hamshere et al. [2009] reported a strong association between a SNP in \textit{A2BP1} (rs4786811) and Schizoaffective Disorder. Also, two intronic SNPs within \textit{A2BP1}, rs7204975 and rs10500336 have been found to be significantly associated (p-values= 0.001 and 0.018, respectively) with BPD in a German population group by the GWAS conducted by Baum et al. [2008].
1.5.6 MBP
Myelin basic protein or MBP encodes a protein which composes approximately 30% of the myelin membrane surrounding neuronal tissue [Chambers and Perrone-Bizzozero, 2004]. This gene is located on chromosome 18q22-q23 [Kamholz et al., 1987]. Myelination continues within the human brain, with the culmination of myelinogenesis occurring in the frontal and temporal lobes in late adolescence and early adulthood [Chambers and Perrone-Bizzozero, 2004]. Interestingly, myelination of these lobes coincide with the supposed time of disease onset in SCZ patients [Chambers and Perrone-Bizzozero, 2004] which is also the age of disease onset for some of the individuals affected with BPD [Miller, 2006]. It is therefore likely that a relationship between myelination and the development of such psychiatric disorders might exist [Chambers and Perrone-Bizzozero, 2004]. Genome-wide microarray studies involving the post-mortem brains of BPD and SCZ patients have revealed significant differential expression of genes involved in myelination [Tkachev et al., 2003; Hakak et al., 2001]. Five genes which are involved in the development and maintenance of the myelin membrane were found to be downregulated in SCZ brains when compared to controls. Thus, abnormalities in myelination may be an aetiological factor for BPD based on the purported genetic overlap between BPD and SCZ.

The GWAS conducted by Sklar et al. [2008] found an association between an intronic SNP (rs12967023) in MBP and individuals affected with BPDI. This association has not been replicated in independent studies.
1.5.7 AK3L1

AK3L1 or adenylate kinase 3-like 1 (also known as AK3 and later renamed to AK4) encodes a member of the adenylate kinase family of enzymes. This family of kinases consists of five isoforms [Noma et al., 2001] and are primarily involved in the regulation of adenine and guanine bases [Yoneda et al., 1998]. The protein encoded by AK3L1 has been shown to be localised to the mitochondrial matrix and is classified as a guanosine triphosphate: adenosine monophosphate (GTP:AMP) phosphotransferase which is involved with the phosphorylation of mitochondrial AMP (required for cellular respiration) [Noma et al., 2001; Dzeja and Terzic, 2009]. Expression studies have revealed that AK3L1 is highly expressed in heart and skeletal muscle, liver and only weakly expressed in the brain [Noma et al., 2001]. AK3L1, on chromosome 1p31.3, has not previously been associated with any psychiatric disorder. However, the GWAS conducted by the WTCCC found that an intronic SNP within AK3L1 (rs4916031) was strongly associated with BPD (p-value 9.8 x 10^{-5}) [Stratton, 2007].

1.5.8 GSK3β

GSK3β encodes glycogen synthase kinase 3 beta which is involved in the phosphorylation of serine and threonine. This protein is involved in the WNT signalling pathway [Cruceanu et al., 2009] which has an important function in cell fate and development [Logan and Nusse, 2004]. Expression of genes from the WNT signalling pathway has been previously found to be dysregulated in affected individuals with BPD [Matigian et al., 2007]. In a genome-wide expression study, GSK3β was found to be downregulated in the post-mortem brains of individuals affected with BPD [Nakatani et al., 2006]. Similarly, GSK3β mRNA is significantly downregulated in the prefrontal cortex of
teenagers who have committed suicide (which is the most severe consequence of BPD) compared to normal controls [Pandey et al., 2009]. Recently, it was noted that polymorphisms within $GSK3\beta$ (rs6438552) are associated with regional grey matter volumes in patients with major depressive disorder [Inkster et al., 2009]. $GSK3\beta$ is also an important regulator of the proteins involved in the circadian rhythm pathway [Sahar et al., 2010]. Lithium, the first line treatment for BPD, as well as the mood-stabiliser, sodium valproate, are thought to inhibit $GSK3\beta$ [Klein and Melton, 1996; Nishiguchi et al., 2006]. The -50 T/C polymorphism (rs334558), located in the promoter region of $GSK3\beta$, has been found to have an association with response to lithium treatment for BPD and depression [Benedetti et al., 2005; Adli et al., 2007]. This polymorphism has also been associated with age of onset for BPD patients, which is an indicator of phenotype severity [Benedetti et al., 2004]. Therefore, $GSK3\beta$ seemed to be a reasonable candidate in the aetiology of BPD [Nishiguchi et al., 2006].

1.5.9 $ANK3$
This gene, located on chromosome 10q21, encodes the protein ankyrin 3. Ankyrins are thought to be involved in the linking of membrane proteins to the cell cytoskeleton and are involved in cellular processes such as motility and proliferation [Kordeli et al., 1995]. $Ankyrin 2$ ($ANK2$) obtained a CFG score of 5 in the research conducted by Le-Niculescu et al. [2008]. $ANK3$ is neural-tissue specific and is found at the Node of Ranvier where it is thought to play a role in maintenance of cell-adhesion molecules and ion channels [Kordeli et al., 1995]. The first study to implicate $ANK3$ as having an association with BPD was the GWAS done by Baum et al., [2008]. The SNP that was implicated, in that instance,
was rs9804190 which is located in an intronic region of the gene. Subsequently, a meta-
analysis of three GWAS found an association between a SNP within this gene (rs10994336) and BPD [Ferreira et al., 2008]. Furthermore, several other studies were able to replicate this finding in independent samples [Scott et al., 2009; Smith et al., 2009; Schulze et al., 2008].

1.5.10 CACNA1C
This gene encodes the alpha 1C subunit of the L-type voltage-gated calcium channel and is located on chromosome 12p13.3 [Lory et al., 1997]. Voltage-gated calcium channels are responsible for calcium influx, which in turn mediate cellular processes such as contraction, secretion, neurotransmission and gene expression. There are ten members of the voltage-gated calcium channel family, located in different cell types [Catterall et al., 2005]. Calcium channel, voltage dependent, alpha 1a subunit (CACNA1A) and calcium channel, voltage dependent, beta 2 subunit (CACNB2) had CFG scores of 6 and 5.5, respectively [Le-Niculescu et al., 2008]. CACNA1C has been shown to be associated with BPD in a meta-analysis involving 3 GWAS [Ferreira et al., 2008]. The particular SNP (rs1006737) which was associated in the study was also found to lead to an increased risk of developing major depression and SCZ [Green et al., 2009]. This SNP, located in an intron, has also been found to be a risk factor for developing a more severe depressive phenotype and insomnia [Casamassima et al., 2010]. In another study, rs1006737 was found to have an association with cerebral grey matter volume. Healthy individuals without any history of mental disorders, carrying the A allele of rs1006737 had larger grey matter volumes compared to individuals who were homozygous for the G-allele
[Kempton et al., 2009]. CANB1 (Calcium channel, voltage-dependent, beta-1 subunit) was shown to be upregulated in post-mortem BPDI brains [Nakatani et al., 2006]. The evidence therefore points to CACNA1C as a possible candidate gene for BPD.

1.6 Aims and Objectives
The first aim of the present investigation was to determine whether SNPs, within ten selected candidate genes, have an association with BPD. The second aim was to determine whether the genotypes of these variants were associated with brain imaging volumes. The objectives were the following:

i) To select a list of candidate genes with very strong evidence of involvement in BPD

ii) To identify a cohort of individuals from the BPD registry at the University of Cape Town (UCT)

iii) To genotype the candidate genes using a PCR-based technology

iv) To analyse the genotyping data with the appropriate statistical methods

v) To obtain brain imaging data of a subset of the larger cohort and to perform the appropriate statistical analysis
Table 4: Candidate genes from CFG and GWAS and respective SNPs

<table>
<thead>
<tr>
<th>Gene</th>
<th>Chromosomal Region</th>
<th>CFG Score</th>
<th>SNP</th>
<th>SNP Location</th>
<th>Base Change</th>
<th>Variation Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>KLF12</td>
<td>13q22.1</td>
<td>8</td>
<td>rs9600160</td>
<td>intron</td>
<td>G/A</td>
<td>Non-functional</td>
</tr>
<tr>
<td>ARNTL</td>
<td>11p15.2</td>
<td>8</td>
<td>rs7107287</td>
<td>intron</td>
<td>G/T</td>
<td>Non-functional</td>
</tr>
<tr>
<td>BDNF</td>
<td>11p14.1</td>
<td>8</td>
<td>rs6265</td>
<td>exon 2</td>
<td>C/T</td>
<td>Missense</td>
</tr>
<tr>
<td>ALDH1A1</td>
<td>9q21.13</td>
<td>8</td>
<td>rs1888202</td>
<td>intron</td>
<td>G/C</td>
<td>Non-functional</td>
</tr>
<tr>
<td>A2BP1</td>
<td>16p13.3</td>
<td>7.5</td>
<td>rs7204975</td>
<td>intron</td>
<td>C/T</td>
<td>Non-functional</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>rs10500336</td>
<td>intron</td>
<td>C/T</td>
<td>Non-functional</td>
</tr>
<tr>
<td>MBP</td>
<td>18q23</td>
<td>7.5</td>
<td>rs12967023</td>
<td>intron</td>
<td>C/T</td>
<td>Non-functional</td>
</tr>
<tr>
<td>AK3L1</td>
<td>1p31.3</td>
<td>7</td>
<td>rs4916031</td>
<td>intron</td>
<td>G/A</td>
<td>Non-functional</td>
</tr>
<tr>
<td>GSK3β</td>
<td>3q13.3</td>
<td>7</td>
<td>rs334558</td>
<td>promoter</td>
<td>C/T</td>
<td>Non-functional</td>
</tr>
<tr>
<td>ANK3</td>
<td>10q21</td>
<td>NA</td>
<td>rs9804190</td>
<td>intron</td>
<td>C/T</td>
<td>Non-functional</td>
</tr>
<tr>
<td>CACNA1C</td>
<td>12p13.3</td>
<td>NA</td>
<td>rs1006737</td>
<td>intron</td>
<td>G/A</td>
<td>Non-functional</td>
</tr>
</tbody>
</table>
Chapter 2: Materials and Methods

2.1 BPD Cohort Description
Ethics approval for the “Genetics of Bipolar Disorder” research project was first granted by UCT’s Research Ethics Committee in 1996 (REC Ref 081/1996). The aim of the project was to build a BPD archive, predominantly consisting of families, for the purpose of investigating the genetic complexities of the disorder. Between 1997 and 1999, individuals were recruited to the project from Gauteng, Free State, Eastern Cape and Western Cape. The cases were diagnosed with the Structured Clinical Interview for the DSM-IV (SCID) applied by a qualified psychiatric research nurse and verified by a psychiatrist. As the name suggests, SCID is a semi-structured questionnaire which allows for the diagnosis of DSM-IV axis disorders. Criteria for inclusion included an index patient and an available first-degree relative, both with a BPDI or BPDII diagnosis. The BPDI diagnosis was subsequently confirmed by a qualified psychiatrist. Interaction with healthcare professionals, advocacy groups and genealogical assessment resulted in the further recruitment of family members with a diagnosis of BPDI, BPDII, or recurrent MDE.

More recently, participants have been recruited from the outpatient unit at Valkenburg Hospital, Observatory, in Cape Town. All individuals were recruited when stable (not experiencing a manic or depressive episode), and blood samples were obtained with informed consent (Appendix 1).
2.1.1 Cohort Description for Case-Control Analysis
For the current study, updated ethics approval was obtained in June 2009 (REC Ref 229/2009). The research cohort, a subset of the larger BPD cohort, consisted of: i) an experimental group consisting of 224 individuals specifically diagnosed with BPDI (cases), and ii) 202 background population controls. The cases consisted of individuals of Mixed Ancestry and Caucasian ethnicity. A select number of cases were from the same family. These families were sampled for previous linkage studies. The inclusion of related cases in case control studies, if used appropriately, have been shown to increase study power as related individuals are thought to carry a higher frequency of disease risk alleles as compared to unrelated cases [Thornton and McPeek, 2007; Yoo et al., 2007].

The controls were selected from the Division of Human Genetics database. Most of these individuals were the unaffected spouses of individuals with other genetic disorders such as non-polyposis colorectal cancer and retinitis pigmentosa. All of the controls were unrelated and of unknown phenotype in terms of psychiatric disorders. The aim was to have a matched case–control design whereby controls were similar in terms of gender and ethnicity to case individuals. An advantage of having a matched case-control design is that a much smaller sample size is required to detect statistical significance [Mann, 2003] and that potential confounders such as gender and ethnicity are adjusted for [Niccolai et al., 2007]. A disadvantage is that if controls are too closely matched they do not represent the general population [Mann, 2003].

2.1.2 Cohort Description for Brain Imaging Analysis
The current study involved an investigation of a cohort of individuals affected with BPD, towards the identification of genes which may predispose to the condition. In addition, other modalities of neuropsychiatric investigations were undertaken as a means of
identifying potential endophenotypes which could be investigated genetically. In this regard, a subset of the BPD cohort was investigated by MRI brain imaging and genetics. The brain imaging analysis was done as part of the “Neuroimaging the Effects of Mindfulness Training in Bipolar Disorder” research project, carried out by the Department of Psychiatry (UCT). Separate ethics approval was obtained for this study (REC Ref 078/2009). For recruitment of participants, the study was advertised at the University of Cape Town, in newspapers and at specialized seminars. Participants were also recruited from the existing cohort and referrals from other health professionals. Imaging was performed for a subset (n=22) of the 224 BPDI case group.

2.2 DNA Isolation and Quality Determination
DNA was isolated, from blood samples, using the Gentra Puregene™ DNA Isolation Kit (Gentra Systems, Minneapolis, MN, USA). The protocol involved the following steps: the lysis of red blood cells (RBCs) and nucleated cells, protein precipitation, DNA precipitation, recovery, and rehydration. Following rehydration in DNA hydration solution, the isolated DNA was stored at either -20°C (for short-term storage) or -80°C (for long term storage). A more detailed protocol is included in Appendix 2.

The integrity and quality of the DNA samples were verified using two methods:

(i) Spectrophotometric quantification

(ii) Agarose gel electrophoresis

For quantification, the concentrations of undiluted, stock DNA samples were determined using a Nanodrop 1000 spectrophotometer (Nanodrop Technologies Inc., Wilmington, DE, USA). Samples were diluted with Sabax distilled water (Adcock Ingram,
Johannesburg, SA) to a working concentration of 100ng/µl. For agarose gel electrophoresis, a volume containing 200ng of DNA together with loading buffer (Fermentas Life Sciences, Hanover, USA) was loaded onto a 1% agarose gel (Appendix 3). To estimate integrity of DNA fragments, the first lane of the gel was loaded with a 100bp molecular weight marker (MWM) (Fermentas Life Sciences, Hanover, USA) (Appendix 3). To enable DNA visualization, 30ng of ethidium bromide (EtBr) (Sigma, England), a DNA intercalating agent, was added to the agarose prior to solidification. Gel electrophoresis was carried out in a gel tank with 1x Tris Borate EDTA (TBE) (Appendix 3) electrophoretic buffer, at 160V for approximately 30 minutes. Thereafter, it was placed in the Uvipro Gold transilluminator (UVITec, Cambridge, UK) for the illumination and image capture of the DNA fragments.

2.3 Gene and SNP Selection, Primer Design and Synthesis

2.3.1. Gene and SNP selection
In 2008, Le-Niculescu et al., based on a CFG approach (see Chapter 1: Introduction, pg 32), constructed a prioritised candidate gene list for BPD. The top eight genes on this list were chosen to be genotyped in the current study. Each of these genes obtained a CFG score of seven and above (12 being the maximum). The CFG score is based on previous evidence, therefore SNPs were chosen based on their significant association to BPD in independent studies. The remaining two genes, ANK3 and CACNA1C and their respective SNPs, were chosen based on their association to BPD in a GWAS [Baum et al., 2008; Ferreira et al., 2008]. Table 5 contains all of the candidate genes, their respective SNPs and CFG scores.
Table 5: Selected candidate genes, respective SNPs and CFG scores

<table>
<thead>
<tr>
<th>Gene</th>
<th>SNP</th>
<th>CFG Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>KLF12</td>
<td>rs9600160</td>
<td>8</td>
</tr>
<tr>
<td>ARNTL</td>
<td>rs7107287</td>
<td>8</td>
</tr>
<tr>
<td>BDNF</td>
<td>rs6265</td>
<td>8</td>
</tr>
<tr>
<td>ALDH1A1</td>
<td>rs1888202</td>
<td>8</td>
</tr>
<tr>
<td>A2BP1</td>
<td>rs7204975</td>
<td>7.5</td>
</tr>
<tr>
<td></td>
<td>rs10500336</td>
<td></td>
</tr>
<tr>
<td>MBP</td>
<td>rs12967023</td>
<td>7.5</td>
</tr>
<tr>
<td>AK3L1</td>
<td>rs4916031</td>
<td>7</td>
</tr>
<tr>
<td>GSK3β</td>
<td>rs334558</td>
<td>7</td>
</tr>
<tr>
<td>ANK3</td>
<td>rs9804190</td>
<td></td>
</tr>
<tr>
<td>CACNA1C</td>
<td>rs1006737</td>
<td></td>
</tr>
</tbody>
</table>

2.3.2 External Primer Design

The sequences of each of the ten candidate genes (KLF12, ARNTL, BDNF, ALDH1A1, A2BP1, MBP, AK3L1, GSK3β, ANK3, and CACNA1C) were obtained from NCBI (http://www.ncbi.nlm.nih.gov) and Ensembl (http://www.ensembl.org/index.html) databases. The sequence of each of the genes was then annotated to specify the location of the exons, coding sequence, start/stop codons and SNPs. This was accomplished using the ANNOTV9 annotation programme [Rebello, 2006, Unpublished], which utilises the Perl scripting language [Wall, 1987].

Once the regions of interest were identified, suitable primers were designed to amplify these sequences. The web-based program, Primer3 (http://fokker.wi.mit.edu/primer3/input.htm), which searches for possible primer sequences from the nucleotide sequence, was used for this purpose. The selected primers were then analysed by Integrated DNA Technologies (IDT) oligo analyser.
(http://www.idtdna.com/analyzer/Applications/OligoAnalyzer), a program which predicts any secondary structures such as primer dimers and hairpins which may occur. Secondary structures greatly reduce the efficiency of the PCR reaction [Singh et al., 2000]. To ensure the primer only binds to the sequence of interest, the primer sequences were checked against a database, NCBI BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi). All the primer sequences are contained in Table 6.

The oligonucleotides were synthesised using the Oligo 1000M DNA Synthesiser v4.20 (Beckman Instruments Inc., Department of Molecular and Cell Biology, UCT). The primer stock solutions were diluted with Sabax distilled water (dH₂O) (Adcock Ingram, Johannesburg, SA) to a working concentration of 20µM.

2.4 Polymerase Chain Reaction (PCR)

The target regions of interest in each of the candidate genes were amplified using PCR. To obtain the optimal annealing temperature (Ta), a temperature gradient was used to assay all of the primer pairs. The PCR reaction consists of deoxynucleotide triphosphates (dNTPs), Taq Polymerase, Taq Buffer, DNA, forward primer, reverse primer and distilled water (dH₂O). Based on the optimal Ta for each primer pair and amplicon size, certain primer sets were multiplexed together in one PCR reaction. Multiplexing of primers allows for the efficient usage of reagents and resources. A negative (water) control was included, which consisted of all the reaction components except DNA, to ensure that any product visualized was not due to contamination.
Table 6: External primer sequences for amplification of candidate gene regions using PCR

<table>
<thead>
<tr>
<th>Gene</th>
<th>Primer Direction</th>
<th>Primer Length</th>
<th>Primer Sequence (5'→3')</th>
<th>GC content (%)</th>
<th>Tm (°C)</th>
<th>Amplicon size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>KLF12</td>
<td>Forward</td>
<td>18</td>
<td>acgtagcagcaccagtct</td>
<td>55.6</td>
<td>56.2</td>
<td>180</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>21</td>
<td>gagagagtagcaagccaaatg</td>
<td>47.6</td>
<td>53.4</td>
<td></td>
</tr>
<tr>
<td>ARNTL</td>
<td>Forward</td>
<td>18</td>
<td>gaagccatgagctagca</td>
<td>50</td>
<td>51.5</td>
<td>237</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>18</td>
<td>gtagaaaccctggacagtt</td>
<td>50</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td>BDNF</td>
<td>Forward</td>
<td>24</td>
<td>caagagcttgacatctggctg</td>
<td>50</td>
<td>58.5</td>
<td>338</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>24</td>
<td>cgcacagctcactgactccttt</td>
<td>50</td>
<td>59.6</td>
<td></td>
</tr>
<tr>
<td>ALDH1A1</td>
<td>Forward</td>
<td>19</td>
<td>tctatagttgtgacccac</td>
<td>47.4</td>
<td>50.2</td>
<td>394</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>18</td>
<td>ttttcagctacgccatg</td>
<td>50</td>
<td>52.3</td>
<td></td>
</tr>
<tr>
<td>A2BP1(A)</td>
<td>Forward</td>
<td>17</td>
<td>tctatagttgtgacccac</td>
<td>47.1</td>
<td>47.7</td>
<td>192</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>19</td>
<td>ttttcagctacgccatg</td>
<td>42.1</td>
<td>47.5</td>
<td></td>
</tr>
<tr>
<td>A2BP1(B)</td>
<td>Forward</td>
<td>20</td>
<td>cgtggtctctctcttgctc</td>
<td>50</td>
<td>52.8</td>
<td>250</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>20</td>
<td>ggcaacaagctatactcgc</td>
<td>50</td>
<td>54.5</td>
<td></td>
</tr>
<tr>
<td>MBP</td>
<td>Forward</td>
<td>18</td>
<td>cttggtctcagtaagcc</td>
<td>55.6</td>
<td>52.9</td>
<td>170</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>19</td>
<td>gaatctgctcctggttgttg</td>
<td>52.6</td>
<td>53.6</td>
<td></td>
</tr>
<tr>
<td>AK3L1</td>
<td>Forward</td>
<td>20</td>
<td>cagactggtcgttcacccctc</td>
<td>55</td>
<td>55.3</td>
<td>155</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>20</td>
<td>ggtagaggttgagacgctgg</td>
<td>55</td>
<td>54.3</td>
<td></td>
</tr>
<tr>
<td>GSK3β</td>
<td>Forward</td>
<td>22</td>
<td>cgatcccgagcgtctgtgctg</td>
<td>59.1</td>
<td>60</td>
<td>302</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>18</td>
<td>tcaggaagttgcgcgctg</td>
<td>61.1</td>
<td>59.4</td>
<td></td>
</tr>
<tr>
<td>ANK3</td>
<td>Forward</td>
<td>18</td>
<td>gccagtaacaatgaaacc</td>
<td>50</td>
<td>51.3</td>
<td>227</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>21</td>
<td>gctatactctgttcggattc</td>
<td>47.6</td>
<td>51.8</td>
<td></td>
</tr>
<tr>
<td>CACNA1C</td>
<td>Forward</td>
<td>18</td>
<td>gttccattcactctcagc</td>
<td>50</td>
<td>50.5</td>
<td>220</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>18</td>
<td>aatgtggtggacagcaat</td>
<td>44.4</td>
<td>52</td>
<td></td>
</tr>
</tbody>
</table>
PCR was carried out on a PxE 0.2 Thermal cycler (Thermohybaid, Middlesex, UK) at the conditions stipulated in Table 7. Once completed, the PCR products were maintained at a temperature of 4°C until required for downstream processes.

To determine whether amplification had occurred, the PCR products together with loading buffer were resolved on a 2% agarose gel (Appendix 3), run at 160V, for approximately 40 minutes. The current was set at the maximum value (approximately 110 MA). MWM (100bp) (Appendix 3) was added to the first lane of the gel. The gel was exposed to UV light in the Uvipro Gold transilluminator (UVITec, Cambridge, UK) and the resultant image captured.

<table>
<thead>
<tr>
<th>Table 7: PCR cycling conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>3</td>
</tr>
</tbody>
</table>

2.5 SNaPshot PCR

SNaPshot PCR is a SNP genotyping method and is an extension of the basic PCR process. Internal primers were designed to bind 5’ or 3’ of the SNP of interest in each of the amplified candidate genes. The use of this technique requires SNaPshot™ multiplex ready reaction mix (Applied Biosystems, Warrington, UK) which contains fluorescently-labelled dideoxynucleotide triphosphates (ddNTPs), AmpliTaq polymerase, and reaction
buffer. The *Taq* polymerase catalyses the single base extension of the unlabelled primers by incorporating the complementary differentially-labelled ddNTP [ABI Prism SNaPshot™ Multiplex Kit Protocol, Applied Biosystems] (See **Figure 1**). The ddNTPs lack an acceptor 3’ hydroxyl group. Therefore, the AmpliTaq polymerase is unable to add an additional base to the ddNTP and extension is terminated. The end results of this technique are peaks which represent the alleles at a particular SNP. **Table 8** indicates the dyes attached to each of the ddNTPs and the resulting peak colours.

![Figure 1: An illustration of the SNaPshot PCR principle.](image)

**Table 8: Dyes assigned to each of the ddNTPs**

<table>
<thead>
<tr>
<th>ddNTP</th>
<th>Dye</th>
<th>Colour</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adenine (A)</td>
<td>dR6G</td>
<td>Green</td>
</tr>
<tr>
<td>Cytosine (C)</td>
<td>dTAMRA™</td>
<td>Black</td>
</tr>
<tr>
<td>Guanine (G)</td>
<td>dR110</td>
<td>Blue</td>
</tr>
<tr>
<td>Thymine (T)</td>
<td>dROX™</td>
<td>Red</td>
</tr>
</tbody>
</table>
2.5.1 SNaPshot PCR Internal Primer Design

The criteria for SNaPshot internal primers were the same as those stipulated for the external primers except that only one primer (either forward or reverse) was needed for SNaPshot PCR. The primers were designed by choosing consecutive bases immediately 5’ or 3’ of the SNP of interest. However, for the purposes of multiplexing, each of the primers needed to be of a different length to ensure that the appropriate alleles could be called for the respective gene and there would be no overlap in the SNaPshot product. Therefore, random tails of nucleotide repeats were attached to each of the primers. Another way of ensuring that primers could be multiplexed, even if two primers were the same length, was to ensure that those primers had different expected peak colours. A local BLAST search, using the sequence alignment program, BioEdit Sequence Alignment Editor 7.0.0 (Tom Hall, Isis Pharmaceuticals, Inc.), was done to ensure that the primers did not hybridise to one another. The sequences of the internal primers are displayed in Table 9.

Internal primers (25nmole) were synthesized by IDT® (http://eu.idtdna.com/order/order.aspx) and were received in a lyophilized form. To prevent contamination, each of the primers were desalted and quality was tested by mass spectrometry. Subsequently, the primers were resuspended with TE (10mM Tris pH8.0, 1mM EDTA) to a stock concentration of 100µM. Working solutions of 20µM primer concentration, in a total volume of 50µl, were made by the addition of Sabax dH₂O (Adcock Ingram, Johannesburg, SA).
Table 9: SNaPshot internal primer sequences for genotyping of SNPs within candidate genes

<table>
<thead>
<tr>
<th>Gene</th>
<th>SNP</th>
<th>Primer Direction</th>
<th>Primer Sequence</th>
<th>Expected</th>
<th>Primer Length</th>
<th>Peak Colours</th>
</tr>
</thead>
<tbody>
<tr>
<td>KLF12</td>
<td>rs9600160</td>
<td>Forward</td>
<td>gagtggtaggaattgattagatgt</td>
<td>A/G</td>
<td>24</td>
<td>Green/Blue</td>
</tr>
<tr>
<td>ARNTL</td>
<td>rs7107287</td>
<td>Forward</td>
<td>agctagcaatctgatgtgtt</td>
<td>G/T</td>
<td>20</td>
<td>Blue/Red</td>
</tr>
<tr>
<td>BDNF</td>
<td>rs6265</td>
<td>Reverse</td>
<td>¹ctctctctctctccacagctctcttatca</td>
<td>C/T</td>
<td>30</td>
<td>Black/Red</td>
</tr>
<tr>
<td>ALDH1A1</td>
<td>rs1888202</td>
<td>Forward</td>
<td>tagtagctgattaagaagtctggttaa</td>
<td>C/G</td>
<td>30</td>
<td>Black/Blue</td>
</tr>
<tr>
<td>A2BP1(A)</td>
<td>rs7204975</td>
<td>Forward</td>
<td>cacacacacacacacacacacatagctagcagaggaaca</td>
<td>C/T</td>
<td>35</td>
<td>Black/Red</td>
</tr>
<tr>
<td>A2BP1(B)</td>
<td>rs10500336</td>
<td>Reverse</td>
<td>gcctgttaactcactatctaa</td>
<td>C/T</td>
<td>23</td>
<td>Black/Red</td>
</tr>
<tr>
<td>MBP</td>
<td>rs12967023</td>
<td>Reverse</td>
<td>catcattcagccgcatagtttt</td>
<td>C/T</td>
<td>25</td>
<td>Black/Red</td>
</tr>
<tr>
<td>AK3L1</td>
<td>rs4916031</td>
<td>Reverse</td>
<td>gaacaggaagctctaaag</td>
<td>A/G</td>
<td>19</td>
<td>Green/Blue</td>
</tr>
<tr>
<td>GSK3β</td>
<td>rs334558</td>
<td>Forward</td>
<td>ttagtagtagtagtagttctcagacagcgc</td>
<td>C/T</td>
<td>35</td>
<td>Black/Red</td>
</tr>
<tr>
<td>ANK3</td>
<td>rs9804190</td>
<td>Forward</td>
<td>cctcaagaaggttaagaaa</td>
<td>C/T</td>
<td>19</td>
<td>Black/Red</td>
</tr>
<tr>
<td>CACNA1C</td>
<td>rs1006737</td>
<td>Forward</td>
<td>tagtagtagtagtagtagcattcactccagccgaa</td>
<td>A/G</td>
<td>40</td>
<td>Green/Blue</td>
</tr>
</tbody>
</table>

¹Red highlighted sequences indicate tails
2.5.2 Sap and Exol Clean-up of PCR Products

Before SNaPshot PCR could be carried out, unincorporated dNTPS and primers from the initial PCR needed to be removed, to ensure that there was no interference with the genotyping process [ABI Prism SNaPshot™ Multiplex Kit Protocol, Applied Biosystems]. Shrimp alkaline phosphatase (SAP) (Promega, USA) allows for the dephosphorylation of the 5’ phosphates of DNA. Exonuclease I (Exol) (Biolabs, New England) catalyses the removal of bases from single-stranded DNA in a 3’-5’ direction, promoting the degradation of the excess single-stranded primers (http://www.neb.com/nebecomm/products/productm0293.asp).

The purification process was carried out on an Applied Biosystems thermal cycler (Gene Amp® PCR System 9700) at the conditions listed in Table 10 and using the reagents in Table 11. Firstly, a master mix was made up with the SAP, Exol and dH₂O and aliquoted into each well of the segmented 96-well PCR Microplate (Axygen Scientific, California, USA). The PCR product was then aliquoted into each respective well. Wells were sealed with PCR cap strips (Axygen Scientific, California, USA).

<table>
<thead>
<tr>
<th>Temperature (°C)</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>37</td>
<td>1 hour</td>
</tr>
<tr>
<td>75</td>
<td>15 minutes</td>
</tr>
</tbody>
</table>

Table 10: Conditions for Sap and Exol purification
### Table 11: Sap and ExoI DNA purification

<table>
<thead>
<tr>
<th>Reagents</th>
<th>Recommended</th>
<th>Concentration</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SAP</strong></td>
<td></td>
<td>5 units</td>
<td>1 unit</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.5 units</td>
</tr>
<tr>
<td><strong>ExoI</strong></td>
<td></td>
<td>2 units</td>
<td>2 units</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2 units</td>
</tr>
<tr>
<td><strong>PCR</strong></td>
<td></td>
<td>-</td>
<td>5 µl</td>
</tr>
<tr>
<td><strong>Product</strong></td>
<td></td>
<td></td>
<td>5 µl</td>
</tr>
<tr>
<td><strong>dH₂O</strong></td>
<td></td>
<td>-</td>
<td>13.9 µl</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>13.4 µl</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td>-</td>
<td>20 µl</td>
</tr>
<tr>
<td><strong>Volume:</strong></td>
<td></td>
<td></td>
<td>20 µl</td>
</tr>
</tbody>
</table>

#### 2.5.3 SNaPshot PCR Conditions

Multiplex SNaPshot reactions, whereby an equal ratio of the respective primers were pooled, were carried out on 96-well PCR Microplates (Axygen Scientific, California, USA). For each SNaPshot reaction, a negative water control was also included. The control was included to ensure that the results obtained from the sample reactions were not a result of any form of contamination. This control contained all the components of the multiplex SNaPshot reaction but lacked any purified PCR sample. Table 12 illustrates the volumes of the components of a SNaPshot reaction. The samples were loaded into a Applied Biosystems thermal cycler (Gene Amp® PCR System 9700) and run at the conditions defined in Table 13.
Table 12: The components of the SNaPshot PCR reaction

<table>
<thead>
<tr>
<th>Reagents</th>
<th>Reaction Volumes (µl)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Singleplex Reactions</td>
</tr>
<tr>
<td>Reaction Mix</td>
<td>1</td>
</tr>
<tr>
<td>Internal Primers</td>
<td>1</td>
</tr>
<tr>
<td>PCR Product</td>
<td>1</td>
</tr>
<tr>
<td>dH₂O</td>
<td>7</td>
</tr>
<tr>
<td>Total Volume:</td>
<td>10</td>
</tr>
</tbody>
</table>

Table 13: Conditions for SNaPshot PCR

<table>
<thead>
<tr>
<th>Stage</th>
<th>Temperature (°C)</th>
<th>Duration</th>
<th>No. of Cycles</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>96</td>
<td>10 seconds</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>5 seconds</td>
<td></td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>30 seconds</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>∞</td>
<td></td>
</tr>
</tbody>
</table>

2.5.4 Post-SNaPshot SAP Treatment

Before the samples were genotyped, all the unincorporated ddNTPs needed to be removed by the phosphatase action of SAP. If left untreated, the ddNTPS would co-migrate with the products of interest. Thus, the removal of the phosphatase groups of these nucleotides hindered the movement and prevented any interference [ABI Prism SNaPshot™ Multiplex Kit Protocol, Applied Biosystems].
One unit of SAP was added to the SNaPshot PCR products. The samples were then loaded into a Applied Biosystems thermal cycler (Gene Amp® PCR System 9700) at the conditions listed in Table 10.

2.5.5 Capillary Electrophoresis on ABI PRISM
Purified SNaPshot reactions were resolved on the ABI prism 3100 Genetic Analyzer (Applied Biosystems, CA, USA). A mixture of SNaPshot product, HiDi formamide (Applied Biosystems, Warrington, UK) and GeneScan-120 Liz size standard (Applied Biosystems, CA, USA) (Appendix 4) were loaded into a 96-well PCR Microplate (Axygen Scientific, California, USA). The HiDi functions to maintain the double-stranded DNA in a denatured state and the Liz allows for the sizing of the allele peaks. The dye set was set at E5, the analysis method at GeneScan-120 Liz and the run module at SNP36_POP4.

2.5.6 Analysis of Results
The genotyping results were analysed by the Genemapper 3.0 Genescan Software (Applied Biosystems, CA, USA). This software provides automated sizing of peaks and was used for the analysis of each of the SNPs. It generated a table with all the alleles and user-specified process-based quality values which relayed the integrity of the data. Bins which corresponded to a particular base-pair range for a specific allele were defined. Software algorithms determined whether peaks were representative of an allele at that position.

2.6 Genotype Validation
The results obtained from the SNaPshot assay were validated by cycle sequencing.
2.6.1 PCR of DNA Samples for Cycle Sequencing
A subset of the samples were chosen based on the results shown by SNaPshot PCR. These contained all possible genotypes (homozygous and heterozygous) for each of the 11 SNPs. PCR was carried out on the selected samples using the same cycling conditions as in Table 7. Amplified products were resolved on a 2% agarose gel (Appendix 3) together with loading buffer. MWM (Appendix 3) was added to the first lane of the gel. The gel was run for approximately 40 minutes and placed in the Uvipro Gold transilluminator (UVITec, Cambridge, UK) for visualisation and image capture purposes. PCR products of adequate band intensity were used for the sequencing reaction.

2.6.2 Cycle Sequencing
Cycle sequencing was done on amplified products using the reagents listed in Table 14. Cycle sequencing was carried out on an Applied Biosystems thermal cycler (Gene Amp® PCR System 9700) at the conditions displayed in Table 15. Most SNPs were sequenced using the forward primer. However, ARNTL rs7107287, BDNF rs6265 and CACNA1C rs1006737 were sequenced using the reverse primer as these SNPs were located towards the 5’ end of the respective amplicons. Typically, peaks at the start (5’end) of a sequence are undistinguishable. Therefore, when sequenced with the forward primers, the ARNTL rs7107287, BDNF rs6265 and CACNA1C rs1006737 SNPs were not identifiable.

2.6.3 Ethanol Precipitation
The DNA was purified from the reaction mixture by precipitation with ethanol. This step removed any unincorporated ddNTPs which could potentially interfere with the start of the sequence and base-calling. The steps listed in Appendix 5 were followed.
Table 14: Components of the cycle sequencing reaction

<table>
<thead>
<tr>
<th>Reagents</th>
<th>Recommended Concentration</th>
<th>Volume (µl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BigDye Terminator Mix</td>
<td>2.5x</td>
<td>1</td>
</tr>
<tr>
<td>Sequencing Buffer</td>
<td>5x</td>
<td>4</td>
</tr>
<tr>
<td>Forward/Reverse Primer</td>
<td>-</td>
<td>0.5</td>
</tr>
<tr>
<td>PCR product</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>dH₂O</td>
<td>-</td>
<td>12.5</td>
</tr>
<tr>
<td>Total Volume</td>
<td></td>
<td>20</td>
</tr>
</tbody>
</table>

Table 15: Cycle sequencing PCR conditions

<table>
<thead>
<tr>
<th>Temperature (°C)</th>
<th>Duration</th>
<th>No. of Cycles</th>
</tr>
</thead>
<tbody>
<tr>
<td>96</td>
<td>5 min</td>
<td></td>
</tr>
<tr>
<td>96</td>
<td>30 sec</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>15 sec</td>
<td>25</td>
</tr>
<tr>
<td>60</td>
<td>4 min</td>
<td></td>
</tr>
</tbody>
</table>

2.6.4 Capillary Electrophoresis on ABI Prism
The sequencing samples were run on the ABI Prism 3130XL Genetic Analyzer (Applied Biosystems, CA, USA). A volume of 5µl of the sequencing sample was added to 8µl HiDi Formamide (Applied Biosystems, Warrington, UK).
2.6.5 Analysis of Results
The sequences of each of the candidate genes were visualised in the BioEdit Sequence Alignment Editor version 7.0.0 (Tom Hall, Isis Pharmaceuticals, Inc.). This allowed for the identification of the genotype of the SNP of interest.

2.7 Statistical Analysis

2.7.1 Genotype Frequencies, Hardy-Weinberg Equilibrium Calculations and Logistic Regression with R
Genotype frequencies for all of the SNPs (cases and controls), Hardy Weinberg Equilibrium (HWE) calculations (controls only), and case-control association were determined using the statistical environment R (http://www.R-project.org). The method used to determine whether there was an association between disease status and genotype frequency, was logistic regression. This method fits a statistical model to a data set whereby the dependent variable is dichotomous and the independent variable is either continuous or discrete. In this instance, the dependent variable was disease affection status (i.e. case or control) and the independent variable were the three genotypes for a particular SNP. The logistic regression test computes the odds ratio which is defined as the probability of an event occurring over the probability of the event not occurring [Munro, 2001]. The z-test value is the test statistic and lower and upper confidence intervals are specified. As more than one statistical analysis was performed (disease status versus SNP), the Bonferroni correction was implemented for multiple testing, to avoid spurious significance. Therefore, the “corrected” significance level was set at the initial alpha value (0.05) divided by the number of tests performed.
2.7.2 \textit{MQLS (adjusting for relatedness) and Pedstats}

A number of the cases belonged to the same family. This presented both an advantage and challenge in the statistical analysis. The advantage is that affected individuals within the same family are expected to have a higher frequency of the disease causing allele which may be easier to detect than in unrelated cases. Therefore, the power of the association is increased as well as the effective sample size. The challenge is that the relatedness amongst cases needs to be adjusted for in the analysis [Thornton and McPeek, 2007]. The M\textsubscript{QLS} method is able to perform basic case-control analysis with the addition of adjusting for related individuals. For this analysis, genotyping results of affected individuals were incorporated into a pedigree file to specify relationships within families. To ensure quality of pedigree data, i.e. whether there were any errors within the pedigree file, the program Pedstats was used [Wigginton and Abecasis, 2005].

2.8 Neuroimaging of Subset of Cohort

T1 weighted MRI scans were obtained using a high resolution Siemens Magneton Allegra 3T scanner at the Cross Universities Brain Imaging Centre (http://sun.ac.za/cubic). Hippocampal and total intracranial volumes, from 22 individuals affected with BPDI, were obtained using the automated software package Freesurfer (version 5.0.0) [Fischl et al., 2002] (http://surfer.nmr.mgh.harvard.edu/) running at the Center for High Performance Computing in Cape Town (http://chpc.ac.za). Hippocampal volumes were normalised to total intracranial volume to correct for individual differences in anatomy.
2.8.1 Statistical Analysis of Neuroimaging Results

Analysis of variance (ANOVA) was used to determine whether there were any associations between imaging measures and genotypes. The null hypothesis for ANOVA is that the means between two or more groups are the same, therefore, the alternate hypothesis is that means between groups are significantly different. This statistical measure tests whether the “between group” variation is greater than the “within group” variation. If this is found to be true, it suggests that the means between groups are different. The F-value represents the ratio of the “between group” and “within group” variance. ANOVA is a fairly robust statistical test, requiring that the dependent variable be continuous (e.g. imaging measure) and the independent variables be categorical (e.g. SNP genotype). ANOVA has the added advantage of being able to include more than one independent variable in a single test. The use of ANOVA requires that three assumptions be met, namely, i) the dependent variable is normally distributed, ii) there is independence between groups, iii) and variances are equal (homogeneity of variance) [Munro, 2001]. The Shapiro-Wilk test was used to ensure a normal distribution for the following measures: left and right hemisphere hippocampal volumes and intracranial volumes. In instances where the data was non-normally distributed, the non-parametric Kruskal-Wallis test was utilised, which compares the mean ranks in each group [Munro, 2001]. The Bartlett’s Test was used to ensure that variances were equal across genotype groups.
Chapter 3: Results

3.1 BPD Cohort Description
To date, the BPD cohort consists of 919 individuals from 221 families. A total of 550 females and 366 males (three of unknown gender) have been recruited. The cohort consists of individuals from the Caucasian (British, Afrikaner, Portuguese, Greek, and Ashkenazi Jewish), Mixed Ancestry, Indian and African population groups. Figure 2 represents the distribution of these ethnic groups within the cohort. The most prevalent diagnoses present in the archive includes mood disorders (BPDI, BPDII, MDE) (with BPDI being the primary diagnosis), Borderline Personality Disorder (PD), Post Traumatic Stress Disorder (PTSD) and Alcohol Abuse. Figure 3 indicates the relative distribution of most of the diagnoses present in the BPD archive.

Figure 2: Ethnic group distribution of BPD archive
3.1.1 Cohort Description for Case-Control Analysis

In terms of cohort size, the aim was to have an equal number of cases and controls, with the control samples similar to cases in terms of gender and ethnicity. However, the final cohort consisted of 224 cases affected with BPDI and 202 background control samples. Within the case group, 93 individuals were derived from 35 BPD families and 131 were singletons (i.e. the only person to be represented in the cohort, from a particular family). Figure 4 represents the numbers of females and males and Figure 5 shows the number of individuals from each of the ethnic groups, within the study cohort. A proportion of controls were of unknown gender and age. There was more than two times the number of Caucasians as there were Mixed Ancestry individuals in the control group. Thus, controls were not similar to the case group in terms of gender and ethnicity. However, these variables were adjusted for in the statistical analysis.

Figure 3: Distribution of most prevalent phenotypes of BPD cohort
Figure 4: Number of males and females in case and control groups.

Figure 5: Number of Caucasian and Mixed Ancestry individuals in case and control groups.
3.1.2 Cohort Description for Brain Imaging Analysis
The imaging cohort consisted of 22 individuals affected with BPD, 20 of who were female and two were males. A total of 14 individuals were Caucasian and eight individuals were of Mixed Ancestry ethnicity.

3.2 DNA Quality Determination
To determine the quality of DNA, samples were quantified by means of spectrophotometry and agarose gel electrophoresis. Samples had to have a concentration of greater than 50ng/µl to be deemed satisfactory for further analysis. DNA quality was determined by resolution on a 1% agarose gel (Appendix 3). Bands which appeared thick and bright at the top of the lanes were indicative of non-degraded, good quality DNA. This was shown in lanes 4, 5, 7, 8, and 9 (Figure 6). Degraded DNA samples appeared as a smear, as can be seen in lanes 1, 2, 3 and 6 (Figure 6).

Figure 6: Integrity gel of genomic DNA. Lane M: 100bp MWM, Lanes 1-9: Genomic DNA from subjects in study.
3.3 Gene and SNP Selection, Primer Design and Synthesis
Candidate genes investigated in this study were selected based on results from the CFG method [Le-Niculescu et al., 2008] and the most significant findings in GWAS, related to neuropsychiatric disorders and BPD specifically. The specific SNPs that were finally chosen per gene were those that had been reported with significant results from previous GWAS and candidate-gene association studies. Primers were designed to produce a PCR product incorporating the specific (respective) SNP for each gene.

3.4 Polymerase Chain Reaction (PCR)
3.4.1 PCR Optimisation
The optimum Ta for each PCR reaction was determined by a temperature gradient, whereby 12 reactions were carried out at different temperatures ranging from 50°C to 62°C. The Ta for each of the 12 reactions is displayed in Appendix 6. A separate temperature gradient was carried out for each primer pair. Figure 7 indicates the temperature gradient for KLF12 rs9600160. Non-specificity is seen at lower temperatures (lanes 1-4) and little/no product at higher temperatures (lanes 11 and 12). The optimum annealing temperature for this reaction was chosen as approximately 58°C as there was ample product and an absence of non-specificity. Similarly, for the other primer pairs, an optimum Ta was chosen based on the best product band intensity and a lack of non-specific products (See Appendix 7 for additional gel pictures). These values, for each primer pair, are displayed in Table 16.
Figure 7: KLF12 rs9600160 (180bp) Temperature Gradient. Lane M: 100bp MWM. Lanes 1-12: PCR products amplified at varying annealing temperatures as specified in Table 1. Lane C: water control.

Table 16: Optimal annealing temperature determined for each primer pair

<table>
<thead>
<tr>
<th>Primer</th>
<th>SNP</th>
<th>Approximate Optimum Ta (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>KLF12</td>
<td>rs9600160</td>
<td>58</td>
</tr>
<tr>
<td>ARNTL</td>
<td>rs7107287</td>
<td>57</td>
</tr>
<tr>
<td>BDNF</td>
<td>rs6265</td>
<td>58</td>
</tr>
<tr>
<td>ALDH1A1</td>
<td>rs1888202</td>
<td>57</td>
</tr>
<tr>
<td>A2BP1(A)</td>
<td>rs7204975</td>
<td>53</td>
</tr>
<tr>
<td>A2BP1(B)</td>
<td>rs10500336</td>
<td>53</td>
</tr>
<tr>
<td>MBP</td>
<td>rs12967023</td>
<td>57</td>
</tr>
<tr>
<td>AK3L1</td>
<td>rs4916031</td>
<td>57</td>
</tr>
<tr>
<td>GSK3β</td>
<td>rs334558</td>
<td>62</td>
</tr>
<tr>
<td>ANK3</td>
<td>rs9804190</td>
<td>57</td>
</tr>
<tr>
<td>CACNA1C</td>
<td>rs1006737</td>
<td>62</td>
</tr>
</tbody>
</table>

3.4.2 PCR Conditions for Candidate Genes
Based on optimum Ta and PCR product size, each of the amplicons was amplified in one of five multiplex reactions. The amplified products were then resolved on a 2% agarose gel (Appendix 3) and stained with EtBr. The first lane of each gel (Figures 8-12) contained 100bp MWM which serves as a size standard. The last lane of each gel contains
the negative (water) control. None of the water controls had amplified product indicating that none of the PCR reactions were contaminated.

Group A consisted of three primer pairs, $ARNTL$ rs7107287, $MBP$ rs12967023, and $ALDH1A1$ rs1888202 and had a Ta of 57°C (See Figure 8). Primer dimers formed at the bottom of each lane. The $GSK3β$ rs334558 and $CACNA1C$ rs1006737 primer pairs constituted Group B and had amplicon sizes of 302bp and 220bp, respectively (Figure 9). The Ta of this group was 62°C. Group C also consisted of two primer pairs, $KLF12$ rs9600160 and $BDNF$ rs6265. The Ta of this group was 58°C and had amplicon sizes of 180bp ($KLF12$) and 338bp ($BDNF$) (Figure 10). Groups, D and E, each consisted of two primer pairs. Group D was made up of the two $A2BPI$ primer pairs, $A2BP1(A)$ rs7204975 and $A2BP1(B)$ rs10500336 (Figure 11). The amplicon sizes were 192bp and 250bp, respectively, and had a Ta of 53°C. The primer pairs $AK3l1$ rs4916031 and $ANK3$ rs9804190 formed part of Group E which had a Ta of 57°C (Figure 12). These had amplicon sizes of 155bp and 227bp, respectively. The primer pairs of Group C and Group E were originally placed into one multiplex reaction (due to similar optimal Ta). However, when these products were resolved on a gel, it was noted that amplification of $AK3l1$ rs4916031 (155bp) was very faint and not easily distinguishable from the $KLF12$ rs9600160 (180bp) fragment. Therefore, these four primer pairs were separated into two PCR reactions.
Figure 8: Group A PCR Products for three primer pairs- *MBP* rs12967023 (170bp), *ARNTL* rs7107287 (237bp), *ALDH1A1* rs1888202 (394bp). Lane M: 100bp MWM. Lane C: water control.

Figure 9: Group B PCR Products for two primer pairs- *CACNA1C* rs1006737 (220bp), *GSK3β* rs334558 (302bp). Lane M: 100bp MWM. Lane C: water control.
Figure 10: Group C PCR Products for two primer pairs- *KLF12* rs9600160 (180bp), *BDNF* rs6265 (338bp). Lane M: 100bp MWM. Lane C: water control.

Figure 11: Group D PCR Products for two primer pairs- *A2BP1*(A) rs7204975 (192bp), *A2BP1*(B) rs10500336 (250bp). Lane M: 100bp MWM. Lane C: water control.
3.5 SNaPshot PCR

3.5.1 Sap and Exol Clean-up of PCR Products
PCR products were purified with SAP and Exol. To allow for further multiplexing, which in turn saved reagents and time, product from Groups A and B were multiplexed in the same clean-up reaction i.e. an equal ratio of Group A and Group B PCR products were pooled. Similarly, Groups C, D, and E PCR products were multiplexed in the same clean-up reaction.

3.5.2 Capillary Electrophoresis on ABI PRISM
Due to multiplexing at the SAP and Exol clean-up stage, genotyping was performed for two multiplexed reactions: Group A-B and Group C-D-E. Group A-B consisted of genotypes for five SNPs and Group C-D-E for six SNPs. Thus, the genotypes for 11 SNPs were incorporated into two SNaPshot reactions. SNaPshot reactions for Group A-B and Group C-D-E are represented by Figures 13 and 14, respectively. Homozygous genotypes are represented by a single, large peak and heterozygous genotypes by two,
overlapping or adjacent peaks, of different colour. G-alleles are blue, C-alleles are black, T-alleles are red and A-alleles are green in colour.

3.5.3 Analysis of Results

![Figure 13: SNaPshot results for Group A-B. Yellow pane- ARNTL rs7107287 (genotype: GG), green pane- MBP rs12967023 (genotype: CT), red pane- ALDH1A1 rs1888202 (genotype: GG), pink pane- GSK3β rs334558 (genotype: TT), purple pane- CACNA1C rs1006737 (genotype: GG). Y-axis represents fluorescent units. X-axis represents base position. Yellow peaks represent GeneScan-120 Liz size standard.](image)

![Figure 14: SNaPshot results for Group C-D-E. Blue pane- ANK3 rs9804190 (genotype: CC), purple pane- AK3L1 rs4916031 (genotype: AA), yellow pane- A2BP1(B) rs10500336 (genotype: TT), pink pane- KLF12 rs9600160 (genotype: AA), green pane- BDNF rs6265 (genotype: CC), orange pane – A2BP1(A) rs7204975 (genotype CC). Y-axis represents fluorescent units. X-axis represents base position. Yellow peaks represent GeneScan-120 Liz size standard.](image)
3.6 Genotype Validation

3.6.1 PCR of DNA Samples for Cycle Sequencing

Figure 15 illustrates CACNA1C rs1006737 PCR products (Lanes 1-3) used for sequencing. The first lane of the gel contained the 100bp MWM (Appendix 3) and the second lane the negative (water) control.

![Image of gel with bands at 100bp, 500bp, and 220bp](image)

**Figure 15:** PCR amplification for cycle sequencing reaction. Lane M: 100bp MWM. Lane C: water control. Lanes 1-3: CACNA1C rs1006737 (220bp) PCR product.

3.6.2 Cycle Sequencing

In total, 32 samples were sequenced for validation of SNAPSHOT genotyping. These samples represented the genotypes for each of the 11 SNPs (KLF12 rs9600160 had only two genotypes present in the sample population).

3.6.3 Analysis of Results

Sequencing results are represented by Figures 16 and 17. Figure 16 represents a portion of the sequencing results of the BDNF rs6265 SNP. This particular sample was homozygous C for this SNP, which is represented as a blue peak in the electropherogram at position 250 (as indicated by the red arrow). Figure 17 represents a heterozygous
sample (C/T) for the SNP ANK3 rs9804190. This is seen by the blue and red peaks at position 142 (as indicated by the red arrow).

Figure 16: Sequencing results for BDNF rs6265. This particular sample was homozygous C (blue peak) which is seen at position 250 (as seen indicated by the red arrow).

Figure 17: Sequencing results for ANK3 rs9804190. This particular sample was heterozygous (C/T- blue and red peaks) seen at position 142 (as indicated by the red arrow).
3.7 Statistical Analysis for Case Control Association

3.7.1 Genotype Frequencies, Hardy Weinberg Equilibrium Calculations and Logistic Regression with R

Genotype counts of each SNP, for cases (n=224) and controls (n=202), are displayed in the following graphical outputs Figures 18 (A, B, C, D, E, and F) and 19 (A, B, C, D, and E). KLF12 (rs9600160) had no homozygous G individuals. It is evident from these figures, that there are no large differences in genotype frequencies between cases and controls. For control samples, none of the SNPs were found to be out of HWE. As a number of cases were related to one another, a random affected individual was selected for logistic regression analysis (logistic regression cannot adjust for relatedness). This resulted in a case cohort of 168 unrelated individuals. Based on the results from logistic regression, both gender (p-value 0.0151) and ethnicity (p-value 0.0107) had an association with disease affection status. Therefore, all subsequent association analyses were adjusted for gender and ethnicity. Case-control analysis showed that ANK3 rs9804190 and A2BP1 rs7204975 were the only SNPs to have obtained a p-value of less than 0.05 (Table 17), in a recessive genetic model. This suggests that having two copies of the C-allele of the ANK3 rs9804190 increased the risk of being diagnosed with BPD by an odds ratio of 1.68. Also, being homozygous for the C-allele of A2BP1 rs7204975, increased the risk of having BPD by an odds ratio of 1.77. Therefore, being homozygous C at both SNPs increases the risk for BPD nearly two fold. However, after adjusting for multiple testing using Bonferroni correction (new p-value cutoff 0.0038), these SNPs were no longer significantly associated with BPD.
Figure 18: Genotype counts of cases and controls for the following SNPs A) *ARNTL* rs7107287, B) *MBP* rs12967023, C) *ALDH1A1* rs1888202, D) *GSK3β* rs334558, E) *CACNA1C* rs1006737, and F) *ANK3* rs9804190.
Figure 19: Genotype counts of cases and controls for the following SNPs of A) *AK3L1* rs4916031, B) *KLF12* rs9600160, C) *A2BP1*(A) rs7204975, D) *A2BP1*(B) rs10500336, and E) *BDNF* rs6265.
Table 17: R logistic regression output for case-control data using a recessive genetic model (adjusted for gender and ethnicity)

<table>
<thead>
<tr>
<th>Gene</th>
<th>SNP</th>
<th>Odds Ratio</th>
<th>Lower</th>
<th>Upper</th>
<th>Z-test</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARNTL</td>
<td>rs7107287</td>
<td>1.21</td>
<td>0.77</td>
<td>1.91</td>
<td>0.82</td>
<td>0.4139</td>
</tr>
<tr>
<td>MBP</td>
<td>rs12967023</td>
<td>0.89</td>
<td>0.55</td>
<td>1.45</td>
<td>-0.45</td>
<td>0.6516</td>
</tr>
<tr>
<td>ALDH1A1</td>
<td>rs1888202</td>
<td>1.44</td>
<td>0.81</td>
<td>2.56</td>
<td>1.25</td>
<td>0.2131</td>
</tr>
<tr>
<td>GSK3B</td>
<td>rs334558</td>
<td>0.80</td>
<td>0.48</td>
<td>1.34</td>
<td>-0.85</td>
<td>0.3979</td>
</tr>
<tr>
<td>CACNA1C</td>
<td>rs1006737</td>
<td>1.19</td>
<td>0.57</td>
<td>2.51</td>
<td>0.47</td>
<td>0.6416</td>
</tr>
<tr>
<td>ANK3</td>
<td>rs9804190</td>
<td>1.68</td>
<td>1.04</td>
<td>2.72</td>
<td>2.14</td>
<td>0.0324</td>
</tr>
<tr>
<td>AK3L1</td>
<td>rs4916031</td>
<td>0.87</td>
<td>0.55</td>
<td>1.36</td>
<td>-0.61</td>
<td>0.5437</td>
</tr>
<tr>
<td>A2BP1(B)</td>
<td>rs10500336</td>
<td>0.97</td>
<td>0.31</td>
<td>3.02</td>
<td>-0.04</td>
<td>0.9642</td>
</tr>
<tr>
<td>KLF12</td>
<td>rs9600160</td>
<td>1.17</td>
<td>0.65</td>
<td>2.10</td>
<td>0.51</td>
<td>0.6101</td>
</tr>
<tr>
<td>BDNF</td>
<td>rs6265</td>
<td>0.71</td>
<td>0.44</td>
<td>1.18</td>
<td>-1.31</td>
<td>0.1893</td>
</tr>
<tr>
<td>A2BP1(A)</td>
<td>rs7204975</td>
<td>1.77</td>
<td>1.09</td>
<td>2.89</td>
<td>2.30</td>
<td>0.0215</td>
</tr>
</tbody>
</table>

3.7.2 MQLS (adjusting for relatedness) and Pedstats

MQLS is a software package capable of analyzing case-control data where certain individuals within the study are related to one another. For this analysis, genotypes needed to be represented in a pedigree file which specified familial relationships between individuals (including unaffected, non-genotyped family members). Before MQLS analysis could be performed, Pedstats a freely available program, was used to execute quality checks on pedigree data. For example, it was able to detect Mendelian inconsistencies whereby a mother was homozygous for one allele at a particular SNP and her offspring was homozygous for the other allele at that same SNP. In reality, this scenario is impossible as
the offspring would have to have inherited one of the alleles from the mother and either be homozygous for the “maternal” allele or be heterozygous at that position. This Mendelian inconsistency was due to genotype calling error and could therefore be rectified. Therefore, Pedstats was able to ensure that pedigree data based on genotypes were “error-free”. From the Pedstats results (Table 18), there were 593 individuals overall, in the pedigree file, with 370 families. There were 280 females and 258 males (the genders for a number of individuals were unknown).

From the MQLS output (Table 19), it can be noted that only one of the SNPs, MBP rs12967023 obtained a significant p-value of less than 0.05. However, after Bonferroni correction for multiple testing, this SNP was no longer significant. It is noted that A2BP1(A) rs7204975 had a p-value approaching significance (p-value 0.0678).

No significant interactions were found, between ANK3 rs9804190, A2BP1(A) rs7204975, and MBP rs12967023, after adjusting for Bonferroni Correction (results not shown).

<table>
<thead>
<tr>
<th>Table 18: Pedigree structure from Pedstats output</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No. of individuals in Pedigree File</strong></td>
</tr>
<tr>
<td><strong>No. of Founders</strong></td>
</tr>
<tr>
<td><strong>Gender Distribution</strong></td>
</tr>
<tr>
<td><strong>No. of Families</strong></td>
</tr>
</tbody>
</table>
Table 19: MQLS Output for Case Control Association

<table>
<thead>
<tr>
<th>Gene</th>
<th>SNP</th>
<th>Chi-Squared Test Statistic</th>
<th>LOD</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARNTL</td>
<td>rs7107287</td>
<td>1.85983</td>
<td>0.403857</td>
<td>0.1726</td>
</tr>
<tr>
<td>MBP</td>
<td>rs12967023</td>
<td>4.84863</td>
<td>1.05287</td>
<td>0.0277</td>
</tr>
<tr>
<td>ALDH1A1</td>
<td>rs1888202</td>
<td>0.075409</td>
<td>0.016375</td>
<td>0.7836</td>
</tr>
<tr>
<td>GSK3β</td>
<td>rs334558</td>
<td>0.24238</td>
<td>0.052632</td>
<td>0.6225</td>
</tr>
<tr>
<td>CACNA1C</td>
<td>rs1006737</td>
<td>0.242548</td>
<td>0.052669</td>
<td>0.6224</td>
</tr>
<tr>
<td>ANK3</td>
<td>rs9804190</td>
<td>0.150696</td>
<td>0.032723</td>
<td>0.6979</td>
</tr>
<tr>
<td>AK3L1</td>
<td>rs4916031</td>
<td>0.064394</td>
<td>0.013983</td>
<td>0.7997</td>
</tr>
<tr>
<td>A2BP1(B)</td>
<td>rs10500336</td>
<td>0.450299</td>
<td>0.097781</td>
<td>0.5022</td>
</tr>
<tr>
<td>KLF12</td>
<td>rs9600160</td>
<td>0.014813</td>
<td>0.003217</td>
<td>0.9031</td>
</tr>
<tr>
<td>BDNF</td>
<td>rs6265</td>
<td>0.964785</td>
<td>0.2095</td>
<td>0.3260</td>
</tr>
<tr>
<td>A2BP1(A)</td>
<td>rs7204975</td>
<td>3.33621</td>
<td>0.724448</td>
<td>0.0678</td>
</tr>
</tbody>
</table>

3.8 Statistical Analysis of Neuro-imaging
The following imaging data (reported as cubic millimeters (mm)), were obtained for a small subset of affected individuals (n=22), of the larger cohort: right, left and total hippocampal volumes (HCV) and intracranial volumes (ICV). Each of the HCVs was divided by the ICV to account for inter-individual differences in brain volume. This accounted for differences in cranial volumes across gender and ethnicity. The Shapiro-Wilk test was used to determine whether imaging data were normally distributed. Results from this test indicated that only left HCV/ICV measure was not normally distributed (p-value 0.03). Therefore, the Kruskal-Wallis rank sum test, the non-parametric equivalent to ANOVA, was used to determine whether there was any difference in the median left HCV/ICV values among the genotype
groups. Bartlett’s test was used to ensure that all variances were equal amongst the genotype groups. All calculations indicated non-significant p-values (not shown), therefore, all groups had equal variances. Results from Table 20 (Kruskal Wallis chi-squared, F-values (ANOVA) and p-values) suggests that none of the genotypes have an association with left HCV/ICV, right HCV/ICV, and total HCV/ICV.

Table 20: Kruskal Wallis and ANOVA values for each of the SNPs and imaging measures, Left HCV, Right HCV and Total HCV

<table>
<thead>
<tr>
<th>Gene</th>
<th>SNP</th>
<th>KW chi-squared</th>
<th>p-value</th>
<th>F-value</th>
<th>p-value</th>
<th>F-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARNTL</td>
<td>rs7107287</td>
<td>1.0909</td>
<td>0.2963</td>
<td>0.996</td>
<td>0.3308</td>
<td>0.7478</td>
<td>0.398</td>
</tr>
<tr>
<td>MBP</td>
<td>rs12967023</td>
<td>2.4304</td>
<td>0.2967</td>
<td>1.6437</td>
<td>0.2210</td>
<td>1.3163</td>
<td>0.2927</td>
</tr>
<tr>
<td>ALDH1A1</td>
<td>rs1888202</td>
<td>2.3211</td>
<td>0.3133</td>
<td>1.1948</td>
<td>0.33</td>
<td>1.3019</td>
<td>0.3011</td>
</tr>
<tr>
<td>GSK3β</td>
<td>rs334558</td>
<td>1.9813</td>
<td>0.3713</td>
<td>1.0793</td>
<td>0.3598</td>
<td>1.0866</td>
<td>0.3574</td>
</tr>
<tr>
<td>CACNA1C</td>
<td>rs1006737</td>
<td>4.7621</td>
<td>0.09246</td>
<td>1.9667</td>
<td>0.1674</td>
<td>1.806</td>
<td>0.1914</td>
</tr>
<tr>
<td>ANK3</td>
<td>rs9804190</td>
<td>2.7391</td>
<td>0.2542</td>
<td>1.0134</td>
<td>0.3828</td>
<td>1.1203</td>
<td>0.3479</td>
</tr>
<tr>
<td>AK3L1</td>
<td>rs4916031</td>
<td>2.6</td>
<td>0.2725</td>
<td>1.8067</td>
<td>0.1927</td>
<td>1.623</td>
<td>0.2249</td>
</tr>
<tr>
<td>A2BP1(B)</td>
<td>rs10500336</td>
<td>0.8475</td>
<td>0.6546</td>
<td>0.9849</td>
<td>0.3918</td>
<td>0.8857</td>
<td>0.4288</td>
</tr>
<tr>
<td>KLF12</td>
<td>rs9600160</td>
<td>1.0435</td>
<td>0.307</td>
<td>2.3747</td>
<td>0.1390</td>
<td>1.5087</td>
<td>0.2336</td>
</tr>
<tr>
<td>BDNF</td>
<td>rs6265</td>
<td>1.1938</td>
<td>0.2746</td>
<td>1.5175</td>
<td>0.2323</td>
<td>1.2198</td>
<td>0.2825</td>
</tr>
<tr>
<td>A2BP1(A)</td>
<td>rs7204975</td>
<td>1.4522</td>
<td>0.4838</td>
<td>0.7756</td>
<td>0.4744</td>
<td>0.6812</td>
<td>0.5179</td>
</tr>
</tbody>
</table>
Chapter 4: Discussion

BPD is a severe disorder affecting the lives of millions of individuals worldwide. Currently, the exact aetiology of BPD is unknown, treatment options are limited and a definitive cure seems distant. Due to the high heritability of this disorder, genetics is thought to be a key factor in the discovery of the pathophysiology. However, to date, even with high through-put genomic technologies, this has proven to be challenging. Numerous genes have been implicated in the aetiology of BPD, but have subsequently been difficult to replicate in independent studies. Besides genetic research, brain imaging has also been utilized in an attempt to elucidate biological mechanisms of BPD causality or association. However, as with genetic research, findings have been inconsistent.

In the present study, ten genes were selected from a wide range of candidate genes which were prioritized through a convergent approach and GWAS. It was found that SNPs in the genes ANK3, A2BP1 and MBP may play a role in the aetiology of BPD. In addition, due to the enabling multidisciplinary research environment here at UCT, it was possible to explore a dimension of brain imaging research of BPD subjects for whom genetic studies were also undertaken. In this pilot study of 22 affected individuals, none of the SNPs in ten genes tested were associated with MRI features.

Over the years, in efforts to determine the underlying genetics of BPD, linkage, association, and GWAS have implicated numerous chromosomal regions and genes. Generally, these original findings have proven difficult to replicate. These discrepancies
could be due, at least in part, to aetiological heterogeneity [Faraone and Tsuang, 2003]. The Division of Human Genetics at UCT has a well characterized BPD archive which serves as a rich resource for genetic studies of the disorder. This cohort is predominantly familial (total of 221 families), and limited to a few ethnic groups, which is surmised to reduce the aetiological heterogeneity. The offspring of individuals affected with BPD are at higher genetic risk for developing the disorder and offer a significant source of convergent information regarding the development of this disorder. Gene-environment interactions may also be assessed in this circumstance as these offspring may have been affected by mood episodes experienced by the affected parent [Chang et al., 2003]. Another way of attempting to reduce genetic and aetiological heterogeneity i.e. to have a more homogeneous study group, is to sample geographically and genetically isolated populations. These derived cohorts may be advantageous as they may have a higher prevalence of certain diseases, an endogamously-related probability of mapping recessive genes, and less migration [Peltonen et al., 2000]. In our cohort, the Afrikaner Caucasians and indigenous African Xhosa-speaking populations are relatively homogenous and genetically isolated, and may be ideal candidates for psychiatric genetic research. However, due to the small numbers of indigenous Africans and other isolated ethnic groups in our cohort, individuals from this population group were not included in the current study.

The first aim of this investigation was to determine whether variants in ten selected and previously associated candidate genes are risk factors for the development of BPD. Logistic regression is a statistical method commonly used in the analysis of case-control data. Using this statistical method, and indicating a recessive genetic model, only ANK3 rs9804190 (OR 1.68) and A2BP1(A) (rs7204975) (OR 1.77) showed nominal
associations with BPD. Thus, individuals who were homozygous C at both SNPs had an increased risk for BPD. This suggests that BPD is a recessive disorder whereby individuals require two copies of the disease-predisposing allele to be affected. However, the mode of inheritance of BPD is thought to be non-Mendelian and a result of the interaction of numerous genes with the environment [Serretti and Mandelli, 2008].

After adjusting for multiple testing using the Bonferroni correction method, none of the SNPs were statistically significant. This situation may reflect i) that there are no genuine associations, or ii) that there is some association and this disappears upon correction for multiple testing. The Bonferroni correction decreases Type I error (the probability of rejecting the null hypothesis, when it is true). This method is often used as it is relatively simple to calculate and can be applied to most multiple-testing scenarios [Bender and Lange, 2001]. However, the Bonferroni correction has often been criticized as being overly conservative and increases the chances of Type II error (the probability of rejecting the alternate hypothesis, when it is true) [Perneger, 1998]. Thus, the debate continues as to whether correction for multiple-testing, particularly using the Bonferroni method, is necessary for genetic studies. In light of this, it is worth re-examining the biological role of the two candidates: ANK3 and A2BP1.

ANK3 encodes a neurally expressed, structural protein that is thought to be involved in the maintenance of ion channels and cell adhesion molecules of axonal segments [Kordeli et al., 1995]. Four GWAS, in various population groups, have shown ANK3 associated with BPD [Baum et al., 2008; Ferreira et al., 2008; Scott et al., 2009; Smith et al., 2009]. These findings were replicated in a meta-analysis study consisting of Caucasian subjects (n=1668 cases and 1604 controls) [Schulze et al., 2008]. The associated SNP is located within an intronic region of the gene and has no known
functional significance. ANK3 is located on chromosome 10, in a region that was originally linked with late onset Alzheimer's disease [Myers et al., 2000]. Subsequently, a candidate gene association study detected a nominal association of this gene and late onset Alzheimer's disease [Morgan et al., 2008].

A2BP1 encodes a RNA-binding protein, which interacts with the gene ataxin-2 which has been shown to underlie the aetiology of the neurodegenerative disorder, SCA type 2 [Shibata et al., 2000; Kiehl et al., 2001]. This gene was previously found to be associated with several neuropsychiatric disorders, including attention deficit hyperactivity disorder (ADHD) [Elia et al., 2009], autism [Martin et al., 2007], and Schizoaffective Disorder [Hamshere et al., 2009]. The A2BP1 SNP, rs7204975, is located within an intron of the gene, and seemingly, has no known functional significance. However, it may be in linkage disequilibrium with the causal variant or it may play a role in the regulation of the gene. This SNP was found to be associated with BPD, in a German population group, by a previous GWAS (p-value 0.001) [Baum et al., 2008]. However, A2BP1(B) rs10500336 which had a p-value of 0.018 in the same GWAS, did not attain significance in the current study.

From the MQLS analysis, only MBP rs12967023 showed a nominal association with a p-value of less than 0.05. However, after correction for multiple testing, this finding was no longer significant. The product encoded by MBP comprises 30% of the protein in CNS myelin. Myelin is found around the axons of neurons and increases the velocity of nerve impulses [Kamholz et al., 1987]. MBP expression has been shown to be reduced in SCZ postmortem brains [Tkachev et al., 2003] and a decreased intensity of myelin staining has been observed in BPD brain tissue [Regenold et al., 2007]. Similarly, a BPD twin
study has shown that an affected twin (n=24) has significantly reduced white matter volumes compared to the unaffected co-twin (n=15) and demographically matched control twins (n=27) [Kieseppä et al., 2003]. *MBP* has also been implicated in the recovery from neuronal insults. In animal models, autoimmune T-cells, activated against the CNS antigen MBP, reduce degeneration after optic nerve injury [Hauben et al., 2000]. Furthermore, rats that have been immunized with MBP show more severe neuronal loss after CNS injury [Jones et al., 2004]. This suggests that MBP has a possible role in the progression of neurodegeneration. The *MBP* SNP, rs12967023, is located within an intron, therefore, the exact function is unknown. This SNP has however been previously associated with BPD (p-value 0.001) in a GWAS involving Caucasian subjects [Sklar et al., 2008].

Surprisingly, none of the other candidate genes (or at least the SNPs assayed) was found to be associated with BPD. *BDNF*, and particularly the Val66Met polymorphism, has often been found to be associated with BPD [Neves-Pereira et al., 2002; Sklar et al., 2002; Lohoff et al., 2005; Muller and De, 2006; Okada et al., 2006]. However, the results of the current study are in accordance with many studies that have not found an association between BPD and this neurotrophic factor [Nakata et al., 2003; Kunugi et al., 2004]. *CACNA1C* which has been strongly associated with BPD in a GWAS [Ferreira et al., 2008] and replicated in independent studies [Scott et al., 2009; Schulze et al., 2008] was not associated with BPD in the present study. The difference in results between the current and previous studies could be due to any one, or a combination, of several factors including the different ethnic groups under investigation, differences in study designs and/or sample numbers in the respective studies.
The final cohort consisted of 224 subjects affected with BPDI and 202 background control individuals. The controls formed part of the background population and were not screened for any psychiatric illness. Thus, in accordance with the estimated prevalence of BPD, 1% of these controls may have been predisposed to BPD. The cases and controls were not “matched” with respect to gender and ethnicity. However, it is acceptable to adjust for potential confounders such as gender and ethnicity in the regression analysis [Bland and Altman, 1994].

An association was found between gender and BPD. This is not in accordance with previous findings, where it has been shown that BPD affects both males and females at equal frequencies [Merikangas et al., 2007; American Psychiatric Association. Task Force on DSM-IV, 1994]. However, it has been found that men, generally, have an earlier age of onset than females [Kennedy et al., 2005]. In the current study, differences in genotype frequencies between Caucasian and Mixed Ancestry individuals were observed. This finding does not agree with a previous U.S. study which found no difference in BPD prevalence amongst ethnic groups (n=9282 study subjects) [Merikangas et al., 2007]. However, different ethnic groups (viz. Hispanic, Caucasian, and African American) were investigated by Merikangas et al. [2007], compared to the current study which included Caucasians (British and Afrikaner ancestry) and individuals of Mixed Ancestry. The South African Mixed Ancestry group is an admixed population group consisting of individuals of Khoesan, Bantu-speaking African, European and Asian ethnicity [de Wit et al., 2010].

The second aim of the current study was to determine whether the endophenotypic hippocampal volumes have an association with any of the selected candidate genes. This
component of the project should be considered a pilot study as there were only 22 participants with both imaging and genotyping data available. Nonetheless, previous imaging genetic studies have had similar sample sizes [Bueller et al., 2006; Chepenik et al., 2008]. As with other similar studies, the brain area under investigation was the hippocampus, involved in long-term memory [Squire, 1992], stress responses [Sapolsky et al., 1984] and spatial navigation [Maguire et al., 2000]. In this study, no association was found between any of the genotypes and hippocampal volumes (left, right and total), corrected for total intracranial volumes. This finding could be expected, as the study was a pilot, testing the technologies and analytical capacities between the clinical phenotype, imaging endophenotypes and genetic data. From the outset, it was obvious that the small sample size and potential number of tests would render the study inadequately powered.

Previous studies have implicated the BDNF rs6265 (Val66Met) polymorphism with hippocampal volume. In healthy controls, the Met-allele was associated with reduced hippocampal volumes [Bueller et al., 2006; Chepenik et al., 2008]. These findings have also been shown for individuals affected with major depression (n=60 cases and 60 controls) [Frodl et al., 2007]. However, results from the current study suggest that there is no association between BDNF rs6265 and brain hippocampal and intracranial volumes of individuals affected with BPDI. This finding is similar to that shown by Dutt et al. [2009] where the BDNF rs6265 SNP has no association with hippocampal volumes of individuals diagnosed with psychosis (n= 128 cases, 194 unaffected relatives, and 61 controls).
4.1 Future Directions
Future studies should consider the following aspects: The functionality of the following SNPs should be investigated: ANK3 rs9804190, A2BP1 rs7204975 and MBP rs12967023. Statistical significance of genotype with diagnosis does not illustrate the biological mechanism for risk. Therefore, the underlying biology needs to be determined [Bigos et al., 2010]. The effect of the SNPs on expression of the gene can be investigated by the use of Allelic Expression Imbalance (AEI). This method relies on the differentiation of maternal and paternal alleles and to determine the relative mRNA expression of each [Buckland, 2006]. An individual needs to be heterozygous at the SNP of interest and any deviation from a 50:50 expression of each allele suggests an effect of the genetic variation [Quinn et al., 2010].

It is possible that other variants, within the candidate genes investigated, have an association with BPD. High throughput technologies, such as next generation re-sequencing, could be used to get a more dense coverage of variants, in a wide range of subjects, of a particular candidate gene. For example, exome sequencing has been able to identify the candidate gene and variants underlying the aetiology of a rare Mendelian disorder [Ng et al., 2009].

Subsequent to their original CFG publication, Le-Niculescu et al. [2008] did a follow-up study whereby they re-generated a prioritized BPD list using four instead of three GWAS results and included other, more recent lines of evidence [Patel et al., 2010]. The top three candidate genes were the same as those from the 2008 study. They also developed a genetic risk prediction score (GRPS) whereby individuals affected with BPD could be distinguished from controls. This score was based on the most significant panel
of SNPs from the CFG list (which included the MBP rs12967023 SNP), providing proof of principle whereby the candidate gene list could be used to predict risk for developing BPD at a population level [Patel et al., 2010]. Although it may not be used clinically yet, this predictive value is enormously important in dismantling large heterogeneous cohorts into more aetiologically homogeneous entities, towards identifying the underlying range of molecular pathologies.

Future work could also investigate alternate candidate genes. Interestingly, genes that have been previously implicated in the aetiology of cancer have also been shown to be significant for BPD. The WTCCC GWAS found a significant signal for the locus containing the PALB2 gene [Stratton, 2007]. This finding was replicated in a subsequent study whereby both PALB2 and breast cancer 2 (BRCA2) were found to be positively associated to BPD in a Scandinavian population group [Tesli et al., 2010].

Recently, work done by our group on a subset of the larger BPD cohort, have shown that genes involved in glutamatergic signaling have an association with BPD severity. The gene for glutamate receptor, metabotropic 3 (GRM3) was associated with a four-fold increased risk of having psychotic symptoms and the genes, ionotropic glutamate receptor, N-methyl D-aspartate 2B (GRIN2B) and D-amino acid oxidase activator (DAOA) interact to increase the risk of being admitted to hospital for mania [Dalvie et al., 2010] (Appendix 8). Therefore, these genes and others involved in glutamatergic signaling, should be investigated more extensively.
With regard to the imaging genetics, it is obvious that the technologies are at hand to carry out the work in this promising area. Apart from a significant increase in numbers of subjects, imaging data should be obtained from control individuals to determine whether hippocampal volumes have an association with BPD. Changes in hippocampal volumes have often been described in psychiatric illnesses such as borderline PD, PTSD, major depressive disorder (MDD) [Bremner et al., 2000] and BPD. Chepenik et al. [2009] have found that BPD-affected individuals have smaller hippocampal volumes compared to controls. However, results have varied across studies. Some studies have not found differences in hippocampal volumes between individuals affected with BPD and controls [Hajek et al., 2009].

Since the implicated candidate genes (i.e. ANK3, A2BP1 and MBP) have been shown to have a relationship with neurodegeneration, it would be of interest to determine whether any of these genes have an association with any neurodegenerative disorders. The Division of Human Genetics has a long history of investigating neurodegenerative disorders such as SCA [Bryer et al., 2003; Greenberg et al., 2006] and Huntington’s Disease [Greenberg et al., 1991; Squitieri et al., 1994; Bardien et al., 2007; Scholefield and Greenberg, 2007]. A cursory study may involve investigating neurodegenerative symptoms or disorders in individuals or families where the BPD phenotype is associated with these genes.

**4.2 Conclusions**

This study showed that intronic SNPs within ANK3, A2BP1 and MBP may have an association with BPD and that polymorphisms within the ten investigated candidate genes do not show an association with endophenotypic measures of BPD. The three implicated candidate genes have an indirect relationship with neurodegenerative
disorders [Hauben et al., 2000; Myers et al., 2000; Shibata et al., 2000; Kiehl et al., 2001; Jones et al., 2004; Morgan et al., 2008]. Neurodegenerative disorders such as Parkinson’s and Huntington’s are characterized by a loss of neurons, BPD has often been shown to be accompanied by a decrease in neuronal plasticity as well as decreases in various regional volumes [Rajkowska, 2002]. This disorder has also been shown to be associated with minor cognitive impairments [Martinez-Aran et al., 2004], as well as an increased risk for developing dementia in the elderly [Nunes et al., 2007]. This may suggest that BPD could be considered to have features in common with the classical neurodegenerative disorders. Therefore, it can be postulated that the pathophysiology of BPD has a link with neurodegeneration.

The present study highlights the importance of examining complex phenotypes in a familial setting where the traditional view of comorbidities may be examined across individuals “at risk”. Phenotypic studies in such high risk subjects and families continue to be important in future studies of complex psychiatric disorders, including BPD.
Websites Referenced

2) http://www.ensembl.org/index.html
3) http://fokker.wi.mit.edu/primer3/input.htm
4) http://www.idtdna.com/analyzer/Applications/OligoAnalyzer
5) http://blast.ncbi.nlm.nih.gov/Blast.cgi
6) http://eu.idtdna.com/order/order.aspx
7) http://www.neb.com/nebecomm/products/productm0293.asp
8) http://www.R-project.org
9) http://sun.ac.za/cubic
10) http://surfer.nmr.mgh.harvard.edu/
11) http://chpc.ac.za
References


Craddock N, O'Donovan M, Owen M. 2009. Psychosis genetics: Modeling the relationship between schizophrenia, bipolar disorder, and mixed (or" schizoaffective") psychoses. Schizophr Bull. 35:482-490.


Elia J, Gai X, Xie H, Perin J, Geiger E, Glessner J, D'arcy M. 2009. Rare structural variants found in attention-deficit hyperactivity disorder are preferentially associated with neurodevelopmental genes. Mol Psychiatry 15:637-646.


Stratton M. 2007. Genome-wide association study of 14,000 cases of seven common

Tesli M, Athanasiu L, Matingsdal M, Kähler AK, Gustafsson O, Andreassen BK, Werge T,
Hansen T, Mors O, Mellerup E. 2010. Association analysis of PALB2 and BRCA2 in
bipolar disorder and schizophrenia in a scandinavian case–control sample. American

4:53-91.

Thornton T, McPeek MS. 2007. Case-control association testing with related individuals:
A more powerful quasi-likelihood score test. The American Journal of Human Genetics
81:321-337.

Tkachev D, Mimmack ML, Ryan MM, Wayland M, Freeman T, Jones PB, Starkey M,
Webster MJ, Yolken RH, Bahn S. 2003. Oligodendrocyte dysfunction in schizophrenia

van der Schot AC, Vonk R, Brans RGH, van Haren NEM, Koolschijn P, Nuboer V, Schnack
HG, van Baal GCM, Boomsma DI, Nolen WA. 2009. Influence of genes and environment
on brain volumes in twin pairs concordant and discordant for bipolar disorder. Arch
Gen Psychiatry 66:142-151.

Val66Met polymorphism: Relation to familiar risk of affective disorder, BDNF levels and
salivary cortisol. Psychoneuroendocrinology 34:1380-1389.


Appendices

Appendix 1: Patient Consent Form
University of Cape Town
Division of Human Genetics
IIDMM, LEVEL 3
UCT Medical School, Observatory 7925
Tel: (021) 406 6425  Fax: (021) 406 6826

REQUEST FOR MOLECULAR STUDIES (DNA)

Blood should be drawn in 2 plastic EDTA Tubes
(Purple top) +/- 10ml each using a yellow barrel.
Each tube should be inverted to mix and should be
clearly labelled with the patient's name and DOB.
Keep blood in fridge at 4°C until able to send to laboratory.
Please DO NOT send specimens on ice or frozen.

Please fill in all the information requested:

Surname: ___________________________ First Name(s): ___________________________

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: ______________________ Fax: ______________________

Referring Doctor/Sister: ______________________ Town: ______________________ Tel: ______________________

Hospital or Address: ______________________ Fax: ______________________

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 □ Waardenburg Syndrome □

Additional disorders (apparent or previously treated):

Additional family history

Clinical Details:

Physical disability □ Mental retardation □ Dwarfism □ 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: ______________________

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:

1. ___________________________

2. ___________________________

3. ___________________________

4. ___________________________

5. ___________________________

6. ___________________________

7. ___________________________

8. ___________________________

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:

   (a) ___________________________

   (b) ___________________________

   (c) ___________________________

   (d) ___________________________

   (e) ___________________________

   (f) ___________________________

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

   (a) ___________________________

   (b) ___________________________

   (c) ___________________________

3. I request that no portion of the sample be stored for later use. □ (MARK IF APPLICABLE)

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

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

6. I have been informed that:

   (a) ___________________________

   (b) ___________________________

   (c) ___________________________

   (d) ___________________________

   (e) ___________________________

   (f) ___________________________

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 2: DNA Consent Form
Appendix 2: Gentra Puregene™ DNA Isolation Protocol

- Added 900µl red blood cell lysis solution to microcentrifuge tube
- Added 500µl whole blood and mixed by inverting
- Incubated for one minute at room temperature
- Centrifuged for 20 seconds at 13 000rpm to pellet white blood cells. Discarded supernatant
- Added 300µl cell lysis solution and vortexed
- Incubated at 37°C for ten minutes
- Added 100µl protein precipitation solution and inverted
- Centrifuged for one minute at 13 000rpm
- Added supernatant to clean microcentrifuge tube with 300µl isopropanol
- Mixed by inversion until DNA was visible as threads or clumps
- Centrifuged for one minute at 13 000rpm
- Discarded supernatant
- Added 300µl 70% ethanol and inverted several times
- Centrifuged for one minute 13 000rpm
- Discarded supernatant
- Rehydrated with 50µl DNA hydration solution
Appendix 3: Reagents, Buffers and Solutions

Molecular weight marker- GeneRuler™ 100 bp Plus DNA Ladder (Fermentas Life Sciences, Hanover, USA)

- 25µl stock 100bp DNA ladder solution
- 225µl dH2O
- 250µl loading dye

(Adapted from: http://www.fermentas.com/catalog/electrophoresis/generulers.htm)

10x TBE Buffer (stock)
- 216g (0.89M) Tris (B&M Scientific cc, Cape Town, SA)
- 110g (0.89M) Boric Acid (MP Biomedicals Inc., USA)
- 14.8g (0.04M) EDTA (BDH Electron® Laboratory Supplies, UK)
- Made up to 2L with Sabax dH2O (Adcock Ingram, Johannesburg, SA)
1x TBE Buffer (working stock)
- 1:10 dilution was made with stock TBE buffer with dH_{2}O

Loading Buffer (Fermentas Life Sciences, Hanover, USA)

1% Agarose Gel (100ml)
- 1g agarose (Whitehead Scientific, SA)
- 100ml 1X TBE (working stock)
- 6µl (5ng/µl) EtBr (Sigma, England)

2% Agarose gel (100ml)
- 2g agarose (Whitehead Scientific, SA)
- 100ml 1X TBE (working stock)
- 6µl (5ng/µl) EtBr (Sigma, England)
Appendix 4: GeneScan -120 Liz size standard
Appendix 5: Ethanol Precipitation

- The sequencing reaction was transferred to a 1.5ml eppendorf tube
- Added 50µl absolute ethanol and 2µl sodium acetate (3mM, pH5.5)
- Solution was mixed and left at -20 °C for more than 1 hour
- Spun @ 10 000rpm for 10 minutes
- Supernatant was discarded
- Added 50µl of 70% ethanol, vortexed
- Spun @ 10 000rpm for 10 minutes
- Supernatant was discarded
- Air dried for approximately 1 hour to evaporate the ethanol
- The DNA was resuspended in 10µl dH2O
Appendix 6: Annealing Temperatures (Ta) for Temperature Gradient Reactions

<table>
<thead>
<tr>
<th>Lane</th>
<th>Temperature (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>50.0</td>
</tr>
<tr>
<td>2</td>
<td>50.7</td>
</tr>
<tr>
<td>3</td>
<td>51.3</td>
</tr>
<tr>
<td>4</td>
<td>52.9</td>
</tr>
<tr>
<td>5</td>
<td>54.7</td>
</tr>
<tr>
<td>6</td>
<td>56.4</td>
</tr>
<tr>
<td>7</td>
<td>57.2</td>
</tr>
<tr>
<td>8</td>
<td>58.7</td>
</tr>
<tr>
<td>9</td>
<td>60.1</td>
</tr>
<tr>
<td>10</td>
<td>61.3</td>
</tr>
<tr>
<td>11</td>
<td>61.6</td>
</tr>
<tr>
<td>12</td>
<td>62.0</td>
</tr>
</tbody>
</table>
Appendix 7: Additional Temperature Gradient Gel Pictures for Candidate Gene SNPs

**ARNTL (237bp) Temperature Gradient.** Lane M: Molecular Weight Marker (MWM). Lanes 1-12: PCR products amplified at varying annealing temperatures as specified in Table 1. Lane C: water control.

**BDNF (338bp) Temperature Gradient.** Lane M: Molecular Weight Marker (MWM). Lanes 1-12: PCR products amplified at varying annealing temperatures as specified in Table 1. Lane C: water control.
**ALDH1A1** (394bp) Temperature Gradient. Lane M: Molecular Weight Marker (MWM). Lanes 1-12: PCR products amplified at varying annealing temperatures as specified in Table 1. Lane C: water control.

**A2BP1(A) (192bp)** Temperature Gradient. Lane M: Molecular Weight Marker (MWM). Lanes 1-12: PCR products amplified at varying annealing temperatures as specified in Table 1. Lane C: water control.

**A2BP1(B) (250bp)** Temperature Gradient. Lane M: Molecular Weight Marker (MWM). Lanes 1-12: PCR products amplified at varying annealing temperatures as specified in Table 1. Lane C: water control.
**MBP (170bp)** Temperature Gradient. Lane M: Molecular Weight Marker (MWM). Lanes 1-12: PCR products amplified at varying annealing temperatures as specified in Table 1. Lane C: water control.

**AK3L1 (155bp)** Temperature Gradient. Lane M: Molecular Weight Marker (MWM). Lanes 1-12: PCR products amplified at varying annealing temperatures as specified in Table 1. Lane C: water control.
GSK3β (302bp) Temperature Gradient. Lane M: Molecular Weight Marker (MWM). Lanes 1-12: PCR products amplified at varying annealing temperatures as specified in Table 1. Lane C: water control.

ANK3 (227bp) Temperature Gradient. Lane M: Molecular Weight Marker (MWM). Lanes 1-12: PCR products amplified at varying annealing temperatures as specified in Table 1. Lane C: water control.

CACNA1C (220bp) Temperature Gradient. Lane M: Molecular Weight Marker (MWM). Lanes 1-12: PCR products amplified at varying annealing temperatures as specified in Table 1. Lane C: water control.
Appendix 8: Psychosis and relapse in bipolar disorder are related to GRM3, DAOA, and GRIN2B genotype. Dalvie et al. [2010].
Psychosis and relapse in bipolar disorder are related to GRM3, DAOA, and GRIN2B genotype

S Dalvie¹, N Horn², C Nossek¹, L van der Merwe³,⁴, DJ Stein², R Ramesar¹

¹Division of Human Genetics, MRC/UCT Human Genetics Research Unit, Faculty of Health Sciences, University of Cape Town Medical School, Observatory, Cape Town, South Africa
²Department of Psychiatry and Mental Health, University of Cape Town Medical School, Observatory, Cape Town, South Africa
³Biostatistics Unit, Medical Research Council, Cape Town, South Africa
⁴Department of Statistics, University of Western Cape, Bellville, South Africa

Abstract

Objective: Dysfunction in glutamate signalling is thought to play a role in the pathophysiology of bipolar disorder (BD). There is evidence of associations between single nucleotide polymorphisms (SNPs) in GRM3, GRIN2B, and DAOA genes and the diagnosis of BD. In this pilot study, we investigated the frequency of SNP variants in these 3 genes within South African population groups, and assessed interactions between genes and phenotypes of BD disease severity. Method: Multiplex SNPsShot™ PCR was used to genotype 191 case and 188 control samples. Cases comprised of 191 individuals in a South African cohort of mixed ancestry and Caucasians, with BD Type 1. Phenotypes of BD disease severity were: age of onset, number of illness episodes, number of hospitalisations for depression or mania and history of psychotic symptoms. Results: There were no significant differences in SNP allele frequencies between cases and controls. In the case-only analysis, the GRM3 rs6465084 heterozygote was associated with a 4-fold increased risk of lifetime history of psychotic symptoms, and the specific variants within the gene pair, DAOA and GRIN2B, had a significant interaction with the number of hospitalisations for mania, with lowest admission rates associated with both pairs of ancestral alleles. Conclusion: In BD, variations in glutamatergic genes may influence phenotypes related to the severity of illness. Speculatively, newly derived genes associated with various evolutionary advantages, may also increase the risk for more severe BD. These preliminary findings deserve validation in a larger cohort.

Key Words: Manic-Depressive Psychosis; Glutamate; GRIN2B receptor; mGlur3; G72 protein; Human

Received: 20-04-2010
Accepted: 02-08-2010

Introduction

The lifetime prevalence of bipolar disorder Type 1 (BD) is approximately 1% in the US and Europe, and likely equally high in Southern Africa.²³ BD is associated with a high level of mortality (between 10-20% of individuals affected with BD successfully commit suicide) and morbidity (BD is estimated as one of the top ten causes of disability worldwide).³⁶ An earlier age of disease onset, increased number of episodes and hospitalisations for mania and depression, and the presence of psychotic symptoms are markers of a more severe disease phenotype.³⁶ Previous research has not examined associations between phenotype severity and glutamate signalling in BD, although variants within brain derived neurotrophic factor (BDNF) may be associated with indices of severity in BD.⁹¹¹ Glutamate signalling is involved in brain development and synaptic plasticity, both of which are modified in individuals affected with BD, and have been implicated in its aetiology.¹²¹³ A range of psychotropics involved in the treatment of BD alter glutamatergic function; for example, lamotrigine inhibits the release of glutamate and lithium affects glutamate receptor function.¹⁴ It has also been observed that variation in the glutamate receptor encoding gene, GRK4 (rs1954787) makes subjects with depressive episodes more likely to respond to citalopram.¹⁵ Fallin et al. found that the glutamatergic genes most likely to be associated with BD rather than schizophrenia (SZ) were GRM3 and GRIN2B, and replicated previous findings of strong associations between DAOA and BD diagnosis.¹⁶ Glutamate receptor, metabotropic 3 (GRM3) encodes subunit 3 of the Group II metabotropic glutamate receptors.
In glutamate receptor, metabotropic 2 (GRM2) and GRM3 knockout mice, Lyon et al. showed a compensatory up-regulation of remaining group II mGlUTs and the N-methyl D-aspartate (NMDA) receptor subunit 2A (NR2A) expression, and down-regulation of glutamate transporter expression. Post-mortem studies have suggested altered GRM3 expression in BD. Crespi et al. identified GRM3 as one of 14 genes with significant evidence of selection in human populations. GRM3 modulates glutamate neurotransmission by regulating the excitatory amino acid transporter 2 (EAAT2). Egan et al. have demonstrated that decreased levels of EAAT2 are associated with a GRM3 SNP (rs6465084), which has a weak association with SZ. Egan et al. have also demonstrated that individuals who are homozygous A for GRM3 rs6465084 have lower levels of N-acetylaspartate, an indirect indicator of synaptic function and glutamate levels and suggests that GRM3 rs6465084 has an effect on GRM3 transcript expression.

Ionotropic glutamate receptor, N-methyl D-aspartate 2B (GRIN2B) encodes the NR2B subunit of the NMDA receptor. Mice with enhanced NR2B function show supernormal learning and memory function, while BD sufferers consistently show impairment in working memory. Associations have been observed between several polymorphisms in the GRIN2B gene and the 1p13.1-p13.2 region (containing the GRIN2B gene) and BD. Miyatake et al. have observed that the T allele of the GRIN2B rs1019385 SNP results in increased luciferase transporter activity in the presence of nerve growth factor (NGF). The presence of a G allele had no effect on reporter activity, suggesting that GRIN2B rs1019385 may have a role in transcriptional control of the GRIN2B transcript.

The G72 gene product, D-amino-acid oxidase activator (DAOA), activates the peroxisomal protein D-amino-acid oxidase (DAO). This protein (G30) degrades D-serine, which acts, similar to glycine, as a coactivator on the “glycine binding site” of the glutamatergic NMDA receptor. Hashimoto et al. showed that the percentage of D-serine in the cerebrospinal fluid of drug naive SZ patients was significantly lower than that of control samples. Genetic variants of G72/G30 have been found to be associated with BD and SZ. In some populations, the at-risk haplotypes are shared between SZ and BD. The pathogenic mutations have not yet been identified but might be located in the vicinity of the G72/G30 gene complex or in the regulatory region.

### Method

#### Subjects

The cohort comprised of two groups: a group of BD Type 1 affected individuals (cases) and the background population group (controls). Blood samples were previously obtained, with the appropriate written informed consent. This study was approved by the University of Cape Town Health Science Faculty Research Ethics Committee (081/96) and is in accordance with the guidelines of the Helsinki Declaration of 2000.

The Division of Human Genetics has a BD database consisting of 883 individuals comprising 185 families. For the purposes of this study, the cases consisted of 191 individuals specifically diagnosed with BD Type 1 and the Structured Clinical Interview for the DSM-IV (SCID). This group comprised of 87 (46%) males and 104 (54%) females of ages ranging from 22 to 91 years. Subjects included 103 (54%) individuals of mixed ancestry and 88 (46%) Caucasian individuals. The cases included 95 individuals from 36 families that were related to at least one other case. The age of onset, number of depressive episodes, number of hospitalisations due to mania and depression of each of the cases were noted, as well as history of psychotic symptoms. The control group consisted of 188 unrelated individuals not diagnosed with any psychiatric disorder, and similar to the cases in terms of age, gender and ethnicity.

### DNA Analysis

The following SNPs were chosen for genotyping based on previous association to either BD or SZ: GRM3 rs6465084 (intronic), GRIN2B rs1019385 (5’ upstream), and DAOA rs701567 (intronic). Gene sequences were obtained from the National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov) and Ensembl (http://www.ensembl.org/index.html) databases and the SNP locations were determined. DNA concentration and integrity was determined by spectrophotometric quantification and agarose gel electrophoresis, respectively. Polymerase chain reaction (PCR) was performed on genomic DNA, according to standard procedures, for each of the SNPs in the candidate genes. The polymorphisms were genotyped using the SNaPshot™ multiplex ready reaction mix (Applied Biosystems, Warrington, UK) and the ABI Prism 3100 Genetic Analyzer (Applied Biosystems, CA, USA). Genotyping results were validated by cycle sequencing using BigDye Terminator Mix (Applied Biosystems, CA, USA) and the ABI Prism 3100 Genetic Analyzer (Applied Biosystems, CA, USA).

### Statistical Analysis

For the case-control analysis, one individual from each family was randomly selected, leaving 132 unrelated cases. For this analysis, logistic regression was used.

The case-only analyses involved all 191 cases, which were adjusted for relatedness. We used generalised linear models for all the association tests. For age of onset, we used a normal distribution for all the association tests. For age of onset, we used a normal distribution for all the association tests. For all the association tests, we used a normal distribution for all the association tests.
Results
Case-control analysis (132 cases; 188 controls)
The differences in the genotype and allele frequencies between cases and controls were not significant. Neither were any SNP-interactions on case-control status significant (results not shown). The allele frequencies for each of the SNPs obtained from the cases and controls in the current study, are displayed in Table I.

Case-only analysis (n=191)
All the results presented here were adjusted for gender, ethnic group, duration of disease and family-membership.

We did not find any significant association between the SNPs GRM3 rs6465084, GRIN2B rs1019385, DAOA rs701567 and main effects or interactions on age of onset (n=124), number of depressive episodes (n=130), or number of hospitalisations for depression (n=171). However, a significant association between GRM3 rs6465084 and history of psychosis (63% with a history and 37% without history of psychosis) was found. The GRM3 rs6465084 G-allele was found to increase the odds of having had a psychotic episode (OR=3.9; p=0.004) considerably. Even after adjusting for family-relatedness only, a significant association (OR=2.2; p=0.0137) was observed.

Reported number of hospitalisations for mania (n=167) varied between 0 and 11. The interaction between DAOA rs701567 and GRIN2B rs1019385 on number of hospitalisations for mania was significant (p=0.0108), and this interaction was independently significant after including other interactions in the model. The nature of the modelled interaction is illustrated in Figure 1 and can be described as follows: For individuals with the GRIN2B rs1019385 genotype G/G, their expected number of hospitalisations for mania increased with their number of DAOA rs701567 G alleles. For individuals with the GRIN2B rs1019385 genotype G/T, there were smaller decreases in number of hospitalizations and for those individuals with the T/T genotype, each DAOA rs701567 G allele resulted in larger decreases. Individuals who are homozygous A for DAOA rs701567 and homozygous T for GRIN2B rs1019385 had the largest number of hospitalisations, and those who are homozygous G for DAOA rs701567 in combination with being homozygous G for GRIN2B rs1019385, had the second largest. Those with DAOA rs701567 G/G and GRIN2B rs1019385 T/T had the least number of hospitalisations.

Discussion
The first set of findings related to SNP frequency showed that all 3 SNPs examined had a similar distribution between Caucasian and mixed ancestry groups, and that BD diagnosis was not associated with any of the SNPs examined. The second set of findings, related to BD severity, was that the non-ancestral GRM3 G allele at rs6465084 carries a 4 times greater risk of developing psychosis in BD subjects, and that alleles at DAOA rs701567 and GRIN2B rs1019385 interact such that the risk of hospital admission is higher with a greater number of non-ancestral alleles.

This is the first time that allele frequencies in the candidate genes studies here have been described for the mixed ancestry population of South Africa. SNP frequencies were broadly similar to reported frequencies for Caucasian groups described elsewhere (NCBI http://www.ncbi.nlm.nih.gov).22 No difference in frequency was detected for the ancestral allele for GRM3 rs6465084 and of GRIN2B rs1019385 in individuals of both Caucasian and mixed ancestry. The A/G alleles of DAOA rs701567 were found to occur at almost equal frequency between 0.43 and 0.48, and the A/T alleles of GRIN2B rs1019385 were found to occur at almost equal frequency between 0.46 and 0.56.

Table I: SNP Allele Frequencies For Each of the Candidate Genes

<table>
<thead>
<tr>
<th>Gene</th>
<th>SNP</th>
<th>Ethnic Group</th>
<th>No. of Cases</th>
<th>No. of Controls</th>
<th>Ancestral Allele Frequency Cases</th>
<th>Ancestral Allele Frequency Controls</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>GRM3</td>
<td>rs6465084</td>
<td>Caucasian, Mixed Ancestry</td>
<td>132</td>
<td>188</td>
<td>0.75</td>
<td>0.77</td>
<td>0.5963</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Caucasian only</td>
<td>69</td>
<td>89</td>
<td>0.73</td>
<td>0.78</td>
<td>0.3054</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mixed Ancestry only</td>
<td>63</td>
<td>99</td>
<td>0.77</td>
<td>0.76</td>
<td>0.7887</td>
</tr>
<tr>
<td>GRIN2B</td>
<td>rs1019385</td>
<td>Caucasian, Mixed Ancestry</td>
<td>132</td>
<td>188</td>
<td>0.35</td>
<td>0.34</td>
<td>0.8208</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Caucasian only</td>
<td>69</td>
<td>89</td>
<td>0.41</td>
<td>0.40</td>
<td>0.9181</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mixed Ancestry only</td>
<td>63</td>
<td>99</td>
<td>0.29</td>
<td>0.28</td>
<td>0.8304</td>
</tr>
<tr>
<td>DAOA</td>
<td>rs701567</td>
<td>Caucasian, Mixed Ancestry</td>
<td>132</td>
<td>188</td>
<td>0.43</td>
<td>0.48</td>
<td>0.2000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Caucasian only</td>
<td>69</td>
<td>89</td>
<td>0.41</td>
<td>0.41</td>
<td>0.8961</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mixed Ancestry only</td>
<td>63</td>
<td>99</td>
<td>0.45</td>
<td>0.56</td>
<td>0.0725</td>
</tr>
</tbody>
</table>
frequencies in both population groups.

We did not demonstrate that any of the putative risk alleles were associated with BD diagnosis. This is consistent with the lack of consistently reproducible associations between individual genes and BD diagnosis even in large data sets.31 There is increasing evidence that BD is a disorder where many genes of small effect contribute to risk of developing the disorder.20

In this study, we observed that the G allele at SNP rs6465084 GRM3 carries a 4 times greater risk of developing psychosis in subjects with BD. This is unexpected as the A allele in this SNP has been weakly associated with SZ diagnosis—and might be expected to be a risk factor for psychosis in BD.20 This finding by Egan et al., however, was not replicated by others.34,35 Egan et al. also reported an association between the A allele, impaired list learning and verbal fluency and low prefrontal levels of glutamate transporter EEAT2 in normal subjects and people with SZ.20 It is possible to speculate that the A allele confers a risk of a dysexecutive syndrome typical of certain types of SZ, while the G allele is associated with normal or enhanced executive function, consistent with the finding by Diller et al. that GRM3 is a gene which has evolved favourable new (derived) alleles.36 The G allele may, however, together with other risk factors, be associated with more severe mania. Crespi et al. identified GRM3 as a gene with significant evidence of selection in human populations, and although not enough is known about this SNP in global populations or its function in normal or disease states, it may be one which differentiates Kraepelinean “dementia praecox” from “manic depression”1,19,37.

The rs701567 SNP in DAOA was not associated with BD diagnosis in this study, contrary to the finding of Fallin et al.18 The G allele was not associated with any phenotype or severity measures; however there was a significant gene–gene interaction with GRIN2B rs1019385 SNP on the number of hospitalisations for mania. In the South African context, hospitalisation is most often required for emergency treatment of severe episodes. The combination of both ancestral alleles (G for DAOA and T for GRIN2B) conferred the lowest risk of admission while homozygotes with a derived allele/ancestral allele combination had the highest risk for hospitalisation. Homozygotes for both derived alleles had a relatively low risk of hospitalisation. DAOA activates DAO which determines synaptic glutamate and serine levels, while NMDA receptor function varies according to glutamate levels and GRIN2B genotype. Both DAOA and GRIN2B form part of the phosphoinositide-3 kinase/ AKT-signalling network (PI3K/AKT), involved in the regulation of cell proliferation, growth, development and apoptosis.38,39 We may speculate that heterozygotes have intermediate glutamate function and homozygotes an inflexible regulatory capacity for glutamate load or deficiency states, compromising normal neural development and plasticity.

A major limitation of this study was the relatively small sample size. To avoid this, recent studies have focussed on obtaining large consortia for genetic association tests. An example of this would be the Wellcome Trust Case Control Consortium which investigated the genetic basis of seven common diseases, including BD, using 2000 affected individuals for each disorder and 3000 shared controls. The outcome of this study was that modest genetic effects were able to be detected using larger sample sizes.40

The most constraining limitation of the current study is the lack of information about the in vivo and in vitro function of the specific SNPs examined. Future work is needed to determine the functionality of glutamate SNPs in terms of gene expression and physiological correlates, and whether these link to phenotypes for BD. The interaction between DAGO rs701567 and GRIN2B rs1019385 may be indirect or direct and if direct could be examined with a protein interaction assay such as yeast-2-hybrid screen. Other SNPs in the genes we examined and other genes involved in glutamate metabolism including glutamate receptor, ionotropic, N-methyl D-aspartate 1 (GRIN1), glutamate receptor, metabotropic 1 (GRM1) and glutamate receptor, ionotropic, kainate 1 (GRK1) and their interactions with BD phenotypes should also be explored. Another consideration for future work would be to have greater detail regarding cognitive phenotype. This could be helpful to test (post-hoc) hypotheses related to cognitive function in, for example, controls with GRM3 rs6465084 derived allele homozygotes.

Conclusion
To our knowledge this is the first study investigating these 3 glutamatergic SNPs and their interactions with BD diagnosis and phenotype. The study provided preliminary evidence that ancestral alleles in the 3 SNPs examined have a protective effect by decreasing the chances of having more severe phenotypes of BD, specifically psychosis and repeat hospital admission. It is possible to speculate that newly derived genes associated with various evolutionary advantages (including perhaps cognitive advantages), may also increase the risk for more severe bipolar disorder. However, findings from this pilot study warrant further research and validation.

Acknowledgements
We would like to thank Sr Gamede Benefeld for conducting the psychiatric interviews and Ms Alvera Vorster for intellectual input. We acknowledge the funding from the Medical Research Council (MRC), the National Research Foundation (NRF), the University of Cape Town and the Benfara Scholarship.

References


